Theor Appl Genet (2005) 111: 1457–1465
DOI 10.1007/s00122-005-0027-y

ORIGINAL PAPER

Min Xu · Reid G. Palmer

# Molecular mapping of *k2 Mdh1-n y20*, an unstable chromosomal region in soybean [*Glycine max* (L.) Merr.]

Received: 23 March 2005 / Accepted: 22 June 2005 / Published online: 9 August 2005


**Abstract** In the soybean genome, a chromosomal region covering three tightly linked genes, *k2*, *Mdh1-n*, and *y20*, was found very unstable. It was suspected that the instability of the *k2 Mdh1-n y20* chromosomal region was caused by a non-autonomous transposable element residing adjacent to or in this region. In this study, we located and mapped this region with simple sequence repeat (SSR) markers on the soybean integrated map using five mapping populations. The *k2 Mdh1-n y20* chromosomal region was located on molecular linkage group H. The integrated map from five mapping populations consisted of 13 loci in the order Satt541, Satt469, Sat_122, Satt279, Satt253, Satt314, *Mdh1-n*,*y20*, *k2*, Satt302, Satt142, Satt181, and Satt434. The *k2 Mdh1-n y20* chromosomal region was very close to Satt314, Satt253, and Satt279. The genetic distance between the *Mdh1-n* gene and Satt314 was less than 1 cM. The results of the mapping study were consistent with the results from previous studies that the *Mdh1-n* mutation in T261 (*k2 Mdh1-n*) and the *Mdh1-n y20* mutation in T317 (*Mdh1-n y20*) were caused by deletions. In addition, another putative deletion was found in the genome of T261 which covered three SSR markers (Satt314, Satt253, and Satt279).

Communicated by S. J. Knapp

This is a joint contribution of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa, Project No. 3769, and from the USDA, Agricultural Research Service, Corn Insects and Crop Genetics Research Unit, and supported by the Hatch Act and the State of Iowa. The mention of a trademark or proprietary product does not constitute a guarantee or warranty of the product by Iowa State University or the USDA, and the use of the name by Iowa State University or the USDA implies no approval of the product to the exclusion of others that may also be suitable.

M. Xu · R. G. Palmer (✉)
USDA ARS CICGR, Department of Agronomy,
and Interdepartmental Genetics Graduate Program,
Iowa State University, Ames, IA, 50011 USA
E-mail: rpalmer@iastate.edu

## Introduction

Three recessively inherited loci, *k2*, *Mdh1-n*, and *y20*, were tightly linked in the soybean genome. The *k2* locus conditions tan-saddle seed coat, the *Mdh1-n* locus conditions mitochondrial malate dehydrogenase 1 (MDH1) null, and the *y20* locus conditions chlorophyll-deficient (CD) foliage. The recombination rate between *k2* and *Mdh1-n* was detected as $1 \pm 1.36\%$ in the cross of T261 (*k2 Mdh1-n*) × wild-type cultivars (*K2 Mdh1*) (Chen and Palmer 1996). The recombination rate between *k2* and *Mdh1-n y20* was zero in the coupling phase [T253 (*k2 Mdh1-n y20*) × Clark-*w1* (L69-4776) (*K2 Mdh1 Y20*)] (Palmer 1984), and $3.04 \pm 0.48\%$ in the repulsion phase [T317 (*Mdh1-n y20*) × T239 (*k2*) or Clark-*k2* (*k2*)] (Chen and Palmer 1998a). But no recombination was found between *Mdh1-n* and *y20* in any of the three cross-combinations. The co-segregation between the *Mdh1-n* and *y20* loci is possibly caused by a chromosome deletion. Southern blot results using watermelon *Mdh* cDNA as a probe showed that all the *Mdh1-n y20* mutants tested were missing a 5.5 kb *Eco*RI band, which corresponded to soybean *Mdh1* gene [gi:5929963] (Pittig et al. 1994; Imsande et al. 2001).

The *k2 Mdh1-n y20* chromosomal region is very unstable (Chen and Palmer 1998b). So far, 31 mutants have been reported from this region (Table 1) (Palmer et al. 2004). Among these mutants, four CD lines (T323, T324, T325, and T346) with *Mdh1-n y20* alleles were found in the descendants from germinal revertants of the *w4*-mutable line, T322, which was proposed to contain an active transposable element at the *W4* locus (Palmer et al. 1989; Chen et al. 1999). Two mutants with *Mdh1-n y20* alleles (T317 and T361) were found from tissue culture-derived plants (Amberger et al. 1992; Palmer et al. 2000). Twelve mutants with *k2 Mdh1-n y20* alleles, T334, T335, T336, T337, T338, T339, T340, T341, T342, T343, T344, and T345, were identified in the $F_2$ descendants derived from crosses of T239 (*k2*) or T261 (*k2 Mdh1-n*) with the wild-type parental strains of the

**Table 1** Summary of soybean lines mutated at the *k2 Mdh1-n y20* chromosomal region

| Mutant genes | *k2* | *k2 Mdh1-n* | *Mdh1-n* | *Mdh1-n y20* | *k2 Mdh1-n y20* | |
|---|---|---|---|---|---|---|
| Mutant lines | T239 (L63-365)<br>Clark-*k2* (L67-3483) | T261/S56-26 | PI 567.391<br>PI 567.630A<br>Mandell | T234<br>T317/LA45-1-5-1<br>T323/CD-1<br>T324/CD-2<br>T325/CD-3<br>T346/CD-9<br>T361 | T253<br>T334/X-197<br>T335/X-203<br>T336/X-217<br>T337/X-219<br>T338/X-241<br>T339/X-451<br>T340/M-7-2<br>T341/M-11-4 | T342/M-11-7<br>T343/M-15-23<br>T344/M-19-3<br>T345/M-20-11<br>T347/X-193<br>T348/X-194<br>T349/RP-95-649<br>T350<br>T351 |

*w4-m* and *Y18-m* mutable lines that were suspected to contain an active transposon in their genomes (Chen and Palmer 1998b). The hypothesis for the instability of the *k2 Mdh1-n y20* chromosomal region is that a non-autonomous transposable element could reside adjacent to or in this region (Chen and Palmer 1998b). The non-autonomous transposable element could excise from its chromosomal position and cause mutations such as chromosome deletions, when it is activated by tissue culture or by transposases provided *in trans* through transposon tagging experiments or crossing experiments.

In the present study, the objective was to position the *k2 Mdh1-n y20* region in the soybean integrated genetic linkage map (Song et al. 2004) and to determine the order of the *k2*, *Mdh1-n*, and *y20* genes with simple sequence repeat (SSR) markers.

## Materials and methods

### Plant material and DNA extraction

Five $F_2$ mapping populations named POP-1, −2, −3, −4, and −5, were used to map the *k2 Mdh1-n y20* chromosomal region. They were constructed from five different crosses, each derived from a single $F_1$ seed. The details are described in Table 2. $F_2$ seeds in POP-1 and −2 were placed on germination paper at 32°C in a dark growth chamber for 3 days. Two punches of samples were taken from cotyledons of 3-day-old seedlings by a 100-µl micropipette. Samples were analyzed using starch gel electrophoresis to determine malate dehydrogenase isozyme banding patterns (Cardy and Beversdorf 1984). The seedlings were placed into a 32°C growth chamber with light, and transplanted in the field at the Bruner Farm (Ames, IA) after 4 days. All plants were recorded for seed-coat color at maturity and threshed individually. $F_2$ plants in POP-3, −4, and −5 were grown at the Bruner Farm. Leaf color was scored after 3 weeks, and seed-coat color was scored at maturity. Plants were threshed individually. The genotypes of $F_2$ individuals in POP-1, −3, and −5 were evaluated for the target gene (or genes) by analyzing their $F_{2:3}$ families.

About 3–4 g young leaves were sampled from each $F_2$ plant and their parental lines of the five mapping populations, and ground into powders after they were freeze-dried. Genomic DNA was extracted from the leaf powders using CTAB method (Keim et al. 1988), diluted into 10 ng/µl, and stored at 4°C.

### Bulked segregant analysis (BSA)

POP-1 [cultivar Minsoy (PI 27890)(wild-type) × T261 (*k2 Mdh1-n*)] was used in BSA (Michelmore et al. 1991)

**Table 2** The $F_2$ populations for mapping the *k2 Mdh1-n y20* chromosomal region

| Population | Parental lines | Genotype | Description | Population size | Target gene or genes |
|---|---|---|---|---|---|
| POP-1[a] | Minsoy (♀) | *K2K2 Mdh1Mdh1 Y20Y20* | Wild-type cultivar | 118 | *k2, Mdh1-n* |
| | T261 (♂) | *k2k2 Mdh1-nMdh1-n Y20Y20* | Tan-saddle seed coat and MDH1 null | | |
| POP-2 | Williams (♀) | *K2K2 Mdh1Mdh1 Y20Y20* | Wild-type cultivar | 122 | *k2, Mdh1-n* |
| | T261 (♂) | *k2k2 Mdh1-nMdh1-n Y20Y20* | Tan-saddle seed coat and MDH1 null | | |
| POP-3[a] | Williams (♀) | *K2K2 Mdh1Mdh1 Y20Y20* | Wild-type cultivar | 84 | *k2* |
| | T239 (♂) | *k2k2 Mdh1Mdh1 Y20Y20* | Tan-saddle seed coat | | |
| POP-4 | Minsoy (♀) | *K2K2 Mdh1Mdh1 Y20Y20* | Wild-type cultivar | 113 | *k2* |
| | Clark-*k2*(♂) | *k2k2 Mdh1Mdh1 Y20Y20* | Tan-saddle seed coat | | |
| POP-5[a] | T317 (♀) | *K2K2 Mdh1-nMdh1-n y20y20* | Yellow seed coat, MDH1 null, and yellow foliage | 84 | *k2, y20* |
| | T261 (♂) | *k2k2 Mdh1-nMdh1-n Y20Y20* | Tan-saddle seed coat, MDH1 null, and green foliage | | |

[a]Populations whose $F_2$ individuals had genotype-evaluation in the $F_{2:3}$ generation

to identify the markers linked to the *k2 Mdh1-n* region. Two bulks were made. Bulk 1 (B1) contained DNA aliquots from ten $F_2$ individuals in POP-1 homozygous for wild-type phenotypes (yellow seed coat and MDH1 present). Bulk 2 (B2) contained DNA aliquots from ten $F_2$ individuals in POP-1 homozygous for tan-saddle seed coat and MDH1 null. SSR analysis was conducted with B1, B2, and the two parental lines (Minsoy and T261) to screen for candidate markers linked to the target region.

### SSR analysis

SSRs were amplified by polymerase chain reaction (PCR) in 30 µl mixture containing: 50 ng genomic DNA, 1× PCR buffer, 1.75 mM $MgCl_2$, 150 µM dNTP, 0.15 µM of each primer (Song et al. 2004), and 3 U Taq DNA polymerase (Promega, Madison, WI, USA). The PCR was performed in a MJR PTC-100 thermal cycler (MJ Research, Inc., Waltham, MA, USA) for 32 cycles of 45 s at 94°C, 45 s at 47°C, and 45 s at 68°C. Amplified products were resolved and evaluated by electrophoresis on 2% Agarose 3:1 (AMRESCO, Solon, OH, USA) gels in 1× TBE (Tris/borate/EDTA) buffer, or on 8% (w/v) denaturing gels [29:1 acrylamide-bisacrylamide, 5.6 M urea, and 30% (v/v) formamide) in 1× TAE (Tris/acetate/EDTA) buffer.

### Linkage analysis

The molecular linkage maps for five mapping populations first were calculated individually with Mapmaker 2.0 (Lander et al. 1987). The thresholds for linking two markers together were LOD 4.0 and recombination value 0.4. The genetic distances were converted from recombination rates using the Kosambi map function (Kosambi 1944). Then, the five maps were integrated with Joinmap 3.0 (Van Olijen and Voorrips 2001).

## Results

### Molecular mapping of the *k2 Mdh1-n* chromosomal region with mapping population POP-1

The mapping population POP-1 contained 118 $F_2$ individuals derived from a single $F_1$ seed from the cross of Minsoy × T261 (*k2 Mdh1-n*) (Table 2). The $F_{2:3}$ segregation for both the *k2* and *Mdh1-n* genes fit a 1 wild-type homozygote:2 heterozygotes:1 recessive homozygote ratio with $\chi^2 = 0.02$, $P = 0.99$ and $\chi^2 = 0.02$, $P = 0.99$, respectively (Table 3).

BSA analysis was used to identify the SSR markers linked with the *k2 Mdh1-n* chromosomal region. A total of 182 SSR markers from the 20 soybean molecular linkage groups (MLGs) (Song et al. 2004), were selected initially to screen the two bulks (B1 and B2) constructed from POP-1 and the two parental lines (Minsoy and T261). The average genetic distances between any two adjacent markers were about 25 cM. The results showed that Satt253 in MLG H was polymorphic between the two parental lines and the two bulks. Its primers could amplify a ~138 bp band from Minsoy and B1, but none from T261 and B2, which indicated that the SSR marker Satt253 could be tightly linked with the *k2 Mdh1-n* chromosomal region. To confirm this, Satt253 was used to screen the entire POP-1 population. Since Satt253 could only be detected from Minsoy, it was a dominant marker in the POP-1 mapping population. The genotypes of homozygotes of Minsoy allele and heterozygotes could not be distinguished in $F_2$ generation. In both cases, a ~138 bp band could be amplified by the primers of Satt253. If "A" was used for homozygotes of Minsoy allele of one locus, "H" for heterozygotes of this locus, and "B" for homozygotes of T261 allele of this locus, "D" was used to represent either A or H when they could not be distinguished. The $F_2$ D:B segregation ratio of Satt253 in POP-1 fit a 3:1 ratio ($\chi^2 = 0.01$, $P = 0.92$) (Table 3). The result from Mapmaker 2.0 showed that

**Table 3** Segregation of *k2, Mdh1-n*, and linked SSR markers in POP-1 derived from Minsoy × T261(*k2 Mdh1-n*)

| Locus | Distance[a] (cM) | Segregation[b] | | | | | $\chi^2_{(1:2:1)}$ | $\chi^2_{(3:1)}$ | *P* |
|---|---|---|---|---|---|---|---|---|---|
| | | A | H | D | B | Total | | | |
| Sat_122 | 7.1 | 31 | 57 | | 30 | 118 | 0.15 | | 0.93 |
| Satt314 | 1.7 | | | 89 | 29 | 118 | | 0.01 | 0.92 |
| Satt279 | 1.7 | | | 89 | 29 | 118 | | 0.01 | 0.92 |
| Satt253 | 1.7 | | | 89 | 29 | 118 | | 0.01 | 0.92 |
| *Mdh1-n* | 0.9 | 29 | 59 | | 30 | 118 | 0.02 | | 0.99 |
| *k2* | 0 | 30 | 59 | | 29 | 118 | 0.02 | | 0.99 |
| Satt302 | 18.9 | 36 | 52 | | 30 | 118 | 2.27 | | 0.32 |
| Satt142 | 24.1 | 38 | 52 | | 28 | 118 | 3.36 | | 0.19 |
| Satt181 | 28.4 | 37 | 54 | | 27 | 118 | 2.54 | | 0.28 |
| Satt434 | 45.6 | 37 | 58 | | 23 | 118 | 3.36 | | 0.19 |

[a]Distance from the *k2* locus

[b]A: homozygous for the allele from Minsoy at this locus; B: homozygous for the allele from T261 at this locus; H: heterozygous for the allele at this locus; D: A or H. The genotypes of the *Mdh1-n* and *k2* loci were determined in the $F_{2:3}$ generation. The genotypes of SSR markers were determined in the $F_2$ generation

**Table 4** Segregation of *k2, Mdh1-n*, and linked SSR markers in POP-2 derived from Williams × T261(*k2 Mdh1-n*)

| Locus | Distance[a] (cM) | Segregation[b] | | | | | $\chi^2$ (1:2:1) | $\chi^2$ (3:1) | *P* |
|---|---|---|---|---|---|---|---|---|---|
| | | A | H | D | B | Total | | | |
| Satt541 | 7.0 | 33 | 58 | | 30 | 121 | 0.36 | | 0.84 |
| Sat_122 | 4.9 | 33 | 55 | | 32 | 120 | 0.85 | | 0.65 |
| Satt314 | 0.8 | | | 90 | 32 | 122 | | 0.10 | 0.75 |
| Satt279 | 0.8 | | | 90 | 31 | 121 | | 0.02 | 0.87 |
| Satt253 | 0.8 | | | 90 | 32 | 122 | | 0.10 | 0.75 |
| *Mdh1-n* | 0.8 | | | 90 | 32 | 122 | | 0.10 | 0.75 |
| *k2* | 0.0 | | | 91 | 31 | 122 | | 0.01 | 0.92 |
| Satt302 | 23.9 | 30 | 58 | | 34 | 122 | 0.56 | | 0.76 |
| Satt142 | 32.6 | 33 | 59 | | 29 | 121 | 0.34 | | 0.84 |
| Satt181 | 37.3 | 39 | 54 | | 28 | 121 | 3.40 | | 0.18 |

[a]Distance from the *k2* locus

[b]A: homozygous for the allele from Williams at this locus; B: homozygous for the allele from T261 at this locus; H: heterozygous for the allele at this locus; D: A or H. The genotype of each locus was determined in the $F_2$ generation

Satt253 was linked to the *k2* locus with a LOD of 23.83 and the *Mdh1-n* locus with a LOD of 26.02.

An additional 14 markers, Satt353, Satt568, Satt192, Satt442, Satt541, Satt469, Sat_122, Satt052, Satt314, Satt279, Satt302, Satt142, Satt181, and Satt434, from MLG H were used to screen the two parental lines of POP-1. These 14 SSR markers and Satt253 also were used to screen the parental lines of the other four populations described in the following sections.

As a result, seven markers, Sat_122, Satt279, Satt314, Satt302, Satt142, Satt181, and Satt434, showed polymorphism between the two parental lines, Minsoy and T261 (*k2 Mdh1-n*). Satt279 and Satt314 were dominant markers. Like Satt253, they can only be amplified from Minsoy, but not from T261. The others were co-dominant markers. The dominant markers segregated in POP-1 as a 3D:1B ratio, and the co-dominant markers segregated in POP-1 as a 1A:2H:1B ratio (Table 3). Data were subjected to analysis with Mapmaker 2.0. The most plausible position of the *k2 Mdh1-n* chromosomal region was located between Satt302 and three co-segregating SSR markers, Satt253, Satt279, and Satt314, with genetic distances of 18.9 cM (distance between *k2* and Satt302) and 0.8 cM (distance between *Mdh1-n* and Satt253, Satt279, and Satt314), respectively (Table 3). The *k2* gene was closer to Satt302, and the *Mdh1-n* gene was closer to the three co-segregating SSR markers. The recombination rate between *k2* and *Mdh1-n* was detected as 0.9%.

## Molecular mapping of the *k2 Mdh1-n* chromosomal region with mapping population POP-2

The mapping population POP-2 consisted of 122 $F_2$ plants derived from a single $F_1$ seed of cultivar Williams × T261 (*k2 Mdh1-n*) (Table 2). The $F_2$ segregation for both the *k2* and *Mdh1-n* genes fit a 3 wild-type:1 homozygous recessive ratio with $\chi^2 = 0.01$, $P = 0.92$ and $\chi^2 = 0.10$, $P = 0.75$, respectively (Table 4).

Eight markers that could detect polymorphism between the two parental lines were used to generate the map. Dominant markers Satt314, Satt279, and Satt253 segregated in the mapping population as a 3:1 ratio (Table 4). Co-dominant markers segregated for homozygous for the Williams allele:heterozygous:homozygous for the T261 (*k2 Mdh1-n*) allele in the mapping population as a 1:2:1 ratio (Table 4). The results showed that the *Mdh1-n* gene co-segregated with Satt314, Satt279, and Satt253. The *k2* gene was located between the *Mdh1-n* gene and the Satt302 marker with 0.8 cM genetic distance to the *Mdh1-n* gene, and 23.9 cM genetic distance to the Satt302 marker (Table 4).

## Molecular mapping of the *k2* gene with mapping population POP-3 and −4

The mapping population POP-3 was formed by crossing Williams with T239 (*k2*) (Table 2). The population contained 84 $F_2$ individuals, descendants from a single $F_1$ seed. The genotype of each $F_2$ individual was determined by $F_{2:3}$ family analyses. The $F_{2:3}$ segregation ratio for the *k2* gene fit a 1 wild-type homozygote:2 heterozygotes:1 recessive homozygote ratio ($\chi^2 = 0.19$, $P = 0.91$) (Table 5).

The mapping population POP-4 consisted of 113 $F_2$ plants derived from a single $F_1$ seed of Minsoy × Clark-*k2* (*k2*) (Table 2). The $F_2$ segregation for the *k2* gene fit a 3 wild-type:1 homozygous recessive ratio ($\chi^2 = 0.24$, $P = 0.62$) (Table 6).

SSR markers that were polymorphic between the parental lines of each population were co-dominant, and segregated in the mapping populations as the ratio of 1A [homozygous for alleles from wild-types (Williams or Minsoy)]:2H (heterozygous):1B [homozygous for alleles from mutants (T239 or Clark-*k2*)] (Tables 5, 6). The *k2* gene was mapped between Satt279 and Satt302 with 3.0 cM (POP-3) or 2.6 cM (POP-4) genetic distance to

**Table 5** Segregation of *k2* and linked SSR markers in POP-3 derived from Willams × T239 (*k2*)

| Locus | Distance[a] (cM) | Segregation[b] | | | | $\chi^2_{(1:2:1)}$ | *P* |
|---|---|---|---|---|---|---|---|
| | | A | H | B | Total | | |
| Satt541 | 14.0 | 19 | 43 | 22 | 84 | 0.26 | 0.88 |
| Sat_122 | 11.0 | 19 | 42 | 23 | 84 | 0.38 | 0.83 |
| Satt314 | 4.2 | 22 | 44 | 18 | 84 | 0.57 | 0.75 |
| Satt253 | 3.6 | 22 | 42 | 20 | 84 | 0.10 | 0.95 |
| Satt279 | 3.0 | 23 | 41 | 20 | 84 | 0.26 | 0.88 |
| *k2* | 0.0 | 20 | 44 | 20 | 84 | 0.19 | 0.91 |
| Satt302 | 24.5 | 17 | 44 | 23 | 84 | 1.05 | 0.59 |
| Satt142 | 36.1 | 18 | 42 | 24 | 84 | 0.86 | 0.65 |

[a]Distance from the *k2* locus
[b]A: homozygous for the allele from Williams at this locus; B: homozygous for the allele from T239 at this locus; H: heterozygous for the allele at this locus. The genotype of the *k2* locus was determined in the $F_{2:3}$ generation. The genotypes of SSR markers were determined in the $F_2$ generation

**Table 6** Segregation of *k2* and linked SSR markers in POP-4 derived from Minsoy × Clark-*k2*

| Locus | Distance[a] (cM) | Segregation[b] | | | | | $\chi^2_{(1:2:1)}$ | $\chi^2_{(3:1)}$ | *P* |
|---|---|---|---|---|---|---|---|---|---|
| | | A | H | D | B | Total | | | |
| Satt541 | 9.4 | 26 | 59 | | 28 | 113 | 0.29 | | 0.86 |
| Satt469 | 9.4 | 26 | 59 | | 28 | 113 | 0.29 | | 0.86 |
| Sat_122 | 6.7 | 25 | 61 | | 27 | 113 | 0.79 | | 0.67 |
| Satt253 | 3.5 | 27 | 60 | | 26 | 113 | 0.45 | | 0.80 |
| Satt279 | 2.6 | 28 | 60 | | 25 | 113 | 0.59 | | 0.74 |
| *k2* | 0.0 | | | 87 | 26 | 113 | | 0.24 | 0.62 |
| Satt142 | 24.0 | 27 | 55 | | 31 | 113 | 0.36 | | 0.83 |
| Satt181 | 30.4 | 24 | 59 | | 30 | 113 | 0.86 | | 0.65 |
| Satt434 | 52.3 | 24 | 51 | | 38 | 113 | 4.54 | | 0.10 |

[a]Distance from the *k2* locus
[b]A: homozygous for the allele from Minsoy at this locus; B: homozygous for the allele from Clark-*k2* at this locus; H: heterozygous for the allele at this locus; D: A or H. The genotype of each locus was determined in the $F_2$ generation

**Table 7** Segregation of *k2, y20*, and linked SSR markers in POP-5 derived from T317 (*Mdh1-n y20*) × T261 (*k2 Mdh1-n*)

| Locus | Distance[a] (cM) | Segregation[b] | | | | | $\chi^2_{(1:2:1)}$ | $\chi^2_{(3:1)}$ | *P* |
|---|---|---|---|---|---|---|---|---|---|
| | | A | H | D | B | Total | | | |
| Sat_122 | 13.6 | 21 | 47 | | 16 | 84 | 1.79 | | 0.41 |
| Satt253 | 1.3 | | | 67 | 17 | 84 | | 1.02 | 0.31 |
| Satt279 | 1.3 | | | 67 | 17 | 84 | | 1.02 | 0.31 |
| Satt314 | 1.3 | | | 67 | 17 | 84 | | 1.02 | 0.31 |
| *y20* | 0.6 | 19 | 47 | | 18 | 84 | 1.21 | | 0.55 |
| *k2* | 0.0 | 20 | 46 | | 18 | 84 | 0.86 | | 0.65 |
| Satt302 | 21.6 | 25 | 35 | | 24 | 84 | 2.36 | | 0.31 |
| Satt142 | 28.4 | 25 | 40 | | 19 | 84 | 1.05 | | 0.59 |
| Satt181 | 32.0 | 21 | 42 | | 21 | 84 | 0.00 | | 1.00 |

[a]Distance from the *k2* locus
[b]A: homozygous for the allele from T317 at this locus; B: homozygous for the allele from T261 at this locus; H: heterozygous for the allele at this locus; D: A or H. The genotypes of the *k2*and *y20* loci were determined in the $F_{2:3}$ generation. The genotypes of SSR markers were determined in the $F_2$ generation

Satt279, and 24.5 cM (POP-3) or 24.0 cM (POP-4) genetic distance to Satt302 (Tables 5, 6)

## Molecular mapping of the *k2* and *y20* genes with mapping population POP-5

The mapping population POP-5 was constructed from the cross of T317 (*Mdh1-n y20*) × T261 (*k2 Mdh1-n*) (Table 2). The population consisted of 84 $F_2$ individuals derived from a single $F_1$ seed. The genotype of each $F_2$ individual was determined by $F_{2:3}$ family analyses. The $F_{2:3}$ segregation ratio for both the *k2* and *y20* genes fit a 1 wild-type homozygote:2 heterozygotes:1 recessive homozygote ratio with $\chi^2 = 0.86$, $P = 0.65$ and $\chi^2 = 1.21$, $P = 0.55$, respectively (Table 7).

Seven out of 15 SSR markers detected polymorphisms between T317 and T261. Satt314, Satt253, and

Satt279 were dominant markers since they could be amplified only from T317. These three markers segregated in the mapping population as a 3:1 ratio (Table 7). The remaining four markers, Sat_122, Satt302, Satt142, and Satt181, were co-dominant. The segregation pattern was 1 homozygous for the T317 (*Mdh1-n y20*) allele:2 heterozygous:1 homozygous for the T261 (*k2 Mdh1-n*) allele in the mapping population (Table 7). The recombination rate between the *k2* gene and the *y20* gene was 0.6%. The *k2* and *y20* genes were mapped between Satt302 and three co-segregating markers, Satt314, Satt253, and Satt279. The *y20* gene was closer to the three co-segregating markers with 2.8 cM genetic distance between each other, and the *k2* gene was closer to Satt302 with 21.6 cM genetic distance to Satt302 between each other.

### Construction of an integrated map for the five mapping populations

The data from the five mapping population, were loaded into Joinmap 3.0 (Van Olijen and Voorrips 2001) to make an integrated map for the *k2 Mdh1-n y20* chromosomal region. The program first created a map for each mapping population. The gene order of each map was the same as the corresponding map created by Mapmaker 2.0 (data not shown). Then the maps were combined by calculating the mean recombination frequencies and combined LOD scores of each pair of loci from all five mapping populations as described by Stam (1993). The integrated map contained 13 loci in the order; Satt541, Satt469, Sat_122, Satt279, Satt253, Satt314, *Mdh1-n*,*y20*, *k2*, Satt302, Satt142, Satt181, and Satt434 (Fig. 1A).

## Discussion

### Position of the *k2 Mdh1-n y20* chromosomal region on the soybean genetic map

SSRs are PCR-based markers. They can be very easily detected through standard PCR reactions. In soybean, SSR markers are usually multiple alleles, single locus, and inherited as co-dominant alleles (Akkaya et al. 1992). The integrated genetic linkage map (Cregan et al. 1999) includes 606 SSRs loci, 689 restriction fragment length polymorphisms (RFLP), 79 random amplified polymorphic DNAs (RAPD), and 11 amplified fragment length polymorphisms (AFLP) distributed on 20 MLGs which are assumed to represent the 20 chromosomes of the soybean genome, and covers about 2,400 cM. Recently, a new integrated genetic linkage map of the soybean was released (Song et al. 2004). It included a total of 1,849 markers, in which 1,015 markers were SSRs, and covered about 2,500 cM of the chromosome genome. Thus, SSR markers are a good choice for positioning new genes on the soybean linkage map. By using SSR markers, we have mapped several genes, such as female partial-sterile genes *Fsp1* (Kato and Palmer 2003a), *Fsp2*,*Fsp3*, *Fsp4*, and *Fsp5* (Kato and Palmer

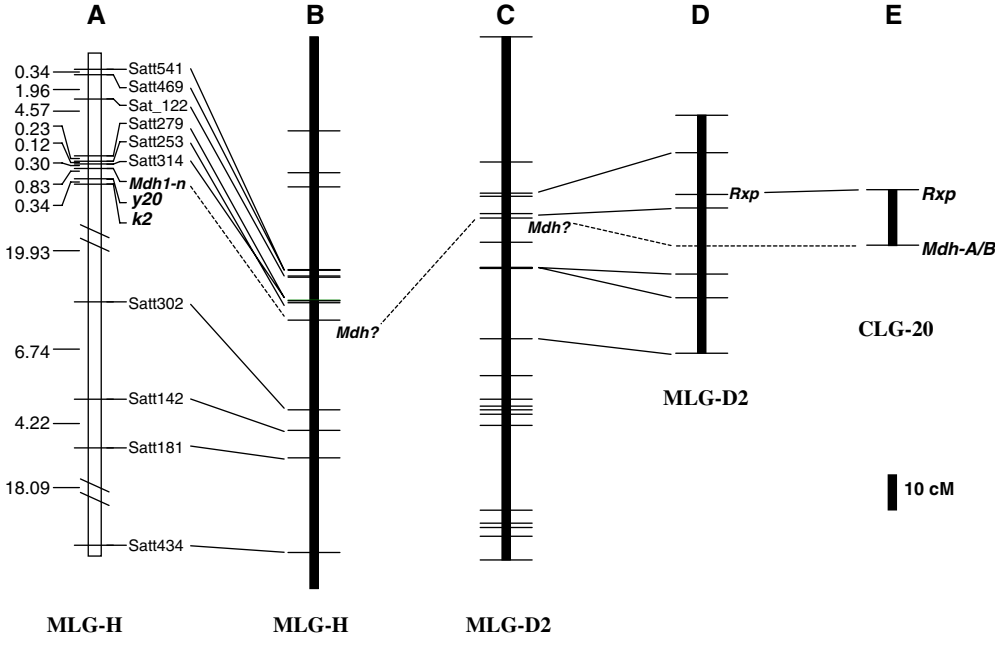


**Fig. 1** Comparison of five linkage maps. Distances are shown in centiMorgans (cM). The alleles between two maps were connected with *straight lines*, and the putative alleles between two maps were connected with *dotted lines*. *A* The integrated molecular map for the *k2 Mdh1-n y20* unstable chromosomal region constructed in this study. *B* MLG H from the USDA/Iowa State University map (Cregan et al. 1999). *C* MLG D2 from the University of Nebraska map (Cregan et al. 1999). *D* Molecular map of BP (bacterial pustule) resistant gene *Rxp* (Narvel et al. 2001). *E* Classical linkage group 20 (Palmer et al. 1992)

2004a), male-sterile and female-sterile gene *st8* (Kato and Palmer 2003b), lethal yellow gene *y18* (Kato and Palmer 2004b), and the mutable gene *w4-m* (Xu and Palmer 2005).

The primary goals of this study were to locate the *k2 Mdh1-n y20* chromosomal region in the soybean integrated genetic linkage map (Song et al. 2004) and to determine the order of the *k2*,*Mdh1-n*, and *y20* loci by SSR markers. Since the *k2, Mdh1-n*, and *y20* loci are very closely linked and no recombination between the *Mdh1-n* and *y20* loci was ever detected in previous genetics studies (Palmer 1984; Chen and Palmer 1996; Chen and Palmer 1998), their order on the chromosome could not be determined using only one mapping population. Thus five mapping populations POP-1, −2, −3, −4, and −5 were constructed to map the *k2* and *Mdh1-n* loci, the *k2* locus, and the *k2* and *y20* loci, respectively. Then data from these five populations were combined to generate one integrated map for the three loci.

The individual maps of the five mapping populations were created by Mapmaker 2.0 (Lander et al. 1987). The results were very consistent. All the overlapping loci in the five mapping populations were in the same order on all the maps. The *k2 Mdh1-n y20* chromosomal region was located on MLG H, and tightly linked with three SSR markers Satt279, Satt253, and Satt314. The *Mdh1-n* and *y20* loci were positioned between the *k2* locus and the SSR markers (Tables 3, 4, 5, 6, and 7). The maps were integrated by Joinmap 3.0 (Van Olijen and Voorrips 2001). The result was basically consistent with the maps from Mapmaker 2.0. Although there was no recombination ever found between the *Mdh1-n* locus and *y20* locus in experiments, the genetic distance between these two loci was estimated as 0.83 cM by Joinmap 3.0 according to the results from five mapping populations. The order of the three loci on the chromosome was determined as *Mdh1-n*,*y20*, and *k2* (Fig. 1A).

## Double chromosomal deletions in T261 (*k2 Mdh1-n*)

Previous genetics analyses (Palmer 1984; Chen and Palmer 1998a) and Southern analysis (Pittig et al. 1994; Imsande et al. 2001) indicated that in soybean, most of the *Mdh1-n* or *Mdh1-n y20* mutations, including T261 (*k2 Mdh1-n*) and T317 (*Mdh1-n y20*), were caused by a chromosome deletion(s). The results from the mapping studies are in agreement with these results. The mapping experiments showed that the *Mdh1-n* and *y20* loci were located between the *k2* locus and Satt279, Satt253, and Satt314 (Tables 3, 4, 5, 6, and 7; Fig. 1A). With the *Mdh1-n* (T261) or *Mdh1-n y20* (T317) mutations (Tables 3, 4, and 7), the genetic distance between the *k2* locus and Satt279 was shorter. It was 1.7 cM, 0.8 cM, or 1.9 cM in mapping populations POP-1, −2, or −5, while 3.0 cM or 2.6 cM in POP-3 or −4 (Tables 3, 4, 5, 6, and 7). As the genetic distance is usually positively related to the physical distance, the shortness of genetic distance between the *k2* locus and Satt279 in POP-1, −2, and −5 was assumed to be the result of the physical deletions at the *Mdh1-n* (T261) or *Mdh1-n y20* (T317) loci.

Besides the putative *Mdh1-n* deletion, MLG H in T261 may have another deleted region that covers three SSR markers (Satt253, Satt279, and Satt314). Firstly, in T261, there were no SSR alleles for these three SSR markers. That is, the SSR alleles could not be amplified from T261. Secondly, these three SSR markers co-segregated in the mapping populations using T261 as a parent (POP-1, −2, and −5 in Tables 3, 4, and 7), but not in the other mapping populations (POP-3 and −4 in Tables 5 and 6). These two phenomena indicated that the chromosomal region which covered Satt253, Satt279, and Satt314 may have been deleted. And, this deleted region was separate from the *Mdh1-n* deletion. If they were on the same deleted chromosomal segment, there would be no recombination between the SSR markers and the *Mdh1-n* locus because genes on the same deleted segment would co-segregate. However, the mapping results showed that there were meiotic crossovers between the SSR markers and the *Mdh1-n* or *Mdh1-n y20* deletions (Tables 1, 7). So, in T261, there were two closely linked chromosome deletions on MLG H. The size of the deleted region covering three SSRs was estimated as approximately 1 Mb according to the positions of the three SSRs on the physical map by Southern Illinois University, Carbondale (http://bioinformatics.siu.edu/cgi-bin/gbrowse/SoyV4R4). The size of the *Mdh1-n* deletion in T261 or the *Mdh1-n y20* deletion in T317 could not be determined in this study.

## Comparisons of the integrated molecular map of the *k2 Mdh1-n y20* chromosomal region and the existing maps of linkage groups with *Mdh* genes

In this study, the final map of the *k2 Mdh1-n y20* chromosomal region was made with Joinmap 3.0, which integrated all the segregation data from the five mapping populations. The integrated map was similar to the MLG H maps of Cregan et al. (1999) and Song et al. (2004). Interestingly, at the chromosomal location where we placed the *k2, Mdh1-n*, and *y20* loci (MLG H), there was a *Mdh* gene already placed on the USDA/Iowa State University map (Cregan et al. 1999) (Fig. 1B). Is the *Mdh* gene on the USDA/Iowa State University map the same as the *Mdh1-n* gene, or is it a different *Mdh* gene that is linked to the *Mdh1-n* gene? To answer these questions, we checked the MDH isozyme migration patterns on a 12% starch gel of the two parental lines used in the mapping population for constructing the original USDA/Iowa State University map, PI 468.916 (*Glycine soja*) and A81-356022 (*Glycine max*) (Shoemaker and Olsen 1993) (Fig. 2, lanes 5 and 6).

Both parents were MDH1 positive, having the first and second bands. However, PI 468.916 (*G. soja*) was MDH-A pattern, having the sixth band counting from the cathode, while A81-356022 (*G. max*), was the MDH-

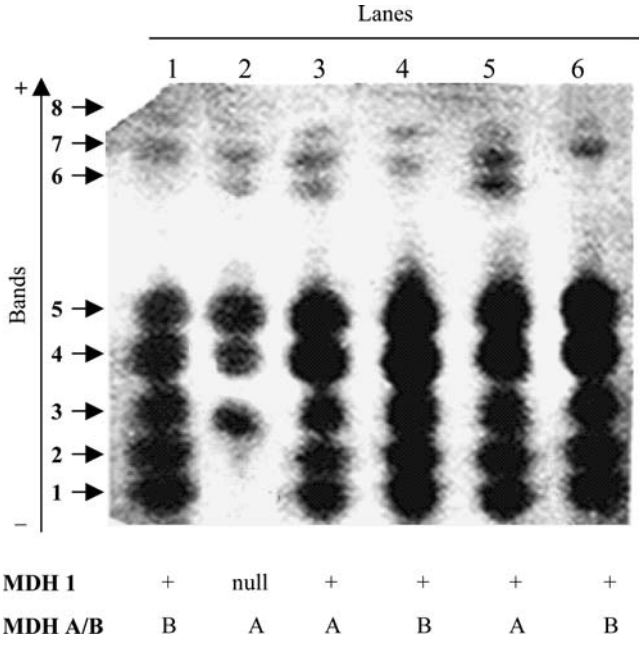


**Fig. 2** The MDH patterns of six soybean lines on a 12% starch gel. Samples were run from the cathode to the anode. The mutant lines with MDH 1 null pattern were missing two MDH bands, band 1 and band 2. The soybean lines with MDH pattern A had band 6, while the soybean lines with MDH pattern B were missing band 6. *Lane 1* Minsoy (*Mdh1Mdh1 Mdh-BMdh-B*); *lane 2* T261 (*Mdh1-nMdh1-n Mdh-AMdh-A*); *lane 3* Harosoy (*Mdh1Mdh1 Mdh-AMdh-A*); *lane 4* Clark(*Mdh1Mdh1 Mdh-BMdh-B*); *lane 5* PI 468.916 (*G. soja*) (*Mdh1Mdh1 Mdh-AMdh-A*); *lane 6* A81-356022 (*G. max*) (*Mdh1Mdh1 Mdh-BMdh-B*)

B pattern, missing the sixth band. MDH-A/MDH-B phenotypes were controlled by a single locus, with the MDH-A variant dominant to the MDH-B variant (Palmer et al. 1992). Here we use *Mdh-A* to represent the allele that conditions the MDH-A pattern, and *Mdh-B* to represent the allele that conditions the MDH-B pattern. Given these phenotypes, the *Mdh* gene on the USDA/Iowa State University map could not be the *Mdh1*/*Mdh1-n* gene but must be the *Mdh-A*/*B* gene, and according to the location it was placed on the USDA/Iowa State University map, it should be closely linked to the *Mdh1-n* gene on MLG H.

However, the same MDH-A/B polymorphism was detected between Harosoy and Clark, two parental lines of the mapping population used by the University of Nebraska (Fig. 2, lanes 3 and 4), but in that population the *Mdh-A*/*B* gene was mapped on MLG D2 (Cregan et al. 1999) (Fig. 1C). Many experiments supported this result. For example, a bacterial pustule resistant gene, *Rxp*, which was identified on the same classical linkage group with the *Mdh-A*/*B* gene (Palmer et al. 1992), was mapped onto MLG D2 (Narvel et al. 2001) with the position close to the *Mdh-A*/*B* gene mapped by the University of Nebraska (Fig. 1C, D).

The reason could be that the *Mdh-A* genes located on different MLGs had different origins. The *Mdh-A* allele positioned on MLG H was from *G. soja*, whereas, the *Mdh-A* allele placed on MLG D2 was from *G. max*. It is inferred that the *Mdh-A* genes in *G. soja* and *G. max* are located on different chromosomes. However, they could be orthologous or paralogous loci, since they condition the same phenotype. But the results of comparative mapping showed that there was no synteny or duplicate loci between MLG D2 and MLG H (Shoemaker et al. 1996; Soybase: http://129.186.26.94/). We think this is possibly due to the absence of RFLP markers around *Mdh-A* genes on MLG D2 and H (Cregan et al. 1999; Song et al. 2004). So, to verify if these two *Mdh-A* genes are orthologous or paralogous loci, either new RFLP markers need to be developed, or more molecular work on gene cloning and sequencing should be done.

**Acknowledgements** The authors would like to thank Dr. R. C. Shoemaker, USDA ARS and the Department of Agronomy, Iowa State University, Ames, IA, for providing the seeds of *G. soja* (PI 468.916) and A81-356022. The authors would also like to thank Dr. D. Grant, USDA ARS and the Department of Agronomy, Iowa State University, Ames, IA, for guidance in constructing maps using Mapmaker2.0.

## References

Akkaya MS, Bhagwat AA, Cregan PB (1992) Length polymorphism of simple sequence repeat DNA in soybean. Genetics 132:1131–1139

Amberger LA, Shoemaker RC, Palmer RG (1992) Inheritance of two independent isozyme variants in soybean plants derived from tissue culture. Theor Appl Genet 84:600–607

Burzlaff JD, Palmer RG (1999) Soybean linkage studies: *y18* and *y20*. Soybean Genet Newslett 26 (Online journal). http://www.soygenetics.org/articles/sgn1999-011.html (posted 23 April 1999)

Cardy BJ, Beversdorf WD (1984) A procedure for the starch gel electrophoretic detection of isozymes in soybean [*Glycine max* (L.) Merr.]. Dep Crop Sci Tech Bull 119–8401. Univ of Guelph, ON

Chen XF, Imsande J, Palmer RG (1999) Eight new mutants at the *k2 Mdh1-n y20* chromosomal region in soybean. J Hered 90:339–403

Chen XF, Palmer RG (1996) Inheritance and linkage with the *k2* and *Mdh1-n* loci in soybean. J Hered 87:433–437

Chen XF, Palmer RG (1998a) Recombination and linkage estimation between the *k2* and *Mdh1-n y20* loci in soybean. J Hered 89:488–494

Chen XF, Palmer RG (1998b) Instability at the *k2 Mdh1-n y20* chromosomal region in soybean. Mol Gen Genet 260:309–318

Cregan PB, Jarvik T, Bush A, Shoemaker RC, Lark KG, Kahler A, Kaya N, VanToai T, Lohnes DG, Chung J, Specht JE (1999) An integrated genetic linkage map of the soybean genome. Crop Sci 39:1464–1490

Imsande J, Pittig J, Palmer RG, Wimmer C, Gietl C (2001) Independent spontaneous mitochondrial malate dehydrogenase null mutants in soybean are the result of deletions. J Hered 92:333–338

Kato KK, Palmer RG (2003a) Genetic identification of a female partial-sterile mutant in soybean. Genome 46:128–134

Kato KK, Palmer RG (2003b) Molecular mapping of male-sterile, female-sterile mutant gene (*st8*) in soybean. J Hered 94:425–428

Kato KK, Palmer RG (2004a) Molecular mapping of four ovule lethal mutants in soybean. Theor Appl Genet 108:577–585

Kato KK, Palmer RG (2004b) Duplicate chlorophyll-deficient loci in soybean. Genome 47:190–198

Keim P, Olson TC, Shoemaker RC (1988) A rapid protocol for isolating soybean DNA. Soybean Genet Newslett 15:150–152

Kosambi DD (1944) The estimation of map distance from recombination values. Ann Eugen 12:172–175

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci USA 88:9828–9832

Narvel JM, Jakkula LR, Phillips DV, Wang T, Lee SH, Boerma HR (2001) Molecular mapping of *Rxp* conditioning reaction to bacterial pustule in soybean. J Hered 92:267–270

Palmer RG (1984) Pleiotropy or close linkage of two mutants in soybeans. J Hered 75:445–447

Palmer RG, Burzlaff JD, Shoemaker RC (2000) Genetic analyses of two independent chlorophyll-deficient mutants identified among the progeny of a single chimeric foliage soybean plant. J Hered 91:297–303

Palmer RG, Hedges BR, Benavente RS, Groose RW (1989) The *w4* mutable line in soybean. Dev Genet 10:542–551

Palmer RG, Lim SM, Hedges BR (1992) Testing for linkage between the *Rxp* locus and nine isozyme loci in soybean. Crop Sci 32:681–683

Palmer RG, Pfeiffer TW, Buss GR, Kilen TC (2004) Qualitative genetics. In: Boerma HR, Specht JE (eds) Soybeans: improvement, production, and uses, 3rd edn. ASA, CSSA, SSSA, Madison, pp 137–214

Pittig J, Guggenberger C, Palmer RG, Imsande J, Gietl C (1994) Molecular analysis of a deletion-prone chromosome segment in soybean [abstract 67]. Fourth international congress of plant molecular biology. International Society of Plant Molecular Biology, Amsterdam

Shoemaker RC, Olson TC (1993) Molecular linkage map of soybean (*Glycine max* L. Merr.). In: O'Brien SJ (ed) Genetic maps: locus maps of complex genomes. Cold Spring Harbor Laboratory Press, New York, pp 6.131–6.138

Shoemaker RC, Polzin K, Labate J, Specht J, Brummer EC, Olson T, Young N, Concibido V, Wilcox J, Tamulonis JP, Kochert G, Boerma HR (1996) Genome duplication in soybean (*Glycine* subgenus *soja*). Genetics 144:329–339

Song QJ, Marek LF, Shoemaker RC, Lark KG, Concibido VC, Delannay X, Specht JE, Cregan PB (2004) A new integrated genetic linkage map of the soybean. Theor Appl Genet 109:122–128

Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant J 3:739–744

Van Ooijen JW, Voorrips RE (2001) JoinMap® 3.0, software for the calculation of genetic linkage maps. Plant Research International, Wageningen

Xu M, Palmer RG (2005) Genetic analysis and molecular mapping of a pale flower allele at the *W4* locus in soybean. Genome 48:334–340