



| | |
|------------------|---|
| Title | Hybrid Breeding Skewed the Allelic Frequencies of Molecular Variants Derived from the Restorer-of-fertility 1 Locus for Cytoplasmic Male Sterility in Sugar Beet (<i>Beta vulgaris</i> L.) |
| Author(s) | Taguchi, Kazunori; Hiyama, Hajime; Yui-Kurino, Rika; Muramatsu, Aki; Mikami, Tetsuo; Kubo, Tomohiko |
| Citation | Crop Science, 54(4), 1407-1412 https://doi.org/10.2135/cropsci2013.11.0731 |
| Issue Date | 2014 |
| Doc URL | http://hdl.handle.net/2115/56484 |
| Type | article (author version) |
| File Information | Taguchi.pdf |



[Instructions for use](#)

1 **Title: Hybrid breeding skewed the allelic frequencies of molecular variants derived**
2 **from the *restorer-of-fertility 1* locus for cytoplasmic male sterility in sugar beet**
3 **(*Beta vulgaris* L.)**

4
5 *Authors:* Kazunori Taguchi¹, Hajime Hiyama², Rika Yui-Kurino², Aki Muramatsu²,
6 Tetsuo Mikami², Tomohiko Kubo²

7
8 *Affiliations:*

9 ¹Hokkaido Agricultural Research Center (HARC), National Agriculture and Food
10 Research Organization (NARO), Memuro, Hokkaido, Japan 082-0052

11 ²Research Faculty of Agriculture, Hokkaido University, Sapporo, Japan, 060-8589

12
13 *Corresponding author:*

14 Kazunori Taguchi, NARO/HARC, Memuro, Hokkaido, Japan 082-0052

15 ktaguchi@affrc.go.jp

16 tel, +81-11-155-62-9271

17

18

19

20

20 Abstract

21 Hybrid breeding of crops may involve the selection of reproductive traits, such as
22 cytoplasmic male sterility (CMS), whose expression is controlled by cytoplasmic and
23 nuclear genes. Intense selection of a single cytoplasm and the consequent lack of
24 cytoplasmic divergence is a potential danger, the so-called genetic vulnerability.
25 However, little is known about the relationship between hybrid breeding and diversity
26 of nuclear genes that suppress the expression of CMS (*Rf*). Despite the multi-allelic
27 nature of *Rf* at the molecular level, a common *Rf* variant was previously found to
28 predominate in sugar beet maintainer lines that were selected for a specific genotype for
29 propagating the CMS line. The question was raised as to the frequency of the common
30 *Rf* variant before the hybrid-breeding era. As the origin of Japanese maintainer lines can
31 be genealogically traced back to seven non-hybrid cultivars, we investigated the allelic
32 diversity of *Rf* in the seven cultivars using molecular markers. Our results indicated that
33 *Rf* diversity differs among the cultivars but exceeds that of the maintainers in total, and
34 the common variant in the maintainers is infrequent in all the cultivars. Therefore,
35 maintainer selection has involved selecting a small number of *Rf* variants in the founder
36 population in Japan.

37

38

38 Introduction

39

40 Selection results in the altered allelic frequency of specific gene(s) involved in the
41 selected trait (e.g. Yamasaki et al., 2007). Hence, a change in the crop-breeding method
42 may alter the allelic frequency of genes that were less attractive to breeders before the
43 change. For example, hybrid breeding of maize using cytoplasmic male sterility (CMS)
44 resulted in the increased frequency in commercial cultivars of T-type cytoplasm, one of
45 the four cytoplasm of this crop (Duvick, 1965; Levings, 1993). Having >85% of maize
46 in the production bearing the T-type cytoplasm was ended in 1970 by the
47 southern-corn-leaf-blight epidemic that specifically damaged maize with the T-type
48 cytoplasm (Laughnan and Gabay-Laughnan, 1983). On the other hand, the allelic
49 frequency of a nuclear gene that suppresses CMS (termed *restorer of fertility*, *Rf*) has
50 had less attention. Although the classical genetic model provided only two alleles (i.e. a
51 restoring allele and a nonrestoring allele), the molecular variation of *Rf* has suggested
52 that there are multiple alleles of the *Rf* locus (Wang et al., 2006; Kato et al., 2007). How
53 hybrid breeding impacts the allelic diversity of the *Rf* locus and the possible related
54 consequences are unknown.

55 Sugar beet is an out-crossing crop that often exhibits genetic heterogeneity even
56 within a breeding line and is known to show a high heterotic response (heterosis) (De
57 Biaggi and Skaracis, 2005). Because of this, almost all of the current commercial sugar
58 beets are hybrid-based cultivars using CMS, which was first described by Owen (Owen,
59 1942; Bosemark, 2006). Owen proposed that sugar beet CMS is genetically conditioned
60 by a cytoplasmic factor (S) and two recessive genes (*x* and *z*), but he had noticed that
61 expression of sugar beet CMS is sometimes too complex to be explained by this simple
62 genetic model because of the emergence of semi-male sterile phenotypes in the progeny
63 of crossing experiments (Owen, 1945). According to Owen's genetic model, male-sterile
64 sugar beets have an [S]*xxzz* genotype. To propagate CMS lines, maintainer lines that are
65 devoid of the restoring allele but with normal fertile cytoplasm to ensure pollen
66 production (i.e. [N]*xxzz*) have been developed. Identification of the maintainer genotype
67 is the most important process in this hybrid breeding method because the CMS line is

68 produced by repeated backcrossing of the maintainer line to the CMS progenitor
69 (Skaracis and De Biaggi, 2005). In this crop, the maintainer genotype appears to be rare,
70 making maintainer selection very difficult (Bosemark, 2006). Single-cross hybrids of
71 sugar beet are possible, but because elite CMS lines tend to express inbreeding
72 depression, they are usually unsuitable as final seed parents due to lower seed yield and
73 quality (Skaracis and De Biaggi, 2005). To overcome this shortcoming, three-way
74 hybrids have been developed in which the final seed parent is an F1 plant produced
75 between a CMS line and an unrelated maintainer line, and the pollinators are from a
76 population with a broad genetic base (Skaracis and De Biaggi, 2005).

77 Molecular cloning of sugar beet *Rfl* (an allele of the *X* locus) revealed a gene
78 cluster consisting of metalloprotease-like genes (Matsuhira et al., 2012). Moritani et al.
79 (2013) investigated the molecular variation of *Rfl* in Japanese sugar beet lines and
80 found that *Rfl* organization varies among restorer lines that often exhibit copy-number
81 variation of the metalloprotease-like genes. The metalloprotease-like gene is not
82 clustered but is present as a single copy in many maintainer lines, in which the same
83 metalloprotease-like variant is shared as a nonrestoring allele, termed *bvORF20L*.
84 Identification of a common variant raised a question about the frequency of *bvORF20L*
85 in the ancestral population of Japanese maintainer lines. This is an important question
86 because the answer would elucidate the dynamics of diversity and frequency of *Rf*
87 alleles during the transition to hybrid breeding.

88 Japanese maintainer lines received significant genealogical contributions
89 (sometimes ~100%) from seven open-pollinated cultivars (OPCs) that were introduced
90 from the United States or European countries before the 1960s (Taguchi et al., 2006).
91 We hypothesized that the molecular diversity of *Rfl* in these seven OPCs would provide
92 the answer to our question.

93 In this study, we first improved one of the two *Rfl* markers developed by Moritani
94 et al. (2013) for large-scale genotyping. The seven OPCs were examined for their *Rfl*
95 diversity with special interest on the frequency of *bvORF20L*. We discuss the impact of
96 maintainer selection on genetic diversity.

97

98

99 Materials and methods

100

101 Plant materials

102

103 Sugar beet (*Beta vulgaris* L.) cultivars Tmm-1 (originated in the United States),
104 Tmm-14 (US), TA-15 (Poland), TA-27 (England), TA-30 (Poland), TA-36 (Germany),
105 and TA-37 (Sweden) are self-incompatible OPCs, introduced by the Japan Sugar Beet
106 Improvement Foundation (succeeded by HARC-NARO) before the 1960s. Plants of
107 TA-36 were different from those used in Moritani et al. (2013). Seeds were sown in
108 plastic pots filled with vermiculite.

109

110

111 Polymerase chain reaction (PCR), nucleotide sequence analysis, and
112 restriction-endonuclease digestion

113

114 Total cellular DNA was isolated from seedlings by the modified method of Rogers and
115 Bendich (1985). Nucleotide sequences of PCR primers are
116 5'-CAATCTGTGGTGCTGACCAA-3' (T1) and
117 5'-GATTAAAGAGGGCTGCTGAAGCCGAGA-3' (T2). Reactions for s17 PCR
118 contained 0.1 U of Ex-Taq (Takara Bio, Ohtsu, Japan), 0.25 mM of each dGTP, dATP,
119 dTTP and dCTP, 0.2 μ M of each T1 primer and T2 primer, and ~10 ng of total cellular
120 DNA in a 20- μ L solution. The PCR protocol was 94 °C for 5 min, then 35 cycles of 94
121 °C for 1 min, 56 °C for 1 min and 72 °C for 1 min. PCR for 20L-int was detailed in
122 Moritani et al. (2013). PCR products were subcloned into the pBluescript (SK+) vector
123 and sequenced using an ABI3130 Genetic Analyzer (Applied Biosystems, Foster City,
124 CA) or were directly sequenced. Sequence analysis was conducted using GENETYX
125 (GENETYX CORPORATION, Tokyo, Japan) or Sequencher (Hitachi Software
126 Engineering, Tokyo, Japan). Sequence alignment was modified manually. Nucleotide
127 sequence data were deposited in DDBJ/EMBL/GenBank under the following accession

128 numbers: AB830620, AB830621, AB830622, AB830623, and AB830624. For cleaved
129 amplified polymorphic sequence (CAPS) detection, PCR products were digested with
130 *HapII* and *HindIII* (Takara Bio) in a 20- μ L solution containing 10 μ L of PCR-product
131 solution with T buffer (supplied by the manufacturer) at the recommended concentration.
132 The digests were electrophoresed in 1.2%-agarose gels or 5%-polyacrylamide gels.

133

134 Statistical analyses

135

136 Data were analyzed with Fisher's exact test using the website of Gunma University,
137 Japan (<http://aoki2.si.gunma-u.ac.jp/exact/fisher/getpar.html>) (accessed on 23rd July,
138 2013). Confidence intervals were calculated according to the Wald confidence interval
139 (e.g. Agresti and Coull, 1998). When the observed ratio was 0, the upper limit of the
140 probability was given by $1-\alpha^{1/n}$, where α is the significant level and n is the size of the
141 sample. Allelic differentiation was examined by GENEPOP version 4.2 (Raymond and
142 Rousset, 1995) available at the website (<http://genepop.curtin.edu.au/index.html>)
143 (accessed on 23rd July, 2013) using the default parameters.

144

145

146 Results and Discussion

147

148 Development of a novel polymorphic marker targeting the *bvORF17-bvORF20L*
149 intergenic region

150

151 *bvORF17* and *bvORF20L* are two genes found within the *Rfl* locus of a maintainer line,
152 in which *bvORF17* is located downstream of *bvORF20L* (Matsuhira et al., 2012) (Fig.
153 1). Marker 17-20L, one of the two DNA markers that tags *bvORF20L* (Moritani et al.,
154 2013), targeted the entire *bvORF17-bvORF20L* intergenic region, which is ~5.5 kbp in
155 length. Difficulty in PCR amplification of a ~5.5-kbp sequence is a shortcoming of
156 marker 17-20L. Therefore, we first improved this marker by confining the polymorphic
157 region in this study.

158 As detailed in Figure S1, a *bvORF17*-upstream region (~1.8 kbp in length, see Fig.
159 1) was identified as the polymorphic region, which is the target region of PCR with
160 primers T1 and T2. Our sequence analysis of this PCR-targeted region from several
161 sugar beet plants revealed five different nucleotide sequences (Fig. 1; see Fig. S2),
162 whose differences can be visualized by agarose gel electrophoresis of the PCR products
163 after digestion with *HapII* and *HindIII* (Fig. 2): pattern '1' generated 1.0- and 0.8-kbp
164 fragments; pattern '2' generated 1.2- and 0.5-kbp fragments; pattern '3' generated a
165 1.3-kbp fragment; pattern '4' generated 1.0- and 0.7-kbp fragments; and pattern '5'
166 generated a 1.7-kbp fragment. Hereafter, the patterns 1 to 5 refer to CAPS (cleaved
167 amplified polymorphic site) marker s17. The target sequence of s17 is single copy in the
168 **sugar beet** genome. Therefore, s17 is a codominant marker.

169

170

171 Molecular diversity of *Rfl* in cultivars developed in the pre-hybrid breeding era

172

173 To analyze *Rfl* diversity, we used two *Rfl* markers, s17 and 20L-int, the latter of which
174 is another *Rfl* marker targeting the first intron of *bvORF20L* and other variants of the
175 metalloprotease-like genes clustered in the *Rfl* locus (Matsuhira et al., 2012; Moritani et
176 al., 2013). Patterns of 20L-int are denoted as S, LS, and L, designations that indicate the
177 presence of a short intron, the presence of both a short intron and a long intron, and the
178 presence of a long intron, respectively (Moritani et al., 2013). Note that types of 20L-int
179 cannot be converted into genotypes because of complex gene clustering (Moritani et al.,
180 2013). Plants homozygous for *bvORF20L* are expected to exhibit pattern-4 homozygous
181 for s17 and type L for 20L-int. Hereafter, we refer to this marker type as 44/L. We
182 genotyped plants that had been inferred to be homozygous for *bvORF20L* by Moritani
183 et al., (2013), and confirmed their marker type as 44/L (Table S1).

184

185 A total of 334 plants from the seven OPCs that significantly contributed to
186 Japanese maintainer lines were analyzed (Table 1). The seven OPCs were developed by
187 mass selection, an orthodox breeding method for sugar beet. Of the seven OPCs, seeds
of Tmm-14 germinated poorly, perhaps due to inadequate preservation. By analyzing

188 s17 and 20L-int, we found a total of 16 marker types. Among the seven OPCs, marker
189 types were diversified in TA-30 (14 types) and TA-37 (11 types), whereas the number of
190 marker types was small in Tmm-1 (4 types). The number of marker types in Tmm-14
191 was also four, but this may be due to the small number of samples.

192 We compared the number of observed marker-types with those of the 22 Japanese
193 maintainer lines analyzed in Moritani et al. (2013), where the 17-20L marker was used
194 instead of s17. Marker pattern 'a', pattern 'c', and pattern 'd' of 17-20L correspond to
195 pattern 1, pattern 3, and pattern 4 of s17, respectively (see the legend of Fig. S1). As
196 described in the legend of Fig. S1, pattern 'b' of 17-20L includes pattern 2 and pattern 5
197 of s17. We reanalyzed the plants previously genotyped as having pattern b of 17-20L in
198 Moritani et al. (2013). Using the s17 marker, we found that all the pattern b alleles
199 corresponded to pattern 5 (Table S1). Therefore, the total number of marker types in the
200 22 Japanese maintainer lines is six (33/S, 34/LS, 35/LS, 44/L, 45/LS and 55/LS;
201 summarized in Table 1). The most diversified line exhibited four marker types (one line),
202 whereas one line had three types, five lines had two types and 14 lines had one type;
203 hence, only one or two marker types occurred in 19 of 22 lines (Table S1), confirming
204 the reduced allelic diversity in the maintainer lines.

205 In the seven OPCs, the frequency of marker type 44/L is 0 to 0.06 (see Table 1).
206 We tested the null hypothesis that the frequency of 44/L is the same among the seven
207 OPCs. Based on the presence/absence of a short intron in the metalloprotease-like gene
208 of the *Rfl* locus, **sugar beet** plants can be divided into two classes, those without the
209 short intron (44/L plants) and those with the short intron (the other marker-type plants)
210 (Table S2). Differences in the occurrence of plants without the short intron among the
211 seven OPCs were statistically marginal (Fisher's exact test; $p=0.054$) (Table S2). Note
212 that frequency of the homozygote for *bvORF20L* in the 22 Japanese maintainer lines
213 was 0.68 on average (see Table 1) and, in 15 out of the 22 lines, reached 0.95–1.00
214 (Table S1).

215 Of the seven OPCs, 44/L plants were found only in the TA-36 line; however, it is
216 unlikely that TA-36 is the only 44/L source for Japanese maintainer lines because some
217 maintainer lines that are genealogically independent of TA-36 have the 44/L marker

218 type. For example, plants of NK-252mm-O and NK-310mm-O exclusively had 44/L
219 (Table S1), but their genealogy indicates no contribution from TA-36; according to
220 Taguchi et al. (2006), NK-252mm-O is 100% Tmm-14-origin, and NK-310mm-O is
221 100% Tmm-1-origin. We inferred that the *bvORF20L* allele exists at a low frequency in
222 other OPCs, and subpopulations made from these OPCs can generate 44/L plants by
223 random mating. To test this hypothesis, we examined the s17-allelic frequency in the
224 seven OPCs, with particular interest in the frequency of pattern 4, which is linked to
225 *bvORF20L* (Table 2). The seven OPCs tended to exhibit a high frequency of pattern 5
226 (0.51–0.88, mean=0.67). The pattern-4 frequency varied from 0.00 to 0.17 (mean=0.09),
227 which was not as major as pattern 5. We estimated the 95%-confidence intervals for the
228 frequency of pattern 4 in each OPC as the following: Tmm-1, 0.00–0.03; TA-15,
229 0.06±0.04; TA-27, 0.08±0.05; TA-30, 0.09±0.05; TA-36, 0.17±0.07; TA-37, 0.13±0.07;
230 and Tmm-14, 0.07±0.13. Therefore, the allelic frequency of *bvORF20L* appeared to
231 remain low in the seven OPCs.

232 Variation in the pattern-4 frequency among the seven OPCs could be explained by
233 something unrelated to *Rfl*, for example, genetic drift or a hitchhiking effect,
234 considering that each OPC was developed by independent mass selection. If so, allelic
235 differentiation of *Rfl* among the seven OPCs would be expected. Allelic differentiation
236 of s17 for each OPC pair was analyzed by GENEPOP (Table 3). The difference between
237 TA-30 and TA-37 was not significant ($p=0.12$). Also, five out of the six OPC pairs
238 involving Tmm-14 were not significantly different ($p=0.08$ – 0.50), but this may be due
239 to the small sample size of Tmm-14. Differences of the other OPC pairs were significant
240 ($p=0.00$ – 0.02). Therefore, it is likely that the allelic frequency of *Rfl* had more or less
241 fluctuated before the 1960s.

242

243

244 Conclusion

245

246 All the maintainer lines were developed after the launch of hybrid breeding in Japan.
247 The seven OPCs used in this study can be considered as ancestral populations because

248 of their significant contribution to the genealogy of Japanese maintainer lines (i.e. many
249 Japanese maintainer lines originated from the seven OPCs). The allelic frequency of
250 *bvORF20L* in the seven OPCs remained low but more or less fluctuated during the mass
251 selection for each of the seven OPCs. Marker types and alleles that are missing from the
252 maintainer lines were found in the seven OPCs. In contrast, an increase in the frequency
253 of *bvORF20L* is apparent in the Japanese maintainer lines, and notably, this increase in
254 *bvORF20L* frequency has occurred in at least two genealogically independent
255 maintainer lines (e.g. NK-252mm-O is exclusively derived from Tmm-14, whereas
256 NK-310mm-O is from Tmm-1 [Taguchi et al., 2006]). Therefore, breeders have
257 preferred *bvORF20L* during maintainer selection, and the consequence is predominance
258 of the 44/L-marker type in Japanese maintainer lines.

259 Three implications are deduced from this conclusion. First, because *bvORF20L*
260 was repeatedly selected from independent populations, the repertoire of nonrestoring
261 alleles of *Rfl* is limited in **sugar beet** populations. We are currently seeking another
262 nonrestoring alleles from *B. vulgaris* genetic resources. Second, as was seen in the case
263 of *bvORF20L*, the allelic frequency of the nonrestoring allele can fluctuate by mass
264 selection. As such, the nonrestoring allele may be lost from some subpopulations, from
265 which no maintainer line would be expected. This scenario can explain the experience
266 of one of the co-authors in that a population appeared to lack any maintainer genotype
267 (K. Taguchi, unpublished data). Third, the selection intensity for a maintainer genotype
268 is very high because a specific, infrequent allele is selected among the vast majority of
269 multiple alleles. Consequently, the genetic diversity of other loci, especially those linked
270 to *Rfl*, could be reduced after maintainer selection. This concern can be negligible in the
271 commercial varieties if the genetic heterogeneity of the three-way hybrid is warranted
272 by using pollinators with a wide genetic base. However, care should be taken if the
273 pollinator is a maintainer genotype in the case of a single-cross hybrid or three-way
274 hybrids, in which genetic homogeneity around the *X* locus, on which *Rfl* is located, is
275 expected. We speculate that genetic linkage between the *X* locus and several disease
276 resistance genes (e.g. Taguchi et al., 2011) may be a point of concern. It may be
277 important to monitor the genetic diversity of maintainer lines to prevent genetic

278 vulnerability.

279

280

281 Acknowledgements

282 This work was supported in part by: Grants-in-Aid for Scientific Research from the
283 Ministry of Education, Culture, Sports, Science, and Technology, Japan; Grants-in-Aid
284 for Scientific Research from the Program for Promotion of Basic and Applied
285 Researches for Innovations in Bio-oriented Industry (BRAIN); and Grants-in-Aid for
286 Scientific technique research promotion program for agriculture, forestry, fisheries and
287 food industry, Japan.

288

289

290 References

291 Agresti, A., and Coull, B.A. 1998. Approximate is better than "exact" for interval
292 estimation of binomial proportions. *Am. Stat.* 52: 119-126.

293 Bosemark, N.O. 2006. Genetics and Breeding. *In Sugar Beet. Edited by A.P. Draycott.*
294 Blackwell Publishing, Oxford, UK. pp. 50-88.

295 De Biaggi, M., and Skaracis, G.N. 2005. Selection methods. *In Genetics and Breeding*
296 *of Sugar Beet. Edited by E. Biancardi, L.G. Campbell, G.N. Skaracis, and M. De*
297 *Biaggi.* Science Publishers, Enfield, NH. pp. 169-191.

298 Duvick, D.N. 1965. Cytoplasmic pollen sterility in corn. *Adv. Genet.* 13: 1-56.

299 Kato, H., Tezuka, K., Feng, Y.Y., Kawamoto, T., Takahashi, H., Mori, K., and Akagi, H.
300 2007. Structural diversity and evolution of the *Rf-1* locus in the genus *Oryza*.
301 *Heredity* 99: 516-524.

302 Laughnan, J.R., and Gabay-Laughnan, S. 1983. Cytoplasmic male sterility in maize.
303 *Ann. Rev. Genet.* 17: 27-48.

304 Levings, C.S.III 1993. Thoughts on cytoplasmic male sterility in *cms-T* maize. *Plant*
305 *Cell* 5: 1285-1290.

306 Matsuhira, H., Kagami, H., Kurata, M., Kitazaki, K., Matsunaga, M., Hamaguchi, Y.,
307 Hagihara, E., Ueda, M., Harada, M., Muramatsu, A., Yui-Kurino, R., Taguchi, K.,

- 308 Tamagake, H., Mikami, T., and Kubo, T. 2012. Unusual and typical features of a
309 novel *restorer-of-fertility* gene of sugar beet (*Beta vulgaris* L.). *Genetics* 192:
310 1347-1358.
- 311 Moritani, M., Taguchi, K., Kitazaki, K., Matsuhira, H., Katsuyama, T., Mikami, T., and
312 Kubo, T. 2013. Identification of the predominant nonrestoring allele for
313 Owen-type cytoplasmic male sterility in sugar beet (*Beta vulgaris* L.):
314 development of molecular markers for the maintainer genotype. *Mol. Breed.* 32:
315 91-100.
- 316 Owen, F.V. 1942. Male sterility in sugar beets produced by complementary effects of
317 cytoplasmic and mendelian inheritance. *Am. J. Bot.* 29: 692.
- 318 Owen, F.V. 1945. Cytoplasmically inherited male-sterility in sugar beets. *J. Agr. Res.*
319 71: 423-440.
- 320 Raymond, M., and Rousset, F. 1995. GENEPOP Version 1.2: population genetics
321 software for exact tests and ecumenicism. *J. Hered.* 86: 248-249.
- 322 Rogers, S.O., and Bendich, A.J. 1985. Extraction of DNA from milligram amounts of
323 fresh, herbarium and mummified plant tissues. *Plant Mol. Biol.* 5: 69-76.
- 324 Skaracis, G.N, and De Biaggi, M. 2005. Production of commercial varieties. *In Genetics*
325 *and Breeding of Sugar Beet. Edited by E. Biancardi, L.G. Campbell, G.N.*
326 *Skaracis, and M. De Biaggi. Science Publishers, Enfield, NH. pp.191-220.*
- 327 Taguchi, K., Nakatsuka, K., Takahashi, H., Okazaki, K., and Yoshida, T. 2006.
328 Relationship between the coefficient of parentage and sugar yield in sugar beet F1
329 hybrid (in Japanese). *Breed. Res.* 8: 151-159.
- 330 Taguchi, K., Kubo, T., Takahashi, H., and Abe, H. 2011. Identification and precise
331 mapping of resistant QTLs of *Cercospora* leaf spot resistance in sugar beet (*Beta*
332 *vulgaris* L.). *G3.* 1: 283-291.
- 333 Wang, Z., Zou, Y., Li, X., Zhang Q., Chen, L., Wu, H., Su, D., Chen, Y., Guo, J., Luo,
334 D., Long, Y., Zhong, Y., and Liu, Y.-G. 2006. Cytoplasmic male sterility of rice
335 with Boro II cytoplasm is caused by a cytotoxic peptide and is restored by two
336 related PPR motif genes via distinct modes of mRNA silencing. *Plant Cell* 18:
337 676-687.

338 Yamasaki, M., Wright, S.I., and McMullen, M.D. 2007. Genomic screening for artificial
339 selection during domestication and improvement in maize. *Ann. Bot.* 100:
340 967-973.
341

Table 1 Marker types and numbers of plants in seven open-pollinated cultivars (OPCs) and Japanese maintainer lines.

| Name of cultivars | Type of s17/20L-int | | | | | | | | | | | | | | | Total | |
|---|---------------------|------|-------|-------|------|-------|------|------|-------|-------|------|-------|------|-------|------|-------|-------|
| | 11/S | 13/S | 13/LS | 14/LS | 15/S | 15/LS | 25/S | 33/S | 33/LS | 34/LS | 35/S | 35/LS | 44/L | 45/LS | 55/S | | 55/LS |
| Tmm-1 | 2 | | | | 7 | | | | | | 12 | | | | 30 | | 51 |
| TA-15 | | | | | 3 | | | | | | 5 | 1 | | 8 | 31 | 24 | 33 |
| TA-27 | 2 | | | | 8 | | 1 | | | | 2 | | | 8 | 20 | 8 | 49 |
| TA-30 | 3 | 2 | 3 | 2 | 6 | 2 | | 2 | 6 | 4 | 7 | 2 | | 4 | 8 | 3 | 54 |
| TA-36 | | | | | | | | 1 | 2 | 2 | 7 | 6 | 3 | 10 | 21 | | 52 |
| TA-37 | 2 | 4 | | 1 | 9 | 1 | | | | 5 | 7 | 1 | | 7 | 11 | 1 | 49 |
| Tmm-14 | | | | | 1 | | | | | | 4 | | | 1 | | 1 | 7 |
| Summary of seven OPCs | 9 | 6 | 3 | 3 | 34 | 3 | 1 | 3 | 8 | 11 | 44 | 10 | 3 | 38 | 121 | 37 | 334 |
| Summary of 22 maintainer lines ¹ | | | | | | | | 6 | | 23 | | 5 | 245 | 31 | | 52 | 362 |

¹Based on Moritani et al. (2013) and this study. For details, see Table S1.

Table 2 Number of alleles and allelic frequencies (in brackets) of s17 in OPCs

| Name of OPC | Alleles of s17 | | | | | Total |
|-------------|----------------|-------------|--------------|--------------|---------------|------------|
| | 1 | 2 | 3 | 4 | 5 | |
| Tmm-1 | 11 [0.11] | 0 [0] | 12 [0.12] | 0 [0] | 79 [0.77] | 102 [1] |
| TA-15 | 3 [0.02] | 0 [0] | 6 [0.04] | 8 [0.06] | 127 [0.88] | 144 [1] |
| TA-27 | 12 [0.12] | 1 [0.01] | 2 [0.02] | 8 [0.08] | 75 [0.77] | 98 [1] |
| TA-30 | 21 [0.19] | 0 [0] | 34 [0.31] | 10 [0.09] | 43 [0.41] | 108 [1] |
| TA-36 | 0 [0] | 0 [0] | 21 [0.20] | 18 [0.17] | 65 [0.63] | 104 [1] |
| TA-37 | 19 [0.19] | 0 [0] | 17 [0.17] | 13 [0.13] | 49 [0.51] | 98 [1] |
| Tmm-14 | 1 [0.07] | 0 [0] | 4 [0.29] | 1 [0.07] | 8 [0.57] | 14 [1] |
| Total | 67 [0.10] | 1 [0.00] | 96 [0.14] | 58 [0.09] | 446 [0.67] | 668 [1] |

Table 3 Probabilities of allelic differentiation for each pair of seven OPCs calculated by GENEPOP¹

| | Tmm-1 | TA-15 | TA-27 | TA-30 | TA-36 | TA-37 | Tmm-14 |
|--------|-------------------|-------------------|-------------------|-------------------|-------------------|-------|--------|
| Tmm-1 | - | - | - | - | - | - | - |
| TA-15 | 0.00 ² | - | - | - | - | - | - |
| TA-27 | 0.00 ² | 0.02 ² | - | - | - | - | - |
| TA-30 | 0.00 ² | 0.00 ² | 0.00 ² | - | - | - | - |
| TA-36 | 0.00 ² | 0.00 ² | 0.00 ² | 0.00 ² | - | - | - |
| TA-37 | 0.00 ² | 0.00 ² | 0.00 ² | 0.12 | 0.00 ² | - | - |
| Tmm-14 | 0.09 | 0.01 ² | 0.08 | 0.60 | 0.17 | 0.50 | - |

¹The exact G test was used.

²Columns with values less than 0.05 are shaded.

Figure captions

Figure 1 Organization of the *bvORF17-bvORF20L* intergenic region and target region of s17. *Panel a.* Arrangement of *bvORF17* and *bvORF20L* (adopted from Matsuhira et al., 2012). Arrows indicate transcription direction. The target region of s17 is indicated by a bracket. *Panel b.* Organizational comparison of PCR fragments generating patterns 1 to 5 for s17 (for details, see Fig. S2). Wedges indicate deletions. *HindIII* restriction site and *HapII* restriction sites are shown by filled and open triangles, respectively. Scale bar is shown below.

Figure 2 Fragment patterns of s17. PCR products from plants with 1/1, 2/2, 3/3, 4/4, and 5/5 marker types were digested with *HapII* and *HindIII*, and subsequently electrophoresed in a 1.2% agarose gel. M1 and M2 are lambda DNA digested with *StyI* and pBluescript (SK+) digested with *HapII*, respectively. Size markers are shown in kbp.

a.

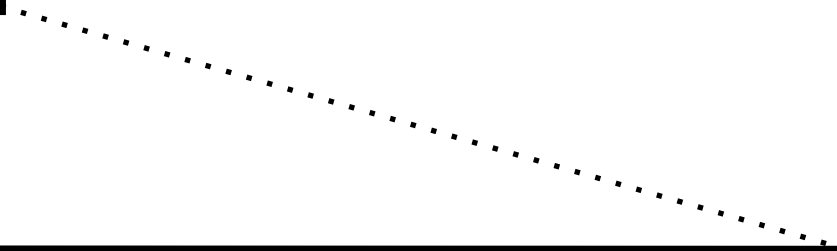
bvORF17



bvORF20L



s17 (~1.8 kbp)



b.

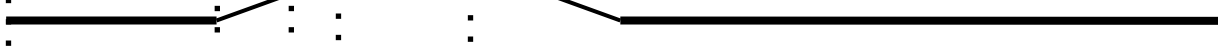
Pattern 1



Pattern 2



Pattern 3



Pattern 4



Pattern 5



300 bp



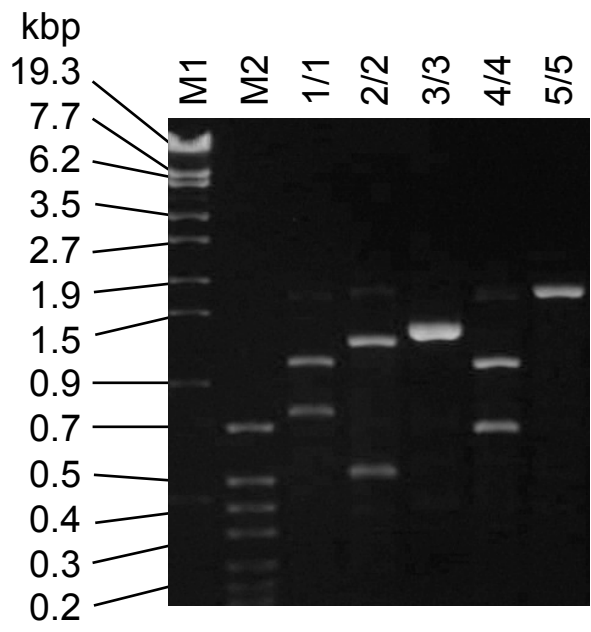


Table S1 Marker-type coordinates and number of plants¹

| | $\frac{cc(S)^2}{33/S^3}$ | $\frac{cd(LS)^2}{34/LS^3}$ | $\frac{bb(LS)^2}{22/LS^3 \ 55LS^3}$ | | $\frac{bc(LS)^2}{23/LS^3 \ 35/LS^3}$ | | $\frac{bd(LS)^2}{24/LS^3 \ 45/LS^3}$ | | $\frac{dd(L)^2}{44/L^3}$ | Total |
|--------------|--------------------------|----------------------------|-------------------------------------|----|--------------------------------------|---|--------------------------------------|---|--------------------------|-------|
| | | | | | | | | | | |
| NK-169mm-O | | | | | | | | | 28 | 28 |
| NK-172BRmm-O | | | | | | | 8 | | 12 | 20 |
| NK-183BRmm-O | | | | | | | | | 4 | 4 |
| NK-204mm-O | | | | | | | | | 4 | 4 |
| NK-205mm-O | | | | | | | | | 4 | 4 |
| NK-206mm-O | | | | | | | | | 6 | 6 |
| NK-208mm-O | 4 | | | | | | | | 1 | 5 |
| NK-219mm-O | | | | 47 | | | | | | 47 |
| NK-222BRmm-O | | | | 4 | | 5 | | 9 | 2 | 20 |
| NK-226BRmm-O | | | | | | | | | 7 | 7 |
| NK-252mm-O | | | | | | | | | 7 | 7 |
| NK-280mm-O | | | | | | | | | 10 | 10 |
| NK-294mm-O | | | | | | | | | 8 | 8 |
| NK-296mm-O | | | | | | | | | 20 | 20 |
| NK-300mm-O | | | | | | | | 1 | 19 | 20 |
| NK-310mm-O | | | | | | | | | 27 | 27 |
| NK-325mm-O | | 1 | | | | | | | 19 | 20 |
| NK-341mm-O | | | | 1 | | | | 8 | 11 | 20 |
| NK-343mm-O | | | | | | | | 5 | 15 | 20 |
| TA-33BB-O | | | | | | | | | 11 | 11 |
| TK-76mm-O | | | | | | | | | 9 | 9 |
| TK-81mm-O | 2 | 22 | | | | | | | 21 | 45 |

¹Based on Moritani et al. (2013) and this study.

²Marker type of 17-20L and 20L-int.

³Marker type of s17 and 20L-int.

Table S2 Number of plants containing a short intron in the metalloprotease-like gene in the *Rf1* locus.

| Name of open pollinated cultivar | With short intron ¹ | Without short intron ² | Total |
|-------------------------------------|--------------------------------|-----------------------------------|-------|
| Tmm-1 | 51 | 0 | 51 |
| TA-15 | 72 | 0 | 72 |
| TA-27 | 49 | 0 | 49 |
| TA-30 | 54 | 0 | 54 |
| TA-36 | 49 | 3 | 52 |
| TA-37 | 49 | 0 | 49 |
| Tmm-14 | 7 | 0 | 7 |
| Total | 331 | 3 | 334 |

¹20L-int types of S and LS.

²20L-int types of L.

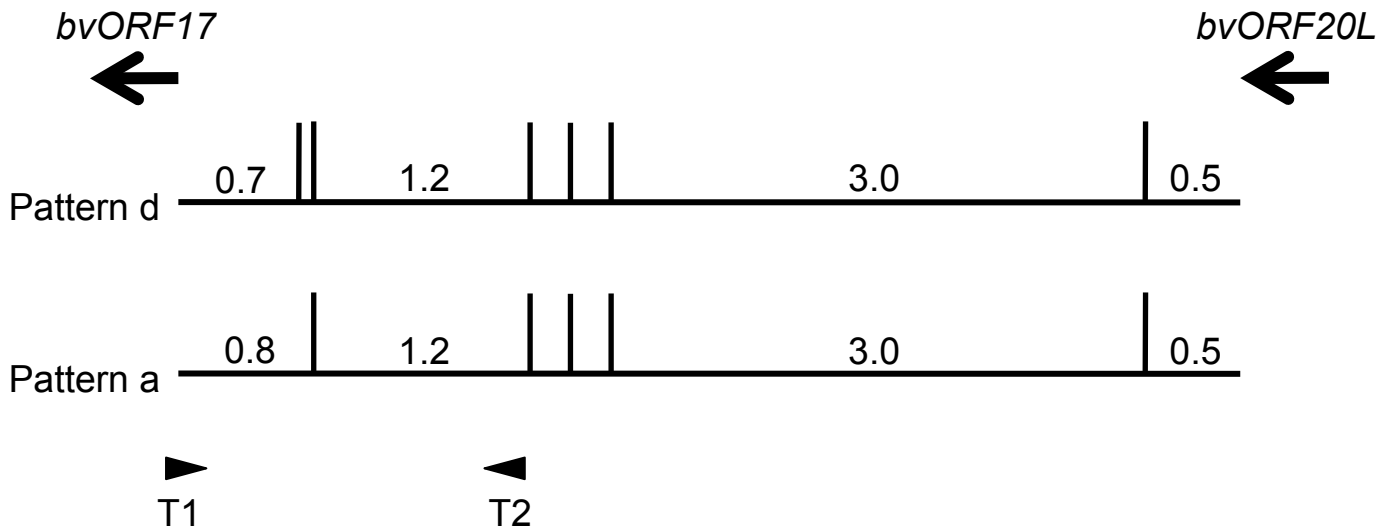


Fig. S1 *HapII*-restriction map of the *bvORF17*-*bvORF20L* intergenic region (the target of 17-20L). The map is drawn based on the nucleotide sequences of DDBJ/EMBL/GenBank accession numbers AB646133 and AB646136 that correspond to NK-198 (pattern a) and TK-81mm-O (pattern d), respectively. *HapII*-restriction sites are shown by vertical lines. Fragment sizes on agarose gel are specified (kbp). Arrows indicate the arrangement of *bvORF17* and *bvORF20L*. Positions of the T1 and T2 primers are shown by arrowheads.

Marker 17-20L is known to generate four patterns, named 'a' to 'd' (Moritani et al. 2013). We identified two nucleotide sequences of the *bvORF17*-*bvORF20L*-intergenic region from nucleotide-sequence entries of DDBJ/EMBL/GenBank accession numbers AB646133 and AB646136 that correspond to pattern a and pattern d, respectively. Because the polymorphism of 17-20L can be detected after *HapII* digestion of the *bvORF17*-*bvORF20L*-PCR products, we mapped the *HapII* sites in the *bvORF17*-*bvORF20L* intergenic region (see above). This analysis revealed that the polymorphic *HapII* site is located 0.7-kbp upstream of *bvORF17*. In addition, because the other two patterns (pattern 'b' and pattern 'c') share 3.0- and 0.5-kbp *HapII* fragments with pattern a and pattern d, the polymorphic region of pattern b and pattern c appeared to be located upstream of *bvORF17* (see above).

We compared the nucleotide sequences of the *bvORF17* upstream regions (~1.8 kbp) that were PCR amplified with primers T1 and T2 from DNA samples of sugar-beet plants known to exhibit patterns a to d (Fig. S2). Since we obtained two different nucleotide sequences from pattern b plants, a total of five different nucleotide sequences were identified that were named patterns '1' to '5' (Fig. S2 and Fig. 1). Patterns 1, 3 and 4 were derived from plants of patterns a, c and d, respectively. Patterns 2 and 5, both were derived from pattern-b plants, differing in single-nucleotide substitutions and indels, one of which generates a polymorphic *HindIII* site (Fig. S2 and Fig. 1).

1_NK-198 GATTAAAGAGGGCTGCTGAAGCCGAGA AAGAAGATAAGAAGTAGAGAATAATATCTTGAT 60
2_NK305 GATTAAAGAGGGCTGCTGAAGCCGAGA AAGAAGACAAGAAGTAGAGAAAACATTTTGAT 60
3_TK81-0 GATTAAAGAGGGCTGCTGAAGCCGAGA AAGAAGACAAGAAGTAGAGAATAATATCTTGAT 60
4_TA33-0 GATTAAAGAGGGCTGCTGAAGCCGAGA AAGAAGATAAGAAGTAGAGAATAATATCTTGAT 60
5_NK219-0 GATTAAAGAGGGCTGCTGAAGCCGAGA AAGAAGACAAGAAGTAGAGAATAATATCTTGAT 60
* * * *

1_NK-198 ACGACCATTTTTTAAATTTTTACCTCTTTGTTGATCTTGTGTTTTCTTCAGCTGGTGTGTTT 120
2_NK305 ACGACCATTTTTTAAATTTTTACCTCTTTGTTGATCT-GTGTTTGCCTTCAGCTGGTGTGTTT 119
3_TK81-0 ACGACCATTTTTTAAATTTTTACCTCTTTGTTGATCTTGTGTTTGCCTTCAGCTGGTGTGTTT 120
4_TA33-0 ACGACCATTTTTTAAATTTTTACCTCTTTGTTGATCTTGTGTTTGCCTTCAGCTGGTGTGTTT 120
5_NK219-0 ACGACCATTTTTTAAATTTTTACCTCTTTGTTGATCTTGTGTTTGYTTCAGCTGGTGTGTTT 120
* * * **

1_NK-198 TCTATACTCTTTTT-ACTGTGAGTTTCATATCCATGATTTTCAGAACTGGTATTTATTTTA 179
2_NK305 TTTATACTCTTTTT-ACTGTGAGTTTCATATCCATGATTTTCAGAACTGGTATTTATTTTA 178
3_TK81-0 TTAATACTTTTTTTACTGTGAATTTTCATATCCATGATTTTCAGAACTGGTATTTATTTTA 180
4_TA33-0 TCTATACTCTTTTT-ACTGTGAGTTTCATATCCATGATTTTCAGAACTGGTATTTATTTTA 179
5_NK219-0 T-TATACTCTTTTT-ACTGTGAGTTTCATATCCATGATTTTCAGAACTGGTATTTATTTTA 178
* * * * *

1_NK-198 TTGGGTGTTTTAAATGAAGAATGATTTGTCAGGTGCGGATGTGTTTCATGTGATGTTGAGCAA 239
2_NK305 TCAGGTGTTTTAAATGAATAATGATTTGTCAGGTGCGGATGTGTTTATGTGATGTTGAGCAA 238
3_TK81-0 TCGGGTGTTTTTAAATGAATAATGATTTGTCAGGTGCGGATGTGTTTATGTGATGTTGAGCAA 240
4_TA33-0 TCGGGTGTTTTTAAATGAAGAATGATTTGTCAGGTGCGGATGTGTTTATGTGATGTTGAGCAA 239
5_NK219-0 TCGGGTGTTTTTAAATGAATAATGATTTGTCAGGTGCGGATGTGTTTATGTGATGTTGAGCAA 238
* * * * *

1_NK-198 AAGAAGGTAGAGTTGAATATATTTATGAAATTTATCCATTTGTTTTTCATTTTGCTGTCTTA 299
2_NK305 AAGAAGGTAGAGTTGAATATATTTATGAAATTTATCCATTTGTTTTTCATTTTGTTGTCTTA 298
3_TK81-0 AAGAAGGTAGAGTTGAATATATTTATGAAATTTATCCATTTGTTTTTCATTTTGTTGTCTTA 300
4_TA33-0 AAGAAGGTAGAGTTGAATATATTTATGAAATTTATCCATTTGTTTTTCATTTTGCTGTCTTA 299
5_NK219-0 AAGAAGGTAGAGTTGAATATATTTATGAAATTTATCCATTTGTTTTTCATTTTGTTGTCTTA 298
* * * *

1_NK-198 CAATTTTAACTT----- 314
2_NK305 CAACTTTTAAATGCTTCCTCCATCTCGTTTTAAACGCAACAATTGGAAATTTTATATCTCA 358
3_TK81-0 CAACTTTTAA-----CGT----- 313
4_TA33-0 CAATTTTAACTT----- 314
5_NK219-0 CAACTTTTAAATGCTTCCTCCATCTCGTTTTAAACGCAACAATTGGAAATTTTATATCTCA 358
* *****

1_NK-198 ----- 314
2_NK305 CAAATAATACCCTTTTGTTGCATTTGAAACGAAACGGAGGAAGTACTTTGGTAAAATTTCT 418
3_TK81-0 ----- 313
4_TA33-0 ----- 314
5_NK219-0 CAAATAATACCCTTTTGTTGCATTTGAAACGAAACGGAGGAAGTACTTTGGTAAAATTTCT 418

1_NK-198 ----CGAATTGGAGCTAGGATTTATAACGCTTTTTCTTTTTGAGGCAAGGATTTAC-TTC 369
2_NK305 TTTACGAATTGGAGCTAGGATTTATATCGTTTTTTCTTTTTGAGGAAAGGATTTATATTTG 478
3_TK81-0 ----- 313
4_TA33-0 ----CGAATTGGAGCTAGGATTTATAACGCTTTTTCTTTTTGAGGCAAGGATTTAC-TTC 369
5_NK219-0 TTTACGAATTGGAGCTAGGATTTATATCGTTTTTTCTTTTTGAGGAAAGGATTTATATTTG 478

1_NK-198 ATCC-GTTTCAT-AATACTCGCTACACTTTTCATTATGGGAAGTTTCACATTATTTGCTA 427
2_NK305 ATCTTGATTAATCAAA----- 494
3_TK81-0 ----- 313
4_TA33-0 ATCC-GTTTCAT-AATATTCGCTACACTTTTCATTATGGGAAGTTTCACATTATTTGCTA 427
5_NK219-0 ATTTTGATTTATCGAAA----- 495

1_NK-198 CACTTTTTTTTTGGGCAACAACTCTCTCTATCTCTTTCTCTCACATGGGTCCCTTTTC 487
2_NK305 ----- 494
3_TK81-0 ----- 313
4_TA33-0 CACTTCTTTTTTTGGGCAACAACTCTCTCTATCTCTTTCTCTCACATGGGTCCCTTTTC 489
5_NK219-0 ----- 495

1_NK-198 TTTTACTCTCTATCTCTCTCTTTTCCAAAAATTGCC TAGGACTCTGAATTGAAAAGTAT 547
2_NK305 ----- 494
3_TK81-0 ----- 313
4_TA33-0 TTTTACTCTCTATCTCTCTCTTTTCCAAAAATTGCC TAGGAGTCTGAATTGAAAAGTAT 547
5_NK219-0 ----- 495

1_NK-198 AGCGAATAATATGAAACGGATGAATTTATCAAGTCAAACTGAAAAAGCATGATGGTATA 607
2_NK305 -----GTCAAAACTGAAAAAGCTTGATGGTATA 522
3_TK81-0 ----- 313
4_TA33-0 AGCGAATAATATGAAATGGAGGAATTTATCAAGTCAAACTGAAAAAGCATGATGGTATA 607
5_NK219-0 -----GTGAAAACTGAAAAAGCATGATGGTATA 523

1_NK-198 C-ACTTGCAATTAAAACTCTCAAAACATTATGTGGATTTTGGTGG-GGAATATACTATTG 665
2_NK305 CCACTTGCAATTAAAACTCTCAAAACATTATGTGGATTTTGGCGGGGAATATACTATTG 582
3_TK81-0 ----- 313
4_TA33-0 C-ACTTGCAATTAAAACTCTCAAAACATTATGTGGATTTTGGTGG-GGAATATACTATTG 665
5_NK219-0 CCACTTGCAATTAAAACTCTCAAAACATTATGTGGATTTTGGCGGGAGGAATATACTATTG 583

1_NK-198 CTGAACTAAACTGGCTTATCCTGTGATGAGTTCCTCTATGAAAGTTGACGAATTAAGATGT 725
2_NK305 TTGAACTAAACTGGCTTATCCTGTGATGAGTTCCTCTATGAAAGTTGACGAATTAAGATGT 642
3_TK81-0 ----- 313
4_TA33-0 TTGAACTAAACTGGCTTATCCGGTGATAAGTTCCTCTATGAAAGTTGACGAATTAAGATGT 725
5_NK219-0 TTGAACTAAACTGGCTTATCCTGTGATGAGTTCCTCTATGAAAGTTGACGAATTAAGATGT 643

1_NK-198 GAAATCAACTAAGAAATGAATCTAAGTTGAGGAGTATGCCGGTGATATTAATTAGAATGGA 785
2_NK305 GAAATCAACTAAGAAATGAATCTAAGTTGAGGAGTATGCCGATGATATTAATTAGAATGGA 702
3_TK81-0 ----- 313
4_TA33-0 GAAATCAACTAAGAAATGAATCTAAGTTGAGGAGTATGCCGGTTATATTAATTAGAATGGA 785
5_NK219-0 GAAATCAACTAAGAAATGAATCTAAGTTGAGGAGTATGCCGATGATATTAATTAGAATGGA 703

1_NK-198 ACATATACCAGAAATCAAGTTGGTATCTTAATAATACATGTCAGAAATTTCTTCGACGAAG 845
2_NK305 ACCTATACCAGAAATTTAAGTTGGTATCTTAATAATACATGTCAGAAATTTCTTCGACGAAC 762
3_TK81-0 -----AG-----AGTTGGTATCTTAATAAT----GTCACAAATTTCTTCGACGAAG 354
4_TA33-0 ACATCTACCAGAAATTTAAGTTGGTATCATAATAATACATGTCAAAAATTTCTTCGACGAAG 845
5_NK219-0 ACCTATACCAGAAATTTAAGTTGGTATCTTAATAATACATGCCAGAAATTTCTTCGACGAAG 763
***** * * * * *

1_NK-198 GAAATTTGTAACAAAAATATACTGAAAAATCATGTTGTCTGTAACAGCAACATGATCAAAA 905
2_NK305 GAAATTTATAAACAAAAATCTACTGAAAAATCATGTTGTCTGTAACAGCAACATGATCAAAA 822
3_TK81-0 GAAATTTGTAACAAAAATATACTGAAAAATCATGTTG-CTGTTACAGAAACATGATCAAAA 413
4_TA33-0 GAAATTTGCAAAACAAAAATATACTGAAAAATCATGCTGTCTGTAACAGCAACATGATCAAAA 905
5_NK219-0 GAAATTTGTAACAAAAATATACTGAAAAATCATGTTGTCTGTAACAGCAACATGATCAAAA 823
** * * * * *

1_NK-198 TGGTTCTTACAACGAAATCTCGCAATGTATCTACTCTGAAATTTGTGATATAGGCGTAGA 965
2_NK305 TGGTTCTTACAACAGAAATCTCGCAATGTATCTACTCTGAAATTTGTGATATAGGCGTAGA 882
3_TK81-0 TGGTTCTTACAACAGAAATCTCGCAATGTATCTACTCTGAAATTTGTGATATAGGCGTAGA 473
4_TA33-0 TGGTTCTTACAACAGAAATCTCGCAATGTATCTACTCTGAAATTTGTGATATAGGCGTAGA 965
5_NK219-0 TGGTTCTTACAGCAGAAATCTCGCAATGTATCTACTCTGAAATTTGTGATATAGGCGTAGA 883
* *

1_NK-198 GCAGATGAATCAGTCATTAAGGCTTAGTTCAACTGTTCAACCCACTTTAATTATTAGAG 1025
2_NK305 GCTGATGAATCAGTCATTTAATGCTTA-----GTTCAACCCACTTTATTTATTAGAG 934
3_TK81-0 GCTGATGAATCAGTCATTTAATGCTTA-----CTTCAACCCACTTTATTTACTAGAG 525
4_TA33-0 GCTGATGAATCAGTCATTTAATGCTTA-----GTTCAACCCACTTTATTTATTAGAG 1017
5_NK219-0 GCAGATGAATCAGTCATTAAGGCTTA-----GTTCAACCCACTTTAATTATTAGAG 935
* * * * *

1_NK-198 T-----ACG--T-CAGAATAATTATACTGTTGTTTGATAGAAAATGCAGTTTGATAGTAG 1077
2_NK305 T-----ACG--T-CACAATAATTATACTGTTGTTTGATAGAAAATGCAATTTGATAGTAG 986
3_TK81-0 TGCGTCACAAATTCACAATAATTATACTGTTGTTTGACAGAAAATGCAGTTTGATAGTAG 585
4_TA33-0 T-----ACG--T-CACAATAATTATACTGTTGTTTGATAGAAAATGCAATTTGATAGTAG 1069
5_NK219-0 C-----ACG--T-CAGAATAATTATACTGTTGTTTGATAGAAAATGCAGTTTGRTAGTAG 987
***** * * * * *

1_NK-198 CAATTGATTAACGTTAGTTGTATATATTTCTTTATTGATCTTGCTTACATTTGTATGACACA 1137
2_NK305 CAATTGATTAACGTTAGTTGTATATATTTCTTTATTGATCTTGCTTACATTTGTATGACACA 1046
3_TK81-0 CAATTGATTAACGTTAGTTGTATATATTTCTTTATTGATCTTGCTTACATTTGTATGACACA 645
4_TA33-0 CAATTGATTAACGTTAGTTGTATATATTTCTTTATTGATCTTGCTTACATTTGTATGACACA 1129
5_NK219-0 CAATTGATTAACGTTAGTTGTATATATTTCTTTATTGATCTTGCTTACATTTGTATGACACA 1047
* *

1_NK-198 GAAATAAAAGTACTTGCTTTCAATCATTAGACAATGACATAGCAATTGAAGAATATAGGT 1197
2_NK305 GAAATAGAAGTACTTGCTTGCAATTACCAGACAATGACATAGCAATTGAAGAATATAGGT 1106
3_TK81-0 GAAATAAAAGTACTTGCTTTCAATCATTAGACAATGACATAGCAATTGAAGAATATAGGT 705
4_TA33-0 GAAATAGAAGTACTTGCTTGCAATTACCAGACAATGACATAGCAATTGAAGAATATAGGT 1189
5_NK219-0 GAAATAAAAGTACTTGCTTTCAATCATTAGACAATGACATAGCAATTGAAGAATATAGGT 1107
* * **

1_NK-198 CAAATGTGATTTTAAAGTAATACAGGAACACCA-AGTTTTATTACATTGTGTTACTATACT 1256
2_NK305 CAAATGTGATTTTAAAGTAATA-AGGAACACTTTAGTTTTATTACATTGTGTTACTATACT 1165
3_TK81-0 CAAATGTGATTTTAAAGTAATACAGGAACACTA-AGTTTTATTACATTGTGTTACTATACT 764
4_TA33-0 CAAATGTGATTTTAAAGTAATA-AGGAACACTTTAGTTTTATTACATTGTGTTACTATACT 1248
5_NK219-0 CAAATGTGATTTTAAAGTAATACAGGAACACTA-AGTTTTATTACATTGTGTTACTATACT 1166
* * ***

1_NK-198 TTTTCTTTGGAATTTGTGTGTTTGAGGATTTTCATCGAAACATAGGCTCCTCATAGCAAC 1316
2_NK305 TTT-CTTTGGAATTTGTGTGTTTGAGGATTTTCATCGAAACATAGGCTCCTCATAGCAAC 1224
3_TK81-0 TTTTCTTTGGAATTTGTGTGTTTGAGGATTTTCATCGAAACATAGGCTCCTCATAGCAAC 824
4_TA33-0 TTTTCTTTGGAATTTGTGTGTTTGAGGATTTTCATCGAAACATAGGCTCCTCATAGCAAC 1308
5_NK219-0 TTTTCTTTGGAATTTGTGTGTTTGAGGATTTTCATCGRAACATAGGCTCCTCATAGCAAC 1226
* * *

1_NK-198 ATGTATGATTAGAGCAGCAGGATGTGCAGGGCCAATCGCAGTCGTTGGAGTCCATGCAGG 1376
2_NK305 ATGTATGATTAGAGCAGCAGGATGTGCAGGGCCAATCGCAGTCGTTGGAGTCCATGCAGG 1284
3_TK81-0 ATGTATGATTAGAGCAGCAGGATGTGCAGGGCCAATCGCAGTCGTTGGAGTCCATGCAGG 884
4_TA33-0 ATGTATGATTAGAGCAGCAGGATGTGCAGGGCCAATCGCAGTCGTTGGAGTCCATGCAGG 1368
5_NK219-0 ATGTATGATTAGAGCAGCAGGATGTGCAGGGCCAATCGCAGTCGTTGGAGTCCATGCAGG 1286
*

1_NK-198 GGCCATAGC-ACAACCATGGATTTGATACGGCGGGAATTCGTATCCTGGTGGTGGAAAGTT 1435
2_NK305 GGCCATAGCCACAACCATGGATTTGATACGGCGGGAATTCGTATCCTGGTGGTGGAAAGTT 1344
3_TK81-0 GGCCATAGCCACAACCATGGATTTGATACGGCGGGAATTCGTATCCTGGTGGTGGAAAGTT 944
4_TA33-0 GGCCATAGCCACAACCATGGATTTGATATGGCGGGAATTCGTATCCTGGTGGTGGAAAGTT 1428
5_NK219-0 GGCCATAGCCACAACCATGGATTTGATATGGCGGGAATTCGTATCCTGGTGGTGGAAAGTT 1346
* *

1_NK-198 CCGCAG-CGTTCCATATATCATAAGCAGGTTGTGGTGCAAACATATCATCATTTTCTACAT 1494
2_NK305 CCGCAGGCGTTCCATATATCATAAGCAGGTTGTGGTGCAAACATATCATCATTTTCTACAT 1404
3_TK81-0 CCGCAGGCGTTCCATATATCATAAGCAGGTTGTGGTGCAAACATATCATCATATTTCTACAT 1004
4_TA33-0 CCGCAGGCGTTCCATATATCATAAGCAGGTTGTGGTGCAAACATATCATCATTTTCTACAT 1488
5_NK219-0 CCGCAGGCGTTCCATATATCATAAGCAGGTTGTGGTGCAAACATATCATCATTTTCTACAT 1406
* * * *

1_NK-198 CCTCTGTAGATGGAGGTGGAGATGGTGGAAAGAGGTAATAATTTTCGGTGGCTCAGAGGGTG 1554
2_NK305 CCTCTGTAGATGGAGGTGGAGATGGTGGAAAGAGGTAATAATTTTCGGTGGCTCAGAGGGTG 1464
3_TK81-0 CCTCTGTAGATGGAGGTGGAGATGGTGGAAAGAGGTAATAATTTTCGGTGGCTCAGAGGGTG 1064
4_TA33-0 CCTCTGTAGATGGAGGTGGAGATGGTGGAAAGAGGTAATAATTTTCGGTGGCTCAGAGGGTG 1548
5_NK219-0 CCTCTGTAGATGGAGGTGGAGATGGTGGAAAGAGGTAATAATTTTCGGTGGCTCAGAGGGTG 1466

1_NK-198 ACATAGGTGGTAGCATGTACTCAGGATCAGCCACATCATAACAGCAGGTATAGTTTGCAC 1614
2_NK305 ACATAGGTGGTAGCATGTACTCAGGATCAGCCACATCATAACAGCAGGTATAGTTTGCAC 1524
3_TK81-0 ACATAGGTGGTAGCATGTACTCAGGATCAGCCACATCATAACAGCAGGTATAGTTTGCAC 1124
4_TA33-0 GCATAGGTGGTAGCATGTACTCAGGATCAGCCACATCATAACAGCAGGTATAGTTTGCAC 1608
5_NK219-0 ACATAGGTGGTAGCATGTACTCAGGATCAGCCACATCATAACAGCAGGTATAGTTTGCAC 1526
* *

1_NK-198 AACAGACTGTACAGGGTGCATCACAGTCATTGGAGTCGTTACATGGGTAACCACAACCAT 1674
2_NK305 AACAGACTGTACAGGGTGCATCACAGTCATTGGAGTCGTTACATGGGTAACCACAACCAT 1584
3_TK81-0 AACAGACTGTACAGGGTGCATCACAGTCATTGGAGTCGTTACATGGGTAACCACAACCAT 1184
4_TA33-0 AACAGACTGTACAGGGTGCATCACAGTCATTGGAGTCACTACATGGGTAACCACAACCAT 1668
5_NK219-0 AACAGACTGTACAGGGTGCATCACAGTCATTGGAGTCGTTACATGGGTAACCACAACCAT 1586
* *

```

1_NK-198  GAATCTGATATGGTGGTGCCGCGATGCCATTACCAGGTACCCATTGTGGAATGGGTGCTT 1734
2_NK305   GAATCTGATATGGTGGTGCCGCGATGCCATTACCAGGTACCCATTGTGGAATGGGTGCTT 1644
3_TK81-O  GAATCTGATATGGTGGTGCCGCGATGCCATTACCAGGTACCCATTGTGGAATGGGTGCTT 1244
4_TA33-O  GAATCTGATATGGTGGTGCCGCGATGCCRTTACCAGGTACCCATTGTGGAATGGGTGCTT 1728
5_NK219-O GAATCTGATATGGTGGTGCCGCGATACCATTACCAGGTACCCATTGTGGAATGGGTGCTT 1646
          *      *

1_NK-198  CTGTGAGAGGCAATGGAGCAGGCGAGGGTGATGTGAAGAAGGGCTCATCAAAAAGTACAGG 1794
2_NK305   CTGTGGGAGGTAATGGAGCAAGCGAGGGTGATGTAAAGAAGGGCTCATCAAAAAGTACAGG 1704
3_TK81-O  CTGTGGGAGGTAATGGAGCAAGCGAGGGTGATGTAAAGAAGGGCTCATCAAAAAGTACAGG 1304
4_TA33-O  CTGTGAGAGGTAATGGAGCAAGCGAGGGTGATGTAAAGAAGGGTTCATCAAAAAGTACAGG 1788
5_NK219-O CTGTGGGAGGTAATGGAGCAAGCGAGGGTGATGTAAAGAAGGGCTCATCAAAAAGTACAGG 1706
          *      *      *      *      *

1_NK-198  TTTGGTCAGCACCACAGATTG 1815
2_NK305   TTTGGTCAGCACCACAGATTG 1725
3_TK81-O  TTTGGTCAGCACCACAGATTG 1325
4_TA33-O  TTTGGTCAGCACCACAGATTG 1809
5_NK219-O TTTGGTCAGCACCACAGATTG 1727

```

Figure S2 Nucleotide-sequence alignment of PCR fragments amplified with primers T1 and T2 from plants of NK-198 (NK-198, s17 pattern is 1/1), NK-305 (NK305, 2/2), TK-81mm-O (TK81-O, 3/3), TA-33BB-O (TA33-O, 4/4), and NK-219mm-O (NK219-O, 5/5). In Moritani et al. (2013), these plants were typed as aa, bb, cc, dd, and bb, respectively, by the 17-20L marker. The numbers of nucleotide residues are shown on the right. Dashes are incorporated for maximum matching. The positions of T1 and T2 primers are indicated by blue and red fonts, respectively. Restriction sites are indicated by single- and double underlines for *HapII* and *HindIII*, respectively. Asterisks denotes positions of nucleotide divergence.