SHORT PAPER

A DNA marker for the duplicated cytosolic PGI genes in sheep's fescue (*Festuca ovina* L.)

LENA GHATNEKAR AND BENGT O. BENGTSSON*

Department of Genetics, Lund University, Sölvegatan 29, SE-223 62 Lund, Sweden

(*Received 20 January and in revised form 27 June 2000*)

Summary

Active duplicate *PgiC* genes in sheep's fescue, *Festuca ovina*, are associated with a PCR marker of specific length (about 370 bp, of which 231 are in an intron). Using this marker, the frequency of plants with duplicate genes is estimated to be about 10 % in a population from southern Sweden. The close molecular similarity between the electrophoretically different duplicated genes is in accordance with the conclusion reached earlier that they are indeed alleles at the same locus.

1. Introduction

The enzyme phosphoglucose isomerase (PGI; EC 5.3.1.9) regulates one of the key processes in glycolysis. Plants have at least two sets of PGI enzyme, one in the cytosol and one in the chloroplasts (Gottlieb, 1982). The cytosolic form is often electrophoretically highly polymorphic, and the processes maintaining this variation have been much discussed (see e.g. the recent article by Filatov & Charlesworth, 1999, and references therein).

Cytosolic PGI is normally coded for by a single gene in diploid plants, but in both *Clarkia* and *Persea* some species carry a second, duplicated locus (Gottlieb, 1977; Goldring *et al.*, 1985). In sheep's fescue (*Festuca ovina* L.) the situation is even more complex. Here plants with multibanded PGI zymograms, i.e. with more allozymes than can be attributed to the standard locus, have been found in several populations in southern Sweden (Prentice *et al.*, 1995; Bengtsson *et al.*, 1995). Based on an extensive set of crosses with such multibanded plants, Ghatnekar (1999) reached the conclusion that in *F. ovina* a second locus for cytosolic PGI exists, *PgiC-2*, unlinked to the standard locus, *PgiC-1*. To the second locus map two alleles (b_2 and c_1) with electrophoretic properties identical to common alleles at the standard locus. A complex variant coding for both electrophoretic types

(b_2c_1) was also found to map to *PgiC-2*. Most chromosomes, however, have no active *PgiC* gene at this second locus. The frequency of active genes at the duplicate locus is difficult to determine from electrophoretic information alone, since the duplicate alleles produce isozyme bands indistinguishable from bands produced by common variants at the standard locus. Since no other fescue species investigated show any sign of a *PgiC* duplication (Ghatnekar, 1999), the situation in *F. ovina* probably represents the spread of an evolutionarily recent gene duplication that has rapidly accumulated allelic variation.

Here we describe a PCR-based DNA marker associated with the active alleles at the duplicate *PgiC* gene of *F. ovina* in southern Sweden. The marker identifies a length variant for one of the introns in the gene, and thus also gives evidence about the nature and origin of the duplication.

2. Materials and methods

(i) *Plant material and its electrophoretic characterization*

Festuca ovina is a self-incompatible grass, very common in northern Europe. All plants reported on here belonged to diploid *Festuca ovina* ssp. *vulgaris* (Koch) Sch. & Kell (2n = 14). In an earlier population survey of southern Sweden the plants designated *K*, *L*, *M*, *N* and *P* were found to have multibanded zymograms for cytosolic PGI (Bengtsson *et al.*, 1995;

* Corresponding author. Tel: +46 46 222 98 90. Fax: +46 46 14 78 74. e-mail: bengt_olle.bengtsson@gen.lu.se

Ghatnekar, 1999). Plant **D**, found in the same survey, showed no active gene duplication but was homozygous for a rare, characteristic allele (*d*) at the standard locus that made it useful as a test plant. Electrophoretic analyses of PGI isozymes were performed using standard procedures and recipes (Tanksley & Orton, 1983). The genotypes of the experimental plants with respect to the two *PgiC* loci were deduced by Ghatnekar (1999) to be as follows:

K: $a_2/a_2\ b_2c_1/-$
L: $a_2/a_2\ b_2c_1/-$
M: $a_1/b_1\ c_1/-$
N: $a_2/b_1\ c_1/c_1$
P: $a_2/c_2\ b_2/-$
D: $d/d\ -/-$

A new sample was collected from Kabusa, the locality on the south coast of Sweden from which the multibanded plants **K** and **L** came. Panicles with mature seed were collected in 1995 and single-offspring plants were derived from each such seed family. At Kabusa, *F. ovina* grows in very large numbers in an uninterrupted area of many square kilometres; the sampled plants grew well separated from each other within a radius of 200 m.

(ii) *DNA methods*

Leaf material from individual plants was ground under liquid nitrogen to a fine powder and DNA was isolated using QIAGEN DNeasy Plant Mini Kit. PCR amplifications were performed according to standard techniques with 1 min at 95 °C, followed by 30 cycles of 30 s at 95 °C, 1 min at 55 °C, 45 s at 72 °C, ending with 7 min at 72 °C. All PCR runs reported on here were performed with forward primer LG9 (5′-CTGCGAAGCTCAAGGAGAAG-3′) and reverse primer LG10 (5′-GACCACATTCACCCCATCAC-3′). The PCR products obtained were visualized by agarose gel electrophoresis and stained with ethidium bromide. Gels were analysed with Gel-Pro analyzer (Media Cybernetics). In our runs, the apparatus assigns the size of DNA fragments with an accuracy of about ± 10 bp.

For sequencing, PCR products were purified (QIAquick PCR Purification Kit) and used as templates for direct sequencing by an autosequencer (ABI 310). Reactions were performed using 30–60 ng of template DNA, 2·0 pmol of primer and 4 μl of fluorescent dideoxy terminator mix (BigDye, PE Applied Biosystems) in a total of 10 μl reaction volume. The cycle sequencing reactions were performed with the same primers as the PCR reaction, with 35 cycles of 20 s at 96 °C, 1 °C/s to 52 °C, 15 s at 52 °C, 1 °C/s to 60 °C and 4 min at 60 °C. The sequencing data were analysed with Sequencher 3.0 (Genes Codes Corporation). Detailed information on

the sequences of different *PgiC* genes will be presented in a later publication.

3. Results

(i) *An informative primer combination*

The published cDNA sequence for the cytosolic PGI gene in *Zea mays* (Lal & Sachs, 1995) was used to develop primer combinations that were tested on *Festuca ovina* material. Many, but not all, primer combinations led to the production of distinct DNA fragments, indicating a close homology between the *PgiC* genes in *Zea* and *Festuca*.

Particularly interesting results were obtained with the pair of primers denoted LG9 and LG10 (see above). For the homozygous test plant **D**, with no active duplicate *PgiC* gene, the combination yielded a single DNA fragment of about 440 bp length (Fig. 1*b*). A comparison of its sequence with the published information from other species, in particular *Zea*, suggested that the fragment came from an active *PgiC* gene. It covered, as planned, 137 bp from exons 4 and 5, plus an intervening intron of 303 bp (data not shown). (The exact number of exons in the *PgiC-1* gene in *Festuca* is not yet known but is here assumed to be the same as in *Clarkia lewisii*; see Thomas *et al.* (1992).) When the primers were used in PCR with other *F. ovina* plants lacking duplicate *PgiC* genes, similar PCR products were obtained, though some variation in fragment length was seen. Sequence analysis showed that this variation was due to differences in the length of the intron, which in some plants was reduced to only 179 bp. All the sequenced introns had characteristic GT–AG boundaries and showed sequence variation typical of non-coding regions (data not shown).

When tested on experimental plants known to carry active duplicate *PgiC* genes (Fig. 1*a*), a characteristic
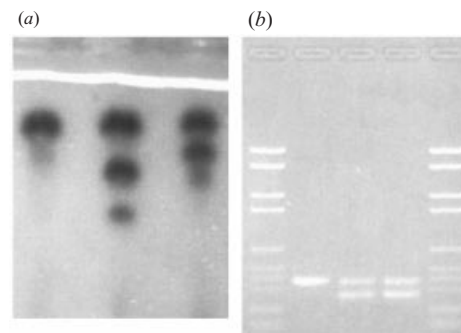

Fig. 1. Isozyme pattern (*a*) and PCR products (*b*) of *F. ovina* plants with the following genotypes at the *PgiC* loci: d/d, $-/-$ (lane 1); d/d, $b_2/-$ (lane 2); d/d, $c_1/-$ (lane 3). The lengths of the PCR produced bands in (*b*) are 440 bp (all three lanes) and 370 bp (the two lanes with active genes at *PgiC-2*). Flanking lanes with Boehringer molecular weight marker VI are included in (*b*).

pattern of fragment lengths was obtained. In addition to fragments that could be identified as coming from alleles at the standard *PgiC-1* locus, a 370 bp fragment was always seen (Fig. 1*b*), with intron length 231 bp when sequenced. To study whether this particular PCR fragment could be used as a marker for the duplicate *PgiC* genes, the following investigations were performed.

(ii) *Do all active duplicate genes produce this marker fragment?*

All multibanded plants investigated by Ghatnekar (1999) (**K**, **L**, **M**, **N** and **P** as well as their offspring with duplicate *PgiC* genes) were tested and all were found to produce the 370 bp marker fragment, in addition to one or two other PCR fragments obviously coming from the standard locus. The marker fragment appeared always to have the same length, independent of which duplicate gene was tested (b_2, c_1 or b_2c_1).

To extend the analysis, 584 independently derived plants from Kabusa were scored for their *PgiC* isozymes, and 16 of them were found to be clearly multibanded. When tested with PCR amplification, all these 16 plants produced the 370 bp DNA marker band. Thus, all plants known to have a duplicate *PgiC* gene were found to produce the 370 bp marker band.

(iii) *Do plants without the duplication produce fragments of similar length?*

In our original experimental material, no plant without a duplicate gene produced the 370 bp fragment. To extend this observation, the material from Kabusa was re-examined. This time, 24 plants were randomly chosen from the 568 ($= 584 - 16$) plants earlier judged as not having zymograms with a clear multibanded pattern. Due to the electrophoretic similarity between the duplicate genes and the alleles at the standard locus, the genotypic status of these plants was unknown. When tested for their PCR fragments, two plants produced bands of approximate length 370 bp while one plant produced a band slightly but distinctly longer than this; the rest of the plants produced no band of this approximate size. When further tested by sequencing, it was found that the marker band in the first pair of plants had an intron of exactly 231 bp, the same as other plants with duplications. In addition, the zymograms of these plants were compatible with their carrying one of the known duplicate genes. The third plant, however, was found to produce a fragment with an intron of 246 bp; in addition it had a zymogram that could not be accounted for by any of the known duplicate genes. Thus, the 24 plants investigated were judged to include 22 plants without and 2 with a *PgiC* duplication, the last type of plants being recognized by their PCR marker band alone.

4. Discussion and conclusions

The primer combination described here produces a PCR marker characteristic for active duplicate *PgiC* genes in *Festuca ovina*. All plants known to have a duplicate *PgiC* gene produced this 370 bp marker with an intron of 231 bp, while (as far as we can tell) all plants without the duplication lacked this marker. There is thus a strong population association between the 370 bp marker and the presence of active duplicate *PgiC* genes. Due to the extensive variation in intron size at the standard locus (Ghatnekar & Bengtsson, in preparation), it is likely that *somewhere* in the range of *F. ovina* one can find plants whose intron at the *PgiC-1* locus is exactly 231 bp. This would then weaken the correlation between the 370 bp PCR marker band and the duplicate *PgiC* genes. However, no such plants have yet been found, and the association between the marker and the duplicate genes must be considered strong in *F. ovina* populations from southern Sweden.

The 370 bp PCR marker can thus be used to estimate the frequency of plants with active duplicate *PgiC* genes. This frequency could, in principle, be calculated from the distribution of zymograms alone, using likelihood methods, but given the overlapping electrophoretic mobilities at the two loci such estimates would be highly uncertain (Bengtsson & Ghatnekar, unpublished results). Hence, a marker for the duplication is most welcome. Already from the data reported above, a rough calculation of the frequency of duplication-carrying plants in Kabusa can be made. Among the 584 electrophoretically investigated plants, 16 had multibanded zymograms, and among the 24 plants investigated by PCR from the rest of the sample (568), two were found to have duplicate genes. This gives an estimate of about 10 % for the frequency of duplication-carrying plants ($[16 + (2/24)\,568]/584 = 0\cdot108$). It is interesting to note that most duplicate *PgiC-2* alleles occur in combination with alleles at the *PgiC-1* locus that makes their zymogram look normal and not multibanded.

Duplicate *PgiC* genes have been found over a large geographic area (Prentice *et al.*, 1995; Bengtsson *et al.*, 1995) and their frequency may at least locally become substantial, as the data obtained from Kabusa indicate. Among the duplicate genes, electrophoretic and structural variability occurs (Ghatnekar, 1999). All this indicates that the *PgiC* duplication in *F. ovina* is not an ephemeral, altogether deleterious phenomenon, but that it must be associated with a non-negligible evolutionary history.

Besides providing us with a useful tool for further studies on the population frequency of the duplication, the results reported here shed light on some of the molecular properties of the duplicate genes. The close molecular similarity in intron length between the electrophoretically different duplicate genes b_2, c_1 and

b_2c_1 gives, first of all, strong support to our earlier conclusion (Ghatnekar, 1999) that these genes are alleles at the very same locus. Secondly, the fact that all active alleles at *PgiC-2* carry an intron shows that the duplication cannot have originated via retro-transposition. The duplication in *Festuca* is thus similar to the *PgiC* duplication in *Clarkia* (Gottlieb & Ford, 1996) but different from, for example, the *jingwei* duplication of *Adh* in *Drosphila* (Long & Langley, 1993).

References

Bengtsson, B. O., Weibull, P. & Ghatnekar, L. (1995). The loss of alleles by sampling: a study of the common outbreeding grass *Festuca ovina* over three geographic scales. *Hereditas* **122**, 221–238.

Filatov, D. A. & Charlesworth, D. (1999). DNA polymorphism, haplotype structure and balancing selection in the *Leawenworthia PgiC* locus. *Genetics* **153**, 1423–1434.

Ghatnekar, L. (1999). A polymorphic duplicated locus for cytosolic PGI segregating in sheep's fescue (*Festuca ovina* L.). *Heredity* **83**, 451–460.

Goldring, A., Zamir, D. & Degani, C. (1985). Duplicated phosphoglucose isomerase genes in avocado. *Theoretical and Applied Genetics* **71**, 491–494.

Gottlieb, L. D. (1977). Evidence for duplication and divergence of the structural gene for phosphoglucose isomerase in diploid species of *Clarkia*. *Genetics* **86**, 289–308.

Gottlieb, L. D. (1982). Conservation and duplication of isozymes in plants. *Science* **216**, 373–380.

Gottlieb, L. D. & Ford, V. S. (1996). Phylogenetic relationships among the sections of *Clarkia* (Onagraceae) inferred from the nucleotide sequences of *PgiC*. *Systematic Botany* **21**, 45–62.

Lal, S. K. & Sachs, M. M. (1995). Cloning and characterization of an anaerobically induced cDNA encoding glucose-6-phosphate isomerase from maize. *Plant Physiology* **108**, 1295–1296.

Long, M. & Langley, C. H. (1993). Natural selection and the origin of *jingwei*, a chimeric processed functional gene in *Drosophila*. *Science* **260**, 91–95.

Prentice, H. C., Lönn, M., Lefkovitch, L. P. & Runyeon, H. (1995). Associations between allele frequencies in *Festuca ovina* and habitat variation in the alvar grasslands on the Baltic island of Öland. *Journal of Ecology* **83**, 391–402.

Tanksley, S. D. & Orton, T. J. (eds.) (1983). *Isozymes in Plant Genetics and Breeding*. Amsterdam: Elsevier.

Thomas, B. R., Laudencia-Chingcuanco, D. & Gottlieb, L. D. (1992). Molecular analysis of the plant gene encoding cytosolic phosphoglucose isomerase. *Plant Molecular Biology* **19**, 745–757.