Heredity **83** (1999) 451–459 *Received 24 November 1998, accepted 12 May 1999*

A polymorphic duplicated locus for cytosolic PGI segregating in sheep's fescue (*Festuca ovina* L.)

LENA GHATNEKAR*

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

Normally sheep's fescue (*Festuca ovina*) carries a single locus encoding the cytosolic enzyme phosphoglucoisomerase (PGI). However, at several localities in southern Sweden plants with more than two different alleles have been found by isozyme studies. Crossing experiments and subsequent progeny analyses revealed two unlinked loci. The additional locus was polymorphic for two alleles with different electrophoretic mobilities. Two of the numerous alleles at the primary locus had the same mobility. A complex gene coding for both electrophoretic types was also found to map to the additional locus, but no such complex gene has been found at the primary locus. The genes at the additional locus can be lacking, or present in hemizygous or homozygous form. Allozymes encoded by the additional locus readily form heterodimers with those encoded by the primary locus, thus giving rise to multibanded isozyme patterns. This segregating, polymorphic gene duplication may well represent a rare case where the initial stage of the evolution of a new gene can be directly studied.

Keywords: duplication, *Festuca ovina*, gene evolution, genetic polymorphism, isozymes, PGI.

Introduction

Gene duplications play an important role in the evolution of eukaryote genomes. However, very little is known about the two key processes that initiate the evolution of a new gene. First, one may ask about the molecular origin of the duplication. Whereas tandem gene duplications may be explained by errors in the replication or recombination processes, gene duplications to new sites in the genome are more difficult to understand, unless they result from a retrovirus-like insertion of gene transcripts. Other molecular issues concern how errors caused by mispairing are avoided during meiosis, and about the degree of 'gene-talk' that occurs between the duplicated sites as a result of gene conversions. The second question raised by the evolution of a gene duplication is the nature of the selective processes which help the duplicated gene to become established in the population and prevent it from being lost through chance effects. Again, a tandemly duplicated gene is easier to analyse — at least theoretically — because it behaves as a new allele at an already existing locus, whereas a duplicated gene at an unlinked locus gives rise to many more interactions between the different genes. In very few natural cases has it been possible to study these questions during the initial evolutionary stages of new duplicated genes.

In higher plants the number of genes present for any particular enzyme is normally highly conserved (in diploid species) and it often reflects the number of different cell compartments where the same catalytic reaction takes place (Gottlieb, 1982). Nevertheless, several cases of fixed duplications of genes encoding enzymes that are active in the same cell compartment have been observed (Gottlieb, 1982). Recent duplications which have not yet become fixed but show intraspecific variation have been reported from *Drosophila* (Lange *et al*., 1990; Lootens *et al*., 1993) but no well documented case is known from any higher plant.

Phosphoglucoisomerase (PGI; EC 5.3.1.9) is an indispensable enzyme to glycolysis where it catalyses the reversible isomerization of glucose 6-phosphate and fructose 6-phosphate. In plants the gene is often highly polymorphic, and characterization of its variation pattern at the DNA sequence level has begun (Terauchi *et al*., 1997; Liu *et al*., 1999). A duplication of the locus encoding cytosolic PGI has been well studied in *Clarkia*, where different diploid species have either one or two loci (Gottlieb, 1977). A gene duplication which occurred early in the evolution of the genus *Clarkia*, resulting in two unlinked loci which has been retained in all subsequently evolved species, is the most likely cause

*Correspondence. E-mail: lena.ghatnekar@gen.lu.se

of this variation in gene number. A duplication of the cytosolic *Pgi* gene resulting in two unlinked copies has also been reported from avocado (Goldring *et al*., 1985).

In a previous study of *Festuca ovina* (Bengtsson *et al*., 1995) an electrophoretic pattern, similar to the patterns found in duplicated *Clarkia* species and in polyploids, was seen for cytosolic PGI in about 2% of the investigated plants. These plants grew at a number of sites in Scania, the southern province of Sweden. From the large island Öland in the Baltic Sea, a similar multibanded isozyme pattern for PGI has also been reported (Prentice *et al*., 1995). The deviant plants had, for all other investigated enzyme systems, the isozyme patterns expected from a diploid outbreeder.

The present investigation is an analysis of the genetic background to the complex electrophoretic phenotypes of cytosolic PGI in *Festuca ovina*. Because of the complexity of the system, an intricate series of crosses had to be analysed before the underlying genetic situation could be understood. From the results it appears as if this system may well represent a rare instance where the initial stages of the evolution of a new gene can be directly studied.

Materials and methods

Plant material

Twelve *F. ovina* plants with a multibanded PGI phenotype were found in a study of genetic differentiation in southern Sweden (Bengtsson *et al*., 1995). Of these, five plants — designated **K**, **L**, **M**, **N**, **P** — were chosen for further analysis to determine the genetic determinants of the multibanded trait. Two of the plants (**K**, **L**) came from a population at Kabusa on the south coast of Scania, one plant (**P**) from an inland population at Skogshejdan 20 km north of Kabusa, and two plants (**M**, **N**) from Högsma, an inland population in the northern part of the province (about 100 km from Kabusa and Skogshejdan). The geographical positions of the collection sites are given in Bengtsson *et al*. (1995, fig. 1).

Morphologically, all multibanded plants belonged to *Festuca ovina* ssp. *vulgaris* (Koch) Sch. & Kell., and their isozyme patterns were normal for other enzyme systems investigated (DIA-1, GOT-2, PGD-1, PGM, TPI-1, TPI-2). To ensure that the plants represented single genotypes, they were twice vegetatively cloned into very small tufts. In no instance did this change their multibanded PGI phenotypes. Chromosome counts were made from root tip preparations.

Two test plants from our collection of *F. ovina* genotypes were used in the crossing experiments. Test plant **C** was originally isolated from the commercial variety Barok of Dutch origin (bred by Barenbrug Holland B.V., Arnhem, The Netherlands), and test plant **D** was derived from a seed collected at Hagestad on the south coast of Scania. Both plants were known to be homozygous for specific and characteristic PGI alleles, and had earlier been clonally multiplied.

Other Festuca *species*

To compare *F. ovina* with related species, plants of *F. altissima*, *F. polesica*, *F. pratensis* and *Lolium perenne* were also investigated. Forty plants of *F. altissima* were collected from Rövarkulan and 20 plants from Lemmeströ, two localities in Scania about 20 km apart. From Kivik, situated on the east coast of Scania, 16 plants of *F. polesica* were collected. About one hundred plants of diploid *F. pratensis* were screened, together with some plants of *L. perenne*. These plants were derived from commercial varieties and were obtained from the Svalöf Weibull plant breeding company at Svalöv, Sweden.

Cultivation

Seeds were sown and grown on moist filter paper in Petri dishes until the seedlings were well developed. At that stage they were either directly analysed by electrophoresis or replanted into clay pots in the greenhouse for later analysis. As *F. ovina* needs a period of vernalization to flower, plants intended for crossing experiments were kept in an outdoor bench yard during the winter.

Crosses

All crosses were performed under controlled conditions but without emasculation of the female parent. In April the plants intended for crosses were transferred to a common greenhouse. Parental plants were placed next to each other in a plastic isolation cage. *Festuca ovina* is normally self-incompatible, but some seeds were later found to be derived from self-fertilization (see tables below). The seed-set in the crosses was, in general, normal.

Electrophoresis

All electrophoretic analyses were performed by running 12% starch gels (Connaught) at 75 mA for 3–4 h. A Tris/citric acid buffer, pH 7.0, was used and the staining procedures followed standard protocols (Weibull *et al*., 1988). Young leaves or seedlings of 1–2 cm length were homogenized in only small amounts of extraction buffer (Weibull *et al*., 1988) and the electrophoretic analysis was performed immediately after sample preparation.

For both kinds of tissue, this procedure yielded high enzyme activities and led to clear and consistent bands.

PGI

Most higher plants have one PGI isozyme in plastids and a different one in the cytoplasm. The two sets of isozymes have different migration rates, are clearly separable on electrophoretic gels, and never form interlocus hybrid bands.

The cytosolic enzyme is, in virtually all diploid plants, encoded by a gene with a single copy per genome (*PgiC*; later in this paper to be denoted *PgiC-1*). The enzyme is a dimer, so heterozygous plants produce three electrophoretic bands, whereas homozygous plants produce one band. *Festuca ovina* is known to be highly variable for cytosolic PGI, and at least 10 different electrophoretic alleles can be recognized in the species (Weibull *et al*., 1991).

The PGI zymograms of the multibanded plants chosen for analysis were very similar but not identical. Because five or six bands of cytosolic PGI were distinguishable for each of them, all plants must have carried at least three different genes for the enzyme.

Nomenclature

The electrophoretic bands of cytosolic PGI were divided into four areas based on their migration rates, and the areas were designated A, B, C and D. Fast dimers migrated to the A-region, slow ones to the D-region, and intermediary ones to the B- and C-regions. With this nomenclature, all five plants chosen for further analysis had phenotypes ABC.

Gene variants that code for homodimers in the A-region will be called *a*, and the *b* and *c* and *d* genes are defined analogously. More than one homodimeric variant existed in all the different areas. When in some later crosses it became important to distinguish between different genes coding for allozymes in the same region, a numeral was added to the letter describing the exact genetic type. The numeral was so chosen that gene a_1 denoted a homodimer migrating faster than a_2, etc.

Test plant genotypes

From earlier crosses it was known that test plant **C** was homozygous for a gene producing a band in the C region. Its genotype was thus designated *c*/*c*, and this *c* allele was in the analyses of the crosses used to define the standard cytosolic locus (*PgiC*).

Similarly, test plant **D** was homozygous *d*/*d* for an allele at the same locus with an unusually slow migration rate.

Results

To analyse the genetic constitution of the multibanded plants, a series of crosses was carried out. All plants were crossed to the test plants, but crosses between lineages were also performed to ascertain the allelic relationships between the additional genes detected.

The zymograms were normally of an excellent quality and could, in certain situations, also be used to determine the number of genes encoding a certain band. However, the analyses of the crosses were occasionally disturbed by two kinds of problems. The first was an irregular frequency of offspring derived from self-fertilization. The second was occasional deviations from the expected Mendelian ratios. Both types of problems could be expected to occur in analyses of the present kind. The second problem is specifically considered at the end of the Results section, whereas the first is reported on only in the tables. Taking into account the possibility of these effects, the structure of the results was always clear and could be used to determine the genetic constitution of the multibanded plants.

Crosses with plant **N**

When crossed to test plant **C**, plant **N** produced offspring of two phenotypic kinds, AC and BC, in a 1:1 ratio (Table 1). All offspring plants had phenotype C, but as C was contributed by both parents, this cross was uninformative with respect to the genetic determinant of phenotype C in plant **N**.

Two classes of offspring were obtained also in the cross to plant **D** (Table 1). Half of the progeny had phenotype ACD and the other half BCD, again in a 1:1 ratio. Phenotype C was inherited by all offspring also in this cross, although now it was contributed by only one of the parents, plant **N**. Thus, plant **N** behaved as if it were heterozygous for *a* and *b* at one locus and homozygous for *c* at another locus.

To investigate this suggestion, F_1 plants from the cross **D** × **N** were test-crossed further. Offspring with

Table 1 Segregation of PGI phenotypes in *Festuca ovina*. Plant **N** (phenotype ABC) crossed to plants **C** and **D** (phenotypes C and D, respectively). Segregation of offspring phenotypes

	Offspring phenotypes				
Cross	AC	BC	ACD	BCD	Total
C × ABC	14	10			24
ABC × D			3	3	6
D × ABC			10	7	17

Table 2 F_1 *Festuca ovina* with phosphoglucoisomerase (PGI) phenotypes ACD and BCD, derived from **D** × **N** (Table 1), crossed to plants **C** and **D**. Segregation of phenotypes, where duplicated letters denote bands of strong intensity. In parentheses is the genetic interpretation of the results, based on plant **N** being homozygous for allele c' at a separate, additional locus (such additional genes are designated by a prime). In the statistical tests it is assumed that the loci are unlinked and that all segregation is Mendelian

	Offspring phenotypes (genotypes)								
Cross	ACC ($a/c\ c'/-$)	BCC ($b/c\ c'/-$)	CCD ($c/d\ c'/-$)	AC (a/c)	BC (b/c)	CD (c/d)	Total	χ^2_3	P
C × ACD ($c/c \times a/d\ c'/-$)	8		11	13		7	39	2.333	NS
C × BCD ($c/c \times b/d\ c'/-$)		10	12		13	6	41	2.805	NS
	ACD ($a/d\ c'/-$)	DDC ($d/d\ c'/-$)	AD (a/d)	D (d/d)					
ACD × D ($a/d\ c'/- \times d/d$)	13	15	11	3			42	7.905*	< 5%
D × ACD ($d/d \times a/d\ c'/-$)	11	16	14	12			53	1.113	NS

phenotypes ACD and BCD were crossed to plant **C** (Table 2). In the scoring of the progeny phenotypes it was possible to distinguish between plants carrying one copy or two copies of the *c* gene, recognized by a difference in staining intensity of the dimeric C band relative to the other bands. As a precaution, only seeds matured on test plant **C** were analysed. All such seedlings carried, by necessity, a single copy of the known *c* gene from their mother, giving robustness to the interpretation of staining intensity differences. In the offspring of both types of paternal plants (ACD and BCD) the strong-intensity C phenotype segregated to half the progeny. Phenotypes A and D, in the first cross, and B and D in the second cross segregated as Mendelian alternatives independently with respect to the presence/absence of the strong-intensity C phenotype.

In a back-cross to test plant **D** only one phenotype, ACD, was investigated, but offspring from both parental plants were analysed (Table 2). The ACD-parental plant produced 28 offspring with phenotype C and 14 without it, whereas the D-parental plant yielded 27 plants with phenotype C and 26 without it. A and D segregated as alternatives independently with respect to the presence/absence of phenotype C.

Preliminary genetic interpretation

These results support the suggestion that plant **N** was heterozygous for *a* and *b* at one locus, and homozygous for *c* at a second locus. From the second set of crosses it can be deduced that this second locus is distinct from *PgiC* in **C** and **D**.

Alleles not situated at the standard *PgiC* locus will be designated by a prime sign ($'$), and the deduced genotype of plant **N** can thus be written $a/b\ c'/c'$. According to the proposed genetic model, the offspring of **N** × **D** with phenotypes ACD and BCD had genotypes $a/d\ c'/-$ and $b/d\ c'/-$, respectively. That the two PGI loci are unlinked follows from the independent segregation of the different genotypes in the latter crosses.

The only result which did not completely support the deduced interpretation was the significantly skewed segregation ratio of presence/absence of c' in the offspring of the cross between the plant with phenotype ACD and test plant **D** (Table 2). We do not take this to imply a mistake in our interpretation, as such skewness may be caused by many factors unrelated to the gene duplication of primary interest here. Instead, what is of importance is the close agreement between the kinds of phenotypes expected and the kinds of phenotypes observed in the different test cross progenies.

In Table 2 the genotypic interpretation of the different phenotypes is given in parentheses, as well as the χ^2 values which test the interpretation. For simplicity, later tables use only this genotypic notation.

*Crosses with plants **M** and **P***

When crossed to test plant **D** (Table 3), plant **M** gave rise to four phenotypic classes of offspring: AD, BD, ACD and BCD. The phenotypes segregated as 1:1:1:1. Using the same kind of reasoning as above, the results indicate that the genotype of plant **M** was $a/b\ c'/-$, i.e. that plant **M** was hemizygous for allele c' at an additional *Pgi* locus.

Offspring of **P** × **D** also fell into four classes (Table 3). The analysis and the interpretation were very similar to the preceding case, but here the combination of alleles was different. Plant **P** was heterozygous for

Table 3 Segregation of phosphoglucoisomerase (PGI) phenotypes in *Festuca ovina*. Plants **M** and **P** crossed to plant **D** (*d*/*d*). Hypothesis: the genotype of **M** is *a*/*b c*′/–, of **P** is *a*/*c b*′/–; hemizygous alleles segregate to half the progeny

	Offspring genotypes							
Cross	*a*/*d c*′/–	*b*/*d c*′/–	*a*/*d*	*b*/*d*	Others	Total	χ^2_3	*P*
a/*b c*′/– × *d*/*d*	3	3	2	0		8		
d/*d* × *a*/*b c*′/–	15	10	10	13	5†	48 (53)	1.500	NS
	a/*d b*′/–	*c*/*d b*′/–	*a*/*d*	*c*/*d*				
a/*c b*′/– × *d*/*d*	10	10	7	10		37	0.730	NS
d/*d* × *a*/*c b*′/–	8	5	12	15		40	5.800	NS

† Of which four were *d*/*d* (probably derived by selfing) and one was *c*/*d* (discussed further in the text).

alleles *a* and *c* at the standard locus, and hemizygous for a *b* gene (thus designated *b*′) at a second, unlinked locus.

Homology relationship between the additional genes

The three plants, **P**, **M** and **N**, had a similar genetic organization in that they carried standard alleles at the locus normally encoding cytosolic PGI at the same time as they carried additional unlinked genes.

So far, the analysis has not addressed the question as to whether these additional genes were located at a homologous locus, or whether they resided at different places in the genome. To this end, a second generation cross was performed which combined gene *b*′ from plant **P** and gene *c*′ from plant **N** (the genotype of the resulting plant was *a*/*d b*′/*c*′ or *a*/*d b*′/– *c*′/–, depending on the correct homology relationship between the additional genes). This plant was then crossed to test plant **D**.

If *b*′ and *c*′ were alternative alleles at a single locus, all offspring plants would carry one of these alleles, in addition to the alleles at the main locus. If, on the other hand, the additional genes resided at two different loci, then two additional classes of offspring should be found: those having both extra genes and those without either of them. The data from the cross are given in Table 4. All analysed offspring, a total of 47 plants, were of the four types (*a*/*d b*′/–, *a*/*d c*′/–, *d*/*d b*′/–, *d*/*d c*′/–) expected for the segregation at two, and not three, unlinked loci. The χ^2 value for the goodness-of-fit between the data and the two-locus model was 2.106 (3 d.f.; $P > 0.05$), whereas the corresponding value for the three-locus model was 51.213 (7 d.f.; $P < 0.001$).

An analogous second generation cross was performed to combine gene *c*′ from plant **M** and *b*′ from plant **P** in order to clarify their allelic relationship. In this cross a small mobility difference between the alleles at the standard locus, classed as *a*, was utilized. Thus, the plant that was test-crossed to **D** had either genotype a_1/a_2 *b*′/*c*′ or a_1/a_2 *b*′/– *c*′/–, where a_1 gave a faster migrating homomeric band than a_2. As can be seen from Table 4, the additional alleles *b*′ and *c*′ segregated as alleles at a single locus. A goodness-of-fit test to the two-locus model, in which is included the independent segregation of the *a* alleles at the standard locus, gave a χ^2 value of 0.634 (3 d.f.; $P > 0.05$), whereas the value for the corresponding three-locus model was 83.269 (7 d.f.; $P < 0.001$).

Thus, the additional alleles in plants **N**, **M** and **P** behaved as if they belonged to the same locus, unlinked to the standard locus.

*Crosses with plants **L** and **K***

When plant **L** was crossed to test plant **C**, half of the 31 investigated offspring plants were multibanded ABC-plants and the other half triple-banded AC-plants

Table 4 Segregation of phosphoglucoisomerase (PGI) phenotypes in *Festuca ovina*. Test of the hypothesis that the additional genes *b*′ and *c*′ are allelic. Segregation of offspring from crosses between **D** and (i) *a*/*d b*′/*c*′, and (ii) a_1/a_2 *b*′/*c*′. Plant *a*/*d b*′/*c*′ was derived from the cross [(**P** × **D**) × (**N** × **D**)] and a_1/a_2 *b*′/*c*′ from the cross [(**P** × **D**) × (**M** × **D**)]

	Offspring genotypes						
Cross	*a*/*d b*′/–	*a*/*d c*′/–	*d*/*d b*′/–	*d*/*d c*′/–	Total	χ^2_3	*P*
(i) *a*/*d b*′/*c*′ × *d*/*d*	10	11	10	16	47	2.106	NS
	a_1/*d b*′/–	a_2/*d b*′/–	a_1/*d c*′/–	a_2/*d c*′/–			
(ii) a_1/a_2 *b*′/*c*′ × *d*/*d*	21	20	23	18	82	0.634	NS

Table 5 Segregation of phosphoglucoisomerase (PGI) phenotypes in *Festuca ovina*. Plants **L** (i) and **K** (ii) crossed to plants **C** and **D**. Hypothesis: both plants **L** and **K** have genotype *a*/*a b′c′*/–; the hemizygous complex additional gene segregates to half the progeny

	Offspring genotypes							
Cross	*a*/*c b′c′*/–	*a*/*c*	*a*/*d b′c′*/–	*a*/*d*	Others	Total	χ_1^2	*P*
(i)								
a/*a b′c′*/– × *c*/*c*	2	6				8		
c/*c* × *a*/*a b′c′*/–	10	7			6†	17 (23)	0.529	NS
a/*a b′c′*/– × *d*/*d*			18	13		31	0.807	NS
d/*d* × *a*/*a b′c′*/–			38	42		80	0.200	NS
	a/*c b′c′*/–	*a*/*c*	*a*/*d b′c′*/–	*a*/*d*				
(ii)								
a/*a b′c′*/– × *c*/*c*	54	30				84	6.857**	<1%
a/*a b′c′*/– × *d*/*d*			336	139		475	81.703***	<0.05%
d/*d* × *a*/*a b′c′*/–			57	36	74‡	93 (167)	4.742*	<5%

† Five *c*/*c* (probably derived by selfing) and one *c*/*d* (further discussed in the text).
‡ 74 *d*/*d* (probably derived by selfing).

(Table 5). When the same plant was crossed to test plant **D**, equal proportions of ABCD and AD plants were found among the 111 investigated offspring (Table 5). Thus, plant **L** seemed to be homozygous for allele *a* while at the same time being hemizygous for a complex *b′c′* gene at an additional locus.

The same offspring classes were found when plant **K** was crossed to the two test plants (Table 5), but with considerable deviations from the expected segregation ratios. The most likely explanation was, however, that plant **K** had a genotype similar to plant **L** and that the segregation distortion was caused by some other effect.

To test the degree of linkage between the additional B and C determinants, a larger number of offspring than in the other progenies was investigated. However, despite a total of 678 offspring being scored from the cross **K** × **D**, not once was a plant observed that had only inherited *b′* and not *c′*, or the reverse, from its multibanded parent.

*Two unlinked loci in plants **K** and **L***

From the analysis made so far, the genotype of plants **L** and **K** could be either *a*/*a b′c′*/– or *a*/*ab′c′*. The first alternative would mean that the additional genes were clearly separated from the standard locus, whereas the second alternative would imply a local triplication of the standard locus. For further analysis an ABCD plant, obtained in the F_1 of **L** × **D**, was crossed to plant **C**. Four classes of offspring, ABC, BCD, AC and CD, were found among the 61 plants screened (Table 6). The fact that four classes with a good fit to a 1:1:1:1 segregation were obtained, ruled out the second alternative which would have produced only two offspring types, ABC and CD, in equal proportions. Thus, the most likely interpretation was that in plants **K** and **L** there were two unlinked loci coding for cytosolic PGI.

To ascertain that these plants actually were homozygous at the standard locus, some further crosses were performed. (The details of these crosses are not shown.) The result was as expected: the complex *b′c′* gene mapped to an additional locus precisely as the earlier characterized alleles *b′* and *c′*, whereas the *a* alleles mapped to the standard *PgiC* locus.

Homology relationship between additional genes

To clarify whether or not the complex *b′c′* gene belonged to the same locus as the earlier studied *b′*

Table 6 Segregation of phosphoglucoisomerase (PGI) phenotypes in *Festuca ovina*. Test of the hypothesis that plant **L** has two unliked loci, one of which carries a complex gene. Plant ABCD (from the cross **L** × **D**) crossed to test plant **C**

	Offspring genotypes							
Cross	*a*/*c b′c′*/–	*c*/*d b′c′*/–	*a*/*c*	*c*/*d*	Others	Total	χ_3^2	*P*
c/*c* × *a*/*d b′c′*/–	14	15	14	9	9†	52 (61)	1.692	NS

† Nine *c*/*c* (probably derived by selfing).

Table 7 Segregation of phosphoglucoisomerase (PGI) phenotypes in *Festuca ovina*. Test of the allelic relationship between gene b' from plant **P** and the complex gene $b'c'$ from plant **L**. Plant $d/d\ b'/b'c'$ was derived from the cross $[(\mathbf{P} \times \mathbf{D}) \times (\mathbf{D} \times \mathbf{L})]$

	Offspring genotype				
Cross	$d/d\ b'/-$	$d/d\ b'c'/-$	Total	χ^2_1	P
$d/d\ b'/b'c' \times d/d$	26	23	49	0.184	NS

and c' alleles, a second generation cross was performed which combined b' from plant **P** and $b'c'$ from plant **L**. The genotype of the resulting plant would be either $d/d\ b'/b'c'$ or $d/d\ b'/-\ b'c'/-$, depending on the correct homology relationship between the additional genes. This plant was then crossed to test plant **D** and the results are given in Table 7. The χ^2 value for the goodness-of-fit between the data and the two-locus model was 0.184 (1 d.f.; $P > 0.05$), whereas the corresponding value for the three-locus model was 49.367 (3 d.f.; $P < 0.001$). Thus, the $b'c'$-complex appeared to be homologous to the earlier recognized b' and c' alleles at the extra locus.

Segregation ratios

In most of the crosses the segregation ratios were as expected under the simple genetic models assumed, but distortions were occasionally found. In all such cases exactly the same genetic material which in one cross gave rise to segregation distortion, behaved in a perfectly Mendelian way in other crosses. This indicates that the observed distortion was caused by some biological effect not necessarily associated with the additional Pgi genes as such. Two examples will be given.

The first concerns the cross between a plant with phenotype ACD and test plant **D**, reported in Table 2, with excess of progeny carrying the c' allele. This deviation was, however, only seen when the multibanded plant was used as mother plant; in the reciprocal cross no such skewness was observed.

The second example concerns plant **K** which gave very skewed results when crossed both to plants **C** and **D**, as seen in Table 5. A significant excess of progeny with additional genes was found from crosses to both test plants, irrespective of whether the extra gene was transmitted by the egg cell or by the pollen. However, in an additional cross ($a/a\ b'c'/- \times a/a$) exactly the same complex gene was transmitted to 37 out of 77 investigated offspring, thus producing a clear 1:1 segregation ($\chi^2_1 = 0.117$; $P > 0.05$).

Phenotype comparisons

Three specificities were detected at the additional PGI locus: b', c' and $b'c'$. As far as can be judged from zymograms obtained after electrophoretic runs with different buffers (results not shown), the enzymes encoded by the genes in the complex $b'c'$ have exactly the same electrophoretic mobilities as those encoded by the simple alleles b' and c'.

Allele c' produces a monomeric band with the same migration rate as one c allele at the standard *PgiC* locus, as seen, for example, in plant **C**.

The mobility of the monomeric band encoded by the b' allele was slightly slower than that of the bands produced by the b alleles at the main locus in plants **M** and **N**. However, another b allele at the standard *PgiC* locus, with exactly the same migration properties as the b' allele could be seen in other plants from southern Sweden.

Other species

The isozyme patterns produced by related diploid species (*F. altissima*, *F. polesica*, *F. pratensis* and *L. perenne*) did not indicate the regular presence of any additional genes for cytosolic PGI. *Festuca pratensis* and *F. polesica* both had zymogram patterns very similar to *F. ovina*. Both species had a large number of PGI alleles and heterozygosity was high. In *F. altissima*, however, only two alleles for cytosolic PGI were found in the material investigated. In each of the two sampled populations a specific allele had gone to fixation and no heterozygote was thus found.

Cytology

The cytological investigation left no doubt about the five multibanded plants having a diploid constitution, as they all had a normal set of 14 A chromosomes. However, in three of the investigated plants (**K**, **M** and **P**), as well as in test plant **D**, small B chromosomes were seen. No B chromosomes were found in plants **L** and **M**. There is thus no reason to suggest that the B chromosomes are in any way associated with the gene duplications discussed above.

Conclusion

The analysis of the multibanded plants has indicated the existence of a second, polymorphic locus for cytosolic PGI in *F. ovina*. This locus will be designated *PgiC-2* (and the ′ sign on its alleles will be dropped); the main locus is renamed *PgiC-1*. Giving the genotypes of the investigated plants at *PgiC-1* and *PgiC-2*, the result of the genetic analysis can be summarized as follows:

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

Thus, in the present restricted sample of five investigated plants, three alleles besides a null-allele were recognized at *PgiC-2*. Of these alleles, one appeared to be a closely linked complex of a combination of the other two alleles.

Discussion

The polymorphic duplication found in *F. ovina* is interesting in that it relates to both questions raised in the Introduction about duplications in general: the molecular and chromosomal processes important for their creation and sequence evolution, and the fitness interactions leading to their spread and maintenance in the species.

That there is an occasional extra locus encoding cytosolic PGI in this species has been amply demonstrated in the crosses reported on above. The genetic interpretation of the results has been confirmed with respect to its basic structure, and the problems that appeared in the analysis can be given satisfactory explanations as follows.

1 Among the 1467 offspring plants analysed, 92 plants were considered as being derived from selfing (see Tables 3, 5 and 6) and were therefore excluded from the analysis. *Festuca ovina* is in nature known to be largely self-sterile, but an increased rate of selfing in the greenhouse is not surprising given the known positive effect of high temperature on pseudo-compatibility (Lundqvist, 1990).

2 A second problem was caused by two plants (seen in Tables 3 and 5) that did not fit the expected outcome of the crosses. The most plausible explanation of these is that they were derived from 'illegal' fertilization events, not involving the intended paternal plants. Considering that *F. ovina* is a wind-pollinated species with very light, mobile pollen grains (Urquhart, 1971), it seems reasonable to accept this level of experimental mishap (two out of 1467 offspring plants).

3 Problems were also created by the occasional instances of non-Mendelian segregation (see above). This phenomenon could be explained in many different ways but most probably it has something to do with the self-incompatibility system of the species. In several grass species *PgiC* is known to be linked to one of the two self-incompatibility loci (Hayman, 1992), with a reported recombination frequency between the loci of about 20% (Cornish *et al.*, 1980; Leach & Hayman, 1987). Occasional instances of skewed segregation ratios are therefore to be expected, in particular because an accumulation of common incompatibility alleles in the crosses was unavoidable, especially in later generations, owing to the limited number of plants used.

The genetic similarity between the two loci encoding cytosolic PGI indicates that *PgiC-2* represents a recent duplication of the ancestral locus *PgiC-1*. An argument favouring this view is that the enzymatic subunits produced by *PgiC-2* readily form active dimeric enzymes with the products encoded by *PgiC-1*. Furthermore, the two alleles at the duplicate locus produced enzymes with electrophoretic migration rates indistinguishable from products produced by the standard locus. The absence of duplicated loci for cytosolic PGI in any of the closely related species, also points to a recent origin of the duplication in *F. ovina*. However, from our electrophoretic results it cannot formally be excluded that at one time the progenitors of the *F. ovina* of today all carried active genes at two loci and that later null alleles have accumulated at one of them, until now only a small minority of plants carry active genes at this locus. There are, however, no specific results indicating that this suggestion should be true. Those plants that do not have active genes at the duplicate locus are thus not expected to carry any inactive homologue, but instead show the original chromosomal nucleotide sequence into which an active gene once moved.

This interpretation raises the question of how the genetic variation at *PgiC-2* has originated. Two simple suggestions can be made. The first is that a single active gene was transposed and that this gene later, by mutations, gave rise to the isozyme variation seen today. That the electrophoretic migration pattern of the resulting mutational products exactly corresponds to variants at the primary locus would then be because of the intrinsic properties of the amino acids in the enzyme, rather than because of any interesting genetic process. The second explanation is that the two loci, *PgiC-1* and *PgiC-2*, have been in 'genetic contact' with each other after the duplication event, whereby allelic information at the first locus has been transferred to the second, e.g. by gene conversion.

Drawing from the well-defined material produced in the crossing experiments, DNA investigations have been initiated using primers based on sequence information from other grass species. Of particular interest are the preliminary results which indicate that all expressed genes at the *PgiC-2* locus are associated with a specific, unusually short, intron between exons 4 and 5 which has so far not been seen in any *PgiC-1* allele (L. Ghatnekar & B. O. Bengtsson, unpubl. data). The sheer existence of

an intron indicates that *PgiC-2* has not originated via a retrotransposition event (unlike, e.g. *jingwei*; see Long & Langley, 1993). The absence of the specific intron from plants not showing active genes at *PgiC-2* adds evidence to the suggestion that this gene is indeed the result of a gene duplication into a new site. The question of whether gene conversion has occurred between the two *PgiC* genes after the duplication event remains as yet unknown.

When it comes to the evolutionary processes by which the duplicated gene is maintained in the population, it is easy to speculate that there is a selective advantage to *F. ovina* plants that carry many different active isozyme variants of cytosolic PGI (Riddoch, 1993). Having two polymorphic loci is then, obviously, selectively favoured. In a first step towards assessing the strength of this suggestion, we will study the frequency of the duplicated genes in some natural populations, for which purpose the existence of a promising DNA marker for active genes at *PgiC-2* would be very useful.

Acknowledgements

This paper is dedicated to the memory of Professor Arne Lundqvist. I thank Professor Bengt O. Bengtsson for his support and help with the manuscript. Peder Weibull collected the seeds and Bengt Jacobsson took care of the plants in the greenhouse. This work was financially supported by a grant to Bengt O. Bengtsson from the Swedish Natural Science Research Council.

References

BENGTSSON, B. O., WEIBULL, P. AND 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.

CORNISH, M. A., HAYWARD, M. D. AND LAWRENCE, M. J. 1980. Self-incompatibility in ryegrass III. The joint segregation of *S* and *PGI-2* in *Lolium perenne* L. *Heredity*, **44**, 55–62.

GOLDRING, A., ZAMIR, D. AND DEGANI, C. H. 1985. Duplicated phosphoglucose isomerase genes in avocado. *Theor. Appl. Genet.*, **71**, 491–494.

GOTTLIEB, L. D. 1977. Evidence for duplication and divergence of the structural gene for phosphoglucoisomerase in diploid species of *Clarkia*. *Genetics*, **86**, 289–307.

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

HAYMAN, D. L. 1992. The S–Z incompatibility system. In: Chapman, G. P. (ed.) *Grass Evolution and Domestication*, pp. 117–137. Cambridge University Press, Cambridge.

LANGE, B. W., LANGLEY, C. H. AND STEPHAN, W. 1990. Molecular evolution of *Drosophila* metallothionein genes. *Genetics*, **126**, 921–932.

LEACH, C. R. AND HAYMAN, D. L. 1987. The incompatibility loci as indicators of conserved linkage groups in the Poaceae. *Heredity*, **58**, 303–305.

LIU, F., CHARLESWORTH, D. AND KREITMAN, M. 1999. The effect of mating system differences on nucleotide diversity at the phosphoglucose isomerase locus in the plant genus *Leavenworthia*. *Genetics*, **151**, 343–357.

LONG, M. AND LANGLEY, C. H. 1993. Natural selection and the origin of *jingwei*, a chimeric processed functional gene in *Drosophila*. *Science*, **260**, 91–95.

LOOTENS, S., BURNETT, J. AND FRIEDMAN, T. B. 1993. An intraspecific gene duplication polymorphism of the urate oxidase gene of *Drosophila virilis*: a genetic and molecular analysis. *Mol. Biol. Evol.*, **10**, 635–646.

LUNDQVIST, A. 1990. The complex S-gene system for control of self-incompatibility in the buttercup genus *Ranunculus*. *Hereditas*, **113**, 29–46.

PRENTICE, H. C., LÖNN, M., LEFKOVITCH, L. P. AND RUNYEON, H. 1995. Associations between allele frequencies in *Festuca ovina* and habitat variation in the alvar grasslands on the Baltic island of Öland. *J. Ecol.*, **83**, 391–402.

RIDDOCH, B. J. 1993. The adaptive significance of electrophoretic mobility in phosphoglucose isomerase (PGI). *Biol. J. Linn. Soc.*, **50**, 1–17

TERAUCHI, R., TERACHI, T. AND MIYASHITA, N. T. 1997. DNA polymorphism at the *Pgi* locus of a wild yam, *Dioscorea tokoro*. *Genetics*, **147**, 1899–1914.

URQUHART, C. 1971. Genetics of lead tolerance in *Festuca ovina*. *Heredity*, **26**, 19–33.

WEIBULL, P., GHATNEKAR, L., FRYKMAN, I. AND BENGTSSON, B. O. 1988. Electrophoretic variation in *Festuca ovina* L. *Agri. Hortique Genetica*, **44**, 25–37.

WEIBULL, P., GHATNEKAR, L. AND BENGTSSON, B. O. 1991. Genetic variation in commercial varieties and natural populations of Sheep's Fescue, *Festuca ovina* L. *Pl. Breed.*, **107**, 203–209.