A horizontally transferred nuclear gene is associated with microhabitat variation in a natural plant population

Honor C. Prentice1, Yuan Li1, Mikael Lönn2, Anders Tunlid1 and Lena Ghatnekar1

1Department of Biology, Lund University, 223 62 Lund, Sweden
2School of Natural Sciences, Technology and Environmental Studies, Södertörn University, 141 89 Huddinge, Sweden

Horizontal gene transfer involves the non-sexual interspecific transmission of genetic material. Even if they are initially functional, horizontally transferred genes are expected to deteriorate into non-expressed pseudogenes, unless they become adaptively relevant in the recipient organism. However, little is known about the distributions of natural transgenes within wild species or the adaptive significance of natural transgenes within wild populations. Here, we examine the distribution of a natural plant-to-plant nuclear transgene in relation to environmental variation within a wild population. *Festuca ovina* is polymorphic for an extra (second) expressed copy of the nuclear gene (*PgiC*) encoding cytosolic phosphoglucose isomerase, with the extra *PgiC* locus having been acquired horizontally from the distantly related grass genus *Poa*. We investigated variation at *PgiC* in samples of *F. ovina* from a fine-scale, repeating patchwork of grassland microhabitats, replicated within spatially separated sites. Even after accounting for spatial effects, the distributions of *F. ovina* individuals carrying the additional *PgiC* locus, and one of the enzyme products encoded by the locus, are significantly associated with fine-scale habitat variation. Our results suggest that the *PgiC* transgene contributes, together with the unlinked ‘native’ *PgiC* locus, to local adaptation to a fine-scale mosaic of edaphic and biotic grassland microhabitats.

1. Introduction

There is widespread unease about the artificial use of horizontal gene transfer (HGT) in the process of creating genetically modified crops [1], yet the acquisition of genetic material by means other than sexual reproduction and hybridization (vertical gene transfer) also occurs naturally. Most examples of HGT to eukaryotes involve genes from bacteria [2,3], but reports of distantly related multicellular eukaryotes as donors are steadily increasing [4–6]. In flowering plants (Magnoliophyta), there are many known instances of interspecific exchange of mitochondrial genes and a few cases involving plastid genes [4,7,8]. However, reports of natural examples of the horizontal transfer of nuclear genes, with flowering plants as both donor and recipient, are still scarce and, in most cases, the function of the acquired genes is unknown [4,8].

The transfer of a *Mu*-like transposon between the grass genera *Setaria* (millet) and *Oryza* (rice) [9] and the transfer of retrotransposons between *Oryza* species [10] are two examples where nuclear genetic material of unknown function has been transferred between flowering plants. The high frequency of horizontal transfer of mitochondrial genes between parasitic plants and their hosts is attributed to the direct, long-term cell-to-cell contact between the parasite’s haustoria and the host plant [4,7,11], and there are also a few reports of the exchange of nuclear genes between host plants and their distantly related plant parasites. For example, the crop pest *Striga hermonthica* has acquired a nuclear gene of unknown function from a host grass (*Sorghum*) [11] and another devastating...
root-parasitizing crop pest, Orobanche aegyptiaca, has acquired a strictosidine synthase-like gene from the Brassicaceae [12].

The two well-supported examples of the horizontal transfer of nuclear genes between flowering plants, where the function of the transferred gene is retained, both come from the grass family Poaceae. The first example is the acquisition, by taxa of the transferred gene is retained, both come from the grass family Poaceae. The first example is the acquisition, by taxa of the Brassicaceae [12]. The second example comes from the species that is the focus of the present study, Festuca ovina (sheep’s fescue), where an extra, full-sequence and fully functional copy of the locus encoding cytosolic phosphoglucose isomerase (PGIC) has been horizontally acquired from a species in the genus Poa [14,15].

Horizontally acquired genetic material may disrupt regulatory or physiological functions within the recipient species [16] and, unless it confers a selective advantage, a transferred, functional gene is expected to degenerate into a non-expressed pseudogene within the recipient species [2,17,18]. Although the selective value of many horizontally acquired genes is unclear [4], the adaptive importance of a transferred gene can, in some cases, be inferred from its metabolic role together with the ecology and life history of the recipient species [3].

In the coffee berry borer (Hypothenemus hampei), the acquisition of a bacterial gene coding for a mannanase (a type of enzyme that is otherwise lacking in insects) appears to be clearly adaptive in that it has allowed the beetle to exploit a novel niche and to feed exclusively within coffee berried—causing major economic losses [19]. Another example where an adaptive role for a transgene may be inferred from the ecology of the recipient species is the case of the pea aphid (Acyrthosiphon pisum), where populations contain both red and green clones. A horizontally acquired, fungal gene, carotenoid desaturase, is responsible for the production of the carotenoid that is the basis of the red body colour [20,21].

The body colour polymorphism is predominantly under genetic control [20,22], influences susceptibility to predators and parasites, and may be maintained by frequency-dependent selection [20,23]. It seems likely that the fungal transgene contributes to fitness differences between red and green aphids. In the grass genus Alloteropsis, there is a strong presumption of an integrated, overall adaptive role for four, independent, acquisitions of two key genes involved in C4 photosynthesis. The genes were optimized for C4 function within distantly related grass lineages before their transfer to Alloteropsis [13]. It is suggested that the superior performance of the C4 transgenes led to their rapid fixation in the recipient Alloteropsis taxa, where the corresponding vertically acquired genes were less well adapted to a function in the C3 pathway [13].

The detection of a functionally important transgene within an organism may allow inferences about the putative adaptive value of the acquisition or provide major insights into macroevolutionary processes [3,8]. However, there are, so far, no studies that have demonstrated associations between the within-individual presence of horizontally acquired genes, their expressed products, and habitat variation within natural populations of plants or, to the best of our knowledge, animals. Indeed, little is known about the frequency distributions of natural transgenes within and between populations in the wild.

PGIC (EC 5.3.1.9) is a dimeric enzyme which catalyses the interconversion of glucose-6 phosphate and fructose-6-phosphate and plays a key regulatory role in glucose metabolism [24,25]. PGIC has been studied extensively in natural populations, and represents a classic example of a metabolically important and notably polymorphic enzyme, which has also been shown (in mammals) to have moonlighting functions that are separate from its role in energy metabolism [24,26]. There is compelling evidence, from a wide range of organisms, that PGIC isozyme polymorphisms are strongly associated with environmental variation in the wild, and that these habitat associations reflect fitness differences [27–30].

In parts of north-east Europe, isozyme analyses reveal that two, unrelated [31], copies of the locus coding for PGIC may be present in diploid populations of the grass F. ovina [31–33]. Recent genetic and phylogenetic studies provide substantial evidence that the second copy of the PgIC locus in F. ovina has been horizontally derived from the distantly related grass genus Poa [14,15]; the acquisition of this extra copy is estimated to have occurred at 600 000 years ago [14]. Whereas all individuals contain the ‘native’ PgIC1 locus, the presence/absence of the functional, transgenic copy of the locus, PgIC2(f), is polymorphic and the locus segregates within F. ovina populations [15,31,33]. A non-expressed pseudogene, PgIC2(ϕ), also derived from Poa, is usually present in individuals carrying PgIC2(f) [15,33,34]. The pseudogene contains a 29 bp deletion (electronic supplementary material, table S1) and, in contrast to PgIC1 and PgIC2(f), does not encode a PGIC isozyme product [34]. In individuals that contain the functional, transgenic PgIC2(f) locus, the monomeric products coded for by both PgIC loci form, in addition to homodimers, all possible, metabolically active, heterodimeric combinations [31]. The overall function of PGIC in individuals with the transgenic locus reflects, therefore, the properties of the interlocus heterodimers as well as those of the homodimers produced by both loci and the intralocus heterodimers.

Festuca ovina is perennial, widespread and common, strongly outbreeding [35,36], and has wind-dispersed pollen and seeds. Many studies have demonstrated fine-scale adaptive differentiation within local populations of grasses in the face of extensive gene flow [37–39], and earlier studies of natural populations of F. ovina suggest that PGIC isozymes may be involved in local adaptation [32,40]. Festuca ovina occupies the full range of moist and dry, neutral and base-rich grassland microhabitats which form a fine-scale mosaic that is spatially replicated in sites throughout the area of steppe-like ‘alvar’ grasslands on the Baltic island of Öland, Sweden [32]. The species shows high levels of PGIC isozyme polymorphism in the alvar grasslands and, despite large population sizes and extensive gene flow, different PGIC isozyme electromorphs are significantly associated with different grassland microhabitats [32]. Electromorph frequencies change, as predicted, in response to experimental habitat manipulation [40].

Although the horizontally acquired PgIC2(f) locus is expressed in F. ovina, and codes for functional isozymes, there has so far been no report that its acquisition is adaptive [5]. However, the fact that the segregating PgIC2(f) locus persists at moderate-to-high frequencies in some F. ovina populations in south-east Sweden [33] suggests that the transgene may play an adaptive role.

In this study, we developed two pairs of primers to identify the presence/absence of the PgIC2(f) locus in F. ovina individuals, and used data from five, geographically
separated sites on Öland to investigate whether the presence of PgiC2(f) was associated with abiotic and biotic microhabitat variables within the fine-scale mosaic of grassland communities. The unique, natural replication of the different microhabitats within sites throughout the Öland study area allows the effects of purely spatial variation to be separated from the effects of fine-scale abiotic and biotic variation between the microhabitats. Significant associations between the presence of the PgiC2(f) locus and microhabitat variation in wild populations, after accounting for the effects of spatial variation, would provide support for the suggestion that an involvement in local adaptation may have contributed to the persistence of this natural transgene.

2. Material and methods

Full descriptions of material and methods are provided in electronic supplementary material, appendix S1.

(a) Study area

The 26 000 ha of steppe-like grasslands of the Great Alvar on the Baltic island of Öland (Sweden) are characterized by a high diversity of edaphic microhabitats that form a fine-scale mosaic of moist and dry, neutral and base-rich soils which is replicated at sites throughout the area [32]. The different microhabitats support distinct plant communities and differ in their biotic properties.

(b) Study species

Festuca ovina L. is a common, outbreeding, perennial grass with wind-dispersed pollen and seeds. Populations in the ‘alvar’ grasslands are diploid and tussock-forming. The species occupies the full range of moist and dry, neutral and base-rich microhabitats within sites throughout the Great Alvar [32].

(c) Field sampling and environmental variables

A total of 181 F. ovina individuals were sampled from microhabitats representing the extremes of the orthogonal gradients of soil pH and soil moisture that determine the variation in plant community composition within the alvar grasslands. The four microhabitats (low pH/moist, low pH/dry, high pH/moist, high pH/dry) were defined by their plant communities [32]: three to five individuals were collected from each of two 1 × 1 m plots in each microhabitat, within each of five geographically separated sites on Öland (electronic supplementary material, table S2). The replicated, balanced sampling design, with the same mosaic of sampled microhabitat categories in each of the sites, allowed purely geographical site effects to be accounted for in the analyses. The 40 sampling plots were characterized by their site-membership and by nine environmental variables: the within-plot proportions of within-individual presence of PgiC1 and PgiC2(f) electromorphs and of the PgiC2 pseudogene.

(d) Locus-specific primers for PgiC1 and PgiC2(f)

Two locus-specific primer pairs were used to detect the within-individual presence of PgiC1 and two locus-specific primer pairs were used to detect the presence of the horizontally transferred PgiC2(f) locus (electronic supplementary material, table S1) in the 181 wild-sampled individuals, and in 12 reference individuals with known PgiC1 and PgiC2(f) genotypes (electronic supplementary material, tables S2 and S6). The primer pairs for PgiC2(f) were designed to preclude amplification of the PgiC2ψ pseudogene.

(e) Isozyme electrophoresis

Starch gel electrophoresis [32] was used to score PGI isozyme electromorph phenotypes in the 181 wild-sampled individuals and the 12 reference individuals (electronic supplementary material, tables S2 and S6). The electrophoretic mobilities of the allelic products coded for by the two loci overlap [31] and PgiC1 and PgiC2(f) genotypes could not be inferred from the electrophoretic phenotypes of the wild-collected material in this study. Isozyme data used in the statistical analyses were therefore based on within-individual presence/absence of specific electromorph bands (electronic supplementary material, table S2), without assumptions about which of the two PgiC loci coded for the bands. Six electromorphs were represented in the analysed material (electronic supplementary material, table S7).

(f) Statistical analyses

All analyses were based on data from 181 individuals. A series of generalized linear models (GLMs) were used to analyse: (i) the relationships between within-plot counts of PgiC2(f) (binomial response variable) and the 10 site/environmental explanatory variables, and (ii) the relationships between within-plot counts of individual isozyme electromorphs (binomial response variables) and the within-plot proportions of PgiC2(f), together with the 10 site/environmental variables (explanatory variables). Explanatory variables in the final models were selected using a stepwise procedure, based on the Akaike information criterion and including two-way interactions, followed by further model simplification. Only variables that gave a significant change in deviance if deleted from the model were retained in the final model. A further series of GLMs was used (iii) to analyse the relationships between the within-individual presence of PgiC2(f) (Bernoulli response variables) and each of the individual isozyme electromorphs. Correspondence analysis and principal components analysis were used to summarize gradients of variation in plant community composition among the sampling plots and to display relationships between genetic and environmental variables. All analyses were carried out in the R environment [41].

3. Results

Validation of primer-pair specificity, using the 12 reference individuals with known PgiC1 and PgiC2(f) genotypes [31], showed that the loci detected by the locus-specific PgiC1 and PgiC2(f) primers agreed in all cases with the genotypes of the reference individuals (electronic supplementary material, table S6). Both primer pairs for PgiC2(f) gave identical presence/absence results for PgiC2(f) in each of the 181 individuals.

The PgiC2(f) locus was present in 58 (32%) of the 181 individuals. The native PgiC1 locus, also identified by two pairs of locus-specific primers (electronic supplementary material, table S1), was present in all individuals (electronic supplementary material, table S2). Although the sampling sites were separated from each other by up to 30 km (electronic supplementary material, table S2), within-plot counts of PGI electromorphs and of the PgiC2(f) locus were not significantly related to the geographical location (site-membership) of the samples (tables 1 and 2).

The analysis of deviance in table 1 shows that within-plot counts of the presence of the transgenic PgiC2(f) locus are positively associated with the biotic variables vegetation height
and the percentage cover of lichens (an indicator of dry microhabitats: cf. figure 1), and there is a significant interaction between the percentage cover of mosses and site-membership. Data on the isozyme electromorph phenotypes of the studied individuals allow associations between the presence of PgiC2(f) and microhabitat variables to be viewed in the context of the active enzyme products encoded by both PgiC loci. The within-individual presence of two of the PGIC isozyme electromorphs (out of the six present in the material; electronic supplementary material, table S7) was significantly associated with the within-individual presence of the transgenic PgiC2(f) locus (table 3). Of these two electromorphs, electromorph 7 was present in only four individuals, all of which contained PgiC2(f). In contrast, electromorph 4, which was highly significantly associated with the presence of PgiC2(f), occurred frequently in the sampled material and was present in 63 (35%) of the individuals. The 79% probability of it occurring in individuals also carrying PgiC2(f) indicates that this isozyme is encoded predominantly by the transgene.

Table 1. Within-plot associations between PgiC2(f) and environmental variables. Analysis of deviance: generalized linear model assuming a binomial error distribution. The response variable is within-plot counts of F. ovina individuals containing PgiC2(f) (n = 40). Explanatory variables in the full model were site-membership and nine within-plot environmental variables. (+) indicates a positive association.

<table>
<thead>
<tr>
<th>explanatory variable</th>
<th>d.f.</th>
<th>likelihood ratio</th>
<th>significance probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>site</td>
<td>4</td>
<td>7.60</td>
<td>0.107</td>
</tr>
<tr>
<td>moss cover</td>
<td>1</td>
<td>0.93</td>
<td>0.334</td>
</tr>
<tr>
<td>lichen cover</td>
<td>1</td>
<td>4.36</td>
<td>0.037* (+)</td>
</tr>
<tr>
<td>veg height</td>
<td>1</td>
<td>4.16</td>
<td>0.041* (+)</td>
</tr>
<tr>
<td>site × moss</td>
<td>4</td>
<td>13.34</td>
<td>0.010**</td>
</tr>
<tr>
<td>residual</td>
<td>28</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < 0.05, **p < 0.01.
Table 2. Within-plot associations between PGIC electromorphs, $PgiC2(f)$, and environmental variables. Analyses of deviance: separate generalized linear models for individual PGIC electromorphs. Response variables are within-plot counts of the electromorph (binomial error distribution, $n = 40$ in all four models). Explanatory variables are within-plot proportions of individuals with $PgiC2(f)$, site-membership, and within-plot habitat variables. The direction (+ or −) of an association is given in parentheses.

<table>
<thead>
<tr>
<th>explanatory variable</th>
<th>d.f.</th>
<th>electromorph</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>4$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$PgiC2(f)$</td>
<td>1</td>
<td></td>
<td>&lt;0.001*** (+)</td>
<td>not used$^d$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>site</td>
<td>4</td>
<td></td>
<td>0.629</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>moisture</td>
<td>1$^b$</td>
<td></td>
<td>0.021* (−)</td>
<td>0.042* (+)</td>
<td>0.003** (−)</td>
<td></td>
</tr>
<tr>
<td>pH</td>
<td>1$^c$</td>
<td></td>
<td>0.012* (+)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bare ground</td>
<td>1</td>
<td></td>
<td>&lt;0.001*** (−)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>grass cover</td>
<td>1</td>
<td></td>
<td>0.040* (+)</td>
<td>0.006** (−)</td>
<td>0.078</td>
<td></td>
</tr>
<tr>
<td>forb cover</td>
<td>1</td>
<td></td>
<td>&lt;0.001** (−)</td>
<td>0.416</td>
<td></td>
<td></td>
</tr>
<tr>
<td>moss cover</td>
<td>1</td>
<td></td>
<td>0.977</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lichen cover</td>
<td>1</td>
<td></td>
<td>0.194</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>soil depth</td>
<td>1</td>
<td></td>
<td>0.116</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>site × moss</td>
<td>4</td>
<td></td>
<td>0.014* (−)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>site × veg height</td>
<td>4</td>
<td></td>
<td>&lt;0.001***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>moisture × lichen</td>
<td>1</td>
<td></td>
<td>0.020*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>moisture × forb</td>
<td>1</td>
<td></td>
<td>0.005**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>grass × veg height</td>
<td>1</td>
<td></td>
<td>0.004**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>residual d.f.</td>
<td></td>
<td></td>
<td>34</td>
<td>35</td>
<td>38</td>
<td>19</td>
</tr>
</tbody>
</table>

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.
$^a$An additional analysis was carried out for electromorph 4, omitting $PgiC2(f)$ as an explanatory variable.
$^b$Categorical variable: (+), moist; (−), dry.
$^c$Categorical variable: (+), high; (−), low.

Table 3. Associations between the presence of PGIC electromorphs and presence of the $PgiC2(f)$ locus in Festuca ovina individuals. Analyses of deviance: separate generalized linear models for individual PGIC electromorphs (Bernoulli error distribution, $n = 181$, residual d.f. = 179 in all models). The response variable in each model is the within-individual presence/absence of an electromorph; the explanatory variable is the presence/absence of $PgiC2(f)$ in the same individual.

<table>
<thead>
<tr>
<th>electromorph</th>
<th>deviance change$^a$</th>
<th>significance probability</th>
<th>fitted value$^b$ (s.e.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.74</td>
<td>0.053</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.25</td>
<td>0.616</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>75.98</td>
<td>&lt;0.001***</td>
<td>0.793 (0.053)</td>
</tr>
<tr>
<td>5</td>
<td>0.01</td>
<td>0.911</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.01</td>
<td>0.932</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>9.30</td>
<td>0.002**</td>
<td>0.069 (0.033)</td>
</tr>
</tbody>
</table>

*$p < 0.01$, **$p < 0.001$.
$^a$Tested by comparison with a $\chi^2$ distribution.
$^b$The probability of finding the electromorph in an individual where $PgiC2(f)$ is present (shown for models with significant effects).

Analysis of deviance models showing associations between PGIC electromorphs 1, 2, and 4, respectively, and within-plot proportions of the $PgiC2(f)$ locus, together with site and environmental variables are presented in table 2. No explanatory variables were significant in a model for electromorph 5, and models for the rare electromorphs 6 and 7 were severely underdispersed: these three models were not interpreted and are not presented. The first three analysis of deviance models (for electromorphs 1, 2, and 4; table 2) included the proportion of individuals carrying the $PgiC2$ locus as an explanatory variable in the full model—in addition to the 10 site and environmental descriptors. The presence of the $PgiC2$ locus was not retained in the final models for electromorphs 1 or 2. As in earlier studies [32,40], both these electromorphs show significant associations with abiotic (pH and/or moisture) and biotic variables (the percentage covers of bare ground and different categories of vegetation cover, respectively) characterizing the grassland plant communities in the sampled microhabitats (cf. figure 1 and electronic supplementary material, table S4).

In contrast to the models for electromorphs 1 and 2, where the presence of the $PgiC2(f)$ locus did not make a significant contribution as an explanatory variable, the within-plot proportion of individuals in which the locus was present was the only significant explanatory variable in the model for electromorph 4 (table 2)—supporting the conclusion (cf. table 3) that the PGIC enzyme electromorph 4 is strongly associated with the presence of the transgenic locus. An additional model for electromorph 4, in which $PgiC2(f)$ was not included as an explanatory variable, reveals multiple,
significant associations between within-plot counts of the electromorph and environmental variables (table 2). Electromorph 4 is associated with dry microhabitats and shallow soils (table 2 and figure 1c), with significant interaction effects involving environmental variables that reflect aspects of moisture status and vegetation height.

4. Discussion

In *F. ovina*, the presence/absence of the PgiC2(f) transgene and the PGIC isozyme electromorph 4, with which it is associated, are significantly related to local microhabitat variation—after accounting for spatial variation. The relationships between the within-individual presence of the segregating PgiC2(f) locus, and variables characterizing the four grassland microhabitats suggest that the transgenic locus may be contributing, together with the unlinked, native PgiC1 locus, to local adaptation within the fine-scale complex of edaphic habitats in the alvar grasslands.

Whereas PgiC1 and PgiC2(f) are unlinked [31], we cannot exclude the possibility that the significant habitat associations shown by the transgene reflect an adaptive advantage of other loci or genetic material (linked or unlinked) that is in linkage disequilibrium with the PgiC2(f) locus. If this is the case, then the PgiC2(f) locus, nevertheless, represents a marker for a transgenic insertion that is significantly associated with local environmental variation.

There are a number of considerations which suggest that high levels of genome-wide linkage disequilibrium are unlikely in *F. ovina*. The species is a near-obligate outbreeder [35,36] and closely related grasses have been shown to have a gametic self-incompatibility system [42]. Linkage disequilibrium is expected to be low in obligate outbreeders [43,44], and studies of other self-incompatible grasses reveal low levels of linkage disequilibrium, even within genes [44–47]. Results from a grassland experiment suggest low levels of genomic linkage disequilibrium in British *F. ovina* [48], and a recent study of the *F. ovina* PgiC1 locus reveals high levels of recombination ($R_{sd} = 22$) and low levels of linkage disequilibrium within the gene—suggesting frequent recombination and low levels of linkage disequilibrium also in other parts of the genome [49].

The sampling design in the present study allows site effects to be separated from the effects of within-site microhabitat variation. The fact that no significant main effects of site were detected for any of the PgiC electromorphs, or for the presence of the transgene, supports the conclusion that gene flow among the Öland populations is extensive, that the overall effective population size within the study area is large—as predicted by the species’ breeding system and dispersal syndrome—and that the geographically well-separated sites do not show a local limitation of population size that would result in a Wahlund effect and locally raised levels of linkage disequilibrium.

A transgene that is not adaptively relevant will decay into a pseudogene [2,17,18], and genes that are disadvantageous or deleterious will be purged from a large, highly outbreeding population [50]. The acquisition of PgiC2(f) from *Poa* occurred ca 600 000 years ago [14], yet the transgene persists at a frequency of around 30% in Öland populations of *F. ovina*—suggesting that it (or a closely linked genetic factor) may be being maintained by selection.

Although the two PgiC loci in *F. ovina* are unlinked and segregate independently [31], the function of the PgiC2(f) transgene is not independent from that of the native PgiC1 locus, and the presence of the transgene in an individual results in the formation of all possible combinations of functional dimeric products involving the monomers coded for by both loci [31]. The fact that the within-individual presence of isozyme electromorph 4 is associated with the presence of PgiC2(f) (table 3), and because this electromorph is itself strongly associated with microhabitat variables (table 2), it is possible that the relationship between PgiC2(f) and drier microhabitats mainly reflects isozyme–habitat associations involving the specific properties of electromorph 4. However, a possible adaptive role for the PgiC2(f) transgene is more likely to reflect the integrated effects of variation at both PgiC loci than an independent role for the transgene and its products. The presence of an additional copy of PGIC may be advantageous in that it contributes to a higher overall production [51] of the metabolically important PGI in individuals with the transgene. In addition, the within-individual presence of the transgene not only enhances the overall level of within-individual allelic polymorphism, but also confers a higher diversity of both homodimeric and intra- and interlocus heterodimeric enzyme products which may enhance the ability to fine-tune individual function in different microhabitats [31].

PGI plays a central and essential metabolic role [24] and, because the overall function of the glycolytic pathway would be compromised if the interlocus PGIC heterodimers were not fully functional, strong selection against the transgene would be expected, unless the additional gene made a positive, adaptive contribution to the overall function of PGIC.

Successful examples of HGT are often associated with major niche shifts, with the acquired genes allowing organisms to exploit novel niches or extreme habitats [3,18,52]. For example, a bacterial enzyme allows the coffee berry borer to exploit a novel niche and feed on a resource that is not available to other insects [19], and it is suggested that an unconventional photoreceptor, derived from hornworts, may have promoted the evolutionary diversification of ferns under low-light conditions within shaded habitats [6].

The present study has a finer resolution, with a focus on the within-individual occurrence of a segregating transgene in wild populations rather than on a horizontally acquired gene that is fixed within a species or group of organisms. The study suggests that, at the intraspecific level, natural transgenes may also be involved in fine-scale adaptation within locally heterogeneous environments. Rather than being involved in a major niche shift, the presence of the fully functional PgiC2(f) locus appears to be enhancing the niche breadth of *F. ovina* by contributing to the ability of populations to exploit a wide range of microhabitats within an ecologically diverse, local complex of grassland plant communities.

The locus coding for PGI represents a strong ‘universal’ candidate gene that is related to fitness in a wide range of eukaryotes [27,30]. In *F. ovina*, the PgiC2(f) transgene is intricately involved in the overall relationship between variation in PGI and microhabitat variation, and the fact that enzyme products coded for by both the unlinked PgiC loci are significantly associated with environmental variation supports the suggestion that the PgiC2(f) transgene may be contributing, together with the native PgiC1 locus, to local adaptation. Together, the two PgiC loci endow *F. ovina* with a high level of isozyme polymorphism that is available for short-term [2] fine-tuning to the different environmental niches that are available within a fine-scale habitat mosaic.
References


14. L. C. P. from The Swedish Research Council. We thank Björn Canback, Deborah Charlesworth, Norman C. Ellstrand, Jenny Flageamblad, Bengt Hansson, Christopher Wheat, and Raj Whitlock for valuable discussions and for their comments on the manuscript, and Tomas Johansson for technical advice. The ‘Station Linne’ research station provided a base for the fieldwork on Oland.

Funding. The work was supported by grant no. 621-2008-5617 (to H.C.P.) from The Swedish Research Council.

Acknowledgements. We thank Björn Canback, Deborah Charlesworth, Norman C. Ellstrand, Jenny Flageamblad, Bengt Hansson, Christopher Wheat, and Raj Whitlock for valuable discussions and for their comments on the manuscript, and Tomas Johansson for technical advice. The ‘Station Linne’ research station provided a base for the fieldwork on Oland.

Data accessibility. The datasets supporting this article have been uploaded as part of the electronic supplementary material.

