The Plant Journal (2008) **53**, 950–959 doi: 10.1111/j.1365-313X.2007.03388.x

Evidence of multiple horizontal transfers of the long terminal repeat retrotransposon *RIRE1* within the genus *Oryza*

Anne Roulin[1], Benoît Piegu[1], Rod A. Wing[2] and Olivier Panaud[1,*]
[1]*Laboratoire Génome et Développement des Plantes, UMR 5096 CNRS-IRD-Université de Perpignan. 52, avenue Paul Alduy. 66860 Perpignan. France, and*
[2]*Arizona Genomics Institute, Department of Plant Sciences and BIO5, University of Arizona, Tucson, AZ 85 721. USA*

Received 19 September 2007; revised 9 November 2007; accepted 14 November 2007.
*For correspondence (fax +33 04 4686 64899; e-mail panaud@univ-perp.fr).

Summary

Horizontal gene transfer, defined as the transmission of genetic material between reproductively isolated species, has been considered for a long time to be a rare phenomenon. Most well-documented cases of horizontal gene transfer have been described in prokaryotes or in animals and they often involve transposable elements. The most abundant class of transposable elements in plant genomes are the long terminal repeat (LTR) retrotransposons. Because of their propensity to increase their copy number while active, LTR retrotransposons can have a significant impact on genomics changes during evolution. In a previous study, we showed that in the wild rice species *Oryza australiensis*, 60% of the genome is composed of only three families of LTR retrotransposons named *RIRE1*, *Wallabi* and *Kangourou*. In the present study, using both *in silico* and experimental approaches, we show that one of these three families, *RIRE1*, has been transferred horizontally between *O. australiensis* and seven other reproductively isolated *Oryza* species. This constitutes a new case of horizontal transfer in plants.

Keywords: horizontal transfer, *Oryza*, LTR retrotransposons, BAC end sequences, *RIRE1*.

Introduction

The transmission of genetic information among living organisms, whether they reproduce sexually or not, usually occurs vertically from one generation to another (i.e. from parents to their progeny). However, there are several well-documented cases of transfer of genetic material between reproductively isolated species, referred to as horizontal transfer (HT; Brown, 2003; Gogarten and Townsend, 2005; Syvanen, 1994). For a long time, HT was considered to be a rare phenomenon, but several reports have shown that it is quite common in prokaryotes and has occurred recurrently during evolution (Brown, 2003; Koonin *et al.*, 2001). In fact, HTs are now considered to be a major source of evolution and speciation in bacteria (Lawrence, 2002). Because of their mobility and capacity for integration into DNA, transposable elements are particularly prone to HT. Moreover, some elements are known to bear an envelope-like coding domain (Laten *et al.*, 2003; Vicient *et al.*, 2001) and are thus structurally similar to retroviruses. Transposable elements are, therefore, good candidates for the study of HT. This could explain why many studies implicating transposable elements in this process can be found in the literature. In animals, several cases of HT have been reported, but the mechanisms remain largely unknown, particularly concerning the mode of transmission and whether or not it involves a biotic vector (Casse *et al.*, 2006; Herédia *et al.*, 2004; International Human Genome Consortium, 2001; Jordan *et al.*, 1999; Salzberg *et al.*, 2001; Scholl *et al.*, 2003; Shen *et al.*, 2003). However, in plants few cases of HT have been reported, and most involve mitochondrial genes (Mower *et al.*, 2004; Richardson and Palmer, 2007). Nevertheless, the recent demonstration of the HT of a *Mu-like* transposon between foxtail millet (*Setaria italica*) and rice (Diao *et al.*, 2006) constitutes a well described case of a nuclear-encoded gene transfer in plants.

Transposable elements are the main components of large eukaryotic genomes. Due to their 'copy and paste' mechanism of retrotransposition, one particular class, the retrotransposons, are mainly responsible for genome size increase (Vitte and Panaud, 2005). It is well known that retrotransposons, and more particularly long terminal

repeat (LTR) retrotransposons, can compose a large part of the genome of many grass species (Ammiraju *et al.*, 2007; Bennetzen and Kellogg, 1997; SanMiguel *et al.*, 1996; Zuccolo *et al.*, 2007). Recently, we showed that in the wild rice species *Oryza australiensis* 60% of the genome is composed of only three LTR retrotransposon families that have undergone extensive bursts of transposition (Piégu *et al.*, 2006). One of them, *RIRE1*, has accumulated more than 30,000 copies and thus represents about 27% of the genome, making it one of the most highly repeated transposable elements within a plant genome.

The genus *Oryza*, to which the Asian cultivated rice *Oryza sativa* belongs, is composed of 24 reproductively isolated species widely distributed throughout the world. Based on interspecific crosses and cytogenetic analysis (Morinaga *et al.*, 1964), ten genome types, six diploids (AA, BB, CC, EE, FF, GG) and four tetraploids (BBCC, CCDD, HHJJ, HHKK), have been defined, (Khush, 1997). Because of the economic importance of cultivated rice and the genetic diversity in related wild species which is putatively exploitable at the agronomic level, large amounts of genomic resources have recently been generated for 12 wild rice species by the Oryza Map Alignment Project (OMAP; http://www.omap.org/). These are composed of bacterial artificial chromosome (BAC) libraries, several hundred thousand BAC end sequences (BES) and fingerprints of the BAC clones, allowing reconstruction of the physical map of these 12 *Oryza* species (Ammiraju *et al.*, 2006). These resources, particularly the BES, thus provide the opportunity to study the evolution of plant genomes at a whole-genus level on an unprecedented scale. In this paper, we present the results of *in silico* and molecular analysis which provide strong evidence that multiple HT of the LTR retrotransposon *RIRE1* have occurred within the genus *Oryza*.

Results and discussion

Identification of the HT by in silico *analyses*

Previously we showed that the LTR retrotransposon family *RIRE1* has undergone massive transposition bursts in the genome of *O. australiensis* (EE), reaching a number of approximately 30 000 copies (Piégu *et al.*, 2006). We also tentatively studied the evolutionary dynamics of this family in the *Oryza* genus and showed that it is present in most of the *Oryza* species we surveyed. Our interpretation was that the *RIRE1* family was probably present in the ancestral genome of the *Oryza* genus and was subsequently vertically transmitted to the different evolutionary lineages of this taxonomic group. One would therefore expect that the divergence between the *RIRE1* elements found in two *Oryza* species should be congruent with the time of their radiation, which is what we observed in our present study. As an example, the average identity between the *RIRE1* elements found in the genomes of *Oryza punctata* (BB) and the reference *RIRE1* sequence from *O. australiensis* (EE) was 84%. This result corresponds to a radiation time of at least 4 Myr (using a molecular clock of 2×10^{-8}; Vitte *et al.*, 2004), which is in accordance to the radiation time estimated with the *Adh2* gene (i.e. 8.5 Myr; Figure 1). However, while performing additional *in silico* analyses for *Oryza minuta* (BBCC) and *Oryza granulata* (GG) we identified *RIRE1* homologs with surprisingly high sequence identity (i.e. 97.5%) with the reference *RIRE1* sequence from *O. australiensis* (EE). Using the same molecular clock as above, this 97.5% identity could be translated into a radiation date of 600 000 years, which is obviously incongruent with the estimated radiation dates either between *O. minuta* (BBCC) and *O. australiensis* (EE) (8.5 Myr) or between *O. granulata* (GG) and *O. australiensis* (EE) (20 Myr; Table 1 and Figure 1). The distribution of all pairwise distances between the reference sequence of *O. australiensis RIRE1* and the homologous BES found in the other *Oryza* species was also examined. This distribution was measured using sequences distributed throughout the element, including both gag and LTR regions, which are known to evolve faster than the pol region (Figure 2) and they clearly show that *O. minuta* (BBCC) and *O. granulata* (GG) are the only two species in the genus *Oryza* showing a bimodal distribution, with several identity values higher than 95% between *O. australiensis* (EE) and either *O. minuta* (BBCC) or *O. granulata* (GG) copies. All other *Oryza* species show only sequences with low identity to the *RIRE1* elements of *O. australiensis* (EE). We also performed sequence alignment and built phenetic trees with *RIRE1* homologs isolated from the OMAP BES dataset. The full length of *RIRE1* is 8.3 kb (Noma *et al.*, 1997). The BES are too short to cover such a large sequence, therefore all our *in silico* analyses are based on partial data concerning small portions of the element (subregions of 400 bp, see Experimental Procedures). All the sequences from

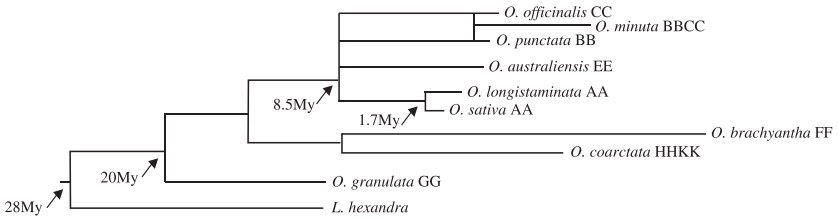


Figure 1. Phylogenetic tree of the *Oryza* genus. Radiation dates between species is based on divergence of the *Adh2* gene. Only nodes with a bootstrap value >70% are shown.

Table 1 Interspecific percentage sequence identity for *Adh2*

	aus	min	gr	punct	coar	brachy	longi
aus							
min	84.3						
gr	74.2	72.2					
punct	84.4	94.9	74.2				
coar	75.7	72.5	72.6	71.5			
brachy	59.1	55.6	65	55.5	66.3		
longi	87.6	85.5	75.9	86.0	76.7	61.8	
hex	59.4	61.4	73.4	64.8	68.2	50.5	68.2

O. minuta (BBCC) and *O. granulata* (GG) with high sequence identity to *O. australiensis* (EE) *RIRE1* form a distinct cluster which is embedded within the larger cluster of the copies corresponding to the retrotransposition burst of *RIRE1* in *O. australiensis* (EE; Figure 3a). Transposable elements are known to mutate at a much higher rate than genes, through small deletions and mutations (Ma *et al.*, 2004; Vitte *et al.*, 2007). In addition, we established in a previous report that in rice their molecular clock runs on average three times faster than that of the *Adh2* gene (Vitte *et al.*, 2004). Consequently, this phylogenetic incongruity could only be explained by the occurrence of several HTs between these three species.

It is known that the insertion of a transposable element can alter gene expression and be selectively advantageous, which leads to its fixation in the population where the mutation occurred. However, this conservation usually involves only some parts of the tranposable element. For example, in humans it has been shown that two genes, *Apolipoprotein-C1* (*apo-C1*) and *Endothelin B Receptor* (*EBR*) use the LTR of an endogenous retrovirus as alternative promoters (Medstrand *et al.*, 2001). While the insertions are 20–30 and 30–40 Myr old for apo-C1 and EBR respectively, the LTR of both loci are highly conserved, showing 88% sequence identity. In contrast, the *paternally expressed 10* (*PEG10*) gene is an LTR retrotransposon-derived gene, but even if *PEG10* is expressed and conserved in mammals no LTR has been found in this gene (Ono *et al.*, 2001).

These examples show that the domestication process usually involves only partial sequences of the transposable element. However, in the present study we show that, at least for *O. minuta* (BES data), the phylogenetic incongruity concerns subregions of *RIRE1* covering the whole element, including both the LTRs and the internal region. For these reasons, we consider that the most parsimonious explanation of our results is an exchange of genetic material between several *Oryza* species, after their radiation. Moreover, the high sequence conservation of a *Mu-like* transposable element between rice and foxtail millet (Diao *et al.*, 2006) suggests unequivocally that HT has occurred between the *Panicoideae* and *Bambusoideae* subfamilies of

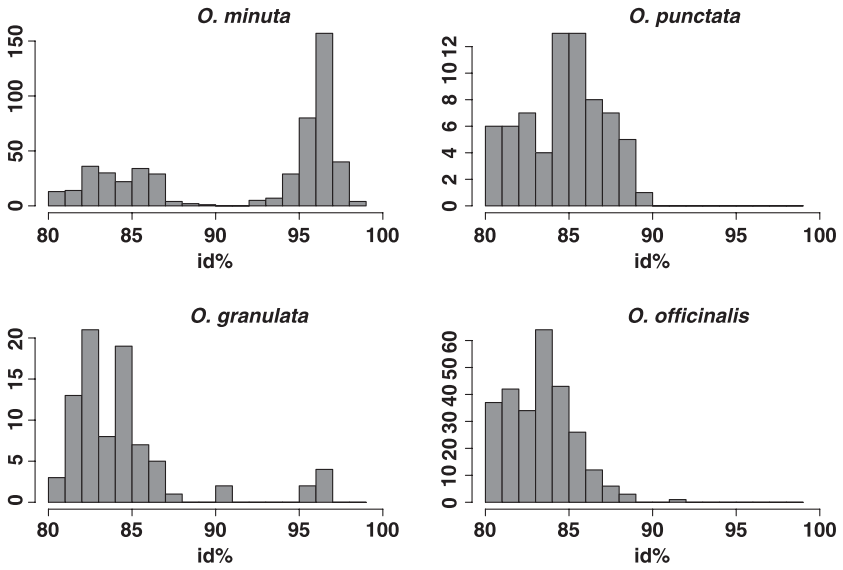
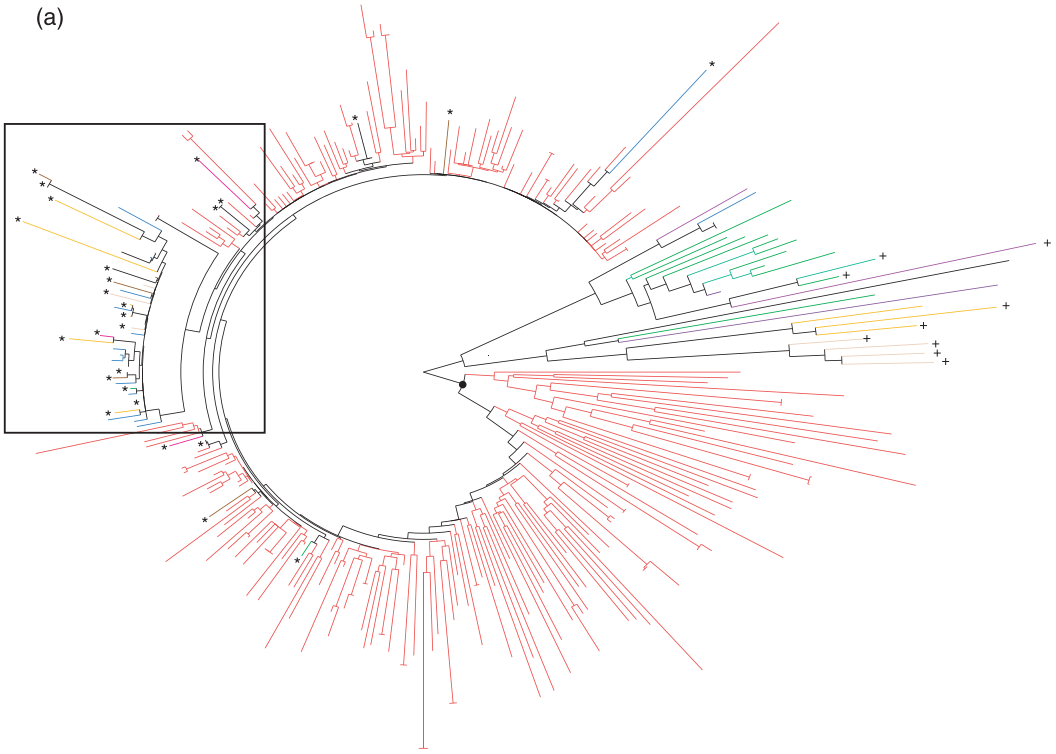
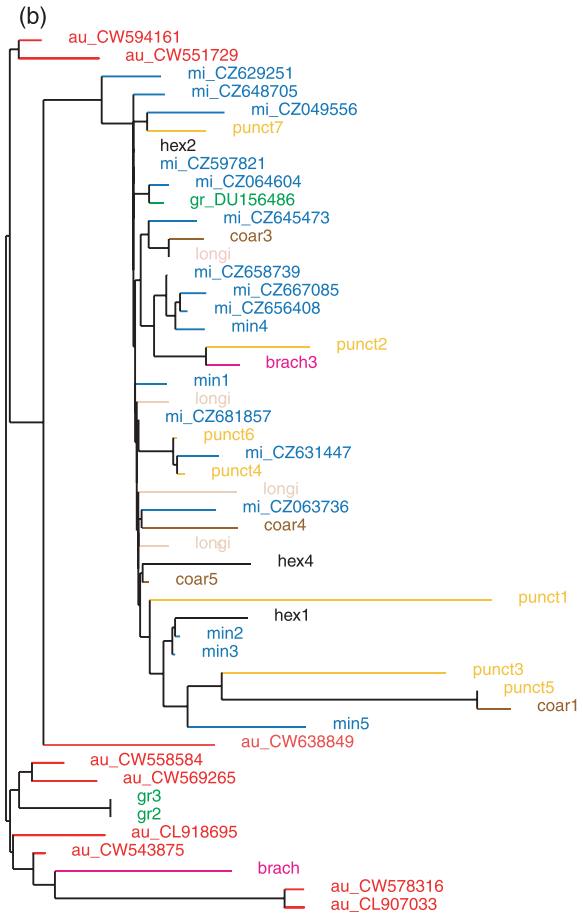


Figure 2. Distribution [number of bacterial artificial clone (BAC) end sequences (BES)] of pairwise identities between the reference sequence of *RIRE1* of *O. australiensis* (EE) and all sequences of four *Oryza* species matching with *RIRE1*.

Figure 3. Phenetic tree of *RIRE1* in the genus *Oryza*.
(a) Phenetic relationships of *RIRE1* long terminal repeat (LTR) regions in the genus *Oryza* calculated using bacterial artificial clone (BAC) end sequences (BES) and our sequences: the neighbor-joining tree was constructed based on the alignments given in Appendix S1. The dot shows the last significant bootstrap value. Color coding: pink for *O. sativa* and *O. longistaminata* (AA genome); red for *O. australiensis* (EE), blue for *O. minuta* (BBCC), green for *O. granulata* (GG), orange for *O. punctata* (BB), deep purple for *O. officinalis* (CC), brown for *O. coarctata* (HHKK), magenta for *O. brachyantha* (FF), grey for *O. alta* (CCDD), black for *L. hexandra.* *Sequences obtained with specific primers A5/A3. +, sequences obtained with non-specific primers A5/F3. The others sequences correspond to BES data.
(b) Subtree corresponding to the main putative transferred copies. Sequences followed by a number from 1 to 5 were obtained by cloning using specific primers A5/A3. Other sequences correspond to BES. Species abbreviation: au for *O. australiensis*; mi or min for *O. minuta*; gr for *O. granulata*; coar for *O. coarctata*; longi for *O. longistaminata*, punct for *O. punctata*; hex fo *L. hexandra*; brach for *O. brachyantha*.





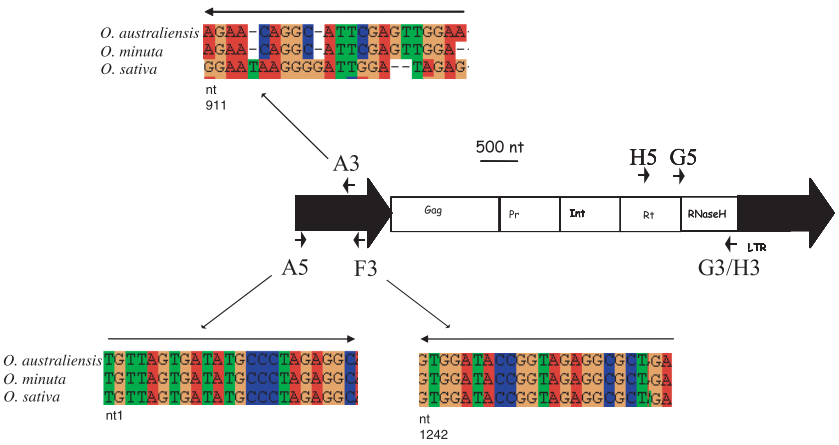


Figure 4. Schematic *RIRE1* and position of primers used for amplification.
Alignments show *RIRE1* sequence conservation or divergence between *O. australiensis* (EE), *O. minuta* (BBCC) and *O. sativa* (AA) where primers for the long terminal repeat (LTR) region were designed. Similar alignments were obtained for the internal region (not shown).

the *Poaceae* and therefore between more distantly isolated species. In our case, we chose to use the term 'horizontal transfer' to describe the transmission of *RIRE1* within the genus, mainly because it is well known that *Oryza* species with distinct genome types exhibit strong reproductive barriers (Brar and Khush, 1997), thus excluding the possibility that gene flows occurred through sexual reproduction.

Confirmation of the HT by molecular cloning and sequencing

The OMAP BES datasets are valuable resources for studying the evolutionary dynamics of a genome at a genus level. However, these BES represent only about 10% (from 7.5% to 11.5%) of the genome of a given species (Ammiraju *et al.*, 2006). In order to confirm our *in silico* data based on these sequences and to study the dynamics of *RIRE1* in the genus in more detail, we performed several PCR amplifications of *RIRE1* homologs in nine distinct *Oryza* species as well as in one species from the genus *Leersia*, which is closely related to *Oryza*. Amplifications were performed with primers which were either non-specific (A5/F3 and H5/H3) or specific (A5/A3 and G5/G3) to the putative transferred copy (see Experimental Procedures and Figure 4). The PCR amplifications were performed for both the LTR and the internal subregion (Figure 4). The species were chosen according to their genome types (AA, *O. sativa* and *Oryza longistaminata*; BB, *O. punctata*; CC, *Oryza officinalis*; BBCC, *O. minuta*; EE, *O. australiensis*; FF, *Oryza brachyantha*; GG, *O. granulata*; HHKK, *Oryza coarctata* and *Leersia hexandra* as an outgroup). The PCR amplification results for the LTR region are summarized in Table 2. Similar results were obtained for the internal region except for *O. granulata* (GG) where no sequence of the transferred copy was obtained. For each species, we obtained sequences of 1 kb for both the LTR and internal region that we aligned with the *O. australiensis* (EE) consensus *RIRE1* sequence. A total of 400 bp of these sequences was used to complete the phenetic trees obtained from the *in silico* data (Figure 3a,b, Appendices S1, S2 and S3).

All the 1 kb sequences generated with specific primers from our cloning experiments using both *O. minuta* (BBCC) and *O. granulata* (GG) show high identity with *RIRE1* of *O. australiensis* (EE; data not shown). Phenetic trees built both from BES and from our sequences clearly show that they are clustered within the *O. australiensis RIRE1* group, as well as the sequences that we previously mined from the OMAP BES dataset. This observation confirms our results based on the *in silico* analyses and therefore rules out the possibility of contamination in the preparation of the BAC libraries. High sequence identity (up to 97%) is observed among these copies (Table 3). We obtained no amplification from genomic DNA of *O. sativa* (AA) and *O. officinalis* (CC) using the primers specific to the transferred copy. However, we were able to successfully amplify and subsequently

Table 2 Results of amplification and cloning with non-specific (A5/F3) and specific (A5/A3) primers for the LTR/internal regions

	sat	longi	punct	off	min	aus	brachy	coar	gr	hex
Non-specific	+/+	+/+	+/+	+/+	+/−	+/−	−/−	+/+	+/+	−/−
Specific	−/−	+/+	+/+	−/−	+/+	+/+	+/+	+/+	+/−	+/+

Table 3 Percentage sequence identity for RIRE1 clones obtained with A5/A3 (LTR region)

	aus	min	gr	punct	coar	brachy	longi	hex
aus	94.3							
min	96.7	98.6						
gr	97.6	97	98.9					
punct	96.2	97.6	97	97.9				
coar	96.6	97.4	97.1	97.1	96.9			
brachy	96.7	97	97.5	96.5	96.9	98.5		
longi	97.3	98.3	97.8	97.7	97.8	97.6	99	
hex	97.1	98	97.6	97.5	97.6	97.4	98.4	98.8

clone fragments (three to seven for the LTRs and two to four for the internal region) corresponding to the transferred copy of *RIRE1* from *O. punctata* (BB), *O. coarctata* (HHKK), *O. brachyantha* (FF), *O. longistaminata* (AA) and *L. hexandra*. As for *O. minuta* (BBCC) and *O. granulata* (GG), the 1 kb sequences obtained for each species show high sequence identity with the *O. australiensis RIRE1* reference copy. Similarly, all 400-bp subsequences were clustered within the *O. australiensis RIRE1* group (Figure 3a) with high sequence identity (up to 96%, Table 3). These results strongly suggest that the HT first identified in *O. minuta* (BBCC) and *O. granulata* (GG), based on the *in silico* studies, may have occurred between seven species of the genus and even with the more distant species, *L. hexandra.* Since we were also able to amplify and clone old copies of *RIRE1* with non-specific primers from the same DNA source in almost all species (sequence divergence in agreement with species radiation and the expected positions in the tree, Figure 3a and Appendices S1 and S2), we can strongly claim that the results obtained with specific primers are not due to DNA contamination.

As previously explained, our *in silico* analyses are based on partial data concerning small portions of the element (subregions of 400 bp, see Experimental Procedures). However, more than 1 kb of *RIRE1* was sequenced for both the LTR and the internal region in almost all the species, and in addition our results show clear phylogenetic incongruities for *O. minuta* (BBCC) BES throughout the sequence of the element (data not shown). Moreover, several distinguishable sequences of the putative transferred copy were obtained for each species. This strongly suggests that an active, and therefore complete, copy was horizontally transferred.

However, we were not able to clone any sequence corresponding to the internal region of the putatively transferred copy of *RIRE1* from *O. granulata* (GG; Appendix S2). It is known that deletion of LTR retroelements via unequal homologous recombination occurs concomitantly with retrotransposition and leads to the formation of solo LTR (Kalendar *et al.*, 2000; Vitte and Panaud, 2003). From our previous study, we showed that the extent of the formation of solo LTR varies among the families of LTR retrotransposons in rice (Vitte and Panaud, 2003). However, we do not have any evidence that the rate of formation of solo LTRs varies among the evolutionary lineages of a plant genus. The high activity of *RIRE1* in *O. granulata* (GG) still remains to be explained, because transposable elements usually need their internal region to transpose. However, it is known that some non-autonomous elements (absence of some or all coding domains), using autonomous element machinery for their transposition, are still active. In the rice genome, the non-autonomous transposable element *Dasheng*, using putatively *RIRE2* machinery, has recently been amplified so that it is now one of the highest-copy-number LTRs in the rice genome (Jiang *et al.*, 2002a,b). Whether the absence of a genomic sequence homologous to the internal region of *RIRE1* in *O. granulata* (GG) could be explained simply by a higher rate of ectopic recombination between the LTRs of *RIRE1* still remains to be proven. This could be achieved by the analysis of a larger sample of genomic sequences.

Possible origin and direction of the horizontal transfer

The fact that the cluster of all horizontally transferred *RIRE1* copies is included in the much larger cluster of the *RIRE1* copies that originated from the recent transpositional burst of this element in *O. australiensis* (EE) suggests that this could be the donor species and that the transfer is after the burst of *RIRE1* in *O. australiensis*. Since this burst occurred within the last 750 000 years (Piégu *et al.*, 2006) we propose that the HT between *O. australiensis* (EE) and the seven species [*O. longistaminata* (AA), *O. punctata* (BB), *O. minuta* (BBCC), *O. brachyantha* (FF), *O. granulata* (GG), *O. coarctata* (HHKK) and *L. hexandra*] may have occurred during the same period, which is in agreement with our estimation based on the divergence between the transferred copies (i.e. between 700 000 and 600 000 years ago). Our data, however, do not allow us to make a conclusion about the exact scenario of these transfers, i.e. whether all transfers occurred concomitantly from *O. australiensis* (EE) to all seven species or whether one occurred first between *O. australiensis* and one of the seven species, followed by transfers to the other species.

Both *in silico* analyses and cloning experiments concern only one accession of all the species studied (cf. Experimental Procedures). We surveyed a collection of 26 *O. minuta* (BBCC) accessions to test the dynamics of *RIRE1* at the species level by Southern blot hybridization. Our estimation of the copy number of the horizontally transferred copy of *RIRE1* in the genome of *O. minuta* (BBCC) is 3000, based on *in silico* analysis (see Experimental Procedures), while there are about 1000 'older' copies (probably originating from an ancient burst) inherited vertically from a common ancestor of the *Oryza* genus. We observed that the horizontally transferred copies have a higher copy number than the old ones, therefore we anticipated that if some accessions of *O. minuta* (BBCC) were not subject to the transfer, then a difference in the hybridization signal using a probe corresponding to *RIRE1* of *O. australiensis* (EE) would be observed. Figure 5 clearly shows that no polymorphism was observed among accessions, which suggests that the transferred copy has now invaded the genome of the plants from all the populations surveyed. These results are further supported by PCR amplification data obtained for all accessions with the primers specific to the transferred copy (data not shown). This could therefore represent the second reported case of such a massive invasion, the first being the well-documented invasion of the *Drosophila melano-*

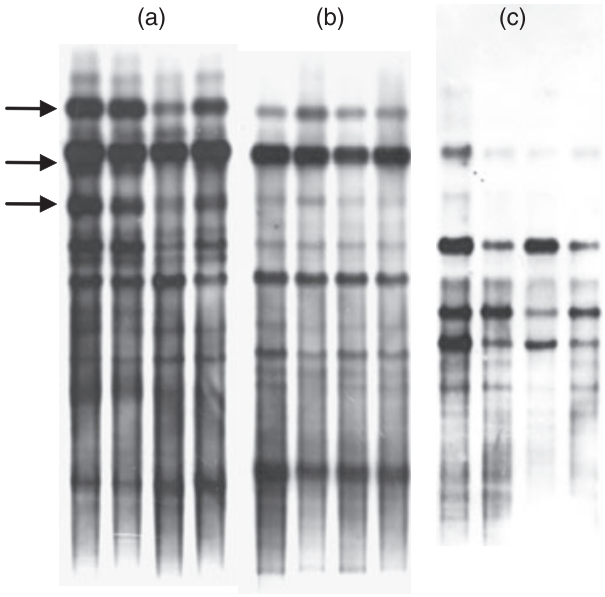


Figure 5. Southern hybridization of the *RIRE1* retrotransposons on total genomic DNA digested with *Rsa*I.
(a) *Oryza australiensis* (EE), 2 min exposure; (b) *O. minuta* (BBCC), 10 min exposure; (c) *O. granulata* (GG), 20 min exposure. Arrows indicate the bands shared by the three species. The blot only shows four of the 26 accessions of *O. minuta*

gaster genome by the *P element*, after its HT from the distantly related species *Drosophila willistoni* (Daniels *et al.*, 1990). The same results were obtained with the accessions of *O. granulata* (GG, Figure 5) except that the number of accessions available for our survey was relatively low for us to consider comparable conclusion in this species.

Conclusions

The results presented in this paper clearly show that several HTs of the LTR retrotransposon *RIRE1* have occurred between eight *Oryza* species. Figure 6 presents the distribution of all species in this genus and shows that most of the species for which we provide evidence for HT are not sympatric today. However, one can expect that the distribution of *Oryza* species has undergone many changes over the past million years, especially during the last 100 000 years, mainly because of paleoclimatic variations (i.e. the glaciation cycles). For example, the tetraploid BBCC genome species *O. minuta* is endemic to the Philippines, while one of its progenitor species *O. punctata* (BB) is endemic to Africa. We have tentatively dated the HTs before 700 000 years BC, i.e. long before the beginning of the last glaciations. Vaughan *et al.* (2005) suggested that a region in Southeast Asia, centered on Papua New Guinea and including eastern Indonesia, may be one of the centres of diversification of the *Oryza* genus, implying that most present-day *Oryza* species may have been sympatric in this region in the past.

The actual mechanisms of HT remain largely unknown in plants. Diao *et al.* (2006) proposed that because of their capacity to integrate into a host genome, transposable elements are particularly prone to HT. Bacteria, fungi or sap-sucking insects are often believed to be the vectors of HT (Won and Renner, 2003). Gene transfer from a pathogen to its host has been well described and, for example, a fragment of 11 kb has been transferred from the endosymbiotic bacterium *Wolbachia* to the X chromosome of its host insect *Callosobrochus chinensis* (Kondo *et al.*, 2002). Even if this hypothesis is purely speculative, it remains attractive to explain the transfer of *RIRE1* between wild rice species. Further experiments are needed in order to validate it. In any case, one first needs to demonstrate that pathogens or insects may infect several *Oryza* species in sympatric situations. However, gene capture through hybridization and introgression (i.e. transfer of genes from one genome to the other in the interspecific hybrid followed by restoration of one of the parental genome by successive backcrosses) has been well described in different genera (see Rieseberg and Soltis, 1991 for review). The existence of several allopolyploid species in the genus *Oryza* shows that interspecific hybridizations occurred in the past. We thus cannot

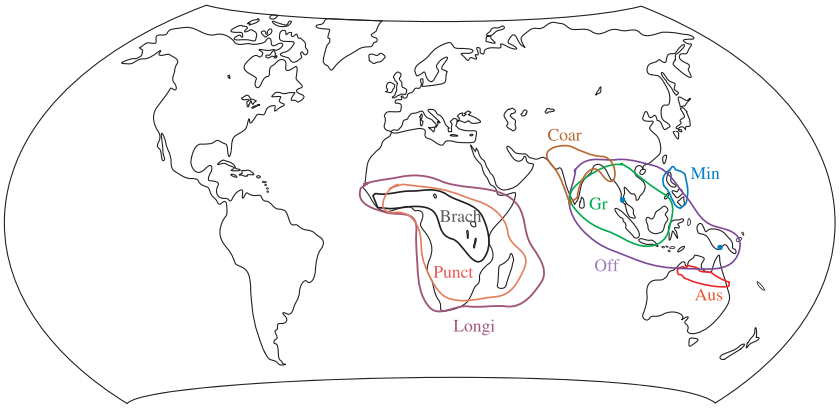


Figure 6. Present-day distribution of wild *Oryza* species (adapted from Vaughan and Morishima, 2003).
Species abbreviations: Aus for *O. australiensis* (EE); Min for *O. minuta* (BBCC); Gr for *O. granulata* (GG); Coar for *O. coarctata* (HHKK); Longi for *O. longistaminata* (AA), Punct for *O. punctata* (BB); Brach for *O. brachyantha* (FF). The worldwide distributed species *Leersia hexandra* is not represented on the map.

exclude the hypothesis that hybridizations and introgressions within the genus would have been responsible for the transfer of *RIRE1* between the different species studied. However, because *Oryza* species exhibit strong reproductive barriers, the HT hypothesis seems to be more parsimonious.

After their HT, transposable elements still have a high probability of elimination from the genome by genetic drift if no burst of transposition occurs (Le Rouzic and Capy, 2005). BES distribution in *O. minuta* (BBCC) and *O. granulata* (GG) and Southern blot hybridization show that in both species, *RIRE1* was still active after the transfer (Figures 2 and 5). We estimate that about 3000 and 300 copies are now present in the genomes of *O. minuta* (BBCC) and *O. granulata* (GG) respectively, which could explain why we were able to detect the HT by *in silico* analysis. We conclude that the transferred copy was still active after horizontal transfer in *O. minuta* (BBCC) and *O. granulata* (GG). It also suggests that *RIRE1* is particularly prone to massive amplification and that, as in *O. australiensis* (EE), it has undergone a recent burst of retrotransposition in *O. minuta* (BBCC) and *O. granulata* (GG). Since we were able to clone several distinguishable sequences for each species (see intraspecies identity in the diagonal of Table 3), we suspect that the transferred copy was also still active in the other species.

In this study, we have shown that *RIRE1* has been horizontally transferred in many species of the genus *Oryza* during a short evolutionary time period. If similar observations are made for other LTR retrotransposon families then the process of horizontal transmission of mobile elements between distinct species within the same genus may play an important role in plant genome evolution, particularly if it involves active elements that can be at the origin of genomic expansions through bursts of transposition, as is the case in *O. australiensis* (EE) and *O. minuta* (BBCC) for *RIRE1*.

Experimental procedures

In silico *approach and phenetic analyses*

The complete sequence of *RIRE1* (Noma *et al.*, 1997) was split into subsequences of 400 bp, in order to fit with the average size of the OMAP BES (700 bp). Each subsequence was used as a query for a BLASTN search against all the OMAP BES released in GenBank. Only hits showing a similarity of over at least 90% of the total length of the query were kept. All subsequences corresponding to the match were aligned using CLUSTALX (Thompson *et al.*, 1997) and the alignments modified by hand using the SEAVIEW software (Galtier *et al.*, 1996). Final alignments were used to construct a neighbor-joining dendrogram with the CLUSTALX software, using the observed divergence distance and performing 100 bootstrap replicates. A circular classification tree was drawn only for one subregion in the LTR and one in the internal region using the Treedyn package (http://www.treedyn.org/). The observed divergence was translated into an insertion date based on a substitution rate of 2×10^{-8} substitutions site^{-1} year^{-1} (Vitte *et al.*, 2004). For each species, we calculated the distribution of BES according to their identity with the *O. australiensis RIRE1* reference sequence (Noma *et al.*, 1997, GenBank accession no. D85597).

Copy number estimation

The copy number of the element in the genome was estimated by calculating the fraction of BES in the corresponding database completely matching the element (lengthwise) and dividing this fraction by the estimated coverage of the genome by the BES (Ammiraju *et al.*, 2006). This fraction was then multiplied by the estimated genome size (Ammiraju *et al.*, 2006).

Molecular approach and phenetic analysis

Total DNA was extracted from 10 species [accession 105 690 for *O. punctata* (BB), 101 116 for *O. officinalis* (CC), 105 089 for *O. minuta* (BBCC), 100 882 for *O. australiensis* (EE), 102 118 for *O. granulata* (GG), 104 502 for *O. coarctata* (HHKK), 110 404 for *O. longistaminata* (AA), 101 232 for *O. brachyantha* (FF), Nipponbare for *O. sativa* (AA) (no accession number available for *L. hexandra*)] provided by the International Rice Research Institute, Manila, Philippines. In order to draw phenetic trees, the *O. australiensis RIRE1* sequence was aligned with those of *O. sativa* (AA) and with BES from *O. minuta* (BBCC) identified as phylogenetically incongruent. Sequences conserved between species were used to design primers which enabled us to amplify old copies of *RIRE1* (F3: AGCGCCTCTACCGGTATCCAC and H5: ATTTCGATTATGAGATATGGC/H3: CCACTTGCACTAGAG-TCAATCA). Sequences conserved between *O. minuta* and *O. australiensis* but not with *O. sativa* were used to design primers specific to the transferred copy (A5: TGTTAGTGATATGCCCTAGAGGC/A3: CCAACTCGAATGCCTGTTCT and G5: GTGGGAGCGCAATTGTATTT/G3: CAACATCTCCATGTTGCCTCT). Positions of primers and alignments (for the LTR region) of sequences where primers were designed are given in Figure 4.

The PCR products were cloned in the pGEM-T Easy vector (Promega, http://www.promega.com/) and sequenced. For each clone, sequences were obtained both from the 3′ and 5′ ends to give contigs of around 1 kb. For each species, three to seven (for the LTR) and two to four (for internal region) clones were sequenced. These sequences were then aligned with the *RIRE1* consensus sequence and with BES from the corresponding region from all rice species available in the OMAP BES library using CLUSTALX. Because BES length is on average 400 bp, we only kept part of our sequences for this alignment. The same procedure as for the *in silico* approach was used to modify alignments and draw circular classifications.

Matrices of nucleotide distances were calculated using all sequences obtained through the cloning of PCR products obtained with specific primers A5/A3. The sequence CW638849 was used as a reference sequence for *O. australiensis* (EE). The observed divergence was translated into an insertion date using a substitution rate of 2×10^{-8} substitutions site^{-1} year^{-1} (Vitte *et al.*, 2004).

Phylogenetic analyses

The sequence of the *Adh2* gene was retrieved from GenBank for seven diploid and one tetraploid *Oryza* species. We used the GenBank accession no. AF148623 for *O. australiensis* (EE), AF148632 for *O. brachyantha* (FF), AF148613 for *O. officinalis* (CC), AF148611 for *O. punctata* (BB), AF148602 for *O. sativa* (AA), AF148628 and AF148627 for *O. coarctata* (HHKK), AF148607 for *O. longistaminata* (AA), AF148616 and AF148612 for *O. minuta* (BBCC), AY792580 for *L. hexandra* and AF148631 for *O. granulata* (GG). For each accession, the coding sequence (CDS) was extracted and aligned using

CLUSTALX. A pairwise distance matrix was computed using the Nei and Gojobori method (Nei and Gojobori, 1986). Only synonymous substitutions were then taken into account. A neighbor-joining tree was built using 100 bootstrap replicates. We estimated the date of the nodes using a rate of 6.5×10^{-9} synonymous substitutions site^{-1} year^{-1}.

Southern blot analysis

Southern blots were prepared using 2 μg of total genomic DNA digested with *Rsa*I and transferred to Hybond-N+ membranes. A non-radioactive procedure for probe labelling and signal detection was used (Panaud *et al.*, 1993). The probes were prepared by cloning the PCR product from *RIRE1* obtained with primers A5/A3 (LTR region) on total *O. australiensis* (EE) genomic DNA. This procedure was performed on 26 accessions of *O. minuta* (BBCC) 9only accessions 101 096, 101 097, 82 048 and 101 141 were used for Figure 5), four accesssions of *O. granulata* (GG; 102 118, 8923, 100 879 and 104 563) and four accessions of *O. australiensis* (EE; 86 526, 86 527, 86 528, and 86 529). The DNA used in this study was provided by the International Rice Research Institute, Manila, Philippines.

Acknowledgements

This work was supported by CNRS funding and NSF OMAP grant. We thank R. Cooke, M. A. Grandbastien, A. Zuccolo, J. S. S. Ammiraju and J. L. Goicoechea for their useful comments on the manuscript.

Supplementary Material

The following supplementary material is available for this article online:

Appendix S1. Sequences alignment file of the tree for the LTR region.

Appendix S2. Sequence alignment file of the tree for the internal region.

Appendix S3. Phenetic relationships of *RIRE1* internal regions in the genus *Oryza* drawn using BES and our sequences: the neighbor-joining tree was constructed based on the alignments given in Appendix S2.

This material is available as part of the online article from http://www.blackwell-synergy.com.

Please note: Blackwell publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

References

Ammiraju, J.S.S., Luo, M., Goicoechea, J.L. *et al.* (2006) The oryza bacterial artificial chromosome library resource: construction and analysis of 12 deep-coverage large-insert bac libraries that represent the 10 genome types of the genus oryza. *Genome Res.* **16**, 140–147.

Ammiraju, J.S.S., Zuccolo, A., Yu, Y. *et al.* (2007) Evolutionary dynamics of an ancient retrotransposon family provides insights into evolution of genome size in the genus Oryza. *Plant J.* **52**, 342–351.

Bennetzen, J.L. and Kellogg, E.A. (1997) Do plants have a one-way ticket to genomic obesity? *Plant Cell*, **9**, 1509–1514.

Brar, D.S. and Khush, G.S. (1997) Alien introgression in rice. *Plant Mol. Biol.* **35**, 35–47.

Brown, J.R. (2003) Ancient horizontal gene transfer. *Nat. Rev. Genet.* **4**, 121–132.

Casse, N., Bui, Q.T., Nicolas, V., Renault, S., Bigot, Y. and Laulier, M. (2006) Species sympatry and horizontal transfers of mariner transposons in marine crustacean genomes. *Mol. Phylogenet. Evol.* **40**, 609–619.

Daniels, S.B., Peterson, K.R., Strausbaugh, L.D., Kidwell, M.G. and Chovnick, A. (1990) Evidence for horizontal transmission of the p transposable element between drosophila species. *Genetics*, **124**, 339–355.

Diao, X., Freeling, M. and Lisch, D. (2006) Horizontal transfer of a plant transposon. *PLoS Biol.* **4**, e5.

Galtier, N., Gouy, M. and Gautier, C. (1996) Seaview and phylo_win: two graphic tools for sequence alignment and molecular phylogeny. *Comput. Appl. Biosci.* **12**, 543–548.

Gogarten, J.P. and Townsend, J.P. (2005) Horizontal gene transfer, genome innovation and evolution. *Nat. Rev. Microbiol.* **3**, 679–687.

Herédia, F., Loreto, E.L.S. and Valente, V.L.S. (2004) Complex evolution of gypsy in drosophilid species. *Mol. Biol. Evol.* **21**, 1831–1842.

International Human Genome Consortium (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860–921.

Jiang, N., Bao, Z., Temnykh, S., Cheng, Z., Jiang, J., Wing, R.A., McCouch, S.R. and Wessler, S.R. (2002a) Dasheng: a recently amplified nonautonomous long terminal repeat element that is a major component of pericentromeric regions in rice. *Genetics*, **161**, 1293–1305.

Jiang, N., Jordan, I.K. and Wessler, S.R. (2002b) Dasheng and RIRE2. A nonautonomous long terminal repeat element and its putative autonomous partner in the rice genome. *Plant Physiol.* **130**, 1697–1705.

Jordan, I.K., Matyunina, L.V. and McDonald, J. (1999) Evidence for the recent horizontal transfer of long terminal repeat retrotransposon. *Proc. Natl Acad. Sci. USA*, **96**, 12621–12625.

Kalendar, R., Tanskanen, J., Immonen, S., Nevo, E. and Schulman, A.H. (2000) Genome evolution of wild barley (*hordeum spontaneum*) by bare-1 retrotransposon dynamics in response to sharp microclimatic divergence. *Proc. Natl Acad. Sci. USA*, **97**, 6603–6607.

Khush, G.S. (1997) Origin, dispersal, cultivation and variation of rice. *Plant Mol. Biol.* **35**, 25–34.

Kondo, N., Nikoh, N., Ijichi, N., Shimada, M. and Fukatsu, T. (2002) Genome fragment of wolbachia endosymbiont transferred to x chromosome of host insect. *Proc. Natl Acad. Sci. USA*, **99**, 14280–14285.

Koonin, E.V., Makarova, K.S. and Aravind, L. (2001) Horizontal gene transfer in prokaryotes: quantification and classification. *Annu. Rev. Microbiol.* **55**, 709–742.

Laten, H.M., Havecker, E.R., Farmer, L.M. and Voytas, D.F. (2003) Sire1, an endogenous retrovirus family from glycine max, is highly homogeneous and evolutionarily young. *Mol. Biol. Evol.* **20**, 1222–1230.

Lawrence, J.G. (2002) Gene transfer in bacteria: speciation without species? *Theor. Popul. Biol.* **61**, 449–460.

Le Rouzic, A. and Capy, P. (2005) The first steps of transposable elements invasion: parasitic strategy vs. genetic drift. *Genetics*, **169**, 1033–1043.

Ma, J., Devos, K.M. and Bennetzen, J.L. (2004) Analyses of ltr-retrotransposon structures reveal recent and rapid genomic DNA loss in rice. *Genome Res.* **14**, 860–869.

Medstrand, P., Landry, J.R. and Mager, D.L. (2001) Long terminal repeats are used as alternative promoters for the endothelin b

receptor and apolipoprotein c–i genes in humans. *J. Biol. Chem.* **276**, 1896–1903.

Morinaga, T., Tsunoda, S. and Takahashi, N. (1964) *Rice Genetics and Cytogenetics*. Amsterdam: Elsevier.

Mower, J.P., Stefanović, S., Young, G.J. and Palmer, J.D. (2004) Plant genetics: gene transfer from parasitic to host plants. *Nature*, **432**, 165–166.

Nei, M. and Gojobori, T. (1986) Simple methods for estimating the numbers of synonymous and non synonymous nucleotide substitutions. *Mol. Biol. Evol.* **3**, 418–426.

Noma, K., Nakajima, R., Ohtsubo, H. and Ohtsubo, E. (1997) *RIRE1*, a retrotransposon from wild rice O*ryza australiensis*. *Genes Genet. Syst.* **72**, 131–140.

Ono, R., Kobayashi, S., Wagatsuma, H., Aisaka, K., Kohda, T., Kaneko-Ishino, T. and Ishino, F. (2001) A retrotransposon-derived gene, peg10, is a novel imprinted gene located on human chromosome 7q21. *Genomics*, **73**, 232–237.

Panaud, O., Magpantay, G. and McCouch, S.R. (1993) A protocole for nonradioactive DNA labelling and detectionin the RFLP analysis of rice and tomato using single copy probes. *Plant Mol. Biol. Rep.* **11**(1), 54–59.

Piégu, B., Guyot, R., Picault, N. *et al.* (2006) Doubling genome size without polyploidization: dynamics of retrotransposition-driven genomic expansions in oryza australiensis, a wild relative of rice. *Genome Res.* **16**, 1262–1269.

Richardson, A.O. and Palmer, J.D. (2007) Horizontal gene transfer in plants. *J. Exp. Bot.* **58**, 1–9.

Rieseberg, L.H. and Soltis, D.E. (1991) Phylogenetic consequences of cytoplasmic gene flow in plants. *Evol. Trends plant.* **5**(1), 65–84.

Salzberg, S.L., White, O., Peterson, J. and Eisen, J.A. (2001) Microbial genes in the human genome: lateral transfer or gene loss? *Science*, **292**, 1903–1906.

SanMiguel, P., Tikhonov, A., Jin, Y.K. *et al.* (1996) Nested retrotransposons in the intergenic regions of the maize genome. *Science*, **274**, 765–768.

Scholl, E.H., Thorne, J.L., McCarter, J.P. and Bird, D.M. (2003) Horizontally transferred genes in plant-parasitic nematodes: a high-throughput genomic approach. *Genome Biol.* **4**, R39.

Shen, Z., Denton, M., Mutti, N., Pappan, K., Kanost, M.R., Reese, J.C. and Reek, G.R. (2003) Polygalacturonase from *Sitophilus oryzae*: possible horizontal transfer of a pectinase gene from fungi to weevils. *J. Insect Sci.* **3**, 24.

Syvanen, M. (1994) Horizontal gene transfer: evidence and possible consequences. *Annu. Rev. Genet.* **28**, 237–261.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) The clustal_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876–4882.

Vaughan, D.A. and Morishima, H. (2003) Biosystematics of the genus *Oryza*. In: Wayne Smith, C., eds. *Rice Origin, History, Technology and Production*. New York: Wiley and sons, Inc. pp. 27–65.

Vaughan, D.A., Kadowaki, K., Kaga, A. and Tomooka, N. (2005) On the phylogeny and biogeography of the genus *Oryza*. *Breeding Sci.* **55**, 113–122.

Vicient, C.M., Kalendar, R. and Schulman, A.H. (2001) Envelope-class retrovirus-like elements are widespread, transcribed and spliced, and insertionally polymorphic in plants. *Genome Res.* **11**, 2041–2049.

Vitte, C. and Panaud, O. (2003) Formation of solo-LTRs through unequal homologous recombination counterbalances amplifications of LTR retrotransposons in rice *Oryza sativa* L. *Mol. Biol. Evol.* **20**(4), 528–540.

Vitte, C. and Panaud, O. (2005) Ltr retrotransposons and flowering plant genome size: emergence of the increase/decrease model. *Cytogenet. Genome Res.* **110**, 91–107.

Vitte, C., Ishii, T., Lamy, F., Brar, D.S. and Panaud, O. (2004) Genomic paleontology provides evidence for two distinct origins of asian rice (O*ryza sativa l.*). *Mol. Genet. Genomics*, **272**, 504–511.

Vitte, C., Panaud, O. and Quesneville, H. (2007) LTR retrotransposons in rice (*Oryza sativa,* L.): recent burst amplifications followed by rapid DNA loss. *BMC Genomics.* **8**, 218.

Won, H. and Renner, S.S. (2003) Horizontal gene transfer from flowering plants to gnetum. *Proc. Natl Acad. Sci. USA*, **100**, 10824–10829.

Zuccolo, A., Sebastian, A., Talag, J., Yu, Y., Kim, H., Collura, K., Kudrna, D. and Wing, R.A. (2007) Transposable element distribution, abundance and role in genome size variation in the genus Oryza. *BMC Evol. Biol.* **7**, 152.

