



# Heterogeneous evolution of Ty3-gypsy retroelements among bamboo species

M.M. An, C. Guo, P.P. Lin and M.B. Zhou

The Nurturing Station for the State Key Laboratory of Subtropical Silviculture, Department of Silviculture, Zhejiang A&F University, Lin'an, Zhejiang, China

Corresponding author: M.B. Zhou  
E-mail: zhoutingbing@zafu.edu.cn

Genet. Mol. Res. 15 (3): gmr.15038515  
Received February 1, 2016  
Accepted March 28, 2016  
Published August 18, 2016  
DOI <http://dx.doi.org/10.4238/gmr.15038515>

Copyright © 2016 The Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution ShareAlike (CC BY-SA) 4.0 License.

**ABSTRACT.** Ty3-gypsy long-terminal repeat retroelements are ubiquitously found in many plant genomes. This study reports the occurrence of heterogeneous Ty3-gypsy retroelements in four representative bamboo species: *Phyllostachys heterocyclus* (Carr.) Mitford cv. *pubescens*, *P. heterocyclus* (Carr.) Mitford cv. *heterocyclus*, *Dendrocalamopsis oldhami*, and *Pleioblastus fortunei*. Using degenerate oligonucleotide primers corresponding to the conserved domains of reverse transcriptase (*rt*) genes of Ty3-gypsy retroelements, 165 distinct sequences were amplified from genomic DNA. The length of the nucleotide sequences varied from 366 to 438 bp. The sequences demonstrated a high heterogeneity, with homology ranging from 52.2 to 99.8%. A phylogenetic tree was constructed, including *Arabidopsis thaliana* and *Oryza sativa*. Bamboo Ty3-gypsy sequences formed three distinct retroelement clusters (gypsy I-III). Further analysis indicated that there were not only nearly identical Ty3-gypsy retroelements found in distantly related species, but also highly diverse Ty3-gypsy retroelements observed in closely related species. The results of this

study provide genetic and evolutionary information about the bamboo genome that could contribute to further studies of repetitive elements in bamboo as well as in other species.

**Key words:** Ty3-*gypsy*; Retroelement; Reverse transcriptase; Heterogeneity; Horizontal transfer

## INTRODUCTION

Long-terminal repeat (LTR) retroelements are mobile genetic elements that are ubiquitously present in eukaryotes, both in animals and plants (Bennetzen, 2000). LTR retroelements are subdivided based on their structure and transposition mechanism into Ty1- *copia*, Retroviridae, Bel, and the Ty3-*gypsy* group, which are widely distributed in the plant kingdom (Flavell et al., 1994; Suoniemi et al., 1998; Feschotte et al., 2002; Eickbush and Jamburuthugoda, 2008). LTR retroelements mainly encode two genes (*gag* and *pol*). The *gag* gene encodes structural proteins that form a virus-like particle while *pol* encodes the enzymatic regions such as aspartic protease, reverse transcriptase, RNase H, and integrase.

Retroelements are highly variable in their sequences, yet the reverse transcriptase (*rt*) region of Ty3-*gypsy* retroelements is highly conserved (Doolittle et al., 1989), owing to the importance of five highly conserved domains for enzymatic function. Homologous *rt* exist in many divergent species and have discontinuous distributions among populations (Llorens et al., 2009), which could be attributed to different rates of evolution, or vertical transmission followed by horizontal transfer and stochastic loss (Du et al., 2010; Domingues et al., 2012; Kolano et al., 2013).

The Bambusoideae (bamboo) is one of 12 sub-clades of Poaceae (grass family), which comprises 115 genera and 1439 species (Bamboo Phylogeny Group, 2012). Bamboo is subclassified based on its rhizome structure into monopodial, sympodial, and amphipodial bamboo types. Sympodial bamboos have short thick rhizomes that form clumps and are typically native to tropical or subtropical climates. Monopodial bamboos have long adventitious rhizomes that are cylindrical and segmented like the culms and typically grow in temperate climates. Amphipodial bamboos (with both sympodial and monopodial rhizomes), either occupy open habitats with spreading rhizomes or hold particular niches with clumping rhizomes in temperate or subtropical climates (Bamboo Phylogeny Group, 2012).

Previous studies have shown that several types of transposable elements are widely distributed in the bamboo genome, including *PIF*-like elements (Zhou et al., 2010a), *Mariner*-like elements (Zhou et al., 2010b), *Pong*-like elements (Zhong et al., 2010), and Ty1- *copia* retroelements (Zhou et al., 2010c). In addition, the distribution of some intragenic transposable elements in moso bamboo is correlated with transcript profiles (Zhou et al., 2011). Although Ty3-*gypsy* retroelements are one of the most commonly found plant transposable elements, their distribution and evolutionary pattern in bamboo genomes have not been investigated to date. In this study, Ty3-*gypsy* retroelement diversification was investigated in four representative bamboo species: *Phyllostachys heterocyclus* (Carr.) Mitford cv. *pubescens*, *P. heterocyclus* (Carr.) Mitford cv. *heterocyclus*, *Dendrocalamopsis oldhami*, and *Pleioblastus fortunei*, representing monopodium, sympodium, and amphipodium types. In detail, *P. heterocyclus*, (Carr.) Mitford cv. *pubescens* and *P. heterocyclus* (Carr.) Mitford cv. *heterocyclus*

are monopodium type, *D. oldhami* is sympodium type and *P. fortunei* is amphipodium type. Utilizing the availability of degenerate primers for the *rt*-conserved regions, we isolated and sequenced 165 Ty3-gypsy *rt* fragments. The distribution, heterogeneity, and evolutionary patterns of these *rt* sequences were investigated.

## MATERIAL AND METHODS

### Plant materials and genomic DNA isolation

Four representative species of bamboo were collected in this study, including *Phyllostachys heterocyclus* (Carr.) Mitford cv. *pubescens* and *P. heterocyclus* (Carr.) Mitford cv. *heterocyclus* (both representing the monopodial type), *D. oldhami* (a sympodial type), and *P. fortunei* (an amphipodial type) (Bamboo Phylogeny Group, 2012). The young leaves of each species were collected and genomic DNA was extracted using a DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany).

### Polymerase chain reaction (PCR) amplification, cloning, and sequencing

Degenerate oligonucleotide primers were designed based on the description by Kumar and Bennetzen (1999). Forward TY3-5 (5'-AGMGRATGTGYGTSGATYAT-3') and reverse TY3-3 (5'-GTKGGKYTTRWGTGTRAA-3') primers were synthesized by Sangon Biotech Co., Ltd., Shanghai, China. PCR amplifications were performed in 20- $\mu$ L reaction volumes containing 150 ng DNA, 2  $\mu$ L 10X PCR buffer, 1.2  $\mu$ L 2.0 mM MgCl<sub>2</sub>, 1.8  $\mu$ L 0.3 mM dNTPs, 0.5  $\mu$ L 10  $\mu$ M primers (TY3-F and TY3-R), 0.1  $\mu$ L Ex *Taq* enzyme, and 13.9  $\mu$ L sterile water. The reaction was preheated to 94°C for 3 min, followed by 35 cycles at 94°C for 30 s, 50°C for 30 s, 72°C for 90 s, and a final extension step at 72°C for 5 min.

Ty3-gypsy *rt* amplification products were resolved on a 1% agarose gel and purified using the UNIQ-10 Spin Column DNA Gel Extraction Kit (Sangon Biotech Co., Ltd.) and cloned into the pMD<sup>TM</sup> 18-T Vector (TaKaRa, Japan).

### Sequence analysis

BLASTn (<http://www.ncbi.nlm.nih.gov/>) was employed to test the authenticity and homology of the DNA sequences. The DNAMAN software (Lynnon Corporation, USA) was used to translate the nucleotide sequences into peptide sequences with reading-frame shifts when necessary. The heterogeneities of the nucleotide and peptide sequences were investigated using DNAMAN. The conserved domains of amino acid sequences were tested using the NCBI CCD analysis (<http://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>). Tajima tests of the nucleotide sequences and the nonsynonymous-synonymous substitution rate ratio (dN/dS) analysis were performed using the DnaSP software (Librado and Rozas, 2009).

In order to investigate evolutionary pattern of bamboo Ty3-gypsy retroelements, we also downloaded representative sequences of Ty3-gypsy retroelements of *Arabidopsis thaliana* and *Oryza sativa* from PGSB Repeat Element Database (<http://pgsb.helmholtz-muenchen.de/plant/recat/>). The sequences corresponding to the conserved domains of the Ty3-gypsy

retroelement *rt* gene were extracted. The multiple-sequence alignments, Model Test analysis, and NJ tree reconstruction were performed using MEGA 5.0 (Tamura et al., 2011).

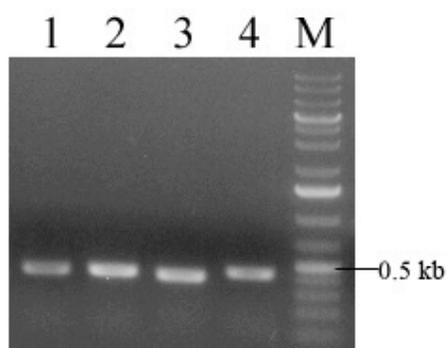
## RESULTS

### Identification and polymorphic analysis of the *rt* domain

More than 50 clones were randomly sequenced from each bamboo species until only redundant sequences were found. After removing repetitive sequences and filtering low-quality sequences, 165 sequences were obtained (KT715847-KT716011; Table 1). Each amplified fragment was 360-430 bp long (Figure 1). The BLASTn analyses showed that all the sequences had significant homology to known plant Ty3-*gypsy* retroelements. Each bamboo sequence was named according to the first letter of the genus and species, followed by the clone number, such as *P. heterocykla* (Carr.) Mitford cv. *pubescens*: Php-#.

**Table 1.** Species names, rhizome types, clone number and GenBank accession Nos.

Taxa	Abb.	Rhizome structure	Clone number	GenBank accession No.
<i>Phyllostachys heterocykla</i> (Carr.) Mitford cv. <i>heterocykla</i>	Ph	monopodium	40	KT715939-KT715978
<i>P. heterocykla</i> (Carr.) Mitford cv. <i>pubescens</i>	Php	monopodium	33	KT715979-KT716011
<i>Pleurocalamopsis fortunei</i>	Pf	amphipodium	41	KT715847-KT715887
<i>Dleioblastus oldhami</i>	Do	sympodium	51	KT715888-KT715938



**Figure 1.** PCR amplification of *rt* sequence from *Phyllostachys heterocykla* (Carr.) Mitford cv. *pubescens*, *Dendrocalamopsis oldhami*, *Pleioblastus fortunei*, and *P. heterocykla* (Carr.) Mitford cv. *heterocykla*. Lane 1 = *P. heterocykla* (Carr.) Mitford cv. *pubescens*, lane 2 = *D. oldhami*, lane 3 = *P. fortunei*, lane 4 = *P. heterocykla* (Carr.) Mitford cv. *heterocykla*, and lane M = 100-bp Plus ladder.

The Ty3-*gypsy* retroelement *rt* gene sequences were found to be highly polymorphic in the tested bamboo species. Pairwise comparisons of nucleotide sequences showed 52.2-99.8 and 51.1-100% identity at the amino acid level (Tables S1 and S2). Pairwise identities of Ty3-*gypsy* retroelement *rt* gene sequences within each species also revealed a spectrum of

diversity among the four species (53.6-100% identity at the amino acid level in *P. heterocyclus* (Carr.) Mitford cv. *heterocyclus*, 56.5-99.3% in *P. heterocyclus* (Carr.) Mitford cv. *pubescens*, 51.1-100% in *D. oldhami*, and 54.3-100% in *P. fortunei*) (Table S2).

### Tests for neutrality

To gain insight into the evolution of Ty3-gypsy retroelements in bamboo, Tajima's D neutral test was applied. The test including 165 Ty3-gypsy retroelement sequences resulted in a D-value of 2.13 ( $P < 0.001$ ; Table 2). The excess of intermediate frequency alleles of bamboo Ty3-gypsy retroelements may be indicative of past population bottlenecks, structure, and/or balancing selection (Biswas and Akey, 2006). The dN/dS for 75.36% of the codons were found to be greater than 1 ( $P < 0.001$ ), which suggests that the majority of the sites in the bamboo *rt* sequences are under positive selection (Miyata et al., 1979; Li et al., 1985, Yang et al., 2005).

**Table 2.** Tajima's neutrality test of 165 bamboo Ty3-gypsy retroelement sequences.

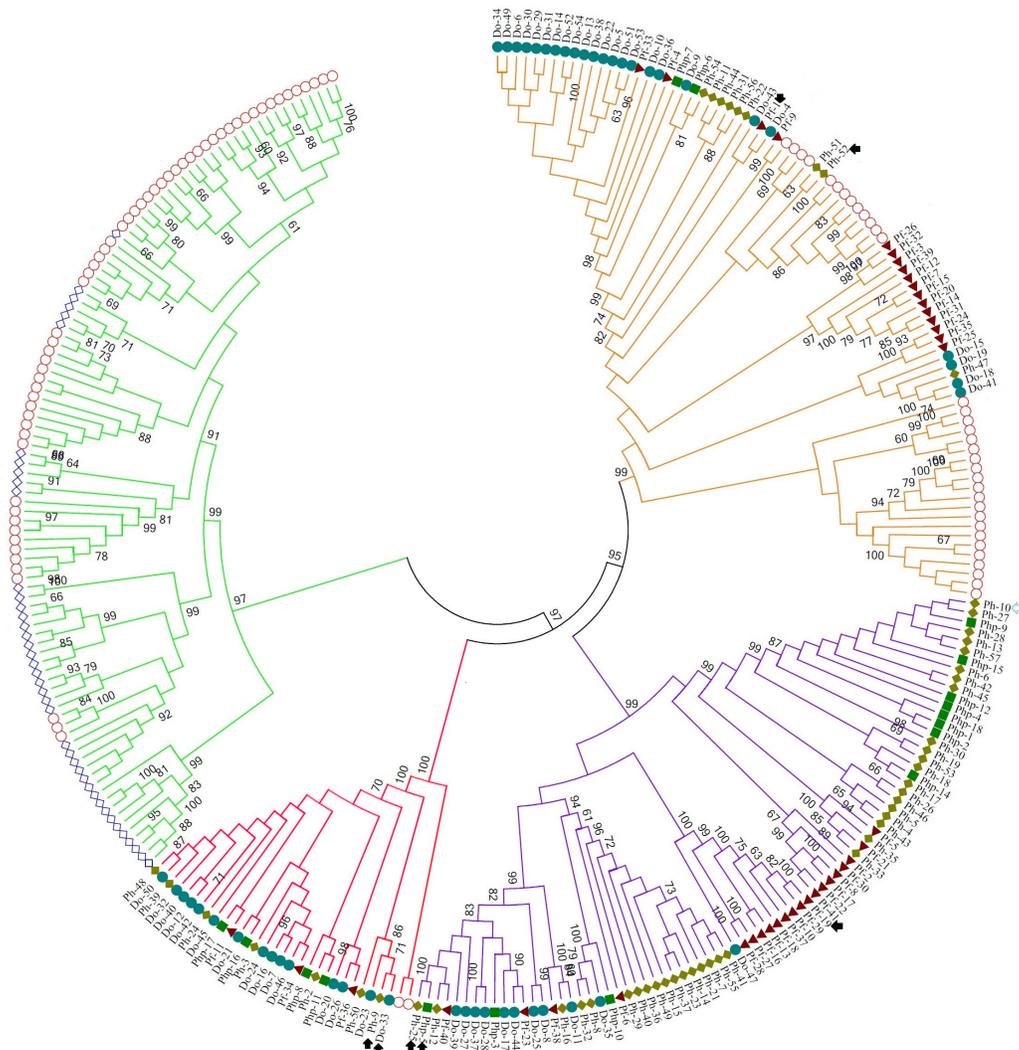
m	S	Ps	$\theta$	$\Pi$	D
165	262	0.821317	0.146895	0.243267	2.132570

m = number of sequences, S = number of segregating sites, Ps = S/m,  $\theta = Ps/a_1$ ,  $\Pi$  = nucleotide diversity, and D is the Tajima test statistic.

### Phylogenetic analysis of the *rt* domains of the bamboo Ty3-gypsy retroelements

Four clusters (gypsy I-IV) were the largest and best supported monophyletic groups, based on the NJ method (bootstrap values >60%; Figure 2). Most of the bamboo Ty3-gypsy retroelements clustered into three branches of the tree, representing gypsy I-III. Two of these clusters (gypsy I and gypsy III) contained both bamboo and rice Ty3-gypsy retroelements, gypsy II included exclusively bamboo retroelements, whereas gypsy IV contained both *A. thaliana* and *O. sativa* retroelements (Figure 2).

Among the bamboo Ty3-gypsy retroelements, distantly related bamboo species shared closely related Ty3-gypsy retroelements. For instance, the Ty3-gypsy retroelement family Do-33 (from *D. oldhami*) and Ph-9 (from *P. heterocyclus* (Carr.) Mitford cv. *heterocyclus*) showed 93.6% nucleotide identity, although these two species are spatially separated and represent distinctly different bamboo types (sympodial and monopodial, respectively). The reciprocal situation was also observed in which diverse Ty3-gypsy retroelements were present in the same species, e.g., in *P. fortunei* Pf-1 from cluster gypsy I and Pf-19 from cluster gypsy II, which shared only 64.6% identity. Similar phenomena were observed in the intra-species relationships of *P. heterocyclus*. Multiple divergent Ty3-gypsy retroelement fragments could be present in the same individual genome (e.g., Ph-52 and Ph-10 from *P. heterocyclus* (Carr.) Mitford cv. *heterocyclus*). Meanwhile, identical fragments of the Ty3-gypsy retroelement were shared by multiple cultivars (e.g., Ph-25 from *P. heterocyclus* (Carr.) Mitford cv. *heterocyclus* and Php-5 from *P. heterocyclus* (Carr.) Mitford cv. *pubescens*).



**Figure 2.** Phylogenetic tree of nucleotide sequences of 165 *rt* genes among four bamboo species plus 95 sequences in *Oryza sativa* and 41 sequences in *Arabidopsis thaliana* with the best-fit model of substitution of JTT+G and 1000 bootstrap replicates. Yellow branches represent gypsy I, purple branches represent gypsy II, red branches represent gypsy III, and green branches represent gypsy IV. The bamboo *rt* gene sequences are emphasized by capital letters, clone numbers, and different colored symbols, respectively. Squares: *Phyllostachys heterocyclus* (Carr.) Mitford cv. *pubescens*, circles: *Dendrocalamopsis oldhami*, triangles: *Pleioblastus fortunei*, and losenges: *P. heterocyclus* (Carr.) Mitford cv. *heterocyclus*. The *O. sativa* and *A. thaliana* *rt* genes are represented by the two open symbols, circles and losenges, respectively. The clones mentioned in the main text are emphasized by black arrows.

## DISCUSSION

Bamboo is an economically important plant cultivated in China. The available fossil evidence and the surviving basal lineages suggest that Bambusoideae evolved in

Gondwanaland during the Paleogene period more than 30 mya (Guo and Li, 2002; Ruiz-Sanchez, 2011). In the present study, using available degenerate oligonucleotide primers, 165 sequences representing *rt* domains of Ty3-*gypsy* retroelements were isolated from four bamboo species. The tested bamboo species covered monopodial, sympodial, and amphipodial types of bamboo, representing different rhizome structures. Ty3-*gypsy* retroelements are abundant in bamboo, and our phylogenetic analyses showed that bamboo Ty3-*gypsy* retroelements formed three clusters, all of which originated from a common ancestor (Figure 2). Among four clusters of the phylogenetic tree (Figure 2), one cluster (*gypsy* IV) was comprised of only *A. thaliana* and *O. sativa* retroelements. A possible reason for this is that the degenerate oligonucleotide primers used were too specific and did not successfully amplify bamboo retroelements that belong to the *gypsy* IV cluster. In contrast, one cluster (*gypsy* II) consisted exclusively of bamboo retroelements, which may suggest a bamboo-specific origin.

Ty3-*gypsy* group retroelements in the bamboo genomes are highly heterogeneous, like those observed in other plant species (Jiang et al., 2013; Kolano et al., 2013). There are four possible explanations for this. First, the retroelement “copy-and-paste” mechanism is error-prone, owing to the low fidelity and lack of proofreading activity in the *rt* gene, which leads to sequence diversity (Sun et al., 2008). Second, according to Kumar and Bennetzen (1999), retroelement heterogeneity could gradually be maintained over generations, through vertical transmission. Third, repetitive DNA could be deleted from the genome through unequal crossing-over and illegitimate recombination (Devos et al., 2002; Pereira, 2004), a process that might occur in one species but perhaps not in another (Jiang et al., 2013; Kolano et al., 2013). Fourth, transposable elements undergo more stringent natural selection than other sequences, due to their potential impact on genic regions. There are at least three deleterious effects of transposable elements in the host genome including mutations resulting from insertions into genes or regulatory sequences (Finnegan, 1992), chromosomal rearrangements caused by ectopic recombination between elements in non-homologous insertion sites (Devos et al., 2002; Pereira, 2004), and direct costs due to their transposition activity (Brookfield, 1991). Therefore, transposable elements undergo more stringent natural selection than other sequences, as the host could eliminate active transposable elements through stochastic loss or vertical extinction mechanisms to reduce the deleterious effects of transposition (Pritham, 2009; Jiang et al., 2013; Kolano et al., 2013). The positive value of Tajima’s *D* obtained for the 165 bamboo Ty3-*gypsy* sequences (2.13) indicates that these retroelements indeed underwent strong natural selection. The dN/dS values of each codon suggest that the majority of the sites (75.36%) are under positive selection.

Horizontal transmission can also cause sequence heterogeneity in retroelements (Kumar and Bennetzen, 1999). In our phylogenetic analysis, the bamboo Ty3-*gypsy* retroelements did not irrefutably show the presence of horizontal transfer, although some evidence suggesting this was observed. For example, the Ty3-*gypsy* retroelement phylogenetic tree was incongruent with the bamboo taxonomy reflected by the presence of near identical Ty3-*gypsy* retroelement sequences in distantly related species and the presence of relatively diverse Ty3-*gypsy* retroelement sequences in closely related species. This phenomenon indicates the possibility of horizontal transfer events between phylogenetically distant species during bamboo evolution (El Baidouri et al., 2014; Davis and Xi, 2015).

Artifacts in the phylogenetic reconstruction may occur as a result of differential rates of evolution among the retroelement clusters (Fortune et al., 2008). Notably, not all Ty3-*gypsy* retroelement sequences might be amplified in all four bamboo species due to primer

specificity. The resulting loss of Ty3-*gypsy* retroelement sequence information might have an impact on the subsequent phylogenetic analysis.

In conclusion, we show high heterogeneity, phylogenetic relationships, and evolutionary patterns for Ty3-*gypsy* retroelements in multiple bamboo species. The results of this study provide genetic and evolutionary information concerning the bamboo genome. This could contribute to future studies of repetitive elements in bamboo as well as other species.

### Conflicts of interest

The authors declare no conflict of interest.

### ACKNOWLEDGMENTS

Research supported by the Program of Natural Science Foundation of Zhejiang Province (grant #LR12C16001) and the National Natural Science Foundation of China (grant #31470615 and #31270645). We thank LetPub ([www.letpub.com](http://www.letpub.com)) for its linguistic assistance during the preparation of this manuscript.

### REFERENCES

- Bamboo Phylogeny Group (2012). An updated tribal and subtribal classification of the bamboos (Poaceae: Bambusoideae). *J. Am. Bamboo Soc.* 24: 1-10.
- Bennetzen JL (2000). Transposable element contributions to plant gene and genome evolution. *Plant Mol. Biol.* 42: 251-269. <http://dx.doi.org/10.1023/A:1006344508454>
- Biswas S and Akey JM (2006). Genomic insights into positive selection. *Trends Genet.* 22: 437-446. <http://dx.doi.org/10.1016/j.tig.2006.06.005>
- Brookfield JF (1991). Models of repression of transposition in P-M hybrid dysgenesis by P cytotype and by zygotically encoded repressor proteins. *Genetics* 128: 471-486.
- Davis CC and Xi Z (2015). Horizontal gene transfer in parasitic plants. *Curr. Opin. Plant Biol.* 26: 14-19. <http://dx.doi.org/10.1016/j.pbi.2015.05.008>
- Devos KM, Brown JK and Bennetzen JL (2002). Genome size reduction through illegitimate recombination counteracts genome expansion in *Arabidopsis*. *Genome Res.* 12: 1075-1079. <http://dx.doi.org/10.1101/gr.132102>
- Domingues DS, Cruz GM, Metcalfe CJ, Nogueira FT, et al. (2012). Analysis of plant LTR-retrotransposons at the fine-scale family level reveals individual molecular patterns. *BMC Genomics* 13: 137. <http://dx.doi.org/10.1186/1471-2164-13-137>
- Doolittle RF, Feng DF, Johnson MS and McClure MA (1989). Origins and evolutionary relationships of retroviruses. *Q. Rev. Biol.* 64: 1-30. <http://dx.doi.org/10.1086/416128>
- Du J, Tian Z, Hans CS, Laten HM, et al. (2010). Evolutionary conservation, diversity and specificity of LTR-retrotransposons in flowering plants: insights from genome-wide analysis and multi-specific comparison. *Plant J.* 63: 584-598. <http://dx.doi.org/10.1111/j.1365-3113X.2010.04263.x>
- El Baidouri M, Carpentier MC, Cooke R, Gao D, et al. (2014). Widespread and frequent horizontal transfers of transposable elements in plants. *Genome Res.* 24: 831-838. <http://dx.doi.org/10.1101/gr.164400.113>
- Eickbush TH and Jamburuthugoda VK (2008). The diversity of retrotransposons and the properties of their reverse transcriptases. *Virus Res.* 134: 221-234. <http://dx.doi.org/10.1016/j.virusres.2007.12.010>
- Feschotte C, Jiang N and Wessler SR (2002). Plant transposable elements: where genetics meets genomics. *Nat. Rev. Genet.* 3: 329-341. <http://dx.doi.org/10.1038/nrg793>
- Finnegan DJ (1992). Transposable elements. In: *The genome of Drosophila melanogaster* (Lindsley DL and Zimm GG, eds.). Academic Press, New York, 1096-1107.
- Flavell AJ, Pearce SR and Kumar A (1994). Plant transposable elements and the genome. *Curr. Opin. Genet. Dev.* 4: 838-844. [http://dx.doi.org/10.1016/0959-437X\(94\)90068-X](http://dx.doi.org/10.1016/0959-437X(94)90068-X)
- Fortune PM, Roulin A and Panaud O (2008). Horizontal transfer of transposable elements in plants. *Commun. Integr. Biol.* 1: 74-77. <http://dx.doi.org/10.4161/cib.1.1.6328>

- Guo Z and Li D (2002). Advances in the systematics and biogeography of the Bambusoideae (Gramineae) with remarks on some remaining problems. *Acta Bot. Yunnanica* 24: 431-438.
- Jiang B, Liu W, He X, Peng Q, et al. (2013). Characterization and chromosomal distribution of Ty3-gypsy-like retrotransposons in wax gourd (*Benincasa hispida*). *Sci. Asia* 39: 466-471. <http://dx.doi.org/10.2306/scienceasia1513-1874.2013.39.466>
- Kolano B, Bednara E and Weiss-Schneeweiss H (2013). Isolation and characterization of reverse transcriptase fragments of LTR retrotransposons from the genome of *Chenopodium quinoa* (Amaranthaceae). *Plant Cell Rep.* 32: 1575-1588. <http://dx.doi.org/10.1007/s00299-013-1468-4>
- Kumar A and Bennetzen JL (1999). Plant retrotransposons. *Annu. Rev. Genet.* 33: 479-532. <http://dx.doi.org/10.1146/annurev.genet.33.1.479>
- Li WH, Wu CI and Luo CC (1985). A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol. Biol. Evol.* 2: 150-174.
- Librado P and Rozas J (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452. <http://dx.doi.org/10.1093/bioinformatics/btp187>
- Llorens C, Muñoz-Pomer A, Bernad L, Botella H, et al. (2009). Network dynamics of eukaryotic LTR retroelements beyond phylogenetic trees. *Biol. Direct* 4: 41. <http://dx.doi.org/10.1186/1745-6150-4-41>
- Miyata T, Miyazawa S and Yasunaga T (1979). Two types of amino acid substitutions in protein evolution. *J. Mol. Evol.* 12: 219-236. <http://dx.doi.org/10.1007/BF01732340>
- Pereira V (2004). Insertion bias and purifying selection of retrotransposons in the *Arabidopsis thaliana* genome. *Genome Biol.* 5: R79. <http://dx.doi.org/10.1186/gb-2004-5-10-r79>
- Pritham EJ (2009). Transposable elements and factors influencing their success in eukaryotes. *J. Hered.* 100: 648-655. <http://dx.doi.org/10.1093/jhered/esp065>
- Ruiz-Sanchez R (2011). Biogeography and divergence time estimates of woody bamboos: insights in the evolution of Neotropical bamboos. *Bol. Soc. Bot. Mex.* 88: 67-75.
- Sun HY, Dai HY, Zhao GL, Ma Y, et al. (2008). Genome-wide characterization of long terminal repeat -retroelements in apple reveals the differences in heterogeneity and copy number between Ty1-copia and Ty3-gypsy retroelements. *J. Integr. Plant Biol.* 50: 1130-1139. <http://dx.doi.org/10.1111/j.1744-7909.2008.00717.x>
- Suoniemi A, Tanskanen J and Schulman AH (1998). Gypsy-like retrotransposons are widespread in the plant kingdom. *Plant J.* 13: 699-705. <http://dx.doi.org/10.1046/j.1365-313X.1998.00071.x>
- Tamura K, Peterson D, Peterson N, Stecher G, et al. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28: 2731-2739. <http://dx.doi.org/10.1093/molbev/msr121>
- Yang Z, Wong WS and Nielsen R (2005). Bayes empirical bayes inference of amino acid sites under positive selection. *Mol. Biol. Evol.* 22: 1107-1118. <http://dx.doi.org/10.1093/molbev/msi097>
- Zhong H, Zhou M, Xu C and Tang D (2010). Diversity and evolution of *Pong*-like elements in Bambusoideae subfamily. *Biochem. Syst. Ecol.* 38: 750-758. <http://dx.doi.org/10.1016/j.bse.2010.06.010>
- Zhou MB, Lu JJ, Zhong H, Liu XM, et al. (2010a). Distribution and diversity of *PIF*-like transposable elements in the Bambusoideae subfamily. *Plant Sci.* 179: 257-266. <http://dx.doi.org/10.1016/j.plantsci.2010.05.012>
- Zhou MB, Lu JJ, Zhong H, Tang KX, et al. (2010b). Distribution and polymorphism of *Mariner*-like elements in the Bambusoideae subfamily. *Plant Syst. Evol.* 289: 1-11. <http://dx.doi.org/10.1007/s00606-010-0323-0>
- Zhou MB, Zhong H, Zhang QH, Tang KX, et al. (2010c). Diversity and evolution of Ty1-copia retroelements in representative tribes of Bambusoideae subfamily. *Genetica* 138: 861-868. <http://dx.doi.org/10.1007/s10709-010-9469-5>
- Zhou MB, Liu XM and Tang DQ (2011). Transposable elements in *Phyllostachys pubescens* (Poaceae) genome survey sequences and the full-length cDNA sequences, and their association with simple-sequence repeats. *Genet. Mol. Res.* 10: 3026-3037. <http://dx.doi.org/10.4238/2011.December.6.3>

## Supplementary material

**Table S1.** Pairwise similarities between nucleotide sequences from four representative bamboo species. Identities are given in percent. Sequences were aligned using the CLUSTAL W software. The alignments were then transferred to DNASTar, to obtain the identity percentage. Sequences were named according to the first letter of the genus name and the first letter of the species name followed by the clone number.

**Table S2.** Pairwise similarities between amino acid sequences from four representative bamboo species. Identities are given in percent. Sequences were aligned using CLUSTAL W. The alignments were then transferred to DNASTar, to obtain the identity percentage. Sequences were named according to the first letter of the genus name and the first letter of the species name followed by the clone number.