Supplementary information, Data S1

Materials and Methods

BAC DNA sequencing and analysis

BAC libraries were constructed from nuclei extracted from etiolated seedling of maize inbred lines W22 and Yu87-1, or etiolated leaf of a single Zea mays ssp. Parviglumis (PI 384061) plant, as described[1]. BAC clones were screened with a PCR based protocol by two pairs of primer. BAC clones were sequenced and gaps were finished as described before[2]. Homology sequences were aligned using a BLAST 2 program of the National Center of Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov). LTR_FINDER program was used for full-length LTR retrotransposon prediction[3].

PCR and sequencing

The 58-69kb upstream region of tb1 was divided into two or three fragments to amplify by using TaKaRa Ex taq with several primer sets. PCR products were purified with TaKaRa DNA fragment purification kit, and direct sequenced with DYEnamic ET terminator cycle sequencing kit and analyzed on a MegaBACE 4500 DNA capillary sequencer. Sequence data was processed by Phrep/Phrap/Consed software package[4,5], and aligned using Mega software[6].

Polymorphism analysis

The maize association panel was described in Yang et al[7], there are 12 additional lines in this study, and the complete list of maize inbred lines see supplementary table. The teosinte

collection contained 189 accessions (Supplementary table 2), the analysis of each accession was based on two independent DNA samples extracted from single seedling.

Insertional polymorphism of the LTR retrotransposon was amplified with the PCR mixture consisted of 30 ng of genomic DNA, 0.25 μM of each primer, 0.2 mM of each dNTP, 2×GC buffer Ⅰ (TaKaRa, Japan) and 0.5 unit of TaKaRa Ex Taq (TaKaRa, Japan) in a total volume of 20 μl, under the following thermal cycling conditions: 5 min denaturation at 95℃; 40 cycles of 30 sec denaturation at 95℃, 30 sec annealing at 64℃ and 4 min extension at 72℃; followed by a final extension of 10 min at 72℃. PCR products were visualized on 1% agarose gel.

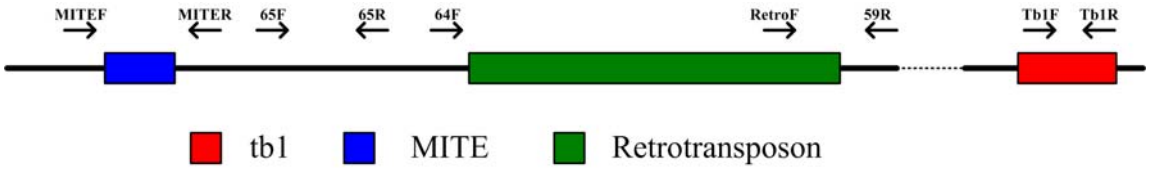
Insertional polymorphism of the MITE was amplified with the PCR mixture consisted of 30 ng of genomic DNA, 0.25 μM of each primer, 0.2 mM of each dNTP, 10× Taq buffer with Mg^{2+} and 1 unit of Taq DNA Polymerase (Tiangen, China) in a total volume of 20 μl, under the following thermal cycling conditions: 5 min denaturation at 95℃; 40 cycles of 30 sec denaturation at 95℃, 30 sec annealing at 60℃ and 1 min extension at 72℃; followed by a final extension of 5 min at 72℃. PCR products were visualized on 2% agarose gel.

Primers

Primer name	Primer sequence (5' To 3')
Tb1F	CGACAAGGCAAGCAAGACGGTAC
Tb1R	GCGTGAGTTCTGCTGAAAGACGA
65F	ACTCAACGGCAGCAGCTACCTA
65R	TAGCGGTAATGTTTGTTGGACTGG
64F[1]	CAGTCCAACAAACATTACCGCTAT
59R[1,2]	CGCGGTCGATCGTTCTGA
RetroF[2]	CAAATCTTGTCTCATGGAGTGC
MITEF[3]	CAGAGGGTACACACGCTAGACAG
MITER[3]	GCCGTTGAGTGTCGCCTAGAC

[1] primer set 1
[2] primer set 2
[3] primer set 3

Approximate location of primers



Sequence data deposition

All sequences were deposited in Genbank under accession numbers JF791320-JF791322, JF791174-JF791242.

References

1 Meng, X. et al., Construction of a Coix BAC library and isolation of the 22 kDa α-coixin gene cluster. *Genome* **53** (9), 667.

2 Zhou, L. et al., The amplification and evolution of orthologous 22-kDa alpha-prolamin tandemly arrayed genes in coix, sorghum and maize genomes. *Plant Mol Biol* **74** (6), 631.

3 Xu, Z. and Wang, H., LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res* **35** (Web Server issue), W265 (2007).

4 Ewing, B., Hillier, L., Wendl, M. C., and Green, P., Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* **8** (3), 175 (1998).

5 Gordon, D., Abajian, C., and Green, P., Consed: a graphical tool for sequence

finishing. *Genome Res* **8** (3), 195 (1998).

6 Tamura, K., Dudley, J., Nei, M., and Kumar, S., MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24** (8), 1596 (2007).

7 Yang, X. et al., Characterization of a global germplasm collection and its potential utilization for analysis of complex quantitative traits in maize. *Mol Breeding*.