Supporting information for Yao *et al.* (April 16, 2002) *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.082562199.

The Structures of the *A1-LC Sh2*, *a1-mum2 Sh2*, and *A1-LH82 Sh2* Haplotypes Are Identical or Nearly Identical in the Region between the *a1* and *yz1* Loci. For example, the sequences of the *A1-LC*, *a1-mum2* and *A1-LH82* alleles are identical (GenBank accession nos. X05068, AF363390, AF363391, AF347696, and U46063). In addition, sequence analysis of the *a1-mum2 Sh2* and *A1-LH82 Sh2* haplotypes and PCR analysis of the *A1-LC Sh2* haplotype has established that all three haplotypes contain the 1.1-kb tandem duplication (TD2) that includes the *Gnat1* insertion (Fig. 2). Further, sequence analysis of the *A1-LH82 Sh2* haplotype and extensive DNA gel blotting, PCR and sequencing analyses of the *a1-mum2 Sh2* and *A1-LC Sh2* haplotypes has established that all three contain the *Ozymandias* and *Machiavelli* retrotransposons (unpublished work). There is, however, a single nucleotide polymorphism (SNP) in *Ozymandias* between the *a1-mum2 Sh2* and *A1-LH82 Sh2* haplotypes. Finally, DNA sequencing of portions of the *A1-LC Sh2* haplotype and many of its derivative recombinant haplotypes has established that the *A1-LC Sh2* haplotype is identical to GenBank accession nos. AF347696 and AF434192 at every SNP or small indel that is polymorphic between the *a1::rdt sh2* haplotype and GenBank accession nos. AF347696/AF434192. Hence, a 21,230-bp sequence assembled from GenBank accession nos. AF434192, AF347696, AF363390, X05068, and AF363391 has been designated the *A1-LC Sh2* haplotype. Positions 1 - 15,783 of "*A1-LC Sh2*" were derived from positions 1 - 15,783 of GenBank accession no. AF434192; positions 15,784 - 16807 of *A1-LC Sh2* were derived from positions 1,321 - 2,344 of GenBank accession no. AF347696; positions 16,808 - 17,075 of *A1-LC Sh2* were derived from positions 1 - 268 of GenBank accession no. AF363390; positions 17,076 - 20,659 of *A1-LC Sh2* were derived from positions 313 - 3,896 of GenBank accession no. X05068; positions 20,660 - 21,209 of *A1-LC Sh2* were derived from positions 1 - 550 of GenBank accession no. AF363391; positions 21,210 - 21,230 of *A1-LC Sh2* were derived from positions 4,447 - 4,467 of GenBank accession no. X05068.

Identification of the Maize *x1* Gene. Computational analysis of the sequences of the rice *a1-sh2* intervals revealed a predicted gene (gene *X*, ref. 1). A 2.1-kb rice cDNA clone of gene *X* (ID # R2277) was obtained from the Japanese Rice Genome Research Program (Tsukuba, Ibaraki 305-0854, Japan) and sequenced. Based on the finding that a shotgun plasmid clone (p2-32H9) from the maize *a1-sh2* interval exhibits a high degree of sequence similarity to exon 6 of the rice gene *X*, this rice cDNA (R2277) was used to screen maize cDNA libraries. Three maize cDNAs (2.6, 1.75, and 1.4 kb) were identified that hybridize to the rice gene *X* and serve as templates for PCR using primers designed based on the sequence of clone p2-32H9. The 2.6-kb maize *x1* cDNA clone (X-V1) was isolated from a library prepared from immature tassels of the inbred W22 (2) and was shown to be full length by means of 5' and 3' Rapid Amplification of cDNA Ends (RACE) (GIBCO BRL, Life Technologies) experiments (data not shown); two other clones contain partial *x1* cDNA clones (X-V3 and X-V5) were isolated from a library prepared from seedlings of the inbred CI31A.

The Sequences of the Oligonucleotides Used as Primers for PCR and Sequencing. QZ1543: 5’-AAACATAAAAACAATACGTAATCCAG-3’; A1.2: 5’-GATTGTTGCTTAAGCGCCAATCGT-3’; XX026: 5’-GAGGTCGTCGAGGTGGATGAGCTG-3’; AE4EI: 5’-CGAATTCCGCCAGGGTTTTAGACA-3’; XX390: 5’-TCGGCTTGATTACCTCATTCT-3’; A1.1: 5’-GTCTTCATTGCACATGCACTGCAC-3’; XX231: 5’-GCCAAACTCTGATTCGCTCCGTG-3’; XX653: 5’-CGAGCCAGGAGCCGACGAAG-3’; AE1SP: 5’-GACTAGTGCCGGTGCGAGCGAGA-3’; XX025: 5’-GGTAGTTGCAGCGTGTGGTGTT-3’; A1522: 5’-GGGAGTTTGGAGTTGGAGAGG-3’; QZ1001: 5’-GATACAGAAGTATATATAAGGGCCAA-3’; QZ1002: 5’-TATTCGTAATGATGTTTAT-3’; QZ3470: 5’-CATCTGAGTGGGAGGCTAAA-3’; QZ2976: 5’-ACTTGTCTCCATCGCTCT-3’; HYilpU7: 5’-AGACGATTGATGATGATTT-3’; a1rdt3273: 5’-GATTGTCTTTAGGGAACTG-3’; HYilpU6: 5’-GCAGTTCCCTAAAGACA-3’; a1rdt2912: 5’-AACACCCCGCTAACAC-3’; HYilpU5: 5’-GTGTTAGCGGGGTGTT-3’; HYilpL4: 5’-ATCTTGATCCTCTTGAAT-3’; HYilpU4: 5’-CGATGATTCAAGAGG-3’; HYilpL3: 5’-GCTTGCTTGCTTCTGGATGT-3’; HYilpU3: 5’-CAAGCATAAGCATCCATC-3’; a1rdt2381: 5’-TCAACCGTGCTACCAACT-3’; a1rdt2332: 5’-CCGAGTGATAGTAAAGACC-3’; a1rdt1885: 5’-AAAACCAAACGAACATACC-3’; HYilpL2: 5’-ATTCGGTATGTTCGTTTGGTT-3’; HYilpU1: 5’-CAGCCTGTACCAACC-3’; HYilpL1: 5’-CGAAACAGTTACCGAGATAG-3’; a1rdt1541: 5’-CGCTAACTATCTCGGTAACT-3’; QZ684: 5’-GGTTTTTGGGAAGCGTCT-3; YZ4725: 5’-AAATGGTCAGGATAGCTTAGTT-3’; IDPyzrdt: 5’-GAAGTTATGTTCGCGGTG-3’; ZH2617: 5’-CGAACAGGGAAGAATGG-3’; ZH2587: 5’-GCCTGGTTAGCGAAGTTG-3’; HYyz2222L: 5’-CGCCAAAAAAAAAAAACA-3’; YZ1: 5’-GCGGCGTTGCTGCTGTA-3’; ZH1748: 5’-CACATCCCCGTCTCCT-3’; ZH1384: 5’-GCCATCTCTACTGTTACCTT-3’; HYx6488L: 5’-ATCTGGGGAAGGGTATCT-3’; XL1: 5’-ATGTTCTTCTTTGAGTG-3’; H9-forward: 5’-ATCGAGGATGATGCAAAG-3’; XL6: 5’-AAAATCCCCTCGCTGTG-3’; XL3: 5’-ATGAGCGGGAGCCTATG-3’; x302: 5’-CTCTCCCATTCTCTTGATTCCT-3’; XL2: 5’-TGTTCAAAGTGGGAGG-3’; x502: 5’-AGGAATAATAGCGGACCACTTG-3’.

References:

1. Chen, M. & Bennetzen, J. L. (1996) *Plant Mol. Biol.* **32,** 999-1001.
2. DeLong, A., Calderon-Urrea, A. & Dellaporta, S. L. (1993) *Cell* **74,** 757-768.