Science
AAAS

Tuned for Transposition: Molecular Determinants Underlying the Hyperactivity of a Stowaway MITE

Guojun Yang, *et al.*

Science **325**, 1391 (2009);
DOI: 10.1126/science.1175688

Tuned for Transposition: Molecular Determinants Underlying the Hyperactivity of a *Stowaway* MITE

Guojun Yang,[1,2,3] Dawn Holligan Nagel,[1] Cédric Feschotte,[4] C. Nathan Hancock,[1] Susan R. Wessler[1]*

Miniature inverted repeat transposable elements (MITEs) are widespread in eukaryotic genomes, where they can attain high copy numbers despite a lack of coding capacity. However, little is known about how they originate and amplify. We performed a genome-wide screen of functional interactions between *Stowaway* MITEs and potential transposases in the rice genome and identified a transpositionally active MITE that possesses key properties that enhance transposition. Although not directly related to its autonomous element, the MITE has less affinity for the transposase than does the autonomous element but lacks a motif repressing transposition in the autonomous element. The MITE contains internal sequences that enhance transposition. These findings suggest that MITEs achieve high transposition activity by scavenging transposases encoded by distantly related and self-restrained autonomous elements.

Most eukaryotic genomes contain large numbers of many different types of transposable elements (TEs). The vast majority of these TEs cannot replicate and mobilize themselves into new regions of the genome (they are nonautonomous). These elements thus depend on transposases encoded by other, autonomous, elements. Miniature inverted repeat TEs (MITEs) are a type of nonautonomous element found in both prokaryotic and eukaryotic genomes, where they are often located in or near genes (*1*–*3*). MITEs resemble typical DNA transposons, although they tend to have a small size [<500 base pairs (bp)] and high copy number and contain terminal inverted repeats (TIRs) flanked by target site duplications. Although MITEs lack coding capacity, in plants they are classified as either *Tourist*-like or *Stowaway*-like (*4*). Although the high copy numbers of MITEs in plant genomes suggest that they have high transposition activity, only the *Tourist*-like rice element *mPing*, an internal deletion derivative of its autonomous partner *Ping*, has been shown to be currently active (*5*–*7*). However, the mechanism by which rice strains accumulate 100- to 1000-fold more *mPing* than *Ping* elements is unknown (*8*).

Unlike *mPing*, the vast majority of characterized MITEs are not deletion derivatives of existing autonomous transposons (*9*, *10*). Furthermore, because none of these MITEs are active, their origin, success, and source of transposase have been a mystery. The most logical model to explain the movement of these nondeletion derivative MITEs is that they can borrow the transposase of distantly related elements (in a process referred to as cross-mobilization) and amplify within the genome (*4*).

The sequenced genome of *Oryza sativa* (rice) contains more than 22,000 *Stowaway* MITEs belonging to at least 25 subfamilies (*10*). Furthermore, the structure of MITE phylogenetic trees indicates that subfamilies are derived from the amplification of one or a few individual elements to hundreds or thousands of copies (*10*). *Stowaway* elements have not previously been shown to be active in rice or in any other genome. Surprisingly, none of the *Stowaway* families appeared to be deletion derivatives of any transposase-encoding element found within sequenced rice genomes. It was predicted that rice *Mariner*-like elements (called *Osmar*s) were the most likely source of transposase because *Stowaway* and *Osmar* share

[1]Department of Plant Biology, University of Georgia, Athens, GA 30602, USA. [2]Department of Biology, University of Toronto at Mississauga, Mississauga, ON L5L 1C6, Canada. [3]Cell and Systems Biology, University of Toronto, Toronto, ON M5S 3G5, Canada. [4]Department of Biology, University of Texas at Arlington, Arlington, TX 76019, USA.

*To whom correspondence should be addressed. E-mail: sue@plantbio.uga.edu

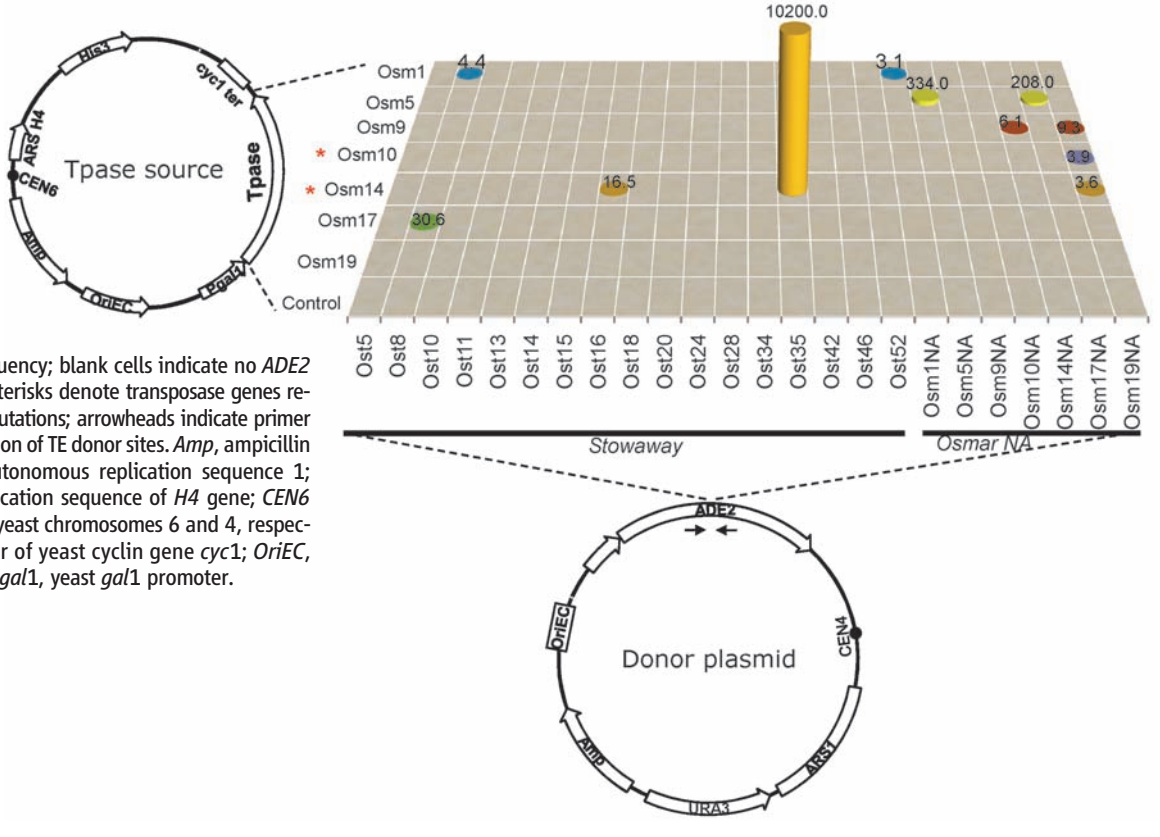

Fig. 1. Yeast excision assay. *Stowaway*s (Osts) and nonautonomous *Osmar*s (OsmNAs) cloned into the *ade*2 coding sequence are shown on the *x* axis; *Osmar* transposases are shown on the *y* axis; and *ADE2* reversion frequency (10^{-9} per cell) (table S1) is shown on the *z* axis. Numbers on cylinders show average *ADE2* reversion frequency; blank cells indicate no *ADE2* reversion was detected. Asterisks denote transposase genes repaired for frame-shifting mutations; arrowheads indicate primer positions for PCR amplification of TE donor sites. *Amp*, ampicillin resistance gene; *ARS1*, autonomous replication sequence 1; *ARS H4*, autonomous replication sequence of *H4* gene; *CEN6* and *CEN4*, centromeres of yeast chromosomes 6 and 4, respectively; *cyc*1 ter, terminator of yeast cyclin gene *cyc*1; *OriEC*, *E. coli* replication origin; P*gal*1, yeast *gal*1 promoter.

terminal inverted repeats (of ~10 bp) and the same target site duplication (the dinucleotide TA) (fig. S1).

On the basis of their transposase protein sequences, *Osmar*s were classified into 25 families and placed in three major clades (*10*). Among the 34 *Osmar*s, only five contain intact transposase-coding regions. Here, we refer to *Osmar*s with complete ends and intact coding sequences as potentially autonomous elements, those with mutated or truncated coding sequences as nonautonomous elements, and small elements lacking any coding sequences as deletion derivatives of *Osmar* (fig. S1).

To determine whether any *Osmar* transposases could catalyze the transposition of rice *Stowaway* MITEs, we modified a yeast assay previously developed to demonstrate transposition of a nonautonomous version of *Osmar*5 (Osm5NA) by its own transposase (*11*). The assay has two plasmid components (Fig. 1): One is a transposase (Tpase) source that expresses one of seven *Osmar* transposases (abbreviated Osm1, Osm5…) representing each *Osmar* subclade under the control of an inducible yeast promoter (P*gal1*) (*12*). The second plasmid in our assay contains the *ade2* reporter gene disrupted by one of 24 nonautonomous elements, including 17 *Stowaway*s (abbreviated Ost5, Ost8…) chosen to represent the diversity of *Stowaway* families in rice, and seven direct-deletion derivatives of each *Osmar* element (abbreviated OsmXNA, where NA means nonautonomous) (*12*). In yeast cells containing both plasmids, potentially successful transposase-transposon interactions were scored on the basis of *ADE2*-revertant colonies (Fig. 1).

These tests revealed that six of the seven transposases showed evidence of activity and uncovered several instances of cross-mobilization (Fig. 1 and table S1). The nonautonomous versions of *Osmar*1 and *Osmar*5 (Osm1NA and Osm5NA) were excised in the presence of their cognate transposases, and cross-mobilization occurred for Osm17NA by Osm5; Osm19NA by Osm9, Osm10, and Osm14; and Osm14NA by Osm9. We also identified interactions between three *Stowaway* elements and *Osmar* transposases: Ost8 by Osm1 and Osm17, and Ost16 and Ost35 by Osm14 (Fig. 1 and table S1). In contrast to the low excision activities observed for Ost8 and Ost16, the excision of Ost35 by Osm14 was the highest interaction recorded by this assay (Fig. 1 and fig. S2A). Polymerase chain reaction (PCR) amplification confirmed independent excision events of Ost35 by the *Osmar*14 transposase (fig. S2B), and sequencing of the PCR products revealed excision footprints (fig. S2C) similar to those retrieved for *Osmar*5 in this and a previous study (fig. S3A) (*11*). Additionally, reinsertions of Ost35 into chromosomal loci were confirmed by means of Southern hybridization of yeast genomic DNA (figs. S2D and S3B). The fact that *Stowaway*s were mobilized by *Osmar* transposases provides a functional mechanism for the mobilization of *Stowaway* MITEs in the rice genome. The observed excision of one *Stowaway* by two distinct transposases and the excision of two different *Stowaway*s by the same transposase suggest that cross-mobilization may be a major mechanism for the amplification of rice MITEs.

The differences in excision frequency between Ost35 and Osm14NA mirrored the difference in copy number between *Stowaway* MITEs and *Osmar* elements in the rice genomes (*10*). Several properties of the MITEs could explain the differential activity of Ost35 as compared with Osm14NA.

First, it is possible that the differences in size between Osm14NA and Ost35 (1004 versus 239 bp) may be a factor because increases in TE size has been shown to reduce transposition efficiency (*13*, *14*). Second, because transposition of *Mariner* elements requires the binding of transposase to transposon TIRs and/or subterminal regions to form synaptic complexes for subsequent excision and reintegration (*15*), any differential binding of *Osmar*14 transposase to the ends of Osm14NA and Ost35 may have affected their transposition activity. Ost35 has shorter TIRs than *Osmar*14 (20 versus 32 bp) and different subterminal sequences. Finally, Ost35 and Osm14NA also differ in their internal sequences, which may enhance or repress excision.

To investigate whether element size affects excision frequency, we shortened Osm14NA (1004 bp) to the length of Ost35 (239 bp) (now called Osm14NAS) (Fig. 2 and fig. S4). Relative to Osm14NA, the shorter Osm14NAS displayed weak but detectable excision activity (Fig. 2). Thus, size alone cannot account for the magnitude of excision activity of Ost35.

To investigate how transposase binding differed between elements, we examined the interaction of purified recombinant *Osmar*14 transposase with synthesized oligonucleotides corresponding to the terminal (32 bp) and subterminal (32 bp) regions of Ost35 and *Osmar*14 (Fig. 3A and fig. S4) in electrophoretic mobility shift assays. These experiments revealed differences in transposase affinity but not in the expected direction. For Ost35, the 5′ and 3′ terminal regions (abbreviated 5′T and 3′T, respectively, in Fig. 3A and fig. S4) each bound transposase equivalently. There was also some binding, albeit less efficiently, to each of the subterminal regions of Ost35 (5′ subT and

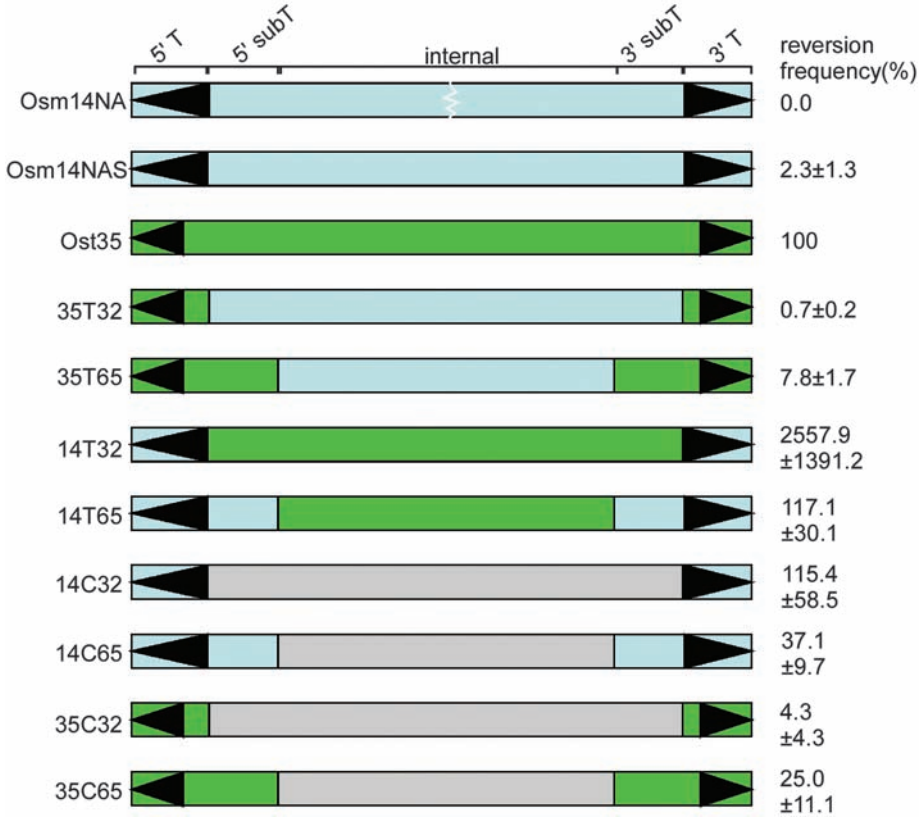

Fig. 2. Contribution of different regions of Osm14NAS and Ost35 to excision activity. *ADE2* reversion frequency is shown to the right of each construct. The chimeric constructs contain the following: 35T32, terminal 32 bp of Osm14NAS replaced by that of Ost35; 35T65, terminal 65 bp of Osm14NAS replaced by that of Ost35; 14T32, terminal 32 bp of Ost35 replaced by that of Osm14NAS; 14T65, terminal 32 bp of Ost35 replaced by that of Osm14NAS; 14C32, control sequence (coding sequence of *gfp*) flanked by terminal 32 bp of Osm14NAS; 14C65, control sequence flanked by terminal 65 bp of Osm14NAS; 35C32, control sequence flanked by terminal 32 bp of Ost35; and 35C65, control sequence flanked by terminal 65 bp of Ost35. Green represents regions derived from Ost35, blue represents regions derived from Osm14NAS, and gray represents regions derived from *gfp*.

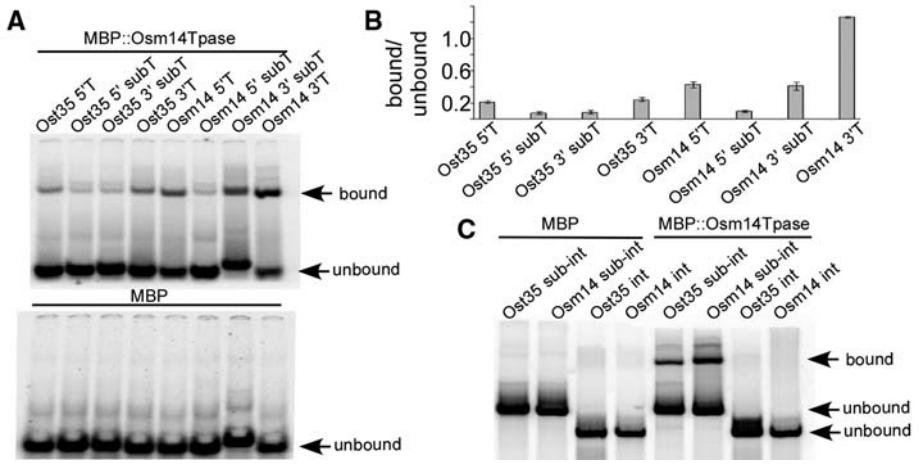

Fig. 3. *Osmar*14 transposase shows binding to regions of Ost35 and Osm14NAS. (**A**) Electrophoretic mobility shift assay (EMSA) of terminal and subterminal regions of Osm14NAS and Ost35. MBP, maltose binding protein control; MBP::Osm14Tpase, fusion protein of MBP and *Osmar*14 transposase; T, terminal 32 bp; subT, subterminal 32 bp. (**B**) Ratio of bound to unbound DNA. Error bars show SE of the mean for three independent events. (**C**) EMSA of the internal sequences of Ost35 and Osm14NAS (*12*). Int, internal sequences; sub-int, internal sequences with subterminal sequences attached were used as positive controls.

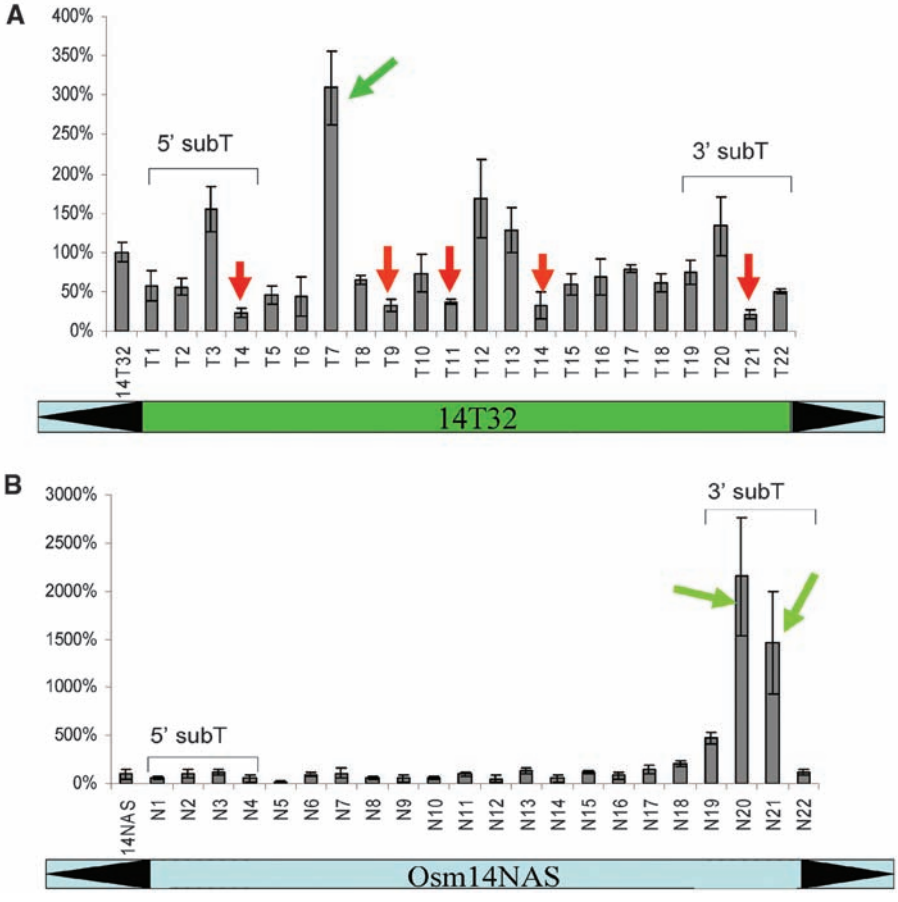

Fig. 4. Mutagenesis analyses of Ost35 and Osm14NAS internal sequences. (**A**) *ADE2* reversion frequency for mutations in the subterminal and internal sequences of Ost35. The template for mutagenesis was 14T32. (**B**) *ADE2* reversion frequency for mutations in the subterminal and internal sequences of Osm14NAS. Positions of mutations correspond to that on the diagram shown beneath each chart. T followed by a number indicates 14T32 mutated at the indicated site; N followed by a number indicates Osm14NAS mutated at the indicated site. Green arrows indicate positions of mutations that resulted in increased activity, and red arrows indicate mutations that resulted in decreased activity.

the 3′ subT in Fig. 3A and fig. S4). In contrast, both the terminal and subterminal regions of *Osmar*14 bound to transposase more strongly than the equivalent regions in Ost35. In addition, binding to *Osmar*14 was asymmetrical because sequences from the 3′ end of the element showed more binding affinity than those from the 5′ end. We detected no binding to the sequences internal to the subTs of either Ost35 or Osm14NA (Fig. 3, B and C, and fig. S4). Furthermore, when unlabeled Osm14NAS or Ost35 was used to compete with radiolabeled Osm14NAS, Ost35 did not outcompete Osm14NAS for transposase binding (fig. S5). These results suggest that the higher excision frequency of Ost35 is not due to an increased affinity for the *Osmar*14 transposase.

We swapped multiple regions between Osm14NAS and Ost35 and confirmed that the central region of Ost35 is involved in enhancing the frequency of excision (Fig. 2). However, this experiment showed, surprisingly, that despite the poor function of Osm14NAS in the excision assay, the replacement of the Ost35 TIRs by those of *Osmar*14 (a chimeric element we called 14T32) resulted in a ~25-fold increase in excision as compared with the original Ost35 transposon or a modified Ost35 with a different internal sequence of the same length (in this case, derived from the *gfp* gene). These data suggest that the function of the TIRs is normally suppressed in Osm14NAS (and also presumably in the natural *Osmar*14 element), whereas the function of the Ost35 TIRs is enhanced by the sequence between them. Further supporting this theory, we observed that the subTs of Osm14NAS represses excision when combined with either the *Osmar*14 TIRs, the equivalent regions of Ost35, or a different internal sequence (Fig. 2). Also, the region of Ost35 between the terminal sequences (32 bp) enhanced excision of adjacent TIRs by ~20-fold relative to the random internal sequence.

Taken together, these data support two general conclusions. For Osm14NAS, the extremely low excision frequency suggests that there is repression of the otherwise optimal *Osmar*14 TIRs by one or both of its subterminal regions. For Ost35, the high excision frequency most likely indicates that there is enhancement of the otherwise suboptimal Ost35 TIRs by part of or the entire Ost35 internal region.

We next performed site-directed mutagenesis so as to more precisely localize the regions responsible for the enhancement and repression of excision in the internal region of Ost35 and the Osm14NAS subTs, respectively. We used 14T32 and Osm14NAS constructs as the templates for mutagenesis. For each construct, we replaced consecutive blocks of eight nucleotides by the sequence ATTTAAAT (*Swa*I restriction site), resulting in 22 derivative constructs from each starting template. Among the mutant constructs derived from 14T32, a 75 to 80% decrease in excision activity was caused by mutations distributed throughout the Ost35 internal region, whereas a single mutation showed an approximately three-

fold increase (Fig. 4A). These data suggest that multiple motifs throughout the Ost35 internal region may promote excision by transposase probably through, for example, favorable DNA curvature or chromatin structure. In contrast, mutagenesis of Osm14NAS pinpointed the repressive region to the 3′ subterminal region where mutations resulted in up to a 20-fold increase in excision activity (Fig. 4B).

Our data support a model for the cross-mobilization of *Stowaway* elements by *Osmar* transposase and suggest how *Stowaway* MITEs may arise and amplify in the genome. Most plant genomes characterized to date, including rice, harbor relatively few *Mariner*-like elements (such as *Osmar*) but hundreds to tens of thousands of *Stowaway* MITEs (*10*, *16*, *17*). Our data suggest that the low copy number of *Osmar*14, and possibly other *Osmar*s, is due in part to a self-regulatory mechanism involving a repressive motif in their 3′ subterminal region. We speculate that this strategy limits the amplification of the elements, thereby attenuating their deleterious effects and facilitating their persistence in the genome.

The vertical persistence of DNA transposons has been hypothesized to be accompanied by diversification of transposase DNA-binding activities (*10*, *18*, *19*), suggesting that their sequence specificity undergoes episodic relaxation. This view is supported by the fact that *Osmar* transposases have surprisingly weak binding specificity (*20*) and can apparently cross-mobilize distantly related *Osmar*s (Fig. 1). Although the apparent promiscuity of *Osmar* transposases may facilitate their survival by buffering any potential impact of inactivating mutations (*18*), it may also allow parasitism by simpler transposons with similar binding sites, such as *Stowaway* MITEs. One factor underlying the success of *Stowaway* MITEs may be the fact that they have minimal cis-requirements for recognition by *Osmar* transposases, coupled to the presence of internal sequences that enhance excision. As short non-coding elements, *Stowaway*s may rapidly increase their copy number by tapping into a transposase source if and when it becomes available. In summary, MITE amplification may differ from previously postulated models of transposon invasion and provides an illustration of the complex ecosystem deployed within the genome. Because MITEs are widespread in eukaryotes, the fundamental principles outlined herein may be applicable to a broad range of organisms.

References and Notes

1. S. R. Wessler, T. E. Bureau, S. E. White, *Curr. Opin. Genet. Dev.* **5**, 814 (1995).
2. G. Yang *et al.*, *Plant Cell* **17**, 1559 (2005).
3. J. Piriyapongsa, I. K. Jordan, *PLoS One* **2**, e203 (2007).
4. C. Feschotte, X. Zhang, S. R. Wessler, in *Mobile DNA II*, N. Craig, R. Craigie, M. Gellert, A. Lambowitz, Eds. (American Society for Microbiology Press, Washington, DC, 2002), pp. 1147–1158.
5. T. Nakazaki *et al.*, *Nature* **421**, 170 (2003).
6. K. Kikuchi, K. Terauchi, M. Wada, H. Y. Hirano, *Nature* **421**, 167 (2003).
7. N. Jiang *et al.*, *Nature* **421**, 163 (2003).
8. K. Naito *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 17620 (2006).
9. N. Jiang, C. Feschotte, X. Zhang, S. R. Wessler, *Curr. Opin. Plant Biol.* **7**, 115 (2004).
10. C. Feschotte, L. Swamy, S. R. Wessler, *Genetics* **163**, 747 (2003).
11. G. Yang, C. F. Weil, S. R. Wessler, *Plant Cell* **18**, 2469 (2006).
12. Materials and methods are available as supporting material on *Science* Online.
13. L. R. Tosi, S. M. Beverley, *Nucleic Acids Res.* **28**, 784 (2000).
14. J. C. Way, N. Kleckner, *Genetics* **111**, 705 (1985).
15. C. Auge-Gouillou, B. Brillet, M. H. Hamelin, Y. Bigot, *Mol. Cell. Biol.* **25**, 2861 (2005).
16. C. Feschotte, S. R. Wessler, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 280 (2002).
17. G. Menzel *et al.*, *Chromosome Res.* **14**, 831 (2006).
18. D. L. Hartl, A. R. Lohe, E. R. Lozovskaya, *Annu. Rev. Genet.* **31**, 337 (1997).
19. D. J. Lampe, K. K. Walden, H. M. Robertson, *Mol. Biol. Evol.* **18**, 954 (2001).
20. C. Feschotte, M. T. Osterlund, R. Peeler, S. R. Wessler, *Nucleic Acids Res.* **33**, 2153 (2005).

21. We thank C. F. Weil and D. J. Garfinkel for materials and technical assistance and J. Bennetzen, X. Zhang, and J. Leebens-Mack for discussion and insightful comments. This work was supported by NIH, the NSF Plant Genome Program, the University of Georgia Research Foundation, and the University of Toronto. Sequences for *Osmar* transposase coding sequences and their nonautonomous elements were deposited in GenBank (accession numbers GQ379705 to GQ379718 and GQ382183).

Supporting Online Material

www.sciencemag.org/cgi/content/full/325/5946/1391/DC1
Materials and Methods
Figs. S1 to S5
Table S1
References
Sequences and Primers

1 May 2009; accepted 16 July 2009
10.1126/science.1175688

The RNA-Binding Protein NANOS2 Is Required to Maintain Murine Spermatogonial Stem Cells

Aiko Sada,[1] Atsushi Suzuki,[2] Hitomi Suzuki,[3] Yumiko Saga[1,3,4]*

Stem cells give rise to differentiated cell types but also preserve their undifferentiated state through cell self-renewal. With the use of transgenic mice, we found that the RNA-binding protein NANOS2 is essential for maintaining spermatogonial stem cells. Lineage-tracing analyses revealed that undifferentiated spermatogonia expressing *Nanos2* self-renew and generate the entire spermatogenic cell lineage. Conditional disruption of postnatal *Nanos2* depleted spermatogonial stem cell reserves, whereas mouse testes in which *Nanos2* had been overexpressed accumulated spermatogonia with undifferentiated, stem cell–like properties. Thus, NANOS2 is a key stem cell regulator that is expressed in self-renewing spermatogonial stem cells and maintains the stem cell state during murine spermatogenesis.

Stem cells are essential for tissue homeostasis and regenerative responses to injury and disease. In the spermatogenic stem cell system, germ cell–intrinsic factors have an essential role in the maintenance of stem cells for the continuation of spermatogenesis throughout life (*1*–*5*). However, the previous loss-of-function studies have some limitations in terms of understanding the mechanism by which stem cells are lost upon the gene deletion, as it could be caused by cell death, defective self-renewal, premature differentiation, or other mechanisms.

For decades, the mammalian spermatogenic stem cell has been characterized by the morphological features of the spermatogonia. The spermatogonial types A_{single} (A_s; isolated single cells), A_{paired} (A_{pr}; chains of 2 cells), and $A_{aligned}$ (A_{al}; chains of 4, 8, 16 or 32 cells) are the most primitive germ cells observed in mature testes and are collectively described as undifferentiated spermatogonia. They give rise to differentiating spermatogonia, which undergo additional divisions and enter a differentiation pathway. It has been proposed that only A_s spermatogonia represent the stem cells (*6*–*8*); however, there is no A_s-specific molecular marker, and the presence of stem cells is assayed by long-term colony formation after the transplantation of candidate cells into recipient testes (*9*). For this reason, undifferentiated spermatogonia containing A_s to A_{al} are the smallest population proven to have the properties of stem cells.

Recently, two functionally distinct spermatogonial stem cell populations were identified in mice (*10*). One is the population that acts as the self-renewing stem cells (actual stem cells), and the other population possesses the potential to self-

[1]Department of Genetics, SOKENDAI, 1111 Yata, Mishima, Shizuoka 411-8540, Japan. [2]Interdisciplinary Research Center, Yokohama National University, 79-1 Tokiwadai, Hodogaya-ku, Yokohama 240-8501, Japan. [3]Department of Biological Sciences, Graduate School of Science, University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo, 113-0033, Japan. [4]Division of Mammalian Development, National Institute of Genetics, 1111 Yata, Mishima 411-8540, Japan.

*To whom correspondence should be addressed. E-mail: ysaga@lab.nig.ac.jp