Supporting Information

Li et al. 10.1073/pnas.1121543109

SI Methods

DNA Synthesis. The *Space Invaders* ($SPIN_{ON}$) transposase ORF flanked by NcoI and SnaBI sites and a mammalian codon-optimized *TcBuster*1 transposase ORF ($TcBuster_{CO}$) flanked by KpnI and NotI sites were synthesized by DNA2.0 and cloned in their pJ201 vector. A *SPIN* transposon segment containing 250 bp L and 250 bp R ends flanked by XhoI sites and separated by NcoI, BamHI, SpeI, NdeI, and EcoRI sites was synthesized by DNA2.0 and cloned in their pJ241 vector to give pJ241-*SPIN*-D.

Mammalian Cell Transposition. Transposition was measured in HeLa cells by cotransfecting a donor plasmid containing a minitransposon containing segments from the ends of the element flanking an antibiotic resistance gene or other marker (designated as p-source-transposon-D-marker) and a helper plasmid expressing a transposase under a human CMV promoter (designated as p-source-CMV-transposon).

Mammalian piggyBac plasmids for Figs. 1 and 4–6, Fig. S3, and Table S4. ***Mammalian helper plasmid* pXL-CMV-*piggyBac*.** The *Tribolium ni* transposase ORF sequence was PCR-amplified from *piggyBac* [*PB*-Helper (a gift from David O'Brochta, University of Maryland, College Park, MD)] using the primer PB_{for} with a KpnI site (5′-GA *GGTACC* GCCACCATGGGTAGTTCTTTAGACGATGAG; KpnI is in italics and start codon is underlined) and PB_{rev} containing a *NotI* site (5′CAG *GCGGCCGC* TCAGAAACAACTTTGGCACATATCAATATTATG; NotI is in italics and stop codon is underlined) digested with KpnI and NotI and cloned into KpnI and NotI sites of pcDNA3.1/myc-HisA (Invitrogen) such that transposase expression was driven by the CMV promoter.

***Mammalian donor plasmid* pXL-*PB*-D-GFP/Bsd.** The donor plasmid containing a *T. ni. piggyBac*-GFP/Bsd element was made by digesting pCMV/Zeo (Invitrogen) with MluNI and EcoRI to release the Zeo gene and 3′ UTR and ligating in its place an MluNI-EcoRI fragment of the Bsd gene and the 3′ UTR from pTracer/CMV-Bsd (Invitrogen) made by PCR amplification using primers Bsd_{for} containing an MluNI site (5′-GAC *TGGCCA* AGCCTTTGTCTCAAGAAG; MluNI is in italics) and Bsd_{rev} containing an EcoRI site (5′-CA *GAATTC* AGACATGATAAGATACATTGATGAGTTTGG; EcoRI is in italics) to make pCMV-Bsdx. The left end of *piggyBac* was then cloned between the ApaI and XhoI sites of pCMV/Bsdx by PCR-amplifying the *piggyBac* left end from *pB*ac-TIR (terminal inverted repeat) (a gift from David O'Brochta) using primers *PB*-$Left_{for}$ containing an ApaI site (5′-CAG *GGGCCC* TTAA CCCTAGAAAGATAGTCTGC; ApaI site is in italics and TIR is underlined) and *PB*-$Left_{rev}$ (5′-GAC *CTCGAG* GACAATGTTCAGTGCAGAGACTC; XhoI is in italics), yielding a 673-bp *PB*-left end fragment flanked by TTAA, digesting the fragment with ApaI and XhoI, and cloning it between the ApaI and XhoI sites of pCMV/Bsdx. The right end of *PB* was PCR-amplified using primers *PB*-$Right_{for}$ containing a NarI site (5′-CAG *GGCGCC* TTAACCCTAGAAAGATAATCATATTG; NarI is in italics, TIR is underlined) and *PB*-$Right_{rev}$ containing an XbaI site (5′-CAG*TCTAGA* GTCGAGAGCATAATATTGATATGTG; XbaI site is in italics), yielding a 400-bp *PB*-right end fragment that was digested with NarI and XbaI and cloned between the NarI and XbaI sites of pCMV/Bsdx containing the left end TIR and generating pCMV-*PB*-D-Bsd. A DNA fragment extending from within 5′ UTR end to the middle of the GFP/Bsd fusion gene from pTracer/CMV-Bsd was isolated by digestion with RsrII and BlpI restriction enzymes and then cloned between the RsrII and BlpI sites of pXL-CMV-*PB*-D-Bsd, thus creating a GFP-Bsd fusion gene to give the donor plasmid pXL-*PB*-D-GFP/Bsd in which GFP/Bsd is expressed from a CMV promoter and flanked downstream by an SV40 polyA site.

Mammalian SPIN plasmids for Figs. 1 and 4–6, Fig. S3, and Table S3. ***Mammalian helper plasmid* pHE-CMV-*SPIN*$_{ON}$.** The $SPIN_{ON}$ transposase gene was PCR-amplified from pJ201-$SPIN_{ON}$ using a KpnI-containing primer from the 5′ end of the gene (5′-CGG *GGTACC* ATGACAATGGATCGTGTAGAGAAGAATGTTAAGAAAAG; KpnI is in italics and start codon is underlined) and a XhoI-containing primer from the 3′ end of the gene (5′-CCG *CTCGAG* TCAGTGCGAAGGTTGAGATTGCTTCTTTCTCACCAGATCAG; XhoI is in italics and the termination codon is underlined). This KpnI-XhoI fragment was cloned between the KpnI and XhoI sites of pcDNA3.1 myc-HisA (Invitrogen) such that transposase expression was driven by the *CMV* promoter.

***Mammalian donor plasmid* pXL-*SPIN*-D-GFP/Bsd.** The GFP/Bsd cassette was PCR-amplified from pXL-*PB*-D-GFP/Bsd using primers containing SpeI sites, $SpeI_{for}$ (5′-CTAG *ACTAGT* CCGTTACATAACTTACGGTAAATGGCCCGCCTG; SpeI is in italics) and $SpeI_{rev}$ (5′-CTAG *ACTAGT* ATCCCCGGGAATTCAGACATGATAAGATACATTGATGAGTTTGGAC; SpeI is in italics), and cloned into the SpeI site between the *SPIN* ends in the donor plasmid pJ241-*SPIN*-D plasmid to give p*SPIN-D*-GFP/Bsd. The p*SPIN*-D-GFP/Bsd element was then amplified with a *SPIN* left primer containing a NheI site and an 8-bp flanking imperfect target duplication (gtcttgaa) as described in the work by Pace et al. (1) (5′-GCGATATC *GCTAGC* TCGAG gtttgaa CAGTGGTTCTCAACCTTCCTAATG; NheI is in italics and the TIR sequence is underlined) and a *SPIN* right primer containing a PstI site and the other 8-bp flanking imperfect target site duplication (ctttaatg) described in the work by Pace et al. (1) (5′-CAA *CTGCAG* ctttaatg CAGCGGTTCTCAACCT GTGGG; PstI is in italics and TIR is underlined). This *SPIN*-GFP/Bsd fragment was digested with NheI and PstI and cloned between the NheI and PstI sites of pCMV/Zeo (Invitrogen) to create pXL-*SPIN*-D-GFP/Bsd.

Mammalian TcBuster plasmids for Figs. 2 and 3. ***Mammalian helper plasmid* pPWA-CMV-*TcBuster*.** The *TcBuster1* ORF was amplified from *T. castanaeum* Ga-2 DNA (obtained from Sue Brown, Kansas State University, Manhattan, KS). PCR was performed with the primers *TcBuster1* ORF-For containing an EcoRV site (5′-AAT*GATATC*AGAAATATGATGCTGAATTGGCTCAAAAGTGG; EcoRV is in italics and the start codon is underlined) and *TcBuster1* ORF-Rev containing an EcoRV site (5′-GAT*GATATC*TTAATGACTTTTTTGCGCTTGCTTATTATTGCAC; EcoRV is in italics and the stop codon is underlined). The PCR product was digested with EcoRV, and cloned into plasmid pKHsp70 (2) digested with SmaI to generate pKHsp70*TcBuster*1. The sequence of the ORF is identical to the sequence from the genome search file except for a single base change, which results in a Ser-Thr substitution at amino acid 551. Thr at this position was found in all of the clones sequenced for the *TcB*1 ORF. The pPWA CMV-*TcBuster* helper plasmid was constructed by digestion of pCMV-HSB16 (a *Sleeping Beauty* helper plasmid encoding a hyperactive transposase) (3) with KpnI and XbaI to remove the HSB16 ORF followed by ligation to the KpnI-XbaI *TcBuster* ORF fragment from pKHsp70*TcBuster*.

***Mammalian donor plasmid* pPWA-*TcB*-D-Neo.** *TcBuster1* ends were amplified from *T. castanaeum* Ga-2 DNA (obtained from Sue Brown, Kansas State University, Manhattan, KS). pBs-*TcBuster*L was made by PCR amplification using primers that

encompassed the region from the target site duplication of *TcBuster* to the sequence immediately upstream of the ATG of the *TcBuster* ORF such that the L end was 328 bp. The primers used were *TcBuster* L_{for} (5′-AAT *GGTACC* CTTTAGGC CAG-TGTTCTTCAACCTG; KpnI is in italics and TIR is underlined) and *TcBuster* L_{Rev} (5′-CAT*CTCGAG*ATTTCTGAACGATT-CTAGGTTAGGATCAAAC; XhoI is in italics and transposon end is underlined). After amplification, the DNA digested with KpnI and XhoI and ligated into the KpnI-XhoI site of *pB*luescript SK+ (Stratagene) to make *p-TcBuster*-L.

The right end of *TcBuster* was amplified in a similar manner using the primers *TcBuster* $Right_{for}$ containing an XbaI site (5′-GAT *TCTAGA* CAACTGATCCATCCCGATATTGATAAT-TTGTGC; XbaI in italics) and *TcBuster1* R_{rev} containing a SacI site (5′-AAT *GAGCTC* GTATAAG CAGTGTTTTCAACCT-TTGCCATCC; SacI is in italics and TIR is underlined). The R end contained 145 bp of the transposon. After amplification, the PCR products were purified, digested with SacI and XbaI, quantified, and ligated to *pPWA-TcBuster*L digested with SacI and XbaI clone to make pPWA-*TcBuster-D*. The mammalian donor vector pT-PWAcB-D-Neo was made by ligating the 1,670-bp BamHI-HpaI fragment of pT-SVNeo (3), which contains the SV-40 promoter, and Neomycin resistance ORF to the BamHI-EcoRV fragment of pPWA-*TcBuster*-D-Neo.

***Mammalian piggyBac plasmids used in Figs. 2 and 3. Mammalian helper plasmid* pPWA-CMV-*piggyBac*.** pPWA-CMV-*PB* was constructed by blunt end ligation of the vector fragment of pCMV-*HSB*16 (3) obtained by digestion with BamHI and XhoI, and the ends were filled in with Klenow polymerase with a *piggyBac* transposase fragment.

***Mammalian donor plasmid* pPWA-*PB*-D-Neo.** The pPWA-*PB*-D-*Neo* donor was constructed by ligation of the BamHI-HpaI fragment of pT-SVNeo (3) with the pPWA-*PB*-D plasmid containing 353 bp from the left end of *PB* and 357 bp of the right end (4).

***Mammalian Sleeping Beauty plasmids used in Fig. 2* Upper *and* Lower.** The mammalian helper plasmid was pCMV-HSB16, and the mammalian *SB* donor plasmid was pT-SVNeo, which were obtained from James Baus (3).

***Mammalian TcBuster plasmids used in Figs. 4–6, Fig. S3, and Table S3. Helper plasmid* pXL-CMV-*TcBuster*.** The helper plasmid pXL-CMV-*TcBuster* was constructed by PCR amplification of the *TcBuster* ORF from pKHsp70*TcBuster* with a primer $TcBuster_{for}$ containing an KpnI site (5′-GC *GGTACC* GCCACC ATG ATG-CTGAATTGGCTCAAAAGTGG; KpnI is in italics and start codon is underlined) and $TcBuster1_{rev}$ containing a NotI site (5′-CAG *GCGGCCGC* TTA ATGACTTTTTTGCGCTTG-CTTATTATTG; NotI is in italics and the stop codon is underlined) followed by cloning between the KpnI and NotI sites of pcDNA3.1/myc-HisA.

***Mammalian helper plasmid* pXL-CMV-$TcBuster_{CO}$.** The mammalian helper plasmid pXL-CMV-$TcBuster_{CO}$ was constructed by excision of the mammalian codon optimized $TcBuster_{CO}$ fragment from pJ201:$TcBuster_{CO}$ by digestion with KpnI and NotI and ligation of the resulting $TcBuster1_{CO}$ ORF fragment between the KpnI and NotI sites of pcDNA3.1/myc-HisA.

***Mammalian helper plasmid* pXL-CMV-$TcBuster_{CO}$ V596A.** The mutant $TcBuster1_{CO}$ V596A was made by site-directed mutagenesis using the QuikChange Site-Directed Mutagenesis Kit (Stratagene) and the oligonucleotide CCTTTGTGAAAAATC-TTTTTCTGTATAC gcc (GTG to GCC) GCGACTAAGACC-AAATACCGG and its complementary strand.

***Mammalian donor plasmid* pXL-*TcB*-D-GFP/Bsd.** The mammalian plasmid pXL-*TcBuster*-GFP/Bsd was made by PCR amplification of the *TcBuster*1 left end from pTcB-D-Neo using a *TcBuster*1-L_{for} primer with a NarI site (5′-CAG *GGCGCC* CC CTTTA-GGC CAGTGTTC; NarI is in italics and TIR is underlined) and a *TcBuster*1-$Left_{rev}$ primer with PstI site (5′-GG *CTGCAG* GA-ATTCGATATCAAG; PstI is in italics), which was then cloned into pXL-GFP/Bsd digested with NarI and PstI to make pXL-*TcBuster*-Left-GFP/Bsd. The *TcBuster* right end was PCR-amplified using a *TcBuster*1-$Right_{for}$ primer with an ApaI site (5′-CAG *GGGCCC* GTATAAAG CAGTGTTCTTCAAC; ApaI is in italics and TIR underlined) and a *TcBuster*-$Right_{rev}$ end primer with a BglII site (5′-GA *AGATCT* ACA ACGATCCA-TCCCGATATTG; BglII is italics), and it was cloned between the ApaI and BglII sites of pXL *TcB*-Left-GFP/Bsd to generate pXL-*TcB*-D-GFP/Bsd.

***Sleeping Beauty plasmids used in the experiment in Figs. 5 and 6, Fig. S3, and Table S4. Mammalian helper plasmid* pXL-CMV-*Sleeping Beauty*.** The mammalian helper plasmid *pXL-CMV-Sleeping Beauty* was made by excising the *Sleeping Beauty* ORF from pCMV/HSB16 (3) by digestion with KpnI and NotI and cloning it between the KpnI and NotI sites of pcDNA3.1/myc-HisA such that transposase expression was driven by the CMV promoter.

***Mammalian donor plasmid* pXL-*SB*-D-GFP/Bsd.** The mammalian *Sleeping Beauty* transposon donor plasmid *pXL-SB-D-GFP/Bsd* was constructed by PCR amplification of the *Sleeping Beauty* left end from pT-SVNeo (3) with a *Sleeping Beauty* $Left_{for}$ primer containing a BglII site (5′-GA *AGATCT* GAGCTCGGTACC-CTACAGTTG; BglII is in italics and TIR is underlined) and a *Sleeping Beauty*-$Left_{rev}$ primer containing a NotI site (5′-GAC *GCGGCCGC* CTTAGTGTATGTAAACTTCTGACC; NotI is in italics and TIR is underlined) from pT-SVNeo to give a 228-bp left end that was cloned between the BglII and NotI sites of pXL-GFP/Bsd to give pXL-*SB*-L-GFP/Bsd. The *Sleeping Beauty* right end was PCR-amplified from pT-SVNeo with a *Sleeping Beauty*-$Right_{for}$ primer containing a HindIII site (5′-CGTC *AA-GCTT* GGATCCCCTACAGTTGAAGTCG; HindIII is in italics and TIR is underlined) and a *Sleeping Beauty* $Right_{rev}$ primer containing a XbaI site (5′-GC *TCTAGA* CTTCTGACCCACT-GGGAATGTG; XbaI is in italics,) to give a 210-bp right end, and it was cloned between the HindIII and XbaI sites of pXL-*SB*-L-GFP/Bsd to give pXL-*SB*-D-GFP/Bsd.

HeLa Cell Integration Assays for Figs. 2 and 3. Cells were transfected in six-well plates at 300,000 cells/well with FuGENE 6 (Roche) according to the manufacturer's protocol. Two days posttransfection, cells were removed with trypsin and counted. Cells were seeded at multiple dilutions in triplicate into new plates, and G418 was added to 1 mg/mL. Fresh media with antibiotic was added at 1 wk, and cells were incubated for 1 additional wk. Cells were then rinsed with PBS, fixed with 3.7% formaldehyde for 15 min, and stained for 30 min with 0.35% methylene blue in PBS before counting colonies.

All Other HeLa Cell Integration Assays. HeLa cells were maintained in DMEM (Invitrogen) with 5% FBS (HyClone) and 1× penicillin-streptomycin (Invitrogen), which contained 100 units penicillin (base) and 100 μg streptomycin (base)/mL. For transfection, cells on a 100-mm plate were washed with PBS (10 mL) and trypsinized with 1 mL Trypsin-EDTA (Invitrogen). The cells were then seeded at about 300,000 cells/well in a six-well plate and incubated at 37 °C for 0.5–2 h; 5 μl FUGENE 6 (Roche) was added to 95 μl DMEM without FBS or antibiotics, vortexed briefly. 0.4 μg helper DNA and 1.6 μg donor DNA were added. pcDNA3.1/ Myc-His was used as the negative control. The mixtures were vortexed briefly and incubated for 15–45 min at room temperature. FuGENE 6-DNA complex was added to cells dropwise, and the solution was swirled well. The plates were then incubated at 37 °C. After 2 d, the cells were washed and trypsinized (100 μL), and 0.9 mL DMEM supplemented with 5% FBS and 3.5 μg/mL Bsd were added. The cells were diluted in medium (usually 1:200, 1:100, or 1:50) and seeded into a six-well plate. The cells were then incubated at 37 °C for 18–21 d, and the medium was changed every 2–3 d. The cells were then washed with 2 mL PBS, fixed with 1.5 mL 10% formaldehyde in PBS for

10 min, stained for 1 h with 0.2% methylene blue in PBS, rinsed with distilled water to destain, and air-dried, and the colonies were counted.

Transposon Integration Library. HeLa cells were seeded at 2×10^6 in 10-cm plates and transfected with FuGENE 6 with 2.0 μg helper plasmid and 8.0 μg donor plasmid as described above; 48 h after transfection, cells were trypsinized and diluted into 6–8 10-cm plates and cultured in DMEM with 5% FBS containing 3.5 μg/mL blasticidine. Drug selection was continued for 18–21 d, and the media were changed every 2–3 d. Surviving cells were harvested, and genomic DNAs were prepared using DNeasy Blood & Tissue Kit (Qiagen).

Integration Site Recovery, 454 Sequencing, and Analysis. Integration sites were recovered as described (5) Briefly, genomic DNA was extracted from an integration transfection library using the DNeasy tissue kit (Qiagen). Two micrograms genomic DNA were digested overnight with ApoI or BstYI and then ligated to linkers overnight at 16 °C. Nested PCR was then carried out under stringent conditions using end-specific primers complementary to transposon sequences and linker-specific primers complementary to the DNA linker. DNA barcodes were included in the second-round PCR primers to track sample origin. The PCR products were gel-purified, pooled, and sequenced using 454 sequencing platform. Only sequences that uniquely aligned to the human genome by BLAT (hg18, version 36.1, >98% match score) and began within 3 bp of the LTR end were used in downstream analyses.

Oligonucleotides for ApoI and BstYI Linkers for 454 Libraries. The oligonucleotide 454L-2 5′-GTAATACGACTCACTATAGGGCTCCGCTTAAGGGAG was annealed with 454S-Apo-2 5′-P-AATTCTCCCTTAAGCGGAG-NH_2 to make a linker with an ApoI sticky end. The oligonucleotide 454L-2 5′-GTAATACGACTCACTATAGGGCTCCGCTTAAGGGAG was annealed with 454S-BstY-2 5′-P-GATCCTCCCTTAAGCGGAG-NH_2 to make a linker with a BstYI sticky end.

Primers for PCR Amplification from the ApoI and BstYI Linkers for All 454 Libraries. The first round PCR primer was 454-link GTAATACGACTCACTATAGGGCTC, and the second-round nested PCR linker primer was 454A2-2 5′-gcctccctcgcgccatcagAGGGCTCCGCTTAAGGGAG (454 adaptor is in lowercase).

***SPIN*-Specific 454 Library Primers.** For the first round, priming of 50–72 bp *SPIN* R end *Spin*-R1-2, 5′-GTCTTAGGCGACCCCTGTGAAAG was used. For the second-round PCR primers, priming of 11–29 bp of *SPIN* R end was used: SpR454-1 *SPIN*1/ApoI 5′-GCCTTGCCAGCCCGCTCAG**GACATCTG**GGTCGCGACCCACAGGTTG (bold is the barcode and underlined letters are the transposon end). SpR454-2: *SPIN*2/ApoI = 5′-GCCTTGCCAGCCCGCTCAG**GGTAGCTT**GGTCGCGACCCACAGGTTG. SpR454-3: *SPIN*3/ApoI 5′-GCCTTGCCAGCCCGCTCAG**GTAGTGCT**GGTCGCGACCCACAGGTTG. SpR454-4: *SPIN*1/BstYI 5′-GCCTTGCCAGCCCGCTCAG**AGCACATC**GGTCGCGACCCACAGGTTG. SpR454-5: *SPIN*2/BstYI 5′-GCCTTGCCAGCCCGCTCAG**AGGTCTAC**GGTCGCGACCCACAGGTTG. SpR454-6: *SPIN*3/BstYI 5′-GCCTTGCCAGCCCGCTCAG**CTGAGAGA**GGTCGCGACCCACAGGTTG.

***piggyBac*-Specific 454 Library Primers.** For first-round PCR primer, priming of 46–91 bp *PB* R end was used for PBR454-1 5′-AAACCTCGATATACAGACCGATAAAACACATGCGTCAATTTTACGC. For second-round nested PCR, priming of 10–42 bp *PB* R end was used for PBR454-3 for *PB*1/BstYI 5′-GCCTTGCCAGCCCGCTCAGTGTCACAGATTATCTTTAACGTACGTCACAATATGATTATC, PBR454-4 for *PB*2/BstYI 5′-GCCTTGCCAGCCCGCTCAG**TCTGACAG**ATTATCTTTAACGTACGTCACAATATGATTATC, PBR454-5 for *PB*3/BstYI 5′-GCCTTGCCAGCCCGCTCAG**CATGACAG**ATTATCTTTAACGTACGTCACAATATGATTATC, PBR454-20 for *PB*1/ApoI 5′-GCCTTGCCAGCCCGCTCAG**GACTCTGT**ATTATCTTTAACGTACGTCACAATATGATTATC, PBR454-21 for *PB*2/ApoI 5′-GCCTTGCCAGCCCGCTCAG**GACACAGT**ATTATCTTTAACGTACGTCACAATATGATTATC, and PBR454-22 for *PB*3/ApoI 5′-GCCTTGCCAGCCCGCTCAG**GACTCAGT**ATTATCTTTAACGTACGTCACAATATGATTATC.

***TcBuster*-Specific 454 Library Primers.** For first-round PCR, priming of 41–63 bp *TcBuster* L end was used for TcL454-1, 5′-GGAACCCCGAAAGCCTTTGGGTG. For second-round PCR primers, priming L-end 15–35 bp was used for TcL454-3 for both *TcB*/ApoI and *TcB*/BstYI 5′-GCCTTGCCAGCCCGCTCAG**CAGTGACT**CTAGGGTTCCGCGGAACACAG.

***Sleeping Beauty*-Specific 454 Library Primers.** For first-round PCR priming, primer 78–105 bp *SB* R end was used for SBR454-1, 5′-GACCTTAAGACAGGGAATCTTTACTCGG. For second-round PCR primers, priming 12–40 bp *SB* R end was used for SBR454-3 for *SB*/ApoI 5′-GCCTTGCCAGCCCGCTCAG**AGACCAGT**GTATTTGGCTAAGGTGTATGTAAACTTCC and SBR454-5 for *SB*/BstYI 5′-GCCTTGCCAGCCCGCTCAG**TGACCAGT**GTATTTGGCTAAGGTGTATGTAAACTTCC.

Bioinformatic Analysis of Target Site Duplications and Other Genomic Features. A 32-bp target DNA sequence surrounding each integration site was extracted from the draft human genome (hg18), and aligned using WebLogo (http://weblogo.Berkeley.edu/logo.cgi). Detailed bioinformatic methods for analysis of association with chromosomal features are described in the work by Berry et al. (6). The methods for generating heat maps based in receiver operating characteristic curves are as described in the work by Berry et al. (6). A detailed description of the methods used to generate the genomic features heat map in Fig. 5*A* can be found at http://microb230.med.upenn.edu/assets/doc/HeatMapGuide_v12_formatted.doc.

Yeast Plasmids. The *TcBuster* transposase ORFs were expressed under the control of the *GALS* promoter in the *TRP*1 p*GALS* plasmid p414GalS (7). The $TcBuster_{CO}$ and $TcBuster_{CO}$ V596A ORFs were PCR-amplified from the mammalian *TcBuster* expression vectors using a primer for the 5′ end of the ORF containing an XmaI site (5′-CTC *CCCGGG* ATGATGCTGAATTGGCTGAAAAGC; XmaI is in italics and start codon is underlined) and one for the 3′ end containing an XhoI site (5′-GAC *CTCGAG* TCA GTG AGA TTT CTG GGC CTG C; stop codon is underlined) and cloned between the XmaI and XhoI sites of p*GALS*.

The *TcBuster* donor plasmid in the integration assay was a derivative of pRS416, which contains a *URA3* backbone (8). First, the KpnI-SacI fragment containing the *TcBuster*-Neo element from pPWA-*TcB*-D-Neo that contains a 328-bp left *TcBuster* end segment and a 145-bp right *TcBuster* end segment was cloned between the KpnI and SacI sites of pRS416 that had a point mutation in the plasmid CEN to facilitate plasmid loss on removal of selection (9). The element is flanked by imperfect 8-bp repeats of **ctttaggc**CAGTG....CACTG**ctttatac**. The Neo gene was excised by digestion with SalI and SpeI and replaced with the SalI-SpeI NatMX fragment from pAG/NatMX to give the mini *TcBuster*-ClonNAT^R element. This element was then inserted into the XhoI site on the plasmid backbone by amplification of the *TcBuster*-miniClonNAT element with a 5′ pRS416-*TcBuster* left primer (5′-CTCACTATAGGGCGAATTGGGTACCGGGCCCCCCC*CTCGAG***ctttaggc**CAGTGTTCTTCAACCTGTGTTCCGCGGAAC-3′; XhoI is in italics and TIR is underlined) and a 3′ pRS416-*TcBuster* right primer (5′-GGAATTCGATA-

TCAAGCTTATCGATACCGTCGAC*CTCGAG***gtataaag**CAG-TGTTCTTCAACCTTTGCCATCCGGCGG; XhoI is in italics and TIR is underlined), by homologous recombination with pRS416 linearized by XhoI digestion. The integration assay (see below) selects for cells lacking the *URA3* donor plasmid (i.e., are FOA^R but retain mini*TcBuster*-$ClonNat^R$).

The donor plasmid in the excision assay was a Ura^- derivative of the HIS^+ plasmid pRX1 that contains the *URA3*::actin intron (9) containing the *TcBuster* element in the actin intron. The mini *TcBuster* element containing a 328-bp left-end segment and a 145-bp right-end segment flanking a *ClonNAT* gene was generated by PCR amplification of the mini *TcBuster*-ClonNAT element in the integration donor vector, including the imperfect flanking repeats using a primer that contained homology to the *TcBuster* ends and the portion of the actin intron 5′ of the intron XhoI site [CACCATCCCATTTAACTGTAAGAATTGCACG-GTCCCAATGC*CTCGAG***ctttaggc**CAGTGTTCTTCAACCTGT-GTTC; XhoI is in italics and TIR is underlined; the imperfect terminal repeats found in the work by Pace et al. (1) are shown in bold and to the 3′ portion of the actin intron-*TcBuster* primer (5′-GAGGTTATGGGAGAGTGAAAAATAGTAAAAAAAGG-TAAAAGAGAAATCT*CTCGAG***gtataaag**CAGTGTTCTTCA-ACCTTTGCCA; XhoI is in italics and TIR is underlined)]. The element was introduced into the *URA3*::actin intron plasmid of pRX1 by homologous recombination with pRX1 linearized by XhoI digestion. The mRNA of the *URA3*::actin intron::mini-*TcBuster* donor is too large to splice, resulting in Ura^- auxotrophy but after mini *TcBuster* excision, splicing can occur, resulting in Ura^+ prototrophy. Thus, excision can be followed by looking for Ura^- to URA^+.

***TcBuster* Integration Assays.** Integration assays were performed in BY4727 (*MAT α-his3Δ200 leu2Δ0 lys2Δ0 met150 trp1Δ63 ura3Δ0*) (10) after transformation with the Trp^+ p*GALS TcBuster* transposase helper plasmid and the pRS416 *URA3* mini *TcBuster*-ClonNAT donor plasmid. Single colonies were grown on SC-TRP-URA + 2% galactose plates for 5 d, resuspended in water, diluted, and plated onto SC + glucose + 1 mg/L 5-FOA plates to measure cells lacking the donor plasmid and on SC + glucose + 1 mg/L 5-FOA + 100 mg/L ClonNAT to measure cells lacking the donor plasmid but containing a chromosomal mini *TcBuster* insertion. The integration frequency is the number of cells lacking the donor plasmid and containing a chromosomal mini *TcBuster* insertion/cells lacking the donor plasmid.

Analysis of *TcBuster* Insertion Sites in Yeast. Genomic DNA was isolated from several 5-FOA^R $ClonNAT^R$ yeast strains using the Epicentre yeast MasterPure yeast DNA purification kit and used as a template for rapid amplification of genomic DNA sequences adjacent to *TcBuster* insertions (11) using a semirandom, two-step nested PCR using genomic DNA as a template. For isolation of sequences adjacent to the left end of *TcBuster1*, a first round of PCR were carried out using P1-Left1 and P1. The end-specific linker was P-Left1, 5′-GAAACTCCGATTGACGCAT-GTGCATTCTGAAG; P1 is a random primer that contains a specific 20-nt sequence followed by 10 bases of degenerate sequence and a specific 5-nt sequence (5′-GGCCACGCGTCGA-CTAGTACNNNNNNNNNNGATAT).

Isolation of sequences adjacent to the *TcBuster* R end involved a first round of PCR with a right end-specific primer P-Right1, 5′-GCACGGGCTCACCTTGTTCGTAACAAGTCAAC, and the random P1 primer. For the second round of PCR, PCR products from the first reaction serve as the template, and new pairs of primers (P-Left2, 5′-CCGACGTCTCTCGAATTGAA-GCAATGACTCGC, P2 5′-GGCCACGCGTCGACTAGTAC, P-Right2, 5′-GCTGTCCCTAAAATCTCATCTGGGTGTATT-AC, and P2 direct amplification) were used. The resulting PCR products were gel-purified (Qiagen) and topocloned into pCR2.1 (Invitrogen) for sequencing. The target site duplication was, thus, verified from sequence analysis from both transposon ends.

Yeast Excision Assay. Excision assays were performed in BY4727 (*MAT α-his3Δ200 leu2Δ0 lys2Δ0 met150 trp1Δ63 ura3Δ0*) (10) transformed with the TRP+ p*GALS* $TcBuster_{CO}$ transposase helper plasmids and the HIS+ *URA3*::actin intron::*TcBuster*-ClonNat excision donor plasmid. Excision resulted in a change from Ura^- auxotrophy to Ura^+ prototrophy. Cells containing both plasmids were streaked on SC-HIS-TRP plates containing 2% glucose, and single colonies were resuspended, diluted, and grown on SC-HIS-TRP plates containing 2% galactose for 5 d. Colonies were resuspended in water, diluted, and plated on SC-HIS-TRP-URA + 2% glucose plates containing measured excision events and SC + 2% glucose plates to count all cells. The excision frequency is the number of Ura^+ cells/total cells.

Excision footprints at the donor site in Ura^+ colonies from SC-HIS-TRP-URA plates were determined by PCR amplification of the *URA3*::actin intron using the yeast colony PCR. Colonies were resuspended into 25 μL 20 mM NaOH, boiled for 15 min, briefly vortexed for 3 min, and spun at 20,800 × *g* for 10 min; 1 μL resulting supernatant was used in 50 μL PCR using 10 nM forward primer that anneals at 5′ of the *URA3* gene (5′-ATG-TCGAAAGCTACATATAAGGAACG; start codon is underlined) and 10 nM reverse primer that anneals at 3′ of the *URA3* gene (5′-TTAGTTTTGCTGGCCGCATCTTCTCA; stop codon is underlined); the PCR products were cloned into pCR2.1 (Invitrogen) and sequenced.

$TcBuster_{CO}$ and $SPIN_{ON}$ Transposase Expression and Purification from Yeast. Strep-HA tags (IBA) were added to the N termini of the $TcBuste_{CO}$ and $SPIN_{ON}$ transposases, allowing for purification by affinity chromatography on Strep-Tactin (IBA). The $TcBuster_{CO}$ transposase gene was PCR-amplified from the yeast p*GAL*S vector using an XmaI site-containing primer for the 5′ end of the gene (5′-TCC*CCCGGG*ATGATGCTGAATTGGCTGAAAA-GCGGTAAA; XmaI is in italics and start is underlined) and a XhoI containing primer at the 3′ end of the gene (5′-CCG-*CTCGAG*TCAGTGAGATTTCTGGGCCTGCTTGTTGTT; XhoI is in italics and stop is underlined), and it was cloned between the XmaI and XhoI sites of the pKS1-ST expression vector (Dualsystems Biotech). The $SPIN_{ON}$ transposase gene was PCR-amplified from the pJ201 *SPIN* ORF vector using a BamHI site-containing primer for the 5′ end of the gene (5′- CG*G-GATCC*ATGACAATGGATCGTGTAGAGAAGAACGTT; BamHI is in italics and start is underlined) and a XhoI-containing primer at the 3′ end of the gene (5′-CCG-*CTCGAG*TCAGTGCGAAGGTTGAGATTGCTTCTTTCT; XhoI is in italics and stop is underlined), and it was cloned between the BamHI and XhoI sites of the pKS1-ST expression vector. In pKS1-ST, expression of the inserted gene is driven by the *ADH2* promoter that is induced when glucose becomes depleted in the medium (12). Expression clones were transformed into *Saccharomyces cerevisiae* strain DSY-5 (*MAT α-leu2 trp1 ura3-52 his3::GAL1-GAL4 pep4 prb1-11222*) (Dualsystems Biotech) for protein expression.

A 5-mL starter culture was prepared by inoculating a single colony into YPD + 60 mg/L G418 and then incubating for 8 h at 30 °C with shaking at 250 rpm; 20 mL fresh YPD + 60 mg/L G418 were added, and incubation continued overnight at 30 °C with shaking at 250 rpm. Two liters YPD + 60 mg/L G418 were then added to the overnight culture to an OD_{600} of 0.2, and the culture was then incubated for 20 h at 30 °C with shaking at 250 rpm. The cells were harvested by centrifugation at 4,000 × *g* for 20 min, the supernatant was discarded, and the cell pellet washed one time with 50 mL water;. Two milliliters extraction buffer [25 mM Tris·HCl, pH 8.0, 150 mM NaCL, 10% V/V glycerol, 5 mM DTT, 1 mM phenylmethylsulphonyl fluoride (PMSF)] per gram

of wet weight pellet were added, the pellet was resuspended by gentle pipeting, and the suspension was transferred to a 50-mL Falcon tube; 500 μL glass beads/gram pellet wet weight were added, and the tube was vortexed for 1 min. Vortexing was repeated 10 times each, placing the tube on ice for 1 min between vortexing steps. The sample was then centrifuged at 4,000 × *g* for 30 min in a cooled centrifuge to remove unbroken cells. The resulting soluble extract was transferred to a beaker, 0.23 g solid $(NH_4)_2SO_4$ were added per ml extract, and the extract was incubated for 1 h at 4 °C. The protein precipitate was collected by centrifugation for 15 min at 38,000 × *g* and resuspended in 20 mL column buffer (50 mM Tris·HCl, pH 8.0, 150 mM NaCl, 10% V/V glycerol, 5 mM DTT, and 1 mM PMSF). The resolubilized protein pellet was incubated with 40 μg/mL avidin for 1 h to ensure complete depletion of free cellular biotin and some biotinylated proteins that will interfere with binding of target protein to the resin. The resolubilized protein fraction was loaded onto strep-tactin column, and the column bed volume was adjusted to 1 mL. The column was washed with 10 volumes extraction buffer, and transposase protein was eluted with extraction buffer containing 2.5 mM desthiobiotin (Sigma-Aldrich); 1 mL fractions were collected. Fractions containing transposase were pooled, and glycerol was added to a final concentration of 25% and stored at −80 °C. More than 60% of transposase protein was recovered and was 90% pure as judged from protein gels.

In Vitro Strand Transfer Using a Precleaved Substrate. The strand transfer reaction was performed with a 145-bp *TcBuster* right-end fragment generated by PCR amplification from yeast integration donor and a 250-bp *SPIN* right-end fragment that was generated by PCR amplification from pJ241-*SPIN*-D. The substrates were radiolabeled at their 5′ ends using polynucleotide kinase and γ-P^{32}-ATP, and they were purified using a Sephadex G-50 column (Roche); 150 nM transposase was incubated with a 1.5 nM radiolabeled transposon end and 10 nM pUC19 plasmid as the target DNA in 25 mM Mops, pH 7.0, 25 mM Tris·HCl, pH 8.0, 37.5 mM NaCl, 1 mM $MnCl_2$, 5 mM DTT, 20% (vol/vol) glycerol or DMSO, and 100 μg/mL BSA in a final volume of 20 μL at 37 °C in 1 h. The reaction was stopped by adding 2 μL 10% SDS/200 mM EDTA, and it was incubated for 15 min at 65 °C. The reaction products were displayed on a 1% agarose gel in 1× TAE running buffer; the gel was run at 90 V for 2–3 h at room temperature. Radioactivity was analyzed on dried gels using a Typhoon.

In Vitro Coupled Cleavage and Strand Transfer Reactions. For *TcBuster*, the substrate for the coupled cleavage and strand transfer reaction was a 245-bp fragment containing a 145-bp *TcBuster* right-end fragment flanked by 100 bp donor DNA that was PCR-amplified from the donor plasmid used in the yeast integration experiments. For *SPIN*, the substrate was a 350-bp substrate containing 250 bp *SPIN* right-end fragment and 100 bp flanking DNA that was PCR-amplified from the *SPIN*-containing pJ241-*SPIN*-D vector. The substrates were radiolabeled at their 5′ ends using polynucleotide kinase and γ-P^{32}-ATP, and they were purified using a Sephadex G-50 column (Roche); 150 nM transposase was incubated with 1.5 nM radiolabeled transposon end and 10 nM pUC19 plasmid as the target DNA in 25 mM Mops, pH 7.0, 25 mM Tris·HCl, pH 8.0, 37.5 mM NaCl, 1 mM $MnCl_2$, 5 mM DTT, 20% (vol/vol) glycerol or DMSO, and 100 μg/mL BSA in a final volume of 20 μl at 37 °C in 1 h. The reaction was stopped by adding 2 μL 10% SDS/200 mM EDTA, and it was incubated for 15 min at 65 °C. The reaction products were displayed on a 1% agarose gel in 1× TAE running buffer; the gel was run at 90 V for 2–3 h at room temperature. Radioactivity was detected on dried gels using Typhoon.

Analysis of the Covalent Link Between the Transposon and Target DNA. Strand transfer reactions were performed as above except that the 250-bp *SPIN* right-end and the 145-bp *TcBuster* right-end substrates 5′ ends were labeled on only one strand, the one that contains the 3′ tip of the transposon. The products were displayed on a 1% agarose gel run in 50 mM NaOH and 1 mM EDTA at 15 V overnight at 4 °C.

Analysis of *TcBuster* Hairpin Formation. Cleavage and strand transfer reactions were performed with a 173-bp *TcBuster*-R end fragment containing 63 bp *TcBuster* R end DNA and 110 bp flanking DNA derived from the yeast integration donor plasmid. The 174-bp *TcBuster*-R fragment was radiolabeled on the 3′ ends of both strands using Klenow polymerase and α-P^{32}-dATP; 150 nM *TcBuster* was incubated with 1.5 nM radiolabeled *TcBuster*-R DNA in 25 mM Mops, pH 7.0, 10 mM Mg^{2+} or 1 mM Mn^{2+}, 20% DMSO, 2 mM DTT, 100 μg mL^{-1} BSA, and 10 nM pUC19 plasmid as target DNA in a final volume of 20 μL at 37 °C for 0, 30, 60, and 120 min. Reactions were stopped by adding 2 μL 10% SDS/200 mM EDTA and incubated for 15 min at 65 °C. The products were phenol/chloroform-treated, ethanol-precipitated, and resuspended in original reaction volume (20 μL) in nuclease-free water, and one-half of the reaction (10 μL) was displayed on 1% agarose gels to look for double-strand breaks and joining to target DNA. For analysis of DNA nicking and hairpin formation, the other one-half was run on a 5% urea acrylamide gel.

1. Pace JK, 2nd, Gilbert C, Clark MS, Feschotte C (2008) Repeated horizontal transfer of a DNA transposon in mammals and other tetrapods. *Proc Natl Acad Sci USA* 105: 17023–17028.
2. Arensburger P, et al. (2005) An active transposable element, Herves, from the African malaria mosquito Anopheles gambiae. *Genetics* 169:697–708.
3. Baus J, Liu L, Heggestad AD, Sanz S, Fletcher BS (2005) Hyperactive transposase mutants of the Sleeping Beauty transposon. *Mol Ther* 12:1148–1156.
4. Smith R (2007) *Cell, Molecular and Developmental Biology* (University of California, Riverside, CA).
5. Wang GP, Ciuffi A, Leipzig J, Berry CC, Bushman FD (2007) HIV integration site selection: Analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. *Genome Res* 17:1186–1194.
6. Berry C, Hannenhalli S, Leipzig J, Bushman FD (2006) Selection of target sites for mobile DNA integration in the human genome. *PLoS Comput Biol* 2:e157.
7. Mumberg D, Müller R, Funk M (1994) Regulatable promoters of Saccharomyces cerevisiae: Comparison of transcriptional activity and their use for heterologous expression. *Nucleic Acids Res* 22:5767–5768.
8. Sikorski RS, Hieter P (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics* 122:19–27.
9. Gangadharan S, Mularoni L, Fain-Thornton J, Wheelan SJ, Craig NL (2010) DNA transposon Hermes inserts into DNA in nucleosome-free regions in vivo. *Proc Natl Acad Sci USA* 107:21966–21972.
10. Brachmann CB, et al. (1998) Designer deletion strains derived from Saccharomyces cerevisiae S288C: A useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* 14:115–132.
11. Chun KT, Edenberg HJ, Kelley MR, Goebl MG (1997) Rapid amplification of uncharacterized transposon-tagged DNA sequences from genomic DNA. *Yeast* 13: 233–240.
12. Lee KM, DaSilva NA (2005) Evaluation of the Saccharomyces cerevisiae ADH2 promoter for protein synthesis. *Yeast* 22:431–440.

```
Hermes          MQKMDNLEV-----KAKI-------------------------NQ-----
TcBuster1       MM-LNWLKSGKLESQSQEQSSCYLENSNCLPPTLDSTDIIGEENKAGTTS
SPINon          MT-MDRVEK-----N-----------------------------------
                *  :: ::      :

Hermes          -GLYKITPRHKGTSFIWNVLADIQKEDDTLVEGWVFCRKCEKVLKYTTRQ
TcBuster1       RKKRKYDEDYLNFGFTWTGDKDEPNG---------LCVICEQVVNNSSLN
SPINon          VKKRKYSEDFLQYGFTSIITAGIEKP---------QCVICCEVLSAESMK
                    *    .   .*      .  :           *  * :*:.  : :

Hermes          TSNLCRHKCCA--SLK--QSRELKTVSADCKKEAIEKCAQWVVRDCRPFS
TcBuster1       PAKLKRHLDTKHPTLKGK-SEYFKRKCNELNQKKH-TFERYVRDDNK--N
SPINon          PNKLKRHFDSKHPSFAGKDTNYFRSKADGLKKARLDTGGKYH-KQNV--A
                 . :*  **      ::    :. ::  .  ::    .  ::   :

Hermes          AVSGSGFI------------------DMIKFFIKVKAEYGEH--VNVEE
TcBuster1       LLKASYLVSLRIAKQGEAYTIAEKLIKPCTKDLTT--CVFGEKFASKVDL
SPINon          AIEASYLVALRIARAMKPHTIAEDLLLPAAKDIVR--VMIGDEFVTKLSA
                 :..* ::                      * :       *:.   ::.

Hermes          LLPSPITLSRKVTSDAKEKKALISREIKSAVEKDGASATIDLWTDNYIKR
TcBuster1       VPLSDTTISRRIEDMSYFCEAVLVNRLKNAK--CGFTLQMDESTDVAGLA
SPINon          ISLSNDTVRRRIDDMSADILDQVIQEIKSAPL-PIFSIQLDESTDVANCS
                :  *  *: *:: . :      : ..:*.*      :  :*  **

Hermes          NFLGVTLHYHENNELRDLILGLKSLDFERSTAENIYKKLKAIFSQFNVED
TcBuster1       I-LLVFVRYIHESSFEEDMLFCKALPT-QTTGEEIFNLLNAYFEKHSIPW
SPINon          Q-LLVYVRYINDGDFKDEFLFCKPLEM-TTTARDVFDTVGSFLKEHKISW
                    * * ::* .:..:.: :*  *.*    :*..:::. : : :.:..:

Hermes          LSSIKFVTDRGANVV----------KSLA-NNIRINCSSHLLSNVLENSF
TcBuster1       NLCYHICTDGAKAMVGVIKGVIARIKKLVPDIKASHCCLHRHALAVKRIP
SPINon          EKVCGVCTDGAPAMLGCRSGFQRLVLNESPKVIGTHCMIHRQILATKTLP
                     . ** .  ::           .   .    :*  *    . :

Hermes          EETPELNMPILACKNIVKYFKKANLQHRLRSSLKSEC-----------PT
TcBuster1       NA---LHEVLNDAVKMINFIKSRPLNARVFALLCDDLGSLHKNLLLHTEV
SPINon          QE---LQEVMKSVISSVNFVKASTLNSRLFSQLCNELDAPNNALLFHTEV
                :    *:  :    . :::.*   *: *: : * .:             .

Hermes          RWNSTYTMLRSILDNWESVIQILSEAGETQR-IVHINKSIIQTMVNILDG
TcBuster1       RWLSRGKVLTRFWELRDEIRIFFNER---EFAGKLNDTSWLQNLAYIADI
SPINon          RWLSRGKVLKRVFELRDELKTFFNQKARPQFEALFSDKSELQKIAYLVDI
                ** *  .:*  . :  :.:  ::.:    :      :.* :*.:. : *

Hermes          FERIFKELQTCSSPS-LCFVVPSILKVKEICSPDVGDVADIAKLKVNIIK
TcBuster1       FSYLNEVNLSLQGPNSTIFKVNS---------------------RINSIK
SPINon          FAILNELNLSLQGPNATCLDLSE---------------------KIRSFQ
                *  : :   : ..*.   : : .                     ::. ::

Hermes          NVRIIWEENLSIWHYTAFFFYPPALHMQQEKVAQIKEFCLSKMEDLELIN
TcBuster1       SKLKLWEECIT----------------------K---------NNTECFA
SPINon          MKLQLWQKKLD----------------------E---------NKIYMLP
                    :*:: :                       :         :.   :

Hermes          RMSSFNELSATQLNQSDSNSHNSIDLTSHSKDISTT-SFFFPQLTQNNS-
TcBuster1       NLNDFLETSNTALD---PNLK--SNILEHLNGLKNTFLEYFPPTCNN-IS
SPINon          TLSAFFEEHDIEPD---KRITMIISVKEHLHMLADEISSYFPNLPDTPFA
                   :. * *      :    .    .:  .* : :      :**   :.

Hermes          --REPP-VC-PSDEFEF-YRKEIVILSE------DF---KVMEWWNLNSK
TcBuster1       WVENPFNECGNVDTLPIKEREQLIDIRTDTTLKSSFVPDGIGPFWIKLMD
SPINon          LARSPFT--VKVEDVPETAQEEFIELINSDAARTDFSTMPVTKFWIKCLQ
                  ..*       : .   ::::: :        .*    :  :*    .

Hermes          KYPKLSKLALSLLSIPASSAASERTFSLAGNIITEKRNRIGQQTVDSLLF
TcBuster1       EFPEISKRAVKELMPFVTTYLCEKSFSVYVATKTKYRNRLDAED-DMRLQ
SPINon          SYPVLSETVLRLLLPFPTTYLCETGFSSLLVIKSKYRSRLVVED-DLRCA
                .:* :*: .:  *    ::  .*  **      :: *.*:  :  *

Hermes          LNSFY--------KNFCKLDI
TcBuster1       LTTIHPDIDNLCNNKQAQKSH
SPINon          LAKTAPRISDLVRKKQSQPSH
                * .          :: .: .
```

Fig. S1. Comparison of *hAT* elements. Alignment of the *hAT Ac* subfamily element *Hermes* and the *Buster* subfamily elements *TcBuster1* and $SPIN_{ON}$. As determined crystallographically (1), protein domains are color-coded: orange, zinc finger; green, specific end binding; blue, RNase H transposase and integrase core; red, insertion domain. Bold black underlined acid amino acids are the catalytic triad, and the bold black underlined W is involved in hairpinning.

1. Hickman AB, et al. (2005) Molecular architecture of a eukaryotic DNA transposase. *Nat Struct Mol Biol* 12:715–721.

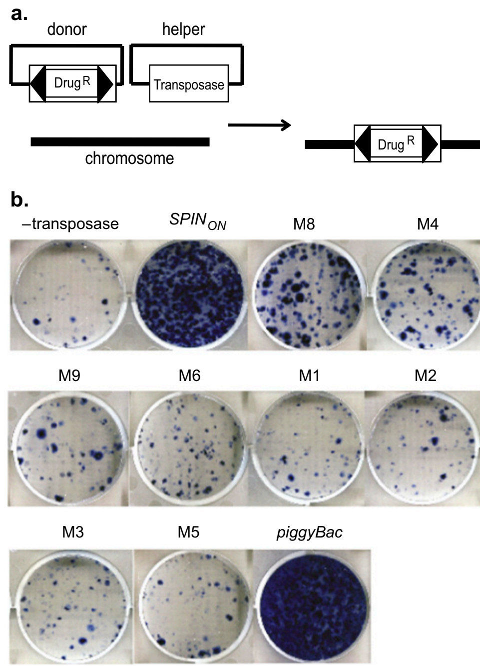
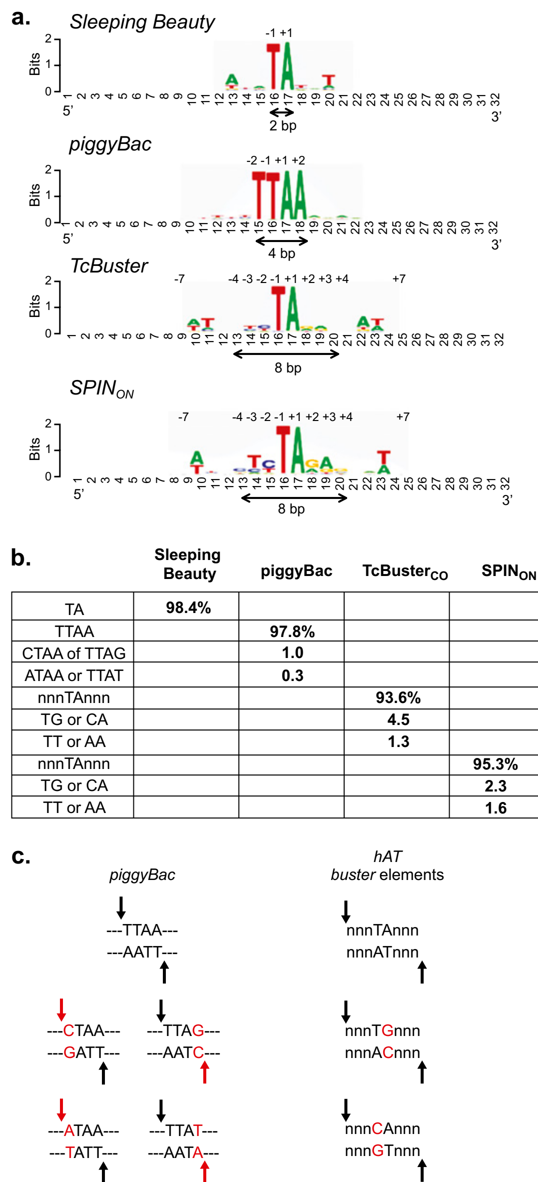
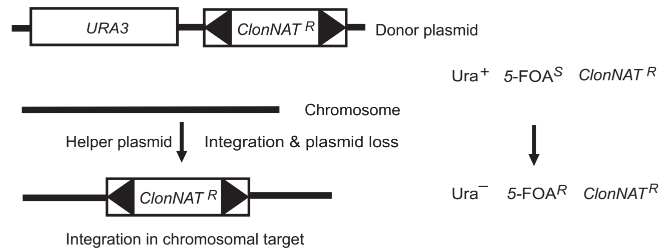


Fig. S2. Transposition of *SPIN* derivatives. (*A*) Transposition is promoted by $SPIN_{ON}$ elements into mammalian cells after transfection of donor and helper plasmids; (*B*) Transposition events are collected by growth in the antibiotic for which the element carries a resistance gene and stained with methylene blue to identify viable cells with integration events.



	Sleeping Beauty	piggyBac	$TcBuster_{CO}$	$SPIN_{ON}$
TA	98.4%			
TTAA		97.8%		
CTAA of TTAG		1.0		
ATAA or TTAT		0.3		
nnnTAnnn			93.6%	
TG or CA			4.5	
TT or AA			1.3	
nnnTAnnn				95.3%
TG or CA				2.3
TT or AA				1.6

Fig. S3. Analysis of transposon target site selection in mammalian cells. Large numbers of de novo insertion sites in HeLa cells for *Sleeping Beauty* (4,490), *piggyBac* (13,494), *TcBuster* (6,390), and $SPIN_{on}$ (8,333) were isolated after cotransfection of transposon donor plasmids and transposase plasmids in HeLa cells, and they were analyzed in a variety of ways. (*A*) Web logos of the 32-bp region surrounding the transposon insertion site are shown, The target site duplication for each element is shown by the two-headed arrow. (*B*) The different target site duplications (TA for *Sleeping Beauty*, TTAA for *piggyBac*, and TA for the central base pairs of the *hAT Buster* elements) recovered are listed along with the percentage of perfect duplications. (*C*) The positions of transposon end joining are indicated by black arrows for *piggyBac* and *hAT Buster* insertions. The sequences variations in the major classes of nonperfect target site duplications are shown with red arrows and base pairs.



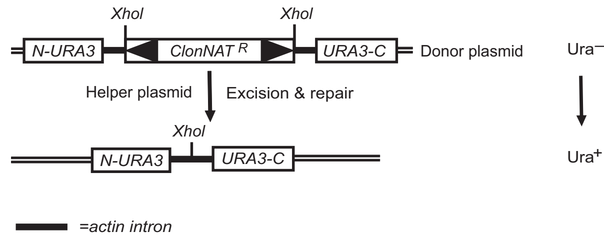


Fig. S4. Yeast assays for integration and excision. Schematic depictions of yeast transposition assays. (*A*) Integration. The transposition of a mini *TcBuster*-ClonNat^R from a *URA3* plasmid to the chromosome is followed by isolation of 5-FOA^R (*ura3*) plasmid-free cells that still contain the mini *TcBuster* ClonNat^R element. (*B*) Excision. The excision of a mini *TcBuster* element from a *URA3*-actin intron::mini *TcBuster* allele that is Ura$^-$ is followed by looking for Ura$^+$ cells. Excision of the transposon allows splicing of the much smaller *URA3*-actin intron.

Fig. S5. In vitro transposition by (*A*) $TcBuster_{CO}$ and (*B*) $SPIN_{ON}$ transposases using different metals.

Table S1. DNA sequences of *$SPIN_{ON}$* and *$TcBuster_{CO}$* transposases

$SPIN_{ON}$

ATGACAATGGATCGTGTAGAGAAGAACGTTAAGAAAAGAAAATATAGTGAGGATTTTTT
ACAGTATGGTTTTACCTCAATAATTACAGCAGGAATTGAGAAACCGCAATGTGTTATTT
GTTGTGAAGTTCTATCAGCCGAATCTATGAAGCCGAACAAACTAAAACGCCATTTTGAT
AGCAAGCATCCGAGCTTTGCCGGCAAGGATACCAACTATTTTAGAAGCAAAGCTGATGG
ACTCAAGAAAGCCAGACTTGACACTGGTGGCAAGTACCACAAACAAAACGTAGCAGCCA
TTGAAGCTTCATATTTGGTGGCACTCAGAATCGCCAGAGCTATGAAACCTCACACCATT
GCTGAGGATTTACTGTTGCCAGCGGCCAAAGACATTGTTCGAGTTATGATCGGAGACGA
ATTTGTTACGAAATTGAGTGCAATTTCCTTATCTAACGACACTGTCCGCAGAAGAATAG
ATGACATGTCTGCTGATATTCTTGATCAGGTAATCCAGGAAATTAAATCTGCTCCACTT
CCAATATTTAGTATCCAGCTTGATGAATCTACAGACGTTGCAAACTGTTCACAGTTACT
GGTTTACGTGAGGTATATTAATGATGGCGACTTTAAAGATGAGTTTCTTTTTTGCAAAC
CTCTTGAAATGACAACTACTGCACGTGATGTATTTGACACAGTTGGTTCATTTCTGAAA
GAGCATAAGATCTCTTGGGAAAAGGTTTGTGGTGTTTGCACAGATGGTGCTCCAGCTAT
GCTAGGATGTCGATCTGGATTTCAACGTTTGGTACTGAATGAGTCACCAAAAGTCATCG
GAACTCACTGTATGATTCATCGGCAAATATTAGCAACGAAGACGCTGCCACAAGAGTTA
CAAGAAGTAATGAAAAGCGTCATAAGTTCTGTCAATTTTGTAAAGGCGAGCACTTTAAA
CAGTCGACTGTTTTCGCAACTGTGCAATGAGTTGGATGCGCCGAACAATGCTCTGCTAT
TTCACACTGAAGTGAGATGGTTGTCGAGAGGAAAAGTTTTAAAACGTGTTTTTGAGCTT
CGTGATGAACTCAAAACGTTTTTTAATCAGAAAGCAAGACCGCAGTTCGAAGCACTTTT
CAGCGATAAAAGTGAACTGCAGAAAATAGCTTACTTGGTTGACATCTTTGCCATCTTGA
ATGAGTTAAATTTATCACTGCAAGGACCAAATGCAACATGCCTCGATTTGTCTGAAAAG
ATCCGATCATTCCAAATGAAACTTCAGCTTTGGCAAAAAAAATTGGATGAAAATAAAAT
TTACATGTTGCCTACCTTATCTGCTTTCTTTGAGGAACATGACATTGAACCAGACAAAA
GGATTACGATGATAATTTCTGTGAAAGAACACTTGCACATGCTTGCAGACGAAATTTCA
TCGTACTTTCCAAATCTACCTGACACCCCATTTGCACTTGCCAGAAGCCCATTCACAGT
CAAAGTTGAAGATGTTCCTGAGACAGCACAAGAGGAGTTCATTGAACTTATTAACAGCG
ATGCAGCGAGAACTGATTTCTCTACAATGCCAGTTACAAAATTCTGGATCAAGTGTTTG
CAGTCATATCCTGTTCTGTCTGAGACTGTGTTGCGCCTTCTTCTTCCATTTCCAACAAC
ATATCTTTGTGAAACAGGGTTTTCCAGCTTGTTGGTTATCAAGTCTAAATACAGAAGTA
GACTTGTTGTGGAAGATGATCTTCGTTGTGCTCTTGCAAAGACTGCCCCGAGAATTTCT
GATCTGGTGAGAAAGAAGCAATCTCAACCTTCGCACTGA

$TcBuster_{CO}$

ATGATGCTGAATTGGCTGAAAAGCGGTAAACTTGAGTCACAGTCCCAAGAGCAGAGCAG
TTGCTACCTTGAGAACAGTAATTGTCTGCCACCCACTCTGGACTCCACGGATATTATTG
GTGAGGAAAACAAAGCTGGCACCACAAGTCGGAAGAAGAGGAAGTATGACGAGGACTAC
CTGAACTTTGGGTTTACCTGGACAGGCGATAAGGATGAACCCAATGGACTCTGCGTGAT
TTGCGAGCAGGTTGTGAACAACAGTTCACTGAATCCCGCCAAGCTTAAACGGCATCTCG
ACACAAAGCACCCTACACTGAAGGGGAAAAGTGAGTACTTCAAAAGGAAGTGCAACGAG
CTGAATCAAAAGAAACATACCTTCGAAAGATATGTGCGCGACGACAACAAAAATCTGCT
TAAGGCAAGCTATCTGGTAAGCCTTCGGATCGCCAAACAGGGAGAAGCCTATACAATTG
CTGAGAAGTTGATTAAACCCTGCACAAAAGACCTGACGACCTGTGTCTTTGGTGAGAAA
TTTGCCTCTAAGGTAGACCTGGTTCCTTTGTCCGACACCACTATCAGCCGACGCATAGA
GGACATGAGTTATTTCTGTGAAGCCGTGCTGGTCAACCGCTTGAAAAACGCCAAGTGCG
GATTCACTTTGCAGATGGACGAGTCCACCGACGTAGCAGGCCTGGCTATTCTTCTCGTC
TTCGTGAGGTACATCCACGAGTCTTCTTTCGAGGAGGATATGCTGTTCTGTAAAGCTCT
GCCTACCCAGACTACCGGGGAAGAAATCTTTAATCTCCTCAATGCCTACTTTGAGAAAC
ACAGCATCCCATGGAACCTCTGCTATCACATTTGCACTGACGGGGCAAAAGCTATGGTT
GGAGTCATCAAAGGGGTGATTGCGAGAATCAAAAAGCTCGTCCCTGACATAAAGGCGAG
CCATTGTTGCCTTCACCGACATGCTCTCGCTGTAAAGCGCATACCGAATGCCCTTCACG
AGGTTCTGAATGACGCAGTCAAGATGATCAATTTTATTAAGTCCCGACCACTCAATGCA
AGAGTGTTTGCCCTGTTGTGCGATGACCTTGGAAGTCTCCATAAAAACCTGCTCCTCCA
CACTGAGGTGAGGTGGTTGTCACGGGGAAAGGTTCTGACTCGATTCTGGGAGCTGCGCG
ATGAGATTAGGATTTTCTTTAACGAGCGAGAATTCGCCGGGAAACTGAACGACACAAGT
TGGTTGCAGAATCTCGCGTACATTGCAGACATCTTCAGCTACCTGAATGAAGTCAACCT
CAGCTTGCAGGGGCCAAATTCCACCATCTTCAAGGTCAACAGCAGGATTAACTCTATCA
AAAGTAAGCTGAAACTGTGGGAGGAATGTATTACGAAAAACAACACTGAATGCTTCGCT
AACCTTAATGACTTCCTTGAGACATCCAATACTGCCTTGGATCCCAATCTTAAGTCTAA
TATCCTGGAGCATCTGAACGGCCTTAAAAACACCTTCCTGGAGTACTTCCCACCTACCT
GCAACAACATCAGCTGGGTGGAAAATCCTTTCAACGAGTGCGGAAACGTTGATACATTG
CCTATTAAAGAAAGGGAACAGCTGATCGATATTAGGACAGATACCACACTGAAGTCCTC
ATTCGTGCCCGACGGCATTGGGCCTTTCTGGATCAAGCTCATGGATGAGTTCCCTGAGA
TCTCCAAGCGCGCTGTGAAAGAACTTATGCCATTTGTGACTACATACCTTTGTGAAAAA
TCTTTTTCTGTATACGTGGCGACTAAGACCAAATACCGGAACCGACTCGACGCCGAAGA
TGATATGCGGCTGCAGTTGACCACAATTCACCCTGATATAGATAATTTGTGTAACAACA
AGCAGGCCCAGAAATCTCACTGA

Table S2. Amino acid substitutions in *SPIN* variants

Variant	Amino acid 220	Amino acid 225	Amino acid 228	Activity
$SPIN_{ON}$	M	R	T	+++++
M8	T	R	A	+
M4	M	R	A	+
M9	T	R	V	—
M6	T	R	M	—
M1	T	R	T	—
M2	M	E	T	—
M3	M	R	M	—
M5	M	R	V	—

Amino acid substitutions made at ambiguous positions in the *SPIN* superconsensus (1).

1. Pace JK 2nd, Gilbert C, Clark MS, Feschotte C (2008) Repeated horizontal transfer of a DNA transposon in mammals and other tetrapods. *Proc Natl Acad Sci USA* 105:17023–17028.

Table S3. Integration datasets

Setname	No. of insertions	Cell type	Transposon/retrovirus type	Source
Burgess-HIV-HeLa	281	HeLa	HIV	PMID:12805549
Burgess-MLV-HeLa	874	HeLa	MLV	PMID:12805549
Miller-AAV-MHF2	432	MHF2	AAV	PMID:16103194
Craig-XL-HeLa-*SPIN*1-ApoI	1540	HeLa	*SPIN*	This work
Craig-XL-HeLa-*SPIN*1-BstYI	1260	HeLa	*SPIN*	This work
Craig-XL-HeLa-*SPIN*2-ApoI	1789	HeLa	*SPIN*	This work
Craig-XL-HeLa-*SPIN*2-BstYI	1386	HeLa	*SPIN*	This work
Craig-XL-HeLa-*SPIN*3-ApoI	1335	HeLa	*SPIN*	This work
Craig-XL-HeLa-*SPIN*3-BstYI	1023	HeLa	*SPIN*	This work
Sunil-HeLa-*PB*1-APOI	3234	HeLa	*piggyBac*	This work
Sunil-HeLa-*PB*1-BSTYI	1788	HeLa	*piggyBac*	This work
Sunil-HeLa-*PB*2-APOI	2262	HeLa	*piggyBac*	This work
Sunil-HeLa-*PB*2-BSTYI	1556	HeLa	*piggyBac*	This work
Sunil-HeLa-*PB*3-APOI	2857	HeLa	*piggyBac*	This work
Sunil-HeLa-*PB*3-BSTYI	1797	HeLa	*piggyBac*	This work
Sunil-HeLa-*SleepingBeauty*-APOI	1954	HeLa	*Sleeping Beauty*	This work
Sunil-HeLa-*SleepingBeauty*-BSTYI	2536	HeLa	*Sleeping Beauty*	This work
Sunil-HeLa-*TcBuster*-APOI	1932	HeLa	*TcBuster*	This work
Sunil-HeLa-*TcBuster*-BSTYI	684	HeLa	*TcBuster*	This work
Sunil-HeLa-*TcBuster*1-Bsty_Apo	3764	HeLa	*TcBuster*	This work

Table S4. Target sites used by *Sleeping Beauty*, *piggyBac*, $TcBuster_{CO}$, and $SPIN_{ON}$ in HeLa cells

Sleeping Beauty		*piggyBac*		$TcBuster_{CO}$		$SPIN_{ON}$	
Total	4,490	Total	13,494	Total	6,380	Total	8,333
TA	4,419	TTAA	13,171	TA	5,973	TA	7,945
CA	16	CTAA	68	TG	136	TG	101
TG	8	TTAG	68	CA	135	CA	93
CT	8	ATAA	21	TT	40	TT	77
AA	8	TTAT	21	AA	37	AA	59
AT	6	TCAA	18	CT	9	TC	13
CG	5	AGTT	16	AC	9	CT	11
GG	4	ATTA	16	GT	8	AG	10
GC	3	GTTA	14	CC	8	GA	7
AG	3	TTGA	14	AG	8	AT	7
TT	2	TTTA	10	TC	6	GT	3
TC	2	TTAC	6	GC	3	CC	3
GA	2	ACTA	4	GA	3	GG	2
CC	2	AGGG	4	GG	2	AC	2
AC	2	CTAG	4	AT	2		
		GTAA	3				
		AGGT	2				
		ATCA	2				
		CTCC	2				
		TAAA	2				
		TCTC	2				
		TGAA	2				
		AAAT	1				
		AATC	1				
		ACAA	1				
		ACAT	1				
		ACTC	1				
		AGTG	1				
		ATAG	1				
		CAAA	1				
		CACA	1				
		CATA	1				
		CCAG	1				
		CCCA	1				
		CGTA	1				
		CTGA	1				
		GTCC	1				
		TAAG	1				
		TCTA	1				
		TGAG	1				
		TGTT	1				
		TTCA	1				
		TTCT	1				
		TTTT	1				

Table S5. ***TcBuster***$_{CO}$ **WT and** ***TcBuster***$_{CO}$ **V596A insertion sites in yeast**

Strain	Chromosome	Target site duplication	Chromosome position
TcBuster$_{CO}$ WT			
1	Chr. X	GTGTAAGC	376143
2	Chr. V	GTTTAAAG	311657
3	Chr. XIV	TTCTATTG	192816
4	Chr. II	CCGTAAAG	100238
5	Chr. VII	CTCTACAT	140482
6	Chr. XV	CTTTAAAA	309140
7	Chr. XI	CCCTAGTT	424352
8	Chr. X	ATCTAGCA	115807
9	Chr. XII	CCCTAAAC	18243
10	Chr. XIII	CTCTAAAT	826982
11	Chr. XVI	GCCTACGC	193241
12	Chr. XV	CTCTAGAT	297092
TcBuster$_{CO}$ V596A			
1	Chr. XII	GTTTAAAC	468808
2	Chr. XVI	TTCTATTG	900530
3	Chr. XII	CTGTAGAG	822006
4	Chr. XV	TGTTAAGG	742495
5	Chr. XIV	CCCTAGCT	230858
6	Chr. VII	ACCTAGAA	490286
7	Chr. VII	ATCTTCTA	490280
8	Chr. X	ATCTAGCA	115807
9	Chr. XIV	CTCTACAC	490201
10	Chr. VII	ATGTAGAG	140276
11	Chr. X	TGCTAGAT	115814
12	Chr. XV	GTGTAGAT	6955