Supplementary Methods

Cell lines

Conditional *Daxx* KO ESCs were generated by CRISPR/Cas9 mediated loxP insertions at introns 1 and 4 in C57BL/6 ESCs. A *Daxx fl/-* (one allele was deleted during loxP insertion by nonhomologous end-joining) ESC line was transiently transfected with a Cre expression plasmid and single colonies were picked to retrieve *Daxx* -/- (KO1) and *Daxx fl/-* (WT1) clones. *Daxx* WT and KO ESCs were transduced with a YFP-tagged H3.3 expression construct using a murine stem cell retroviral vector. A second conditional *Daxx* KO ESC line (WT2/KO2) was generated by CRISPR/Cas9 mediated loxP insertions at introns 1 and 7 in ESCs carrying a 4-Hydroxytamoxifen (4-OHT)-inducible *Rosa Cre-ERT* allele. The parental *Rosa Cre-ERT* ESC line was established from blastocysts derived from B6;129-Gt(ROSA)26Sortm1(cre/ERT)Nat/J mice (Jackson Lab, stock Number: 004847). ESCs were treated with 4-OHT for 48 hours followed by three days of growth in absence of 4-OHT before analysis.

Native ChIP-seq

Native ChIP-seq with *Daxx* WT and KO ESCs carrying a YFP-tagged H3.3 transgene was performed with anti-GFP antibody (A-6455, life technologies) as described previously[2]. ChIP-seq libraries were constructed using the TruSeq ChIP Sample Prep Kit (Illumina) and sequenced as 50 bp single-end reads on a HiSeq2500 machine.

RNA-seq

RNA was harvested from *Daxx fl/fl*; *RosaCreERT* ESCs after 48 hours of tamoxifen treatment followed by three days growth in ESC medium. RNA was purified using the RNeasy Plus mini kit (Qiagen) and strand-specific libraries were generated by the TruSeq RNA Sample Prep Kit v2 (Illumina). Libraries were sequenced on an Illumina HiSeq™ 2000 machine.

Re-analysis of RNA-seq and ChIP-seq data

For analysis of *Daxx* KO RNA-seq data and re-analysis of the data previously published by Karimi *et al.*[2] and Elsasser *et al.*[1], reads were aligned to an extended version of the mouse reference genome (mm9) using JAGuaR v2.1. It uses BWA v0.7.10 to align reads to the genome and a reference transcript model compiled from Ensembl 67 (including annotated exon-exon junctions) specifically allowing the possibility for a single read to span multiple exons. For RNA-seq RPKM normalization, the number of exonic reads were calculated for protein coding genes, excluding mitochondrial and ribosomal genes, as well as the top 0.5% of expressed exons. For re-analysis of published ChIP-seq data[1], reads were aligned to the mouse reference genome (mm9) using BWA v0.7.10. For ChIP-seq RPKM normalization, the total number of uniquely aligned reads was calculated by collapsing the reads having identical coordinates (dups) into a single read and applying mapQ >=5.

For analysis of ERVs, multi-mapped reads from BWA alignment files for both ChIP-seq and RNA-seq were retained, as described in Karimi *et al.*[2] Using the RepeatMasker coordinates for all elements within each ERV family, RPKM expression and ChIP-seq enrichment (over input control) values were calculated for all ERV families with copy numbers above 100. RPKM value for each family was calculated by normalizing agglomerated read coverage of all annotated genomic copies of that family in the reference genome to the agglomerated length of the family, as well as the total number of exonic reads for RNA-seq or uniquely-aligned reads for ChIP-seq.

PCR screening of polymorphic IAP insertions

CD-1 feeder cells were purchased from Stemcell Technologies Inc. 129/SvJ mice were purchased from the Jackson Laboratory. DNA was extracted using Zymo Research Quick-gDNA Miniprep kit following the manufacturer's instructions. PCR was performed in 10 sites using Takara LA Taq (Takara Bio Inc) for long PCR assays (full sites) and with Phusion (NEB) for short ones (empty sites) as per manufacturer's guidelines. Primers are listed in Table S1.

Polymorphic IAP insertion genotyping

Paired-end sequences were obtained from SRA (accessions SRR2014794 and SRR2014795), extracted with sra-tools and aligned to the mouse reference genome mm9 using bwa-mem with default options. Clusters of paired-end reads with one end located in a genomic IAP element based on the mm9 RepeatMasker output were filtered for the presence of a 'switch' in orientation of the 5' and 3' anchoring reads as described in Elsasser *et al.*[1], and additionally filtered against a mappability track obtained from the UCSC Genome Browser mm9 annotation. Clusters with the orientation switch and with an average mappability of greater than or equal to 0.95 were retained. Clusters were annotated against IAP insertions reported in 18 inbred mouse strains[7] and IAP insertions identified in Elsasser *et al.*[1] Scripts, data, and instructions for replicating this analysis can be obtained at https://github.com/adamewing/IAP-Replication.