METHODS

Cell culture and antibodies. K562 cells (ATCC) were grown in Roswell Park Memorial Institute (RPMI) 1640 Medium (11875093, Life Technologies) supplemented with 10% Fetal Bovine Serum (Fisher, Cat# SH30910), 2 mM L-glutamine (Fisher, Cat# SH3003401) and 1% penicillin-streptomycin (Fisher, Cat#SV30010), and cultured at 37 °C with 5% CO_2. HeLa cells (ATCC) were grown in Dulbecco's Modified Eagle's Medium (Life Technologies, Cat# 11995073) supplemented with 10% FBS, 2 mM L-glutamine, and 1% penicillin-streptomycin, and cultured at 37 °C with 5% CO_2. H9 human ES cells were expanded in feeder-free, serum-free medium mTeSR-1 from StemCell technologies, passaged 1:6 every 5–6 days using accutase (Invitrogen) and re-plated on tissue culture dishes coated overnight with growth-factor-reduced matrigel (BD Biosciences). Male mouse embryonic stem cells (R1) were grown as described[28]. Cell cultures were routinely tested and found negative for mycoplasma infection (MycoAlert, Lonza).

Rabbit MORC2 antibody (A300-149A, Bethyl Laboratories), Rabbit MPP8 antibody (16796-1-AP, Protein Technologies Inc), Rabbit TASOR antibody (HPA006735, Atlas Antibodies) were used in Western blots (1:1000 dilution) and ChIP assays. Mouse anti-LINE-1 ORF1p antibody (MABC1152, Millipore)[29], Rabbit HSP90 (C45G5, Cell Signalling, #4877), Beta actin antibody (ab49900, Abcam) were used in Western blots. Histone H3 (tri-methyl K9) antibody (ab8898, Abcam) and RNA Pol II (Santa Cruz Biotechnology, N-20 sc-899) were used in ChIP assays.

L1 reporters. The L1-ORF1-ORF2 sequence is derived from the LRE-GFP[30], a gift from John Moran. To make the L1-GFP reporter, we used Gibson assembly to clone the L1_ORF1/2 fragment and a GFP-B-globin-intron cassette driven by the mammalian promoter EF1a into the pB transgene using a dox inducible promoter (modified from PBQM812A-1, System Biosciences) to drive the L1 sequence and a UBC-RTTA3-ires Blast as a selectable marker for reporter integration. To make the L1-G418^R reporter, we replaced the GFP-B-globin-intron fragment in the L1-GFP reporter with a NEO-intron-NEO cassette driven by the mammalian promoter EF1a. The codon-optimized L1-ORF1-ORF2 sequence in our (opt)-L1 reporter is derived from the SynL1_optORF1_neo, a gift from Astrid Engel[31]. We replaced the self-splicing Tetrahymena NEO-intron-NEO cassette with the neo-B-globin-intron-neo cassette driven by the EF1a promoter or the GFP-B-globin-intron-GFP cassette driven by the EF1a promoter. This L1-syn-ORF1-ORF2-indicator cassette was inserted into the pB transgene using a dox inducible promoter and a UBC-RTTA3-ires Blast, as described above.

Genome-wide screen in K562 cells. The K562 cell line (with a BFP-Cas9 lentiviral transgene) was nucleofected with the pB-tetO-L1-G418^R/Blast construct and the piggyBac transposase (PB210PA-1, System Biosciences) following the manufacturer's instructions (Lonza 2b nucleofector, T-016 program). The nucleofected cells were sorted using limiting dilution in 96-well plates, and positive clones were screened first for sensitivity to Blast, and then the ability to generate G418 resistant cells after dox induction. The Cas9/L1-G418^R cells were lentivirally infected with a genome-wide sgRNA library as described[10], containing ~200,000 sgRNAs targeting 20,549 protein-coding genes and 13,500 negative control sgRNAs at an MOI of 0.3-0.4 (as measured by the mCherry fluorescence from the lentiviral vector), and selected for lentiviral integration using puromycin (1 μg/ml) for 3 days as the cultures were expanded for the screens. In duplicate, 200x10^6 library-infected cells were dox-induced (1 μg/ml) for 10 consecutive days, with a logarithmic growth (500k cells/ml) maintained each day of the dox-induction. After dox-induction, the cells were recovered in normal RPMI complete media for 24 hours, and then split into the G418-selection condition (300 μg/ml G418, Life Technologies, Cat# 11811031) and non-selection conditions. After 7 days of maintaining cells at 500k/ml, 200 M cells under each condition were recovered in normal RPMI media for 24 hours, before they were pelleted by centrifugation for genomic DNA extraction using Qiagen DNA Blood Maxi kit (Cat# 51194) as described[32]. The sgRNA-encoding constructs were PCR-amplified using Agilent Herculase II Fusion DNA Polymerase (Cat# 600675) (See **Table S4** for the primer sequences used). These libraries were then sequenced across two Illumina NextSeq flow cells (~40 M reads per condition; ~200x coverage per library element). Computational analysis of genome-wide screen was performed as previously described[10,11] using CasTLE, which is a maximum likelihood estimator that uses a background of negative control sgRNAs as a null model to estimate gene effect sizes. See **Table S1** for the K562 genome-wide screen results.

Secondary screen in K562 cells. The secondary screen library included the following, non-comprehensive sets of genes (253 genes in total, ~10 sgRNAs per gene, plus 2500 negative control sgRNAs): all genes falling within ~30% FDR from the K562 genome-wide screen (~150 genes), genes known to be functionally related to the 30% FDR genes, genes previously implicated in L1 biology, and genes involved in epigenetic regulation or position effect variegation (see **Table S2** for a complete list). The library oligos were synthesized by Agilent Technologies and cloned into pMCB320 using BstXI/BlpI overhangs after PCR amplification. The Cas9/L1-G418^R (or Cas9/(opt)-L1-G418^R) K562 cell line was lentivirally infected with the secondary library (~4,500 elements) at an MOI of 0.3-0.4 as described previously[33]. After puromycin selection (1 μg/ml for 3 days) and expansion, 40 M (~9,000 coverage per library element) cells were dox-induced for 10 days in replicate, recovered for 1 day, and split for 7-day G418-selection and non-selection conditions, with a logarithmic growth (500k cells/ml) maintained as in the K562 genome-wide screen. 10M cells under each condition were used for genomic extractions, sequenced (~6-10M reads per condition; ~1000-2000x coverage per library element) and analyzed using casTLE as described above[10,11]. See **Table S2** for the K562 secondary screen results with L1-G418^R and (opt)-L1-G418^R.

Genome-wide screen and Secondary screen in HeLa cells. The pB-tetO-L1-G418^R/Blast construct was integrated into Cas9 expressing HeLa cells with piggyBac transposase via nucleofection (Lonza 2b nucleofector, I-013 program) following the manufacturer's instructions. The Cas9/L1-G418^R HeLa cells were blasticidin (10 μg/ml) selected, screened for sensitivity to G418 and the ability to generate G418 resistance cells after dox induction, and lentivirally infected with the genome-wide sgRNA library or with the secondary sgRNA library. Infected cells were then puromycin selected (1 μg/ml) for 5 days and expanded for the screens.

For the genome-wide screen, ~200x10^6 Cas9/L1-G418^R HeLa cells (~1,000x coverage of sgRNA library) were dox-induced for 10 days in replicate, recovered for 1 day, and split for 8-day G418-selection and non-selection conditions, with cells being split every other day to maintain the sgRNA library at a minimum of ~350x coverage. ~200M (1,000x coverage) cells per condition were used for genomic extractions and sequencing as described above for the K562 screens. See **Table S1** for the HeLa genome-wide screen results.

For the secondary screen, ~1x10^7 Cas9/L1-G418^R HeLa cells (~2,000x coverage of sgRNA library) were dox-induced for 10 days in replicate, recovered for 1 day, and split for 8-day G418-selection and non-selection conditions, with cells being split every other day to maintain ~400x coverage. ~5 million (1,000x coverage) cells per condition were used for genomic extractions and sequencing as described above. See **Table S2** for the HeLa secondary screen results.

Validation of individual candidates using the L1-GFP retrotransposition assay. To validate the genome-wide screen hits, we infected clonal Cas9/L1-GFP K562 cells with individual sgRNAs as previously described[32], 3 independent mutant cell lines per gene, each with a different sgRNA (cloned into pMCB320 using BstXI/BlpI overhangs; mU6:sgRNA; EF1a:Puromycin-t2a-mCherry). See **Table S3** for sgRNA sequences. The infected cells were selected against puromycin (1 μg/ml) for 3 days, recovered in fresh RPMI medium for 1 day, and dox-induced for 10 days. Then, the percentage of GFP(+) cells was measured on a BD Accuri C6 Flow Cytometer (GFP fluorescence detected in FL1 using 488 nm laser) after gating for live mCherry(+) cells.

CRISPR-mediated deletion of individual genes and intronic L1s. To delete genes in H9 ESCs, we cloned target sgRNAs in pSpCas9(BB)-2A-GFP (PX458) as described[34]. The sgRNA plasmids were prepared with the Nucleospin plasmid kit (Macherey Nagel) and transfected into H9 ESCs using Fugene following the manufacturer's instructions. After 48-72 hrs, GFP-positive transfected cells were sorted and expanded. Gene depletion effects were validated by western blots.

To delete the L1 from the host gene intron, we designed sgRNAs targeting both upstream and downstream side of the L1 within the intron; one was cloned into pSpCas9(BB)-2A-BFP, while the other into pSpCas9(BB)-2A-GFP. The two sgRNA plasmids were mixed at 1:1 ratio and nucleofected into K562 cells via electroporation following the manufacturer's instructions. After 48-72 hours, BFP/GFP-positive transfected cells were single-cell sorted and expanded. The genetic deletion effects were validated by PCR assay.

Western blotting. Live cells were lysed for 30 min at 4 °C in protein extraction buffer (300 mM NaCl, 100 mM Tris pH 8, 0.2 mM EDTA, 0.1% NP40, 10% glycerol) with protease inhibitors and centrifuged to collect the supernatant lysate. The cell lysate was measured with Bradford reagent (Biorad), separated on SDS-PAGE gels and transferred to nitrocellulose membranes. The L1-reporter containing K562 cells had not been dox-induced when used for western blot assays characterizing endogenous L1_ORF1p levels (Fig. 2d and Extended Data Fig. 4k).

PCR and gel electrophoresis. PCR experiments characterizing the L1-G418^R retrotransposition and the deletion of intronic L1s were performed with Phusion High-Fidelity DNA Polymerase (M0530S, NEB), following the manufacturer's instructions. In general, 30 cycles of PCR reactions were performed at an annealing temperature 5 °C below the Tm of the primer. No 'spliced' PCR products can be detected without dox-induction, even with 40 PCR cycles. PCR reaction products were separated on 1% agarose gels with ethidium bromide. Primer sequences are in **Table S4**.

qRT-PCR and PspGI-assisted qPCR. Total RNA was isolated from live cells using the RNeasy kit (74104, Qiagen) and treated with RNase-Free DNase Set (79254, Qiagen) to remove genomic DNA, according to the manufacturer's instructions. 500 ng total RNA was reverse transcribed with SuperScriptA III First-Strand Synthesis System (18080051, Life Technologies) following the manufacturer's instructions. Beta-actin mRNA was used as internal control within each RNA sample (Figs. 1f and 4d,e). The sequences of PCR primers, including the one targeting the 5'UTR of L1Hs[35–37], are summarized in **Table S4**.