Identification of Spen as a Crucial Factor for Xist Function through Forward Genetic Screening in Haploid Embryonic Stem Cells

Graphical Abstract

Highlights
- A haploid embryonic stem cell screen identifies factors required for Xist function
- The RNA-binding protein Spen is required for gene repression by Xist
- Recruitment of Polycomb group proteins by Xist is affected in the absence of Spen
- Spen binds Xist A-repeat RNA but cannot discriminate functional from mutant motifs

Authors
Asun Monfort, Giulio Di Minin, Andreas Postlmayr, ..., Fabiana Arieti, Stéphane Thore, Anton Wutz

Correspondence
awutz@ethz.ch

In Brief
In mammals, one of the two X chromosomes is inactivated in female cells by the noncoding Xist RNA for dosage compensation. Monfort et al. use haploid embryonic stem cells to identify silencing factors and demonstrate that the RNA-binding protein SPEN is required for gene repression by Xist.
Identification of Spen as a Crucial Factor for Xist Function through Forward Genetic Screening in Haploid Embryonic Stem Cells

Asun Monfort,1 Giulio Di Minin,1 Andreas Postlmayr,1 Remo Freimann,1 Fabiana Arieti,2 Stéphane Thore,3,4 and Anton Wutz1,*
1Institute of Molecular Health Sciences, Swiss Federal Institute of Technology, ETH Hönggerberg, Otto-Stern-Weg 7, 8093 Zurich, Switzerland
2CEITEC–Central European Institute of Technology, Masaryk University, Kamenice 753/5, Brno 62500, Czech Republic
3University of Bordeaux, European Institute for Chemistry and Biology (IECB), ARNA Laboratory, Bordeaux 33000, France
4Institut National de la Sante et de la Recherche Medicale, INSERM, U869, ARNA Laboratory, Bordeaux 33000, France
*Correspondence: awutz@ethz.ch
http://dx.doi.org/10.1016/j.celrep.2015.06.067
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

SUMMARY

In mammals, the noncoding Xist RNA triggers transcriptional silencing of one of the two X chromosomes in female cells. Here, we report a genetic screen for silencing factors in X chromosome inactivation using haploid mouse embryonic stem cells (ESCs) that carry an engineered selectable reporter system. This system was able to identify several candidate factors that are genetically required for chromosomal repression by Xist. Among the list of candidates, we identify the RNA-binding protein Spen, the homolog of split ends. Independent validation through gene deletion in ESCs confirms that Spen is required for gene repression by Xist. However, Spen is not required for Xist RNA localization and the recruitment of chromatin modifications, including Polycomb protein Ezh2. The identification of Spen opens avenues for further investigation into the gene-silencing pathway of Xist and shows the usefulness of haploid ESCs for genetic screening of epigenetic pathways.

INTRODUCTION

In mammals, one of the two X chromosomes in female cells is inactivated for dosage compensation between the sexes (Lyon, 1961); thereby, random chance determines which of the two chromosomes remains transcriptionally active and which is inactivated. X chromosome inactivation has become a paradigm for chromosome-wide gene repression in mammals. The human XIST (Brown et al., 1991a, 1991b) and mouse Xist (Borsani et al., 1991; Brockdorff et al., 1991) genes encode long noncoding RNAs that accumulate over the X chromosome from whence they are transcribed and initiate chromatin modifications and gene silencing. Xist is required for X chromosome inactivation and female mouse development (Marahrens et al., 1997; Penny et al., 1996). Conversely, transgenic expression of Xist is sufficient for initiation of repression in early embryonic cells (Jiang et al., 2013; Lee et al., 1996; Wutz and Jaenisch, 2000).

Studies of Xist have identified a number of factors that contribute to the structure and repression of the inactive X chromosome (Xi). The histone variant macroH2A (Costanzi and Pehrson, 1998) and the nuclear scaffold protein SAF-A/AnrnpU (Helbig and Fackelmayer, 2003) are enriched on the Xi. Polycomb/Trithorax group proteins, including Eed (Wang et al., 2001) and Ash2L (Pullirsch et al., 2010), have been implicated in X inactivation. In addition, multiple functions have been proposed for YY1 in X inactivation, including Xist RNA localization (Jeon and Lee, 2011) and Xist transcription (Makhlof et al., 2014). Whereas Xist is not essential for maintenance of gene repression in somatic cells (Brown and Willard, 1994), loss of Xist in mice has recently been shown to lead to Xi reactivation in blood cells and leukemia (Yildirim et al., 2013). A genetic screen for epigenetic modifiers in mice has identified the SmcHD1 gene, which is required for maintaining DNA methylation and repression of Xi-linked gene promoters (Blewitt et al., 2008). We have previously identified SatB1 as a factor that is required for the initiation of gene repression in a mouse T-cell lymphoma model (Agrelo et al., 2009). However, SatB1 is unlikely a key factor for gene silencing in the mouse embryo, as a mutation is compatible with female development (Nechanitzky et al., 2012). The mechanism that Xist uses for initiation of gene repression remains elusive.

Mouse embryonic stem cells (ESCs) provide a tractable culture system for recapitulating the different steps involved in X inactivation. To advance our understanding of the silencing pathway of Xist, we considered forward genetic screening using recently developed mouse haploid embryonic stem cells (ESCs) (Elling et al., 2011; Leeb and Wutz, 2011). Mutations in a haploid genome are hemizygous and efficiently reveal cellular phenotypes. Screens in haploid cells have been successful in uncovering toxin- and pathogen-related host factors (Jae et al., 2014). The maintenance of pluripotent characteristics of haploid ESCs makes them useful for addressing developmental questions (Leeb et al., 2014) and facilitates the genetic exploration of initiation of X inactivation by Xist.
RESULTS

A Genetic Screen for Silencing Factors

We derived haploid ESC lines from Xist<sup>TX/TX</sup> R26<sup>-SAInlsr-tTA</sup> mice that carry an inducible promoter inserted within the Xist locus and a tetracycline-regulated transactivator targeted to the ROSA 26 locus on chromosome 6 (Savarese et al., 2006) (Figure 1A). We obtained 11 haploid ESC lines (HATX-1 to HATX-11) from a total of 170 activated oocytes (Figure 1B). Addition of doxycycline to HATX cultures induced Xist expression from the single X chromosome and caused cell loss (Figures 1C and 1D). After Xist induction, we observed focal recruitment of the Polycomb protein Ezh2 and the associated histone modification H3K27me3 (Figure 1E). qRT-PCR analysis further showed the repression of X-linked genes, but not of an autosomal β-actin gene as a control (Figure 1F). Flow cytometry profile of HATX3 ESCs before (red) and after (green) infection with gene trap viruses. DNA content (left) and GFP fluorescence (right) of a gene trap encoded reporter are shown. A mixed haploid/diploid DNA content profile is shown for reference in gray.

H3K27me3 (Figure 1E). qRT-PCR analysis further showed the repression of X-linked genes, but not of an autosomal control gene, after Xist induction in HATX ESCs (Figure 1F). These observations are consistent with earlier findings that Xist expression initiates X inactivation in ESCs (Wutz and Jaenisch, 2000; Wutz et al., 2002). In haploid cells, loss of gene expression from the single-X chromosome is not compatible with cell survival. The Xist-induced lethality then provides an efficient selection strategy for recovering mutations abrogating the silencing pathway of Xist and, therefore, preventing Xist-dependent cell death.

For generating a genome-wide set of mutations, we adopted a viral gene trap mutagenesis strategy that has been previously devised for screening in haploid human tumor cells (Carette et al., 2009). We infected 80 million haploid HATX ESCs, which were sorted for a haploid genome content one passage earlier, with a high-titer gene trap vector retrovirus preparation. We confirmed efficient infection by analyzing the fluorescence of a promoter-less EGFP reporter that is encoded by the gene trap virus (Jae et al., 2014) (Figure 1G). The cell pools were subsequently divided into two samples that were either subjected to selection by inducing Xist expression or cultured without doxycycline for obtaining control samples. DNA from both control and selected samples was prepared and used for identification of viral insertion sites through next-generation sequencing (NGS) (Carette et al., 2009). We calculated the number of independent insertions for each gene in selected and control data sets and ranked candidates according to the fold increase in the number of insertions.
in selected samples over controls. As anticipated, *Xist* was found on top of the list of candidates (Tables S1 and S2). Insertions were preferentially observed within the 5'-most region of *Xist* and in an orientation that aligns the poly-adenylation sequences of the gene trap with the transcription unit consistent with the potential termination of *Xist* transcription (Figure 2A). The ROSA26 locus from which the tetracycline responsive transactivator nls-rTTA is expressed was also among the selected candidates. A large number of insertions further mapped within the nls-rTTA cDNA (Figure S1), demonstrating that our screen robustly recovered factors that were anticipated from the technical setup of the screen.

**Figure 2. Mutation of Spen in ESCs**

(A) Schematic representation of viral gene trap insertions in *Xist* and *Spen* locus (not to scale, gene size is indicated). Chr, chromosome.

(B) Schematic representation of the *Spen* gene locus showing the location of CRISPR/Cas9 guide RNAs (gRNAs) used for engineering a deletion and genotyping PCR primers.

(C) Genomic PCR confirming the absence of wild-type *Spen* fragment in *Spen* mutant ESCs.

(D) Quantitative expression analysis of *Spen* using primer sets spanning exons as indicated. Δ*Spen* ESCs lack transcript from the deleted region. Error bars represent SD (n = 3).

(E) Cell survival of control HATX3 ESCs and derived Δ*Spen* ESCs clone 2 (top; n = 2) and clone 3 (bottom; n = 3) after *Xist* induction (+dox). Survival was calculated relative to uninduced cells. Error bars represent SD.

See also Figure S1.

The list of candidates included the chromatin assembly factors Hira and Ubn2, the kinase Cdk8, a subunit of the mediator complex Med25, the RNA-binding protein Spen, and one hypothetical gene Gm3139 (Figure S1; Tables S1–S3). For all candidates except Gm3139, we could observe a strong bias for gene trap insertions in the orientation of the transcription of the gene in selected but not control data sets, which suggests that potential loss-of-function mutations had been enriched. These observations suggest that the selected candidates are involved in *Xist*-mediated gene silencing. However, different mechanisms can be considered: (1) direct interaction with *Xist*, (2) affecting *Xist* expression, and (3) indirect effects from regulating gene expression and signal transduction in pluripotent cells. Indirect effects likely explain the recovery of general regulators of gene expression such as *Hira* and *Med25*. In contrast, *Spen* (also called *Mint* and *Sharp*) possesses RNA-binding domains and has been previously implicated in gene repression (Arieti et al., 2014; Kuroda et al., 2003; Shi et al., 2001), which suggests it as a potential silencing factor of *Xist*. Therefore, we selected *Spen* for further characterization.

**Spen Is Required for Xist-Mediated Gene Repression in ESCs**

*Spen* was enriched in all seven selected samples over controls, indicating that mutations were recovered with high reproducibility. Distribution and orientation of gene trap virus insertions were, furthermore, consistent with a high likelihood of inducing loss-of-function mutations (Figure 2A). To independently validate
Spen as a silencing factor, we used CRISPR/Cas9 nucleases for engineering a deletion within the gene (Figure 2B). We deleted the Spen exons encoding the RNA-binding domains and nuclear localization signals in HATX ESCs and clone 36 ESCs, which carry an inducible Xist transgene on chromosome 11 (Wutz and Jaenisch, 2000). CRISPR/Cas9 nucleases have previously been shown to be very efficient and can yield homozygous mutations (Dou dna and Charpentier, 2014). Our strategy involved the simultaneous use of two guide RNAs for deleting a large fragment of the transcription unit of Spen. Using this strategy, we were successful in obtaining two independent ESC clones with Spen deletions in HATX ESCs and one clone with a homozygous deletion of Spen in clone 36 ESCs. We confirmed the deletion by genomic PCR and Southern blot analysis (Figure 2C; Figures S2A and S2B). qRT-PCR showed an absence of transcript over the deleted region, but the remaining exons appeared to give rise to a transcript of considerably higher abundance than wild-type Spen message (Figure 2D).

Next, we investigated whether mutation of Spen abrogates the ability of Xist to induce X inactivation. For this, we measured cell survival after Xist induction. Whereas induction of Xist caused cell loss in parental HATX ESCs, cells with a deletion of Spen appeared unaffected (Figure 2E). To further confirm that Xist was induced and repression of X-linked genes was abrogated, we performed qRT-PCR analysis. Xist induction was observed 24 and 48 hr after doxycycline addition in Spen mutant and control cells (Figure 3A). Repression of X-linked genes was observed in HATX3 control cells but not in cells carrying a Spen deletion. Autosomal control genes remained unaffected by Xist in both Spen mutant and wild-type cells (Figure 3B). Taken together, these observations showed that a mutation in Spen abrogated the ability of Xist for initiating gene repression. We further analyzed the effect of the Spen deletion in clone 36 ESCs, in which Xist induction represses a puromycin-resistance gene that was co-integrated with the Xist transgene into chromosome 11 and the imprinted Meg1 gene. Xist induction caused efficient repression of the puromycin marker after 48 hr in control cells but less efficient repression in cells lacking Spen (Figure S2C). Consistent with this observation, Spen mutant cells showed an increased survival in the presence of puromycin and less efficient silencing of Meg1 when Xist was induced (Figures S2D–S2F), indicating that loss of Spen also led to a reduced efficiency of Xist-mediated repression in 36 ESCs. Taken together, our results demonstrate that Spen is required for gene repression by Xist in ESCs.

To further explore whether the RNA-binding domains of Spen could bind to Xist, we focused on the A-repeat motif that has been shown previously to be required for initiation of silencing (Wutz et al., 2002). The interaction of three Spen RRM domains with the noncoding steroid receptor RNA activator (SRA) has been recently investigated in detail (Arieti et al., 2014). We performed electrophoretic mobility shift experiments with purified recombinant Spen RRM domains, using the H12-H13 substructure of SRA and a dimer of Xist A-repeat motifs (XCR), which has a similar length as that of SRA, as radiolabeled probe. We found that Spen RRMs can bind XCR (Figure 3C) and that XCR also competes for SRA binding (Figures S2G and S2H). Experiments using in-vitro, T7-transcribed, as well as fully synthetic RNAs indicate that XCR can bind to Spen RRMs similarly to SRA, whereas tRNA does not compete for binding. We also observed that two mutant XCR RNAs (XS1 and XNX) that are not active in initiating gene silencing (Wutz et al., 2002), nonetheless, can compete for binding to Spen RRMs (Figure 3D). These results show that Spen can bind to Xist A-repeat sequences in vitro. However, under our conditions, the binding specificity cannot fully explain the silencing function of Xist A-repeat, potentially reflecting technical limitations of our assay or indicating that other factors cooperate with Spen in binding to Xist A-repeat in vivo.

Mutation of Spen Does Not Affect Xist Localization but Reduces Recruitment of Polycomb Proteins

To further explore whether loss of Spen affected Xist-mediated chromatin modifications, we performed Xist RNA fluorescence in situ hybridization (FISH) and immunofluorescence staining. Xist clusters in control and Spen mutant cells appeared with similar efficiency (Figures S2F, S3A, and S3B), suggesting that Xist localization was not affected by the deletion of Spen. The Polycomb proteins Ezh2 and Ring1b were also recruited by Xist, and focal enrichment of H3K27me3 could be observed in the absence of Spen (Figures 4A and 4C; Figures S3 and S4). Similarly, Xist clusters appeared in a volume that was characterized by low signals of acetylated histone H4 and RNA polymerase II, consistent with the presence of a repressive compartment (Figures 4E and 4F). However, Ezh2 and H3K27me3 foci were less prominent in Spen mutant cells compared to control cells. Combined immunofluorescence staining and RNA FISH showed that the efficiency of Ring1b and Ezh2 recruitment to the Xist cluster was reduced in Spen mutant cells (Figures 4B, 4D, and S4). Taken together, these data demonstrate that loss of Spen resulted in a decreased efficiency of recruitment of chromatin modifications by Xist in ESCs. These results are consistent with previous observations that a mutant Xist RNA lacking A repeat can localize and recruit chromatin modifications but does not initiate gene repression (Wutz et al., 2002). The A-repeat mutant Xist RNA also displays a strong deficit in Polycomb recruitment in ESCs, but recruitment increases upon entry of ESCs into differentiation ( Kohlmaier et al., 2004). However, we did not see an increase in efficiency of recruitment of Ring1b and Ezh2 in Spen mutant cells upon entry in differentiation (Figures 4B and 4D).

DISCUSSION

The mechanism of X chromosome inactivation has been subject to many studies, but the genetic requirements for initiation of gene repression have remained elusive. To advance the genetic basis of the initiation of X inactivation in a developmentally relevant cell system, our study combines haploid ESCs that carry a developmentally validated selection system with saturation mutagenesis as a statistically robust method for the identification of candidate genes.

Our data demonstrate that Spen is required for Xist-mediated silencing in ESCs. Spen possesses four RNA-recognition motif (RRM) domains and is a predominantly nuclear protein. Mutation of Spen in mice causes embryonic lethality (Kuroda et al., 2003).
Figure 3. Spen Is Required for Xist Function in ESCs and Binds Xist A-Repeat Sequences In Vitro

(A and B) qRT-PCR analysis of the X-linked genes (A) and the autosomal control genes (B) as indicated in HATX and Spen mutant ESCs. Error bars represent SD (n = 3). The locations of the genes tested on the X chromosome (chr) is shown at the left. q represents the region of the chromosome. +dox, Xist induction; NT, nontreated.

(C) Electrophoretic mobility shift analysis of Spen RRM domains using a 32P-labeled synthetic XCR RNA probe and cold XCR, SRA, and tRNA competitors as indicated. Asterisk and black triangle indicate position of the RNA-protein complex and free probe, respectively.

(D) Competition with cold XS1 and XNX competitors as in (C).

See also Figure S2.
which might be partly attributable to a repressive function in the Notch signaling pathway (Yabe et al., 2007). Spen is further implicated in nuclear hormone receptor signaling and interacts with the regulatory noncoding SRA RNA through its RRM domains (Arieti et al., 2014; Shi et al., 2001). Its repressor function and the binding of Spen RRM domains to Xist A-repeat are consistent with the expected properties of silencing factors, but we cannot exclude additional factors that regulate the interaction with Xist.

Two recent studies have also identified Spen as an interactor of Xist, using biochemical strategies (Chu et al., 2015; McHugh et al., 2015). Both studies implicate Spen in X inactivation, with slightly different mechanisms. The finding that the Spen protein is detected more abundantly when the Xist A-repeat is present (Chu et al., 2015) is consistent with our observation that Spen RMRMs have binding activity for Xist A-repeat core motifs in vitro. However, our data also suggest that this RNA-binding specificity does not fully explain the silencing function of Xist A-repeat. Notably, depletion of Spen by RNAi has been reported to abrogate Polycomb recruitment by Xist by preventing deacetylation of histones (McHugh et al., 2015). We do observe a reduction, but not a complete block, of Polycomb recruitment in the absence of Spen.

To reconcile the different findings, we consider technical differences between RNAi-mediated depletion and our engineered genetic deletion of Spen exons encoding its RNA-binding domains. Whereas our mutation abrogates gene repression by Xist to a large extent, we observe a moderate and transient reduction of genes close to Xist, including Ftx, Rnf12, and Pkg1 in HATX ΔSpen ESCs (Figure 3A). We find that Xist can recruit Polycomb proteins and form a repressive compartment characterized by low RNA polymerase II and acetylated histone H4 in ΔSpen ESCs, suggesting that multiple interactions between Spen and Xist might exist. The genetic implication of Spen will facilitate further advances on the gene-silencing pathways that operate during X inactivation.

**EXPERIMENTAL PROCEDURES**

**Cell Culture and Generation of Transgenic Cell Lines**

Derivation of haploid ESCs from Xist RX/TX P26 R26nlsrtTA/nlsrtTA mice was performed as previously published (Leeb and Wutz, 2011). For introducing a deletion into the Spen gene locus (Figure 2A) using the CRISPR/Cas system, previously published guide RNA sequences (Koike-Yusa et al., 2014) against Spen exons (gRNA3 Fw: GGGGTGTCTCCTGCGCATT; gRNA5 Fw: CGGACAAGACATACGATC) were inserted into the pX330-U6-Chimeric_BB-CBh-hSpCas9 (Addgene, #42230) vector. Briefly, the vectors were lipofected into HATX3.
cells, together with a GFP reporter vector in a 5:1 ratio. 48 hr later, cells were sorted for green fluorescence and replated to a very low density for isolating individual clones. Spen deletions were identified by PCR on genomic DNA, using the primers 5’-GTCGGTCGATGATATTGCTGC-3’ and 5’-CAGAAGAGGCGACGAGCTTAAAG-3’ (Figure S2A). Fgr2 primers were used as controls (5’-ATTTTGGGACAGACAGG-3’ and 5’-TGGCAACTGCGAGCAGTTT-3’).

Screening in Haploid ESCs
10^3 sorted haploid ESCs were grown in 3 x 145 cm2 dishes for 3 days and harvested. Haploid cells were infected with gene trap virus as previously described (Jae et al., 2014) and plated on 10 cm2 dishes. After 3 days, each dish was split into two 145 cm2 dishes, one used as control and one for selection. For selection, Xist expression was induced by addition of 1 µg/ml doxycycline. Every 3 days, half of the selected cells were passaged and half of the cell pool was used for DNA extraction up to passage 5. Library preparation and computational analysis of sequencing datasets are detailed in the Supplemental Experimental Procedures. The sequencing datasets are deposited in the NCBI Short-Read Archive (http://www.ncbi.nlm.nih.gov/sra) and can be accessed using accession numbers SRX1060416 and SRX1060407.

Immunofluorescence and RNA Analysis
Immunofluorescence staining was performed as previously described (Pullirsch et al., 2010). For further details, see Supplemental Experimental Procedures. For gene expression analysis, RNA was isolated using the QIAshredder (QIAGEN, #79656), purified using the RNeasy Mini Kit (QIAGEN, #74104) and reverse transcribed into cDNA. Real-time PCR was performed using SYBR Green and a LightCycler 480 System (Roche). Elf4a2 was used for normalization. For details, see Supplemental Experimental Procedures.

Electrophoretic Mobility Shift Assay
Binding studies were performed using purified recombinant Spen RRM domain protein and either chemically synthesized or T7 in-vitro-transcribed RNA as previously described (Arieti et al., 2014). For experimental details, see Supplemental Experimental Procedures.

ACCESSION NUMBERS
The sequencing datasets are deposited in the NCBI Short-Read Archive (http://www.ncbi.nlm.nih.gov/sra) and can be accessed using accession numbers SRX1060416 and SRX1060407.

SUPPLEMENTAL INFORMATION
Supplemental Information includes Supplemental Experimental Procedures, four figures, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2015.06.067.

ACKNOWLEDGMENTS
We thank T. Brummelkamp for providing advice on the screen and the gene trap vector; M. Oda, C. Claudio, and M. Kohier for help with reagents and computational analysis; and T. Beyer for critically reading the manuscript. This work was supported by the Swiss National Science Foundation (SNF; grant 31003A_152814/1) and by a Wellcome Trust Senior Research Fellowship (grant reference 087530/Z/08/A), and G.D.M. was supported by the Post-doctoral ETH Fellowship Program. S.T. acknowledges support from the Ligue Genevoise Contre le Cancer, the Worldwide Cancer Research, SNF, and the Canton of Geneva.

Received: March 19, 2015
Revised: May 22, 2015
Accepted: June 23, 2015
Published: July 16, 2015

REFERENCES
Agrelo, R., Souabni, A., Novatchkova, M., Haslinger, C., Leeb, M., Konnenov, V., Kishimoto, H., Gresh, L., Kohwi-Shigematsu, T., Kenner, L., and Wutz, A. (2009). SATB1 defines the developmental context for gene silencing by Xist in lymphoma and embryonic cells. Dev. Cell 16, 507–516.
Arieti, F., Gabus, C., Tambalo, M., Huet, T., Round, A., and Thore, S. (2014). The crystal structure of the Split End protein SHARP adds a new layer of complexity to proteins containing RNA recognition motifs. Nucleic Acids Res. 42, 6742–6752.
Blewitt, M.E., Gentrel, A.V., Pang, Z., Sparrow, D.B., Whitelaw, N., Craig, J.M., Apedaile, A., Hilton, D.J., Dunwoodie, S.L., Brockdorff, N., et al. (2008). SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. Nat. Genet. 40, 663–669.
Borsani, G., Tonlorenzi, R., Simmier, M.C., Dandolo, L., Arnaud, D., Capra, V., Grompe, M., Pizzuti, A., Muzny, D., Lawerree, C., et al. (1991). Characterization of a murine gene expressed from the inactive X chromosome. Nature 351, 325–329.
Brockdorff, N., Ashworth, A., Kay, G.F., Cooper, P., Smith, S., McCabe, V.M., Norris, D.P., Penny, G.D., Patel, D., and Rastan, S. (1991). Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. Nature 351, 329–331.
Brown, C.J., and Willard, H.F. (1994). The human X-inactivation centre is not required for maintenance of X-chromosome inactivation. Nature 368, 154–156.
Brown, C.J., Ballabio, A., Rupert, J.L., Lafreniere, R.G., Grompe, M., Tonlorenzi, R., and Willard, H.F. (1991a). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature 349, 38–44.
Brown, C.J., Lafreniere, R.G., Powers, V.E., Sebasto, G., Ballabio, A., Pettigrew, A.L., Ledbetter, D.H., Levy, E., Craig, I.W., and Willard, H.F. (1991b). Localization of the X inactivation centre on the human X chromosome in Xq13. Nature 349, 82–84.
Carette, J.E., Quinaraes, C.P., Varadarajan, M., Park, A.S., Wuehrich, I., Godarova, A., Kotecki, M., Cochran, B.H., Spooner, E., Ploegh, H.L., and Brummelkamp, T.R. (2009). Haploid genetic screens in human cells identify host factors used by pathogens. Science 326, 1231–1235.
Chu, C., Zhang, Q.C., da Rocha, S.T., Flynn, R.A., Bharadwaj, M., Calabrese, J.M., Magnuson, T., Heard, E., and Chang, H.Y. (2015). Systematic discovery of Xist RNA binding proteins. Cell 161, 404–416.
Costanzi, C., and Pehrson, J.R. (1998). Histone macroH2A1 is concentrated in inactive X chromosome territories through its RGG domain. Nature 393, 599–601.
Doudna, J.A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science 346, 1258096.
Elling, U., Taubenschmid, J., Wimsberger, G., O’Malley, R., Demers, S.P., Vanhaelen, Q., Shukalyuk, A.I., Schmauss, G., Schramek, D., Schnuetgen, F., et al. (2011). Forward and reverse genetics through derivation of haploid mouse embryonic stem cells. Cell Stem Cell 9, 563–574.
Heilig, R., and Fackelmayer, F.O. (2003). Scaffold attachment factor A (SAF-A) is concentrated in inactive X chromosome territories through its RGG domain. Chromosoma 112, 173–182.
Jae, L.T., Raaben, M., Herbert, A.S., Kuehne, A.I., Wirchianski, A.S., Soh, T.K., Stubbs, S.H., Janssen, H., Damme, M., Saflig, P., et al. (2014). Virus entry. Lassa virus entry requires a trigger-induced receptor switch. Science 344, 1506–1510.
Jeon, Y., and Lee, J.T. (2011). YY1 tethers Xist RNA to the inactive X chromosome's nucleation center. Cell 146, 119–133.
Jiang, J., Jing, Y., Cost, G.J., Chiang, J.C., Kolpa, H.J., Cotton, A.M., Carone, D.M., Carone, B.R., Shivak, D.A., Guschin, D.Y., et al. (2013). Translating dosage compensation to trisomy 21. Nature 500, 296–300.
Kohlmaier, A., Savarese, F., Lachner, M., Martens, J., Jenuwein, T., and Wutz, A. (2004). A chromosomal memory triggered by Xist regulates histone methylation in X inactivation. PLoS Biol. 2, E171.
Koike-Yusa, H., Li, Y., Tan, E.P., Velasco-Herrera, Mdel.C., and Yusa, K. (2014). Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. Nat. Biotechnol. 32, 267–273.

Kuroda, K., Han, H., Tani, S., Tanigaki, K., Tun, T., Furukawa, T., Taniguchi, Y., Kurooka, H., Hamada, Y., Toyokuni, S., and Honjo, T. (2003). Regulation of marginal zone B cell development by MINT, a suppressor of Notch/RBP-J signaling pathway. Immunity 18, 301–312.

Lee, J.T., Strauss, W.M., Dausman, J.A., and Jaenisch, R. (1996). A 450 kb transgene displays properties of the mammalian X-inactivation center. Cell 86, 83–94.

Leeb, M., and Wutz, A. (2011). Derivation of haploid embryonic stem cells from mouse embryos. Nature 479, 131–134.

Leeb, M., Dietmann, S., Paramor, M., Niwa, H., and Smith, A. (2014). Genetic exploration of the exit from self-renewal using haploid embryonic stem cells. Cell Stem Cell 14, 385–393.

Lyon, M.F. (1961). Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature 190, 372–373.

Makhlouf, M., Ouimette, J.F., Oldfield, A., Navarro, P., Neuillet, D., and Rouggeule, C. (2014). A prominent and conserved role for YY1 in Xist transcriptional activation. Nat. Commun. 5, 4878.

Marahrens, Y., Panning, B., Dausman, J., Strauss, W., and Jaenisch, R. (1997). Xist-deficient mice are defective in dosage compensation but not spermatogenesis. Genes Dev. 11, 156–166.

McHugh, C.A., Chen, C.K., Chow, A., Surka, C.F., Tran, C., McDonel, P., Pandya-Jones, A., Blanco, M., Burghard, C., Moradian, A., et al. (2015). The Xist IncRNA interacts directly with SHARP to silence transcription through HDAC3. Nature 521, 232–236.

Nechanitzky, R., Davila, A., Savarese, F., Fietze, S., and Grosschedl, R. (2012). Satb1 and Satb2 are dispensable for X chromosome inactivation in mice. Dev. Cell 23, 866–871.

Penny, G.D., Kay, G.F., Sheardown, S.A., Rastan, S., and Brockdorff, N. (1996). Requirement for Xist in X chromosome inactivation. Nature 379, 131–137.

Pullirsch, D., Härtel, R., Kishimoto, H., Leeb, M., Steiner, G., and Wutz, A. (2010). The Trithorax group protein Ash2L and Saf-A are recruited to the inactive X chromosome at the onset of stable X inactivation. Development 137, 935–943.

Savarese, F., Flahndorfer, K., Jaenisch, R., Busslinger, M., and Wutz, A. (2006). Hematopoietic precursor cells transiently reestablish permissiveness for X inactivation. Mol. Cell. Biol. 26, 7167–7177.

Shi, Y., Downes, M., Xie, W., Kao, H.Y., Ordentlich, P., Tsai, C.C., Hon, M., and Evans, R.M. (2001). Sharp, an inducible cofactor that integrates nuclear receptor repression and activation. Genes Dev. 15, 1140–1151.

Wang, J., Mager, J., Chen, Y., Schneider, E., Cross, J.C., Nagy, A., and Magnuson, T. (2001). Imprinted X inactivation maintained by a mouse Polycomb group gene. Nat. Genet. 28, 371–375.

Wutz, A., and Jaenisch, R. (2000). A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. Mol. Cell 5, 695–705.

Wutz, A., Rasmussen, T.P., and Jaenisch, R. (2002). Chromosomal silencing and localization are mediated by different domains of Xist RNA. Nat. Genet. 30, 167–174.

Yabe, D., Fukuda, H., Aoki, M., Yamada, S., Takebayashi, S., Shinkura, R., Yamanoto, N., and Honjo, T. (2007). Generation of a conditional knockout allele for mammalian Spen protein Mint/SHARP. Genes 45, 300–306.

Yildirim, E., Kirby, J.E., Brown, D.E., Mercier, F.E., Sadreyev, R.I., Scadden, D.T., and Lee, J.T. (2013). Xist RNA is a potent suppressor of hematologic cancer in mice. Cell 152, 727–742.
Identification of Spen as a Crucial Factor for Xist Function through Forward Genetic Screening in Haploid Embryonic Stem Cells

Asun Monfort, Giulio Di Minin, Andreas Postlmayr, Remo Freimann, Fabiana Arieti, Stéphane Thore, and Anton Wutz
Selected candidate genes

not enriched during selection

Figure S1
Figure S2
Figure S3

(A) Immunofluorescence images showing Xist RNA, Ezh2, and H3K27me3 in HATX3, HATX3 ΔSPEN clone2, and HATX3 ΔSPEN clone3.

(B) Bar graphs showing the percentage of Xist clusters in HTAX ES cells and ES36.

(C) Bar graphs showing the percentage of Ezh2 foci in HATX ES cells and ES36.

Figure S3
**Figure S4**

### A

|          | 24h RA     |          | 36 ES cells |
|----------|------------|----------|-------------|
|          |            |          |             |
| HATX     | ∆Spen      |          |             |
| clone 2  |            | clone 3  |             |
| Ezh2     | ![Image](image1.png) | ![Image](image2.png) | ![Image](image3.png) |
| Xist     | ![Image](image4.png) | ![Image](image5.png) | ![Image](image6.png) |

### B

|          | 24h RA     |          | 36 ES cells |
|----------|------------|----------|-------------|
|          |            |          |             |
| HATX     | ∆Spen      |          |             |
| clone 2  |            | clone 3  |             |
| Ring1b   | ![Image](image7.png) | ![Image](image8.png) | ![Image](image9.png) |
| Xist     | ![Image](image10.png) | ![Image](image11.png) | ![Image](image12.png) |

### C

|          | 2i         |          | 36 ES cells |
|----------|------------|----------|-------------|
|          |            |          |             |
| Ezh2     | ![Image](image13.png) | ![Image](image14.png) |             |
| Ring1b   | ![Image](image15.png) | ![Image](image16.png) |             |
| Xist     | ![Image](image17.png) | ![Image](image18.png) |             |
Supplemental Figure Legends

**Fig. S1** Schematic representation of viral gene trap insertions in selected candidate genes, related to **Fig. 2** The position and orientation of gene trap insertions is shown over the gene locus (not to scale, gene size is indicated below gene name). The orientation of gene trap insertions is indicated by color. Insertions aligned with the gene transcription unit are shown in green and in antisense orientation in red. Exonic insertions are considered mutagenic in either orientation. A notable bias for sense orientation is observed for all candidates suggesting that predominantly loss of function mutations have been selected. The last line shows insertions in the tetracycline transactivator cDNA, which is protein coding and shows no orientation bias as expected. Above the scheme representing the gene locus insertions recovered from selected cell pools are shown. Below the gene locus insertions in the control cell pools are shown. Intronic insertions in sense orientation are expected to lead to truncation of gene expression and thus have a high mutagenic value. Exonic insertions are considered mutagenic in either orientation. A notable bias for sense orientation is observed for all candidates suggesting that predominantly loss of function mutations have been selected.

**Fig. S2** Analysis of Spen function in Xist mediated gene repression, related to **Fig. 3**

(A) Southern analysis of HindIII digested RNA hybridized with a probe generated from an 800 bp EcoRV fragment of the Spen cDNA containing exon 3 to 6 shows the absence of a 9 + 10 kb double band in ΔSpen mutant cells. (B) Genomic PCR confirming the absence of wild type Spen fragment in Spen mutant ES cells. A genomic Actb primer set is used as a control PCR reaction. (C) Quantitative RT-PCR of puromycin marker gene expression in clone 36 ES cells (wild type) and derived ΔSpen ES cells with (+ dox) and without (no dox) Xist induction for 48 hours. (D) Cell survival of clone 36 ES cells and derived ΔSpen ES cells upon Xist induction in the presence of puromycin. (E) Meg1 expression in wild type (black)
and ΔSpen (white) clone 36 ES cells after 48 hours differentiation in the presence of doxycycline and retinoic acid. (F) Xist RNA FISH in Spen mutant and control clone 36 ES cells after 24 hours of induction with doxycycline. (G) Electrophoretic mobility shift assay (EMSA) using synthetic SRA probe RNA and Spen RRMs. Complex and free probe are indicated by an asterisk and a black triangle, respectively. Competition with cold synthetic SRA RNA is observed. (H) EMSA experiments using T7 in vitro transcribed SRA (above) and XCR RNAs (below) as radiolabelled probes and Spen RRMs. Competition with cold T7 in vitro transcribed SRA and XCR RNAs can be observed for both probes.

**Fig. S3** Characterization of Xist localization chromatin modifications in Spen mutant cells, related to Fig. 4

(A) Xist RNA FISH in Spen mutant and control ES cells after 24 hours of induction with doxycycline. Immunofluorescence staining showing focal recruitment of Ezh2 and H3K27me3. Scale bar represents 10 micrometers. (B) Quantification of Xist expression and localization in Spen mutant and control ES cells. (C) Histogram summarizing Ezh2 recruitment. Cells were counted into three bins: strong focal staining, weak enrichment and absence of apparent focal enrichment (n=100).

**Fig. S4** Analysis of Spen mutant cells, related to Fig. 4

(AB) Combined Ezh2 (A) and Ring1b (B) immunofluorescence and Xist RNA FISH in Spen mutant and control ES cells after 24 hours of differentiation with retinoic acid. (C) Combined Ezh2 and Ring1b immunofluorescence Xist RNA FISH analysis in 36 ES cells shows focal recruitment of Polycomb proteins by Xist in wild type control and Spen mutant cells. Scale bars in all panels represent 5 micrometers.
Table S1. Selected candidate gene list, related to Fig. 2

| Gene name         | Control insertions | Selected insertions | % reads control | % reads selected | fold | p-value |
|-------------------|--------------------|---------------------|-----------------|------------------|------|---------|
|                   | std.               | std.                |                 |                  |      |         |
| Xist*             | 1.29               | 23.43               | 0.000029        | 0.002214         | 18.2 | 0.0009 |
| Gm3139_2          | 0.14               | 2.14                | 0.000000        | 0.000586         | 15.0 | 0.0224 |
| Gt(ROSA)26Sor*    | 0.14               | 2.00                | 0.000000        | 0.000043         | 14.0 | 0.0262 |
| Med25             | 1.14               | 9.43                | 0.000043        | 0.002000         | 8.25 | 0.0022 |
| Hira              | 1.86               | 15.29               | 0.000043        | 0.000600         | 8.23 | 0.0005 |
| Ubn2              | 0.71               | 4.86                | 0.000000        | 0.000143         | 6.80 | 0.0157 |
| Spen              | 2.86               | 18.14               | 0.000057        | 0.002086         | 6.35 | 0.0018 |
| Cdk8              | 2.29               | 11.57               | 0.000043        | 0.027029         | 5.06 | 0.0228 |

The average number of insertions and standard deviation for 7 control and selected cell pools is given for candidate genes (> 5 fold increase in independent insertions; p-value < 0.05). Average percentage of total reads (% reads) for each gene is listed. A fold increase in the average number of insertions in selected relative to control samples and a p-value was calculated. * anticipated hits.

An extended list of genes with additional criteria is provided in Table S3. The computational analysis is detailed in the Supplementary Methods section.
Table S2. Selected candidate genes, related to Fig. 2

Total number of independent virus insertions recovered from NGS libraries for candidate genes in selected and control samples (for each n=7); * indicates anticipated hits.

| Gene name                  | independent viral insertions | control | selected |
|----------------------------|------------------------------|---------|----------|
| nls·rtTA*                  | 19                           |         | 480      |
| Xist*                     | 9                            |         | 164      |
| Gm3139                     | 1                            |         | 15       |
| Gt(ROSA)26Sor*            | 1                            |         | 14       |
| Med25                     | 8                            |         | 66       |
| Hira                      | 13                           |         | 107      |
| Ubn2                      | 5                            |         | 34       |
| Spen                      | 20                           |         | 137      |
| Cdk8                      | 16                           |         | 81       |
| gene         | Number of independent insertion | % of total aligned reads | fold increase of independent insertions | p-value | fold increase of insertions with predicted high mutagenic potential in selected samples | orientation bias in intronic gene traps |
|--------------|---------------------------------|--------------------------|-----------------------------------------|---------|---------------------------------------------------------------------------------|----------------------------------------|
|              | control mean | std | selected mean | std | control | selected |                                   |                                     | control | selected |
| Xist         | 1.29     | 1.38 | 23.43        | 9.98 | 0.000029 | 0.002214 | 18.22                                  | 0.0009                           | 26.00    | 0.33      | 0.50      |
| Gm3139_2     | 0.14     | 0.38 | 2.14         | 1.95 | 0.000000 | 0.000586 | 15.00                                  | 0.0224                           | N/A      | 1.00      |
| Gt(ROSA)26Si | 0.14     | 0.38 | 2.00         | 1.53 | 0.000000 | 0.000043 | 14.00                                  | 0.0262                           | 13.00    | N/A       | 13.00     |
| Med25        | 1.14     | 1.07 | 9.43         | 4.12 | 0.000043 | 0.002000 | 8.25                                   | 0.0022                           | 11.20    | 0.33      | 8.50      |
| Hira         | 1.86     | 1.68 | 15.29        | 4.89 | 0.000043 | 0.000600 | 8.23                                   | 0.0005                           | 10.63    | 1.60      | 6.00      |
| Ubn2         | 0.71     | 0.76 | 4.86         | 2.79 | 0.000000 | 0.00143  | 6.80                                   | 0.0157                           | 10.33    | 1.00      | 7.75      |
| Spen         | 2.86     | 2.48 | 18.14        | 9.34 | 0.000057 | 0.002086 | 6.35                                   | 0.0018                           | 9.14     | 2.00      | 21.00     |
| Cdk8         | 2.29     | 1.50 | 11.57        | 8.30 | 0.000043 | 0.027029 | 5.06                                   | 0.0228                           | 7.00     | 1.67      | 32.50     |
| Ep300        | 0.57     | 0.79 | 2.71         | 2.81 | 0.000000 | 0.00143  | 4.75                                   | 0.0465                           | 5.00     | 0.33      | 0.25      |
| Rbm12        | 1.57     | 1.13 | 7.43         | 3.99 | 0.000029 | 0.000200 | 4.73                                   | 0.0092                           | 5.00     | 1.00      | 2.14      |
| Cpe1         | 1.71     | 1.11 | 7.57         | 4.12 | 0.000029 | 0.00214  | 4.42                                   | 0.0092                           | 4.43     | 1.20      | 2.21      |
| Chadl        | 0.86     | 0.90 | 2.86         | 2.12 | 0.000086 | 0.013086 | 3.33                                   | 0.0327                           | 2.33     | 4.00      | 4.00      |
| Smarcad1     | 10.29    | 6.05 | 33.71        | 17.94| 0.000229 | 0.001314 | 3.28                                   | 0.0128                           | 4.63     | 1.68      | 5.97      |
| Fbxw7        | 7.57     | 2.76 | 24.29        | 15.64| 0.005771 | 0.035457 | 3.21                                   | 0.0459                           | 4.17     | 2.50      | 44.50     |
| Scaf8        | 3.86     | 3.02 | 12.29        | 6.42 | 0.000929 | 0.012457 | 3.19                                   | 0.0382                           | 5.10     | 0.59      | 1.96      |
| Eed          | 5.29     | 3.64 | 16.29        | 7.65 | 0.000186 | 0.000529 | 3.08                                   | 0.0055                           | 4.61     | 1.64      | 12.88     |
| Usp31        | 0.71     | 0.76 | 2.14         | 1.07 | 0.000014 | 0.000043 | 3.00                                   | 0.0465                           | 4.33     | 1.50      | 4.50      |
| Ttl1         | 0.71     | 0.76 | 2.14         | 1.07 | 0.000000 | 0.000157 | 3.00                                   | 0.0353                           | 1.50     | 4.00      | 0.67      |
| Tcf7l1       | 15.29    | 11.13| 41.71        | 19.69| 0.000343 | 0.015800 | 2.73                                   | 0.0265                           | 5.35     | 0.81      | 8.66      |
| Nsd1         | 16.14    | 9.21 | 39.71        | 17.19| 0.005529 | 0.004900 | 2.46                                   | 0.0068                           | 3.45     | 1.30      | 7.00      |
| Kdm5b        | 20.29    | 14.07| 42.43        | 17.48| 0.000471 | 0.001471 | 2.09                                   | 0.0163                           | 3.12     | 1.52      | 13.79     |
| S1c16a2      | 1.00     | 0.58 | 2.00         | 0.82 | 0.000100 | 0.000500 | 2.00                                   | 0.0177                           | 1.33     | 6.00      | 1.33      |
| Cnksr2       | 1.57     | 1.27 | 3.14         | 1.57 | 0.000043 | 0.000086 | 2.00                                   | 0.0171                           | 2.00     | 2.33      | 3.00      |

Legend:
N/A ... not applicable, no insertion of this type occurred
Table S3 (cont): Scores for candidates from the literature in the haploid ES cell screen

**Scores for genes identified in Chu et al. in the haploid ES cell screen**

| Gene   | length (bp) | fold increase in selected samples | Orientation bias of gene traps in intronic insertions selected |
|--------|-------------|-----------------------------------|---------------------------------------------------------------|
|        |             | total independent insertions       |                   | control |               |               |
|        |             |                                  | insertions with high mutagenic potential |               |               |
| Rnf20  | 24827       | 1.00                              | N/A               | 1.00    | N/A            |
| Sap18  | 6792        | 0.20                              | 0.20              | 2.00    | N/A            |
| Wtap   | 25741       | 0.67                              | 0.90              | 0.50    | 3.00           |
| Fus    | 14553       | 1.00                              | 0.50              | 2.00    | 0.33           |
| Rnf2   | 31418       | 0.33                              | 0.50              | 0.40    | 1.00           |
| Myef2  | 37390       | 1.00                              | 1.00              | 0.33    | 0.33           |
| Hnrnpu | 9485        | 0.20                              | 0.25              | 1.00    | N/A            |
| Hnrnpk | 12212       | 0.67                              | 2.00              | N/A     | N/A            |
| Spen   | 70708       | 6.35                              | 9.14              | 2.00    | 21.00          |
| Safb   | 21313       | 0.46                              | 0.33              | 1.50    | 0.67           |
| Pcgf5  | 77138       | 0.76                              | 1.29              | 0.70    | 2.25           |
| Rybp   | 58794       | 0.80                              | 1.00              | 0.88    | 1.40           |
| L1td1  | 11733       | 0.32                              | 0.32              | 1.33    | 1.25           |
| Mybbp1a| 10173       | 0.17                              | 0.33              | 0.33    | N/A            |

**Scores for genes identified in McHugh et al. in the haploid ES cell screen**

| Gene   | length (bp) | fold increase in selected samples | Orientation bias of gene traps in intronic insertions selected |
|--------|-------------|-----------------------------------|---------------------------------------------------------------|
|        |             | total independent insertions       |                   | control |               |               |
|        |             |                                  | insertions with high mutagenic potential |               |               |
| Myef2  | 37390       | 1.00                              | 1.00              | 0.33    | 0.33           |
| Rbm15  | 7180        | 2.00                              | 2.00              | N/A     | 1.00           |
| Ptbp1  | 10004       | 0.25                              | 0.25              | 3.00    | 1.00           |
| Celf1  | 79101       | 0.55                              | 0.38              | 1.86    | 0.83           |
| Hnrnpc | 30649       | 0.83                              | 0.88              | 2.29    | 2.80           |
| Hnrnpm | 39226       | 1.09                              | 0.88              | 2.33    | 1.20           |
| Raly   | 76152       | 1.40                              | 1.50              | 0.67    | 0.75           |
| Eed    | 26323       | 3.08                              | 4.61              | 1.64    | 12.88          |
| Lbr    | 27087       | 1.60                              | 1.50              | 3.00    | 2.00           |
| Spen   | 70708       | 6.35                              | 9.14              | 2.00    | 21.00          |
| Hdac1  | 26543       | 0.63                              | 0.33              | 2.50    | 0.67           |
| Hdac2  | 27345       | 0.38                              | 0.67              | 0.20    | 2.00           |
| Hnrnpu | 9485        | 0.20                              | 0.25              | 1.00    | N/A            |
| Ncor2  | 162062      | 0.27                              | 0.29              | 1.43    | 1.73           |
| Hdac3  | 18018       | 0.20                              | 0.20              | 5.00    | 1.00           |

Legend:
N/A ... not applicable, no insertion of this type occurred
Supplemental Experimental Procedures

Library preparation for insertion site identification

For identification of genomic insertion sites of genetrap viruses next generation sequencing libraries were prepared and sequenced on an Illumina MiSeq instrument. In brief, genomic sequences adjacent to the end of the viral LTR were enriched by linear amplification PCR (LAM-PCR) using a high fidelity DNA polymerase (Invitrogen, #12346-094) with a biotinylated primer as previously published (Carette et al., 2009; Jae et al., 2014). Single stranded LAM-PCR products were captured on Streptavidin coated magnetic particles (Dynabeads M-270, Invitrogen #65305). Subsequently, an oligonucleotide containing the Illumina P7 adaptor sequence was ligated to the 3'-end of the single stranded LAM-PCR fragments using Circligase (Epicentre/Illumina, #CL9025K). A P5 adaptor was added by PCR. The resulting DNA was purified and concentrated using the MinElute PCR Purification Kit (Qiagen, #28004) before loading it on the Illumina MiSeq flow cell. Sequencing was performed using 100 cycles, single-end runs on an Illumina MiSeq sequencer using the MiSeq v2 (Illumina, #MS-102-2002) and MiSeq v3 kits (Illumina, #MS-102-3001).

Computational analysis of NGS datasets

Reads were aligned to the mouse genome (mm10 assembly) using Bowtie2 alignment software (Langmead et al., 2009). Genomic positions corresponding to the 5'-ends of aligned reads were defined as viral integration sites in the genome. Independent insertions were subsequently calculated by collapsing insertions in an interval of 3 bps and excluding insertions supported by less than 2 reads. Using these criteria, we obtained an insertion data table that contains approximately 60,000 independent insertion sites mapped to the mouse genome for each experiment.
Independent insertion sites were then mapped to genes using the HTSeq software package (Anders et al., 2015) and gene annotation of the genome assembly. A ranked list of candidates was obtained by comparing the average number of independent viral insertions per gene in selected relative to control libraries. A total of 7 selected and 7 control samples were used. For each gene a P-value was calculated using the two-sided Fischer exact test. A list of candidates that were considered statistically well supported is provided in Table S1 and the total numbers of insertions in Table S2. We provide an extended list of candidates in Table S3 which also includes additional parameters as follows. The number of mutations that can be predicted to have high mutagenic value was calculated separately as the sum of the number of insertions in exonic regions and the number of intronic insertions in sense orientation. The orientation bias was calculated for gene traps as the ratio between sense and antisense insertions. A high value indicates the preferential occurrence of trapping insertions and an increase in orientation bias in selected over control samples can be used as a further criteria for selecting a candidate. The percentage of total aligned reads in control and selected samples provides a relative measure for the sample size supporting a given gene. A list of these criteria is also provided for the genes identified in Chu et al. and McHugh et al. in Table S3. We do not find strong support for genes in our screen with the exception of Spen. This can likely be explained by two considerations: Firstly, biochemically interacting proteins might not be essential for gene silencing by Xist and therefore not be selected for in our screen. Secondly, mutations in some genes might either occur infrequently or be deleterious to the cells resulting in an insufficient number of insertions for evaluating of the candidate gene.

**Gene expression analysis by quantitative RT-PCR**

Total RNA was isolated from ESCs using the QiAshredder (Qiagen, #79656) and purified using RNeasy Mini Kit (Qiagen, #74104) and on-column Dnase I digestion (Qiagen, #79254). cDNA
for real time PCR was synthesized using the PrimeScript RT Master Mix kit (Takara, #RR036A) from 500 ng of total RNA/10 μl reaction. Real-time quantitative PCR reactions were performed using SYBR Green and a LightCycler 480 System (Roche). Relative expression of the target gene was normalized to *Eif4a2* expression levels. Primer sequences are listed in the Supplemental Methods.

**Primer sequences for gene expression analysis**

Spen: exon1: 5’-GGAAGAGAGATCATCGAGCA -3’, exon2: 5’ TTGGTGACGGAGTTGTGAGC 3’, exon6: 5’-ACCACGACCTCGAAACATCT -3’, exon7-8: 5’- CAAAAACCGCTTGGAGGCGG -3’, exon10: 5’-GCAAATCGGAAAGCCAACTG -3’, exon11: 5’-CTGCACCTCCAGTCTCATGC -3’, exon9: 5’-GGAGACCAAGGGCAGGGAAAA -3’, exon11/12: 5’-TCTTCGTTCTCTCGGCATT -3’, exon15: 5’-GACCCAGGAGGATGTGGTAG -3’, exon16: 5’-CAGGGTGGAAGATCTGAAG -3’; Sdha F: 5’-TTCCGTGTGAGGAGGTGATTGC -3’, Dnmt1 R: 5’-AGGTCTGTGTTCCAACCATTCC- 3’; Xist F: 5’-GCCATCCTCCCTACCTCAGAA -3’, Xist R: 5’-CTGCACATTGTGTTTTTTCCCTAA -3’; Mecp2 F: 5’-CCGGGGACCTATGTATGATG -3’, Mecp2 R: 5’-AGGTCTGTGTTCCAACCATTCC- 3’; Uba1 F: 5’-ATGCACAACCGAATGCAGTA -3’, Uba1 R: 5’-TTGGGTCCAGACCAGAAAG -3’; Usp9x F: 5’-ATGCACAACCGAATGCAGTA -3’, Usp9x R: 5’-CCCTGGAGGAGGGTTTAAGT -3’; Rnf12 F: 5’-TTTGGGTCCAGACCAGAAAG -3’; Rnf12 R: 5’-AGGTCTGTGTTCCAACCATTCC- 3’; Huwe1 F: 5’-TGAGCCTCTGACATTGTTTAAGTTAATGCAGTA -3’, Huwe1 R: 5’-CACAATCTTTTGCTGGAGAT -3’; Rrm2 F: 5’-ATGGGAAGACAACGAAG -3’, Rrm2 R: 5’-CCGGAGCTGGAAAGTAAAGC -3’; ActB F: 5’-CTAAGGCAACCTGGAAAG -3’, ActB R: 5’-GGGAGGTTGTGAGGCTCAAAC -3’; Hprt F: 5’-CAGTCCAGCGTGATTATGGA -3’, Hprt R: 5’-TGGCCTCCCATCTCCTCATC-3’; Pgk1 F: 5’-CCCTTCCGCTACGCAAGTGAAGTGAAG-3’, Pgk1 R: 5’-GATGTGCAATCTCCATGTTGTG-3’; Dnmt1 F: 5’-GTGGACAGTGACACCCCTTT-3’, Dnmt1 R: 5’-CCCTTCCGCTACGCAAGTGAAGTGAAG-3’; Sdha F: 5’-TTCCGTGTGAGGAGGTGATTGC-3’, Sdha R: 5’-AGGTCTGTGTTCCAACCATTCC-3’; Eif4a2 F: 5’-
ACACCATCGGGGTCCATTCC-3’, Eif4a2 R: 5’-CCTGTCTTTTCAGTCGGGC-3’; Meg1(Grb10) F: 5’-CAACCAGGGACGACATCTA -3’, Meg1(Grb10) R: 5’- ACTGCTGGTCTTCTCCTGA -3’.

Electrophoretic mobility shift assay (EMSA)

For experiments using chemically synthesized RNA 50 pmol of synthetic RNA oligonucleotides (Microsynth AG, sequences are detailed in the Supplemental Methods) were labeled using T4 Polynucleotide Kinase (NEB) and 5 μl γ-[32P]-ATP (6000 Ci/mmol; 10 mCi/ml; Hartmann Analytic). Probes were purified using MicroSpin G-25 columns (GE Healthcare). 0.1 pmol of the radioactive RNA and 2.5 μM of the Spen RRMs were incubated during 20 min at room temperature in the binding buffer containing 20 nM HEPES (pH 7.5) with 5 mM MgCl2, 5% Glycerol, 1 mM ß-mercaptoethanol, 150 mM NaCl and 0.5 μl of Superase (Invitrogen) as described (Arieti et al., 2014). Unlabeled RNAs were added in 20x (2 pmol), 40x (4 pmol), 100x (10 pmol) excess in competition experiments. Samples were mixed with 6x loading dye (50% glycerol, 0.2% bromophenol) and run on 10% native polyacrylamide gels in 1X Tris-borate-EDTA buffer at 90V for 90 minutes at 4°C and visualized using autoradiography.

For experiments using in vitro transcribed RNAs oligonucleotides corresponding to XCR-XC dimer, XS1, XNX mutant dimers and H12H13 SRA were cloned into the SacI and KpnI sites of the pBluescript II KS(-) vector (Fermentas). The insert including the T7 promoter was amplified by PCR using M13 primers Fw: 5’ GTAAAACGACGG CCAG 3’ and RV: 5’ CAGGAAACAGCTATGAC 3’. PCR products were checked by gel electrophoresis for purity, and phenol/chloroform extracted. 1 μg of PCR product was used as template for in vitro RNA synthesis using the MAXiScript T7 transcription kit (Ambion) following the recommendations of the supplier. For radiolabelling RNA probes UTP was substituted by 5 μl alpha-[32P]-UTP (3000 Ci/mmol; 10 mCi/ml; Hartmann Analytic) per reaction following the recommendation of the supplier. In vitro transcription products were treated with DNase and purified by
phenol/chloroform extraction. For each reaction 0.05 pmol of the RNAs were incubated with 2.5 μM of the Spen RRM s in the conditions described. For the competition assays increasing concentrations of the different in vitro synthesized cold competitors were added, 20x (1 pmol), 40x (2 pmol), 100x (5 pmol). Samples were run on a 10% native polyacrylamide gel as described above.

RNA sequences: H12H13 SRA: 5’-GCAGGAUGUGAUGACAGCCGACGGCCUGACUGCUUCAUACCUGCUUUAAAUUGCCCAUCGGAUACCUGCUUUAC
CAGGAACAGUGGGCUGGAGGAAAGUUGUCAAUACCUGU

Ac-3’; XCR-Xc dimer: 5’-UGGCCAUCG GGGCCUCGGAUACCUCGUUAAAUUGCCAUCGGGCGCCUCGGAUACCUCGUUUAC-3’; XSI dimer: 5’-UGGCCAUCG GGGCCUCGGAUACCUCGUUAAAUUGCCAUCGGGCGCCUCGGAUACCUCGUUU AC-3’; XNX dimer: 5’-UUGC GCAUCG GAGCCUCCGGAUACCUCGUUAAAUUGCCAUCGG GAGCCUCCGGAUACCUCGUUUAC-3’

Immunofluorescence analysis

For immunofluorescence staining cells were fixed with 4% PFA essentially as previously described (Leeb et al., 2010). Briefly, primary H3K27m3 (Active Motif, #39155) and Ezh2 antibodies (Cell Signaling, #3147) were used at an 1:200 and 1:100 dilution, respectively in PBS/2.5%BSA/0.1%Tween. FITC-labeled secondary antibodies (Jackson Immuno Research) were used at an 1:500 dilution. Nuclei were counterstained with a solution of 1.4 μM DAPI (Molecular Probes, #D1306). After several washes in PBS/0.1% Tween slides were mounted using Vectashield mounting medium for fluorescence (Reactolab, #H-1000). RNA FISH was performed as previously published (Pullirsch et al., 2010).

Combined immunofluorescence and Xist RNA FISH

Multiwell Roboz slides (Cellpoint Scientific) were coated over night with laminin (5μg/ml in PBS). 10^4 cells were seeded per well and grown for 24h. Medium was aspirated, slides rinsed once in PBS and samples fixed in 4% paraformaldehyde in PBS for 10 minutes at room
temperature. After 2 washes in PBS, cells were permeabilized for 5 min at room temperature in 0.1% Sodium Citrate/0.5% TritonX-100, washed twice in PBS/0.1% Tween 20 (PBST) and blocked 30 minutes in PBST+2.5% BSA. Slides were incubated with primary antibodies diluted in PBST containing 2.5% BSA overnight at 4°C. After 3 washes in PBST for 10 minutes each, slides were incubated with secondary antibody 1:500 in PBST for 1 hour at room temperature in the dark. After 3 washes in PBST slides were postfixed for 10 minutes in 4% Paraformaldehyde. For subsequent RNA FISH slides were dehydrated in 70%, 80%, 95% and 100% ethanol series and air dried. 3μl of Cy3-labeled Xist RNA FISH probe generated by random priming (Prime-it-II, Stratagen) and resuspended 1:3 in Hybrisol VII (MP Biomedicals) was applied to each well. After overnight incubation at 37°C in a light-protected humified chamber slides were washed for 5 minutes each 3 times in 2xSSC/0.1% Tween20, 3 times in 2xSSC at 39°C, once 10 minutes at room temperature in 1xSSC, and finally once 5 minutes in 4xSSC. DNA was counterstained by incubation in 4xSSC/0.1% Tween20 + DAPI [1.4 μM] at room temperature. Slides were washed once in 4xSSC and mounted using Vectashield H-1000 (Vector Labs). Samples were analysed on a Zeiss Axio Observer Z.1 fluorescence microscope equipped with a Hamamatsu OrcaFlash4.0 Camera and a Plan Apochromat 100x/1.46 oil DIC objective. Images were processed using ZEISS Zen Software installed on the instrument and figures were prepared using Adobe Photoshop to adjust brightness and contrast.

Antibodies were as follows:

Anti Ring1B (D22F2): Cell Signaling #5694 XP, rabbit monoclonal antibody 1:500

Anti H4ac: Merck Millipore #06-866, rabbit polyclonal antibody 1:500

Anti Pol II (8WG16): Covance #MMS-126R, mouse monoclonal antibody 1:50

Anti Ezh2 (AC22): Cell signaling #3147, mouse monoclonal antibody 1:100

Anti H3K27me3: Active Motif #39155, rabbit polyclonal antibody 1:200
Fluorophore coupled secondary antibodies:

Jackson Lab #715-545-150 Alexa Fluor 488 AffiniPure Donkey anti mouse IgG (H+L) 1:500

Jackson Lab #715-545-152 Alexa Fluor 488 AffiniPure Donkey anti rabbit IgG (H+L) 1:500
Supplemental References

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166-169.

Arieti, F., Gabus, C., Tambalo, M., Huet, T., Round, A., and Thore, S. (2014). The crystal structure of the Split End protein SHARP adds a new layer of complexity to proteins containing RNA recognition motifs. Nucleic acids research 42, 6742-6752.

Carette, J.E., Guimaraes, C.P., Varadarajan, M., Park, A.S., Wuethrich, I., Godarova, A., Kotecki, M., Cochran, B.H., Spooner, E., Ploegh, H.L., et al. (2009). Haploid genetic screens in human cells identify host factors used by pathogens. Science 326, 1231-1235.

Jae, L.T., Raaben, M., Herbert, A.S., Kuehne, A.I., Wirchnianski, A.S., Soh, T.K., Stubbs, S.H., Janssen, H., Damme, M., Saftig, P., et al. (2014). Lassa virus entry requires a trigger-induced receptor switch. Science 344, 1506-1510.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology 10, R25.

Leeb, M., Pasini, D., Novatchkova, M., Jaritz, M., Helin, K., and Wutz, A. (2010). Polycomb complexes act redundantly to repress genomic repeats and genes. Genes & development 24, 265-276.

Pullirsch, D., Hartel, R., Kishimoto, H., Leeb, M., Steiner, G., and Wutz, A. (2010). The Trithorax group protein Ash2l and Saf-A are recruited to the inactive X chromosome at the onset of stable X inactivation. Development 137, 935-943.