Published online 20 February 2019

Nucleic Acids Research, 2019, Vol. 47, No. 8 ***3875–3887***
doi: 10.1093/nar/gkz109

En bloc and segmental deletions of human *XIST* reveal X chromosome inactivation-involving RNA elements

Hyeon J. Lee[1,†], Ramu Gopalappa[2,†], Hongjae Sunwoo[3,†], Seo-Won Choi[1], Suresh Ramakrishna[4,5], Jeannie T. Lee[3], Hyongbum H. Kim[2,6,7,8,*] and Jin-Wu Nam[1,9,*]

[1]Department of Life Science, College of Natural Sciences, Hanyang University, Seoul 04763, Republic of Korea, [2]Department of Pharmacology, Yonsei University College of Medicine, Seoul 03722, Republic of Korea, [3]Department of Molecular Biology, Massachusetts General Hospital, Department of Genetics, Harvard Medical School, Boston MA 02114, USA, [4]Graduate School of Biomedical Science and Engineering, Hanyang University, Seoul 04763, Republic of Korea, [5]College of Medicine, Hanyang University, Seoul 04763, Republic of Korea, [6]Brain Korea 21 Plus Project for Medical Sciences, Yonsei University College of Medicine, Seoul 03722, Republic of Korea, [7]Severance Biomedical Science Institute, Yonsei University College of Medicine, Seoul 03722, Republic of Korea, [8]Center for Nanomedicine, Institute for Basic Science (IBS), Seoul 34126, Republic of Korea and [9]Research Institute for Convergence of Basic Sciences, Hanyang University, Seoul 04763, Republic of Korea

Received September 11, 2018; Revised February 07, 2019; Editorial Decision February 09, 2019; Accepted February 13, 2019

ABSTRACT

The *XIST* RNA is a non-coding RNA that induces X chromosome inactivation (XCI). Unlike the mouse *Xist* RNA, how the human *XIST* RNA controls XCI in female cells is less well characterized, and its functional motifs remain unclear. To systematically decipher the XCI-involving elements of *XIST* RNA, 11 smaller *XIST* segments, including repeats A, D and E; human-specific repeat elements; the promoter; and non-repetitive exons, as well as the entire *XIST* gene, were homozygously deleted in K562 cells using the Cas9 nuclease and paired guide RNAs at high efficiencies, followed by high-throughput RNA sequencing and RNA fluorescence *in situ* hybridization experiments. Clones containing *en bloc* and promoter deletions that consistently displayed no *XIST* RNAs and a global up-regulation of X-linked genes confirmed that the deletion of *XIST* reactivates the inactive X chromosome. Systematic analyses of segmental deletions delineated that exon 5 harboring the non-repeat element is important for X-inactivation maintenance, whereas exons 2, 3 and 4 as well as the other repeats in exon 1 are less important, a different situation from that of mouse *Xist*. This Cas9-assisted dissection of *XIST* allowed us to understand the unique functional domains within the human *XIST* RNA.

INTRODUCTION

Xist is a functionally well-known nuclear long non-coding RNA (lncRNA), the role of which has been studied through its ablation that leads to female-specific lethality in the early embryonic development of mice (1–6). Murine *Xist* encodes a spliced transcript comprising seven exons; the first and last exons contain embedded repeat elements (A–E), whereas the internal exons appear more conserved than others. Xist is specifically localized on the inactive X chromosome (Xi) forming an Xist cloud (7,8), and its expression is sufficient for X chromosome inactivation (XCI), the dosage compensation mechanism of the two X chromosomes during female embryo development in mice (9). A series of deletions of the *Xist* gene demonstrated that repeat A is the most critical element for gene silencing on the X chromosome (10), and repeat E contributes to Xist localization to the Xi (11,12).

The human homolog, *XIST*, encodes an ∼19 kb-long spliced and polyadenylated lncRNA (13). Whereas only 10 or so genes escape from XCI in mice, 100s of genes are expressed from the Xi in humans (14), suggesting that human XCI might be less strictly enforced. This epigenetic instability in human XCI was shown using inducible XIST in human cells (15), raising the possibility that modulation of XIST function from the endogenous locus might reactivate some Xi genes. In addition, XIST-induced chromatin modification status can vary depending on the marker in different human cell lines; in HEK293 cells, H3K27me3 and H4K20me1 appear visibly enriched upon XIST induction. In contrast, enrichment of H3K27me3 was not detected in HT1080 cells (15).

*To whom correspondence should be addressed. Tel: +82 2 2220 2428; Fax: +82 2 2298 0319; Email: jwnam@hanyang.ac.kr
Correspondence may also be addressed to Hyongbum Henry Kim. Tel: +82 2 2228 0879; Fax: +82 2 313 1894; Email: hkim1@yuhs.ac
†The authors wish it to be known that, in their opinion, the first three authors should be regarded as Joint First Authors.

Although deletions of the identified human *XIST* repeat A element were carried out to monitor the expression changes of X-linked genes (15,16), the introduction of segmented *XIST* constructs with a reporter system showed expression changes of proximal reporter genes and only a few distant genes but no other X-linked genes. Furthermore, the ectopic expression of partial *XIST* constructs could be hindered by the expression of endogenous *XIST*, suggesting that segmental deletions of the endogenous *XIST* gene are necessary for unambiguously identifying XCI-related elements. Recently, paired guide RNAs (gRNAs) and the CRISPR/Cas9 nuclease system have been used to efficiently delete genomic regions (17,18) including lncRNA genes (19,20). One study deleting the *XIST* repeat D element by using a single guide RNA (sgRNA) and Cas9 claimed that the deletion of repeat D led to a reactivation of X-linked genes (21). However, the validation for reactivation was carried out for only six X-linked genes and not for global X-linked genes, indicating that global changes of all X-linked genes must be examined with systematic deletions of the endogenous *XIST* gene.

In this study, we designed sgRNA pairs for *en bloc* and 11 segmental deletions of the human *XIST* gene. The analyses of RNA fluorescent *in situ* hybridization (FISH) and RNA-seq on clones containing segmental and whole deletions (WD) unbiasedly revealed functionally important RNA domains of the *XIST* RNA. Although the repeat elements in exon 1 appeared to be less important or not required for XCI maintenance, deletions of the repeats resulted in a decrease in *XIST* RNA levels that might directly mediate a slight elevation in the expression of X-linked genes. Instead, the conserved exon 5 that includes the non-repeat element appeared to be important for X-inactivation maintenance. Repeat E in exon 6 appears important for proper maintenance of the XIST cloud. Our results may provide useful guidelines for the design of sgRNA pairs for the efficient knockout of other lncRNA genes as well as providing a further understanding of the functional mechanism of XIST-induced XCI in human cells.

MATERIALS AND METHODS

Selection of paired sgRNAs for gene knockout

To design sgRNAs for deleting whole and segmental loci of genes, all candidate sgRNA target sites with a protospacer adjacent motif (PAM; 5′-NGG-3′) near the *XIST* were initially identified. For efficient deletion, all sgRNAs for the candidate target sites were evaluated with our scoring system (LINDEL score; see Supplementary Data for more details). sgRNAs with high efficiency, satisfying the following criteria: a Shannon entropy $\geq$0.7, a free energy greater than $-$4 kcal/mol and a LINDEL sgRNA score $\geq$0.3, were selected. Among the potentially efficient sgRNAs, we chose sgRNAs with a low probability of off-target effects (up to six mismatches using a genome-wide alignment and a previously reported off-target prediction program (22)). For each targeted loci, two sets of paired sgRNAs were randomly evaluated and a pair of sgRNAs with a high indel efficiency were finally selected using the T7E1 assay (see Supplementary Data for more details) on genomic DNA isolated from K562 cells transfected with plasmids encoding Cas9 and sgRNAs (Supplementary Table S1). The cells were treated with puromycin to enrich for transfected cells before analysis. The sgRNA pairs that led to higher mutation frequencies, ranging from 21 to 61% were selected for subsequent knockout studies.

Transfection of plasmids encoding Cas9 and sgRNAs

Cas9 and sgRNAs were expressed using the CMV promoter-driven Cas9-2A-mRFP-2A-Puro plasmid (hereafter, Cas9-puro vector) and the hU6-sgRNA plasmid, respectively. The plasmids were purchased from Toolgen (Seoul, South Korea); the full sequences of these vectors are described in Supplementary Data and the vector maps are shown in Supplementary Figure S1b and c. Oligonucleotides containing each target sequences were synthesized (Bioneer, Seoul, South Korea) and annealed *in vitro* using a thermocycler. The sgRNA-encoding vector (pRG2-CT), which was purchased from Toolgen, was digested with BsaI and ligated with annealed oligonucleotides. Oligonucleotide sequences are listed in Supplementary Table S2. Cells were transfected with plasmid mixtures containing Cas9-puro, U6-sgRNA encoding one member of an sgRNA pair and U6-sgRNA encoding the other member of the sgRNA pair at a weight ratio of 1:1:1 using Neon (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. One day after transfection, puromycin (Cat no. A11138-03, Invitrogen, Carlsbad, CA, USA) was added to the culture media at a final concentration of 2.5 μg ml^{-1} and cells were cultured for two days. Three days after transfection, the cells were analyzed.

Cell culture

K562, a human erythromyeloblastoid leukemia cell line was purchased from the American Type Culture Collection (Manassas, VA, USA) and cultured in Roswell Park Memorial Institute medium (Invitrogen, Carlsbad, CA, USA), supplemented with 100 units ml^{-1} penicillin, 100 μg ml^{-1} streptomycin and 10% fetal bovine serum.

Analysis of single cell-derived clones

To obtain single cell-derived clones containing the deletion, we plated the transfected cells into 96-well plates at an average density of 0.25 cells/well. Fourteen days after plating, individual clones were isolated and analyzed using gel electrophoresis of polymerase chain reaction (PCR) products amplified from genomic DNA to check if the clones have a deletion and the wild-type allele. We next sequenced the PCR products derived from the clones containing targeted deletions to check if the two cleavage sites were joined with the generation of small insertions or deletions (indels) in all the clones.

Sequencing analysis

Sequencing of genomic regions including the target sequence was performed as previously described (23). Briefly, PCR amplicons that included the junction regions of the deleted lncRNA target sites were cloned into the T-Blunt

vector (Promega, Madison, WI, USA) and sequenced using the primers used for PCR amplification (Supplementary Table S3).

qRT-PCR

Total RNA was extracted from cells using TRIzol (Ambion by Life Technology) or an RNeasy Kit (QIAGEN), after which complementary DNA (cDNA) synthesis was performed using a DiaStarTM RT Kit (SolGent Co., Ltd. Korea). The synthesized cDNA was subjected to quantitative PCR (qPCR) in triplicate using an Applied Biosystems StepOnePlus Real Time PCR System with PowerSYBR Green PCR Master Mix (Applied Biosystems). Gene expression was normalized to that of the wild-type *XIST* gene (K562). Error bars represent the standard deviation (s.d.) of the mean of triplicate reactions. Primer sequences for qPCR are listed in Supplementary Table S4.

RNA FISH

K562 parental cells and *XIST* knockout cells were grown in suspension. 50 000 cells were cytospun to a slide. After a brief period of air drying, cells were rehydrated in phosphate buffered saline (PBS), fixed in 4% paraformaldehyde for 10 min and permeablized in PBS/0.2% Triton X-100 for 3 min at room temperature. Cells were washed twice in PBS and dehydrated through a series of 70, 90, 100% ethanol. Hybridization buffer containing 2× SSC, 25% formamide, 10% dextran sulfate, 1 mg/ml yeast transfer RNA and fluorescently labeled oligo probes (1–2.5 nM of each oligo) was then applied to the cells. XIST exon 1 probes were used for the SD2, SD4, SD6, SD7 and SD8 clones. XIST exon 6 probes were used for the SD1, SD3, SD5, SD9, SD10 and SD11 clones. When RNA FISH of an X-linked gene (*MAGEB2* or *PIR*) was performed, a whole XIST probe was used. Hybridization was performed in a humidified chamber for 3–4 h at 42°C. Cells were washed in 2× SSC three times and nuclei were stained with Hoechst 33342. Imaging was performed on a Nikon 90i microscope equipped with a 60×, 1.4 N.A. VC objective lens, an Orca ER camera (Hamamatsu) and Volocity software (Perkin Elmer).

Oligo probes recognizing *XIST*, *NEAT1* and the two X-linked genes *MAGEB2* and *PIR* were designed using the Primer3 website and manually selected. Amine-ddUTP was added to pooled single-stranded DNA oligos (synthesized by Integrated DNA Technologies, Coralville, IA, USA) by terminal transferase (New England Biolabs, Ipswich, MA, USA). After phenol extraction and ethanol precipitation, oligos were resuspended in 0.1 M sodium borate. Dye-conjucation was perfomed with NHS-Cy3B (GE Healthcare) or Alexa647 (Thermo Fisher Scientific) at room temperature overnight. Labeled oligos were ethanol precipitated, resuspended in DEPC-H2O and checked by Nanodrop.

RNA-seq data analysis

Strand-specific libraries of poly(A)-selected RNAs from *XIST* knockout and parental cells were constructed using the Truseq stranded messenger RNA (mRNA) prep kit (Illumina, San Diego, CA, USA), after which RNA-seq was performed at Macrogen (Macrogen, Seoul, South Korea). RNA-seq reads were mapped to the human reference genome (hg19) using TopHat version 2.0.6 with mapping parameters '–segment-length 50 –max-multihits 1 –library-type fr-firststrand –mate-inner-dist 125 –mate-std-dev 125 –read-realign-edit-dist 0 –min-intron-length 61 –max-intron-length 265006 –max-segment-intron 265006 –min-segment-intron 61'. To examine gene expression changes between parental (based on GENCODE annotation version 19) and knockout cells (based on the resulting transcript models), the logarithm base 2 of reads per million (RPMs) or reads per kilobase of exons per million (RPKM) for each gene were calculated using an in-house python script. For the mutated XIST gene, the RPKM was calculated with the length of the mutated *XIST* RNA. The genes with <1 RPM in the parental cell line were filtered out and RPM values of <1 in the knockout cells were replaced as 1. Using the junctions reported by TopHat, the splicing patterns were predicted with the number of alignments spanning the junctions.

Reactivation of X chromosomes

The degree of reactivation of the *XIST* knockout cells were calculated by summing up the positive fold-changes of genes on the p-arm of chromosome X, using chromosome 10 as control. Genes with a fold-change above 1.5 were scored as up-regulated genes.

Whole genome sequencing and variant calling

Genomic DNA was extracted from parental K562 cells using the Wizard DNA extraction kit (24) and prepared for DNA sequencing library construction. A paired-end library with inserts of ∼300 bp was constructed using protocols from Illumina (Illumina, San Diego, CA, USA). The constructed library was sequenced using the Illumina Hiseq2000 platform. After whole genome sequencing (WGS) of the parental cells was performed, the reads were mapped using BWA version 0.7.10-r789, duplicates were marked with Picard version 1.96, realigned based on the presence of an insertion or deletion and the quality scores recalibrated using GATK version 3.4. The variants were called using GATK HaplotypeCaller version 3.4.

Allele-specific expression analysis

Among the called single nucleotide polymorphisms (SNPs), only heterozygous SNPs were filtered by allele frequencies from WGS data. The regions of the haplotypes were predicted by using the '–bamOutput' option of GATK HaplotypeCaller. For each position of the filtered SNPs, the expected minor allele frequencies (MAFs) of the *XIST* knockout and parental cells were calculated by the number of reads from the RNA-seq. If the multiple SNPs are located within the same haplotype block, the average of the MAFs were calculated. The difference in MAFs was also calculated by subtracting the MAFs of parental cells from the MAFs of *XIST* knockout cells. Based on the MAFs of the

XIST knockout and parental cells, the genes having haplotype blocks that changed from homozygous to heterozygous were selected.

RESULTS

En bloc deletion of the *XIST* gene

We first attempted to delete *XIST* in its entirety (including transcription factor binding sites (TFBSs)) by inducing paired DNA double-strand breaks using the CRISPR-Cas9 system in human K562 cells (Figure 1A), which are female diploid cells containing two X chromosomes (25) (Supplementary Figure S1a). After transfection with plasmids encoding Cas9 (Supplementary Figure S1b) and the sgRNA pairs selected using the T7E1 assay (Supplementary Figure S1c and d; see 'Selection of paired sgRNAs for gene knockout' in the 'Materials and Methods' section for more details), isolated genomic DNA was subjected to a PCR to detect the targeted deletion. Gel electrophoresis revealed PCR amplicons of ∼1550 bp (Figure 1B, *XIST*), which were not observed in samples from untransfected cells, suggesting that the two target sites were joined after the intervening regions were deleted.

To obtain single cell-derived clones containing the deletion, we plated the transfected cells into 96-well plates, after which individual clones of the transfected cells were isolated and analyzed (see 'Analysis of single cell-derived clones' in the 'Materials and Methods' section). Gel electrophoresis of PCR products amplified from genomic DNA showed that 5 out of 16 randomly examined clones (31%) contained targeted deletions (Supplementary Figure S1e). However, due to incomplete survival of cells during growth to a clonal population, the deletion efficiency could be overestimated. Among the five clones containing the deletions, two of the *XIST* knockouts (clones 10 and 13, $XIST^{-/-}$) did not contain the wild-type allele (Figure 1C), indicating biallelic deletions in these clones. We next sequenced the PCR products derived from the clones containing targeted deletions. The Sanger sequencing results confirmed that the two targeted sites were joined (Supplementary Figure S1f), corroborating that the intervening segment of 30.9 kb was deleted from *XIST* with small insertions or deletions (indels) at the junctions in all of the clones (Supplementary Figure S1f).

To examine the effects of a complete *XIST* deletion, we conducted RNA FISH using a probe recognizing exon 1 in the parental cells and the two $XIST^{-/-}$ clones, 10 and 13. Unlike the situation in parental WT K652 cells, the XIST clouds were absent in the clones with an *en bloc* *XIST* deletion (Figure 1D and E), suggestive of *XIST* loss-of-function. To examine the expression of *XIST* RNA in clones containing the *en bloc* deletion, we isolated poly(A)-containing RNAs from parental cells, mock transfected cells and two biallelic knockout clones (clones 10 and 13), and performed quantitative RT-PCT (qRT-PCR) and RNA-seq. Compared to mock-transfected control cells, which showed a normal level of *XIST* RNA, the two knockout clones displayed no *XIST* expression detectable by qRT-PCR (Figure 1F) and few RNA-seq reads in the targeted regions (Figure 1G), suggesting that the loss of XIST clouds resulted from the deletion of the entire transcriptional unit.

Segmental deletions of human *XIST*

A complete deletion of a lncRNA gene could also remove unknown DNA regulatory elements within the lncRNA gene. It is often difficult to distinguish whether the resulting phenotype is caused directly by the depletion of the lncRNA *per se* or indirectly by loss of the regulatory DNA elements (such as enhancers) that modulate expression of another gene. Moreover, a complete deletion cannot pinpoint the functionally important lncRNA sequences. Therefore, segmental deletions could be a good option for studying any functional domains within a lncRNA. We first set out to delete eight different segments of the *XIST* gene, SD1 to SD8, each of which includes a part of an exon, a whole exon, or multiple exons (Figure 2A and Supplementary Figure S2a). To examine the role of the repeat elements of exon 1, we also targeted three additional repeat elements: A (SD9), D (SD10) and B+C (SD11) of exon 1 (Figure 2A and Supplementary Figure S2e). Repeat F was not targeted due to its close proximity to other repeats (it is located between repeats A and B). Using the same procedure as described for the complete deletions, we transfected sgRNA pairs that resulted in higher mutation frequencies (Supplementary Figure S2b and f), examined the expected sizes of deletion-recognizing PCR amplicons and confirmed the presence of targeted segmental deletions (Supplementary Figure S2c and g). The numbers of clones that contained targeted segmental deletions of the *XIST* gene ranged from 4 to 27, which corresponded to 14–75% of the examined clones (Figure 2B; Supplementary Figure S2d and h). Among the clones with desired deletions, several clones did not contain the wild-type allele, indicating biallelic deletions (Figure 2C and Supplementary Figure S3). For those clones containing biallelic deletions, we next sequenced the deletion junctions in the clones and observed up to two different alleles in each clone (Supplementary Figure S4 and Supplementary Table S5).

To examine the effects of segmental deletions, compared to that of the WD, we also performed single molecule RNA FISH (with probes recognizing XIST exon 1 or 6 that detect regions outside of the deleted region in each clone) on clonal populations of cells containing each deletion (Figure 2A and H; Supplementary Table S6). When we examined about a hundred cells per clone and determined how many contained an XIST cloud, we found that the cell fractions were bimodally distributed (Supplementary Figure S7a). If more than 50% of the cells displayed an XIST cloud, we defined the clone as one with XIST clouds. The XIST clouds displayed in the control clone containing no deletions (Figure 2D) were absent in the clone that contains a segmental deletion of a promoter region (SD5; Figure 2E). On the other hand, many clones containing non-promoter segmental deletions (SD3, 4, 6 and 10) displayed XIST cloud formation (Figure 2F and Supplementary Figure S5). The other deletions (SD1, 2, 8, 9 and 11) affected XIST cloud formation differently, depending on the indel (Figure 2G and Supplementary Figure S5). SD1 clone 14 (designated SD1-14), SD2-2, SD7-10, 30, SD8-11, SD9-33 and SD11-7 lacked XIST clouds (Figure 2H and Supplementary Figure S5).

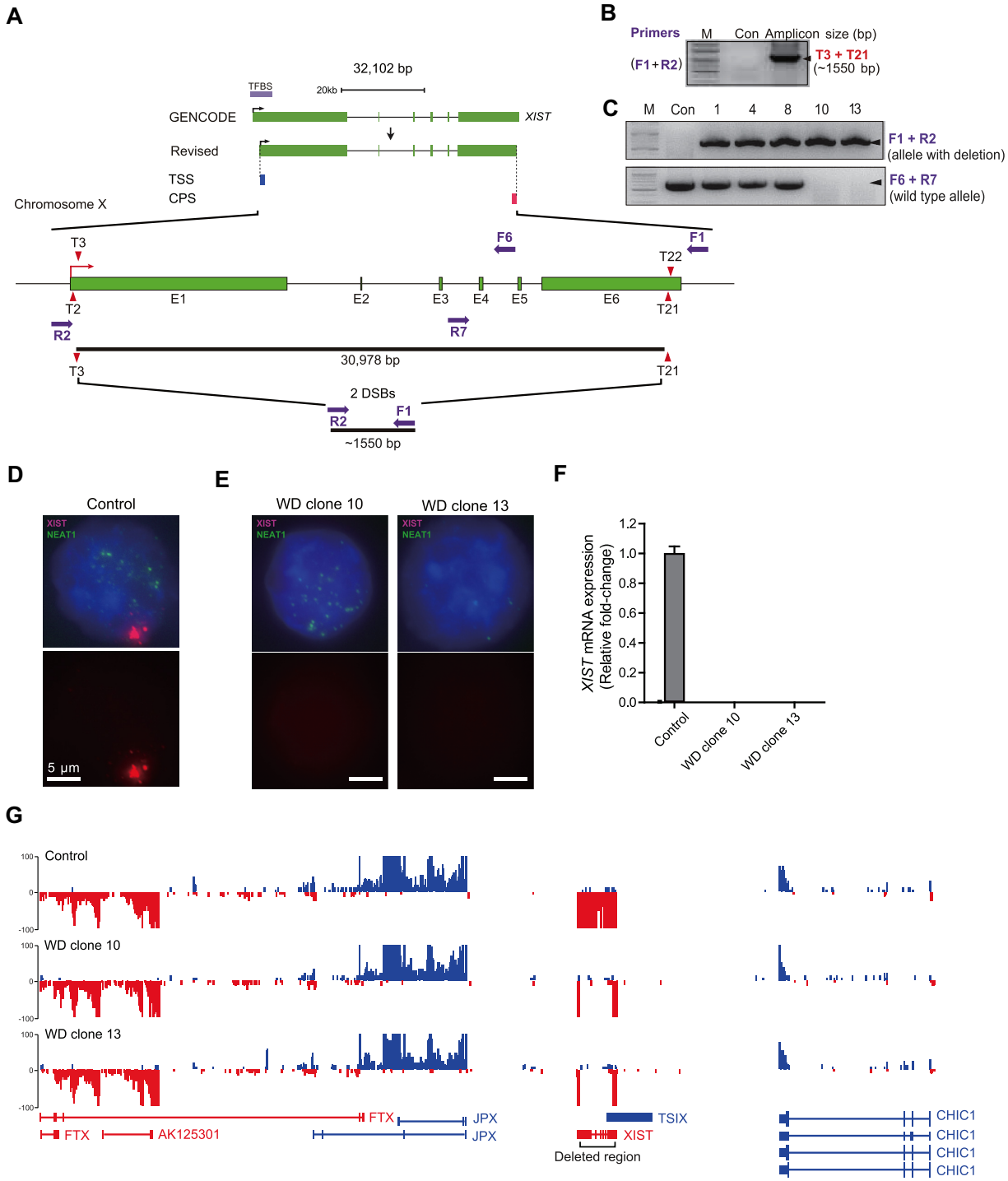

Figure 1. Paired Cas9 nuclease-mediated *en bloc* deletions of *XIST*. (**A**) Schematic depicting the human *XIST* gene structure and targeted deletion sites. The six exons are shown (E1-E6). Red arrowheads indicate the sgRNA-binding sites (T2, T3, T21 and T22). The targeted *en bloc* deletion of 30.9 kb is shown with a black line with red arrow tips (T3 and T21). F1 and R2, and F6 and R7 represent PCR primers (purple arrows) used for the detection of the wild-type allele and the alleles with targeted deletions, respectively. The grayish purple colored box indicates TFBSs. TSS denotes the transcription start site and CPS is the cleavage and polyadenylation site from the BIGTranscriptome annotations. (**B**) PCR-based detection of targeted deletions. Genomic DNA isolated from K562 cells transfected with plasmids encoding Cas9 and the sgRNA pair (T3 and T21) were subjected to PCR to detect the targeted deletion. The arrowhead indicates the expected position of the amplicon from the region containing the targeted deletion. Untransfected cells were used as control (Con). The primer and sgRNA pairs are shown on the left and right, respectively. (**C**) Identification of clones containing the entire *XIST* deletion. The wild-type and deleted alleles were detected by PCR using different primer pairs. The arrowheads indicate the expected positions of the amplicons. (**D** and **E**) RNA FISH of *XIST* (pink) and *NEAT1* (green) in mock transfected cells (D; control) and in clones with Cas9 nuclease-induced WD (E). The white bars in the FISH images indicate the scale of the image (5 micrometers). (**F**) *XIST* mRNA expression levels in control cells and in clones with Cas9 nuclease-induced WD determined using qRT-PCR. (**G**) RNA-seq signals for *XIST* and other genes near *XIST* in mock transfected cells and clones with Cas9 nuclease-induced WD. The color of the signals indicates the orientation of transcribed genes (plus strand as blue and minus strand as red).

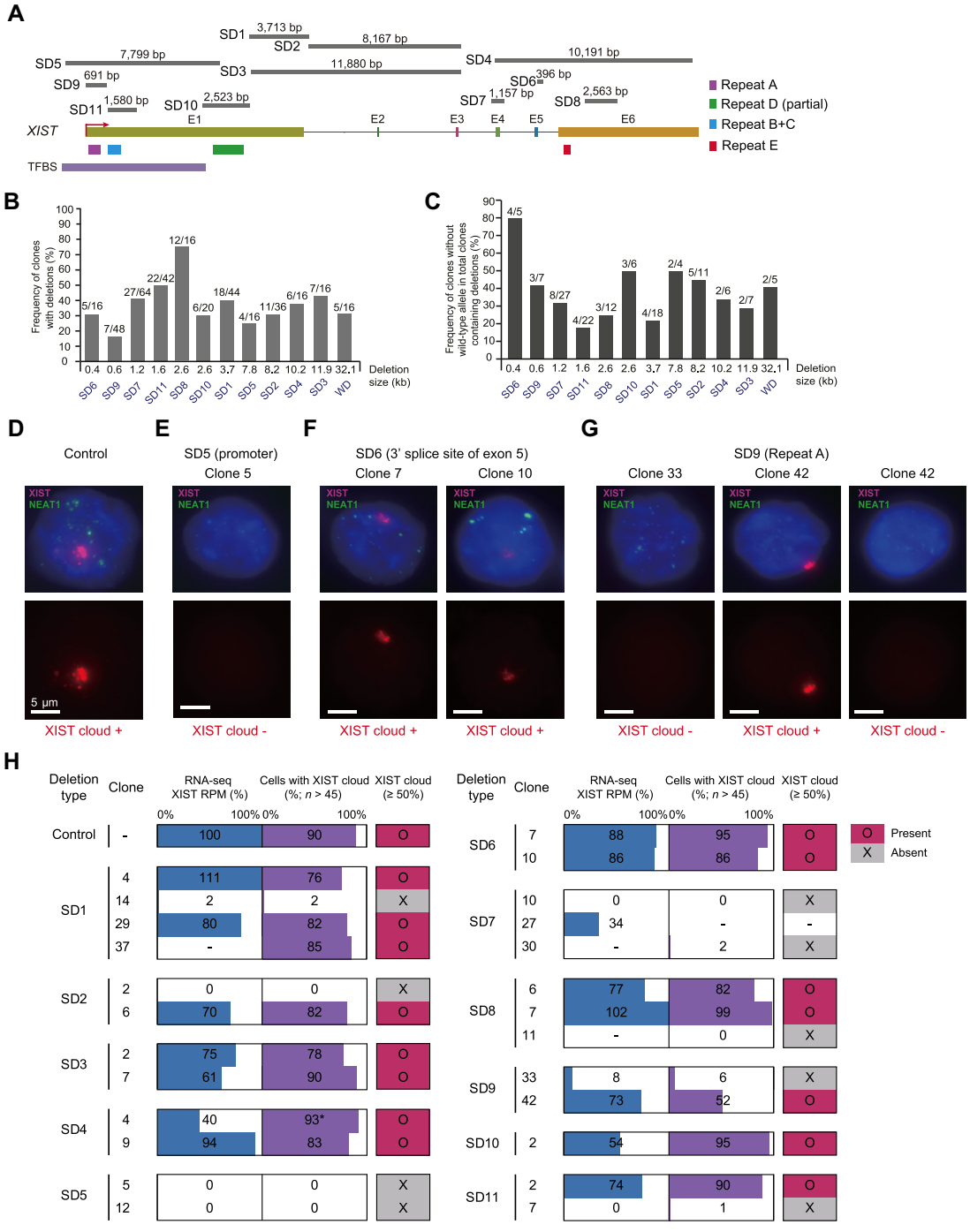

Figure 2. Cas9-induced segmental deletions of the *XIST* gene. (**A**) Schematic depicting the human *XIST* gene structure and targeted segmental deletion (SD) sites. The six exons are shown as E1–E6. The red arrowhead indicates the direction of *XIST* transcription. Targeted deletion sites are shown using gray lines (SD1-SD11) with the deletion sizes. A segment including part of exon 1 and a 3′ splice site (SD1); a segment including exons 2 and 3 (SD2); a segment including SD1 and SD2 (SD3); a segment including exons 4, 5 and 6 with repeat E (SD4); a segment including a promoter (SD5); a segment including the 3′ splice site of exon 5 (SD6); a segment including exon 4 (SD7); a segment including part of exon 6 but not repeat E (SD8), a segment including repeat A (SD9), a segment including repeat D (SD10) and a segment including repeats B and C (SD11). The light grayish purple colored box indicates TFBSs, purple, cyan, green and red boxes indicate repeat elements. (**B**) Frequencies of clones containing segmental and whole (SD and WD) deletions of the *XIST* gene. (**C**) Frequencies of clones containing segmental and whole (SD and WD) deletions of the *XIST* gene that also lack the wild-type allele. (**D**–**G**) RNA FISH of *XIST* (pink) and *NEAT1* (green) in parental cells (control; D) and clones with Cas9 nuclease-induced segmental *XIST* deletions, SD5 (E), SD6 (F), and SD9 (G). (G) About half of the cells in a population of SD9 clone 42 cells display XIST clouds, whereas the other half do not. The white bars in the FISH images indicate the scale of the image (5 micrometers). (**H**) The percentage of XIST expression (reads per million; RPM) in clones with segmental deletions in the *XIST* gene, relative to that of the control, and the percentage of cells with an XIST cloud in clones containing segmental deletions of the *XIST* gene. The XIST cloud is indicated by color codes as described in the legend. 'Present' means that more than 50% of the cells in a clone contain a cloud and 'absent' means that <50% of the cells contain a cloud. A hypen indicates that there is no information and the asterisk (*) in the SD4-4 results indicates that the cells have abnormal XIST cloud patterns.

Non-promoter segmental deletions may mediate either alternative splicing or RNA decay

Because XIST clouds were absent in the clone containing a promoter deletion and in several other clones containing non-promoter deletions, we conducted qRT-PCR and RNA-seq to examine the levels of expression of *XIST* and other X-linked genes in these clones compared to that in the control clone (Figure 3 and Supplementary Figure S6). No signs of XIST transcripts were detected in any of the promoter-deleted clones (SD5) by qRT-PCR or RNA-seq (Figure 3E and Supplementary Figure S6e), suggesting that the promoter deletion leads to transcriptional inactivation of *XIST*. In contrast, the non-promoter segmental deletions seem to either affect RNA decay or induce alternative splicing according to the type of repair (seamless or indel formation) that had occurred during non-homologous end-joining (Figure 3A–D and F–K); these effects perhaps resulted from RNA quality control (26) and alternative splicing using cryptic splicing sites during post-transcriptional regulation of the lncRNA. For instance, the majority of segmental deletions of the *XIST* gene induced alternative splicing, some of which used cryptic splice sites (clones SD1-4,29; SD2-6; SD3-2,7; SD4-4; SD6-7,10; SD7-27; SD9-33,42; SD10-2; SD11-2; Figure 3). Intriguingly, a similar alternative splicing pattern was caused by deletion of repeat A (SD9) in mouse cells (27). Some clones containing segmental deletions exhibited heavily reduced *XIST* levels (SD1-14; SD2-2; SD4-4; SD7-10; SD9-33; SD11-7; Figure 3), whereas other clones containing such deletions expressed *XIST* at levels that were slightly reduced or comparable to that of the control (SD1-4,29; SD2-6; SD3-2,7; SD4-9; SD6-7,10; SD7-27; SD8-6,7; SD9-42; SD10-2; SD11-2; Figure 3).

Analysis of the RNA-seq reads mapped to the *XIST* locus indicated that XIST cloud formation relied on RNA expression (Supplementary Figure S7). For instance, the loss of XIST clouds in SD1-14 resulted from an ∼25-fold reduction of *XIST* RNA compared to that of the parental cells (Figure 3A). However, SD1-4 and -29, containing similar genomic deletions (Supplementary Figure S4a), displayed XIST clouds, which resulted from alternatively spliced forms of *XIST* RNA expressed at levels comparable to that in the parental cells (Figure 3A). Similarly, SD9-33 lost XIST cloud formation as a result of a reduction in *XIST* RNA levels (Figure 3I), whereas SD9-42, containing a similar genomic deletion (Supplementary Figure S4i), formed XIST clouds (Figure 3I). In addition, clones SD2-6, SD3-2,7 SD4-9, SD6-7,10, SD8-6,7, SD10-2 and SD11-2, which displayed XIST clouds, expressed alternatively spliced forms of *XIST* at similar or decreased levels compared to the control (Figure 3).

Although some clones (SD4-4; SD9-33; SD11-7) containing deletions of repeat elements expressed *XIST* at heavily reduced levels, others (SD4-9; SD9-42; SD10-2; SD11-2) expressed *XIST* at levels that were only slightly reduced or comparable to that of the control (Figure 3). A previous mouse *Xist* study, in which conserved repeat elements were deleted, showed that repeats A and B are essential domains for *Xist* localization and silencing functions (10,28). More recent studies showed that the repeat E element of mouse *Xist* is essential for the localization of *Xist* RNA to the inactive X chromosome (Xi) (11,12) and for the expression of escape genes in the inactive X chromosome during XCI (29). In fact, SD4-4, which contains a deletion of exon 4, 5 and 6 including repeat E, expressed a lower level of *XIST* but displayed a scattered *XIST* RNA localization pattern (Figure 3D and Supplementary Figure S5f), as reported in a previous Xist repeat E study (11,12). Because SD6 clones containing deletions of exon 5 and SD8 clones containing partial deletions of exon 6 that do not include repeat E displayed normal XIST clouds, the abnormal XIST clouds in SD4-4 cells could result from the deletion of repeat E. While we failed to detect the wild-type allele in SD4-9, our RNA-seq and RNA FISH results (Supplementary Figure S5f) suggested that SD4-9 may have a heterozygotic deletion on the active X chromosome, explaining the difference of XIST cloud morphology between SD4-4 and SD4-9. In addition, about half of the cells in SD9-42, which contains a deletion of repeat A, formed XIST clouds (Figure 3I), indicating that repeat A is involved in cloud formation. However, other clones containing deletions of repeat D (SD10-2; Figure 3J) or repeat B+C (SD11-2; Figure 3K) exhibited XIST cloud formation, indicating that repeats D and B+C are dispensable for this process.

Functional domain deletions that lead to up-regulation of X chromosome-linked genes

To check whether biallelic *XIST* deletion can lead to reactivation of an inactivated X chromosome, expression differences between X-linked genes across the X chromosome in mock transfected controls and clones with different biallelic deletions were analyzed (Figure 4A). The clones containing a WD of *XIST* displayed globally elevated transcription signals across the X chromosome (except for part of the q-arm) but not for chromosome 10, which is similar in size to the X chromosome (Figure 4B and Supplementary Figure S8). The reactivation of the X chromosome by whole *XIST* deletion was also observed in clone 12, which contains a segmental deletion of a promoter region, although the q-arm region was slightly down-regulated in this clone (Figure 4B and Supplementary Figure S8).

To identify functional domains of the *XIST* RNA involved in XCI, clones expressing *XIST* RNAs with specific segmental deletions were examined for the reactivation of the X chromosome (Supplementary Figure S8). To quantify the reactivation of the X chromosome, the sum of the up-regulation (positive values of $\log_2$ fold-changes) on the p-arm of the X chromosome was calculated in clones containing whole and segmental deletions, and was normalized to that of control clones transfected with scrambled sgRNAs and Cas9 (Figure 4C and D). Clones containing WD and promoter deletions (SD5), and clones in which *XIST* expression was almost completely depleted by segmental deletions (SD11-7, SD1-14 and SD9-33), showed comparable reactivation of X-linked genes (Figure 4D). Clones in which *XIST* RNA was moderately decreased by segmental deletions (SD4-4, SD10-2 and SD2-6) showed moderate reactivation of X chromosomes (Figure 4D). In contrast, clones in which *XIST* is expressed at levels comparable to that of the control (SD1-4, SD1-29, SD3-2, SD3-7, SD8-6, SD8-7 and SD9-42) displayed no reactivation, indicating that the

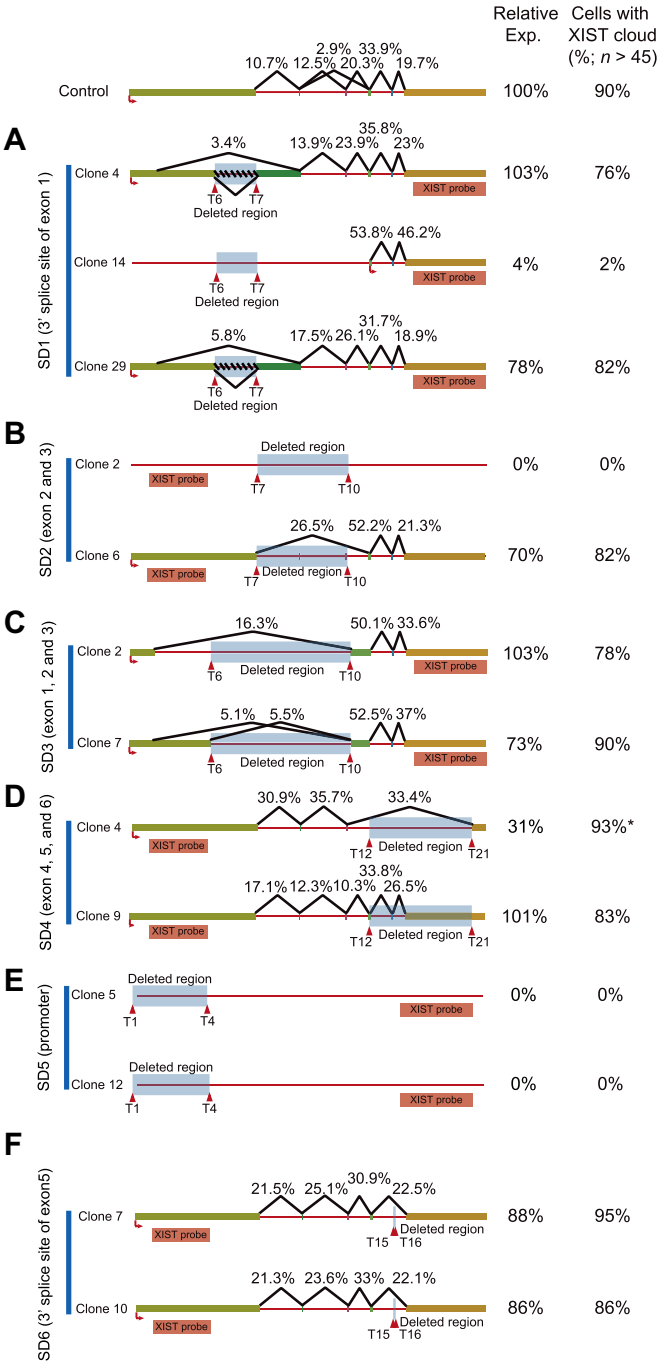

Figure 3. Diagrams summarizing the alternative *XIST* splicing in mock transfected cells and clones transfected with Cas9 and sgRNA pairs (red arrowheads). The gene models were annotated using exon-junction reads mapped to the *XIST* locus. The number over each splice junction indicates the percentage of exon junction reads at the splice junction. A light blue box indicates the deleted region and a light red box indicates the location of the *XIST* probe. (**A**–**K**) Summaries of information about *XIST* in clones in which the segmental regions, SD1 (A), SD2 (B), SD3 (C), SD4 (D), SD5 (E), SD6 (F), SD7 (G), SD8 (H), SD9 (I), SD10 (J) and SD11 (K), have been deleted. The right-side numbers indicate the expression level (RPKM, level (percentage) relative to control) of the *XIST* gene and the percentage of cells that display an XIST cloud. A hypen indicates that there is no information and the asterisk (*) in the SD4-4 results indicates that the cells have abnormal XIST cloud patterns.

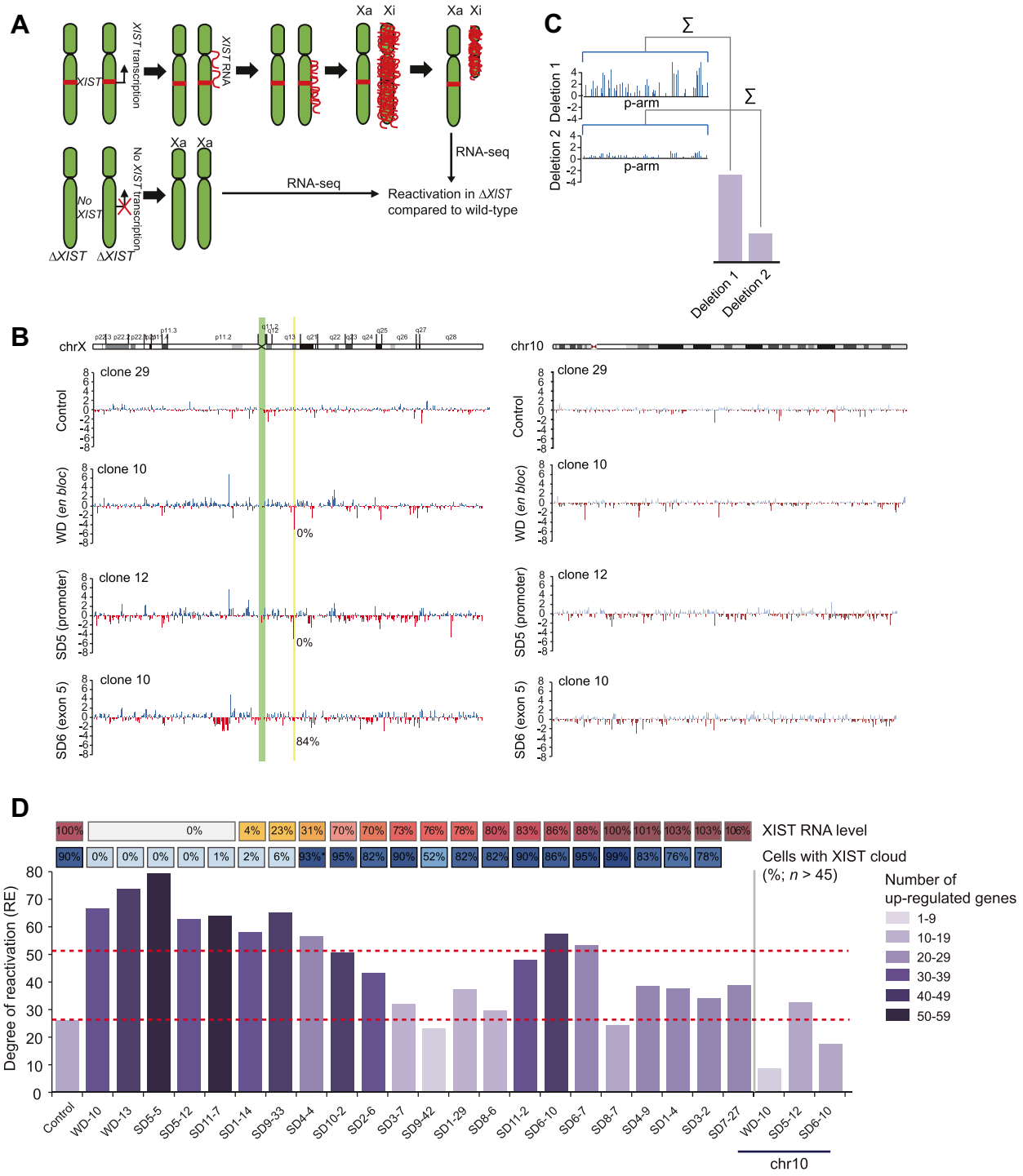

Figure 4. Reactivation of X-linked genes by *XIST* deletions. (**A**) Schematic flow of reactivation analyses with RNA-seqs from control and clones containing whole and segmental deletions. (**B**) Log_2 fold-changes in gene expression on the X chromosome and chromosome 10 in scrambled sgRNA-transfected (control) or sgRNA-transfected (WD, SD5 and SD6) clones relative to mock transfected cells. The green thick line indicates the centrosome and the yellow line indicates the *XIST* locus on the X chromosome. (**C**) A summarized method to measure the extent of reactivation in a clone containing segmental deletions. (**D**) Degree of reactivation of X-linked genes on the p-arm of chromosome X and genes on the p-arm of chromosome 10 in control and clones containing whole and segmental deletions. The *XIST* RNA level in each clone is depicted using a yellow-red color code with percentages relative to the level of the control. XIST cloud formation is presented as the percentage of cells with a cloud, indicated by the intensity of the cyan-colored boxes; the asterisk (*) in the SD4-4 results indicates that the cells have abnormal XIST cloud patterns. The number of up-regulated genes is indicated by a gray-purple color code as described in the legend.

3′ part of exons 1, 2, 3 and 4, as well as repeat A, appear not to be required for XCI maintenance. However, clones (SD6-7 and -10) that express exon5-skipped *XIST* RNAs at levels similar to that of the control and display XIST cloud formation showed a reactivation of X-linked genes (Figure 4D), indicating that exon 5 is required for XCI maintenance but not for XIST cloud formation. In addition, SD11-2, which contains a deletion of repeat B+C, expressed XIST splicing isoforms at levels similar to that of the control (Figure 3K) but showed a moderate reactivation of X-linked gene expression (Figure 4D), indicating that repeat B+C may also be important for XCI maintenance.

Chromosome-wide heterochromatinization of the Xi is one of hallmark of XCI. These modifications include H3K27me3 and H2AK119ub, among others. Understanding how Xist RNA recruits chromatin modifying complexes has been a topic of intense interest in the field. However, K562 cells do not appear to have visible H3K27me3 enrichment on the Xi, compared to elsewhere in the nucleus (Supplementary Figure S9a). To confirm our conclusion, we reanalyzed a dataset from a recent publication in which K562 cells were used in a genomics experiment (30). Although a lack of allelic information might bias our interpretation, we failed to find any significant enrichment of H3K27me3 on the X chromosomes (Supplementary Figure S9b). Instead, X chromosomes appeared to have less H3K27me3 than other chromosomes. In addition, H2AK119ub enrichment on the Xi was rarely detected in K562 cells, unlike in mouse ES cells (Supplementary Figure S9a). Therefore, we did not investigate which motifs in XIST RNA are involved in chromatin modifications.

Allele-specific reactivation of X-linked genes

Our RNA-seq results from clones containing XIST deletions suggest that up-regulation of X-linked genes might be an indication that the Xi has been reactivated. However, due to limited allelic information about our cell line, we could not rule out up-regulation of the Xa. To overcome this problem, we performed RNA FISH of X-linked genes in SD6 clones containing deletions of exon 5 and compared the results to those in control cells. If a nascent transcript is detected and its signal overlaps with the XIST cloud, this finding confirms gene expression from the Xi. On the other hand, an RNA signal that does not overlap with the XIST cloud indicates gene expression from the Xa. For this experiment, we first selected eight genes with RPKM $\geq$ 3 and that were highly up-regulated in both SD6 clones (2-fold greater than in the control). Because BMX and PIR, and MAGEB2 and MAGEB1 are clustered on the X chromosome, we finally chose PIR and MAGEB2 as representative RNA FISH targets. For the PIR gene, an RNA signal was detected from the Xi in 16% of cells in the WT population (Xi + Biallelic expression). In SD6-10, we observed that the PIR transcript overlapped with the XIST cloud in 48% of the cells, indeed suggesting Xi reactivation (Figure 5A). Similar reactivation was seen in the other clone, SD6-7 (Supplementary Figure S10a). On the other hand, the frequency of the MAGEB2 transcript from the Xi did not increase in the SD6 clones. This result might be due to technical limitations in our RNA FISH and RNA-seq; detecting a nascent transcript at the transcription site by RNA FISH requires on-going transcription, whereas RNA-seq is more likely to detect steady gene expression.

We next asked whether the observed reactivation of X-linked genes resulted from the reactivation of Xi and not from the Xa by comparing relative minor allele frequencies (MAF) in parental and mutant clones containing biallelic deletions and displaying global reactivation on X chromosomes (Figure 5B). If the increased expression originated from Xi, then the MAF in mutant clones would be greater than that of the parental clones. To accurately find heterozygous alleles in K562 genomes, we performed a resequencing of the genomes with 30× coverage and called variants using HyplotypeCaller version 3.4 of the genome analysis tool kits (GATKs; see 'Whole genome sequencing and variant calling' in the 'Materials and Methods' section for more details). Of the 7335 resulting heterozygous alleles on the X chromosomes, 770 on the exons from reactivated X-linked genes, expressed at more than 0.1 FPKM were subjected to an analysis of allele-specific expression (see 'Allele-specific expression analysis' in the 'Materials and Methods' section for more details). We found seven reactivated X-linked genes (*MSL3*, *JPX*, *TXLNG*, *HDAC8*, *BTK*, *RP11-706O15.1* and *CXorf40A*) that displayed allele-specific reactivation on Xi in clones in which X-linked genes are reactivated (Supplementary Figure S10b); *MSL3, HDAC8* and *TXLNG* expression was recurrently observed across different clones (Supplementary Figure S10b).

DISCUSSION

Although a large genomic deletion is a straightforward way to completely knockout long non-coding genes, a WD of a non-coding gene could accidently include the deletion of regulatory DNA elements for neighboring genes, such as enhancers. Thus, it is necessary to systematically perform segmental deletions of a non-coding gene, guided by a precise genomic map of the sequences encoding such regulatory DNA elements. In our study, we integrated *XIST* gene annotations and genomic elements of the locus from ENCODE (31) and our BIGTranscriptome (32) maps, which enabled us to unambiguously design paired sgRNAs for segmental deletions. The resulting segmental deletions of the *XIST* RNA domains allowed us to examine the functional effects of the domains.

A pairwise sequence alignment demonstrated that the human *XIST* RNA sequence is highly or moderately conserved compared to the mouse *Xist* RNA except in exon 2, a human-specific exon; human exons 4 and 5 displayed a greater sequence similarity to the mouse *Xist* RNA (Supplementary Figure S10c) than do the other exons. Despite the sequence similarity, in our study, human exon 5 specifically turned out to be important for XCI maintenance, but the function of mouse exon 6 (matched to human exon 5) remains unknown. In the mouse, exon 2, which is homologous with sequences in human *XIST* intron 1, and exon 5, which is absent in human *XIST* are dispensable for XCI (Supplementary Figure S10c). The pairwise sequence alignment also demonstrates where the functional domains for Xist cloud formation and XCI map to human *XIST*. Of the mouse repeat elements in *Xist* RNA that are function-

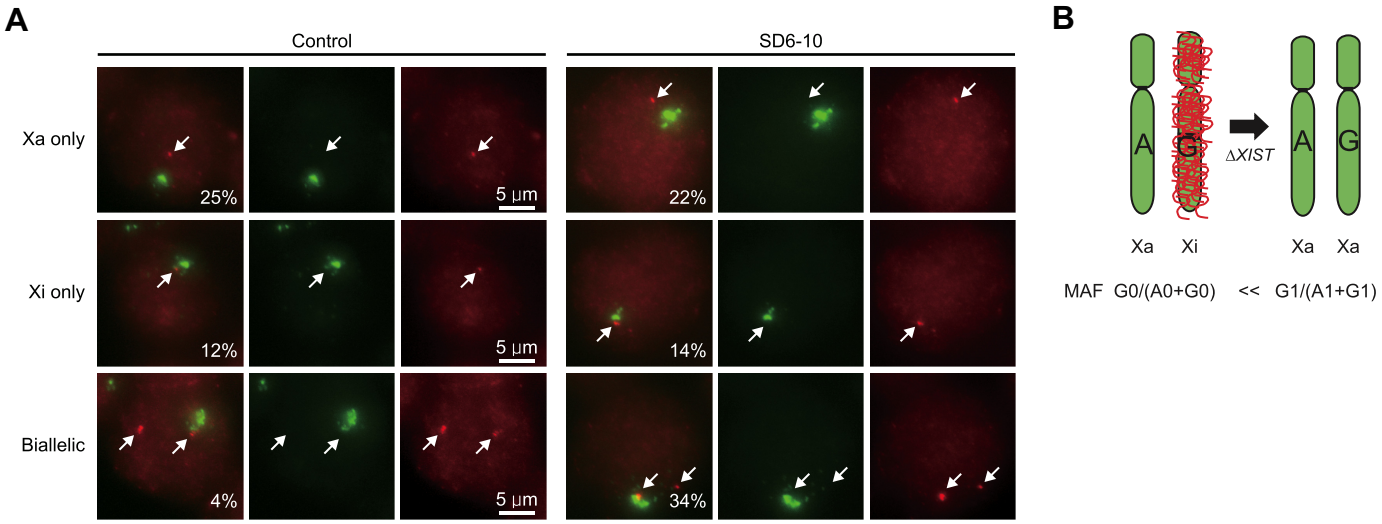

Figure 5. Allele-specific expression of X-linked genes. (**A**) Representative images of XIST (green) and nascent PIR transcripts (red) in WT and SD6-10 clone cells. Numbers indicate the % frequency of each type of image: expression from the Xa only, from the Xi only or from both Xa and Xi in a population of cells. Over 100 cells were counted for each clone. (**B**) Activation (Xa) of the inactive X chromosome (Xi) by *XIST* whole or segmental deletions, which derive a change in the MAF between G0/(A0+G0) in parent cells and G1/(A1+G1) in clones containing *XIST* deletions, where G0 is the minor allele frequency, A0 is the count-allele frequency in parental cells, and G1 is the frequency of the same allele as G0, and A1 is the frequency of the same allele as A0 in clones containing *XIST* deletions.

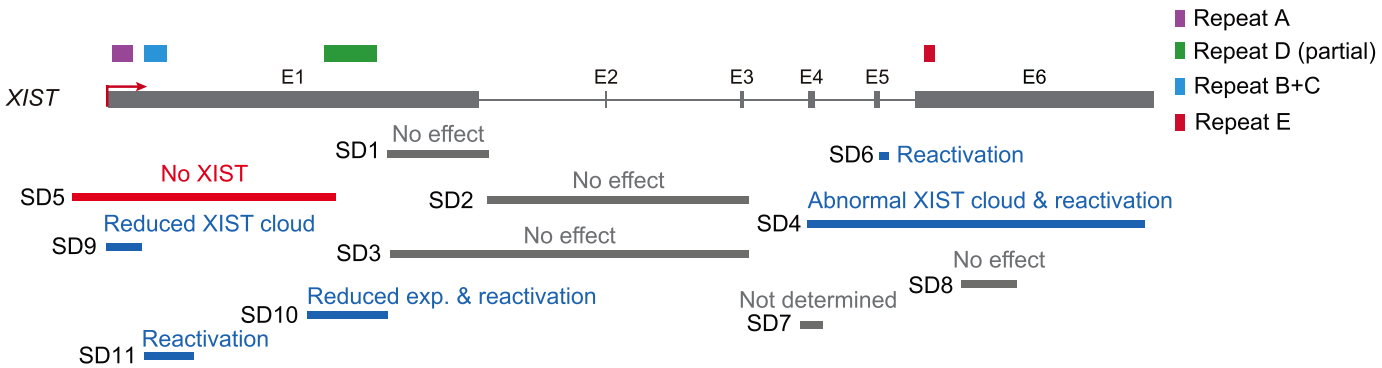

Figure 6. Summary of *XIST* functional elements discovered in this study. The relative location and effects of all 11 segmental deletions are shown in red for those that cause complete ablation of *XIST* expression (SD5), blue for those that either reduce or cause abnormal cloud formation or reactivate X-linked genes (SD4, 6, 9, 10, 11) and gray for those that have no effect (SD1, 2, 3 and 8) or for which the phenotype was not determined.

ally important for gene silencing and *Xist* cloud formation, repeats A and E similarly matched to the 5′ part of exons 1 and 6 of human *XIST*, respectively. In contrast, repeats B and F are less well conserved between mouse Xist and human XIST (33,34). One of our most surprising results was that the segmental deletion of repeat A moderately affected XIST cloud formation but did not affect the global expression of X-linked genes, although previous studies reported that repeat A is important for XCI in mice. This finding is not due to the stability of XCI because our *en bloc XIST* deletions reactivated silenced genes on the Xi. This result may instead suggest that the silencing domain (repeat A) is involved, not in maintaining XCI, but in establishing XCI during early development or in XIST cloud formation. In addition, human repeat E seems to be important for both XIST cloud formation and XCI maintenance; murine repeat E, and repeats B, C and D, also appear to moderately affect XCI maintenance (Figure 6). In summary, both similarities and differences between human XIST and mouse Xist were observed in terms of functional domains involved in cloud formation and XCI maintenance. To gain further understanding about the functional domains of human XIST in a mechanistic sense, RNA-binding proteins associated with the functional domains should be investigated in the near future.

DATA AVAILBILITY

All sequence data can be accessed at the NCBI Gene Expression Omnibus (GEO) using the accession number (GSE65830).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank members of the Kim and Nam labs for helpful comments and discussions.
Author contributions: H.K. and J.W.N. designed the experiments. J.W.N., H.K., S.R., R.G. and H.L. wrote the manuscript. R.G., H.L., H.J.S. and S.W.C. performed the experiments. *XIST* knockout phenotype analysis and related manuscript writing were conducted by H.S. and J.T.L.

FUNDING

National Research Foundation (NRF) of Korea, funded by the Ministry of Science & ICT [2017M3A9G8084539, 2018R1A2B2003782, 2014M3C9A3063541 to J.W.N.; 2017R1A2B3004198, 2017M3A9B4062403 to H.K.]; Korean Health Technology R&D Project, Ministry of Health and Welfare, Republic of Korea [HI15C3224 to J.W.N.; HI17C0676, HI16C1012 to H.K.); Institute for Basic Science [IBS; IBS-R026-D1]; Yonsei University New Faculty Research Seed Money Grant of 2015 (6-2015-0086 to H.K.); faculty research grant of Yonsei University College of Medicine for 2015 [6-2015-0086 to H.K.]; National Institutes of Health (NIH) [5T32HD007396-24 to H.S]; NIH [RO1-GM090278 to J.T.L.]. Funding for open access charge: National Research Foundation (NRF) of Korea, funded by the Ministry of Science & ICT [2017M3A9G8084539, 2018R1A2B2003782, 2014M3C9A3063541 to J.W.N.; 2017R1A2B3004198, 2017M3A9B4062403 to H.K.].
Conflict of interest statement. None declared.

REFERENCES

1. Borsani,G., Tonlorenzi,R., Simmler,M.C., Dandolo,L., Arnaud,D., Capra,V., Grompe,M., Pizzuti,A., Muzny,D., Lawrence,C. *et al.* (1991) Characterization of a murine gene expressed from the inactive X chromosome. *Nature*, **351**, 325–329.
2. Brockdorff,N., Ashworth,A., Kay,G.F., McCabe,V.M., Norris,D.P., Cooper,P.J., Swift,S. and Rastan,S. (1992) The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell*, **71**, 515–526.
3. 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.
4. 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.
5. Lee,J.T. (2011) Gracefully ageing at 50, X-chromosome inactivation becomes a paradigm for RNA and chromatin control. *Nat. Rev. Mol. Cell Biol.*, **12**, 815–826.
6. Yang,L., Kirby,J.E., Sunwoo,H. and Lee,J.T. (2016) Female mice lacking Xist RNA show partial dosage compensation and survive to term. *Genes Dev.*, **30**, 1747–1760.
7. Clemson,C.M., McNeil,J.A., Willard,H.F. and Lawrence,J.B. (1996) XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. *J. Cell Biol.*, **132**, 259–275.
8. Sunwoo,H., Wu,J.Y. and Lee,J.T. (2015) The Xist RNA-PRC2 complex at 20-nm resolution reveals a low Xist stoichiometry and suggests a hit-and-run mechanism in mouse cells. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, E4216–E4225.
9. Augui,S., Nora,E.P and Heard,E. (2011) Regulation of X-chromosome inactivation by the X-inactivation centre. *Nat. Rev. Genet.*, **12**, 429–442.
10. 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.
11. Sunwoo,H., Colognori,D., Froberg,J.E., Jeon,Y. and Lee,J.T. (2017) Repeat E anchors Xist RNA to the inactive X chromosomal compartment through CDKN1A-interacting protein (CIZ1). *Proc. Natl. Acad. Sci. U.S.A.*, **114**, 10654–10659.
12. Ridings-Figueroa,R., Stewart,E.R., Nesterova,T.B., Coker,H., Pintacuda,G., Godwin,J., Wilson,R., Haslam,A., Lilley,F., Ruigrok,R. *et al.* (2017) The nuclear matrix protein CIZ1 facilitates localization of Xist RNA to the inactive X-chromosome territory. *Genes Dev.*, **31**, 876–888.
13. Brown,C.J., Hendrich,B.D., Rupert,J.L., Lafreniere,R.G., Xing,Y., Lawrence,J. and Willard,H.F. (1992) The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell*, **71**, 527–542.
14. Yang,C., Chapman,A.G., Kelsey,A.D., Minks,J., Cotton,A.M. and Brown,C.J. (2011) X-chromosome inactivation: molecular mechanisms from the human perspective. *Hum. Genet.*, **130**, 175–185.
15. Chow,J.C., Hall,L.L., Baldry,S.E., Thorogood,N.P., Lawrence,J.B. and Brown,C.J. (2007) Inducible XIST-dependent X-chromosome inactivation in human somatic cells is reversible. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 10104–10109.
16. Minks,J., Baldry,S.E., Yang,C., Cotton,A.M. and Brown,C.J. (2013) XIST-induced silencing of flanking genes is achieved by additive action of repeat a monomers in human somatic cells. *Epigenetics Chromatin*, **6**, 23–32.
17. Canver,M.C., Bauer,D.E., Dass,A., Yien,Y.Y., Chung,J., Masuda,T., Maeda,T., Paw,B.H. and Orkin,S.H. (2014) Characterization of genomic deletion efficiency mediated by clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 nuclease system in mammalian cells. *J. Biol. Chem.*, **289**, 21312–21324.
18. Vidigal,J.A. and Ventura,A. (2015) Rapid and efficient one-step generation of paired gRNA CRISPR-Cas9 libraries. *Nat. Commun.*, **6**, 8083–8089.
19. Aparicio-Prat,E., Arnan,C., Sala,I., Bosch,N., Guigo,R. and Johnson,R. (2015) DECKO: single-oligo, dual-CRISPR deletion of genomic elements including long non-coding RNAs. *BMC Genomics*, **16**, 846–860.
20. Zhu,S., Li,W., Liu,J., Chen,C.H., Liao,Q., Xu,P., Xu,H., Xiao,T., Cao,Z., Peng,J. *et al.* (2016) Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR-Cas9 library. *Nat. Biotechnol.*, **34**, 1279–1286.
21. Lv,Q., Yuan,L., Song,Y., Sui,T., Li,Z. and Lai,L. (2016) D-repeat in the XIST gene is required for X chromosome inactivation. *RNA Biol.*, **13**, 172–176.
22. Hsu,P.D., Scott,D.A., Weinstein,J.A., Ran,F.A., Konermann,S., Agarwala,V., Li,Y., Fine,E.J., Wu,X., Shalem,O. *et al.* (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.*, **31**, 827–832.
23. Ramakrishna,S., Cho,S.W., Kim,S., Song,M. and Gopalappa,R., (2014) Kim J.S. and Kim H. Surrogate reporter-based enrichment of cells containing RNA-guided Cas9 nuclease-induced mutations. *Nat. Commun.*, **5**, 3378–3387.
24. Miller,S.A., Dykes,D.D. and Polesky,H.F. (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.*, **16**, 1215.
25. Gribble,S.M., Roberts,I., Grace,C., Andrews,K.M., Green,A.R. and Nacheva,E.P. (2000) Cytogenetics of the chronic myeloid leukemia-derived cell line K562: karyotype clarification by multicolor fluorescence in situ hybridization, comparative genomic hybridization, and locus-specific fluorescence in situ hybridization. *Cancer Genet. Cytogenet.*, **118**, 1–8.
26. Popp,M.W. and Maquat,L.E. (2016) Leveraging rules of nonsense-mediated mRNA decay for genome engineering and personalized medicine. *Cell*, **165**, 1319–1322.
27. Royce-Tolland,M.E., Andersen,A.A., Koyfman,H.R., Talbot,D.J., Wutz,A., Tonks,I.D., Kay,G.F. and Panning,B. (2010) The A-repeat links ASF/SF2-dependent Xist RNA processing with random choice during X inactivation. *Nat. Struct. Mol. Biol.*, **17**, 948–954.
28. Colognori,D., Sunwoo,H., Kriz,A.J., Wang,C.Y. and Lee,J.T. (2019) Xist deletional analysis reveals a co-dependency between Xist RNA and Polycomb complexes for spreading along the inactive X. *Mol. Cell*, in press.
29. Yue,M., Ogawa,A., Yamada,N. and Charles Richard,J.L., (2017) Barski A. and Ogawa Y. Xist RNA repeat E is essential for ASH2L

recruitment to the inactive X and regulates histone modifications and escape gene expression. *PLoS Genet.*, **13**, e1006890.
30. Skene,P.J., Henikoff,J.G. and Henikoff,S. (2018) Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nat. Protoc.*, **13**, 1006–1019.
31. Consortium,E.P. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, **489**, 57–74.
32. You,B.H., Yoon,S.H. and Nam,J.W. (2017) High-confidence coding and noncoding transcriptome maps. *Genome Res.*, **27**, 1050–1062.
33. Pintacuda,G., Wei,G., Roustan,C., Kirmizitas,B.A., Solcan,N., Cerase,A., Castello,A., Mohammed,S., Moindrot,B., Nesterova,T.B. *et al.* (2017) hnRNPK recruits PCGF3/5-PRC1 to the Xist RNA B-Repeat to establish polycomb-mediated chromosomal silencing. *Mol. Cell*, **68**, 955–969.
34. Jeon,Y. and Lee,J.T. (2011) YY1 tethers Xist RNA to the inactive X nucleation center. *Cell*, **146**, 119–133.