(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
25 September 2014 (25.09.2014)

(10) International Publication Number
WO 2014/152607 A2

(51) International Patent Classification: Not classified

(21) International Application Number: PCT/US2014/027525

(22) International Filing Date: 14 March 2014 (14.03.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/785,481 14 March 2013 (14.03.2013) US
61/790,917 15 March 2013 (15.03.2013) US

(71) Applicant: UNIVERSITY OF MASSACHUSETTS [US/US]; 225 Franklin Street, 33rd Floor, Boston, Massachusetts 02110 (US).

(72) Inventors: LAWRENCE, Jeanne B.; 650 Cooper Hill Road, Mapleville, Rhode Island 02839 (US). JIANG, Jun; 54C Shrewsbury Green, Shrewsbury, Massachusetts 01545 (US). HALL, Lisa L.; 28 Red Oak Dr., Sudbury, Massachusetts 01776 (US).

(74) Agent: DEYOUNG, Janice Kugler; Fish & Richardson P.C., P.O. Box 1022, Minneapolis, Minnesota 55440-1022 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:
— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:
— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*
— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: DOSAGE COMPENSATING TRANSGENES AND CELLS

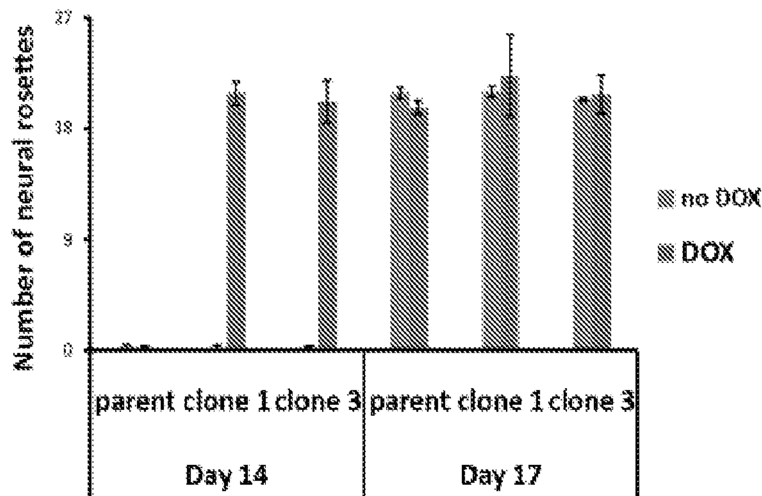


FIG. 6B

(57) Abstract: Methods and compositions for reducing expression of genes on Chromosome 21 ("Chr 21") by targeting an XIST transgene to the Dual specificity tyrosine-phosphorylation-regulated kinase 1A (DYRK1A) gene or a Regulator of calcineurin 1 (RCAN1) gene, and cells and transgenic animals comprising an XIST transgene inserted into a DYRK1A or RCAN1 allele, e.g., cells and animals trisomic for human Chr 21 and mouse Chr 16.

# DOSAGE COMPENSATING TRANSGENES AND CELLS

## CLAIM OF PRIORITY

This application claims the benefit of U.S. Patent Application Serial Nos. 61/785,481, filed on March 14, 2013, and 61/790,917, filed on March 15, 2013. The entire contents of the
5 foregoing are hereby incorporated by reference.

## FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

This invention was made with Government support under Grant No. GM053234, GM085548 and GM096400 RC4 awarded by the National Institutes of Health. The Government has certain rights in the invention.

10 
## TECHNICAL FIELD

This invention relates to methods of effecting dosage compensation in cells trisomic for chromosome 21, and cells produced by said methods.

## BACKGROUND

In the U.S., about 1 in 300 live births carry a trisomy, roughly half of which are
15 trisomy for chromosome 21 (Chr21), which causes Down syndrome (DS). DS is the leading genetic cause of cognitive disability with increasing prevalence, and millions of patients worldwide experience congenital and progressive medical issues that impact multiple organ systems[1,2]. In addition to progressive intellectual impairment and early onset Alzheimer disease, there is greatly increased risk of myeloproliferative disorder, childhood leukemia,
20 heart defects, and both immune and endocrine system dysfunction. DS researchers have sought to define the more "DS critical" genes on Chr21, but this has proven difficult due to high genetic complexity and phenotypic variability of DS, confounded by normal variation between any individuals[1-3]. Much progress has been made in developing DS mouse models[4-6], however there remains a critical need for better ways to understand the underlying cell and
25 developmental pathology of human DS, so key to rationale design of therapeutics of any kind[7].

The last decade has seen great advances in strategies to correct single-gene defects of rare monogenic disorders, beginning with cells *in vitro* and in several cases advancing to *in*

*vivo* and clinical trials. In contrast, genetic correction of the over-dose of genes across a whole extra chromosome in trisomic cells has remained outside the realm of possibility.

## SUMMARY

At least in part, the present invention is based on the discovery that the imbalanced
5 expression of hundreds of genes across an extra chromosome can be *de facto* corrected in DS patient stem cells, by the targeted addition of one gene, XIST, into a specified gene, e.g., the Dual specificity tyrosine-phosphorylation-regulated kinase 1A (DYRK1A) locus, or the Regulator of calcineurin 1 (RCAN1) locus, on Chromosome 21 ("Chr 21"). Using genome editing with zinc finger nucleases, addition of a large, inducible XIST transgene to a precise
10 position in the Chr 21 DYRK1A or RCAN1 loci was achieved in DS iPSCs. This resulted in Chr 21 coating by the non-coding XIST RNA, heterochromatin modifications, chromosome-wide transcriptional silencing and DNA methylation to form a "Chr 21 Barr Body". Silencing became irreversible in differentiated cells. A model to study human chromosome silencing that avoids the selection against silencing of a disomic autosome was created by
15 targeting a trisomic chromosome with an inducible XIST transgene. Such inducible correction of the trisomy provides a system to investigate genomic expression changes and the cellular pathology of trisomy 21, free from genetic and epigenetic noise. Remarkably, a proliferative deficit of DS cells *in vitro* was reversed upon induced silencing of one Chr 21. The present vectors may be useful in "chromosome therapy" for Down syndrome.

20 Accordingly, the present invention features nucleic acid constructs that include a silencing sequence encoding an XIST RNA or fragment thereof that silences a segment of a chromosome), driven by a regulatory sequence comprising a promoter; first and second sequences that direct insertion of the silencing sequence into or near the DYRK1A or RCAN1 genes on chromosome 21; and, optionally, a selectable marker. The first and second
25 sequences that direct insertion of the silencing sequence into DYRK1A or RCAN1 may also be referred to herein as "first and second targeting elements." These sequences or elements can be readily selected and inserted into the nucleic acid constructs using methods well known in the art.

Thus, in one aspect, the invention provides silencing vectors comprising:
30 a silencing element comprising a silencing sequence flanked by first and second targeting sequences, wherein each of the first and second targeting sequences are homologous to at least 50 bp (e.g., 50, 100, 200, or 500) of sequence in or near (e.g., within 1MB, 0.5 MB, 0.1 MB, 0.05 MB, 10000 MB, 5000 MB, 1000 KB, 500 KB, 100KB, 50 KB, 10 KB, 5 KB, or 1

KB) the dual specificity tyrosine-phosphorylation-regulated kinase 1A (DYRK1A) gene, e.g., in the DYRK1A gene, e.g., in intron 1 of DYRK1A, or in or near the Regulator of calcineurin 1 (RCAN1) gene, e.g., in the RCAN1 gene, e.g., in intron 3 of RCAN1; and a promoter operably linked to the silencing element.

5 In some embodiments, the vector is a plasmid or a viral vector. In some embodiments, the viral vector is vaccinia virus, adeno-associated virus (MV), or herpes virus.

In some embodiments, the silencing vector targets intron 1 of human DYRK1A and the first targeting sequence comprises a sequence obtained by performing PCR with a primer pair of: Human Chr 21 DYRK1A left arm primers: forward 5'-
10 GCCGTATACCATTAACTCTTTACTGTTC-3' (SEQ ID NO:1), reverse 5'-TCTGTATACGTAAACTGGCAAAGGGGTGG-3' (SEQ ID NO:2); and the second targeting sequence comprises a sequence obtained by performing PCR with a primer pair of: Human Chr 21 DYRK1A right arm primers: forward 5'-ATTTCGCGAACGGGTGATGAGCAGGCTGT-3' (SEQ ID NO:3), reverse 5'-
15 CCGTCGCGAAAACCAGAAAGTATTCTCAG-3' (SEQ ID NO:4).

In some embodiments, the silencing vector targets intron 3 of human RCAN1 and the first targeting sequence comprises a sequence obtained by performing PCR with a primer pair of: Human Chr 21 RCAN1 left arm primers: forward 5'- ATT GTATAC CCAAGAGCCC TCCTGACCTC -3' (SEQ ID NO:5), reverse 5'- AATGTATACGGGTGGAGGGGCGTGATGCA-
20 3' (SEQ ID NO:6); and the second targeting sequence comprises a sequence obtained by performing PCR with a primer pair of: RCAN1 right arm primers: forward 5'- TAT TCGCGA CC CGCAGTGTCC CAGGAAT -3' (SEQ ID NO:7), reverse 5'-CGCTCGCGACAATGTTTTCAGAAATGTAA-3' (SEQ ID NO:8).

In some embodiments, the silencing element comprises a human XIST cDNA or
25 functional fragment thereof.

In some embodiments, the silencing vector includes a selectable marker sequence, e.g., a selectable marker sequence is operably linked to a promoter.

In another aspect, provided herein are silencing vectors comprising the sequences shown in FIG. 10a, 10c-10j or SEQ ID NOs:14, 16, 17, 18, 19, 20, 21, 22, or 23.

30 In another aspect, the invention provides methods for reducing levels of expression of genes on Chromosome 21 in a cell, the method comprising contacting the cell with a silencing vector described herein, under conditions sufficient for the silencing vector to undergo homologous recombination with the genomic DNA of the cell, wherein the silencing element is inserted into intron 1 of DYRK1A or intron 3 of RCAN1.

In some embodiments, the cell is trisomic for chromosome 21.

In some embodiments, the cell is a human cell.

In some embodiments, the cell is a stem cell or a fibroblast.

In some embodiments, the stem cell is an induced pluripotent stem cell (iPSC), a hematopoietic stem cell, or a neural stem cell.

In another aspect, the invention provides cells produced by a method described herein.

In another aspect, the invention provides methods for reducing the risk of transient myeloproliferative disorder (TMD) in a subject who has Down Syndrome (Trisomy 21). The methods include obtaining a hematopoietic stem cell from the subject;
contacting the cell with a silencing vector described herein, under conditions sufficient for the silencing vector to undergo homologous recombination with the genomic DNA of the cell, wherein the silencing element is inserted into intron 1 of DYRK1A or intron 3 of RCAN1, to produce a modified cell having reduced levels of expression of genes on Chromosome 21;and administering the modified cell to the subject.

In some embodiments, the methods include contacting the cell with a cleavage vector comprising a sequence that enhances or facilitates homologous recombination.

In some embodiments, the cleavage vector comprises a zinc finger nuclease (ZFN) or a transcription activator-like effector nuclease (TALEN).

In some embodiments, the cleavage vector targets a sequence in intron 1 of DYRK1A comprising GCCACCCCTTTGCCAGTTTACACGGGTGATGAGCA GGCTGTT (SEQ ID NO:9).

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Methods and materials are described herein for use in the present invention; other suitable methods and materials known in the art can also be used. The materials, methods, and examples are illustrative only and not intended to be limiting. All publications, patent applications, patents, sequences, database entries, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control.

Other features and advantages of the invention will be apparent from the following detailed description and figures, and from the claims.

## DESCRIPTION OF DRAWINGS

The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawing(s) will be provided by the Office upon request and payment of the necessary fee.

5 **Figure 1. Schematic outline of the trisomy 21 dosage compensation concept**. The natural dosage compensation mechanism (XIST non-coding RNA) is redirected, using ZFN technology, to one trisomic chromosome in iPS cells derived from Down syndrome patient fibroblasts. Subsequent expression of XIST RNA initiates chromosome-wide silencing of the targeted Chr21, producing a stable heterochromatic Chr21-Barr body, and correcting trisomy 10 21 to functional disomy.

**Figures 2A-F. Genome-editing integrates XIST into Chr 21 in trisomic iPS cells derived from DS patient cells.** Both constructs in a and b were introduced simultaneously, to achieve dual-targeted-addition to two loci, in one step. **a**, Top, XIST transgene construct (19 kb) with two homologous Chr 21 arms and 14kb XIST cDNA driven by inducible 15 pTRE3G promoter. The ZFNs cleave intron 1 or intron 3 of DYRK1A locus on Chr 21. Bottom, Four DYRK1A variants. **b**, Construct designed to target a puromycin selection gene and rtTA cassette into the AAVS1 safe harbor locus on Chr 19 by ZFN. **c**. A high resolution G-band karyotype was performed to further verify genome integrity of these subclones. Only Chr21 trisomy was observed, and karyotype is consistent with a male 20 chromosome complement. **d**. Genomic Microarray analysis using the UMass Genomic Microarray platform (Human Genome Build hg19) demonstrated a gain of one chromosome 21(red arrow) (and detected addition of the XIST transgene in these male cells). All other peaks are known human polymorphic variants and are not clinically significant. Note: Chr21 is increased 1.5 fold (from 2 to 3 chromosomes) while the XIST gene is increased 2 fold 25 (from 1 to 2 copies). **e**. Close-up of Chr21 CGH shows full chromosome 21 trisomy with no deletions or duplications. This analysis was done on transgenic clone 3. **f.** Percent of cells showing an XIST "paint" (a large, well-localized nuclear RNA territory), in six independent clones. Mean ± SE from 500 nuclei.

**Figure 3. XIST expression induces a cascade of heterochromatin modifications.** 30 Percentage of XIST territories with associated hallmarks H3K27me3, UbH2A, H4K20me, and MacroH2A. Mean ± SE from 100 nuclei in five or more colonies.

**Figures 4A-F. XIST expression induces long-range transcriptional silencing in transgenic, pluripotent sub-clones. a.** Quantification of APP silencing after 5 d and 20 d Dox induction in two independent sub-clones. Mean ± SE from 100 nuclei. **b**. Four more

Chr21-linked genes were also shown to be effectively silenced by RNA FISH, and scored before and after XIST induction. **c.** The silenced genes assessed by RNA FISH spanned the entire length of Chr21 (USP25 gene is ~21Mb from XIST integration site at DYRK1A; black arrow), suggesting long range silencing of Chr21 by XIST RNA. Mean ± SE from 100 nuclei.
5 **d**. Sequencing analysis of gene transcripts informative for SNPs indicates one of three alleles are silenced by XIST expression. Primer pairs were used that amplified SNP-containing regions of four Chr21 genes to assess allele-specific silencing after sequencing. RNA was amplified from Dox treated and untreated samples of three different clones. Eight of 12 SNPs tested were informative in these cells, and all eight SNPs (in four genes) show reduction in
10 one of the three alleles upon XIST induction. For example, in Clone 3 ADAMTS1 goes from TTC to TT, ETS2 from CCA to CA, TIAM1 from TTC to TC, and HSPA13 from TTC to TT. **e**. In two of the three transgenic clones, the same eight SNP alleles were repressed, consistent with a chromosome-wide mechanism and allows us to extrapolate the haplotype of each chromosome and surmise which carried XIST in each clone. Both clones 2 and 3 silence the
15 far right chromosome and the center chromosome is silenced in Clone 1. **f.** Although XIST RNA is robustly expressed in early time points (3 days) in the double and triple targeted clones, XIST becomes almost entirely silenced in later time points (20 days). Mean ± SE from 100 nuclei.

**Figures 5A-D. Genomic expression profiling with RNA microarray and**
20 **methylation levels shows widespread silencing of genes across Chr21. a.** Microarray: Gene expression of parallel cultures grown with and without Dox treatment (Dox/no Dox) in clones 1,2 &3 was compared to normal male iPS (euploid) and trisomy 21 (untargeted parent line) iPS cells (disomic/trisomic above). Total change in gene expression (N=3) per chromosome shows correction to disomic levels for XIST-expressing sub-clones on Chr21
25 with only limited changes on other chromosomes. Right Y-axis is scaled to reflect percent of gene expression change **b.** Distribution of individual repressed genes across Chr21 and corresponding level of repression for Clone 3 (Dox/no Dox) and Disomic/Trisomic. **c.** Methylation analysis: Genes with CpG island promoters are colored based on the levels of methylation after 22 days of Dox induction. Grey: decrease in methylation, green: no change
30 in methylation, and red: increase in methylation. Ideograms (shown to the left of each heatmap) denote the location of genes (note: no gene probes unique to short arm of Chrs 21 and 22). Length of each chromosome is proportional to the number of gene promoters with CpG islands. Of the 143 individual Chr21 genes that had CpG islands in their promoters, 97-98% in both clones increased methylation by at least 5% (approximately two-fold increase

over the average), compared to none in the parent line. **d**. Relative expression levels of eight genes on Chr21 by qRT-PCR for Clone 3. All 8 genes showed repression. Mean ± SE from triplicate samples.

**Figures 6A-B. "Trisomy correction" *in vitro* has marked effects on cell**
5 **proliferation and neurogenesis, and is stable upon removal of XIST RNA. a.** Changes in cell number for parent line, non-transgenic DS subclone (negative control), and six transgenic lines after 1 week of +/- Dox treatment. Mean ± SE. (n = 4-6). **b**. Quantification of number of neural rosettes at days 14 and 17 for two clones. Mean ± SE from 10-12 random fields in triplicate. **c**. Gene silencing is stable following withdrawal of XIST RNA in cortical neurons.
10 Transgenic cells were treated with Dox for 70 days and then Dox was removed for 30 days. Only two APP RNA transcription foci are present, as seen, with or without Dox. Mean ± SE from 100 nuclei.

**Figures 7A-C. RCAN1 targeting constructs. a.** Schematic and **b.** plasmid map show that the 21.1 kb selectable and inducible human XIST construct contains two
15 homologous arms (left arm, 759 bp; right arm, 758 bp), a hygromycin selection gene, and a 14 kb full length XIST cDNA driven by a tetracycline operator inducible promoter. The specifically designed ZFN cuts the intron 3 of RCAN1 gene on Chr 21. **c.** Plasmid map showing the 14.0 kb selectable and inducible human XIST construct contains two homologous arms (left arm, 759 bp; right arm, 758 bp), a hygromycin selection gene, and a
20 6.8 kb exon 1 of human XIST cDNA driven by a tetracycline operator inducible promoter. The 6.8 kb XIST transgene is targeted the RCAN1 gene on Chr 21 by ZFNs (as shown in schematic of Fig. 7a).

**Figures 8A-F. ZFN Targeting XIST to RCAN1 a.** Cells containing the FL XIST transgene construct shown in Figs. 7a-b targeted to RCAN1 locus on chr. 21 at metaphase.
25 Graph shows quantification of cells containing different integrations. N=145 cells. **b-c,** Cells containing the FL XIST transgene targeted to RCAN1 locus on chr. 21 at interphase. **b,** Graph shows quantification of cells containing different integrations. **c**, Graph shows quantification of cells containing different localizations of XIST RNA. **d-f.** Cells containing the RCAN1 6.8kb XIST transgene at interphase. **d,** Graph shows quantification of cells
30 containing different integrations. **e**, Graph shows quantification of cells containing different integrations in the cells expressing XIST RNA. **f,** Graph shows quantification of cells containing different localizations of XIST RNA. The Distance between chr.21 BAC and RCAN1 gene is 2.3 M b.

**Figures 9A-I show schematic illustrations of some of the constructs used in the present application.**

9a: 3G/FL/hXIST/DYRK1A. The plasmid map show the 18.5 kb inducible human XIST construct consists of two homologous arms (left arm, 690 bp; right arm, 508 bp), and a large XIST cDNA driven by an inducible pTRE3G promoter. The 14 kb XIST cDNA contains exons 1-5 and two fragments of exon 6 of XIST gene. The insert is 15.4 kb. The specifically designed ZFN cleaves the intron 1 of DYRK1A locus on Chr 21 (as shown in schematic of Fig.2a).

9b: puro/rtTA/AAVS1. The plasmid map shows the puro/rtTA construct contains both puromycin (puro) and tetracycline transactivator (rtTA) cassettes with opposite direction. rtTA is driven by a 3G EF1α promoter that is not inactivated in hESCs and hiPSCs. The puro and rtTA plasmid is targeted the AAVS1 locus on Chr 19 by ZFN.

9c: FL/hXIST/DYRK1A. The plasmid map shows that the 20.7 kb selectable and inducible human XIST construct contains two homologous arms (left arm, 690 bp; right arm, 508 bp), a hygromycin selection gene, and a 14 kb full length XIST cDNA driven by a tetracycline operator inducible promoter. The large XIST transgene is targeted the DYRK1A gene on Chr 21 by ZFNs (as shown in schematic of Fig. 2a).

9d: 6.8 kb/hXIST/DYRK1A. The plasmid map shows that the 13.7 kb selectable and inducible human XIST construct contains two homologous arms (left arm, 690 bp; right arm, 508 bp), a hygromycin selection gene, and a 6.8 kb exon 1 of human XIST cDNA driven by a tetracycline operator inducible promoter. The 6.8 kb XIST transgene is targeted the DYRK1A gene on Chr 21 by ZFNs (as shown in schematic of Fig. 2a).

9e. 6.8 kb/hXIST/AAVS1. The plasmid map shows the 15.7 kb selectable and inducible human XIST construct contains two homologous arms (800 bp each arm), a hygromycin selection gene, and a 6.8 kb exon 1 of human XIST cDNA driven by a tetracycline operator inducible promoter. The 6.8 kb XIST transgene is targeted the AAVS1 locus on Chr 19 by ZFNs.

9f: 6.3 kb/mXist/Runx1. The plasmid map shows that the 20.6 kb selectable and inducible mouse Xist construct contains two homologous arms (4 kb each arm), a hygromycin selection gene, and a 6.3 kb exon 1 of mouse Xist cDNA driven by a tetracycline operator inducible promoter. This 6.3 kb mouse Xist transgene is targeted the Runx1 gene on Chr 16 (synteny to human Chr 21) by conventional homologous recombination.

9g: pEF1α/hDYRK1A/FL mXist. The plasmid map shows that the 20.6 kb construct contains full length mouse Xist cDNA, an ampicillin resistance gene, and two homologous arms that target intron 1 of DYRK1A gene on human chromosome 21.

9h: pEF1α/hDYRK1A/6.3kb mXist. The plasmid map shows that the 12.2 kb construct contains 6.3 kb of mouse Xist cDNA that has been reported to function (Wutz et al., Nat Genet 30, 167-174 (2002)) and two homologous arms that target intron 1 of DYRK1A gene on human chromosome 21.

9i: Rosa26/pEF1x-Tet3G/hPGK-PuroR. The plasmid map shows that the 10.4 kb construct contains a puromycin resistance selection gene and rtTA cassette that is targeted to the Rosa26 locus on mouse chromosome 6 by ZFNs.

**Figures 10A-K** set forth some of the characteristics of the sequences of constructs 1-11, respectively, which are described more fully below.

## DETAILED DESCRIPTION

Nature has evolved a mechanism to dosage compensate the difference in X-linked gene copies between mammalian females (XX) and males (XY)[8]. This process is driven by a large (~17 kb in human) non-coding RNA, XIST, which during early development is produced exclusively from the *inactive* X (Xi)[9], and "paints" (accumulates across) the interphase chromosome structure[10]. The RNA induces a cascade of heterochromatin modifications and architectural changes which transcriptionally silence the Xi and manifest cytologically as a condensed Barr Body (reviewed in[11-14]). There is some DNA sequence specificity to XIST function, since many human genes escape X-inactivation[15-18]; however, autosomal chromatin has substantial capacity to be silenced[19-22]. The full potential of an autosome to be silenced, however, needs to be examined under conditions that avoid creation of a deleterious functional monosomy. The strategy pursued here meets that requirement and creates a tractable model to study the distinct biology of *human* chromosome inactivation[21].

As demonstrated herein (see Fig. 1), the present constructs and methods can be used to reroute the human X-chromosome inactivation machinery to a supernumerary Chr 21 in DS cells, and thereby enact its epigenetic transformation, in an controlled fashion. The approach directs, e.g., via zinc finger nuclease-driven targeted gene addition[23], a functional, inducible XIST transgene precisely to the gene-rich core of a trisomic chromosome 21 in induced pluripotent stem cells (iPSCs) derived from a subject with DS. The present results demonstrate (i) an unprecedented efficiency and precision of this addition using the largest transgene used for such an effort to date; (ii) the on-demand heterochromatinization of the

extra Chr 21 by numerous histone modifications and DNA condensation; (iii) long-range essentially uniform transcriptional repression as gauged by *in situ* analyses, genome-wide expression profiling, and CpG promoter methylation status; and finally, (iv) inducible trisomy silencing *in vitro* can also correct a deficit in proliferation of DS stem cells as revealed in this study. Thus, these findings establish a unique system to study DS-related cellular pathologies in a developmental cell context, as well as investigate the initiation of epigenetic chromosome silencing and its relationship to genomic sequence context. In addition, as the present methods result in at least partial correction of the chromosomal imbalance in DS, the methods enable a combined genetic/epigenetic approach to "chromosome therapy" for DS as well as cell therapies using autologous cells.

Unlike random integration into a diploid cell, silencing a trisomic autosome avoids selection against full autosomal silencing and monosomy. Thus, comprehensive analysis demonstrates highly robust competence of Chr21 to be silenced, allowing dosage compensation of trisomy 21 to very near normal disomic levels. This suggests that an RNA evolved for the X-chromosome utilizes epigenome-wide mechanisms. The ability to insert a single XIST transgene in any locus, in multiple isogenic sub-clones, now provides a powerful tool to further study XIST function. The present effort also has almost tripled the size of transgenes compatible with nuclease-driven targeted gene addition, important for a host of other compelling applications that require large sequence insertions.

From a translational perspective, trisomy silencing has immediate impact as a means to define the poorly understood cellular pathways deregulated in DS. Accomplishing this in DS iPSCs provides a means to derive and study various patient-compatible cell-types potentially relevant to DS therapeutics (e.g., hematopoietic, cardiac, neuronal, endocrine, and immune). Inducible "trisomy silencing in a dish" allows discrimination of differences directly due to Chr21 over-expression apart from genetic and epigenetic differences between transgenic sub-clones or rare disomic sub-clones isolated from a trisomic population ([48,49] and this study). Induced XIST expression triggers not only global Chr21 repression, but a defined effect on the genomic expression profile, and, importantly, impacts two major aspects of cell phenotype. This can illuminate the cohort of genes and cognate pathways most consistently impacted in DS, and thus define targets for translational efforts. Our discovery that Chr21 over-expression is linked to a reversible deficit in cell proliferation, and also neural rosette formation, is significant, particularly given that DS individuals show accelerated aging and hypocellularity in certain regions of the brain[42,43]. Understanding the pathways and pathologies of DS will also inform the search for drugs that may rebalance those pathways,

and the impact of whole chromosome silencing can be a benchmark to compare the impact of correcting individual genes (e.g. DYRK1A) to disomy. This general strategy can similarly be extended to study other chromosomal disorders, such as trisomy 13 and 18, so often fatal in the first 1-2 years.

5 Finally, the present methods and compositions can be used for gene therapy to address whole chromosome imbalance.

*Nucleic acid constructs – Silencing Vectors*

Described herein are silencing vectors, nucleic acid constructs that include a silencing sequence driven by a regulatory region comprising a promoter, and one or more targeting 10 sequences (*e.g.*, first and second sequences that flank the silencing sequence and direct insertion of the silencing sequence into a targeted chromosome). The silencing vectors can also include a selectable marker, driven by the same or, more preferably, a different regulatory region.

*XIST Silencing Sequences*

15 In the present application, the term "Xist" refers to an Xist gene or the encoded Xist RNA regardless of the origin of the sequence. For example, the present compositions can include, and the present methods can be carried out with, an Xist gene encoding an Xist RNA from humans or another mammal (e.g., a rodent such as a mouse, dog, cat, cow, horse, sheep, goat, or another mammalian or non-mammalian animal). The scientific literature has adopted 20 a loose convention whereby the term is fully capitalized (XIST) when referring to a human sequence but not fully capitalized (Xist) when referring to the murine sequence. That convention is not used here, and either human or non-human sequences may be used as described herein.

The silencing sequence can be a full-length Xist gene sequence, a full-length Xist 25 cDNA, or any biologically active fragment or other biologically active variant thereof. The sequence is "biologically active" where its activity is sufficient to silence the expression of one or more genes in cis when integrated into chromosome 21. The level of activity of a biologically active fragment or other variant may vary so long as a useful chromosomal silencing RNA is produced. Xist RNA is referred to as a chromosomal silencing RNA 30 because it silences by binding across the chromosome or chromosome segment, and therefore silences at the level of transcription, by inducing repressive changes to chromatin. While Xist RNA is a well-studied example of a chromosomal silencing RNA, other non-coding

RNAs can silence specific clusters of imprinted genes or segments of a chromosome, and in some embodiments a sequence encoding another full-length silencing RNA (examples of which are provided below) or biologically active fragment can be used in place of XIST. These other chromosomal silencing RNAs include Air RNA, HOTAIR RNA, and Kcnq1ot1 RNA (see Goodrich and Kugel, Crit. Rev. Biochem. and Mol. Biol. 44:3-15, 2009), any of which can be formulated and used as described herein for Xist. Other intergenic noncoding RNAs, which may be useful in the present nucleic acid constructs and the silencing methods described herein are described by Khalil et al. (Proc. Natl. Acad. Sci. USA 106:11675-11680, 2009).

The silencing vector further includes at least one regulatory sequence (i.e., a regulatory sequence that promotes expression of the Xist RNA, and a regulatory sequence that promotes expression of a selectable marker, if any). More specifically, the regulatory sequence can include a promoter, which may be constitutively active, inducible, tissue-specific, or a developmental stage-specific promoter. Enhancers and polyadenylation sequences can also be included. For example, the Xist transgene may carry one or more regulatory elements found in the Xic region that are not a part of the Xist coding sequence. For example, deletion of the DXPas34 locus found 3' to the Xist coding sequence eliminates Xist expression in mammalian embryonic stem cells as described in Debrand et al. (Mol. Cell. Bio., 19:8513-8525, 1999) herein incorporated by reference. As a further example, silencing by mouse Xist transgenes have been shown to require a conserved repeat sequence located at the 5' end of Xist (Wutz et al., Nat. Genetics, 30:167-174, 2002).

The silencing sequence can exclude one or more introns (wholly or partially) or one or more exons (wholly or partially). However, the silencing sequence cannot exclude all exons. For example, the silencing sequence can be an Xist gene sequence exclusive of one or more introns or one or more exons (but not all exons). For example, the silencing sequence can include about 6 kb to about 10 kb of exon 1 of an Xist gene sequence (e.g., about 6-7 kb, 7-8 kb, 8-9 kb, 6.5-8.5 kb, or about 7.5 kb). More specifically, the silencing sequence can be or can include the full length human Xist cDNA sequence having accession number M97168.1 or a biologically active fragment or other variant thereof, e.g., the full length XIST shown in SEQ ID NO:10, or the variant shown in SEQ ID NO:11.

The Xist transgene need not include the whole of the Xist gene sequence, although it may. For example, the Xist transgene may be derived from an Xist cDNA cloned from one of multiple naturally occurring splice variants. This cDNA may lack sequences corresponding to one or more introns or exons or portions thereof. Additionally, the Xist transgene may

include non-naturally occurring Xist coding sequences. For example, the Xist coding sequence may be mutated (e.g., truncated) or otherwise variant with respect to naturally occurring Xist coding sequences so long as it includes sequences that are required for transgene function. For example, deletion analysis demonstrates that the first exon of human Xist is sufficient for both transcript localization and the induction of silencing (Chow et al., Proc. Natl. Acad. Sci. USA 104:10104-10109, 2007). Thus, smaller Xist constructs can be generated that are more easily manipulated but still biologically active.

Non-limiting examples of Xist transgenes (derived from mouse and human sequences) that are useful in this invention are described in the following references which are herein incorporated by reference: Chow et al. (Proc. Natl. Acad. Sci. USA 104:10104-10109, 2007); Hall et al. (Proc. Natl. Acad. Sci. USA 99:8677-8682, 2002); Chow et al. (Genomics, 82:309-322, 2003); and Wutz et al. (Nat. Genet., 2002, 30:167-174, 2002).

Integrated Mouse Xist or human Xist transgenes can silence an autosome, as shown by studies in mouse embryonic stem cells (Wutz and Jaenisch, Mol. Cell, 5:695-705, 2000; Savarese et al., Mol. Cell Biol. 26:7167-7177, 2006) and in human somatic (fibrosarcoma) cells (FIG. 3; Hall et al., Hum. Mol. Genet. 11:3157-3165, 2002; Chow et al., Proc. Natl. Acad. Sci. USA 104:10104-10109, 2007). Natural autosomal silencing by Xist was also shown in patient cells, with an autosomal trisomy due to X;autosome translocations (Hall et al., Proc. Natl. Acad. Sci. USA 99:8677-8682, 2002; (FIG. 4)). Although the silencing of autosomal material may not be quite as complete or may vary somewhat between autosomal regions, autosomes studied to date are largely if not entirely silenced in response to Xist RNA.

The silencing sequence can be or can include the sequence of an XIC (X inactivation complex) locus or any portion thereof that encodes an RNA capable of silencing the chromosome into which it has been inserted. For example, the constructs can include an XIC locus lacking the sequences 3' to Xist that trigger the "counting" mechanism. Other constructs can include the Xist gene, with or without some or all of the intronic sequences, or a biologically active variant of the Xist gene (e.g., a fragment or other mutant). For information regarding the structure of XIC, one can consult Wutz and Gribnau (Curr. Opin. Genetics Dev. 17:387-393, 2007).

In some embodiments, the silencing sequence comprises one of the following:

Full length human XIST sequence – SEQ ID NO:10

```
CTAGAACATTTTCTAGTCCCCCAACACCCTTTATGGCGTATTTCTTTAAAAAAATCACCTAAATTCCA
TAAAATATTTTTTTAAATTCTATACTTTCTCCTAGTGTCTTCTTGACACGTCCTCCATATTTTTTTAA
AGAAAGTATTTGGAATATTTTGAGGCAATTTTTAATATTTAAGGAATTTTTCTTTGGAATCATTTTTG
```

Full length human XIST sequence – SEQ ID NO:10

GTGACATCTCTGTTTTTTGTGGATCAGTTTTTTACTCTTCCACTCTCTTTTCTATATTTTGCCCATCG
GGGCTGCGGATACCTGGTTTTATTATTTTTTCTTTGCCCAACGGGGCCGTGGATACCTGCCTTTTAAT
TCTTTTTTATTCGCCCATCGGGGCCGCGGATACCTGCTTTTTATTTTTTTTTCCTTAGCCCATCGGGG
TATCGGATACCTGCTGATTCCCTTCCCCTCTGAACCCCCAACACTCTGGCCCATCGGGGTGACGGATA
TCTGCTTTTTAAAAATTTTCTTTTTTTGGCCCATCGGGGCTTCGGATACCTGCTTTTTTTTTTTTTAT
TTTCCTTGCCCATCGGGGCCTCGGATACCTGCTTTAATTTTTGTTTTTCTGCCCATCGGGGCCGCGGA
TACCTGCTTTGATTTTTTTTTTTCATCGCCCATCGGTGCTTTTTATGGATGAAAAAATGTTGGTTTTG
TGGGTTGTTGCACTCTCTGGAATATCTACACTTTTTTTTGCTGCTGATCATTTGGTGGTGTGTGAGTG
TACCTACCGCTTTGGCAGAGAATGACTCTGCAGTTAAGCTAAGGGCGTGTTCAGATTGTGGAGGAAAA
GTGGCCGCCATTTTAGACTTGCCGCATAACTCGGCTTAGGGCTAGTCGTTTGTGCTAAGTTAAACTAG
GGAGGCAAGATGGATGATAGCAGGTCAGGCAGAGGAAGTCATGTGCATTGCATGAGCTAAACCTATCT
GAATGAATTGATTTGGGGCTTGTTAGGAGCTTTGCGTGATTGTTGTATCGGGAGGCAGTAAGAATCAT
CTTTTATCAGTACAAGGGACTAGTTAAAAATGGAAGGTTAGGAAAGACTAAGGTGCAGGGCTTAAAAT
GGCGATTTTGACATTGCGGCATTGCTCAGCATGGCGGGCTGTGCTTTGTTAGGTTGTCCAAAATGGCG
GATCCAGTTCTGTCGCAGTGTTCAAGTGGCGGGAAGGCCACATCATGATGGGCGAGGCTTTGTTAAGT
GGTTAGCATGGTGGTGGACATGTGCGGTCACACAGGAAAAGATGGCGGCTGAAGGTCTTGCCGCAGTG
TAAAACATGGCGGGCCTCTTTGTCTTTGCTGTGTGCTTTTCGTGTTGGGTTTTGCCGCAGGGACAATA
TGGCAGGCGTTGTCATATGTATATCATGGCTTTTGTCACGTGGACATCATGGCGGGCTTGCCGCATTG
TTAAAGATGGCGGGTTTTGCCGCCTAGTGCCACGCAGAGCGGGAGAAAAGGTGGGATGGACAGTGCTG
GATTGCTGCATAACCCAACCAATTAGAAATGGGGGTGGAATTGATCACAGCCAATTAGAGCAGAAGAT
GGAATTAGACTGATGACACACTGTCCAGCTACTCAGCGAAGACCTGGGTGAATTAGCATGGCACTTCG
CAGCTGTCTTTAGCCAGTCAGGAGAAAGAAGTGGAGGGGCCACGTGTATGTCTCCCAGTGGGCGGTAC
ACCAGGTGTTTTCAAGGTCTTTTCAAGGACATTTAGCCTTTCCACCTCTGTCCCCTCTTATTTGTCCC
CTCCTGTCCAGTGCTGCCTCTTGCAGTGCTGGATATCTGGCTGTGTGGTCTGAACCTCCCTCCATTCC
TCTGTATTGGTGCCTCACCTAAGGCTAAGTATACCTCCCCCCCCACCCCCCAACCCCCCCAACTCCCC
ACCCCCACCCCCCACCCCCCACCTCCCCACCCCCCTACCCCCCTACCCCCCTACCCCCCTCTGGTCTG
CCCTGCACTGCACTGTTGCCATGGGCAGTGCTCCAGGCCTGCTTGGTGTGGACATGGTGGTGAGCCGT
GGCAAGGACCAGAATGGATCACAGATGATCGTTGGCCAACAGGTGGCAGAAGAGGAATTCCTGCCTTC
CTCAAGAGGAACACCTACCCCTTGGCTAATGCTGGGGTCGGATTTTGATTTATATTTATCTTTTGGAT
GTCAGTCATACAGTCTGATTTTGTGGTTTGCTAGTGTTTGAATTTAAGTCTTAAGTGACTATTATAGA
AATGTATTAAGAGGCTTTATTTGTAGAATTCACTTTAATTACATTTAATGAGTTTTTGTTTTGAGTTC
CTTAAAATTCCTTAAAGTTTTTAGCTTCTCATTACAAATTCCTTAACCTTTTTTTGGCAGTAGATAGT
CAAAGTCAAATCATTTCTAATGTTTTAAAAATGTGCTGGTCATTTTCTTTGAAATTGACTTAACTATT
TTCCTTTGAAGAGTCTGTAGCACAGAAACAGTAAAAAATTTAACTTCATGACCTAATGTAAAAAAGAG
TGTTTGAAGGTTTACACAGGTCCAGGCCTTGCTTTGTTCCCATCCTTGATGCTGCACTAATTGACTAA
TCACCTACTTATCAGACAGGAAACTTGAATTGCTGTGGTCTGGTGTCCTCTATTCAGACTTATTATAT
TGGAGTATTTCAATTTTTCGTTGTATCCTGCCTGCCTAGCATCCAGTTCCTCCCCAGCCCTGCTCCCA
GCAAACCCCTAGTCTAGCCCCAGCCCTACTCCCACCCGGCCCCAGCCCTGCCCCAGGCCCAGTCCCCT
AACCCCCCAGCCCTAGGCCCAGTCCCAGTCCTAGTTCCTCAGTCTGTCCAGCTTCTCTCGAAAGTCAC
TCTAATTTTCATTGATTCAGTGCTCAAAATAAGTTGTCCATTGGTATCCTATTATACTGGGATATTCC
GTTTACCCTTGGCATTGCTGATCTTCAGTACTGACTCCTTGACCATTTTCAGTTAAGCATACAATCCC
ATTTGTCTGTGATCTCAGGACAAAGAATTTCCTTACTCGGTACGTTGAAGTTAGGGAATGTCAATTGA
GAGCTTTCTATCAGAGCATTATTGCCCACAATTTGAGTTACTTATCATTTTCTCGATCCCCTGCCCTT
AAAGGAGAAACCATTTCTCTGTCATTGCTTCTGTAGTCACAGTCCCAATTTTGAGTAGTGATCTTTTC
TTGTGTACTGTGTTGGCCACCTAAAACTCTTTGCATTGAGTAAAATTCTAATTGCCAATAATCCTACC
CATTGGATTAGACAGCACTCTGAACCCCATTTGCATTCAGCAGGGGGTCGCAGACAACCCGTCTTTTG
TTGGACAGTTAAAATGCTCAGTCCCAATTGTCATAGCTTTGCCTATTAAACAAAGGCACCCTACTGCG
CTTTTTGCTGTGCTTCTGGAGAATCCTGCTGTTCTTGGACAATTAAAGAACAAAGTAGTAATTGCTAA
TTGTCTCACCCATTAATCATGAAGACTACCAGTCGCCCTTGCATTTGCCTTGAGGCAGCGCTGACTAC
CTGAGATTTAAGAGTTTCTTAAATTATTGAGTAAAATCCCAATTATCCATAGTTCTGTTAGTTACACT
ATGGCCTTTGCAAACATCTTTGCATAACAGCAGTGGGACTGACTCATTCTTAGAGCCCCTTCCCTTGG
AATATTAATGGATACAATAGTAATTATTCATGGTTCTGCGTAACAGAGAAGACCCACTTATGTGTATG
CCTTTATCATTGCTCCTAGATAGTGTGAACTACCTACCACCTTGCATTAATATGTAAAACACTAATTG
CCCATAGTCCCACTCATTAGTCTAGGATGTCCTCTTTGCCATTGCTGCTGAGTTCTGACTACCCAAGT

Full length human XIST sequence – SEQ ID NO:10

TTCCTTCTCTTAAACAGTTGATATGCATAATTGCATATATTCATGGTTCTGTGCAATAAAAATGGATT
CTCACCCCATCCCACCTTCTGTGGGATGTTGCTAACGAGTGCAGATTATTCAATAACAGCTCTTGAAC
AGTTAATTTGCACAGTTGCAATTGTCCAGAGTCCTGTCCATTAGAAAGGGACTCTGTATCCTATTTGC
ACGCTACAATGTGGGCTGATCACCCAAGGACTCTTCTTGTGCATTGATGTTCATAATTGTATTTGTCC
ACGATCTTGTGCACTAACCCTTCCACTCCCTTTGTATTCCAGCAGGGGACCCTTACTACTCAAGACCT
CTGTACTAGGACAGTTTATGTGCACAATCCTAATTGATTAGAACTGAGTCTTTTATATCAAGGTCCCT
GCATCATCTTTGCTTTACATCAAGAGGGTGCTGGTTACCTAATGCCCCTCCTCCAGAAATTATTGATG
TGCAAAATGCAATTTCCCTATCTGCTGTTAGTCTGGGGTCTCATCCCCTCATATTCCTTTTGTCTTAC
AGCAGGGGGTACTTGGGACTGTTAATGCGCATAATTGCAATTATGGTCTTTTCCATTAAATTAAGATC
CCAACTGCTCACACCCTCTTAGCATTACAGTAGAGGGTGCTAATCACAAGGACATTTCTTTTGTACTG
TTAATGTGCTACTTGCATTTGTCCCTCTTCCTGTGCACTAAAGACCCCACTCACTTCCCTAGTGTTCA
GCAGTGGATGACCTCTAGTCAAGACCTTTGCACTAGGATAGTTAATGTGAACCATGGCAACTGATCAC
AACAATGTCTTTCAGATCAGATCCATTTTATCCTCCTTGTTTTACAGCAAGGGATATTAATTACCTAT
GTTACCTTTCCCTGGGACTATGAATGTGCAAAATTCCAATGTTCATGGTCTCTCCCTTTAAACCTATA
TTCTACCCCTTTTACATTATAGAAAGGGATGCTGGAAACCCAGAGTCCTTCTCTTGGGACTCTTAATG
TGTATTTCTAATTATCCATGACTCTTAATGTGCATATTTTCAATTGCCTAATTGATTTCAATTGTCTA
AGACATTTCAAATGTCTAATTGATTAGAACTGAGTCTTTTATATCAAGCTAATATCTAGCTTTTATAT
CAAGCTAATATCTTGACTTCTCAGCATCATAGAAGGGGGTACTGATTTCCTAAAGTCTTTCTTGAATT
TCTATTATGCAAAATTGCCCTGAGGCCGGGTGTGGTGGCTCACACCTGTAATCCCAGCACTTTGGGAG
GCTGAGGTGGGAAGATCCCTTACTGCCAGGAGTTTGAGACCAGCCTGGCCAACATTAAAAAAAAAAAA
AAAAGTAAGACAATTGCCCTGGAATCCCATCCCCCTCACACCTCCTTGGCAAAGCAGCAGGAGTGCTA
ACTAGCTAGTGCTTCTTCTCTTATACTGCTTAAATGCGCATAATTAGCAGTAGTTGATGTGCCCCTAT
GTTAGAGTAGAATCCCGCTTCCTTGCTCCATTTGCATTACTGCAGGAGCTTCTAACTAGCCTGAATTC
ACTCTCTTGGACTGTTAATGTGCATACTTATATTTGCTGCTGTACTTTTTTACCATGTAAGGACCCCA
CCCACTGTATTTACATCCCAGCTGGAAGTACCTACTACTTAAGACCCTTAGACTAGTAAAGTTAGCGT
GCATAATCTTAGGTGTTATATACACATTTTCAGTTGCATACAGTTGTGCCTTTTATCAGGACTCCTGT
ACTTATCAAAGCAGAGAGTGCTAATCAATATTAAGCCCTTCTCTTCGAACTGTAGATGGCATGTAATT
GCAGTTGTCAATGGTCCTTCAATTAGACTTGGGTTTCTGACCTATCACACCCTCTTTGCTTTATTGCA
TGGGGTACTATTCACTTAAGGCCCCTTTCTCAAACTGTTAATGTGCCTAATGACAATTACATCAGTAT
CCTTCCTTTTGAAGGACAGCATGGTTGGTGACACCTAAGGCCCCATTTCTTGGCCTCCCAATATGTGT
GATTGTATTTGTCGAGGTTGCTATGCACTAGAGAAGGAAAGTGCTCCCCTCATCCCCACTTTTCCCTT
CCAGCAGGAAGTGCCCACCCCATAAGACCCTTTTATTTGGAGAGTCTAGGTGCACAATTGTAAGTGAC
CACAAGCATGCATCTTGGACATTTATGTGCGTAATCGCACACTGCTCATTCCATGTGAATAAGGTCCT
ACTCTCCGACCCCTTTTGCAATACAGAAGGGTTGCTGATAACGCAGTCCCCTTTTCTTGGCATGTTGT
GTGTGATTATAATCGTCTGGGATCCTATGCACTAGAAAAGGAGGGTCCTCTCCACATACCTCAGTCTC
ACCTTTCCCTTCCAGCAGGGAGTGCCCACTCCATAAGACTCTCACATTTGGACAGTCAAGGTGCGTAA
TTGTTAAGTGAACACAACCATGCACCTTAGACATGGATTTGCATAACTACACACAGCTCAACCTATCT
GAATAAAATCCTACTCTCAGACCCCTTTTGCAGTACAGCAGGGGTGCTGATCACCAAGGCCCTTTTTC
CTGGCCTGGTATGCGTGTGATTATGTTTGTCCCGGTTCCTGTGTATTAGACATGGAAGCCTCCCCTGC
CACACTCCACCCCCAATCTTCCTTTCCCTTCCGGCAGGAGTGCCCTCTCCATAAGACGCTTACGTTTG
GACAATCAAGGTGCACAGTTGTAAGTGACCACAGGCATACACCTTGGACATTAATGTGCATAACCACT
TTGCCCATTCCATCTGAATAAGGTCCTACTCTCAGACCCCTTTTGCAGTACAGCAGGGGTGCTGATCA
CCAAGGCCCCTTTTCTTGGCCTGTTATGTGCGTGATTATATTTGTCTGGGTTCCTGTGTATTAGACAA
GGAAGCCTTCCCCCCGCCCCCACCCCCACTCCCAGTCTTCCTTTCCCTTCCAGCAGGGAGTGCCCCCT
CCATAAGATCATTACATTTGGACAATCAAGGTGCACAATTATAAGTGACCACAGCCATGCACCTTGGA
CATTATTGGACATTAATGTGCGTAACTGCACATGGCCCATCCCATCTGAATAAGGACCTACTCTCAGA
TGCCTTTGCAGTACAGCAGGGGTACTGAATCACCAAGGCCCTTTTTCTTGGCCTGTTATGTGTGTGAT
TATATTTATCCCAGTTTCTGTGTAATAGACATGAAAGCCTCCCCTGCCACACCCCACCTCCAATCTTC
CTTTCCCTTCCACCAGGGAGTGTCCACTCCATATACCCTTACATTTGGACAATCAAGGTGCACAATTG
TAAGTGAGCATAGGCACTCACCTTGGACATGAATGTGCATAACTGCACATGGCCCATCCCATCTGAAT
AAGGTCCTACTCTCAGACCCTTTTTGCAGTACAGCAGGGGTGCTGATCACCAAGGCCCCTTTTCCTGG
CCTGTTATGTGTGTGATTATATTTGTTCCAGTTCCTGTGTAATAGACATGGAAGCCTCCCCTGCCACA
CTCCACCCCCAATCTTCCTTTCCTTCTGGCAGGAAGTACCCGCTCCATAAGACCCTTACATTTGGACA
GTCAAGGTGCACAATTGTATGTGACCACAACCATGCACCTTGGACATAAATGTGTGTAACTGCACATG

Full length human XIST sequence – SEQ ID NO:10

GCCCATCCCATCTGAATAAGGTCCTACTCTCAGACCCCTTTTGCAGTACAGTAGGTGTGCTGATAACC
AAGGCCCCTCTTCCTGGCCTGTTAACGTATGTGATTATATTTGTCTGGGTTCCAGTGTATAAGACATG
GAAGCCTCCCCTGCCCCACCCCACCCTCAATCTTCCTTTCCCTTCTGGCAGGGAGTGCCAGCTCCATA
AGAACCTTACATTTGGACAGTCAAGGTGCACAATTCTAAGTGACCGCAGCCATGCACCTTGGTCAATA
ATGTGTGTAACTGCACACGGCCTATCTCATCTGAATAAGGCCTTACTCTCAGACCCCTTTTGCAGTAC
AGCAGGGGTGCTGATAACCAAGGCCCATTTTCCTGGCCTGTTATGTGTGTGATTATATTTGTCCAGGT
TTCTGTGTACTAGACAAGGAAGCCTCCTCTGCCCCATCCCATCTACGCATAATCTTTCTTTTCCTCCC
AGCAGGGAGTGCTCACTCCATAAGACCCTTACATTTGGACAATCAAGGTGCACAATTGTAAGTGACCA
CAACCATGCATCTTGGAAATTTATGTGCATAACTGCACATGGCTTATCCTATTTGAATAAAGTCCTAC
TCTCAGACCCCCTTTGCAGTATAGCTGGGGTGCTGATCACTGAGGCCTCTTTGCTTGGCTTGTCTATA
TTCTTGTGTACTAGATAAGGGCACCTTCTCATGGACTCCCTTTGCTTTTCAACAAGGAGTACCCACTA
CTTTTTAAGATTCTTATATTTGTCCAAAGTACATGGTTTTAATTGACCACAACAATGTCCCTTGGACA
TTAATGTATGTAATCACCACATGGTTCATCCTAATTAAACAAAGTTCTACCTTCTCACCCTCCATTTG
CAGTATACCAGGGTTGCTGACCCCCTAAGTCCCCTTTTCTTGGCTTGTTGACATGCATAATTGCATTT
ATGTTGGTTCTTGTGCCCTAGACAAGGATGCCCCACCTCTTTTCAATAGTGGGTGCCCACTCCTTATG
ATCTTTACATTTGAACAGTTAATGTGAATAATTGCAGTTGTCCACAACCCTATCACTTCTAGGACCAT
TATACCTCTTTTGCATTACTGTGGGGTATACTGTTTCCCTCCAAGGCCCCTTCTGGTGGACTATCAAC
ATATAATTGAAATTTTCTTTTGTCTTTGTCAGTAGATTAAGGTCATACCCCATCACCTTTCCTTTGTA
GTACAACAGGGTGTCCTGATCAACCAAAGTCCTGTTGTTTTGGACTGTTAATATGTGCAATTACATTT
GCTCCTGATCTGTGCACTAGATAAGGATCCTACCTACTTTCTTAGTGTTTTTAGCAGGTAGTGCCCAC
TACTCAAGACTGTCACTTGGAATGTTCATGTGCACAAACTCAATTCTCTAAGCATGTTCCTGTACCAC
CTTTGCTTTAGAGCAGGGGGATGATATTCACTAAGTGCCCCTTCTTTTGGACTTAATATGCATTAATG
CAATTGTCCACCTCTTCTTTTAGACTAAGAGTTGATCTCCACATATTCCCCTTGCATCAGGGGCATGT
TAATTATGAATGAACCCTTTTCTTTTAATATTAATGTCATAATTGTATTTGTGGACCTGTGTAGGAGA
AAAAGACCCTATGTTCCTCCCATTACCCTTTGGATTGCTGCTGAGAAGTGTTAACTACTCATAATCTC
AGCTCTTGGACAATTAATAGCATTAATAACAATTATCAAGGGCACTGATCATTAGATAAGACTCCTGC
TTCCTCGTTGCTTACATCGGGGGTACTGACCCACTAAGGCCCCTTGTACTGTTAATGTGAATATTTGC
AATTATATATGTCTCCTTCTGGTAGAGTGGGATATTATGCCCTAGTATCCCCTTTGCATTACTGCAGG
GGCTGCTGACTACTCAAAACTTCTCCTGGGACTGTTAATAGGCACAATGGCAGTTATCAATGGTTTTC
TCCCTCCCTGACCTTGTTAAGCAAGCGCCCCACCCCACCCTTAGTTTCCCATGGCATAATAAAGTATA
AGCATTGGAGTATTCCATGCACTTGTCTATCAAACAGTGGTCCATACTCCCAACCCTTTTGCATTGCG
CCAGTGTGTAAAATCACAGGTAGCCATGGTGTCATGCTTTATATACGAAGTCTTCCCTCTCTCTGCCC
CTTGTGTGCCCTTGGCCCCTTTTTACAGACTATTGCTCACAATCTCAGGTGTCCATATTTGCAGCTAT
TAGGTAAGATTGTGCTGTCTCCCTCTTCCCTTCCCTCTGCCCTGCCCCTTTTGCCTCTTTGCTGGGTA
ATGTTGACCAGACAAGGCCCTTTCTCTTGGACTTAAACAATTCTCAGTTGCACTTTCCTTGGTCCACC
CATTATACATGAACCCCTCTACTTCCTTTCGCATTGCTTCTGAGTATGCTGACTACCCAAAGCCCCTT
CTGTGTTATTAATAAACACAGTACTGATTGTCCCATTTTTCAGCCCATCAGTCCAAGATCTCCCTACC
ACTTTGGTGTGTTGGTGCAGTGTTGACTATGAAAAGCAGGCCTGAACTAGGTGGATAAGCCTTCACTC
ATTTTCTTTCATTTATTAATGATCCTAGTTTCAATTATTGTCAGATTCTGGGGACAAGAACCATTCTT
GCCCACCTGTGTTACTGCTTTACTGTGCAAAATACTGAAGGCAAGTCAGACCCAGGGAGCTGGATTGC
CATCCTTTATTTTGTGTTTCCAGTGTACACTATAAAATTGTCTCCCCAGGAAGGAAGGTTGGCACTTT
CTCTGCATTCTTCTTTCCAGAGCAGATTGCCTGGTTAAGAATCTCTTGTTGTCCCTTCTGTATATTGT
TATTGTAAAGTGCCAAATGCCAGGATACAGCCAGAAAAATTGCTTATTATTATTAAAAAAATTTTTTT
AAGAAAGACATCTGGATTGTAGGGTGGACTCGATAACCTGGTCATTATTTTTTTGAAGCCAAAATATC
CATTTATACTATGTACCTGGTGACCAGTGTCTCTCATTTTAACTGAGGGTGGTGGGTCTGTGGATAGA
ACACTGACTCTTGCTATTTTAATATCAAAGATATTCTAGATCCAGCACAGTGGCGGCCGCTCTAGAGT
GGAACTCTTAAGACCAGTATCTTTGTGTGGGCTTTACCAGCATTCACTTTTAGAAAAACTACCTAAAT
TTTATAATCCTTTAATTTCTTCATCTGGAGCACCTGCCCCTACTTATTTCAAGAAGATTGCAGTAAAA
CGATTAAATGAGGGAACATATGCAGAGGTGCTTTTAAAAAGCATATGCCACCTTTTTTATTAATTATT
ATATAAAATGAAGCATTTAATTATAGTAATAATTTGAAGTAGTTTGAAGTACCACACTGAGGTGAGGA
CTTAAAAATGATAAGACGAGTTCCCTATTTTATAAGAAAAATAAGCCAAAATTAAATATTCTTTTGGA
TATAAATTTCAACAGTGAGATAGCTGCCTAGTGGAAATGAATAATATCCCAGCCACTAGTGTACAGGG
TGTTTTGTGGCACAGGATTATGTAATATGGAACTGCTCAAGCAAATAACTAGTCATCACAACAGCAGT
TCTTTGTAATAACTGAAAAAGAATATTGTTTCTCGGAGAAGGATGTCAAAAGATCGGCCCAGCTCAGG

## Full length human XIST sequence – SEQ ID NO:10

GAGCAGTTTGCCCTACTAGCTCCTCGGACAGCTGTAAAGAAGAGTCTCTGGCTCTTTAGAATACTGAT
CCCATTGAAGATACCACGCTGCATGTGTCCTTAGTAGTCATGTCTCCTTAGGCTCCTCTTGGACATTC
TGAGCATGTGAGACCTGAGGACTGCAAACAGCTATAAGAGGCTCCAAATTAATCATATCTTTCCCTTT
GAGAATCTGGCCAAGCTCCAGCTAATCTACTTGGATGGGTTGCCAGCTATCTGGAGAAAAAGATCTTC
CTCAGAAGAATAGGCTTGTTGTTTTACAGTGTTAGTGATCCATTCCCTTTGACGATCCCTAGGTGGAG
ATGGGGCATGAGGATCCTCCAGGGGAAAAGCTCACTACCACTGGGCAACAACCCTAGGTCAGGAGGTT
CTGTCAAGATACTTTCCTGGTCCCAGATAGGAAGATAAAGTCTCAAAAACAACCACCACACGTCAAGC
TCTTCATTGTTCCTATCTGCCAAATCATTATACTTCCTACAAGCAGTGCAGAGAGCTGAGTCTTCAGC
AGGTCCAAGAAATTTGAACACACTGAAGGAAGTCAGCCTTCCCACCTGAAGATCAACATGCCTGGCAC
TCTAGCACTTGAGGATAGCTGAATGAATGTGTATTTCTTTGTCTCTTTCTTTCTTGTCTTTGCTCTTT
GTTCTCTATCTAAAGTGTGTCTTACCCATTTCCATGTTTCTCTTGCTAATTTCTTTCGTGTGTGCCTT
TGCCTCATTTTCTCTTTTTGTTCACAAGAGTGGTCTGTGTCTTGTCTTAGACATATCTCTCATTTTTC
ATTTTGTTGCTATTTCTCTTTGCTCTCCTAGATGTGGCTCTTCTTTCACGCTTTATTTCATGTCTCCT
TTTTGGGTCACATGCTGTGTGCTTTTTGTCCTTTTCTTGTTCTGTCTACCTCTCCTTTCTCTGCCTAC
CTCTCTTTTCTCTTTGTGAACTGTGATTATTTGTTACCCCTTCCCCTTCTCGTTCGTTTTAAATTTCA
CCTTTTTTCTGAGTCTGGCCTCCTTTCTGCTGTTTCTACTTTTTATCTCACATTTCTCATTTCTGCAT
TTCCTTTCTGCCTCTCTTGGGCTATTCTCTCTCTCCTCCCCTGCGTGCCTCAGCATCTCTTGCTGTTT
GTGATTTTCTATTTCAGTATTAATCTCTGTTGGCTTGTATTTGTTCTCTGCTTCTTCCCTTTCTACTC
ACCTTTGAGTATTTCAGCCTCTTCATGAATCTATCTCCCTCTCTTTGATTTCATGTAATCTCTCCTTA
AATATTTCTTTGCATATGTGGGCAAGTGTACGTGTGTGTGTGTCATGTGTGGCAGAGGGGCTTCCTAA
CCCCTGCCTGATAGGTGCAGAACGTCGGCTATCAGAGCAAGCATTGTGGAGCGGTTCCTTATGCCAGG
CTGCCATGTGAGATGATCCAAGACCAAAACAAGGCCCTAGACTGCAGTAAAACCCAGAACTCAAGTAG
GGCAGAAGGTGGAAGGCTCATATGGATAGAAGGCCCAAAGTATAAGACAGATGGTTTGAGACTTGAGA
CCCGAGGACTAAGATGGAAAGCCCATGTTCCAAGATAGATAGAAGCCTCAGGCCTGAAACCAACAAAA
GCCTCAAGAGCCAAGAAAACAGAGGGTGGCCTGAATTGGACCGAAGGCCTGAGTTGGATGGAAGTCTC
AAGGCTTGAGTTAGAAGTCTTAAGACCTGGGACAGGACACATGGAAGGCCTAAGAACTGAGACTTGTG
ACACAAGGCCAACGACCTAAGATTAGCCCAGGGTTGTAGCTGGAAGACCTACAACCCAAGGATGGAAG
GCCCCTGTCACAAAGCCTACCTAGATGGATAGAGGACCCAAGCGAAAAAGGTATCTCAAGACTAACGG
CCGGAATCTGGAGGCCCATGACCCAGAACCCAGGAAGGATAGAAGCTTGAAGACCTGGGGAAATCCCA
AGATGAGAACCCTAAACCCTACCTCTTTTCTATTGTTTACACTTCTTACTCTTAGATATTTCCAGTTC
TCCTGTTTATCTTTAAGCCTGATTCTTTTGAGATGTACTTTTTGATGTTGCCGGTTACCTTTAGATTG
ACAGTATTATGCCTGGGCCAGTCTTGAGCCAGCTTTAAATCACAGCTTTTACCTATTTGTTAGGCTAT
AGTGTTTTGTAAACTTCTGTTTCTATTCACATCTTCTCCACTTGAGAGAGACACCAAAATCCAGTCAG
TATCTAATCTGGCTTTTGTTAACTTCCCTCAGGAGCAGACATTCATATAGGTGATACTGTATTTCAGT
CCTTTCTTTTGACCCCAGAAGCCCTAGACTGAGAAGATAAAATGGTCAGGTTGTTGGGGAAAAAAAAA
GTGCCAGGCTCTCTAGAGAAAAATGTGAAGAGATGCTCCAGGCCAATGAGAAGAATTAGACAAGAAAT
ACACAGATGTGCCAGACTTCTGAGAAGCACCTGCCAGCAACAGCTTCCTTCTTTGAGCTTAG (SEQ ID NO:10)

## 6.8kb human XIST sequence – SEQ ID NO:11

tctagaacattttctagtcccccaacaccctttatggcgtatttctttaaaaaaatcacc
taaattccataaaatatttttttaaattctatactttctcctagtgtcttcttgacacgt
cctccatatttttttaaagaaagtatttggaatattttgaggcaatttttaatatttaag
gaatttttctttggaatcatttttggtgacatctctgtttttttgtggatcagtttttttac
tcttccactctcttttctatattttgcccatcggggctgcggatacctggttttattatt
ttttctttgcccaacggggccgtggatacctgccttttaattctttttttattcgcccatc
ggggccgcggatacctgctttttattttttttttccttagcccatcggggtatcggatacc
tgctgattcccttcccctctgaaccccccaacactctggcccatcggggtgacggatatct
gctttttaaaaatttttcttttttttggcccatcggggcttcggatacctgcttttttttttt
tttattttccttgcccatcggggcctcggatacctgctttaatttttgtttttctgccca
tcggggccgcggatacctgctttgattttttttttttcatcgcccatcggtgctttttatg
gatgaaaaaatgttggttttgtgggttgttgcactctctggaatatctacacttttttttt

6.8kb human XIST sequence – SEQ ID NO:11

gctgctgatcatttggtggtgtgtgagtgtacctaccgctttggcagagaatgactctgc
agttaagctaagggcgtgttcagattgtggaggaaaagtggccgccattttagacttgcc
gcataactcggcttagggctagtcgtttgtgctaagttaaactagggaggcaagatggat
gatagcaggtcaggcagaggaagtcatgtgcattgcatgagctaaacctatctgaatgaa
ttgatttggggcttgttaggagctttgcgtgattgttgtatcgggaggcagtaagaatca
tcttttatcagtacaagggactagttaaaaatggaaggttaggaaagactaaggtgcagg
gcttaaaatggcgattttgacattgcggcattgctcagcatggcgggctgtgctttgtta
ggttgtccaaaatggcggatccagttctgtcgcagtgttcaagtggcgggaaggccacat
catgatgggcgaggctttgttaagtggttagcatggtggtggacatgtgcggtcacacag
gaaaagatggcggctgaaggtcttgccgcagtgtaaaacatggcgggcctctttgtcttt
gctgtgtgcttttcgtgttgggttttgccgcagggacaatatggcaggcgttgtcatatg
tatatcatggcttttgtcacgtggacatcatggcgggcttgccgcattgttaaagatggc
gggttttgccgcctagtgccacgcagagcgggagaaaaggtgggatggacagtgctggat
tgctgcataacccaaccaattagaaatggggggtggaattgatcacagccaattagagcag
aagatggaattagactgatgacacactgtccagctactcagcgaagacctgggtgaatta
gcatggcacttcgcagctgtctttagccagtcaggagaaagaagtggagggccacgtgt
atgtctcccagtgggcggtacaccaggtgttttcaaggtcttttcaaggacatttagcct
ttccacctctgtcccctcttatttgtcccctcctgtccagtgctgcctcttgcagtgctg
gatatctggctgtgtggtctgaacctccctccattcctctgtattggtgcctcacctaag
gctaagtatacctcccccccaccccccaacccccccaactccccacccccaccccccac
cccccacctccccacccccctacccccctacccccctacccccctctggtctgccctgca
ctgcactgttgccatgggcagtgctccaggcctgcttggtgtggacatggtggtgagccg
tggcaaggaccagaatggatcacagatgatcgttggccaattggcctcccaatatgtgtg
attgtatttgtcgaggttgctatgcactagagaaggaaagtgctcccctcatccccactt
ttcccttccagcaggaagtgcccaccccataagaccctttatttggagagtctaggtgc
acaattgtaagtgaccacaagcatgcatcttggacatttatgtgcgtaatcgcacactgc
tcattccatgtgaataaggtcctactctccgacccctttgcaatacagaagggttgctg
ataacgcagtcccctttcttggcatgttgtgtgtgattataatcgtctgggatcctatg
cactagaaaaggagggtcctctccacatacctcagtctcacctttcccttccagcaggga
gtgcccactccataagactctcacatttggacagtcaaggtgcgtaattgttaagtgaac
acaaccatgcaccttagacatggatttgcataactacacacagctcaacctatctgaata
aaatcctactctcagacccctttgcagtacagcaggggtgctgatcaccaaggcccttt
ttcctggcctggtatgcgtgtgattatgtttgtcccggttcctgtgtattagacatggaa
gcctcccctgccacactccacccccaatcttcctttcccttccggcaggagtgccctctc
cataagacgcttacgtttggacaatcaaggtgcacagttgtaagtgaccacaggcataca
ccttggacattaatgtgcataaccactttgcccattccatctgaataaggtcctactctc
agacccctttgcagtacagcaggggtgctgatcaccaaggccccttttcttggcctgtt
atgtgcgtgattatatttgtctgggttcctgtgtattagacaaggaagccttccccccgc
ccccacccccactcccagtcttcctttcccttccagcagggagtgccccctccataagat
cattacatttggacaatcaaggtgcacaattataagtgaccacagccatgcaccttggac
attattggacattaatgtgcgtaactgcacatggcccatcccatctgaataaggacctac
tctcagatgcctttgcagtacagcaggggtactgaatcaccaaggcccttttttcttggcc
tgttatgtgtgtgattatatttatcccagtttctgtgtaatagacatgaaagcctcccct
gccacaccccacctccaatcttcctttcccttccaccagggagtgtccactccatatacc
cttacatttggacaatcaaggtgcacaattgtaagtgagcataggcactcaccttggaca
tgaatgtgcataactgcacatggcccatcccatctgaataaggtcctactctcagaccct
tttgcagtacagcaggggtgctgatcaccaaggccccttttcctggcctgttatgtgtg
tgattatatttgttccagttcctgtgtaatagacatggaagcctcccctgccacactcca
cccccaatcttcctttccttctggcaggaagtacccgctccataagacccttacatttgg
acagtcaaggtgcacaattgtatgtgaccacaaccatgcaccttggacataaatgtgtgt
aactgcacatggcccatcccatctgaataaggtcctactctcagacccctttgcagtac

6.8kb human XIST sequence – SEQ ID NO:11

agtaggtgtgctgataaccaaggcccctcttcctggcctgttaacgtatgtgattatatt
tgtctgggttccagtgtataagacatggaagcctcccctgccccaccccaccctcaatct
tcctttcccttctggcagggagtgccagctccataagaaccttacatttggacagtcaag
gtgcacaattctaagtgaccgcagccatgcaccttggtcaataatgtgtgtaactgcaca
cggcctatctcatctgaataaggccttactctcagaccccttttgcagtacagcaggggt
gctgataaccaaggcccattttcctggcctgttatgtgtgtgattatatttgtccaggtt
tctgtgtactagacaaggaagcctcctctgccccatcccatctacgcataatctttcttt
tcctcccagcagggagtgctcactccataagacccttacatttggacaatcaaggtgcac
aattgtaagtgaccacaaccatgcatcttggaaatttatgtgcataactgcacatggctt
atcctatttgaataaagtcctactctcagacccccttttgcagtatagctggggtgctgat
cactgaggcctctttgcttggcttgtctatattcttgtgtactagataagggcaccttct
catggactccctttgcttttcaacaaggagtacccactacttttaagattcttatattt
gtccaaagtacatggttttaattgaccacaacaatgtcccttggacattaatgtatgtaa
tcaccacatggttcatcctaattaaacaaagttctaccttctcaccctccatttgcagta
taccagggttgctgacccccctaagtccccttttcttggcttgttgacatgcataattgca
tttatgttggttcttgtgccctagacaaggatgccccacctcttttcaatagtgggtgcc
cactccttatgatctttacatttgaacagttaatgtgaataattgcagttgtccacaacc
ctatcacttctaggaccattatacctcttttgcattactgtggggtatactgtttccctc
caaggccccttctggtggactatcaacatataattgaaattttcttttgtctttgtcagt
agattaaggtcataccccatcacctttcctttgtagtacaacagggtgtcctgatcaacc
aaagtcctgttgttttggactgttaatatgtgcaattacatttgctcctgatctgtgcac
tagataaggatcctacctactttcttagtgtttttagcaggtagtgcccactactcaaga
ctgtcacttggaatgttcatgtgcacaaactcaattctctaagcatgttcctgtaccacc
tttgctttagagcaggggggatgatattcactaagtgccccttcttttggacttaatatgc
attaatgcaattgtccacctcttcttttagactaagagttgatctccacatattcccctt
gcatcaggggcatgttaattatgaatgaaccctttttcttttaatattaatgtcataattg
tatttgtggacctgtgtaggagaaaaagaccctatgttcctcccattaccctttggattg
ctgctgagaagtgttaactactcataatctcagctcttggacaattaatagcattaataa
caattatcaagggcactgatcattagataagactcctgcttcctcgttgcttacatcggg
ggtactgacccactaaggccccttgtactgttaatgtgaatatttgcaattatatatgtc
tccttctggtagagtgggatattatgccctagtatcccctttgcattactgcaggggctg
ctgactactcaaaacttctcctgggactgttaataggcacaatggcagttatcaatggtt
ttctccctccctgaccttgttaagcaagcgccccaccccacccttagtttcccatggcat
aataaagtataagcattggagtattccatgcacttgtctatcaaacagtggtccatactc
ccaacccttttgcattgcgccagtgtgtaaaatcacaggtagccatggtgtcatgcttta
tatacgaagtcttccctctctctgccccttgtgtgcccttggcccctttttacagactat
tgctcacaatctcaggtgtccatatttgcagctattaggtaagattgtgctgtctccctc
ttcccttccctctgccctgccccttttgcctctttgctgggtaatgttgaccGgacaagg
ccctttctcttggacttaaacaattctcagttgcactttccttggtccCacccattatac
atgaacccctctacttcctttcgcattgcttctgagtatgctgactacccaaagcccctt
ctgtgttattaataaacacagtactgattgtcccatttttcagcccatcagtccaagatc
tccctaccactttggtgtgttggtgcagtgttgactatgaaaagcaggcctgaactaggt
ggataagccttcactcatttttctttcatttattaatgatcctagtttcaattattgtcag
attctggggacaagaaccattcttgcccacctgtgttactgctttactgtgcaaaatact
gaaggcaagtcagacccagggagctggattgccatcctttattttgtgtttccagtgtac
actataaaattgtctccccaggaaggaaggttggcactttctctgcattcttctttccag
agcagattgcctggttaagaatctcttgttgtcccCtTtgtatattgttattgtaaagtg
ccaaatgccaggatacagccagaaaaattgcttattattattaaaaaaatttttttaaga
aagacatctggattgtagggtggactcgataacctggtcattattttttttgaagccaaaa
tatccatttatactatgtacctggtgaccagtgtctctcattttaactgagggtggtggg
tctgtggatagaacactgactcttgctattttaatatcaaagatattctagATCCAGCAC

6.8kb human XIST sequence – SEQ ID NO:11

```
AGTGGCggcccgataccgtcgacc
```

*Targeting Sequences*

The nucleic acid constructs described herein include targeting sequences or elements (the terms are used interchangeably herein) that promote sequence specific integration of an Xist transgene into the DYRK1A or RCAN1 gene (e.g., by homologous recombination).
5 Methods for achieving site-specific integration by ends-in or ends-out targeting are known in the art and in the nucleic acid constructs of this invention, the targeting elements are selected and oriented with respect to the Xist transgene according to whether ends-in or ends-out targeting is desired. In certain embodiments, two targeting elements flank the Xist transgene.

A targeting sequence or element may vary in size. In certain embodiments, a
10 targeting element may be at least or about 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000 bp in length (or any integer value in between). In certain embodiments, a targeting element is homologous to a sequence that occurs naturally in a trisomic and/or translocated chromosomal region, including a polymorphic sequence which may be present on just one of the homologous chromosomes.

15 The construct elements as described here may be variants of naturally occurring DYRK1A or RCAN1 sequences. Preferably, any construct element (e.g., an Xist transgene, other non-coding, silencing RNA, or a targeting element) includes a nucleotide sequence that is at least 80% identical to its corresponding naturally occurring sequence (its reference sequence, e.g., an Xist coding region, a human Chr 21 sequence, or any duplicated or
20 translocated genomic sequence). More preferably, the silencing sequence or the sequence of a targeting element is at least 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% identical to its reference sequence (e.g., NG_009366.1, the human refGene Sequence of DYRK1A, or NG_007071.1, the human refGene Sequence of RCAN1).

As used herein, "% identity" of two nucleic acid sequences is determined using the
25 algorithm of Karlin and Altschul (Proc. Natl. Acad. Sci. USA, 87:2264-2268, 1990), modified as in Karlin and Altschul, Proc. Natl. Acad. Sci. USA, 90:5873-5877, 1993). Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul et al. (J. Mol. Biol. 215:403-410, 1990). BLAST nucleotide searches are performed with the NBLAST program, score = 100, wordlength = 12. BLAST protein searches are performed
30 with the XBLAST program, score = 50, wordlength=3. To obtain gapped alignment for comparison purposes GappedBLAST is utilized as described in Altschul et al. (Nucl. Acids

Res., 25:3389-3402, 1997). When utilizing BLAST and GappedBLAST programs the default parameters of the respective programs (e.g., XBLAST and NBLAST) are used to obtain nucleotide sequences homologous to a nucleic acid molecule of the invention.

In preferred embodiments, the targeting elements comprise all or part of the following sequences, with the DYRK1A target sequences shown in bold and underlined:

DYRK1A LEFT ARM:

ATGGTAATTGAGAAATGACAAGAATCATGGAACTCCAAATTCATGACAATATTTGGGTAAGACGTCTA
CCTTTCCCTCCATACCTAAATTAACTAAACGGGTTTCGCTGTGTCTTCAACCATCGATCGATCATTTA
CCGTTTTAACTTAGGTCTGAGGAATACCACAATTAACGATATCGATTTCTACTTTGACCTCAACACGG
TGAGGAACGTGTGAAAATAGACAGTGGGAGAATCCGACAAAATCTTTTAGGGTACAAAATCGAACGGT
AAGACAACTGGGTCGGACGGAAAGATCGGAATTGAATGGGGAGACAGATATAAGATAAAAGGTCGGTT
TATACTCCACTGCAAATTCGACGATGAACTTTCTCTTCACCCTCAATCCGTCTCGTCATCCCCTTAGT
ACAAACCCCTTCTCACTTCTCACATGAACTCTCTCACACCTCCACGGAACCTCCTCGACCTCGGGTCT
CCACGGGGTACTCTTGTTGTGTCCTCCGACGTCCACCTCCACCCACGGACTAACATCTTACGAAAGAT
CAACAGAAGGTGTCCTGTAAAAACCCTCGATAAGTGTTCTAAGTACCGATGGCACGAGATTTTAAACT
ACACTTCAAGTAAAAAGGACCTGAAGAATGAATTAAGGAGACAGAAAACCGGGT**<u>CGGTGGGGAAACGG</u>**
**<u>TCAAA</u>** (SEQ ID NO:12)

DYRK1A RIGHT ARM:

**<u>CCACTACTCGTCCGACAA</u>**ACCTTTCTTGCAGGAGCTCGTCCCACGACAAAGGATTGGGACGCAGAAAA
AGGGGAGACTCTAGTCAAATAGAAATAAGTGAACGTCCACAAGTTGTTAGAACAGAAAATACCCCTTA
AAGATTACACAGAACTCGTGAAAGGGTGGGAGGATAGAACCTCCGTACCAAGTCTCACCTTTTCCCGC
GCCCGGGTGGATGGAGACCGGAAGGGTGGAGTCGGTGGTACGAATCCCGGCACCACCTCACGAACTGG
AGAAACACACATGTTACGTTATGTACGACCTTATTACGGTGGAATACGTATCCCGAAAACACCCACAT
TCCCGTATGGCCTTGTTCAACCGTATCTTATTCTCAAGTCACTTACAACAGTGATGAAAAATAATGAA
AAATTAACACTTTTTGAGTGTCTAAGACATTATTTCCCAGTATCTTTGGACGAAATAGGTATGATAGT
AATGACTCTTATGAAAGACCAAAAGC (SEQ ID NO:13)

In some embodiments, the XIST cDNA is inserted into the silencing vector in the opposite direction in order to avoid generating a fused RNA with DYRK1A exon1. In these embodiments, exemplary targeting arms comprise the sequences of SEQ ID NO:33 and SEQ ID NO: 34, set forth below.

*Selection Markers*

In addition, the nucleic acids may contain a marker for the selection of transfected cells (for instance, a drug resistance gene for selection by a drug such as neomycin, hygromycin, and G418). Such vectors include pMAM, pDR2, pBK-RSV, pBK-CMV, pOPRSV, pOP13, and so on. More generally, the term “marker” refers to a gene or sequence whose presence or absence conveys a detectable phenotype to the host cell or organism. Various types of markers include, but are not limited to, selection markers, screening markers, and molecular markers. Selection markers are usually genes that can be expressed to convey a phenotype that makes an organism resistant or susceptible to a specific set of

environmental conditions. Screening markers can also convey a phenotype that is a readily observable and distinguishable trait, such as green fluorescent protein (GFP), GUS or β-galactosidase. Molecular markers are, for example, sequence features that can be uniquely identified by oligonucleotide probing, for example RFLP (restriction fragment length polymorphism), or SSR markers (simple sequence repeat). To amplify the gene copies in host cell lines, the expression vector may include an aminoglycoside transferase (APH) gene, thymidine kinase (TK) gene, *E. coli* xanthine guanine phosphoribosyl transferase (Ecogpt) gene, dihydrofolate reductase (dhfr) gene, and such as a selective marker.

Expression of the selection marker can be driven by the same regulatory elements (e.g., promoters) as the silencing sequence, or can be driven by a separate regulatory element.

*Recombination Facilitating Elements – Cleavage Vectors*

In some embodiments, the present methods include the use of cleavage vectors, i.e., nucleic acid constructs include a sequence that enhances or facilitates homologous recombination (e.g., a zinc finger nuclease or TALEN). Zinc finger domains and TALENs can recognize and target highly specific chromosomal sequences to facilitate targeted integration of the transgene into the DYRK1A or RCAN1 gene. Alternatively, CRISPR/CAS genome editing could be used. As would be understood in the art, the term “recombination” is used to indicate the process by which genetic material at a given locus is modified as a consequence of an interaction with other genetic material. Homologous recombination indicates that recombination has occurred as a consequence of interaction between segments of genetic material that are homologous or identical. In contrast, “non-homologous” recombination indicates a recombination occurring as a consequence of the interaction between segments of genetic material that are not homologous (and therefore not identical). Non-homologous end joining (NHEJ) is an example of non-homologous recombination.

In some embodiments, targeting the present silencing constructs to DYRK1A or RCAN1 can be facilitated by introducing a chimeric zinc finger nuclease (ZFN), i.e., a DNA-cleavage domain (nuclease) operatively linked to a DNA-binding domain including at least one zinc finger, into a cell. Typically the DNA-binding domain is at the N-terminus of the chimeric protein molecule, and the DNA-cleavage domain is located at the C-terminus of the molecule. These nucleases exploit endogenous cellular mechanisms for homologous recombination and repair of double stranded breaks in genetic material. ZFNs can be used to target a wide variety of endogenous nucleic acid sequences in a cell or organism. The present compositions can include cleavage vectors that target a ZFN to a region within DYRK1A or

RCAN1, and the methods include transfection or transformation of a host cell or organism by introducing a cleavage vector encoding a ZFN (e.g., a chimeric ZFN), or by introducing directly into the cell the mRNA that encodes the recombinant zinc finger nuclease, or the protein for the ZFN itself. One can then identify a resulting cell or organism in which a selected endogenous DNA sequence is cleaved and exhibits a mutation or DNA break at a specific site, into which the transgene will become integrated.

The ZFN can include multiple (e.g., at least three (e.g., 3, 4, 5, 6, 7, 8, 9 or more)) zinc fingers in order to improve its target specificity. The zinc finger domain can be derived from any class or type of zinc finger. For example, the zinc finger domain can include the Cys2His2 type of zinc finger that is very generally represented, for example, by the zinc finger transcription factors TFIIIA or Sp1. In a preferred embodiment, the zinc finger domain comprises three Cys2His2 type zinc fingers.

The ZFN DNA-cleavage domain can be derived from a class of non-specific DNA cleavage domains, for example the DNA-cleavage domain of a Type II restriction enzyme such as FokI. Thus, a chimeric ZFN useful in the present methods can include three Cys2His2 type zinc fingers and a DNA-cleavage domain derived from the Type II restriction enzyme FokI. In this event, each zinc finger contacts three consecutive base pairs of DNA creating a 9 bp recognition sequence for the ZFN DNA binding domain. The DNA-cleavage domain of the embodiment requires dimerization of two ZFN DNA-cleavage domains for effective cleavage of double-stranded DNA. This imposes a requirement for two inverted recognition (target DNA) sites within close proximity for effective targeted genetic recombination. If all positions in the target sites are contacted specifically, these requirements enforce recognition of a total of 18 base pairs of DNA. There may be a space between the two sites. The space between recognition sites for ZFNs may be equivalent to 6 to 35 bp of DNA. The region of DNA between the two recognitions sites may be referred to as the “spacer.”

A linker, if present, between the cleavage and recognition domains of the ZFN can be a sequence of amino acid residues that result in a flexible linker is flexible, although linkerless constructs tend to improve target site specificity. A linkerless construct has a strong preference for binding to and then cleaving between recognition sites that are 6 bp apart. However, with linker lengths of between 0 and about 18 amino acids in length, ZFN-mediated cleavage occurs between recognition sites that are between 5 and 35 bp apart. For a given linker length, there will be a limit to the distance between recognition sites that is consistent with both binding and dimerization. As noted, there may be no linker between the

cleavage and recognition domains, and the target locus can include two nine nucleotide recognition sites in inverted orientation with respect to one another, separated by a six nucleotide spacer.

To target genetic recombination or mutation, two 9 bp zinc finger DNA recognition sequences are identified in the host DNA. These recognition sites will be in an inverted orientation with respect to one another and separated by about 6 bp of DNA. ZFNs are then generated by designing and producing zinc finger combinations that bind DNA specifically at the target locus, and then linking the zinc fingers to a cleavage domain of a Type II restriction enzyme.

A silencing sequence flanked by sequences (typically 400 bp-5 kb in length) homologous to the desired site of integration can be inserted (e.g., by homologous recombination) into the site cleaved by the endonuclease, thereby achieving a targeted insertion. When used in combination with a ZFN construct, the silencing sequence may be referred to as "donor" nucleic acid or DNA.

In some embodiments, the cleavage vector includes a transcription activator-like effector nuclease (TALEN). TALENs function in a manner somewhat similar to ZFNs, in that they can be used to induce sequence-specific cleavage; see, e.g., Hockemeyer et al., Nat Biotechnol. 29(8):731-4 (2011); Moscou et al., 2009, Science 326:1501; Boch et al., 2009, Science 326:1509-1512. Methods are known in the art for designing TALENs, see, e.g., Rayon et al., Nature Biotechnology 30:460–465 (2012).

*Vectors and Transformation*

The various active sequences, including the silencing sequence and the sequence encoding a chimeric ZFN can be introduced into a host cell on the same vector or separately (e.g., on separate vectors or separate types of vectors at the same time or sequentially). Methods for introducing the various nucleic acids, constructs, and vectors are discussed further below and are well known in the art.

Transformation can be carried out by a variety of known techniques which depend on the particular requirements of each cell or organism. Such techniques have been worked out for a number of organisms and cells and are readily adaptable. Stable transformation involves DNA entry into cells and into the cell nucleus. For example, transformation can be carried out in culture, followed by selection for transformants and regeneration of the transformants. Methods often used for transferring DNA or RNA into cells include forming DNA or RNA complexes with cationic lipids, liposomes or other carrier materials, micro-

injection, particle gun bombardment, electroporation, and incorporating transforming DNA or RNA into virus vectors.

Liposomal formulations: In certain embodiments of the invention, the oligo- or polynucleotides and/or expression vectors containing silencing sequences and/or ZFNs may be entrapped in a liposome. Liposomes are vesicular structures characterized by a phospholipid bilayer membrane and an inner aqueous medium. Multilamellar liposomes have multiple lipid layers separated by aqueous medium. They form spontaneously when phospholipids are suspended in an excess of aqueous solution. The lipid components undergo self-rearrangement before the formation of closed structures and entrap water and dissolved solutes between the lipid bilayers. Also contemplated are cationic lipid-nucleic acid complexes, such as lipofectamine-nucleic acid complexes. Lipids and liposomes suitable for use in delivering the present constructs and vectors can be obtained from commercial sources or made by methods known in the art.

Microinjection: Direct microinjection of DNA into various cells, including egg or embryo cells, has also been employed effectively for transforming many species. In the mouse, the existence of pluripotent embryonic stem (ES) cells that can be cultured in vitro has been exploited to generate transformed mice. The ES cells can be transformed in culture, then micro-injected into mouse blastocysts, where they integrate into the developing embryo and ultimately generate germline chimeras. By interbreeding heterozygous siblings, homozygous animals carrying the desired gene can be obtained.

Viral Vectors as Expression Constructs: Other viral vectors may be employed as expression constructs in the present invention. Vectors derived from, for example, vaccinia virus, adeno-associated virus (AAV, e.g., MV), or herpes virus may be employed. Extensive literature is available regarding the construction and use of viral vectors. For example, see Miller et al. (Nature Biotechnol. 24:1022-1026, 2006) for information regarding adeno associated viruses. Defective hepatitis B viruses, may be used for transformation of host cells. In vitro studies show that the virus can retain the ability for helper-dependent packaging and reverse transcription despite the deletion of up to 80% of its genome. Potentially large portions of the viral genome can be replaced with foreign genetic material. The hepatotropism and persistence (integration) are particularly attractive properties for liver-directed gene transfer. The chloramphenicol acetyltransferase (CAT) gene has been successfully introduced into duck hepatitis B virus genome in the place of the viral polymerase, surface, and pre-surface coding sequences. The defective virus was cotransfected with wild-type virus into an avian hepatoma cell line, and culture media

containing high titers of the recombinant virus were used to infect primary duckling hepatocytes. Stable CAT gene expression was subsequently detected.

Non-viral Methods: Several non-viral methods are contemplated by the present invention for the transfer into a host cell of DNA constructs encoding ZFNs and, when appropriate, donor DNA. These include calcium phosphate precipitation, lipofectamine-DNA complexes, and receptor-mediated transfection. Some of these techniques may be successfully adapted for in vivo or ex vivo use.

In one embodiment of the invention, the expression constructs may simply consist of naked recombinant DNA, or in some cases mRNA for the recombinant ZFN. Transfer of the construct may be performed by any of the nuclei acid transfer methods mentioned above which physically or chemically permeabilize the cell membrane. For example, polyomavirus DNA in the form of $CaPO_4$ precipitates was successfully injected into liver and spleen of adult and newborn mice which then demonstrated active viral replication and acute infection. In addition, direct intraperitoneal injection of $CaPO_4$ precipitated plasmid expression vectors results in expression of the transfected genes.

*Pharmaceutical Compositions, RNAs, and Cells*

In another embodiment, the invention features compositions (*e.g.*, pharmaceutically acceptable compositions) that include the nucleic acid constructs or vectors described herein. Various combinations of the constructs and vectors described herein can be formulated as pharmaceutical compositions.

Also within the scope of the invention are RNAs and proteins encoded by the cleavage vector and compositions that include them (*e.g.*, lyophilized preparations or solutions, including pharmaceutically acceptable solutions or other pharmaceutical formulations).

In another embodiment, the invention features cells that include the nucleic acid constructs, vectors (*e.g.*, an adeno associated vector), and compositions described herein. The cell can be isolated in the sense that it can be a cell within an environment other than that in which it normally resides (*e.g.*, the cell can be one that is removed from the organism in which it originated). The cell can be a germ cell, a stem cell (*e.g.*, an embryonic stem cell, an adult stem cell, or an induced pluripotent stem cell (iPS cell or IPSC)), or a precursor cell. Where adult stem cells are used, the cell can be a hematopoietic stem cell, a cardiac muscle stem cell, a mesenchymal stem cell, or a neural stem cell (e.g., a neural progenitor cell). The cell can also be a differentiated cell (*e.g.*, a fibroblast or neuron).

*Methods of Treatment*

The methods of the invention can be used to treat patients who have trisomy 21. Any of the methods can include the step of identifying a patient in need of treatment; any of the patients can be human; and any of the methods can be carried out by either administering the present compositions to the patient, or removing cells from the patient, treating the cells, and "readministering" those cells. For example, the invention features methods of treating a genetic disorder associated with a trisomic chromosome 21 by identifying a patient in need of treatment; and administering to the patient a nucleic acid construct, vector, and/or cleavage vector as described herein. The amount of the construct or vector administered will be an amount sufficient to improve a condition associated with the disorder. Where cells are harvested from a patient to treat a condition or disorder described herein (or an associated symptom), the methods can include the steps of identifying a patient in need of treatment; harvesting cells from the patient; transfecting the cells with one or more of the types of constructs and/or vectors described herein; and administering to the patient a sufficient number of the transfected cells to treat the condition or improve a condition or symptom associated with the disorder. The symptoms associated with many birth defects and other conditions are well known. For example, individuals having Down Syndrome often experience mental retardation, hypotonia, cardiac defects, Alzheimer's Disease, hematological abnormalities and leukemia (*see* Antonarakis and Epstein, *Trends Mol. Med.* 12:473-479, 2006). As noted above, treatment can also be carried out *in vivo* by administering present compositions to the patient via pharmaceutically acceptable compositions.

The cells can include differentiated cells (*e.g.*, white blood cells or fibroblasts) and/or undifferentiated cells (*e.g.*, stem cells or precursor cells). The cells can also be differentiated cells that are induced, *ex vivo*, into iPS cells, or multi-potent stem cells or stem cells of particular lineage, such as neural stem cells. Neural stem cells (also called neural progenitors), are characterized by the ability to form neural rosettes, a neural tube-like structure (see, e.g., Figure 6). The condition can be a neurological or blood disorder such as Alzheimer's Disease and leukemia, respectively, or a muscular defect, including defects of the heart.

To illustrate a particular application, Xist mediated chromosomal therapy could be used to ameliorate transient myeloproliferative disorder (TMD) in Down Syndrome children and possibly prevent the later development of acute leukemia. Successful bone marrow transplants for diseases like leukemia depend upon immune compatibility, to avoid Graft

versus Host Disease (GVHD). To avoid graft rejection, the patient's own cells can be used and transgenically modified prior to transplant. There are two scenarios to acquire and modify stem cells for bone marrow transplant. In the first, the patient's own bone marrow stem cells can be obtained and an Xist transgene as described herein can be introduced and
5 targeted to chromosome 21. When Xist expression silences the trisomic chromosome, these cells can then be transplanted back into the patient following standard bone marrow transplant procedures following the destruction of the patient's bone marrow using irritation. Modified autologous (from the patient) bone marrow cells can be transplanted without first irradiating the patient to destroy the unmodified bone marrow. This would produce a
10 situation where the patient's bone marrow would be mosaic for trisomy 21 (a mixture of modified and unmodified cells). The data presented herein indicate that the modified cells would have a growth advantage over the non-modified fully trisomic cells, and the modified cells would eventually outgrow the non-modified disease-inducing cells (*see* Douillard-Guilloux *et al.*, *J. Gene Med.* 11:279-287, 2009). In the second approach, the patient's
15 fibroblast (skin) cells can be used to produce iPS cells, into which a transgenic Xist gene is inserted and targeted to chromosome 21. IPS cells that silence one of the three trisomic chromosomes will then be differentiated into adult hemopoietic stem cells and introduced back into the patient as described herein.

## EXAMPLES

20 The invention is further described in the following examples, which do not limit the scope of the invention described in the claims.

## MATERIALS & METHODS

The following materials and methods were used in the Examples set forth below.

25 **Cell culture.** HT1080 TetR cells (Invitrogen) and Female DS human primary fibroblast line (Coriell) (AG13902) were cultured as recommended by supplier. DS iPSC parent line (DS1-iPS4) was provided by George Q. Daley (Children's Hospital Boston, USA) and maintained on irradiated mouse embryonic fibroblasts (iMEFs) (R & D Systems, PSC001) in hiPSC medium containing DNEM/F12 supplemented with 20% knockout Serum
30 Replacement (Invitrogen), 1mM glutamine (Invitrogen), 100 μM non-essential amino acids (Invitrogen), 100 μM β-mercaptoethanol (Sigma) and 10 ng/ml FGF-β (Invitrogen,

PHG0024). Cultures were passaged every 5-7 days with 1 mg/ml of collagenase type IV (Invitrogen).

**ZFN design.** ZFNs against the human AAVS1 locus (PPP1R12C) on Chr19 have been previously described[25]. ZFNs against the DYRK1A locus were designed using an archive of pre-validated zinc finger modules[24,53,54], and validated for genome editing activity by transfection into K562 cells and Surveyor endonuclease-based measurement of endogenous locus disruption ("Cel1" [55,56]) exactly as described[54]. Southern blotting for targeted gene addition was performed exactly as described[23,57] on SphI-digested genomic DNA probed with a fragment corresponding to positions Chr 21:38825803+38826056 (hg19).

**iPSC reprogramming of DS fibroblasts.** Three primary DS fibroblast cultures (Coriell: GM02504, AG13902, GM02067) were obtained and cultured. Two of these cultures (GM02504, AG13902) were used for reprogramming with assistance of the UConn Stem Cell Core Laboratory, using retroviral transduction with five reprogramming factors (OSKML).

It was initially noted that two of the three human primary DS fibroblast samples showed very limited proliferation even though age of donor and passage number would not predict this. In addition, a marked deficit in DS mouse tail tip fibroblast proliferation was seen. Additionally, in two attempts at reprogramming human DS fibroblasts, only the AG13902 sample was modestly successful, suggesting DS fibroblasts were more difficult to reprogram than control fibroblasts. Fewer subclones were obtained and most of these showed poor morphology and slower growth than controls.

**XIST and rtTA/puro plasmid construction.** 14 kb human XIST cDNA, a splicing isoform of full length XIST cDNA was subcloned into pTRE3G (Clontech, Cat#: 631167). Two homologous arms (left arm, 690 bp; right arm, 508 bp) of DYRK1A gene on Chr 21 were amplified by PCR from primary DS fibroblasts (AG13902) (Coriell) and cloned into the pTRE3G vector (Human Chr 21 DYRK1A left arm primers: forward 5'-GCCGTATACCATTAACTCTTTACTGTTC-3' (SEQ ID NO:1), reverse 5'-TCTGTATACGTAAACTGGCAAAGGGGTGG-3' (SEQ ID NO:2); Human Chr 21 DYRK1A right arm primers: forward 5'- ATTTCGCGAACGGGTGATGAGCAGGCTGT-3' (SEQ ID NO:3), reverse 5'- CCGTCGCGAAAACCAGAAAGTATTCTCAG-3' (SEQ ID NO:4)).

DYRK1A left arm, reverse:

AAACTGGCAAAGGGGTGGCTGGGCCAAAAGACAGAGGAATTAAGTAAGAAGTCCAGGAAAAATGAACT
TCACATCAAATTTTAGAGCACGGTAGCCATGAATCTTGTGAATAGCTCCCAAAAATGTCCTGTGGAAG
ACAACTAGAAAGCATTCTACAATCAGGCACCCACCTCCACCTGCAGCCTCCTGTGTTGTTCTCATGGG
5 GCACCTCTGGGCTCCAGCTCCTCCAAGGCACCTCCACACTCTCTCAAGTACACTCTTCACTCTTCCCC
AAACATGATTCCCCTACTGCTCTGCCTAACTCCCACTTCTCTTTCAAGTAGCAGCTTAAACGTCACCT
CATATTTGGCTGGAAAATAGAATATAGACAGAGGGGTAAGTTAAGGCTAGAAAGGCAGGCTGGGTCAA
CAGAATGGCAAGCTAAAACATGGGATTTTCTAAAACAGCCTAAGAGGGTGACAGATAAAAGTGTGCAA
GGAGTGGCACAACTCCAGTTTCATCTTTAGCTATAGCAATTAACACCATAAGGAGTCTGGATTCAATT
10 TTGCCATTTACTAGCTAGCTACCAACTTCTGTGTCGCTTTGGGCAAATCAATTAAATCCATACCTCCC
TTTCCATCTGCAGAATGGGTTTATAACAGTACTTAAACCTCAAGGTACTAAGAACAGTAAAGAGTTAA
TGGTA (SEQ ID NO:33)

DYRK1A right arm, reverse:

15 CGAAAACCAGAAAGTATTCTCAGTAATGATAGTATGGATAAAGCAGGTTTCTATGACCCTTTATTACA
GAATCTGTGAGTTTTTCACAATTAAAAAGTAATAAAAAGTAGTGACAACATTCACTGAACTCTTATTC
TATGCCAACTTGTTCCGGTATGCCCTTACACCCACAAAAGCCCTATGCATAAGGTGGCATTATTCCAG
CATGTATTGCATTGTACACACAAAGAGGTCAAGCACTCCACCACGGCCCTAAGCATGGTGGCTGAGGT
GGGAAGGCCAGAGGTAGGTGGGCCCGCGCCCTTTTCCACTCTGAACCATGCCTCCAAGATAGGAGGGT
20 GGGAAAGTGCTCAAGACACATTAGAAATTCCCCATAAAAGACAAGATTGTTGAACACCTGCAAGTGAA
TAAAGATAAACTGATCTCAGAGGGGAAAAAGACGCAGGGTTAGGAAACAGCACCCTGCTCGAGGACGT
TCTTTCCAAACAGCCTGCTCATCACC (SEQ ID NO:34)

The pEF1α-3G rtTA-pA cassette from pEF1α-Tet3G vector (Clontech) was subcloned
25 into a plasmid for targeted gene addition to the PPP1R12C/AAVS1 locus [25], which contains a unique HindIII site flanked by two 800 bp stretches of homology to the ZFN-specified position in the genome.

See Figures 2a & 9a, and 2b & 9b.

30 **Constructs for targeting DYRK1 or RCAN1:**

The following constructs were made and tested. Two constructs for a dual-targeting strategy in human Down Syndrome iPSCs were made as follows:

CONSTRUCT 1 (3G/FL/hXIST/DYRK1A): The 18.5 kb inducible human *XIST* construct that contains 14 kb full length human *XIST* cDNA is targeted to the *DYRK1A* gene
35 on Chr 21 by a dual-targeting strategy in human Down syndrome iPSCs. See Fig. 2a, 9a, 10a and SEQ ID NO:14 for the sequence.

CONSTRUCT 2 (puro/rtTA/AAVS1): The puro/rtTA construct is targeted to the *AAVS1* locus on Chr 19 for the dual-targeting strategy. See Fig. 2b, 9b, 10b, and SEQ ID NO:15 for the sequence.

40 The dual-targeting strategy was specifically designed for Down Syndrome iPSCs. The 18.5 kb inducible human *XIST* transgene (3G/FL/hXIST/DYRK1A) is targeted to the

*DYRK1A* gene on Chr 21, and the puro/rtTA plasmid (puro/rtTA/AAVS1) is targeted to a safe harbor of human genome (*AAVS1* locus) on Chr 19. Puromycin on the puro/rtTA construct is for selection of *XIST*-targeted clones (by 3G/FL/hXIST/DYRK1A) and tetracycline transactivator (rtTA) is for induction of *XIST* transgene expression on Chr 21.

5 Four selectable and inducible *XIST* constructs targeted the *RCAN1* and *DYRK1A* loci on Chr 21 in human somatic cells were made as follows:

CONSTRUCT 3 (FL/hXIST/RCAN1): The 21.1 kb selectable and inducible human *XIST* construct that contains 14 kb full length human *XIST* cDNA is targeted to the *RCAN1* gene on Chr 21 by ZFNs in human somatic cells. See Figs. 7A-B, 10c, and SEQ ID NO:16

10 for the sequence.

CONSTRUCT 4 (FL/hXIST/DYRK1A): The 20.7 kb selectable and inducible human *XIST* construct that contains 14 kb full length human *XIST* cDNA is targeted to the *DYRK1A* gene on Chr 21 by ZFNs in human somatic cells. See Fig. 9c, 10d, and SEQ ID NO:17 for the sequence.

15 CONSTRUCT 5 (6.8 kb/hXIST/RCAN1): The 14.0 kb selectable and inducible human *XIST* construct that contains 6.8 kb exon 1 of human *XIST* cDNA (SEQ ID NO:2, obtained from C. Brown, University of British Columbia) is targeted to the *RCAN1* gene on Chr 21 by ZFNs in human somatic cells. See Fig. 7c, 10e, and SEQ ID NO:18 for the sequence.

20 CONSTRUCT 6 (6.8 kb/hXIST/DYRK1A): The 13.7 kb selectable and inducible human *XIST* construct that contains 6.8 kb exon 1 of human *XIST* cDNA (SEQ ID NO:2, obtained from C. Brown, University of British Columbia) is targeted to the *DYRK1A* gene on Chr 21 by ZFNs in human somatic cells. See Fig. 9d, 10f, and SEQ ID NO:19 for the sequence.

25 One selectable and inducible *XIST* construct targeted AAVS1 locus on Chr 19 in human somatic cells was made as follows:

CONSTRUCT 7 (6.8 kb/hXIST/AAVS1): The 15.7 kb selectable and inducible human *XIST* construct that contains 6.8 kb exon 1 of human *XIST* cDNA (SEQ ID NO:2, obtained from C. Brown, University of British Columbia) is targeted to the *AAVS1* locus on

30 Chr 19 by ZFNs in human somatic cells . See Fig. 9e, 10g, and SEQ ID NO:20 for the sequence.

One selectable and inducible mouse *Xist* construct targeted the *Runx1* gene on mouse Chr 16 was made as follows:

CONSTRUCT 8 (6.3 kb/mXist/Runx1): The 20.6 kb selectable and inducible mouse *Xist* construct that contains a 6.3 kb exon 1 of mouse *Xist* cDNA is targeted the *Runx1* gene on Chr 16 (synteny to human Chr 21) by conventional homologous recombination. See Fig. 9f, 10h, and SEQ ID NO:21 for the sequence.

5 CONSTRUCT 9 (pEF1α/hDYRK1A/FL mXist): The 20.6 kb construct contains full length mouse Xist cDNA, an ampicillin resistance gene, and two homologous arms that target intron 1 of DYRK1A gene on human chromosome 21. See Fig. 9g, 10i, and SEQ ID NO:22 for the sequence.

CONSTRUCT 10 (pEF1α/hDYRK1A/6.3kb mXist). The 12.2 kb construct contains 10 6.3 kb of mouse Xist cDNA that has been reported to function (Wutz et al., Nat Genet 30, 167-174 (2002)) and two homologous arms that target intron 1 of DYRK1A gene on human chromosome 21. See Fig. 9h, 10j, and SEQ ID NO:23 for the sequence.

CONSTRUCT 11 (Rosa26/pEF1x-Tet3G/hPGK-PuroR). The 10.3 kb construct contains a puromycin resistance selection gene and rtTA cassette that is targeted to the 15 Rosa26 locus on mouse chromosome 6 by ZFNs. See Fig. 9i, 10k, and SEQ ID NO:24 for the sequence.

The constructs described above targeting human Chr 21 or mouse Chr 16 constitute the first "dosage compensating transgenes" designed to silence chromosome imbalance involving duplication of chromosomal material, particular trisomy with much clinical import.

20

**Dual-Targeted-addition of human DS iPSCs and generation of stable targeted clones.** The DS iPSC line was cultured in 10 μM of Rho-associated protein kinases (ROCK) inhibitor (Calbiochem; Y27632) 24 h before electroporation. Single cells ($1X10^7$) were harvested using TryPLE select (Invitrogen), resuspended in 1 x PBS and electroporated with 25 a total of 55 μg DNA including five plasmids (XIST, DYRK1A ZFN1, DYRK1A ZFN2, rtTA/puro, and AAVS1 ZFN) with both 3:1 and 5:1 ratios of XIST: rtTA/puro. The electroporation conditions were 220v, and 750 μF (BioRad Gene Pulser II System) [53]. Cells were subsequently plated on puromycin-resistant DR4 MEF feeders (Open Biosystems, Cat#: MES3948) in hiPSC medium supplemented with ROCK inhibitor for the first 24 h. Over 300 30 colonies remained after 12 days of 0.4 μg/ml puromycin selection and 245 randomly chosen individual colonies across 36 pooled wells were examined by interphase DNA/RNA FISH for the presence and expression of XIST, correct targeting and retention of trisomy (since some subclones lacked XIST or showed just two DYRK1A DNA signals). Over 100 individual clones were isolated and characterized, and those of interest, containing targeted XIST on one

of three DYRK1A loci, were frozen. Six single target clones with good pluripotent morphology, OCT4 positive staining, correct targeting to one trisomic chromosome, and good XIST RNA paint were expanded for further characterization. One double and one triple target line, two non-target clones, and one disomic clone were also isolated and frozen.
5 Targeting and correct chromosome number (47) was confirmed by interphase and metaphase FISH and genome integrity by high resolution G-band karyotype and CGH array.

**Chromosome preparation.** iPSCs were treated with 100 ng/ml KaryoMAX colcemid (Invitrogen) for 2-4 h at 37°C in a 5% $CO_2$ incubator. Cells were trypsinized,
10 treated with hypotonic solution, and fixed with methanol:acetic acid (3:1). Metaphases were spread on microscope slides, and at least 20 analyzed per clone. Karyotype analysis was done on pro-metaphase chromosomes using Standard Giemsa-trypsin G band methods.

**CGH array.** CGH was performed in the Cytogenetics Laboratory at UMASS
15 Medical School. 1 ug of DNA was used for Genomic Microarray analysis using UMass Genomic Microarray platform (Human Genome Build hg19). The array contains approximately 180,000 oligonucleotides (60 mers) that represent coding and noncoding human sequences and high density coverage for clinically relevant deletion/duplication syndromes and the telomeric and pericentromeric regions of the genome. Data was analyzed
20 by BlueFuse Multi, v3.1 (BlueGnome, Ltd).

**DNA/RNA FISH, and Immunostaining.** DNA and RNA FISH were carried out as previously described[10,19,21,58]. The XIST probe is a cloned 14 kb XIST cDNA (the same sequence as XIST transgene in Fig. 2a) in pGEM-7Zf(+) (Promega). Six Chr 21 gene probes
25 are BACs from BACPAC Resources (DYRK1A: Rp11-105O24, APP: RP11-910G8, USP25: RP11-840D8, CXADR: RP11-1150I14, ITSN1: RP11-1033C16, COL18A1: RP11-867O18). DNA probes were labeled by nick translation with either biotin-11-dUTP or digoxigenin-16-dUTP (Roche). In simultaneous DNA/RNA FISH (interphase targeting assay), cellular DNA was denatured and hybridization performed without eliminating RNA and also treated with 2
30 U/μl of RNasin Plus RNase inhibitor (Promega). For immunostaining with RNA FISH, cells were immunostained first with RNasin Plus and fixed in 4% paraformaldehyde before RNA FISH. Antibodies: H3K27me3 (Millipore, 07-449), UbH2A (Cell Signaling, 8240), H4K20Me (ABcam, ab9051), MacroH2A (Millipore, 07-219), OCT4 (Santa Cruz, sc-9081), PAX6 (Stemgent, 09-0075), SOX1 (R & D Systems, AF3369).

**Allele-specific SNP analysis**: Primers were designed to amplify 3' UTR regions of chromosome 21 genes reported to contain SNPs (Table 1). Total cDNA was used from three transgenic clones with and without XIST induction for 22 days. RT-PCR products were sequenced by GENEWIZ. Of ~10 genes examined, four were heterozygous and informative in the patient DS iPS cell line used here.

**Table 1. Primers for Chr21 gene amplification (allele-specific SNP silencing analysis)**

| Genes | Forward primer 5' – xxx - 3' | SEQ ID NO: | reverse primer 5' – xxx - 3' | SEQ ID NO: |
|---|---|---|---|---|
| ADAMTS1 | TCTCTGAAACCATAGC AGCCA | 35 | CTTGTGCAGACCATCCCTG C | 39 |
| ETS2 | GCCTTTTGCAACCAGG AACAGC | 36 | ATCACACAGAAGAACGTG GAGC | 40 |
| SPA13 | AACTCTGCTCCAAATG CCGA | 37 | CCTGTACATCATTCTCTGCT TGG | 41 |
| TIAM1 | TGGGGTGATTTGCTTTC CAGTGC | 38 | GTGCAGTGTCTGCCCCAAG C | 42 |

**Microarray analysis.** Three independently targeted subclones plus the parental Chr21 trisomic (non-targeted) iPS cell line were grown ± doxycycline (2 μg/ml) for 22 d. Normal male iPS and disomic isogenic lines were also cultured for 22d and total RNA was extracted with a High Pure RNA extraction kit (Roche) in triplicate for each, processed with a Gene Chip 3' IVT Express Kit (Affymetrix), and hybridized to Affymetrix human gene expression PrimeView arrays. Array normalization was performed with Affymetrix Expression Console Software with Robust Multichip Analysis (RMA)[59]. Probesets with the top 60% of signal values were considered present and "expressed" and were used for all further analysis. Data in Figure 5 has no other threshold applied. When designated, a gene expression change significance threshold was applied using a two-tailed T-test comparing samples ± doxycycline in triplicate (N=3) (Fig. 5d, $p<.01$). For the ~200 genes found to significantly change in all three clones (in text), a T-test with $p<.001$ was applied.

Microarray Data Interpretation: Using extraction-based methods, changes on just one of three alleles (from the XIST-bearing chromosome) will be diluted by the other two. If all three chromosomes are fully expressed, this would predict a 33% reduction in Chr21 expression levels per cell when one Chr21 is fully silenced. However, 33% would apply only if Chr 21 genes are fully over-expressed to start, and prior evidence and results in this study shows this is not the case for many genes. Previous microarray studies have analyzed

expression levels of Chr 21 in DS patient cells, although such analyses are hampered by the extensive genetic and epigenetic differences between any two individuals[60-62]. The fraction of Chr 21 genes detected as over-expressed varies with the study and tissue, but generally is in the 19-36% range[3,34,35,63], with individual gene increases often in the ~1.2-1.4 range (less than the theoretical 1.5). For example, one study of DS embryoid bodies showed only 6-15% of genes appeared significantly up-regulated, but this was comparing non-isogenic samples of different ES cell isolates.

Our trisomy correction system allows direct comparison of the *same cells* grown in identical parallel cultures, with and without XIST-mediated chromosome silencing. Our data shows a ~20% reduction in Chr 21 expression overall; importantly this level of reduction is seen either when the third chromosome is silenced in trisomic cells, or when disomic and trisomic cells are compared. This 20% reduction represents an average per cell for all three chromosomes, but corresponds to a 60% reduction in expression for just one Chr21 (the one silenced by XIST RNA, as shown here).

Apart from our goal here of trisomy dosage compensation, these results add significantly to understanding the extent of Chr 21 over-expression in Down Syndrome, by providing a more comprehensive analysis which shows expression of most genes is increased, but less than the theoretical 1.5 fold.

**qRT-PCR.** qRT-PCR was performed for eight down-regulated Chr 21 genes determined by microarray on an BIO-RAD MyiQ™ Real-Time PCR Detection System in triplicate for clone 3 with/without doxycycline treatment for 22 d. The β-actin gene was used as an internal standard for calculation of expression levels. Primers for eight Chr 21 genes and β-actin were described in Table 2.

**Table 2 Primers for qRT-PCR**

| genes | Forward primer 5' – xxx - 3' | SEQ ID NO: | reverse primer 5' – xxx - 3' | SEQ ID NO: |
|---|---|---|---|---|
| CXADR | TGCGTCTAAACGTTGTCCCT | 43 | AGTGGACGTACGGCTCTTTG | 52 |
| COL6A1 | ATCAGCCAGACCATCGACAC | 44 | GCCCTTCTCTCCCTTGTAGC | 53 |
| PTTG1IP | GTTGGGTGAACTTTGAGGCG | 45 | GTGCTGGAGCGCTTTAGTTG | 54 |
| ADAMTS1 | CCCTCACTCTGCGGAACTTTT | 46 | ATTAAGGCTGGCACACTGCTT | 55 |
| BTG3 | CCCATGTGAGGTGTGCTGT | 47 | AGGGCCCTGGTAACTTTCCT | 56 |
| TIAM1 | TCAAAACCGAGAGCCTTCCC | 48 | CGGAGACGGCATCAGAATCA | 57 |
| USP16 | AGCCTTCAGTTTGGCTGTGT | 49 | GGCTTTGGAGTTGTAATGCTGG | 58 |
| APP | GGAGCGCTCTCGACTTTTCT | 50 | TGTGCATGTTCAGTCTGCCA | 59 |
| β-ACTIN | TTGCCGACAGGATGCAGAAGGA | 51 | AGGTGGACAGCGAGGCCAGGAT | 60 |

**DNA Methylation analysis.** The parent line, and two independent targeted lines were grown with and without doxycycline for 22 d, in duplicate cultures. Genomic DNA was extracted using PureLink Genomic DNA Mini Kit (Invitrogen) and 750 ng bisulfite modified with the Alternative Incubation Conditions from the EZ DNA Methylation Kit (Zymo Research). 160 ng of bisulfite DNA was amplified, fragmented and hybridized to Illumina Infinium HumanMethylation450 array following standard protocol as outlined in the user guide. CpG islands were defined as high and intermediate CpG densities using the CpG density classifications based on those used by[64]. The program CpGIE[65] was used to locate HC and IC islands on the X chromosome and chromosomes 21 and 22. When multiple probes in CpG islands were associated with the same TSS, an average genic methylation value calculated. These average genic values were compared pre and post doxycycline induction using the Mann-Whitney test. Analysis was based on CpG islands within promoters of 143 Chr 21 genes.

The average methylation value was 6% on Chr 21 before XIST induction, and increased to 20-21% in both subclones after induction. Since any methylation increase on the transgenic chromosome would be diluted by the presence of three Chr 21s, this suggests the range of 60% methylation on the one XIST-coated chromosome, which is within the range seen for the inactive X chromosome[37].

**Cell proliferation analysis.** Eight different iPSC lines (parent line, one non-targeted subclone, and six independent targeted subclones) were passaged onto 6-well plates at equal cell densities per well of each line and grown ± doxycycline for 7 d. At least four replicates of each line were analyzed in two independent experiments. Rigorous measures were taken to minimize and control for any minor variations in seeding densities of iPS cells, which cannot be plated as single cell suspensions. First the analysis was done twice for six different transgenic clones, in each case comparing triplicate plates of corrected vs not corrected (Dox vs no Dox). To avoid differences in plating efficiencies of Dox and no Dox cells, we performed the experiments over a time course that did not require passage. For each of the six transgenic clones, the parental line and one negative control (non-targeted) sublcone, a single well of DS iPS cells (without Dox) was used to generate a cell suspension (cells and small disaggregated clumps). Next, equal aliquots of the cell suspension were plated into each of six wells once, then again, then a third time and then a fourth time (not relying on one measurement but the average of four for seeding each well). After plating, Dox was added to three of the six wells, and the cultures were maintained for 7 days. For images, plates were

fixed, stained with 1 mg/ml crystal violet (Sigma) in 70% ethanol for 30 min and scanned to generate TIFF images. For cell counts, single cells were harvested by TryPLE select and counted using Beckman Coulter Z1 Particle Counter.

5 **Differentiation of neural progenitors and irreversibility in cortical neurons.** *Differentiation*: Independent XIST-transgenic iPSC clones, and the parental DS iPS line,were dissociated with Accutase (Innovative Cell Technologies) and 4 X $10^5$ single cells were plated on Matrigel-coated 6-well plates in mTeSR1 medium (Stemcell technologies). Once the cell culture reached 90%-100% confluence, neural induction was initiated by changing
10 the culture medium to neural induction medium, a 1:1 mixture of N2- and B27-containing media supplemented with 500 ng/ml Noggin (R&D Systems), 10 μM SB431542 (Tocris Bioscience), and 1 μM retinoic acid (Sigma, cat#: R2625), with/without treatment of doxycycline for the specified times. The neural rosettes were counted and their diameter measured, for at least 300 rosettes (sampled in random areas from triplicate dishes). At Day
15 14, the dox-induced culture had an average rosette diameter of 142 μm ± 0.55 μm in Clone 1 and 141 μm ± 3.49 μm in Clone 3. Rosettes could not be measured at the same time point in the uncorrected culture, since they had not formed. At day 17, the uncorrected culture had neural rosettes of similar number and size for both Clones 1 (140 μm ± 0.87 μm) and 3 (140 μm ± 1.09 μm). The corrected culture could not be accurately compared for Day 17 because
20 the rosettes had become so mature and often had merged. After 17 d, neural rosettes were collected by dissociation with Dispase and replated on poly-ornithine and laminin-coated plastic dishes in N2- and B27-containing media including 20 ng/ml FGF2. After a further 2 d, FGF2 was withdrawn to promote differentiation of cortical neurons. *Test of the irreversibility of silencing*: Two independent clones were differentiated to cortical neurons in
25 the presence of Dox for 70 days to initiate silencing. They were then split into parallel cultures grown with and without Dox for another 30 days, and XIST and APP expression analyzed by RNA FISH.

**Targeted addition to primary fibroblasts**. Here we used non-immortalized primary human female DS fibroblasts, which like all primary fibroblasts have a limited lifespan in
30 culture (potentially more limited for DS fibroblasts). We reasoned that the robustness of ZFN-driven editing, combined with reduction to disomy for the DRYK1A gene, may make it possible to observe some edited cells before they senesce. We used a transgene carrying an near full length (~14 kb) XIST cDNA under a $TetO_2$ inducible promoter, and a selectable marker on the same construct, with ~600 bp homology arms to the DYRK1A gene (vector is

~21kb total size, with a total insert size of ~17kb) (data not shown). When introduced without the Tet-repressor construct, the $TetO_2$ CMV promoter is constitutively active. Two ZFN containing vectors as well as the 21kb XIST transgene were transfected into the primary DS fibroblasts (ATCC) using Stemfect polymer (Stemgent) (10:1 ratio of XIST to ZFN, and 13
5 ug DNA to 1.3ul Stemfect per well of 6 well plate). Surprisingly, the frequency of stable integrants was such that a sparse monolayer of transgenic fibroblasts emerged, rather than a few individual colonies following selection with hygromycin (75ug/ml). The pooled population of selected cells was analyzed by FISH and IF for targeting, XIST expression and heterochromatin marks. XIST RNA was observed over the DYRK1A locus in ~74% of cells,
10 indicating accurate transgene targeting, which was also verified by metaphase FISH. In many cells there was notable enrichment of H3K27me, H3K20me & UbH2A heterochromatic marks. Due to the limited lifespan of primary cells and the progressive silencing of the CMV promoter used in this construct, these cells were not more fully characterized.
15

**Example 1. Accurate targeted addition of a very large XIST transgene to a trisomic Chr21 in Down syndrome iPSCs:**

Given its large size, neither the *XIST* gene nor its cDNA has previously been integrated in a targeted fashion. Thus our first goal was to demonstrate feasibility of targeted
20 addition of by far the largest transgene targeted to date by nuclease-driven genome editing, orders of magnitude larger than sequences commonly used as templates for homology-directed double-strand break repair [24]. Therefore we first attempted targeted addition of a ~16 kb XIST transgene in an easily manipulated cell line (HT1080 fibrosarcoma cells), using established ZFNs to the AAVS1 locus on Chr19[25]; see Fig. 9e. This proved highly
25 successful. To extend this to Chr21, we chose the DYRK1A locus at Chr21q22 for its interest in DS (reviewed in [26]) and its potential role in pluripotency and senescence[27,28]. From this we reasoned that disrupting one of three DYRK1A alleles may enhance the likelihood of obtaining targeted trisomic pluripotent sub-clones.

Human DYRK1A gene has at least four different splicing isoforms, which differ from
30 each other either in the 5' UTR or in the 3' coding region. We engineered a ZFN heterodimer that binds a 36 bp target sequence in intron 1 of variants 1, 2, 5, or intron 3 of variant 3 (Fig. 2A) of DYRK1A and validated robust activity. Next, an even larger (~21 kb) construct was built containing near full-length XIST cDNA (17 kb), flanked by ~600 bp

homology arms (Fig. 9c). Testing in the HT1080 cells demonstrated efficient, accurate addition of the entire 21 kb transgene to the "DS critical region" of Chr21.

We next determined whether this would be achievable in the technically challenging but translationally relevant iPSCs derived from reprogramming DS patient fibroblasts. These cells have unique therapeutic and developmental potential[29] due to their ability to form a variety of cell types, and thus would represent an important target of any future *ex vivo* cellular therapy efforts. We used a male DS iPSC line from the Daley lab[30], which we confirmed maintains pluripotency markers and trisomy 21. Although a single constitutively transcribed XIST transgene could be used, we engineered an inducible system to maximize utility for investigating the biology of DS. In one step, we targeted a doxycycline-based transgene control component (rtTA) to the AAVS1 safe harbor locus on Chr19[25] (Fig. 2b & 9b), and the Dox-controlled XIST transgene to Chr21 (Fig. 2a & 9a).

We analyzed 245 colonies from the first passage of pooled transformants by dual-color interphase *in situ* RNA/DNA FISH to determine if XIST was present and overlapped one of three DYRK1A alleles. Remarkably, 99% of XIST RNA-positive colonies carried the XIST transgene at this location on Chr21, and also contained rtTA/selection transgene. Efficiency was sufficiently high that, through modifications to editing conditions, we also obtained a few sub-clones with XIST integrated into two or even all three alleles of DYRK1A (see Table 3). Six independent sub-clones were chosen for further study based on: the presence of an XIST transgene on one of three copies of Chr21; pluripotent colony morphology; robust Oct4 staining; and the ability to form embryoid bodies. Southern blotting and FISH to metaphase chromosomes confirmed the interphase FISH analysis and gene addition accuracy, and all six clones retained 47 chromosomes. Selected clones were also examined by high-resolution cytogenetic banding and/or array CGH, which showed no significant abnormalities other than full trisomy for all of Chr21 (Figs. 2c-e).

**Table 3. Accuracy of targeted addition for XIST transgene on Chr 21 in Down Syndrome iPCSs**

| Ratio of XIST to Puro | XIST+ clones (Puro+) | Random Integration | Targeted Integration | Single Target | Double Target | Triple Target |
|---|---|---|---|---|---|---|
| 3:1 | 65 | 1 (1.5%) | 64 (98.5%) | 57 (87.7%) | 7 (10.8%) | 0 (0.0%) |
| 5:1 | 16 | 1 (6.3%) | 15 (93.8%) | 8 (50.0%) | 5 (31.3%) | 2 (12.5%) |

**Example 2. XIST RNA coats the chromosome in cis and induces a heterochromatic Chr21 Barr Body:**

In the panel of six independent genome-edited clones, we induced transgene expression and detected XIST RNA by FISH three days later. XIST RNA expression was
5 consistently robust and localized in a nuclear "territory" over one Chr21, in over 85% of cells in the six clones (Fig. 2f). This mirrored the unique behavior of endogenous XIST RNA which "paints" the inactive X nuclear territory[10].

The Xi in female cells forms a visibly condensed "Barr Body" that carries an epigenetic signature of repressive histone modifications and CpG DNA methylation
10 (reviewed in[13]). Five days after XIST induction, the edited Chr21 became markedly enriched in all heterochromatin marks examined, including H3K27Me3, UbH2A, and H4K20Me in 90%-100% of cells and, later, with macroH2A (Fig. 3). H3K27me could be seen across the metaphase Chr21. Moreover, the chromosomal DNA in many nuclei became notably condensed, further evidence that we successfully generated a heterochromatic "Chr21 Barr
15 Body," which appeared, by multiple criteria, indistinguishable from the Xi in female cells.

**Example 3. XIST RNA drives long-range, allele-specific gene silencing across the targeted Chr21:**

We examined the overall transcriptional impact of XIST RNA "painting" on Chr21
20 using an approach we developed to broadly assay hnRNA by detecting CoT-1 repeat containing RNAs, which clearly distinguishes Xi from Xa[21]. The Chr21 XIST RNA territory is depleted for CoT-1 RNA, suggesting heterochromatic silencing, as on Xi.

We next used multi-color RNA FISH to determine the presence of transcription foci at each allele for six specific Chr21 genes, an established approach we developed to
25 discriminate active versus silenced genes on Xi[31]. Although XIST addition disrupts the large DYRK1A gene (Fig. 2a), without XIST expression, three bright transcription foci remained. However, when XIST RNA was induced, the targeted allele became weaker or undetectable, indicating significant repression of DYRK1A.

Next we examined the APP gene, which encodes amyloid beta precursor protein.
30 Mutations in APP (causing accumulation of $\beta$–amyloid) lead to early onset familial Alzheimer disease (EOFAD)[32], and APP over-expression is linked to AD in DS as well[33]. RNA FISH data for APP are quantified in Figure 4a. Without XIST induction, three bright RNA transcription foci for each allele were readily visualized. Brief XIST expression often

resulted in incomplete repression of the targeted allele, which after 20 days was completely silenced in both independent clones (Fig. 4a).

We extended this analysis to four more genes that ranged from 3 to 21 Mb from the XIST integration site (Figs. 4b-c): ITSN1 (Intersectin-1), USP25, CXADR, and COL18A1.
5 Complete silencing of the allele on the edited Chr21 was seen in ~100% of cells accumulating XIST RNA (Fig. 4b), demonstrating silencing of the XIST-associated allele. Allele-specific silencing was also validated using SNP analysis. RT-PCR products for eight known polymorphic sites (in four genes) were sequenced (ADAMTS1, ETS2, TIAM1, and HSPA13) (Figs. 4d-e). Interestingly, clones 2 and 3 showed the identical pattern of eight
10 SNP alleles repressed, whereas clone 1 showed an alternate pattern. As summarized in Fig. 4e, this chromosome-wide pattern allows extrapolation of the haplotype for each of the three Chr21s, and indirectly identifies for each clone which Chr21 homolog integrated XIST.

We also examined APP silencing in clones carrying XIST on two or all three copies of Chr21. After 20 days of dox, most or all cells carrying XIST on two or three Chr21s,
15 respectively, no longer accumulated XIST RNA across the chromosome, and thus failed to silence the APP gene (Fig. 4f). These data argue there is *in vitro* selection against creating a functional monosomy or nullisomy, consistent with the lethality of any monosomy *in vivo,* and clinical observations that cells monosomic for Chr21 do not persist in mosaic patients.

20 **Example 4. Genome-wide expression analysis demonstrates transcriptional repression across the edited Chr21:**

The above approaches demonstrate XIST RNA induces a heterochromatic Chr21 Barr Body and allele-specific repression for the nine genes examined, yet we extended this to include genome-wide expression profiling. Three independent transgenic clones and the
25 parent line were treated with Dox for three weeks, and their transcriptome compared to parallel cultures without XIST-transcription, all in triplicate. Strikingly, only on Chr21 is there overwhelming change, in all three clones (Fig. 5a), with ~95% of genes significantly expressed showing repression (Fig. 5d).

Figure 5a summarizes the key finding that dosage compensation of trisomy corrects
30 Chr21 expression to near normal disomic levels. This is based on calculation of the change in total output of expressed genes per chromosome after XIST is induced. Since evidence indicates that many Chr21 genes are not increased to the theoretical maximum of 1.5 fold in trisomic cells ([34-36] and further explained above), we included a direct comparison to trisomic versus disomic cells; this provides a baseline to evaluate the degree to which Chr21 over-

expression is corrected by XIST-mediated silencing. After XIST induction, overall Chr21 expression is reduced by 20%, 15%, and 19% for clones 1, 2, and 3, respectively; this mirrors very well the 22% reduction for disomic iPS cells that lack the third Chr21 altogether (Fig. 5a). This disomic iPS line is representative, as a similar (21%) Chr21 difference was seen for
5 another isogenic disomic sub-clone recently isolated from the parental DS iPS cells. Figure 5b shows that individual genes repressed by XIST distribute across Chr21, as do genes over-expressed in trisomic versus disomic cells. In addition, qRT-PCR confirmed repression for all eight Chr21 genes examined (Fig. 5d). Taken together, these results clearly demonstrate that XIST induces robust dosage compensation of most over-expressed genes throughout the
10 length of Chr21.

Trisomy 21 likely has broader impact on genomic expression pathways (e.g.,[36]), but the differences attributable to trisomy 21 are confounded by genetic and epigenetic variability. This inducible trisomy correction system provides a new foothold into that important question. For example, microarray profiles of our three independent transgenic
15 sub-clones reveal that even these isogenic sub-clones show many expression differences (>1000) throughout the genome, but upon XIST induction, a smaller cohort of genes (~200) change in common in all three clones (but not the dox-treated parental line); this cohort is more likely due to Chr21 over-expression. While not our focus here, these findings support the promise for “trisomy correction in a dish” as a means to identify genome-wide pathways
20 perturbed by trisomy of Chr21.

**Example 5. Chromosome-wide methylation of genes on the XIST-carrying Chr21:**

X-inactivation in female cells is further stabilized by hypermethylation of DNA in
25 promoter CpG islands[37-39], which occurs late in the silencing process. Therefore, we examined the promoter methylome in two independent genome-edited clones three weeks after XIST induction. The global promoter methylome remained largely unaltered, with one striking exception (P-value<2.2e-16): the genes on Chr21 (Fig. 5c). Here, 97% of CpG-island-containing genes exhibited a robust increase in promoter DNA methylation on Chr21,
30 within the range of that seen for Xi[37] (when adjusted for the number of active versus silenced chromosomes: see Methods). This change swept across the entire chromosome (Fig. 5c), strongly reinforcing above analyses on gene expression. Interestingly, the fact that a small subset of specific genes “escape” methylation on Chr21 in both clones demonstrates the

impact of DNA sequence on XIST-mediated silencing (as long suggested[15,18,40] and reviewed in[41]).

The sum total of data, from eight different approaches, demonstrates an impressive competence of most sequences across Chr21 to undergo epigenetic modification and silencing in response to XIST RNA, an RNA evolved to silence the X-chromosome.

**Example 6. Chr21 dosage compensation impacts cell phenotype to enhance cell proliferation and neural rosette formation:**

Correction of whole chromosome imbalance by manipulating just one gene presents a new paradigm, with opportunities to advance DS research in multiple directions. Currently, the specific cellular processes perturbed by trisomy 21 that generate patient pathology are largely unknown. Inducing trisomy silencing in parallel cultures of otherwise identical cells may reveal cellular pathologies due to trisomy 21, which could be obscured by differences between cell isolates. To address whether an impact in cell phenotype could be discerned, we examined two properties—cell proliferation and neural rosette formation.

There is some evidence of proliferative impairment in DS[42,43], however we found this was variable between DS fibroblast cell samples, and highly sensitive to culture history and population doublings. However, a clear answer emerged from comparing multiple transgenic clones, grown in the presence or absence of doxycycline for one week. Initial analysis of clones 1 and 2 in triplicate indicated that XIST-induction rapidly resulted in larger, more numerous and more tightly packed cell colonies. This analysis was repeated for six independent transgenic sub-clones, the parental line, and a trisomic sub-clone, each replicated 4-6 times, minimizing technical variations in plating and counting iPS cells (Methods). All transgenic clones showed larger, more tightly packed colonies after just seven days of XIST induction, which contained 18-34% (average 26%) more cells than uninduced cultures (Fig. 6a). In contrast, Dox did not enhance growth of the parental DS cells or sub-clone (Fig. 6a). Thus, a proliferative impairment linked to Chr21 over-expression can be rapidly ameliorated by dosage compensation. Interestingly, this effect is not dependent on DYRK1A silencing[27,28], since the DYRK1A locus is disrupted irrespective of XIST expression.

We next examined differentiation into neural progenitor cells, using a protocol to derive cortical neurons[44]. Six replicate cultures for the parental DS iPSC line and clones 1 and 3 were grown to confluency, placed in neural differentiation medium, and half of the identical samples induced to express XIST. Just 11-12 days after neural induction, all XIST-expressing cultures (in triplicate for both clones) began to form neural rosettes, and in 1-2

days were replete with neural rosettes. These cell structures are a signature of neural progenitors, and were confirmed by expression of Pax6 and Sox1. Remarkably, even at day 14, parallel uninduced cultures were still devoid of any neural rosettes (Fig. 6b). Thus uncorrected cultures required 4-5 more days in neural-induction media to fill with neural rosettes of similar size and number, as they did on day 17 (Fig.6b). This difference is due to XIST, as there was no effect of Dox on neurogenesis in the parental DS line. This marked delay in neural differentiation appears primarily independent of cell proliferation (Methods). A similar difference occurred in repeat experiments with clones 1 and 2. Variability in the kinetics of neural differentiation that exists between various iPS cell lines[45] would likely obscure differences due to trisomy 21. We circumvented this using parallel cultures and on-demand Chr21 silencing, which made clear these important phenotypic differences.

These data highlight the potential of this new experimental model to identify and study cellular pathologies directly attributable to over-expression of Chr21 in iPSCs and their differentiated progeny.

**Example 7. Stable Chr21 silencing and successful targeting of XIST in DS primary fibroblasts:**

Finally, we briefly consider two points relevant to any future potential for *ex vivo* or *in vivo* therapeutic strategies. While a constitutively expressed XIST transgene could be used, it is advantageous if the heterochromatic state induced by XIST RNA is stably maintained, even if XIST is no longer expressed (as reported in mouse[46]). We tested this in our human Chr21 system by removing dox and XIST expression for 30 days, after iPS cells had silenced Chr21 and differentiated to neurons. As shown (Fig 6c), APP gene silencing remained indistinguishable between cultures with and without continued XIST expression, supporting other evidence that in somatic cells multi-layered chromatin modifications triggered by XIST maintain a largely irreversible silent state [39,47].

Finally, we considered the forward-looking question of whether targeted XIST addition could be achieved in primary human cells, as tested in non-immortalized female DS fibroblasts. Surprisingly, in our first attempt we generated not a few sub-clones but a sparse monolayer of edited fibroblasts, most of which carried XIST on Chr21. Due to limited lifespan, these cells were not examined in depth, but notably many showed enrichment of H3K27me3, H3K20me, and UbH2A at the transgene site. This is consistent with evidence that chromosome silencing does not necessarily require the optimal pluripotent cell context. Although pluripotent cells clearly have the optimal capacity to rapidly and fully silence

chromatin in response to XIST RNA (Wutz et al., Mol Cell 5, 695-705 (2000)) , several observations indicate the pluripotent cell context is not necessarily required. For example, random integration of an XIST transgene into human HT1080 cells (a transformed cell line) produced a robust Barr Body (on a Chr4 autosome), although this took longer than in
5 pluripotent cells (Hall et al., Proc Natl Acad Sci U S A 99, 8677-8682. (2002)). Similarly, gene silencing has been seen in other somatic cell lines (Chow et al., Cytogene Genome Research 99, 92-98 (2002); Chow et al., Genomics 82, 309-322 (2003)). Savarese et al. (Mol Cell Biol 26, 7167-7177 (2006)) reported that hematopoietic cells in mouse bone marrow are still capable of Xist-mediated chromosomal inactivation. The Wutz lab also reported that
10 addition of SATB1 to mouse fibroblasts can enhance their ability to silence chromatin in response to XIST RNA (Dev Cell 16, 507-516 (2009)). Data herein suggests that primary human fibroblasts still exhibit significant capacity to induce heterochromatin modifications in response to XIST. In addition, we have data in differentiated mouse and human ES/iPS cells that demonstrate cells in the neuronal pathway can silence chromatin in response to XIST
15 RNA. Finally, our XIST transgene lacks X-chromosome “counting” sequences, and thus is compatible with natural X-inactivation in female cells.

**Example 8. DYRK1A expression is not disrupted by XIST insertion**

Human DYRK1A gene has at least four different splicing isoforms, which differ from
20 each other either in the 5’ UTR or in the 3’ coding region. We inserted XIST transgene into intron 1 of variants 1, 2, 5, or intron 3 of variant 3 (Fig. 2A, bottom panel). To investigate whether DYRK1A expression is disrupted by XIST insertion, we performed RT-PCR and sequencing for parental line and triple target line. Parental line does not have XIST insertion and triple target line has XIST transgene inserted
25 into all three alleles of DYRK1A gene. We designed two sets of primers for RT-PCR. The first set spans the XIST target site that contains 5’ UTR and coding region between exon 1 and exon 2 of variants 1, 2, 5, or between exon 3 and exon 4 of variant 3. The second set amplifies the 3’-end of coding region that spans the sequence between exon 9 and exon 11 of variants 1, 2, 5, or between exon 11 and
30 exon 13 of variant 3. The first set of primers is expected to generate a 202 bp product for all four variants, and the second set of primers is expected to generate a 324 bp product for variant 5 (variant 5 does not contain exon 10), a 449 bp product for variants 1 and 2, and a 580 bp product for variant 3 (Table 4). Genbank accession number: variant 1 (NM_001396.3), variant 2 (NM_130436.2), variant 3
35 (NM_101395.2), and variant 5 (NM_130438.2).

**Table 4**

| Primer sets | product size | amplified variants |
|---|---|---|
| set 1 | 202 bp | variants 1,2,3,5 |
| set 2 | 324 bp | variant 5 |
| | 449 bp | variants 1,2 |
| | 580 bp | variant 3 |

RT-PCR showed that the first set of primers generated a 202 bp band, and the second set of primers generated only one 449 bp of single band in both parental and triple target lines. Sequencing results confirmed that the 202 bp product from the first set of primers in both lines is the sequence spanning exon 1 and exon 2 of variants 1, 2, 5, or spanning exon 3 and exon 4 of variant 3. Sequencing for the second set of primers confirmed that the 449 bp product in both lanes was the sequence spanning exon 9 and exon 11 of variants 1 and 2, indicating that these Down syndrome iPSC lines only contain DYRK1A variant 1 and/or variant 2. These results demonstrate that DYRK1A expression was not disrupted by XIST insertion, which is consistent with the microarray data in which DYRK1A expression level in all three transgenic subclones (without Dox treatment) is not down-regulated compared with that in parental line.

Studies suggest DYRK1A plays an important role in cell proliferation and neurogenesis. This result is important because it rules out the possibility that phenotypic features of the trisomy-silencing cells are impacted by disruption of DYRK1A gene, prior to trisomy silencing by XIST RNA.

**Example 9. Targeting XIST to Alternative Locations on Chr.21: Targeting RCAN1**

Methods as described above were used to create targeting constructs including the 6.8 kb inducible/selectable XIST transgene or the 14 kb full length XIST transgene as shown in Figs. 7a-c. The lengths are shown in Table 5.

Table 5. chr.21 RCAN1 targeting constructs

| | Left arm | Right arm | | Total construct length | Insert length between two arms |
|---|---|---|---|---|---|
| RCAN1 | 759bp | 758bp | 6.8kb | 14026bp | 10108bp |
| | | | FL | 21055bp | 17137bp |

The constructs were introduced into cells as described above. Integration of the transgene and localization of XIST RNA were confirmed by interphase and metaphase FISH; the results are shown in Figures 8a-f. These data demonstrate the feasibility of using ZFN-driven genome editing to direct an entire or active XIST cassette to different loci of the "DS critical region" of Chr21.

**Example 10. Generation of trisomy corrected mouse models of Down's syndrome carrying an extra human chromosome 21**

Several mouse models have been developed that recapitulate many of the phenotypic and anatomical pathologies of Down's syndrome. Two Down's syndrome mouse models (Tc1 and Ts65Dn) that are widely used have a number of well-characterized abnormalities in multiple organ systems. This example uses the Tc1 mouse model to evaluate the effects of silencing the trisomic chromosome in these mice.

The Tc1 mouse strain contains a freely segregating human fragment (~90%) of chromosome 21 containing 269 genes that have been found to contribute to human Down's syndrome. This transchromosomic mouse line represents the most complete model of Down's syndrome, exhibiting deficits in learning and memory, synaptic plasticity, motor coordination, and heart development (O'Doherty et al., Science 309, 2033-2037 (2005); Hernandez et al., Hum Mol Genet 8, 923-933 (1999); Galante et al., Hum Mol Genet 18, 1449-1463 (2009); Alford et al,. Blood 115, 2928-2937). The human chromosome 21 ZFNs and doxycycline-inducible XIST constructs described herein were used to successfully target the Tc1 ES cells which were originally used to create the Tc1 Down's syndrome mouse model. *XIST*-targeted clones with robust *XIST* accumulation are being isolated.

In addition, two mouse Xist constructs were designed to target human chromosome 21. The pEF1α/hDYRK1A/FL mXist construct (20647 bp) contains full length mouse Xist cDNA and two homologous arms that are designed to target intron 1 of DYRK1A gene on human chromosome 21. The construct was generated as follows. The 15 kb of full length mouse Xist cDNA (15 kb of the 17.9 kb sequence at Genbank accession number: NR_001463.3; see SEQ ID NO:62) was subcloned into pTRE3G vector (Clontech, Cat#: 631167). Two homologous arms (left arm, 690 bp; right arm, 508 bp) of *DYRK1A* gene on human chromosome 21 were amplified by PCR from primary DS fibroblasts (AG13902) (Coriell) and cloned into the pTRE3G vector (Human Chr21 *DYRK1A* left arm primers: forward 5'-GCCGTATACCATTAACTCTTTACTGTTC-3' (SEQ ID NO:1), reverse 5'-TCTGTATACGTAAACTGGCAAAGGGGTGG-3' (SEQ ID NO:2); Human Chr21

*DYRK1A* right arm primers: forward 5'- ATTTCGCGAACGGGTGATGAGCAGGCTGT-3' (SEQ ID NO:3), reverse 5'- CCGTCGCGAAAACCAGAAAGTATTCTCAG-3' (SEQ ID NO:4)). The inducible pTRE3G promoter on the pTRE3G vector was replaced by a constitutive promoter PEF1α (from pEF1α-Tet3G vector, Clontech). See Fig. 9g, 10I and 5 SEQ ID NO:22 for sequence information.

In order to generate a trisomy corrected Tc1 mouse model without disturbing other part of the genome, the same dual-targeted-addition strategy described above was used: to target the mouse Xist transgene into human chromosome 21 and a selection gene into chromosome 6. The Rosa26/pEF1x-Tet3G/hPGK-PuroR construct was made, which 10 contains a puromycin resistance selection gene and rtTA cassette that is targeted to the Rosa26 locus on mouse chromosome 6 by ZFNs (Rosa26 locus is the safe harbor of mouse genome). Mouse Rosa26 ZFNs were purchased from Addgene and the pEF1x-Tet3G/hPGK-PuroR cassette from AAVS1/pEF1x-Tet3G/ hPGK-PuroR vector was subcloned into a plasmid for targeted gene addition to the Rosa26 locus of mouse genome (Addgene, Cat#: 15 37200 ), which contains a unique PmeI site flanked by two 800 bp stretches of homology to the ZFN-specified position in the genome. See Fig. 9i, 10K and SEQ ID NO:24 for sequence information. This construct would be useful in any situation in which you do not want the selection gene directly on the transgene. This is important for Xist transgenes (which could silence the selection gene), but also in any inducible transgene which the enhancer in a 20 tandem selection gene could cause "leaky" expression of the transgene.

These constructs were used to silence the human chromosome 21 in the Tc1 Down syndrome mouse model, which contains an extra copy of human chromosome 21 in mouse context. The Tc1 mouse embryonic stem (ES) cells were cultured in mouse ES cell growth medium. Single cells ($1X10^7$) were harvested using trypsin (Invitrogen), resuspended in 1 x 25 PBS and electroporated with a total of 59 μg DNA including five plasmids (pEF1α/hDYRK1A/FL mXist, DYRK1A ZFN1, DYRK1A ZFN2, Rosa26/pEF1x-Tet3G/hPGK-PuroR, and Rosa26 ZFN) with 3:1 ratio of pEF1α/hDYRK1A/FL mXist: Rosa26/pEF1x-Tet3G/hPGK-PuroR. The electroporation conditions were 230v, and 500 μF (BioRad Gene Pulser II System). Cells were subsequently plated on puromycin-resistant 30 MEF feeders in Mouse ES cell growth medium. 288 colonies were picked after 12 days of 3.0 μg/ml puromycin selection and examined by interphase DNA/RNA FISH for targeted clones. In some targeted subclones, the mouse Xist RNA did not accumulate on the targeted human chromosome 21. In another pool of Tc1 cells that contained eight different subclones,

targeted cells showed robust Xist paint that appears to silence the DYRK1A gene on the trisomic human chromosome 21.

**Example 11. Minimal Mouse XIST Silencing Constructs for In Vivo Delivery**

5 In vivo chromosome therapy would have significant impact on the potential development of human therapies for Down's syndrome patients. The large size (15 kb) of the full length of mouse Xist cDNA could complicate delivery into cells using AAV-based gene delivery approach. For in vivo delivery of Xist transgene to cells within the body, a reduced size Xist (5-6 kb) was generated which can be packaged in AAV vector. This second
10 construct, pEF1α/hDYRK1A/6.3kb mXist (12230 bp), contains 6.3 kb of mouse Xist cDNA that has been reported to function (Wutz et al., Nat Genet 30, 167-174 (2002)) and two homologous arms that were designed to target intron 1 of DYRK1A gene on human chromosome 21. The construct was generated as follows. 6.3 kb mouse Xist cDNA that contains exon 1 of Xist gene (SEQ ID NO:61) was subcloned into pTRE3G vector (Clontech,
15 Cat#: 631167). Two homologous arms (left arm, 690 bp; right arm, 508 bp) of DYRK1A gene on human chromosome 21 were amplified by PCR from primary DS fibroblasts (AG13902) (Coriell) and cloned into the pTRE3G vector (Human Chr21 DYRK1A left arm primers: forward 5'-GCCGTATACCATTAACTCTTTACTGTTC-3' (SEQ ID NO:1), reverse 5'- TCTGTATACGTAAACTGGCAAAGGGGTGG-3' (SEQ ID NO:2); Human
20 Chr21 DYRK1A right arm primers: forward 5'-ATTTCGCGAACGGGTGATGAGCAGGCTGT-3' (SEQ ID NO:3), reverse 5'-CCGTCGCGAAAACCAGAAAGTATTCTCAG-3' (SEQ ID NO:4)). The inducible pTRE3G promoter on the pTRE3G vector was replaced by a constitutive promoter PEF1α (from pEF1α-Tet3G vector, Clontech). See Fig. 9h, 10J and SEQ ID NO:23 for sequence
25 information. The smaller Xist transgene is used to test the Xist-mediated chromosome therapy in affected organs *in vivo*.

**Example 12. Transplantation of neural stem cells in Down's syndrome mouse brain.**

30 Children with Down's syndrome have variable but significant levels of cognitive impairment, which limits the independence of Down's syndrome patients and adversely impacts their quality of life. MRI studies reveal that Down's syndrome children and young adults have smaller overall brain volumes with more notable deficits in the hippocampus. Hippocampal volume continues to decrease with age in Down's syndrome individuals, which

is inversely correlated with the degree of cognitive impairment (Smigielska-Kuzia et al., Neurol Neurochir Pol 45, 363-369 (2011); Pinter et al., Neurology 56, 972-974 (2001)). Neurological studies from both Down's syndrome patients and mouse models also suggest a hypocellularity in the hippocampus persists over life, indicating this may be amenable to stem
5 cell therapies (Lorenzi & Reeves, Brain Res 1104, 153-159 (2006); Guidi et al., Brain Pathol 21, 361-373 (2011); Guidi et al., Brain Pathol 18, 180-197 (2008)).

The human iPSC-based trisomy correction model system described herein provides a uniquely powerful resource for the study of transplantation therapies in Down's syndrome. We have successfully differentiated the trisomy corrected and non-corrected cells into neural
10 progenitors with 90% high efficiency and preliminary results indicate the human iPSC-derived neurons are capable of forming synapses in vitro. To study the in vivo function of our in vitro-produced human trisomy corrected neural progenitors, the Tc1 mouse model is used to test if the trisomy corrected human iPSC-derived neurons participate in the established mouse neural network. The trisomy corrected neural progenitors are transplanted
15 to both sides of hippocampi of immunosupressed Tc1 mice (n=10) (group 1) and the control groups will be (1) the trisomic line without XIST induction (group 2); (2) parental trisomic line (group 3); (3) one isogenic disomic line as positive control (group 4). Each group has 10 mice (40 in total). The human-specific nuclear protein, as well as neuronal subtype and synaptic markers will be used to identify the resulting subtype of neurons and synaptic
20 formation between human neurons and mouse neurons

Human embryonic stem cell-derived neurons can functionally integrate into an existing neural circuitry and regulate the activity of an established mouse neural network after transplantation into the mouse brain (Weick et al., Proceedings of the National Academy of Sciences of the United States of America 108, 20189-20194 (2011); Espuny-Camacho et
25 al., Neuron 77, 440-456 (2013); Muotri et al., Proceedings of the National Academy of Sciences of the United States of America 102, 18644-18648 (2005); Acharya et al., Proc Natl Acad Sci U S A 106, 19150-19155 (2009)). Human ESC-derived neural progenitors transplanted into the hippocampus of mice correct learning and memory deficits (Liu et al., Nat Biotechnol 31, 440-447 (2013)). To assess whether the resulting neurons functionally
30 integrate into neural circuits in the hippocampus of Tc1 mice, brain-slice electrophysiological recordings of human iPSC-derived neurons identified by mCherry expression under control of a synapsin promoter, which was introduced into the progenitors using a lentivirus before transplantation as described (Liu et al., Nat Biotechnol 31, 440-447 (2013)). Morris water maze test for all mice in each of the four groups, before and after cell transplantation, are

used to evaluate functional improvement in hippopcampus-related learning and memory deficits.

REFERENCES

5 1 Antonarakis, S. E. & Epstein, C. J. The challenge of Down syndrome. Trends in molecular medicine 12, 473-479 (2006).
2 Megarbane, A. et al. The 50th anniversary of the discovery of trisomy 21: the past, present, and future of research and treatment of Down syndrome. Genetics in medicine : official journal of the American College of Medical Genetics 11, 611-616 (2009).
10 3 Prandini, P. et al. Natural gene-expression variation in Down syndrome modulates the outcome of gene-dosage imbalance. Am J Hum Genet 81, 252-263 (2007).
4 O'Doherty, A. et al. An aneuploid mouse strain carrying human chromosome 21 with Down syndrome phenotypes. Science 309, 2033-2037 (2005).
5 Reeves, R. H. Down syndrome mouse models are looking up. Trends Mol Med 12,
15 237-240 (2006).
6 Liu, C. et al. Mouse models for Down syndrome-associated developmental cognitive disabilities. Dev Neurosci 33, 404-413 (2011).
7 Gardiner, K. J. Molecular basis of pharmacotherapies for cognition in Down syndrome. Trends Pharmacol Sci 31, 66-73 (2010).
20 8 Lyon, M. Gene Action in the X-chromosome of the Mouse (Mus musculus L.). Nature 190, 372-373 (1961).
9 Brown, C. J. et al. 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 (1992).
25 10 Clemson, C. M., McNeil, J. A., Willard, H. F. & Lawrence, J. B. 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 (1996).
11 Hall, L. L. & Lawrence, J. B. The cell biology of a novel chromosomal RNA: chromosome painting by XIST/Xist RNA initiates a remodeling cascade. Semin Cell Dev
30 Biol 14, 369-378 (2003).
12 Heard, E. Delving into the diversity of facultative heterochromatin: the epigenetics of the inactive X chromosome. Curr Opin Genet Dev 15, 482-489 (2005).
13 Wutz, A. Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. Nat Rev Genet 12, 542-553 (2011).
35 14 Lee, J. T. Gracefully ageing at 50, X-chromosome inactivation becomes a paradigm for RNA and chromatin control. Nat Rev Mol Cell Biol 12, 815-826 (2011).
15 Bailey, J. A., Carrel, L., Chakravarti, A. & Eichler, E. E. Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. Proc Natl Acad Sci U S A 97, 6634-6639. (2000).
40 16 Brown, C. J., Carrel, L. & Willard, H. F. Expression of genes from the human active and inactive X chromosomes. Am J Hum Genet 60, 1333-1343 (1997).
17 Carrel, L. & Willard, H. F. X-inactivation profile reveals extensive variability in X-linked gene expression in females. Nature 434, 400-404 (2005).
18 McNeil, J. A., Smith, K. P., Hall, L. L. & Lawrence, J. B. Word frequency analysis
45 reveals enrichment of dinucleotide repeats on the human X chromosome and [GATA]n in the X escape region. Genome Res 16, 477-484 (2006).
19 Hall, L. L., Clemson, C. M., Byron, M., Wydner, K. & Lawrence, J. B. Unbalanced X;autosome translocations provide evidence for sequence specificity in the association of

XIST RNA with chromatin. Hum Mol Genet 11, 3157-3165. (2002).
20 Wutz, A. & Jaenisch, R. A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. Mol Cell 5, 695-705 (2000).
21 Hall, L. L. et al. An ectopic human XIST gene can induce chromosome inactivation in postdifferentiation human HT-1080 cells. Proc Natl Acad Sci U S A 99, 8677-8682. (2002).
22 Lee, J. T., Strauss, W. M., Dausman, J. A. & Jaenisch, R. A 450 kb transgene displays properties of the mammalian X-inactivation center. Cell 86, 83-94 (1996).
23 Moehle, E. A. et al. Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. Proc Natl Acad Sci U S A 104, 3055-3060 (2007).
24 Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S. & Gregory, P. D. Genome editing with engineered zinc finger nucleases. Nat Rev Genet 11, 636-646 (2010).
25 DeKelver, R. C. et al. Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. Genome Research 20, 1133-1142 (2010).
26 Park, J., Song, W. J. & Chung, K. C. Function and regulation of Dyrk1A: towards understanding Down syndrome. Cellular and molecular life sciences : CMLS 66, 3235-3240 (2009).
27 Yabut, O., Domogauer, J. & D'Arcangelo, G. Dyrk1A overexpression inhibits proliferation and induces premature neuronal differentiation of neural progenitor cells. J Neurosci 30, 4004-4014 (2010).
28 Litovchick, L., Florens, L. A., Swanson, S. K., Washburn, M. P. & DeCaprio, J. A. DYRK1A protein kinase promotes quiescence and senescence through DREAM complex assembly. Genes & development 25, 801-813 (2011).
29 Gurdon, J. B. & Yamanaka, S. The Nobel Prize in Physiology or Medicine 2012, <http://www.nobelprize.org/nobel_prizes/medicine/laureates/2012/press.html> (2012).
30 Park, I. H. et al. Disease-specific induced pluripotent stem cells. Cell 134, 877-886 (2008).
31 Clemson, C. M., Hall, L. L., Byron, M., McNeil, J. & Lawrence, J. B. The X chromosome is organized into a gene-rich outer rim and an internal core containing silenced nongenic sequences. Proc Natl Acad Sci U S A 103, 7688-7693 (2006).
32 Tanzi, R. E. & Bertram, L. Twenty years of the Alzheimer's disease amyloid hypothesis: a genetic perspective. Cell 120, 545-555 (2005).
33 Webb, R. L. & Murphy, M. P. beta-Secretases, Alzheimer's Disease, and Down Syndrome. Curr Gerontol Geriatr Res 2012, 362839 (2012).
34 Biancotti, J. C. et al. Human embryonic stem cells as models for aneuploid chromosomal syndromes. Stem Cells 28, 1530-1540 (2010).
35 Lockstone, H. E. et al. Gene expression profiling in the adult Down syndrome brain. Genomics 90, 647-660 (2007).
36 Ait Yahya-Graison, E. et al. Classification of human chromosome 21 gene-expression variations in Down syndrome: impact on disease phenotypes. Am J Hum Genet 81, 475-491 (2007).
37 Cotton, A. M. et al. Chromosome-wide DNA methylation analysis predicts human tissue-specific X inactivation. Human genetics 130, 187-201 (2011).
38 Sharp, A. J. et al. DNA methylation profiles of human active and inactive X chromosomes. Genome research 21, 1592-1600 (2011).
39 Csankovszki, G., Nagy, A. & Jaenisch, R. Synergism of Xist RNA, DNA methylation, and histone hypoacetylation in maintaining X chromosome inactivation. J. of Cell Biol. 153, p. 773-783 (2001).
40 Carrel, L. et al. Genomic environment predicts expression patterns on the human

inactive X chromosome. PLoS Genet 2, e151 (2006).
41 Hall, L. L. & Lawrence, J. B. XIST RNA and architecture of the inactive X chromosome: implications for the repeat genome. Cold Spring Harb Symp Quant Biol 75, 345-356 (2010).
5 42 Guidi, S., Ciani, E., Bonasoni, P., Santini, D. & Bartesaghi, R. Widespread proliferation impairment and hypocellularity in the cerebellum of fetuses with down syndrome. Brain Pathol 21, 361-373 (2011).
43 Haydar, T. F. & Reeves, R. H. Trisomy 21 and early brain development. Trends Neurosci 35, 81-91 (2012).
10 44 Shi, Y., Kirwan, P., Smith, J., Robinson, H. P. & Livesey, F. J. Human cerebral cortex development from pluripotent stem cells to functional excitatory synapses. Nat Neurosci 15, 477-486, S471 (2012).
45 Shi, Y. et al. A human stem cell model of early Alzheimer's disease pathology in Down syndrome. Sci Transl Med 4, 124ra129 (2012).
15 46 Csankovszki, G., Panning, B., Bates, B., Pehrson, J. R. & Jaenisch, R. Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X inactivation [letter]. Nat Genet 22, 323-324 (1999).
47 Ohhata, T. & Wutz, A. Reactivation of the inactive X chromosome in development and reprogramming. Cellular and molecular life sciences : CMLS (2012).
20 48 Li, L. B. et al. Trisomy correction in down syndrome induced pluripotent stem cells. Cell Stem Cell 11, 615-619 (2012).
49 Lavon, N. et al. Derivation of euploid human embryonic stem cells from aneuploid embryos. Stem cells 26, 1874-1882 (2008).
50 Morey, C. & Avner, P. The demoiselle of X-inactivation: 50 years old and as trendy
25 and mesmerising as ever. PLoS genetics 7, e1002212 (2011).
51 Brockdorff, N. Chromosome silencing mechanisms in X-chromosome inactivation: unknown unknowns. Development 138, 5057-5065 (2011).
52 Minkovsky, A., Patel, S. & Plath, K. Concise review: Pluripotency and the transcriptional inactivation of the female Mammalian X chromosome. Stem cells 30, 48-54
30 (2012).
53 Hockemeyer, D. et al. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. Nat Biotechnol 27, 851-857 (2009).
54 Doyon, J. B. et al. Rapid and efficient clathrin-mediated endocytosis revealed in genome-edited mammalian cells. Nat Cell Biol (2011).
35 55 Miller, J. C. et al. An improved zinc-finger nuclease architecture for highly specific genome editing. Nat Biotechnol 25, 778-785 (2007).
56 Guschin, D. Y. et al. A rapid and general assay for monitoring endogenous gene modification. Methods in molecular biology 649, 247-256 (2010).
57 Urnov, F. D. et al. Highly efficient endogenous human gene correction using designed
40 zinc-finger nucleases. Nature 435, 646-651 (2005).
58 Tam, R., Smith, K.P., and Lawrence, J.B. The 4q subtelomere harboring the FSHD locus is specifically anchored with peripheral heterochromatic unlike most human telomeres. Journal of Cell Biology 167, 269-279 (2004).
59 Irizarry, R. A. et al. Summaries of Affymetrix GeneChip probe level data. Nucleic
45 acids research 31, e15 (2003).
60 Gardiner, K. Gene-dosage effects in Down syndrome and trisomic mouse models. Genome Biol 5, 244 (2004).
61 Antonarakis, S. E. & Epstein, C. J. The challenge of Down syndrome. Trends Mol Med 12, 473-479 (2006).
50 62 Antonarakis, S. E., Lyle, R., Chrast, R. & Scott, H. S. Differential gene expression

studies to explore the molecular pathophysiology of Down syndrome. Brain Res Brain Res Rev 36, 265-274 (2001).
63 Tang, Y. et al. Blood expression profiles for tuberous sclerosis complex 2, neurofibromatosis type 1, and Down's syndrome. Ann Neurol 56, 808-814 (2004).
5 64 Weber, M. et al. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet 39, 457-466 (2007).
65 Wang, Y. & Leung, F. C. An evaluation of new criteria for CpG islands in the human genome as gene markers. Bioinformatics 20, 1170-1177 (2004).

## OTHER EMBODIMENTS

10 It is to be understood that while the invention has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

WHAT IS CLAIMED IS:

1. A silencing vector comprising:
a silencing element comprising a silencing sequence flanked by first and second
5 targeting sequences, wherein each of the first and second targeting sequences are homologous to at least 50 bp in, or within 1 MB of, the dual specificity tyrosine-phosphorylation-regulated kinase 1A (DYRK1A) gene or the Regulator of calcineurin 1 (RCAN1) gene; and
a promoter operably linked to the silencing element.

10 2. The silencing vector of claim 1, wherein the vector is a plasmid or a viral vector.

3. The silencing vector of claim 2, wherein the viral vector is vaccinia virus, adeno-associated virus (MV), or herpes virus.

4. The silencing vector of claim 1, wherein the silencing vector targets DYRK1A, and the first targeting sequence comprises a sequence homologous to at least 50 bp of the
15 DYRK1A gene, e.g., of intron 1 of DYRK1A; or
wherein the silencing vector targets RCAN1, and the first targeting sequence comprises a sequence homologous to at least 50 bp of the RCAN1 gene, e.g., of intron 3 of RCAN1.

5. The silencing vector of claims 1-4, wherein the silencing element comprises a human
20 XIST cDNA or functional fragment thereof.

6. The silencing vector of claim 5, further comprising a selectable marker sequence.

7. The silencing vector of claim 6, wherein the selectable marker sequence is operably linked to a promoter.

8. A silencing vector comprising the sequence shown in SEQ ID NOs: 14, 16, 17, 18,
25 19, 20, 21, 22, or 23.

9. A method of reducing levels of expression of genes on Chromosome 21 in a cell, the method comprising contacting the cell with the silencing vector of any of claims 1-8, under conditions sufficient for the silencing vector to undergo homologous

recombination with the genomic DNA of the cell, wherein the silencing element is inserted into intron 1 of DYRK1A or intron 3 of RCAN1.

10. The method of claim 9, wherein the cell is trisomic for chromosome 21.

11. The method of claim 9 or 10, wherein the cell is a human cell.

5 12. The method of claim 11, wherein the cell is a stem cell or a fibroblast.

13. The method of claim 12, wherein the stem cell is an induced pluripotent stem cell (iPSC), a hematopoietic stem cell, or a neural stem cell.

14. A cell produced by the method of claims 8-13.

15. A method of reducing the risk of transient myeloproliferative disorder (TMD) in a
10 subject who has Down Syndrome (Trisomy 21), the method comprising:
obtaining a hematopoietic stem cell from the subject;
contacting the cell with the silencing vector of any of claims 1-8, under conditions sufficient for the silencing vector to undergo homologous recombination with the genomic DNA of the cell, wherein the silencing element is inserted into DYRK1A or
15 RCAN1, to produce a modified cell having reduced levels of expression of genes on Chromosome 21; and
administering the modified cell to the subject.

16. The method of any of claims 9-15, further comprising contacting the cell with a cleavage vector comprising a sequence that enhances or facilitates homologous
20 recombination.

17. The method of claim 16, wherein the cleavage vector comprises a zinc finger nuclease (ZFN) or a transcription activator-like effector nuclease (TALEN).

18. The method of claim 17, wherein the cleavage vector targets a sequence within intron 1 of DYRK1A.

25 19. The method of claim 17, wherein the cleavage vector targets a sequence within intron 3 of RCAN1.

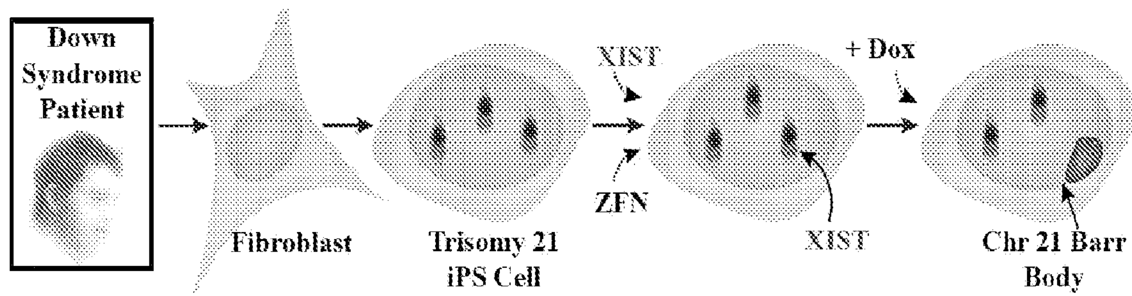


FIG. 1

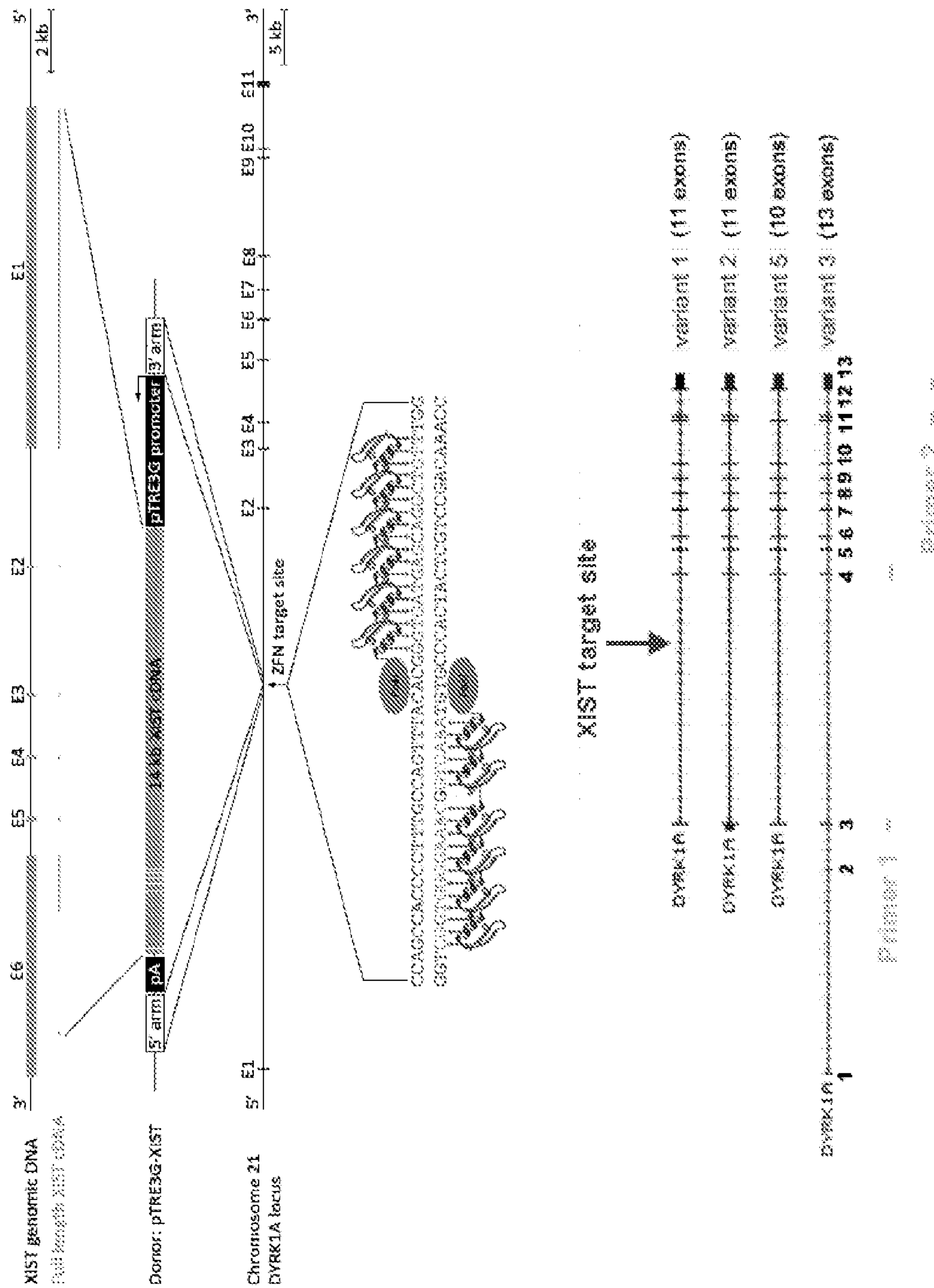


FIG. 2A

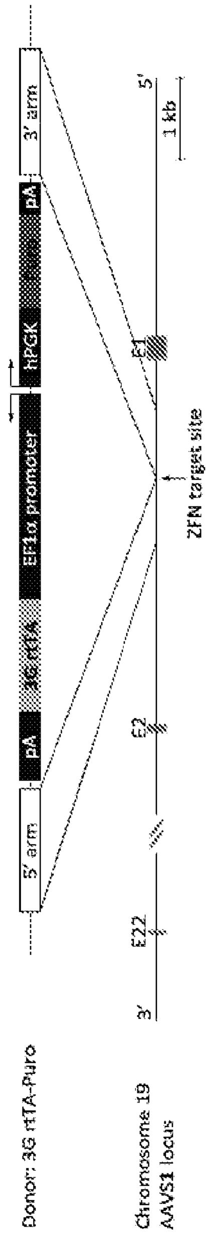


FIG. 2B

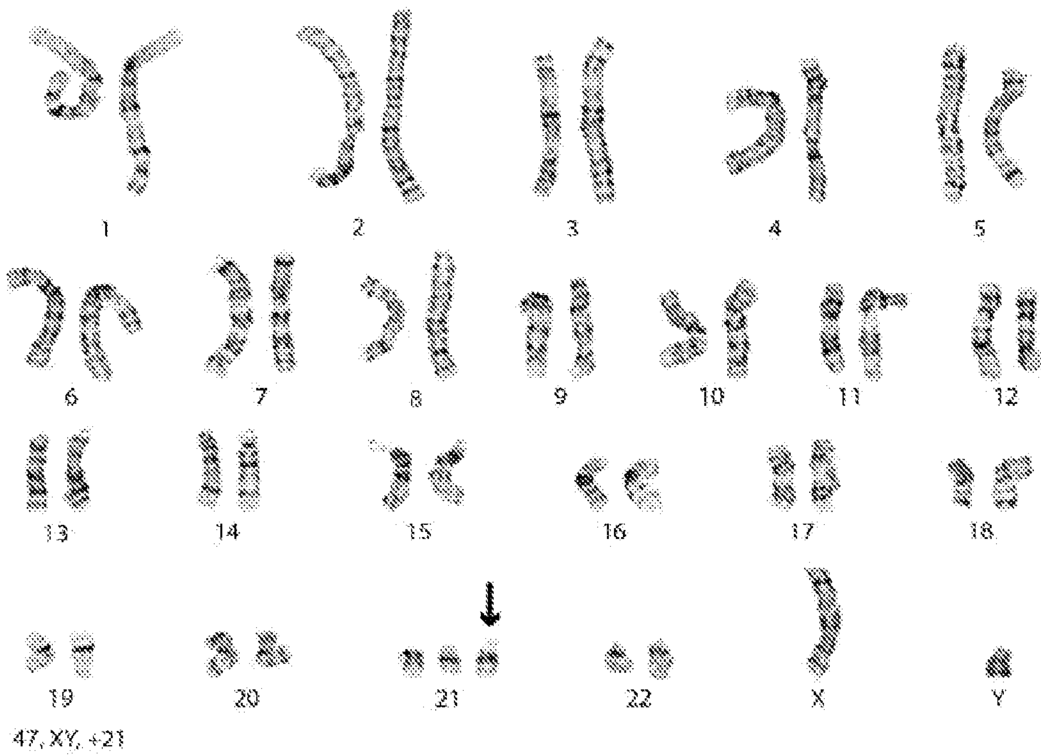


FIG. 2C

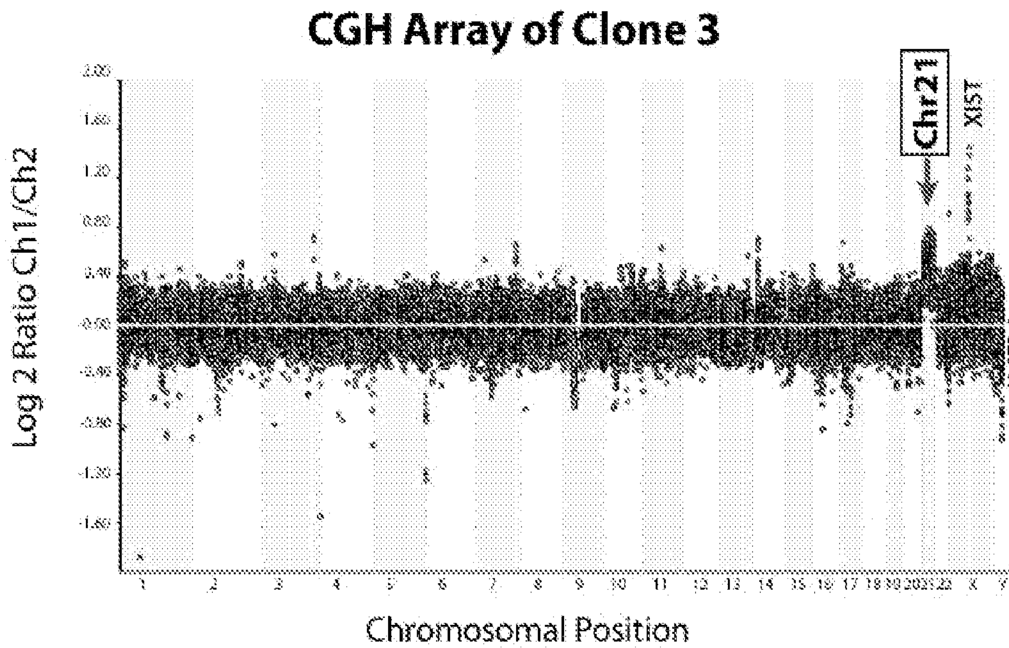


FIG. 2D

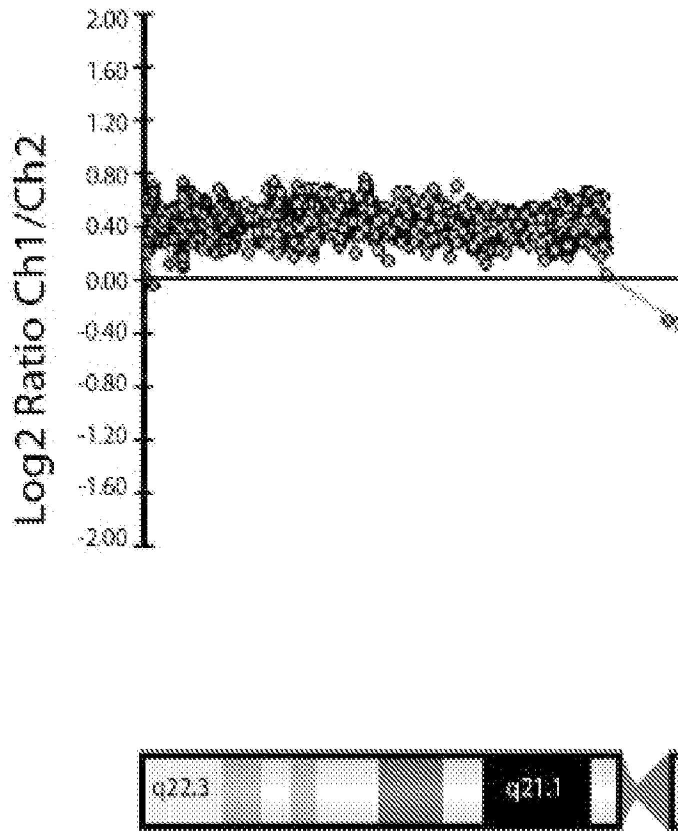


FIG. 2E

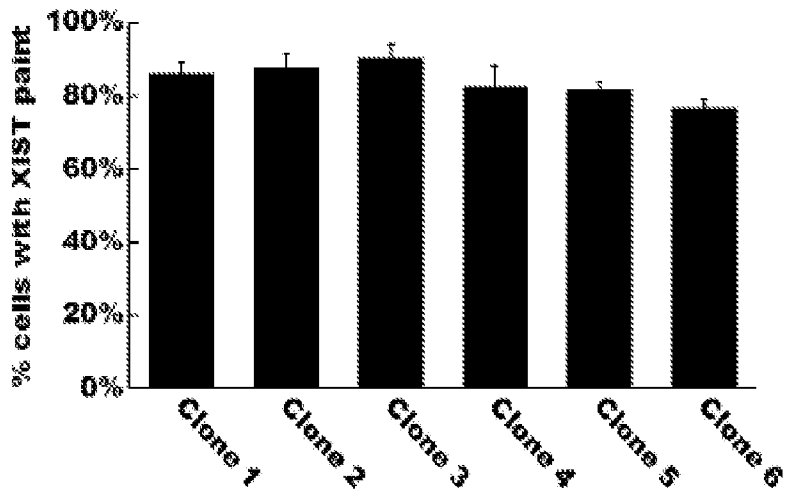


FIG. 2F

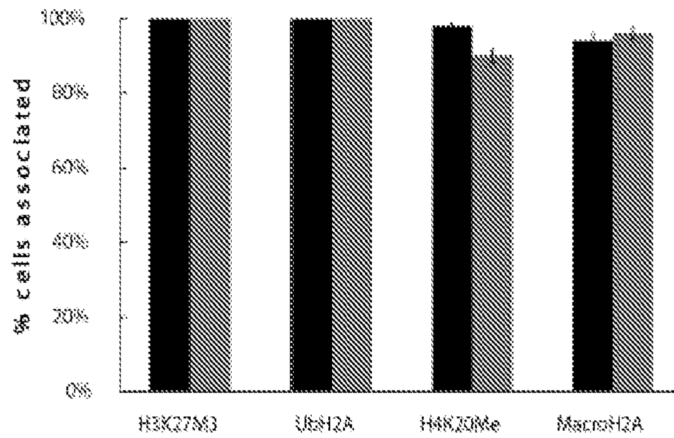


FIG. 3

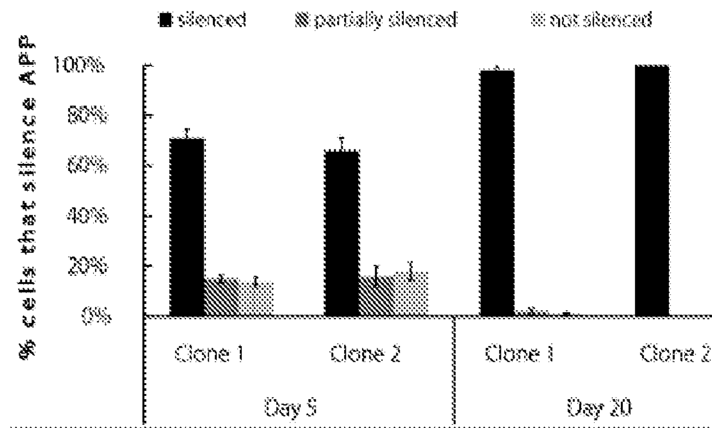


FIG. 4A

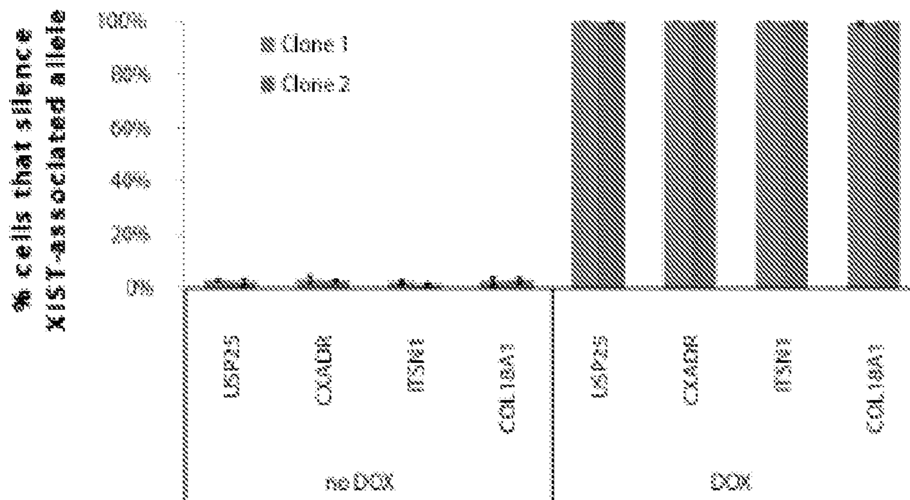


FIG. 4B

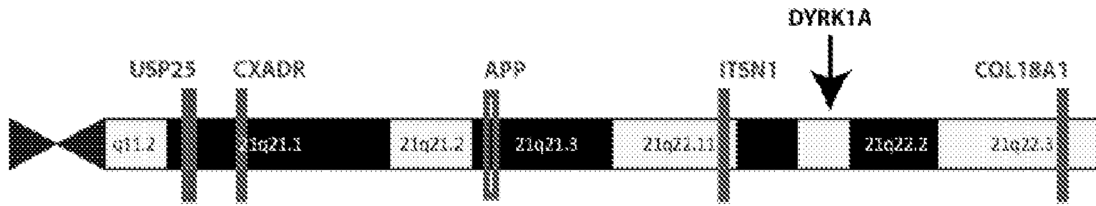


FIG. 4C

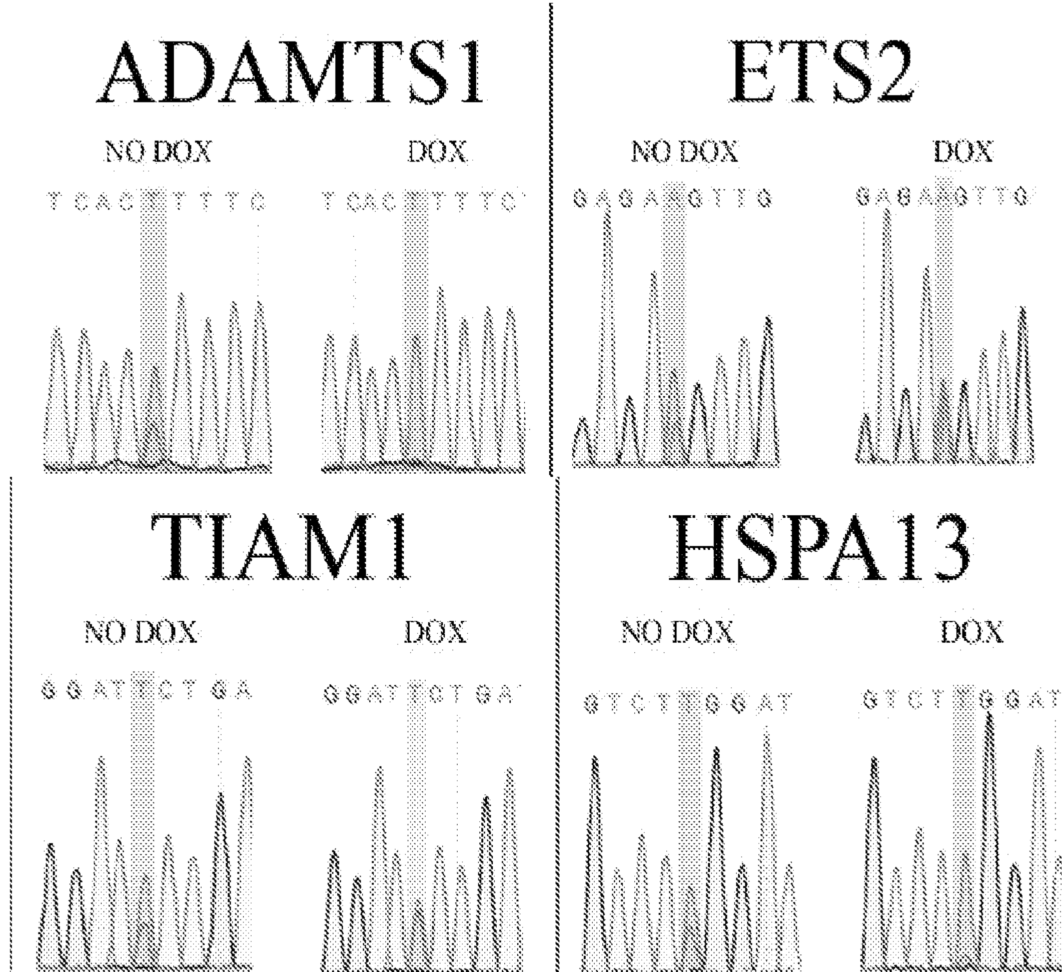


FIG. 4D

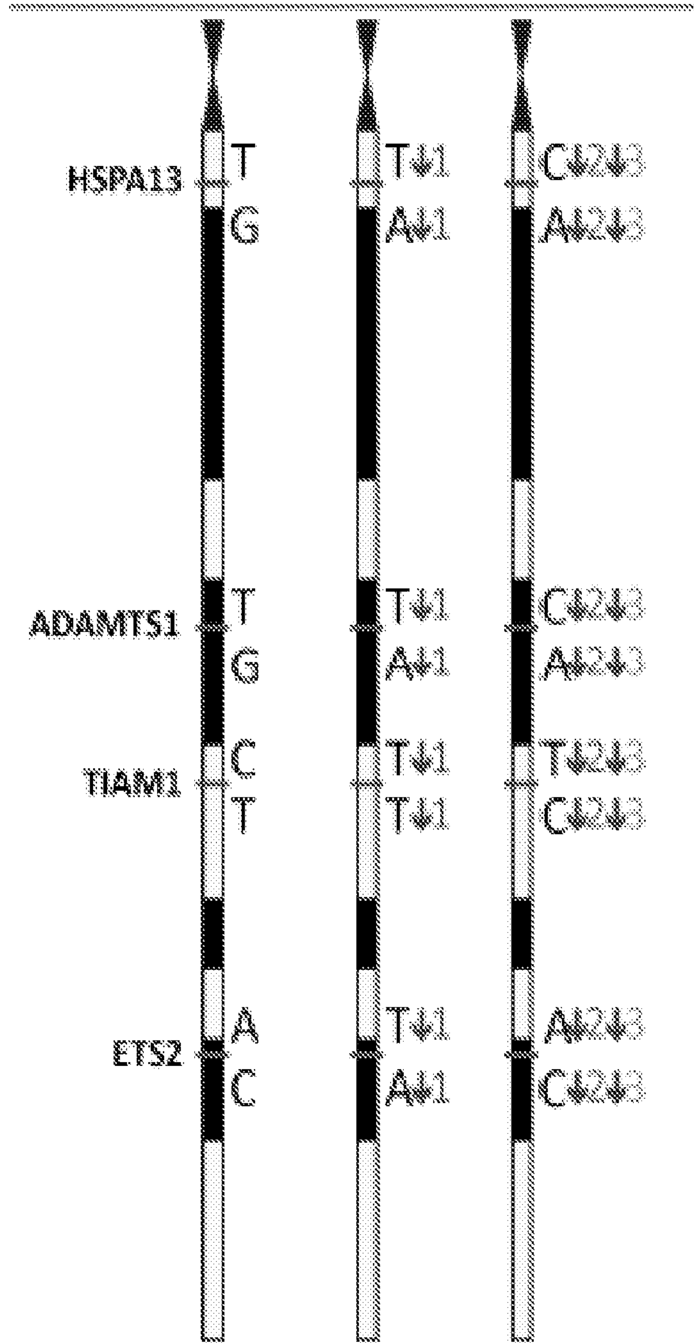


FIG. 4E

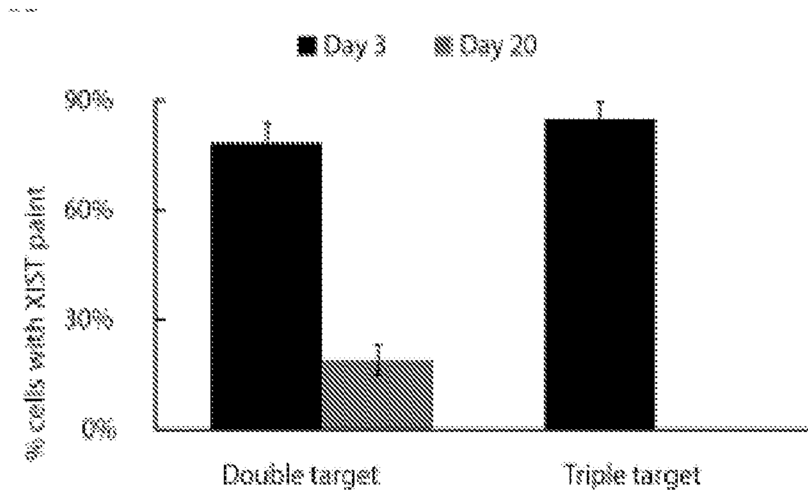


FIG. 4F

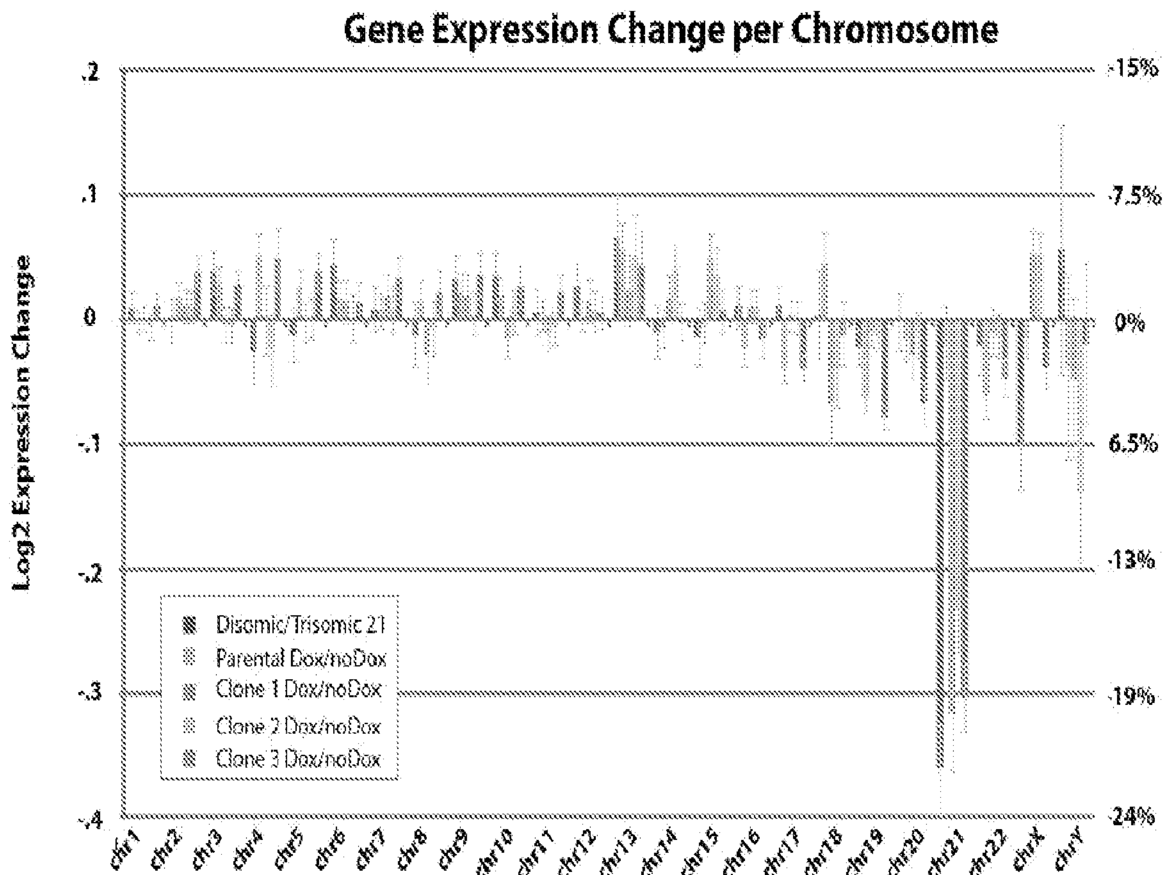


FIG. 5A

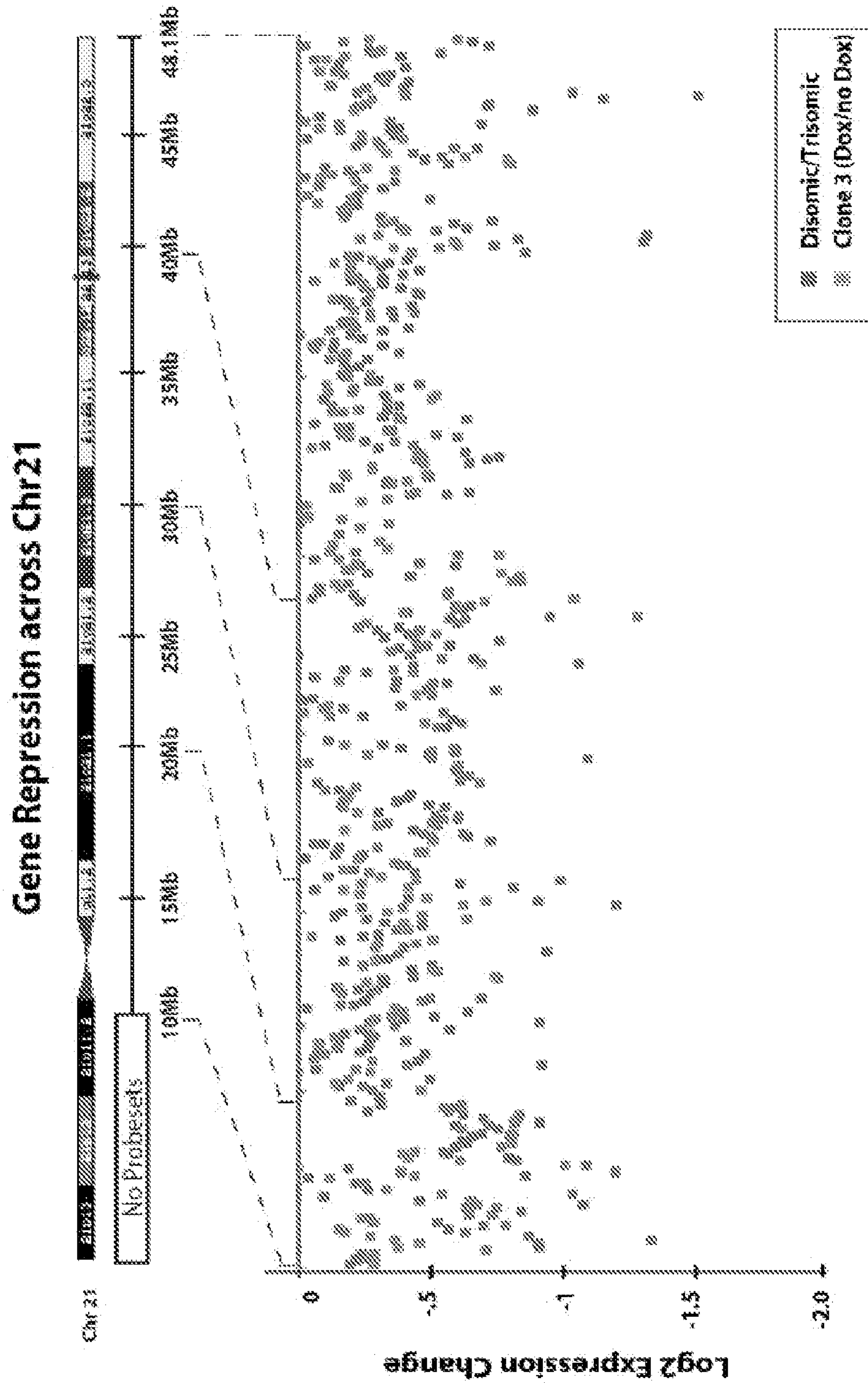


FIG. 5B

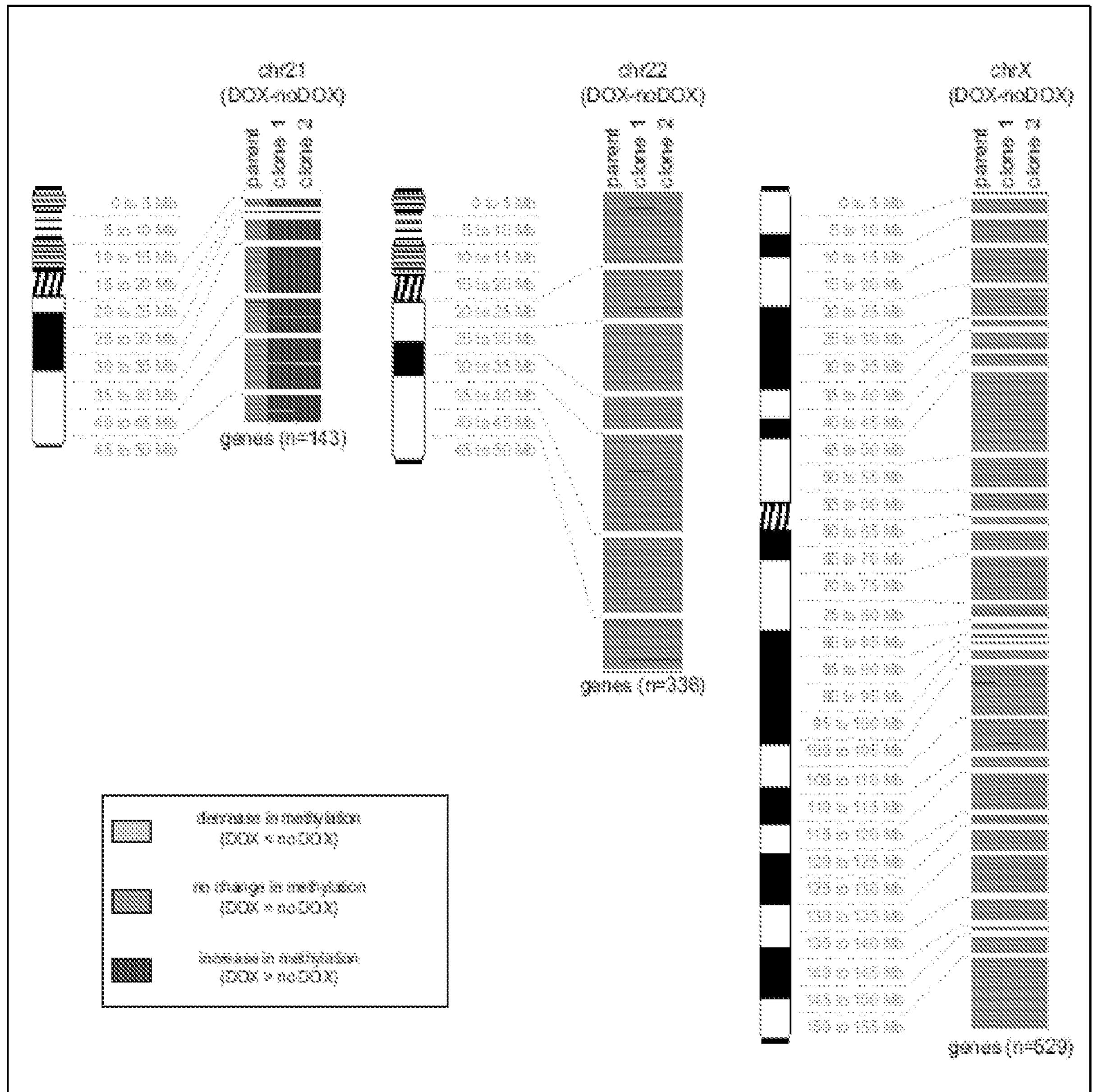


FIG. 5C

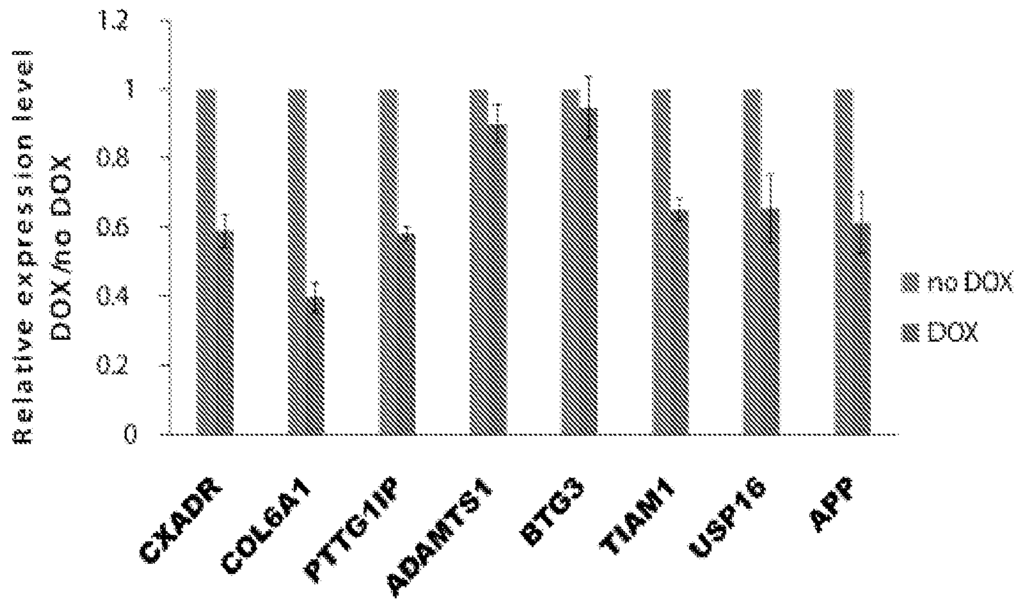


FIG. 5D

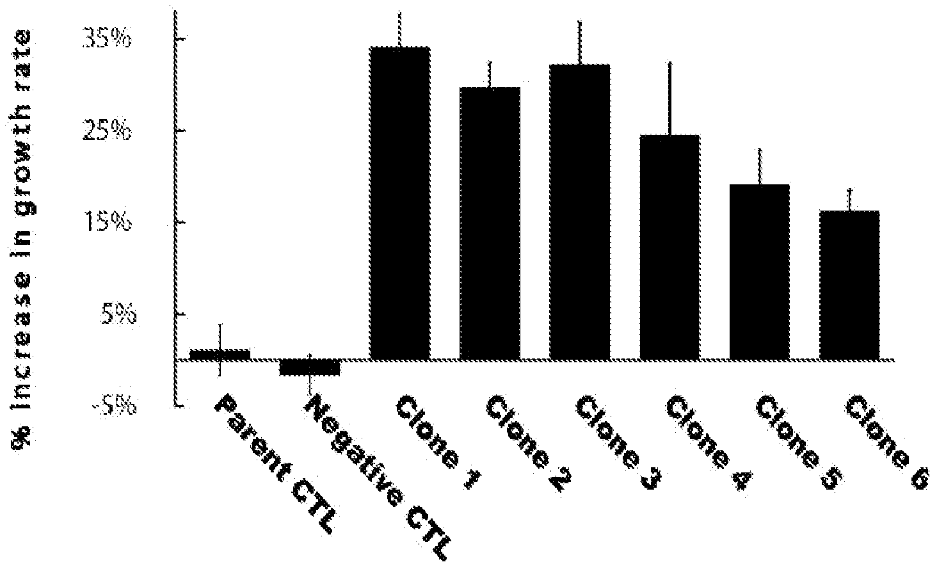


FIG. 6A

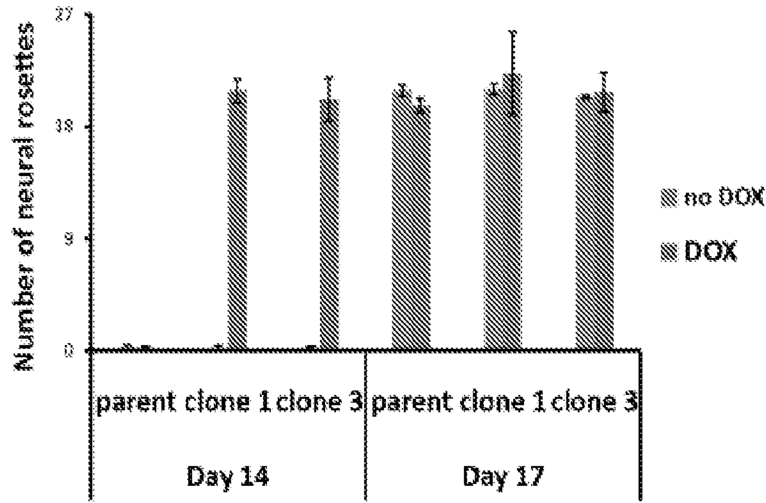


FIG. 6B

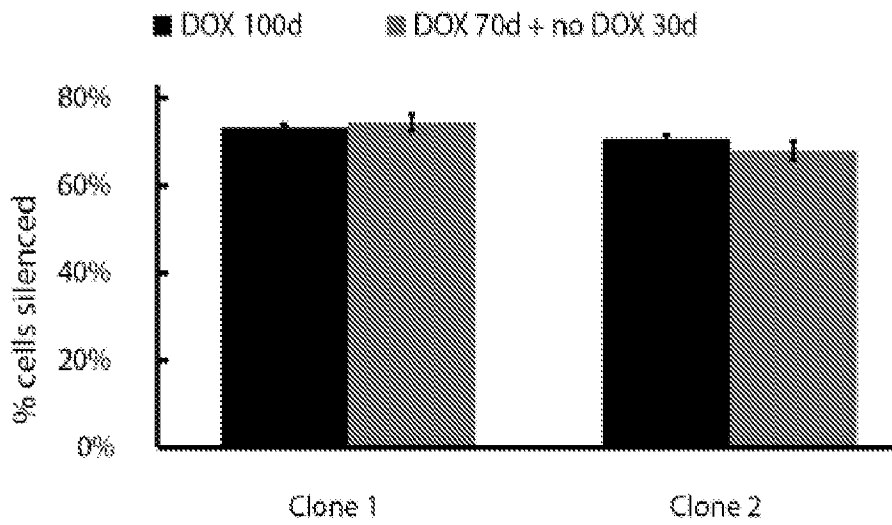


FIG. 6C

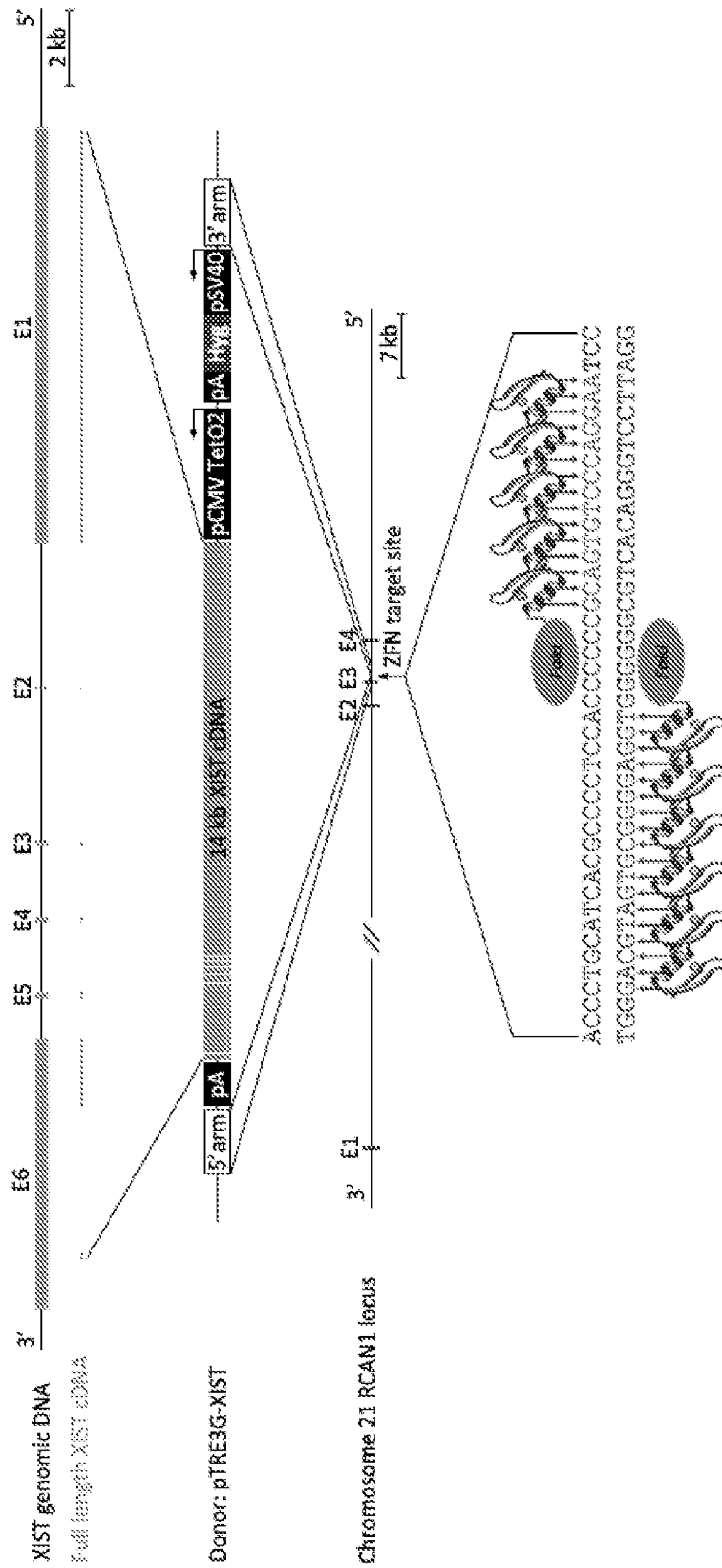


FIG. 7A

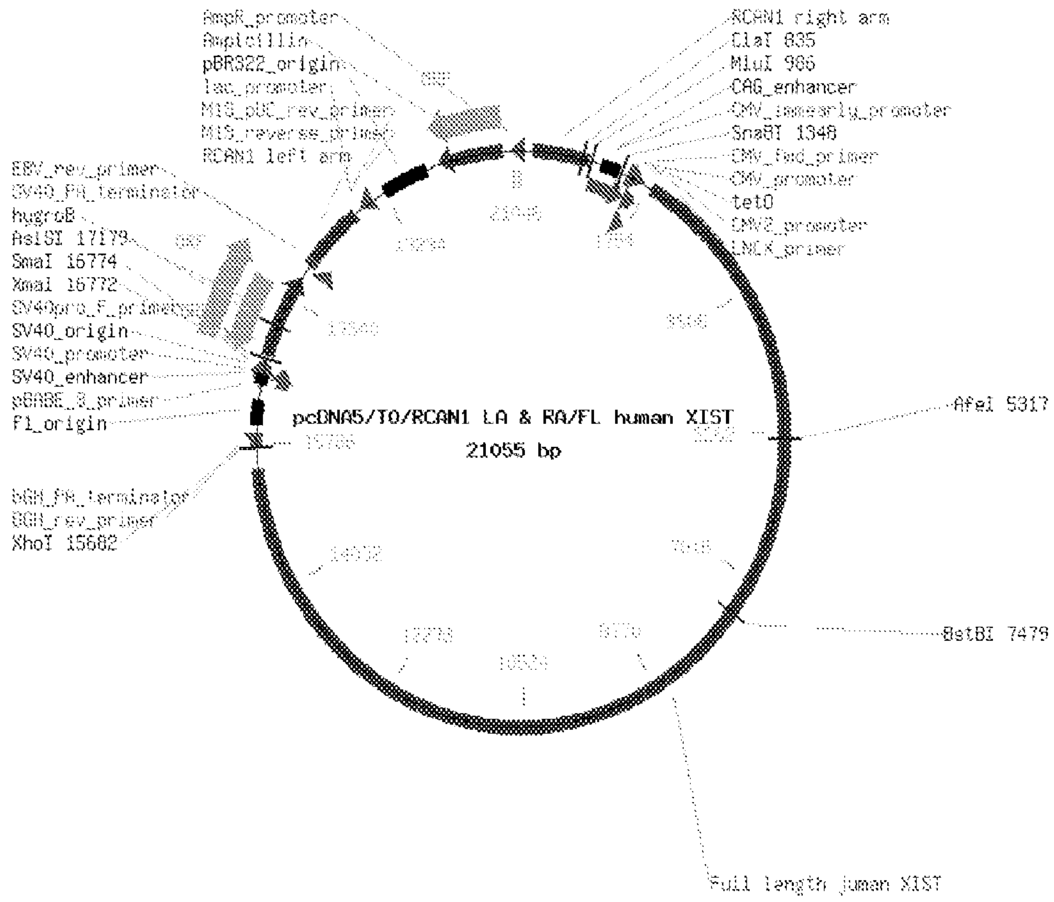


FIG. 7B

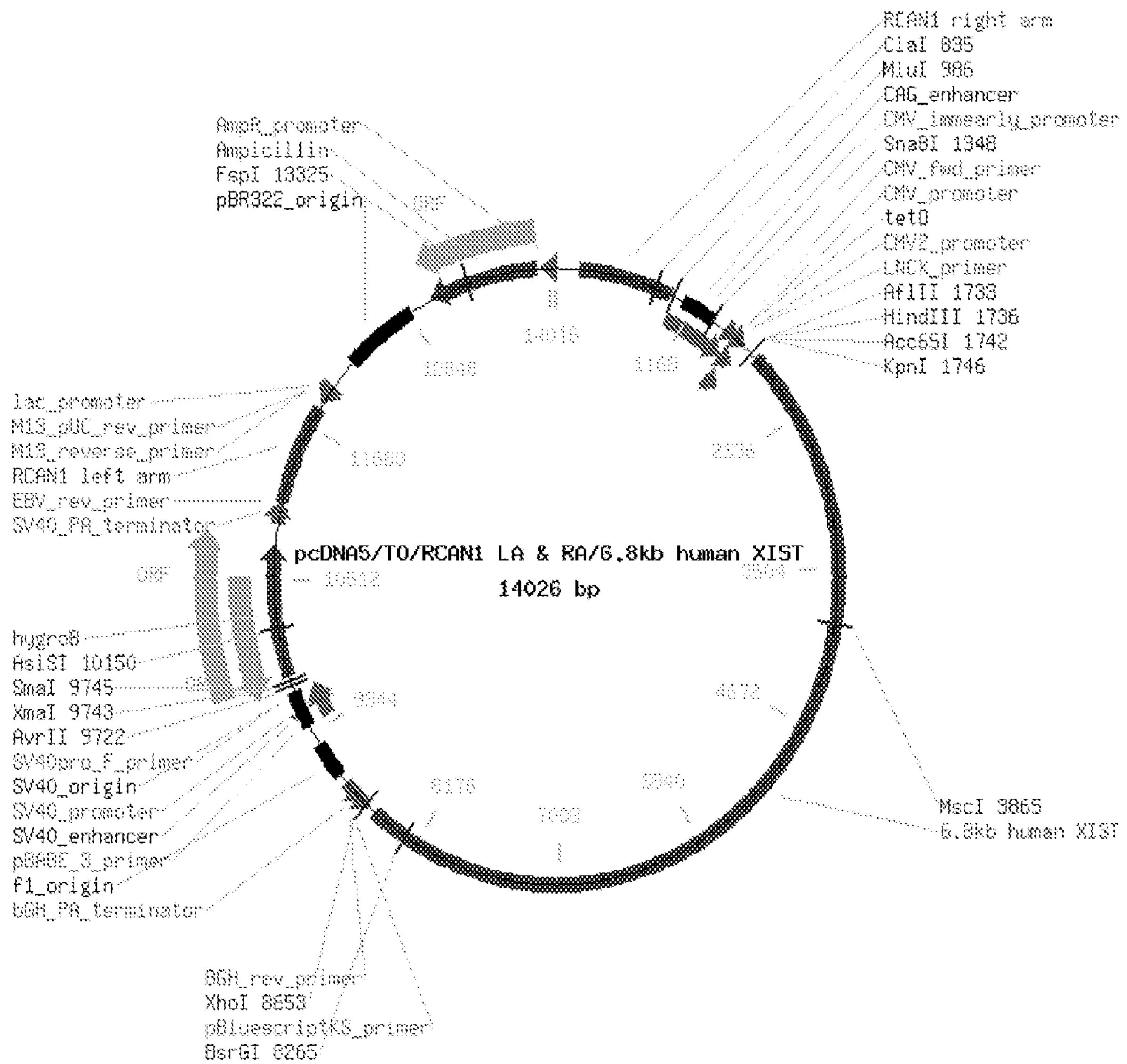


FIG. 7C

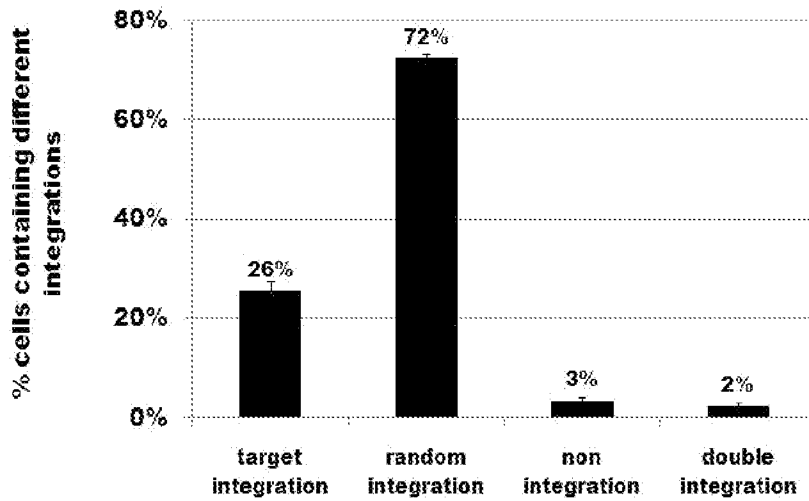


FIG. 8A

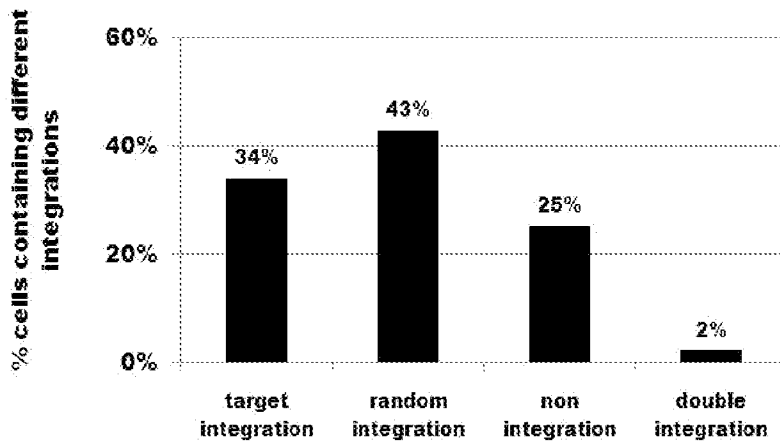


FIG. 8B

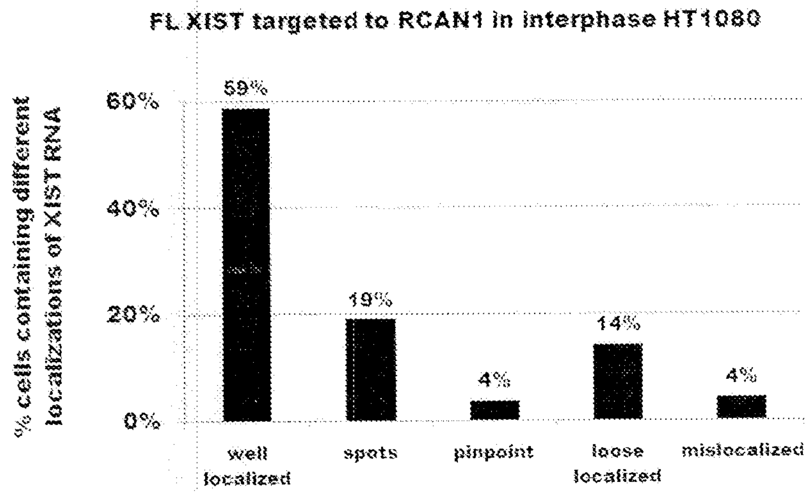


FIG. 8C

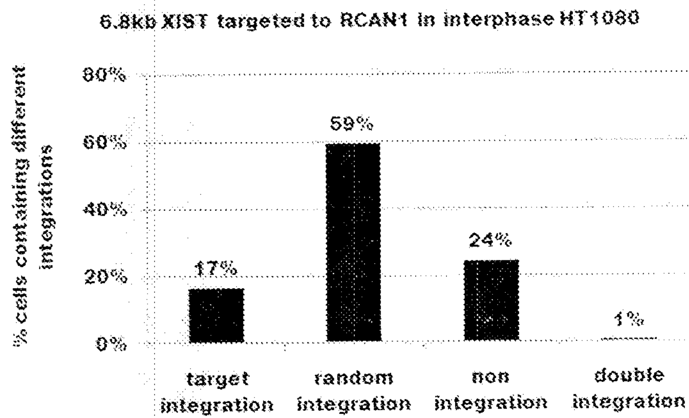


FIG. 8D

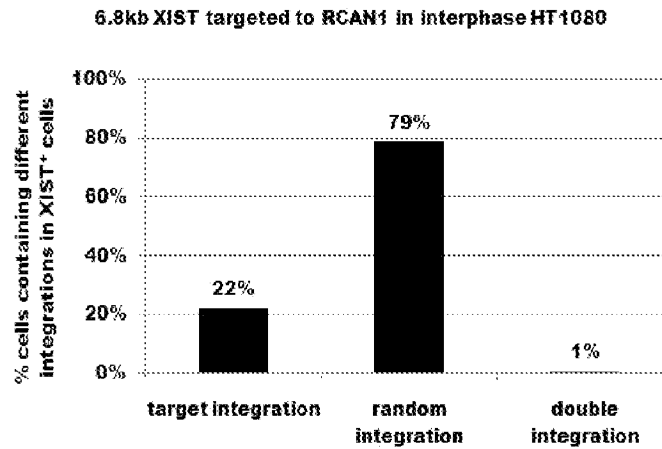


FIG. 8E

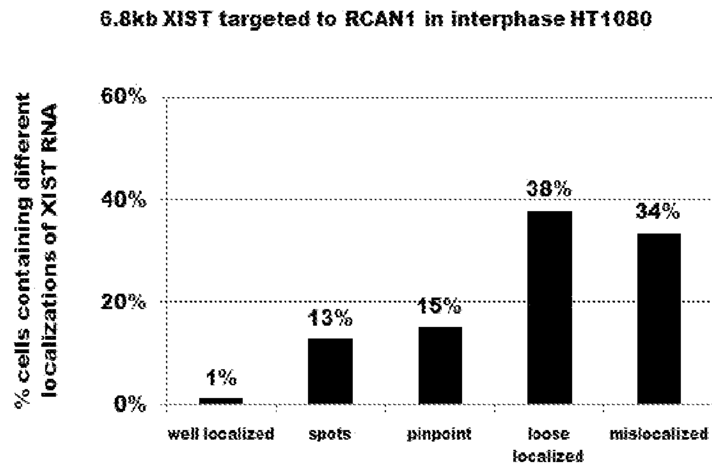


FIG. 8F

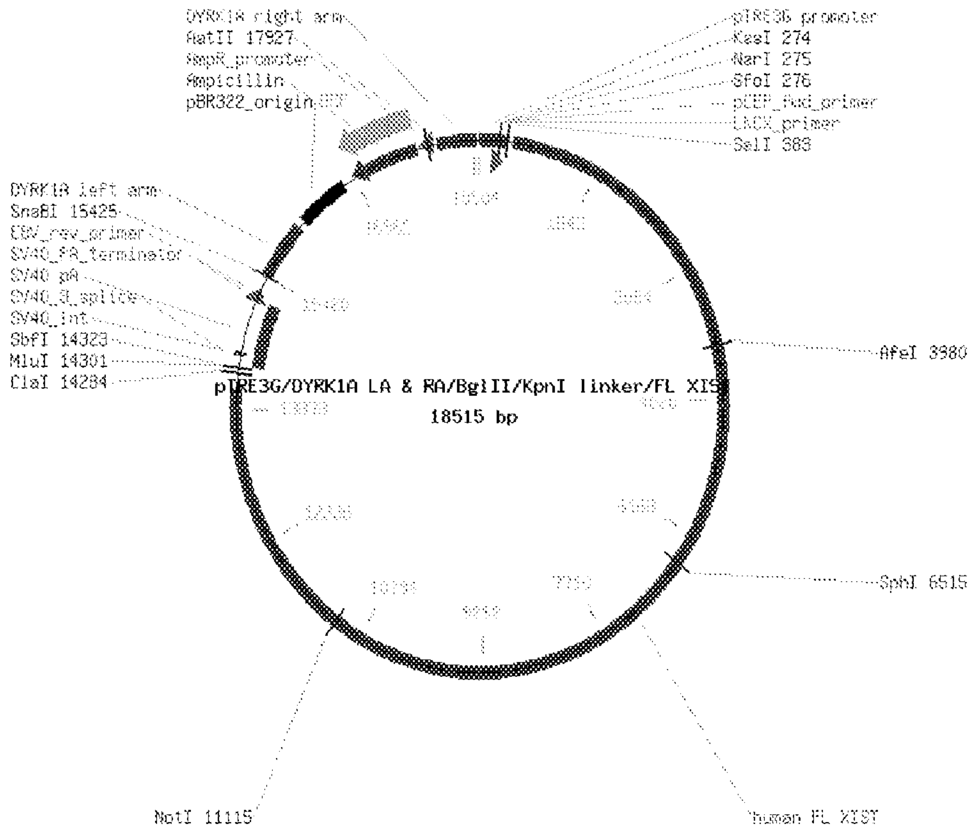


FIG. 9A

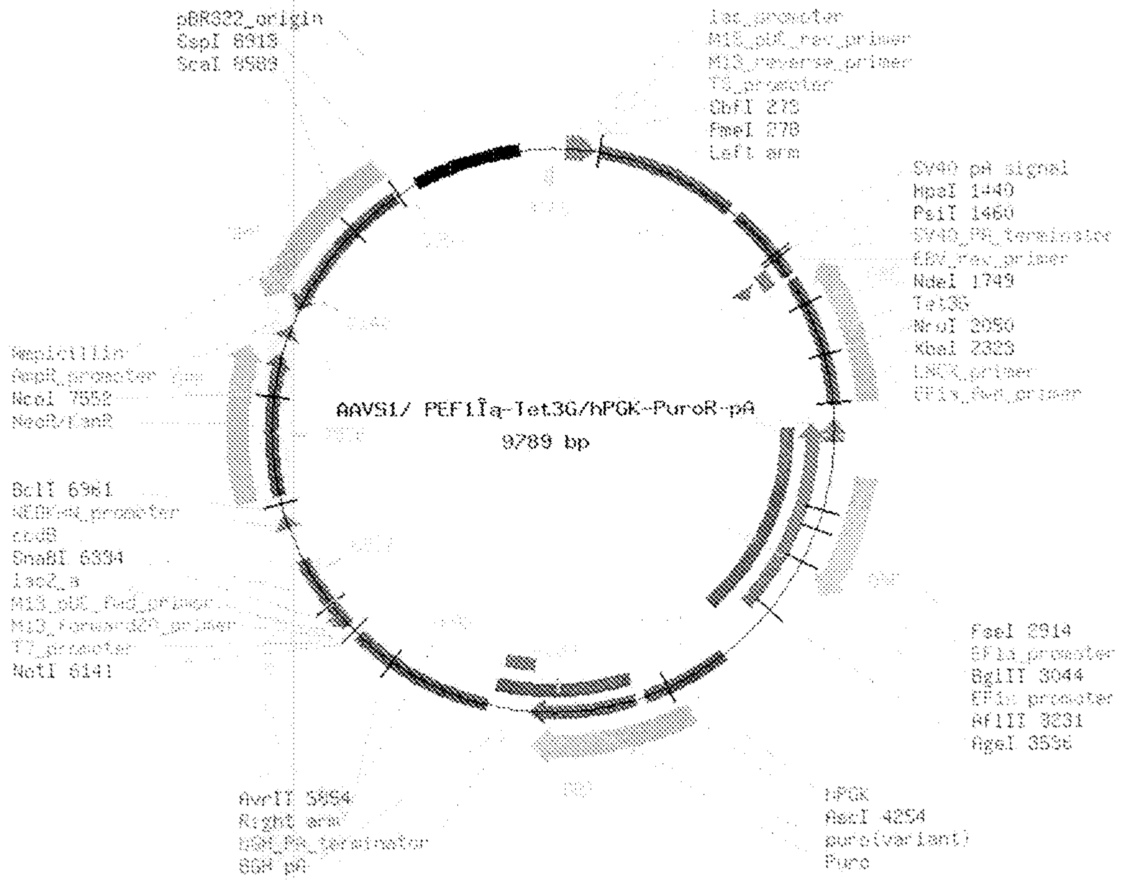


FIG. 9B

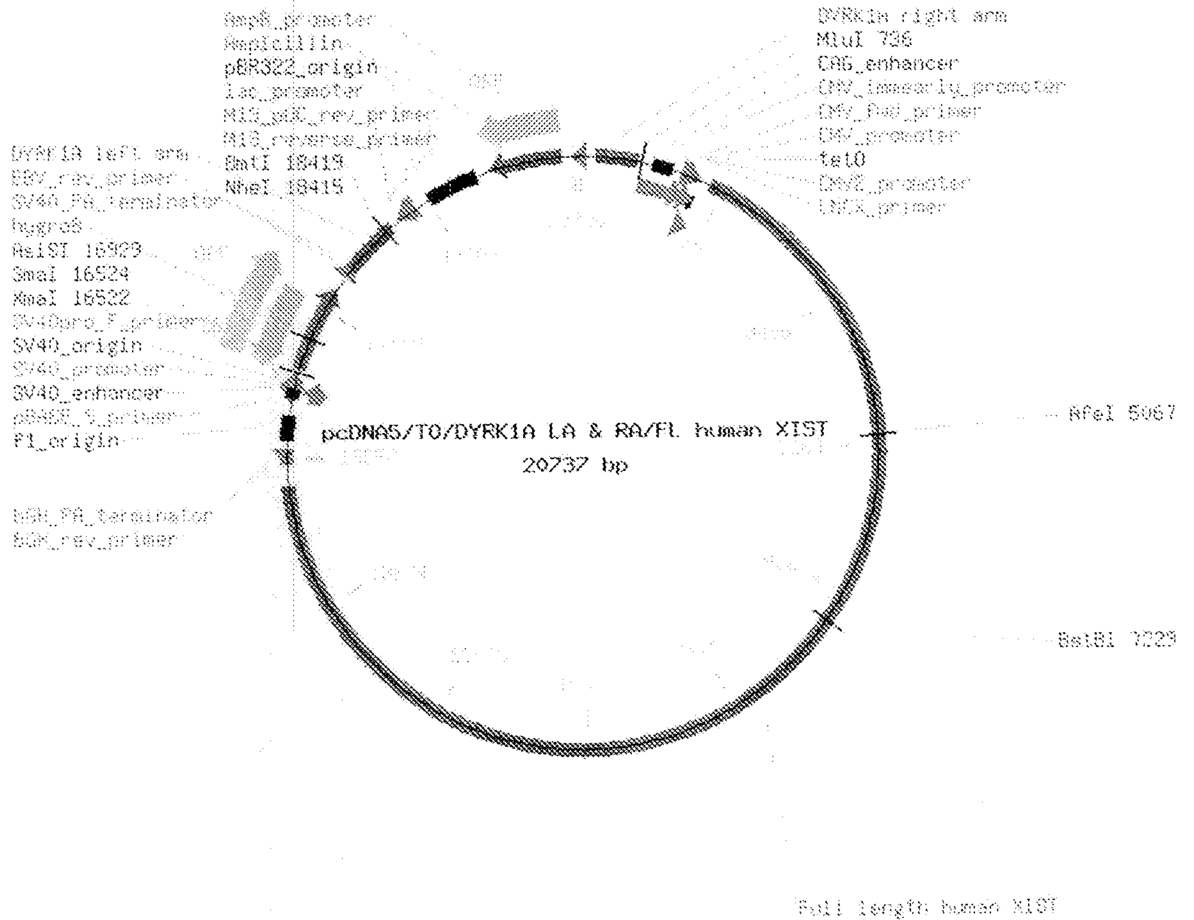


FIG. 9C

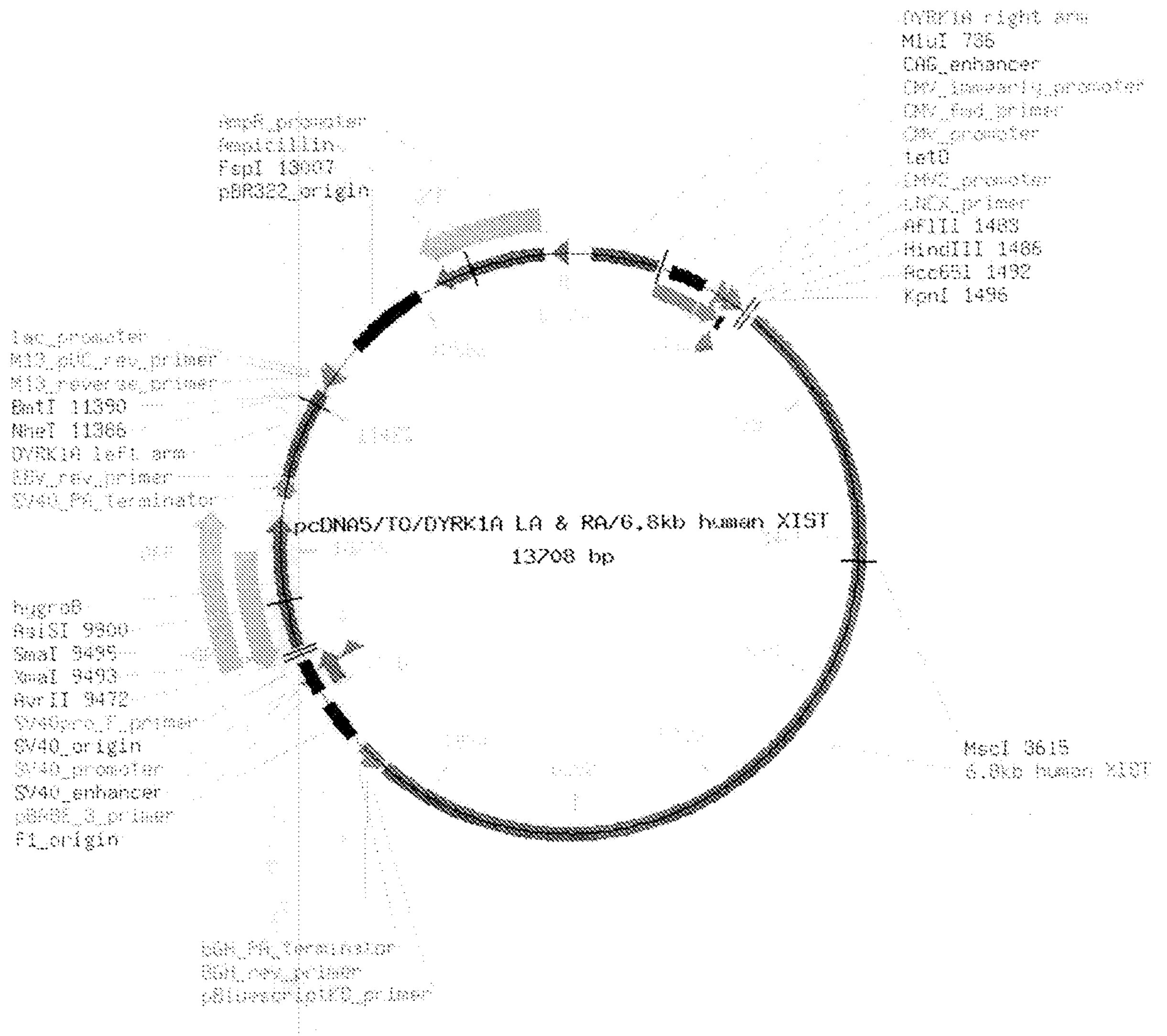


FIG. 9D

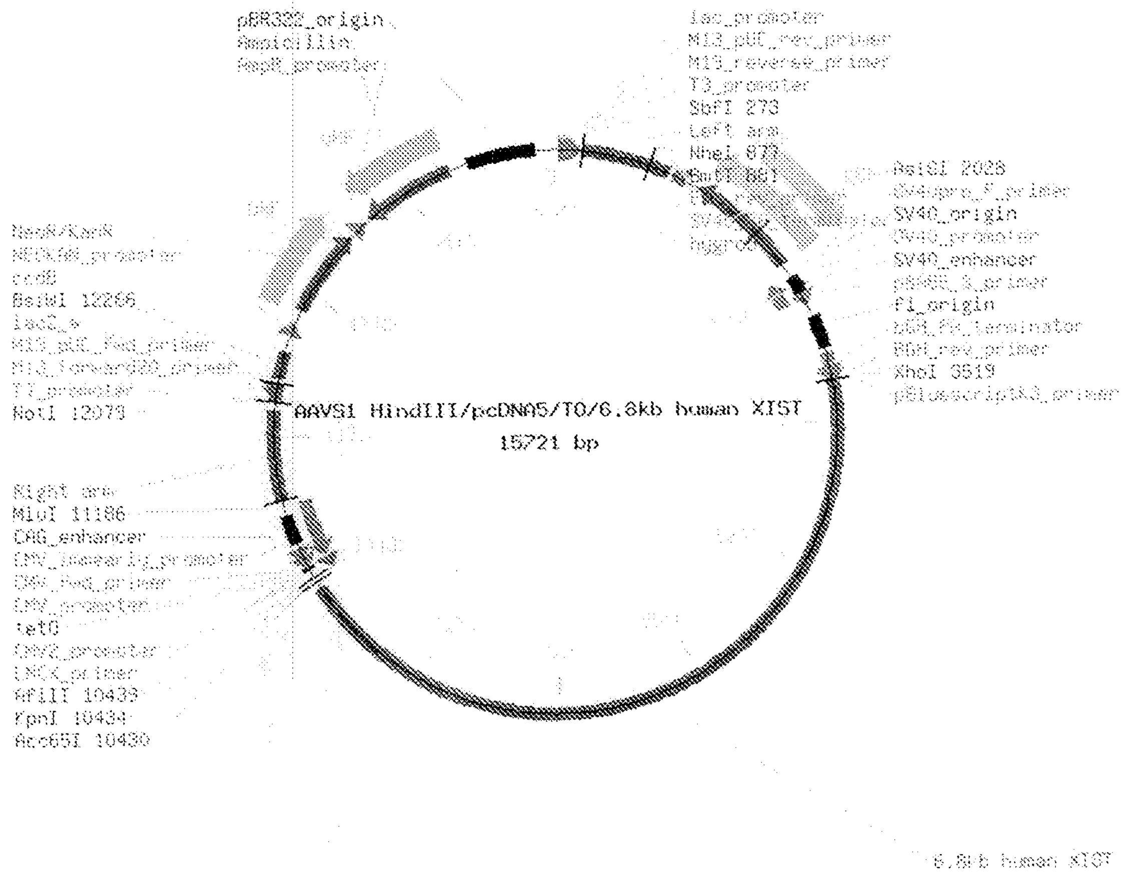


FIG. 9e

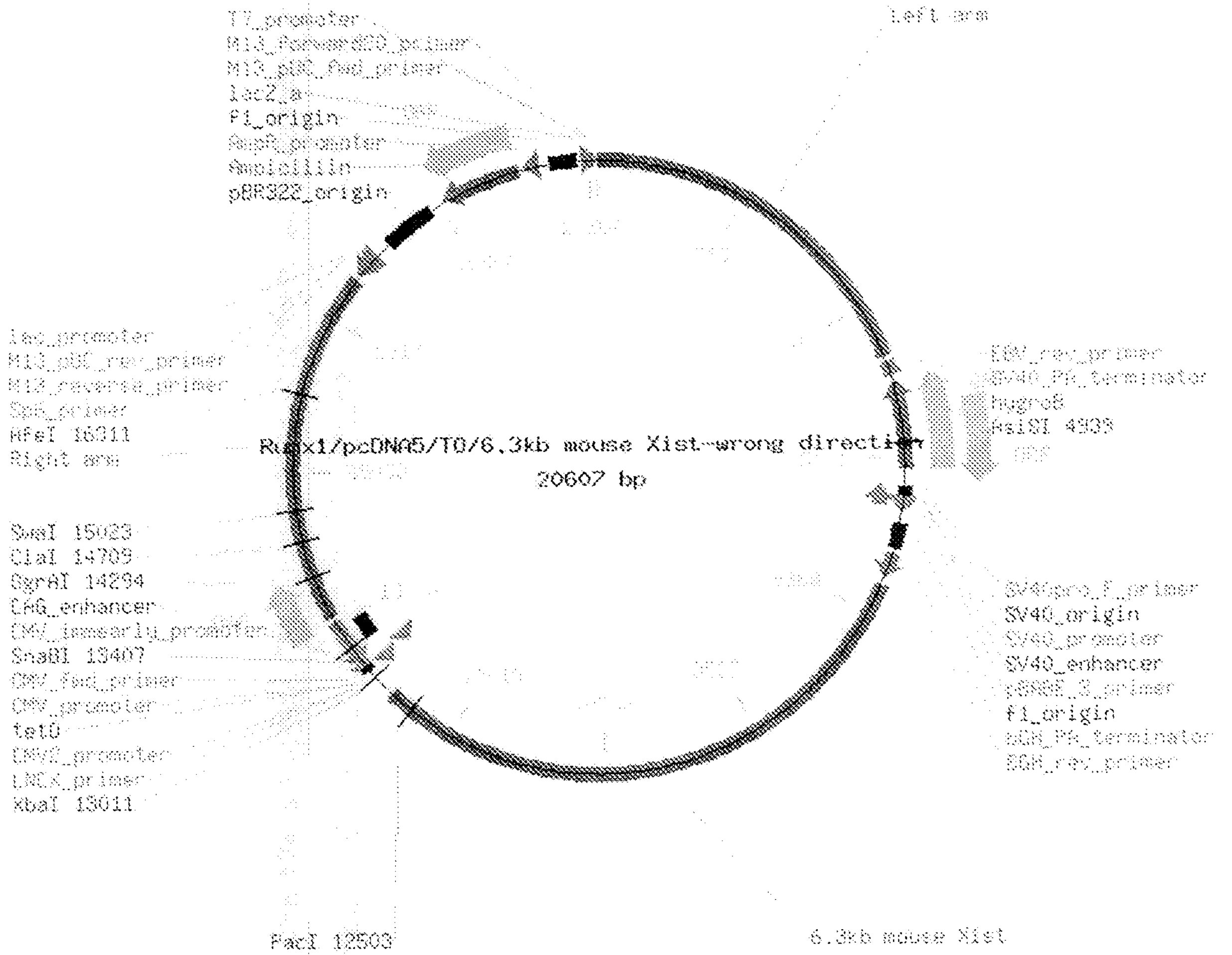


FIG. 9F

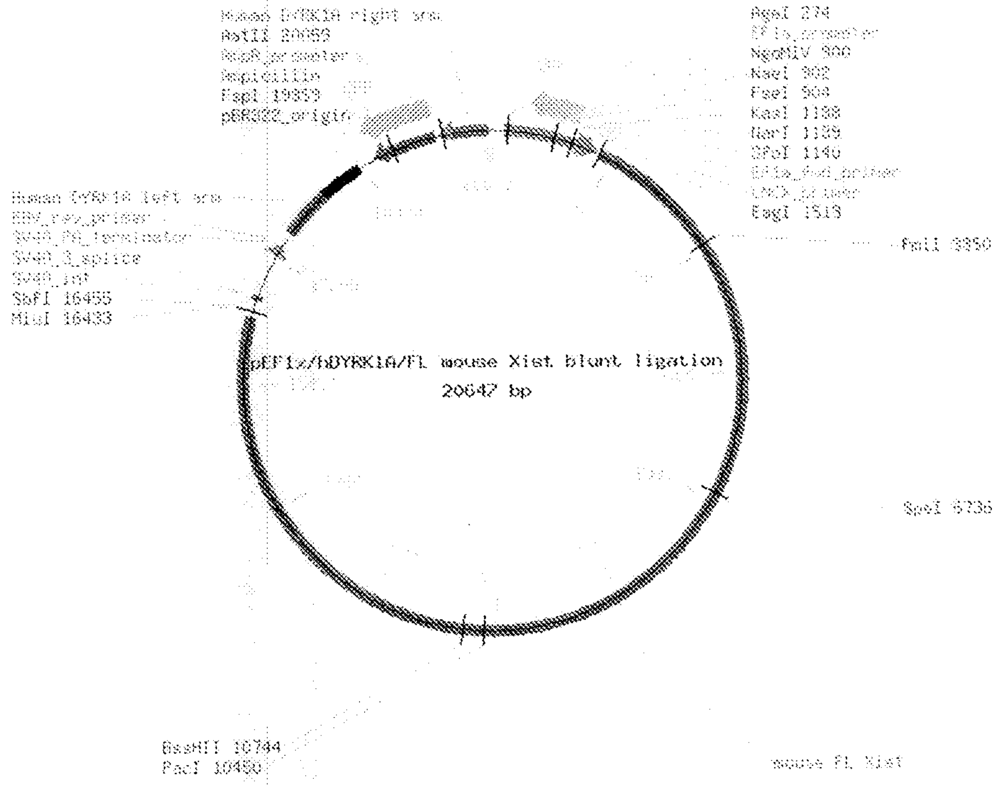


FIG. 9G

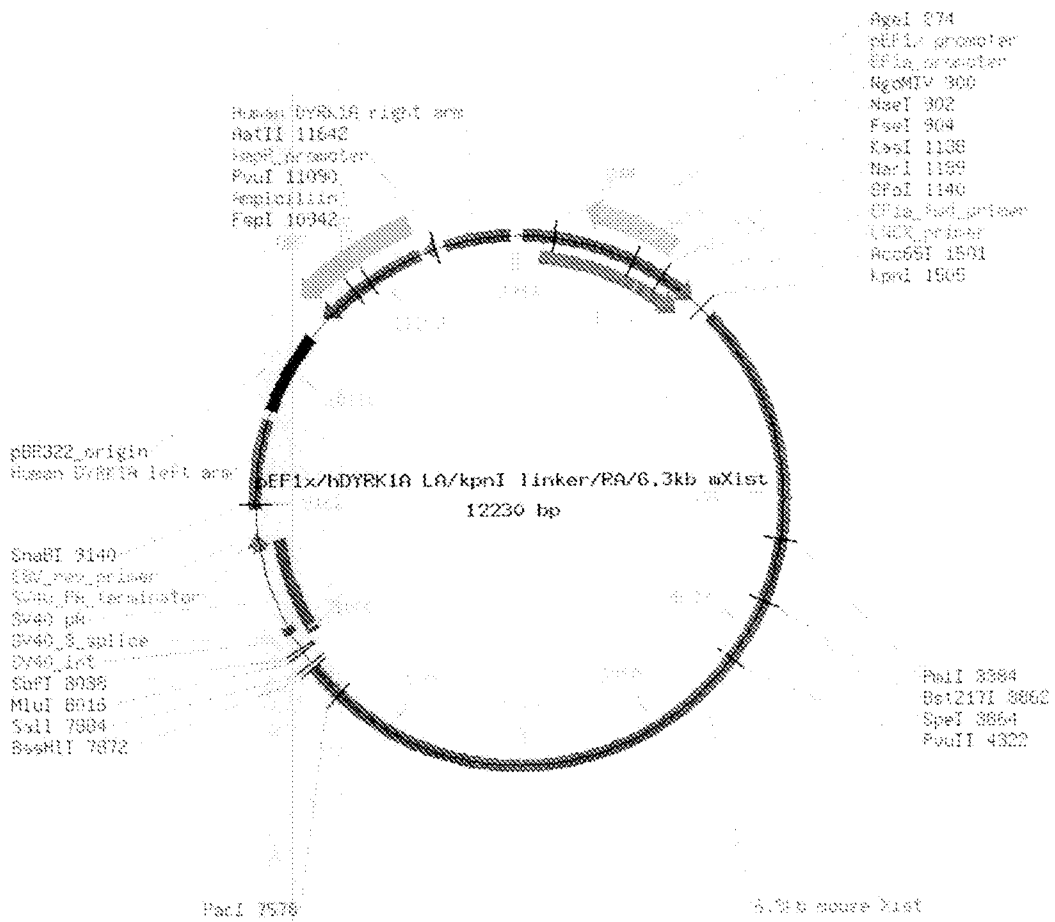


FIG. 9H

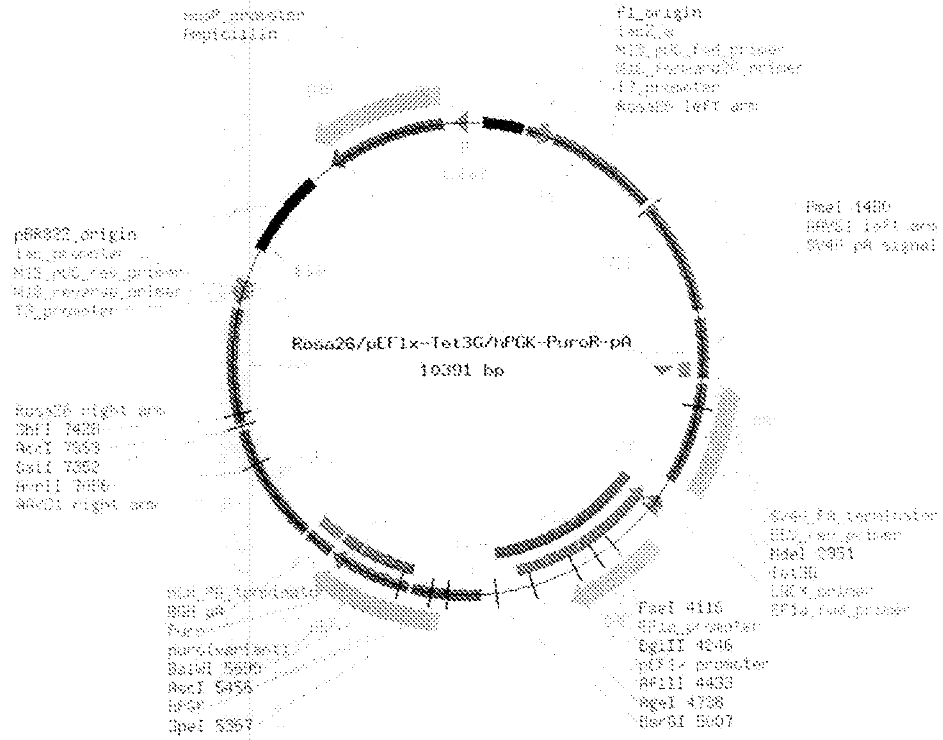


FIG. 9I

# FIG. 10A - CONSTRUCT 1 (3G/FL/hXIST/DYRK1A):

```
LOCUS       pTRE3G/DYRK1A LA &#38; RA/BglII/KpnI linker/FL XIST 18515 bp      DNA     SYN
       29-Nov-2012
DEFINITION  pTRE3G/DYRK1A LA &#38; RA/BglII/KpnI linker/FL XIST
ACCESSION
KEYWORDS
SOURCE
  ORGANISM  other sequences; artificial sequences; vectors.
FEATURES             Location/Qualifiers
     source          1..18515
                     /organism="pTRE3G/DYRK1A LA &#38; RA/BglII/KpnI linker/FL XIST"
                     /mol_type="other DNA"
     promoter        7..382
                     /label="pTRE3G promoter"
     misc_feature    289..308
                     /label="pCEP_fwd_primer"
     misc_feature    291..315
                     /label="LNCX_primer"
     misc_feature    454..14183
                     /label="human FL XIST"
     misc_feature    14460..14475
                     /label="SV40_int"
     misc_feature    14481..14528
                     /label="SV40_3_splice"
     misc_feature    14385..15182
                     /label="SV40 pA"
     terminator      15104..15223
                     /label="SV40_PA_terminator"
     misc_feature    15192..15211
                     /label="EBV_rev_primer"
     misc_feature    15423..16112
                     /label="DYRK1A left arm"
     rep_origin      complement(16158..16777)
                     /label="pBR322_origin"
     CDS             complement(16932..17792)
                     /label="Ampicillin"
                     /gene="Ampicillin"
                     /note="ORF frame 3"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW*"                                (SEQ ID NO:25)
     gene            complement(16932..17792)
                     /label="Ampicillin"
                     /gene="Ampicillin"
     promoter        complement(17834..17862)
                     /label="AmpR_promoter"
     misc_feature    18003..18510
                     /label="DYRK1A right arm"
```

# FIG. 10B - CONSTRUCT 2 (puro/rtTA/AAVS1)

```
LOCUS       AAVS1/ PEF1α-Tet3G/hPGK-PuroR-pA     9789 bpDNA    SYN    14-Mar-2013
DEFINITION  AAVS1/ PEF1α-Tet3G/hPGK-PuroR-pA
ACCESSION
KEYWORDS
SOURCE
  ORGANISM  other sequences; artificial sequences; vectors.
FEATURES             Location/Qualifiers
     source          1..9789
                     /organism="AAVS1/ PEF1α-Tet3G/hPGK-PuroR-pA"
                     /mol_type="other DNA"
     promoter        143..172
                     /label="lac_promoter"
     misc_feature    186..208
                     /label="M13_pUC_rev_primer"
     misc_feature    207..225
                     /label="M13_reverse_primer"
     promoter        242..261
                     /label="T3_promoter"
     misc_feature    295..1095
                     /label="Left arm"
     misc_feature    1120..1562
                     /label="SV40 pA signal"
     terminator      1444..1575
                     /label="SV40_PA_terminator"
     misc_feature    1532..1551
                     /label="EBV_rev_primer"
     CDS             complement(1585..2331)
                     /label="ORF frame 1"
                     /translation="MSRLDKSKVINSALELLNGVGIEGLTTRKLAQKLGVEQPTLYWH
                     VKNKRALLDALPIEMLDRHHTHSCPLEGESWQDFLRNNAKSYRCALLSHRDGAKVHLG
                     TRPTEKQYETLENQLAFLCQQGFSLENALYALSAVGHFTLGCVLEEQEHQVAKEERET
                     PTTDSMPPLLKQAIELFDRQGAEPAFLFGLELIICGLEKQLKCESGGPTDALDDFDLD
                     MLPADALDDFDLDMLPADALDDFDLDMLPG*"      (SEQ ID NO:26)
     misc_feature    1585..2331
                     /label="Tet3G"
     misc_feature    complement(2413..2437)
                     /label="LNCX_primer"
     misc_feature    complement(2469..2489)
                     /label="EF1a_fwd_primer"
     CDS             2697..3344
                     /label="ORF frame 3"
                     /translation="MKRRLRTERPFSFVWVTHPPALPSAASSILSSLQQGREAAIFPL
                     TQLVPTGPALPPRAGRYTAARGQAPEQAGQLETTPVRFSVAALAGPASPNMCAGTHGP
                     RRRPRPQKPKYQCADLGPHLQDYLARKKASQQVIKNFKWLETYRKQRDRREGATRFAR
                     GGPSAQARPQLKHEAKGLLKRKASNSPTHFQPEARDQESRTAARGVEVIQGTQGP*"
                               (SEQ ID NO:27)
     promoter        complement(2437..3624)
                     /label="EF1a_promoter"
     promoter        2437..3771
                     /label="EF1x promoter"
     misc_feature    3913..4419
                     /label="hPGK"
     CDS             4201..5040
                     /label="puro(variant)"
                     /gene="puro(variant)"
                     /note="ORF frame 1"
                     /translation="MAARRPRWAVANSGCSAGRAESSGREGAVREAGCGAVVWALFLP
                     ARCSAFCKPPERTSAVGSLVDRITDLSPQGDPPELTMIEYKPTVRLATRDDVPRAVRT
                     LAAAFADYPATRHTVDPDRHIERVTELQELFLTRVGLDIGKVWVADDGAAVAVWTTPE
                     SVEAGAVFAEIGPRMAELSGSRLAAQQQMEGLLAPHRPKEPAWFLATVGVSPDHQGKG
                     LGSAVVLPGVEAAERAGVPAFLETSAPRNLPFYERLGFTVTADVEVPEGPRTWCMTRK
                     PGA*"           (SEQ ID NO:28)
     gene            4441..5040
                     /label="puro(variant)"
                     /gene="puro(variant)"
     misc_feature    4441..5040
                     /label="Puro"
     misc_feature    5047..5271
```

# FIG. 10B - CONSTRUCT 2 (puro/rtTA/AAVS1)

```
                     /label="BGH pA"
     terminator      5056..5271
                     /label="bGH_PA_terminator"
     misc_feature    5287..6126
                     /label="Right arm"
     promoter        complement(6162..6180)
                     /label="T7_promoter"
     misc_feature    complement(6187..6203)
                     /label="M13_forward20_primer"
     misc_feature    complement(6196..6218)
                     /label="M13_pUC_fwd_primer"
     misc_feature    6184..6327
                     /label="lacZ_a"
     misc_feature    6350..6643
                     /label="ccdB"
     promoter        6854..6903
                     /label="NEOKAN_promoter"
     CDS             6992..7786
                     /label="NeoR/KanR"
                     /gene="NeoR/KanR"
                     /note="ORF frame 2"
                     /translation="MIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAAVFRLSAQGR
                     PVLFVKTDLSGALNELQDEAARLSWLATTGVPCAAVLDVVTEAGRDWLLLGEVPGQDL
                     LSSHLAPAEKVSIMADAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDE
                     EHQGLAPAELFARLKASMPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADRY
                     QDIALATRDIAEELGGEWADRFLVLYGIAAPDSQRIAFYRLLDEFF*"  (SEQ ID NO:29)
     gene            6995..7783
                     /label="NeoR/KanR"
                     /gene="NeoR/KanR"
     promoter        7919..7942
                     /label="AmpR_promoter"
     CDS             complement(8036..8896)
                     /label="Ampicillin"
                     /gene="Ampicillin"
                     /note="ORF frame 2"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW*"   (SEQ ID NO:25)
     gene            complement(8036..8896)
                     /label="Ampicillin"
                     /gene="Ampicillin"
     rep_origin      9004..9623
                     /label="pBR322_origin"
```

# FIG. 10C - CONSTRUCT 3 (FL/hXIST/RCAN1)

```
LOCUS       pcDNA5/TO/RCAN1 LA &#38; RA/FL human XIST    21055 bp      DNA     SYN     29-
Nov-2012
DEFINITION  pcDNA5/TO/RCAN1 LA &#38; RA/FL human XIST
ACCESSION
KEYWORDS
SOURCE
  ORGANISM  other sequences; artificial sequences; vectors.
FEATURES             Location/Qualifiers
     source          1..21055
                     /organism="pcDNA5/TO/RCAN1 LA &#38; RA/FL human XIST"
                     /mol_type="other DNA"
     misc_feature    209..966
                     /label="RCAN1 right arm"
     misc_feature    1073..1360
                     /label="CAG_enhancer"
     promoter        994..1570
                     /label="CMV_immearly_promoter"
     misc_feature    1527..1547
                     /label="CMV_fwd_primer"
     promoter        1528..1651
                     /label="CMV_promoter"
     misc_feature    1578..1617
                     /label="tetO"
     promoter        1540..1713
                     /label="CMV2_promoter"
     misc_feature    1627..1651
                     /label="LNCX_primer"
     misc_feature    1790..15520
                     /label="Full length juman XIST"
     misc_feature    complement(15719..15736)
                     /label="BGH_rev_primer"
     terminator      15722..15949
                     /label="bGH_PA_terminator"
     rep_origin      16012..16318
                     /label="f1_origin"
     misc_feature    complement(16432..16452)
                     /label="pBABE_3_primer"
     misc_feature    complement(16438..16653)
                     /label="SV40_enhancer"
     promoter        16450..16718
                     /label="SV40_promoter"
     rep_origin      16617..16694
                     /label="SV40_origin"
     misc_feature    16679..16698
                     /label="SV40pro_F_primer"
     CDS             complement(16752..17573)
                     /label="ORF frame 3"
                     /translation="MPPLEVARLLLHTSQPRPPEEDVGDLVLGIPEHRLAPVNDRCYA
                     AIVRQDIVGAEIRVHEVPDFGAVLGPKHQLIESLRDGRTDGVVHHSLPVIHMGISNRA
                     YEIITPCSVLTDSLRSEWAEPARLAKIGRSDRIHGLRDRLQNSGQFGFRQVLQRDTLCT
                     AGDAIGQALAEFPNVKHFRNRERGRCKVPINITIFVETIGAAIYPQDISTPSYIEAES
                     TRFFALRELHQVGDAVELFDQKLLDRRRGEFRLFHHVLIRSENGYTSSRELFAKA*"
                                  (SEQ ID NC:30)
     CDS             16815..17840
                     /label="hygroB"
                     /gene="hygroB"
                     /note="ORF frame 3"
                     /translation="MKKPELTATSVEKFLIEKFDSVSDLMQLSEGEESRAFSFDVGGR
                     GYVLRVNSCADGFYKDRYVYRHFASAALPIPEVLDIGEFSESLTYCISRRAQGVTLQD
                     LPETELPAVLQPVAEAMDAIAAADLSQTSGFGPFGPQGIGQYTTWRDFICAIADPHVY
                     HWQTVMDDTVSASVAQALDELMLWAEDCPEVRHLVHADFGSNNVLTDNGRITAVIDWS
                     EAMFGDSQYEVANIFFWRPWLACMEQQTRYFERRHPELAGSPRLRAYMLRIGLDQLYQ
                     SLVDGNFDDAAWAQGRCDAIVRSGAGTVGRTQIARRSAAVWTDGCVEVLADSGNRRPS
                     TRPRAKE*"                (SEQ ID NO:31)
     gene            16830..17837
                     /label="hygroB"
                     /gene="hygroB"
     terminator      17973..18092
```

# FIG. 10C - CONSTRUCT 3 (FL/hXIST/RCAN1)

```
                     /label="SV40_PA_terminator"
     misc_feature    18061..18080
                     /label="EBV_rev_primer"
     misc_feature    18102..18860
                     /label="RCAN1 left arm"
     misc_feature    complement(18894..18912)
                     /label="M13_reverse_primer"
     misc_feature    complement(18911..18933)
                     /label="M13_pUC_rev_primer"
     promoter        complement(18947..18976)
                     /label="lac_promoter"
     rep_origin      complement(19285..19904)
                     /label="pBR322_origin"
     CDS             complement(20059..20919)
                     /label="Ampicillin"
                     /gene="Ampicillin"
                     /note="ORF frame 1"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW*"    (SEQ ID NO:25)
     gene            complement(20059..20919)
                     /label="Ampicillin"
                     /gene="Ampicillin"
     promoter        complement(20961..20989)
                     /label="AmpR_promoter"
```

# FIG. 10D - CONSTRUCT 4 (FL/hXIST/DYRK1A)

```
LOCUS       pcDNA5/TO/DYRK1A LA &#38; RA/FL human XIST   20737 bp     DNA    SYN    29-
Nov-2012
DEFINITION  pcDNA5/TO/DYRK1A LA &#38; RA/FL human XIST
ACCESSION
KEYWORDS
SOURCE
  ORGANISM  other sequences; artificial sequences; vectors.
FEATURES             Location/Qualifiers
     source          1..20737
                     /organism="pcDNA5/TO/DYRK1A LA &#38; RA/FL human XIST"
                     /mol_type="other DNA"
     misc_feature    209..716
                     /label="DYRK1A right arm"
     misc_feature    823..1110
                     /label="CAG_enhancer"
     promoter        744..1320
                     /label="CMV_immearly_promoter"
     misc_feature    1277..1297
                     /label="CMV_fwd_primer"
     promoter        1278..1401
                     /label="CMV_promoter"
     misc_feature    1328..1367
                     /label="tetO"
     promoter        1290..1463
                     /label="CMV2_promoter"
     misc_feature    1377..1401
                     /label="LNCX_primer"
     misc_feature    1540..15270
                     /label="Full length human XIST"
     misc_feature    complement(15469..15486)
                     /label="BGH_rev_primer"
     terminator      15472..15699
                     /label="bGH_PA_terminator"
     rep_origin      15762..16068
                     /label="f1_origin"
     misc_feature    complement(16182..16202)
                     /label="pBABE_3_primer"
     misc_feature    complement(16188..16403)
                     /label="SV40_enhancer"
     promoter        16200..16468
                     /label="SV40_promoter"
     rep_origin      16367..16444
                     /label="SV40_origin"
     misc_feature    16429..16448
                     /label="SV40pro_F_primer"
     CDS             complement(16502..17323)
                     /label="ORF frame 2"
                     /translation="MPPLEVARLLLHTSQPRPPEEDVGDLVLGIPEHRLAPVNDRCYA
                     AIVRQDIVGAEIRVHEVPDFGAVLGPKHQLIESLRDGRTDGVVHHSLPVIHMGISNRA
                     YEITPCSVLTDSLRSEWAEPARLAKIGRSDRIHGLRDRLQNSGQFGFRQVLQRDTLCT
                     AGDAIGQALAEFPNVKHFRNRERGRCKVPINITIFVETIGAAIYPQDISTPSYIEAES
                     TRFFALRELHQVGDAVELFDQKLLDRRRGEFRLFHHVLIRSENGYTSSRELFAKA*"
                                 (SEQ ID NO:30)
     CDS             16565..17590
                     /label="hygroB"
                     /gene="hygroB"
                     /note="ORF frame 2"
                     /translation="MKKPELTATSVEKFLIEKFDSVSDLMQLSEGEESRAFSFDVGGR
                     GYVLRVNSCADGFYKDRYVYRHFASAALPIPEVLDIGEFSESLTYCISRRAQGVTLQD
                     LPETELPAVLQPVAEAMDAIAAADLSQTSGFGPFGPQGIGQYTTWRDFICAIADPHVY
                     HWQTVMDDTVSASVAQALDELMLWAEDCPEVRHLVHADFGSNNVLTDNGRITAVIDWS
                     EAMFGDSQYEVANIFFWRPWLACMEQQTRYFERRHPELAGSPRLRAYMLRIGLDQLYQ
                     SLVDGNFDDAAWAQGRCDAIVRSGAGTVGRTQIARRSAAVWTDGCVEVLADSGNRRPS
                     TRPRAKE* (SEQ ID NO:31)
     gene            16580..17587
                     /label="hygroB"
                     /gene="hygroB"
     terminator      17723..17842
```

# FIG. 10D - CONSTRUCT 4 (FL/hXIST/DYRK1A)

```
                     /label="SV40_PA_terminator"
     misc_feature    17811..17830
                     /label="EBV_rev_primer"
     misc_feature    17853..18542
                     /label="DYRK1A left arm"
     misc_feature    complement(18576..18594)
                     /label="M13_reverse_primer"
     misc_feature    complement(18593..18615)
                     /label="M13_pUC_rev_primer"
     promoter        complement(18629..18658)
                     /label="lac_promoter"
     rep_origin      complement(18967..19586)
                     /label="pBR322_origin"
     CDS             complement(19741..20601)
                     /label="Ampicillin"
                     /gene="Ampicillin"
                     /note="ORF frame 1"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW*"    (SEQ ID NO:25)
     gene            complement(19741..20601)
                     /label="Ampicillin"
                     /gene="Ampicillin"
     promoter        complement(20643..20671)
                     /label="AmpR_promoter"
//
```

# FIG. 10E - CONSTRUCT 5 (6.8 kb/hXIST/RCAN1)

```
LOCUS       pcDNA5/TO/RCAN1 LA &#38; RA/6.8kb human XIST 14026 bp      DNA     SYN     29-
Nov-2012
DEFINITION  pcDNA5/TO/RCAN1 LA &#38; RA/6.8kb human XIST
ACCESSION
KEYWORDS
SOURCE
  ORGANISM  other sequences; artificial sequences; vectors.
FEATURES             Location/Qualifiers
     source          1..14026
                     /organism="pcDNA5/TO/RCAN1 LA &#38; RA/6.8kb human XIST"
                     /mol_type="other DNA"
     misc_feature    209..966
                     /label="RCAN1 right arm"
     misc_feature    1073..1360
                     /label="CAG_enhancer"
     promoter        994..1570
                     /label="CMV_immearly_promoter"
     misc_feature    1527..1547
                     /label="CMV_fwd_primer"
     promoter        1528..1651
                     /label="CMV_promoter"
     misc_feature    1578..1617
                     /label="tetO"
     promoter        1540..1713
                     /label="CMV2_promoter"
     misc_feature    1627..1651
                     /label="LNCX_primer"
     misc_feature    1790..8620
                     /label="6.8kb human XIST"
     misc_feature    complement(8641..8657)
                     /label="pBluescriptKS_primer"
     misc_feature    complement(8690..8707)
                     /label="BGH_rev_primer"
     terminator      8693..8920
                     /label="bGH_PA_terminator"
     rep_origin      8983..9289
                     /label="f1_origin"
     misc_feature    complement(9403..9423)
                     /label="pBABE_3_primer"
     misc_feature    complement(9409..9624)
                     /label="SV40_enhancer"
     promoter        9421..9689
                     /label="SV40_promoter"
     rep_origin      9588..9665
                     /label="SV40_origin"
     misc_feature    9650..9669
                     /label="SV40pro_F_primer"
     CDS             complement(9723..10544)
                     /label="ORF frame 3"
                     /translation="MPPLEVARLLLHTSQPRPPEEDVGDLVLGIPEHRLAPVNDRCYA
                     AIVRQDIVGAEIRVHEVPDFGAVLGPKHQLIESLRDGRTDGVVHHSLPVIHMGISNRA
                     YEITPCSVLTDSLRSEWAEPARLAKIGRSDRIHGLRDRLQNSGQFGFRQVLQRDTLCT
                     AGDAIGQALAEFPNVKHFRNRERGRCKVPINITIFVETIGAAIYPQDISTPSYIEAES
                     TRFFALRELHQVGDAVELFDQKLLDRRRGEFRLFHHVLIRSENGYTSSRELFAKA*"
                               (SEQ ID NO:30)
     CDS             9786..10811
                     /label="hygroB"
                     /gene="hygroB"
                     /note="ORF frame 3"
                     /translation="MKKPELTATSVEKFLIEKFDSVSDLMQLSEGEESRAFSFDVGGR
                     GYVLRVNSCADGFYKDRYVYRHFASAALPIPEVLDIGEFSESLTYCISRRAQGVTLQD
                     LPETELPAVLQPVAEAMDAIAAADLSQTSGFGPFGPQGIGQYTTWRDFICAIADPHVY
                     HWQTVMDDTVSASVAQALDELMLWAEDCPEVRHLVHADFGSNNVLTDNGRITAVIDWS
                     EAMFGDSQYEVANIFFWRPWLACMEQQTRYFERRHPELAGSPRLRAYMLRIGLDQLYQ
                     SLVDGNFDDAAWAQGRCDAIVRSGAGTVGRTQIARRSAAVWTDGCVEVLADSGNRRPS
                     TRPRAKE*"                      (SEQ ID NO:31)
     gene            9801..10808
                     /label="hygroB"
```

# FIG. 10E - CONSTRUCT 5 (6.8 kb/hXIST/RCAN1)

```
                     /gene="hygroB"
terminator           10944..11063
                     /label="SV40_PA_terminator"
misc_feature         11032..11051
                     /label="EBV_rev_primer"
misc_feature         11074..11831
                     /label="RCAN1 left arm"
misc_feature         complement(11865..11883)
                     /label="M13_reverse_primer"
misc_feature         complement(11882..11904)
                     /label="M13_pUC_rev_primer"
promoter             complement(11918..11947)
                     /label="lac_promoter"
rep_origin           complement(12256..12875)
                     /label="pBR322_origin"
CDS                  complement(13030..13890)
                     /label="Ampicillin"
                     /gene="Ampicillin"
                     /note="ORF frame 1"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW*"       (SEQ ID NO:25)
gene                 complement(13030..13890)
                     /label="Ampicillin"
                     /gene="Ampicillin"
promoter             complement(13932..13960)
                     /label="AmpR_promoter"
```

# FIG. 10F - CONSTRUCT 6 (6.8 kb/hXIST/DYRK1A)

```
LOCUS       pcDNA5/TO/DYRK1A LA &#38; RA/6.8kb human XIST      13708 bp      DNA     SYN
       29-Nov-2012
DEFINITION  pcDNA5/TO/DYRK1A LA &#38; RA/6.8kb human XIST
ACCESSION
KEYWORDS
SOURCE
  ORGANISM  other sequences; artificial sequences; vectors.
FEATURES             Location/Qualifiers
     source          1..13708
                     /organism="pcDNA5/TO/DYRK1A LA &#38; RA/6.8kb human XIST"
                     /mol_type="other DNA"
     misc_feature    209..716
                     /label="DYRK1A right arm"
     misc_feature    823..1110
                     /label="CAG_enhancer"
     promoter        744..1320
                     /label="CMV_immearly_promoter"
     misc_feature    1277..1297
                     /label="CMV_fwd_primer"
     promoter        1278..1401
                     /label="CMV_promoter"
     misc_feature    1328..1367
                     /label="tetO"
     promoter        1290..1463
                     /label="CMV2_promoter"
     misc_feature    1377..1401
                     /label="LNCX_primer"
     misc_feature    1540..8403
                     /label="6.8kb human XIST"
     misc_feature    complement(8391..8407)
                     /label="pBluescriptKS_primer"
     misc_feature    complement(8440..8457)
                     /label="BGH_rev_primer"
     terminator      8443..8670
                     /label="bGH_PA_terminator"
     rep_origin      8733..9039
                     /label="f1_origin"
     misc_feature    complement(9153..9173)
                     /label="pBABE_3_primer"
     misc_feature    complement(9159..9374)
                     /label="SV40_enhancer"
     promoter        9171..9439
                     /label="SV40_promoter"
     rep_origin      9338..9415
                     /label="SV40_origin"
     misc_feature    9400..9419
                     /label="SV40pro_F_primer"
     CDS             complement(9473..10294)
                     /label="ORF frame 2"
                     /translation="MPPLEVARLLLHTSQPRPPEEDVGDLVLGIPEHRLAPVNDRCYA
                     AIVRQDIVGAEIRVHEVPDFGAVLGPKHQLIESLRDGRTDGVVHHSLPVIHMGISNRA
                     YEITPCSVLTDSLRSEWAEPARLAKIGRSDRIHGLRDRLQNSGQFGFRQVLQRDTLCT
                     AGDAIGQALAEFPNVKHFRNRERGRCKVPINITIFVETIGAAIYPQDISTPSYIEAES
                     TRFFALRELHQVGDAVELFDQKLLDRRRGEFRLFHHVLIRSENGYTSSRELFAKA*"
                                 (SEQ ID NO:30)
     CDS             9536..10561
                     /label="hygroB"
                     /gene="hygroB"
                     /note="ORF frame 2"
                     /translation="MKKPELTATSVEKFLIEKFDSVSDLMQLSEGEESRAFSFDVGGR
                     GYVLRVNSCADGFYKDRYVYRHFASAALPIPEVLDIGEFSESLTYCISRRAQGVTLQD
                     LPETELPAVLQPVAEAMDAIAAADLSQTSGFGPFGPQGIGQYTTWRDFICAIADPHVY
                     HWQTVMDDTVSASVAQALDELMLWAEDCPEVRHLVHADFGSNNVLTDNGRITAVIDWS
                     EAMFGDSQYEVANIFFWRPWLACMEQQTRYFERRHPELAGSPRLRAYMLRIGLDQLYQ
                     SLVDGNFDDAAWAQGRCDAIVRSGAGTVGRTQIARRSAAVWTDGCVEVLADSGNRRPS
                     TRPRAKE*"                            (SEQ ID NO:31)
     gene            9551..10558
                     /label="hygroB"
```

# FIG. 10F - CONSTRUCT 6 (6.8 kb/hXIST/DYRK1A)

```
                     /gene="hygroB"
terminator           10694..10813
                     /label="SV40_PA_terminator"
misc_feature         10782..10801
                     /label="EBV_rev_primer"
misc_feature         10824..11513
                     /label="DYRK1A left arm"
misc_feature         complement(11547..11565)
                     /label="M13_reverse_primer"
misc_feature         complement(11564..11586)
                     /label="M13_pUC_rev_primer"
promoter             complement(11600..11629)
                     /label="lac_promoter"
rep_origin           complement(11938..12557)
                     /label="pBR322_origin"
CDS                  complement(12712..13572)
                     /label="Ampicillin"
                     /gene="Ampicillin"
                     /note="ORF frame 1"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW*"                         (SEQ ID NO:25)
gene                 complement(12712..13572)
                     /label="Ampicillin"
                     /gene="Ampicillin"
promoter             complement(13614..13642)
                     /label="AmpR_promoter"
```

# FIG. 10G - CONSTRUCT 7 (6.8 kb/hXIST/AAVS1)

```
LOCUS       AAVS1 HindIII/pcDNA5/TO/6.8kb human XIST    15721 bp     DNA    SYN    29-
Nov-2012
DEFINITION  AAVS1 HindIII/pcDNA5/TO/6.8kb human XIST
ACCESSION
KEYWORDS
SOURCE
  ORGANISM  other sequences; artificial sequences; vectors.
FEATURES             Location/Qualifiers
     source          1..15721
                     /organism="AAVS1 HindIII/pcDNA5/TO/6.8kb human XIST"
                     /mol_type="other DNA"
     promoter        143..172
                     /label="lac_promoter"
     misc_feature    186..208
                     /label="M13_pUC_rev_primer"
     misc_feature    207..225
                     /label="M13_reverse_primer"
     promoter        242..261
                     /label="T3_promoter"
     misc_feature    295..1095
                     /label="Left arm"
     misc_feature    complement(1126..1145)
                     /label="EBV_rev_primer"
     terminator      complement(1112..1231)
                     /label="SV40_PA_terminator"
     gene            complement(1369..2376)
                     /label="hygroB"
                     /gene="hygroB"
     CDS             complement(1366..2391)
                     /label="ORF frame 1"
                     /translation="MKKPELTATSVEKFLIEKFDSVSDLMQLSEGEESRAFSFDVGGR
                     GYVLRVNSCADGFYKDRYVYRHFASAALPIPEVLDIGEFSESLTYCISRRAQGVTLQD
                     LPETELPAVLQPVAEAMDAIAAADLSQTSGFGPFGPQGIGQYTTWRDFICAIADPHVY
                     HWQTVMDDTVSASVAQALDELMLWAEDCPEVRHLVHADFGSNNVLTDNGRITAVIDWS
                     EAMFGDSQYEVANIFFWRPWLACMEQQTRYFERRHPELAGSPRLRAYMLRIGLDQLYQ
                     SLVDGNFDDAAWAQGRCDAIVRSGAGTVGRTQIARRSAAVWTDGCVEVLADSGNRRPS
                     TRPRAKE*"    (SEQ ID NO: 31)
     CDS             1633..2454
                     /label="ORF frame 1"
                     /translation="MPPLEVARLLLHTSQPRPPEEDVGDLVLGIPEHRLAPVNDRCYA
                     AIVRQDIVGAEIRVHEVPDFGAVLGPKHQLIESLRDGRTDGVVHHSLPVIHMGISNRA
                     YEITPCSVLTDSLRSEWAEPARLAKIGRSDRIHGLRDRLQNSGQFGFRQVLQRDTLCT
                     AGDAIGQALAEFPNVKHFRNRERGRCKVPINITIFVETIGAAIYPQDISTPSYIEAES
                     TRFFALRELHQVGDAVELFDQKLLDRRRGEFRLFHHVLIRSENGYTSSRELFAKA*"
                           (SEQ ID NO: 30)
     misc_feature    complement(2508..2527)
                     /label="SV40pro_F_primer"
     rep_origin      complement(2512..2589)
                     /label="SV40_origin"
     promoter        complement(2488..2756)
                     /label="SV40_promoter"
     misc_feature    2559..2774
                     /label="SV40_enhancer"
     misc_feature    2754..2774
                     /label="pBABE_3_primer"
     rep_origin      complement(2888..3194)
                     /label="f1_origin"
     terminator      complement(3257..3484)
                     /label="bGH_PA_terminator"
     misc_feature    3470..3487
                     /label="BGH_rev_primer"
     misc_feature    3520..3536
                     /label="pBluescriptKS_primer"
     misc_feature    complement(3524..10387)
                     /label="6.8kb human XIST"
     misc_feature    complement(10526..10550)
                     /label="LNCX_primer"
     promoter        complement(10464..10637)
```

# FIG. 10G - CONSTRUCT 7 (6.8 kb/hXIST/AAVS1)

```
                     /label="CMV2_promoter"
     misc_feature    complement(10560..10599)
                     /label="tetO"
     promoter        complement(10526..10651)
                     /label="CMV_promoter"
     misc_feature    complement(10630..10650)
                     /label="CMV_fwd_primer"
     promoter        complement(10603..11183)
                     /label="CMV_immearly_promoter"
     misc_feature    complement(10816..11103)
                     /label="CAG_enhancer"
     misc_feature    11219..12058
                     /label="Right arm"
     promoter        complement(12094..12112)
                     /label="T7_promoter"
     misc_feature    complement(12119..12135)
                     /label="M13_forward20_primer"
     misc_feature    complement(12128..12150)
                     /label="M13_pUC_fwd_primer"
     misc_feature    12116..12259
                     /label="lacZ_a"
     misc_feature    12282..12575
                     /label="ccdB"
     promoter        12786..12835
                     /label="NEOKAN_promoter"
     CDS             12924..13718
                     /label="NeoR/KanR"
                     /gene="NeoR/KanR"
                     /note="ORF frame 3"
                     /translation="MIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAAVFRLSAQGR
                     PVLFVKTDLSGALNELQDEAARLSWLATTGVPCAAVLDVVTEAGRDWLLLGEVPGQDL
                     LSSHLAPAEKVSIMADAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDE
                     EHQGLAPAELFARLKASMPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADRY
                     QDIALATRDIAEELGGEWADRFLVLYGIAAPDSQRIAFYRLLDEFF*"  (SEQ ID NO: 29)
     gene            12927..13715
                     /label="NeoR/KanR"
                     /gene="NeoR/KanR"
     promoter        13851..13874
                     /label="AmpR_promoter"
     CDS             complement(13968..14828)
                     /label="Ampicillin"
                     /gene="Ampicillin"
                     /note="ORF frame 3"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW*"               (SEQ ID NO:25)
     gene            complement(13968..14828)
                     /label="Ampicillin"
                     /gene="Ampicillin"
     rep_origin      14936..15555
                     /label="pBR322_origin"
```

# FIG. 10H - CONSTRUCT 8 (6.3 kb/mXist/Runx1)

```
LOCUS       Runx1/pcDNA5/TO/6.3kb mouse Xist-right direction   20607 bp    DNA
      SYN   29-Nov-2012
DEFINITION  Runx1/pcDNA5/TO/6.3kb mouse Xist-right direction
ACCESSION
KEYWORDS
SOURCE
  ORGANISM  other sequences; artificial sequences; vectors.
FEATURES             Location/Qualifiers
     source          1..20607
                     /organism="Runx1/pcDNA5/TO/6.3kb mouse Xist-right
direction"
                     /mol_type="other DNA"
     misc_feature    49..4008
                     /label="Left arm"
     misc_feature    complement(4037..4056)
                     /label="EBV_rev_primer"
     terminator      complement(4023..4142)
                     /label="SV40_PA_terminator"
     gene            complement(4280..5287)
                     /label="hygroB"
                     /gene="hygroB"
     CDS             complement(4277..5302)
                     /label="ORF frame 2"
                     /translation="MKKPELTATSVEKFLIEKFDSVSDLMQLSEGEESRAFSFDVGGR
                     GYVLRVNSCADGFYKDRYVYRHFASAALPIPEVLDIGEFSESLTYCISRRAQGVTLQD
                     LPETELPAVLQPVAEAMDAIAAADLSQTSGFGPFGPQGIGQYTTWRDFICAIADPHVY
                     HWQTVMDDTVSASVAQALDELMLWAEDCPEVRHLVHADFGSNNVLTDNGRITAVIDWS
                     EAMFGDSQYEVANIFFWRPWLACMEQQTRYFERRHPELAGSPRLRAYMLRIGLDQLYQ
                     SLVDGNFDDAAWAQGRCDAIVRSGAGTVGRTQIARRSAAVWTDGCVEVLADSGNRRPS
                     TRPRAKE*"     (SEQ ID NO:31)
     CDS             4544..5365
                     /label="ORF frame 2"
                     /translation="MPPLEVARLLLHTSQPRPPEEDVGDLVLGIPEHRLAPVNDRCYA
                     AIVRQDIVGAEIRVHEVPDFGAVLGPKHQLIESLRDGRTDGVVHHSLPVIHMGISNRA
                     YEITPCSVLTDSLRSEWAEPARLAKIGRSDRIHGLRDRLQNSGQFGFRQVLQRDTLCT
                     AGDAIGQALAEFPNVKHFRNRERGRCKVPINITIFVETIGAAIYPQDISTPSYIEAES
                     TRFFALRELHQVGDAVELFDQKLLDRRRGEFRLFHHVLIRSENGYTSSRELFAKA*"
                                       (SEQ ID NO:30)
     misc_feature    complement(5419..5438)
                     /label="SV40pro_F_primer"
     rep_origin      complement(5423..5500)
                     /label="SV40_origin"
     promoter        complement(5399..5667)
                     /label="SV40_promoter"
     misc_feature    5470..5685
                     /label="SV40_enhancer"
     misc_feature    5665..5685
                     /label="pBABE_3_primer"
     rep_origin      complement(5799..6105)
                     /label="f1_origin"
     terminator      complement(6168..6395)
                     /label="bGH_PA_terminator"
     misc_feature    6381..6398
                     /label="BGH_rev_primer"
     misc_feature    complement(6637..12936)
                     /label="6.3kb mouse Xist"
     misc_feature    complement(13105..13129)
                     /label="LNCX_primer"
     promoter        complement(13043..13216)
                     /label="CMV2_promoter"
```

# FIG. 10H - CONSTRUCT 8 (6.3 kb/mXist/Runx1)

```
misc_feature    complement(13139..13178)
                /label="tetO"
promoter        complement(13105..13230)
                /label="CMV_promoter"
misc_feature    complement(13209..13229)
                /label="CMV_fwd_primer"
promoter        complement(13182..13762)
                /label="CMV_immearly_promoter"
misc_feature    complement(13395..13682)
                /label="CAG_enhancer"
CDS             13759..14364
                /label="ORF frame 1"
                /translation="MSTRISGPYIGPRGQEHSLCPTHPPTVGRGTLGNPVCPEPQHSG
                SLGSLCLPDHTLMPSLPLPAHSGSGPDRLRRPTPVPYSAVHLRPAHALPRRLHLLAAR
                HVGHRHRHVSHELGLSLPHRPAAALPRLITGAGRALPDRLALLPSILRRLGRFLPVLH
                GGRREIAPAHPAALHQRIHRRRAAQPQPPQPERRGGDRGQP*"
                      (SEQ ID NO:32)
misc_feature    13796..17681
                /label="Right arm"
misc_feature    complement(17733..17750)
                /label="Sp6_primer"
misc_feature    complement(17764..17782)
                /label="M13_reverse_primer"
misc_feature    complement(17781..17803)
                /label="M13_pUC_rev_primer"
promoter        complement(17817..17846)
                /label="lac_promoter"
rep_origin      complement(18155..18774)
                /label="pBR322_origin"
CDS             complement(18929..19789)
                /label="Ampicillin"
                /gene="Ampicillin"
                /note="ORF frame 2"
                /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                EIGASLIKHW*"      (SEQ ID NO:25)
gene            complement(18929..19789)
                /label="Ampicillin"
                /gene="Ampicillin"
promoter        complement(19831..19859)
                /label="AmpR_promoter"
rep_origin      complement(20104..20410)
                /label="f1_origin"
misc_feature    complement(20439..20587)
                /label="lacZ_a"
misc_feature    20553..20575
                /label="M13_pUC_fwd_primer"
misc_feature    20568..20584
                /label="M13_forward20_primer"
promoter        20591..2
                /label="T7_promoter"
```

# FIG. 10I - CONSTRUCT 9 (pEF1α/hDYRK1A/FL mXist)

```
LOCUS       pEF1x/hDYRK1A/FL mouse Xist 20647 bp DNA  SYN    27-Sep-2013
DEFINITION  pEF1x/hDYRK1A/FL mouse Xist
ACCESSION
KEYWORDS
SOURCE
  ORGANISM  other sequences; artificial sequences; vectors.
FEATURES             Location/Qualifiers
     source          1..20647
                     /organism="pEF1x/hDYRK1A/FL mouse Xist blunt ligation"
                     /mol_type="other DNA"
     promoter        191..1378
                     /label="EF1a_promoter"
     CDS             complement(471..1118)
                     /label="ORF frame 3"
                     /translation="MKRRLRTERPFSFVWVTHPPALPSAASSILSSLQQGREAAIFPL
                     TQLVPTGPALPPRAGRYTAARGQAPEQAGQLETTPVRFSVAALAGPASPNMCAGTHGP
                     RRRPRPQKPKYQCADLGPHLQDYLARKKASQQVIKNFKWLETYRKQRDRREGATRFAR
                     GGPSAQARPQLKHEAKGLLKRKASNSPTHFQPEARDQESRTAARGVEVIQGTQGP*"
                              (SEQ ID NO:27)
     misc_feature    1326..1346
                     /label="EF1a_fwd_primer"
     misc_feature    1378..1402
                     /label="LNCX_primer"
     misc_feature    1545..16406
                     /label="mouse FL Xist"
     misc_feature    16592..16607
                     /label="SV40_int"
     misc_feature    16613..16660
                     /label="SV40_3_splice"
     terminator      17236..17355
                     /label="SV40_PA_terminator"
     misc_feature    17324..17343
                     /label="EBV_rev_primer"
     misc_feature    17555..18244
                     /label="Human DYRK1A left arm"
     rep_origin      complement(18290..18909)
                     /label="pBR322_origin"
     CDS             complement(19064..19924)
                     /label="Ampicillin"
                     /gene="Ampicillin"
                     /note="ORF frame 2"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW*"            (SEQ ID NO:25)
     gene            complement(19064..19924)
                     /label="Ampicillin"
                     /gene="Ampicillin"
     promoter        complement(19966..19994)
                     /label="AmpR_promoter"
     misc_feature    20135..20642
                     /label="Human DYRK1A right arm"
```

# FIG. 10J - CONSTRUCT 10 (pEF1α/hDYRK1A/6.3kb mXist)

```
LOCUS      pEF1α/hDYRK1A/6.3kb mXist 12230 bp DNA    SYN    27-Sep-2013
DEFINITION  pEF1α/hDYRK1A/6.3kb mXist
ACCESSION
KEYWORDS
SOURCE
  ORGANISM  other sequences; artificial sequences; vectors.
FEATURES             Location/Qualifiers
     source          1..12230
                     /organism=" pEF1α/hDYRK1A/6.3kb mXist"
                     /mol_type="other DNA"
     promoter        44..1378
                     /label="pEF1x promoter"
     promoter        191..1378
                     /label="EF1a_promoter"
     CDS             complement(471..1118)
                     /label="ORF frame 3"
                     /translation="MKRRLRTERPFSFVWVTHPPALPSAASSILSSLQQGREAAIFPL
                     TQLVPTGPALPPRAGRYTAARGQAPEQAGQLETTPVRFSVAALAGPASPNMCAGTHGP
                     RRRPRPQKPKYQCADLGPHLQDYLARKKASQQVIKNFKWLETYRKQRDRREGATRFAR
                     GGPSAQARPQLKHEAKGLLKRKASNSPTHFQPEARDQESRTAARGVEVIQGTQGP*"
                              (SEQ ID NO:27)
     misc_feature    1326..1346
                     /label="EF1a_fwd_primer"
     misc_feature    1378..1402
                     /label="LNCX_primer"
     misc_feature    1579..7878
                     /label="6.3kb mouse Xist"
     misc_feature    8175..8190
                     /label="SV40_int"
     misc_feature    8196..8243
                     /label="SV40_3_splice"
     misc_feature    8100..8897
                     /label="SV40 pA"
     terminator      8819..8938
                     /label="SV40_PA_terminator"
     misc_feature    8907..8926
                     /label="EBV_rev_primer"
     misc_feature    9138..9827
                     /label="Human DYRK1A left arm"
     rep_origin      complement(9873..10492)
                     /label="pBR322_origin"
     CDS             complement(10647..11507)
                     /label="Ampicillin"
                     /gene="Ampicillin"
                     /note="ORF frame 3"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW*"                  (SEQ ID NO:25)
     gene            complement(10647..11507)
                     /label="Ampicillin"
                     /gene="Ampicillin"
     promoter        complement(11549..11577)
                     /label="AmpR_promoter"
     misc_feature    11718..12225
                     /label="Human DYRK1A right arm"
```

# FIG. 10K - CONSTRUCT 11 (Rosa26/pEF1x-Tet3G/hPGK-PuroR)

```
LOCUS       Rosa26/pEF1x-Tet3G/hPGK-PuroR   10391 bp    DNA    SYN    03-Oct-
2013
DEFINITION  Rosa26/pEF1x-Tet3G/hPGK-PuroR
ACCESSION
KEYWORDS
SOURCE
  ORGANISM  other sequences; artificial sequences; vectors.
FEATURES             Location/Qualifiers
     source          1..10391
                     /organism="Rosa26/pEF1x-Tet3G/hPGK-PuroR"
                     /mol_type="other DNA"
     rep_origin      complement(135..441)
                     /label="f1_origin"
     misc_feature    complement(459..614)
                     /label="lacZ_a"
     misc_feature    585..607
                     /label="M13_pUC_fwd_primer"
     misc_feature    600..616
                     /label="M13_forward20_primer"
     promoter        626..644
                     /label="T7_promoter"
     misc_feature    658..1452
                     /label="Rosa26 left arm"
     misc_feature    1497..2297
                     /label="AAVS1 left arm"
     misc_feature    2322..2764
                     /label="SV40 pA signal"
     terminator      2646..2777
                     /label="SV40_PA_terminator"
     misc_feature    2734..2753
                     /label="EBV_rev_primer"
     CDS             complement(2787..3533)
                     /label="ORF frame 3"
                     /translation="MSRLDKSKVINSALELLNGVGIEGLTTRKLAQKLGVEQPTLYWH
                     VKNKRALLDALPIEMLDRHHTHSCPLEGESWQDFLRNNAKSYRCALLSHRDGAKVHLG
                     TRPTEKQYETLENQLAFLCQQGFSLENALYALSAVGHFTLGCVLEEQEHQVAKEERET
                     PTTDSMPPLLKQAIELFDRQGAEPAFLFGLELIICGLEKQLKCESGGPTDALDDFDLD
                     MLPADALDDFDLDMLPADALDDFDLDMLPG*"    (SEQ ID NO:26)
     misc_feature    2787..3533
                     /label="Tet3G"
     misc_feature    complement(3615..3639)
                     /label="LNCX_primer"
     misc_feature    complement(3671..3691)
                     /label="EF1a_fwd_primer"
     CDS             3899..4546
                     /label="ORF frame 2"
                     /translation="MKRRLRTERPFSFVWVTHPPALPSAASSILSSLQQGREAAIFPL
                     TQLVPTGPALPPRAGRYTAARGQAPEQAGQLETTPVRFSVAALAGPASPNMCAGTHGP
                     RRRPRPQKPKYQCADLGPHLQDYLARKKASQQVIKNFKWLETYRKQRDRREGATRFAR
                     GGPSAQARPQLKHEAKGLLKRKASNSPTHFQPEARDQESRTAARGVEVIQGTQGP*"
                               (SEQ ID NO:27)
     promoter        complement(3639..4826)
                     /label="EF1a_promoter"
     promoter        3639..4973
                     /label="pEF1x promoter"
     misc_feature    5115..5621
                     /label="hPGK"
     CDS             5403..6242
                     /label="puro(variant)"
```

# FIG. 10K - CONSTRUCT 11 (Rosa26/pEF1x-Tet3G/hPGK-PuroR)

```
                  /gene="puro(variant)"
                  /note="ORF frame 3"
                  /translation="MAARRPRWAVANSGCSAGRAESSGREGAVREAGCGAVVWALFLP
                  ARCSAFCKPPERTSAVGSLVDRITDLSPQGDPPELTMTEYKPTVRLATRDDVPRAVRT
                  LAAAFADYPATRHTVDPDRHIERVTELQELFLTRVGLDIGKVWVADDGAAVAVWTTPE
                  SVEAGAVFAEIGPRMAELSGSRLAAQQQMEGLLAPHRPKEPAWFLATVGVSPDHQGKG
                  LGSAVVLPGVEAAERAGVPAFLETSAPRNLPFYERLGFTVTADVEVPEGPRTWCMTRK
                  PGA*"                    (SEQ ID NO:28)
gene              5643..6242
                  /label="puro(variant)"
                  /gene="puro(variant)"
misc_feature      5643..6242
                  /label="Puro"
misc_feature      6249..6473
                  /label="BGH pA"
terminator        6258..6473
                  /label="bGH_PA_terminator"
misc_feature      6489..7328
                  /label="AAVS1 right arm"
misc_feature      7366..8176
                  /label="Rosa26 right arm"
promoter          complement(8202..8221)
                  /label="T3_promoter"
misc_feature      complement(8238..8256)
                  /label="M13_reverse_primer"
misc_feature      complement(8255..8277)
                  /label="M13_pUC_rev_primer"
promoter          complement(8291..8320)
                  /label="lac_promoter"
rep_origin        complement(8629..9248)
                  /label="pBR322_origin"
CDS               complement(9403..10263)
                  /label="Ampicillin"
                  /gene="Ampicillin"
                  /note="ORF frame 1"
                  /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                  IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                  YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                  DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                  LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                  EIGASLIKHW*"              (SEQ ID NO:25)
gene              complement(9403..10263)
                  /label="Ampicillin"
                  /gene="Ampicillin"
promoter          complement(10305..10333)
                  /label="AmpR_promoter"
```