Published online July 29, 2004

Nucleic Acids Research, 2004, Vol. 32, No. 13 ***3995–4002***
doi:10.1093/nar/gkh746

Gene trap as a tool for genome annotation and analysis of X chromosome inactivation in human embryonic stem cells

Sujoy K. Dhara and Nissim Benvenisty*

Department of Genetics, The Hebrew University, Jerusalem 91904, Israel

Received March 31, 2004; Revised June 24, 2004; Accepted July 17, 2004

ABSTRACT

Human embryonic stem (ES) cells were suggested to be an important tool in transplantation medicine. However, they also play a major role in human genetics. Using the gene trap strategy, we have created a bank of clones with insertion mutations in human ES cells. These insertions occurred within known, predicted and unknown genes, and thus assist us in annotating the genes in the human genome. The insertions into the genome occurred in multiple chromosomes with a preference to larger chromosomes. Utilizing a clone where the integration occurred in the X chromosome, we have studied X-chromosome inactivation in human cells. We thus show that in undifferentiated female human ES cells both X chromosomes remain active and upon differentiation one chromosome undergoes inactivation. In the differentiated embryonic cells the inactivation is random, while in the extra-embryonic cells it is non-random. In addition, using a selection methodology, we demonstrate that in a minority of the cells partial inactivation and XIST expression occur even in the undifferentiated cells. We suggest that X chromosome inactivation during human embryogenesis, which coincides with differentiation, may be separated from the differentiation process. The genetic manipulation of human ES cells now opens new ways of analyzing chromosome status and gene expression in humans.

INTRODUCTION

Human embryonic stem (ES) cells are pluripotent cell lines derived from the inner cell mass (ICM) of blastocyst stage embryos (1). They are characterized by their ability to propagate indefinitely in culture, as undifferentiated cells, while they can be induced to differentiate *in vivo* into teratomas when injected into SCID mice (1,2). They may also differentiate *in vitro* into embryoid bodies (EBs) that contain embryonic cells from the three germ layers (3). Moreover, this differentiation can be somewhat directed by the addition of growth factors into the culture media (4). Human ES cells may be genetically manipulated in culture (5) and the transfected cells remain pluripotent and retain a normal karyotype (6). As a result of their unique features, human ES cells have been suggested to hold the promise of changing the face of cell transplantation. Human ES cells can serve as a putative source of numerous types of differentiated cells needed in different pathologies and also as a component in biomedical engineering. These unique cell lines also provide a valuable tool for the study of early human development. This is because human ES cells express genes common with the inner cell mass (7) and the differentiating EBs express genes that appear during organogenesis (4).

With the availability of the human genome sequence, human ES cells may assist us to better annotate the genome. To achieve this, we employed the gene trapping technique, which is a form of insertional mutagenesis specifically designed to disrupt gene function by producing intragenic integration events (8). This method is considered to be a powerful genetic tool to elucidate molecular mechanisms of complex biological phenomena. Various types of trapping systems were exercised in murine cells. Among these are promoter and enhancer trap (9,10) and polyadenylation [poly(A)] trap (11,12) systems. The poly(A) trap vector is highly advantageous since the occurrence of trapping with poly(A) trap is independent of expression of the target gene. Thus, any gene could potentially be identified at almost equal probability regardless of the relative abundance of its transcripts in target cells (13). By trapping genes in human ES cells *in vitro*, one may identify new genes in the genome, study the relevance of trapped genes to the process of early embryonic differentiation and generate mutant human cells. In addition, if the trapped gene resides on chromosome X the system can be used to study X chromosome inactivation in different cell lineages (both embryonic and extra-embryonic) using human embryonic stem cells. This is an issue because of the inaccessibility of the early human embryo and the ability of using genetically labeled human ES cells to study early developmental processes.

MATERIALS AND METHODS

Cell culture

Human ES H9 cells of passage 45–48 with normal karyotype (see supporting figure http://www.ls.huji.ac.il/~nissimb/Cytogenetic_analysis.htm) were used in the present

*To whom correspondence should be addressed at The Herbert Cohn Chair in Cancer Research, Department of Genetics, Institute of Life Sciences, The Hebrew University of Jerusalem, Givat-Ram, Jerusalem 91904, Israel. Tel: +972 2 6586774; Fax: +972 2 6584972; Email: nissimb@mail.ls.huji.ac.il

experiment (1). These cells were cultured on a Mitomycin-C-treated mouse embryonic fibroblast (MEF) feeder layer (obtained from day 13.5 embryos). The cells were grown in 80% KnockOut^{TM} DMEM medium (Invitrogen, Inc.), supplemented with 20% KnockOut^{TM} SR—a serum-free formulation (Invitrogen, Inc.), 1 mM glutamine (Invitrogen, Inc.), 0.1 mM 2-mercaptoethanol (Sigma), 1% non-essential amino acids stock (Invitrogen, Inc.), penicillin (50 U/ml), streptomycin (50 μg/ml) and 4 ng/ml of basic fibroblast growth factor (bFGF). Embryoid bodies and differentiated ES (DES) cells were obtained as previously described (4). To differentiate human ES cells into extra-embryonic trophoblasts, the cells were grown without MEF in the presence of BMP4 (100 ng/ml, R&D Systems, Inc.) (14).

Transfection and establishment of gene trap clones

Human ES cells underwent stable transfection with the RET gene trap C1010 plasmid (13) by the ExGen 500 transfection system (Fermentas) as previously described (5). Specifically, transfection of human ES cells was carried out in 6-well trays on MEF two days after plating. The plasmid DNA (2 μg) and the transfecting agent ExGen 500 (10 μl) were added to ~3–5×10^5 cells in a final volume of 1 ml media per well. The cells were centrifuged at 280 *g* for 5 min and incubated at 37°C in a moist chamber for an additional 30 min. Residuals of transfecting agent were removed by washing twice with phosphate-buffered saline (PBS). The following day, the cells were trypsinized and approximately 10^6 were re-plated on each of the 10 cm^2 culture dishes containing inactivated MEF^{neo+}. Two days following re-plating, G418 (200 μg/ml; Invitrogen, Inc.) was administered to the growth medium, allowing selective propagation of transfected cells in culture. By day 14, G418 resistant colonies were identified. Single transgenic colonies were manually collected by a micropipette, dissociated into small clumps of cells and transferred into a 2 cm^2 (24-well) culture dish, on a fresh feeder of MEF^{neo+}. The cells continuously proliferated in the presence of G418 and formed a large number of expanding undifferentiated colonies.

Identification of the integration sites in the gene trap clones

RNA was isolated from each of the gene trap clones using guanidium thiocyanate (15). To identify the sequence of the gene into which the neomycin phosphotransferase (*neo^R*) gene has been inserted, 3′ rapid amplification of cDNA ends (RACE) analysis was performed as previously described (13). Briefly, first strand cDNA was synthesized from total RNA using an adaptor primer with poly(T) [AD–poly(T)] and M-MLV reverse transcriptase (Promega, Inc.) enzyme. Then, the cDNA containing the neo^R transcript was amplified twice by PCR with nested primers (first round: NEO 1.5 and AD; second round: NEO 2.0 and AD-plus) (13). The PCR conditions were as follows: (i) 94°C, 3 min (first round) or 1 min (second round), (ii) 8× (94°C, 40 s + 72°C, 4 min), (iii) 27× (94°C, 40 s + 66°C, 2 min + 72°C, 2 min) and (iv) 72°C, 4 min. The PCR product was confirmed and cleaned from an agarose gel (1%) using GFX-Gel Band cleaning kit (Amersham Biosciences) and then, using primer (NEO.SEQ), the product was sequenced in an automated DNA sequencer [ABI prism 3700 (v3.3)] using a fluorescent dye terminator reaction. The sequence was examined for the presence of a splice junction sequence (GAAT) and then the genomic sequence was identified in the human BLAT search (16).

Isolation of clones with inactivated transgenes

Gene trap enables us to have clones with the plasmid containing active *neo^R* and *HSV–TK* genes. To enrich and subsequently isolate clones in which the transgene was inactivated, 10^6 ES cells were grown in the presence of 0.2 μM FIAU (Moravek Biochemicals, Inc.), a drug that selects for HSV–TK$^-$ cells. The experiment was performed on gelatin coated 10 cm^2 Flacon tissue culture plate with human ES cell culture media preconditioned with MEF (scheme shown in Figure 3A). The medium was changed daily, and after about 2 weeks FIAU-resistant colonies appeared. Individual clones were expanded on feeder MEF cells with conditions similar to culture of ES cells as previously described (4).

Analysis of gene expression

RNA was isolated as previously described (15) and cDNA was synthesized from total RNA using hexamer random primer (3). The PCR mixture contained 10 mM Tris–HCl, 50 mM KCl, 3 mM $MgCl_2$, 200 μM of each dNTPs, 0.4 μM each primer and 0.5 U of *Taq* polymersase in a total volume of 25 μl. The cycling reaction was performed at 65°C for 30 to 40 cycles. Expression of X-inactive specific transcript (*XIST*), *OCT4* (ES specific marker), glial cells missing homolog 1 (*GCM1*, a marker for trophoblast cells), chorionic gonadotropin beta polypeptide 5 (*CGB5*, a marker for trophoblast cells) and glyceraldehyde-3-phosphate dehydrogenase (*G3PDH*, a housekeeping gene) were analyzed by PCR using human-specific primers enlisted in Table 1 and PCR products were

Table 1. List of genes and information on their PCR amplification

Gene	Primers
Fragile X mental retardation 1 (*FMR1*)	F: GCTCAGCTCCGTTTCGGTTTCACTTCCGGT R: AGCCCCGCACTTCCACCACCAGCTCCTCCA
X-inactive specific transcript (*XIST*)	F: CTACAAGCAGTGCAGAGAGC R: CTAAGACAAGACACAGACCAC
OCT4	F: GATCCTCGGACCTGGCTAAG R: CTCTCACTCGGTTCTCGATAC
Glial cells missing homolog 1 (Drosophila) (*GCM1*)	F: CTACCAGGCAATTGGACGCC R: CAACTGTGACGGGCCTCTGA
Chorionic gonadotropin, beta polypeptide 5 (*CGB5*)	F: CAGGACCCCACCATAGGCAG R: CTCCTTGGATGCCCATGTCC
HSV–thymidine kinase (*HSV–TK*)	F (T1): CTTCCGGAGGACAGACACAT F (T2): TCGGGGACACGTTATTTACC R (T3): GACCATCCCGGAGGTAAGTT
Neomycin phosphotransferase (*neo^R*)	R (N1): CTCGTCCTGCAGTTCATTCA R (N2): GAAGGCGATAGAAGGCGATG
Glyceraldehyde-3-phosphate dehydrogenase (*G3DPH*)	F: GAGTCAACGGATTTGGTCGT R: TTGATTTTGGAGGGATCTCG

F, forward; R, reverse;
The annealing temperature for all reactions is 65°C.

resolved on 1% agarose gel. Expression of fragile X mental retardation 1 (*FMR1*) gene was performed using the same PCR primers as described in Benjamin *et al*. (17), and the products were resolved in 3% Nusieve GTG agarose gel (Cambrex Bioscience, Inc).

Analysis for chromosomal deletions in transgene-inactivated clones

Genomic DNA was isolated from each of the HSV–TK-inactive clones using EZ-DNA Genomic DNA Isolation Kit (Biological Industries, Inc.). The DNA was analyzed for informative microsatellite markers along the X chromosome: four markers located upstream to the transgene insert (DXS1060, DXS1226, DXS991 and DXS986, which reside 85, 78, 34 and 12 MB from the insert, respectively) and two were downstream to the insert (DXS1106 and DXS8043, which reside 11 and 52 MB from the insert). All primers were available in the ABI panel and end labeled with fluorescent dyes. PCR products were run on an ABI Prism 3700 DNA Analyzer (Perkin Elmer, Inc.). For detection of the HSV–TK gene, PCR was performed with primers listed in Table 1 and conditions described in the previous section.

Staining of cells for expression of OCT4 and alkaline phosphatase

Antibody staining for *OCT4* expression was performed following a published protocol (18). In short, colonies were fixed with 4% paraformaldehyde at room temperature for half an hour. Subsequently, after thorough washing with PBS, blocking of non-specific proteins was achieved with blocker (2% BSA in PBS, 0.1% Triton X-100 and 10% low fat milk) for 1 h in room temperature. Next, cells were stained with primary monoclonal mouse Oct4 antibody (Santa Cruz Biotechnology, Inc.; dilution 1:50) overnight at 4°C and further stained with Cy3 conjugated goat anti-mouse secondary polyclonal antibody (Jackson ImmunoResearch, Inc.; 1:200) for an hour at room temperature. Nuclear staining was performed with Hoechst 33258 (Sigma, Inc.).

Alkaline phosphatase expression was examined with Alkaline Phosphatase kit (Sigma Diagostics, Inc.) following the manufacturer's protocol.

RESULTS AND DISCUSSION

In order to establish the system of human ES cells as a genetic tool to further annotate the human genome and study human development, we aimed at creating a library of clones targeted at different genes in the genome using poly(A)-based gene trap approach. Human ES cells were transfected with the RET gene trap construct (13). Transfection of this construct may enable resistance to G418 only if it has been inserted into an indigenous gene utilizing its poly(A) signal (13,19). Thus, following transfection, we have isolated over 100 different clones (out of 10^9 cells) using G418 selection. This was performed by a series of seven different independent transfection experiments. Therefore, using the RET gene trap plasmid we could show efficiency of about 10^{-7} neo resistance clones. This efficiency of gene trapping is about 10-fold lower than that of transfection into random integration sites which consisted of 10^{-6} resistant clones (5). The integration sites of the clones were identified by 3′ RACE (20). Thus, RNA was isolated from independent clones with guanidium thiocyanate (15), and then cDNA was synthesized followed by two rounds of 35 cycles of PCR to amplify the neo^R transcripts tagged with the genomic sequence. The product was sequenced using a dye terminator reaction in ABI automated sequencer. The sequence search was conducted with BLAT program (16) to identify the insertion sites in the human genome (Table 2).

From over 100 initial clones, 74 could be propagated, frozen and thawed. In 54 of these clones, we achieved amplification by 3′ RACE RT–PCR. Out of the 54 clones, 19 were identified to have only plasmid sequences, resulting from multiple integrations of the transgene, and the use of the internal poly(A) signal. In three clones the PCR product did not have a match in the already sequenced human genome. In the remaining 32 clones, we could identify independent integration sites and show that the neo-resistance transcript resulted from the integration using the splice donor of the construct, and 3′ sequences and the poly(A) signal of the targeted gene. Out of the 32 informative inserted regions, 18 were in known genes that have a Unigene identity (Table 2). By analyzing the expression of these genes by DNA microarray [see Dvash *et al*. (21)] we could show that while most insertions occurred in actively expressed genes, some of the genes involved are not active (or are expressed in very low levels) in human ES cells (data not shown). Thus, we may also trap genes that are not expressed in human ES cells. Of the insertions, 11 occurred in predicted genes not yet verified by existence of an mRNA transcript [or an expressed sequence tag (EST)], and thus may validate their authenticity. In the remaining 3 out of the 32 insertion sites no gene prediction is given, and thus may not reflect a true gene, or may indicate the existence of a gene not predicted by other means.

In our analysis various different genes were targeted. Among them are housekeeping genes such as H2B histone and histidyl-tRNA synthetase. In addition, cell-specific transcripts such as a potassium voltage-gated channel (*KCN2*) were trapped and the targeted genes varied between nuclear, cytosolic or membrane bound proteins. The 32 integration incidents occurred in 17 chromosomes with a preference for larger chromosomes (Figure 1). Thus, in chromosomes 1 through 14, we find at least one integration event. In addition, ~80% of the integration events occurred on the long arms of the chromosomes (Figure 1). An exception to the observation of random integration was 3 insertions that occurred on the short chromosome 18, two of which were on 18q21.

The sequencing of the human genome has provided us with the unprecedented opportunity to deepen the annotation of the genes, which is far from complete. Thus, identifying genes via gene trap may assist us in annotating various genes in the genome even if we do not yet identify cells that express them. This methodology may identify genes that are expressed at very low levels, at a short time window during development, or in cells that exist in very low number. However, there are several limitations in the gene trap methodology, such as lack of effective prescreening of trapped genes, leakiness of poly(A)-based selection, integration of multiple copies of the trap vector, etc.

We have utilized the gene trap method also as a system to analyze the inactivation status of X chromosome. During differentiation, female cells inactivate one of the two

Table 2. Analysis of the integration sites of the gene trap construct in clones of human ES cells

Clone name	Location	Unigene	Gene name
Known genes			
C1010#6	10q21.1	Hs.232819	Protocadherin 15 (*PCDH15*)
C1010#9	18q21.1	Hs.14328	Dymeclin (*FLJ20071*)
C1010#20	12q21.2	Hs.16533	Protein phosphatase 1, regulatory subunit 12A (*PPP1R12A*)
C1010#25	5q31.3	Hs.77798	Histidyl-tRNA synthetase (*HARS*)
C1010#26	9q34.3	Hs.12999	AD038 (*LOC85026*)
C1010#27	6p22.2	Hs.180779	Histone 1, H2BD (*HIST1H2BD*)
C1010#47	7q35	Hs.106552	Contactin associated protein-like 2 (*CNTNAP2*)
C1010#51	5q34	Hs.353229	LOC202400
C1010#64	20q13.32	Hs.182625	VAMP-associated protein B and C (*VAPB*)
C1010#67	18q21.33	Hs.38176	Pleckstrin homology domain, family E member 1(*PLEKHE1*)
C1010#68	3p13	Hs.437983	AW295978
C1010#74	4q13.1	Hs.21917	Latrophilin 3/LEC3 (*LPHN3*)
C1010#78	3p14.2	Hs.89627	Protein tyrosine phosphatase, receptor type, G (*PTPRG*)
C1010#82	7q31.31	Hs.202687	Potassium voltage-gated channel, member 2 (*KCND2*)
C1010#83	11q25	Hs.288433	Neurotrimin (*HNT*)
C1010#87	13q14.11	Hs.80683	Mitochondrial translational release factor 1 (*MTRF1*)
C1010#98	Xq21.31	Hs.374280	Protocadherin 11 X-linked (*PCDH11X*)
C1023#1	7q11.22	Hs.7981	Williams–Beuren syndrome chromosome region 17 (*WBSCR17*)
Predicted genes			
C1010#2	2p25.2	—	NT_005334.4[a]
C1010#5	Xq22.1	—	NT_011651.262[a]
C1010#8	18q23	—	NT_025004.96[a]
C1010#17	8q12.1	—	C8000787[b]
C1010#30	14q31.1	—	NT_026437.1307[a]
C1010#33	13q21.33	—	NT_024524.677[a]
C1010#69	Xq21.31	—	ENS330623.1[c]
C1010#77	4q35.2	—	NT_022792.455[a]
C1010#89	1q21.2	—	NT_032962.3[a]
C1010#103	4p15.1	—	C4000355[b]
C1010#105	3p25.1	—	NT_005927.269[a]
Unidentified genes			
C1010#22	4q22.2	—	—
C1010#61	3q27.3	—	—
C1010#95	1q41	—	—

Cells were transfected with linearized RET construct and selected with G418 to isolate gene targeted clones. RNA was isolated from each clone and 3′ RACE was utilized to identify the sequence of each integration site and chromosomally localized according to BLAT search. Genes predicted by programs: [a]Genscan, [b]Fgenesh++, [c]Ensembl.

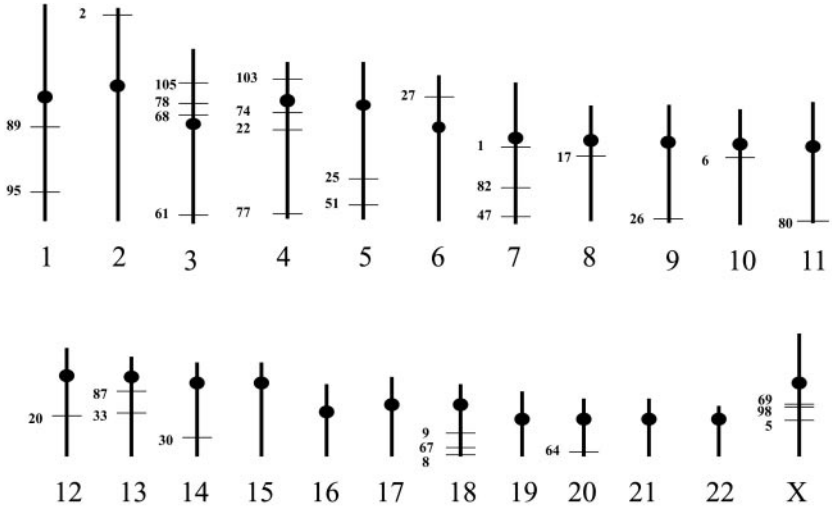

Figure 1. A schematic diagram showing names and locations of the trap clones on different chromosomes. Locations are approximated to the picture obtained from 'view genome' option provided in NCBI BLAST human genome service website.

chromosomes to compensate for the dosage of X-linked genes compared to male cells. The timing and nature of this inactivation in human ES cells is still unknown. In order to analyze this phenomenon we have devised a novel strategy using a gene trap clone where the neo^R gene is integrated in a region that does not escape X chromosome inactivation.

By identifying a polymorphic marker on chromosome X in female human ES cells, we can analyze the expression of the

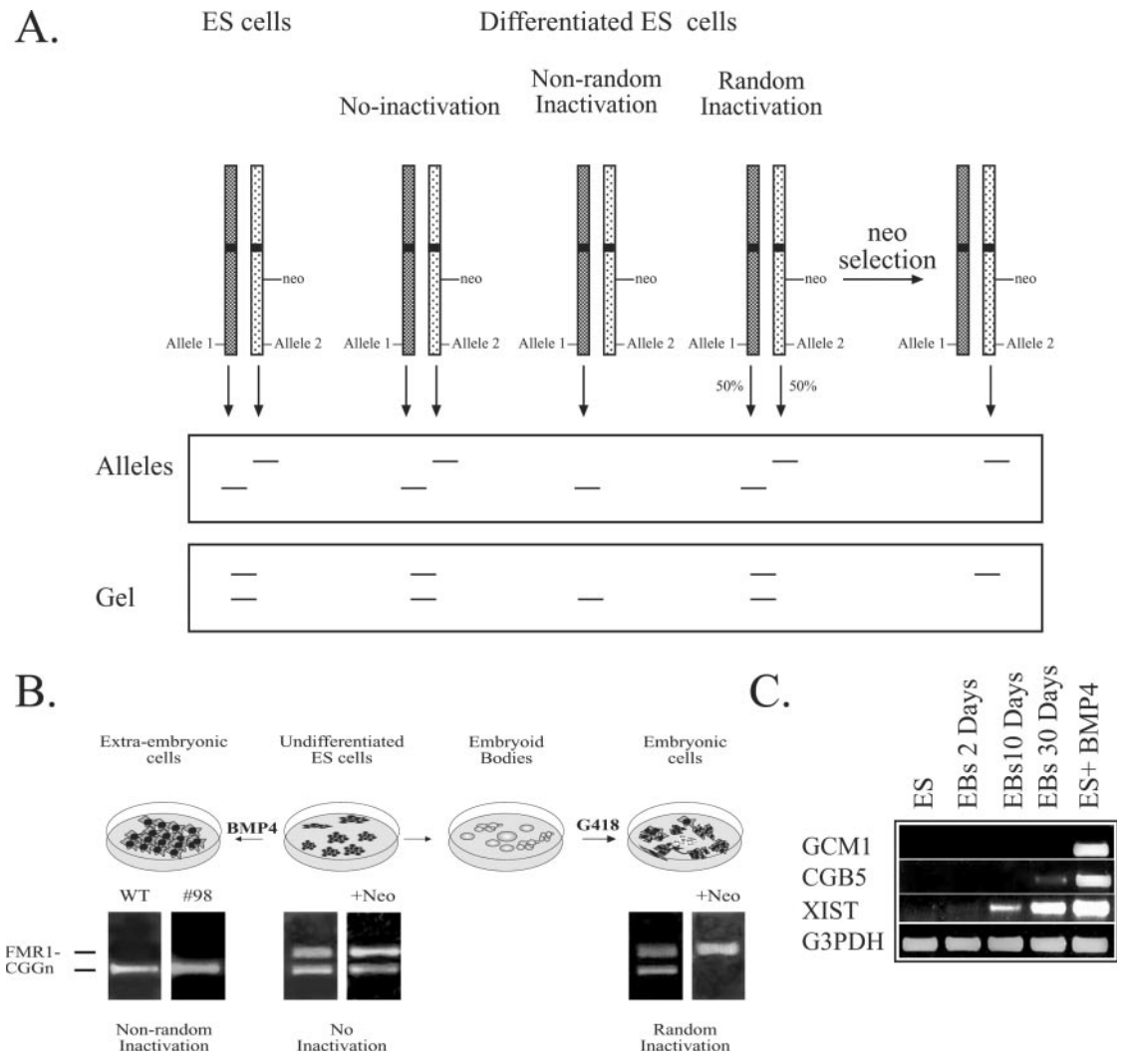

Figure 2. Analysis of X inactivation in a clone with a transgene on X chromosome. (**A**) Schematic diagram showing expression of a polymorphic gene depending on the status of X chromosome inactivation in ES and differentiated ES (DES) cells with or without G418 selection. (**B**) Status of expression of the two alleles of *FMR1*–CGG_n in ES cells in differentiated embryonic cells and in trophoblast (extra embryonic) cells of H9 (wild type) and C1010#98 clone. (**C**) Expression of X-inactive specific transcript (*XIST*), glial cells missing 1 (*GCM1*) and chorionic gonadotropin beta polypeptide 5 (*CGB5*) in wild-type ES cells and EBs 2, 10 or 30 days after differentiation and in trophoblast cells derived from clone C1010#98 using BMP4 (expression of G3PDH is shown as control).

two alleles before and after differentiation. Thus, if the two X chromosomes are active in human ES cells we will identify expression of the two alleles (see Figure 2A). Following differentiation, if no X chromosome inactivation occurs we would still find the two alleles. However, if one chromosome is inactivated in a non-random/skewed manner, then only one allele will be expressed (Figure 2A). However, if the inactivation occurs in a random fashion, then in 50% of the cells one allele is active, and in the other 50% the other allele is active. Since we are analyzing a culture of cells, we would still observe expression of the two alleles, as if no inactivation occurred. In order to recognize the random inactivation process we are using a cell line on which a selection marker resides on chromosome X. After selection, only the cells in which the active X chromosome harbors the selection marker would survive the drug treatment and only one of the polymorphic alleles would be active, allowing us to identify the random X inactivation process (Figure 2A).

To analyze the status of X chromosome inactivation in human ES cells, the highly polymorphic *FMR1* gene was examined. This gene, which resides on Xq27.3, harbors numerous CGG repeats, and in extreme cases of long repeats results in fragile X syndrome (22). Our cell line was indeed polymorphic in this gene as demonstrated by the two bands in RT–PCR analysis of FMR1 expression. The expression of these alleles was also tested in clone C1010#98, where the neo^R gene resides on chromosome X. The two polymorphic alleles were active in the undifferentiated cells with or without G418 selection (Figure 2B, middle panel). However, after differentiation the two alleles appeared in the cells before selection, but one of them disappeared after selection (Figure 2B, right panel). Further, when ES cells were treated with BMP4, they differentiate into true trophoblast cells (14). This is evident from the very high expression levels of *GCM1* and *CGB5* genes that are molecular markers for the trophoectoderm (Figure 2C). In these trophoblast cells, the

expression of only one allele of *FMR1* gene was observed. The monoallelic expression was noted without the need for G418 selection. Moreover, since the chromosome that was inactivated is the one that harbors the transgenes, G418 selection destroyed most of the cells. In addition, the expression of only one allele in the trophoblast cells could also be demonstrated in the wild-type (H9) cells (Figure 2B). Thus, according to the methodology outlined in Figure 2A, we can conclude that X chromosome inactivation occurs in a random fashion in the differentiated cells of embryonic origin, but, in a non-random fashion in trophoblast cells.

The observation that X chromosome inactivation occurs during differentiation of human ES cells is further supported by the analysis of expression of *XIST* gene. The *XIST* gene encodes for an RNA molecule that is responsible for transcriptional silencing of X-linked genes and its expression is correlated with a general shut-off in gene expression on the inactive X chromosome [for a review see (23)]. By RT–PCR analysis we could show that XIST is expressed in extremely low levels in the culture of ES cells and gradually increased during differentiation into embryoid bodies as well as into trophoblast cells. (Figure 2C).

The status of X chromosome inactivation in mouse trophoblast cells is well documented, where the X chromosome is inactivated in a non-random fashion (with preferential paternal X chromosome inactivation) (24). However, the data are very conflicting with regard to human trophoblast cells, where several reports claimed skewed X chromosome inactivation (similar to mouse), while others described random inactivation [(25,26) and references within (27)]. This kind of confusion arises because of a lack of proper experimental materials from developing human embryos. We have overcome this problem with the help of an ES clone and by differentiating it into extra-embryonic trophoblast cells. The availability of human ES cells now allows us to analyze the process of inactivation in culture into the most early trophoblast cells. We could thus suggest that in human the X chromosome inactivation in trophoectoderm is indeed different from embryonic cells and is probably under imprinting. Since the identity of the donors for human ES cells is not disclosed, as of now, it is not possible to ascertain the parental origin of the inactivated X chromosome in the extra-embryonic cells.

Gene trap insertions, harboring HSV–TK and neo^R genes, were identified on several places in the genome including X chromosome. In order to isolate subclones that have lost the expression of the transgenes on X chromosomes, we have grown the cells in the presence of FIAU, a drug that specifically kills cells expressing the HSV–TK gene (Figure 3A). The rate of FIAU-resistant clones was $3.4 \pm 0.2 \times 10^{-4}$ when the HSV–TK transgene resided on X chromosome, while if the transgene resided on an autosomal chromosome, the rate of FIAU-resistant clones dropped to about 10^{-6} cells (results based on series of repeated experiments with more than 20×10^6 cells). We isolated and studied several clones where the HSV–TK gene on the X chromosome overcame inactivation. Interestingly, after FIAU selection, the clones gained resistance to FIAU but became sensitive to G418 selection.

The loss of HSV–TK and neo^R activities may have resulted from genetic or epigenetic processes. To test for chromosomal deletions in the X chromosome (where the transgene resides), a panel of six informative microsatellite markers located around the insert was tested. As shown in Figure 3B, for three representative microsatellite markers, the same bands appear in the original clones and in their FIAU-resistant subclones. We may thus conclude that no chromosomal loss or major deletion (>20 cM) occurred in the clones. We further demonstrated the presence of *HSV–TK* and *neo^R* sequences by PCR at the genomic level as a qualitative assay (Figure 3C). Thus, transgene inactivation seems more likely to result from the epigenetic loss of expression of both *HSV–TK* and *neo^R* genes. Since inactivation of genes on the X chromosome coincides with induction of *XIST* gene, we have examined the expression of *XIST* in the clones where the transgenes were inactivated. As shown in Figure 3C, expression of *XIST* was indeed induced to very high levels in these subclones. The clones that start to express high levels of *XIST* are still undifferentiated. When examined for morphological features, these cells grow in compact colonies very much similar to undifferentiated ES cells. Further, they simultaneously express high levels of OCT4 transcript (Figure 3D), a well-known marker for the undifferentiated cells. These colonies also highly express the *OCT4* protein (staining with monoclonal antibody) and were positive for alkaline phosphatase (histochemical detection) expression (Figure 3E). Beside these evidences, these cells could successfully form matured EBs (data not shown) demonstrating their pluripotent nature. Therefore, onset of XIST expression may represent a phenomenon that occurs before true differentiation is established. Thus, our system should enable a stepwise analysis of X chromosome inactivation during embryonic differentiation.

The observation that X chromosome inactivation occurs during differentiation of human ES cells allows us to analyze this process in a culture system of human cells. Previously, there were informative attempts to study the capacity of human *XIST* transgene to induce chromosome inactivation in mouse ES cells (28,29). As differences in X chromosome activation between human and mouse are evident (30,31), the analysis of RET gene trap in human ES cells allows the study of X chromosome inactivation in the human context. The new methodology of analyzing X chromosome inactivation as outlined in Figure 2 should assist us in the future in studying the possibility of non-random versus random inactivation in specific lineages that are derived from human ES cells. Moreover, by insertion of different selection markers into the two X chromosomes we can try to analyze the consequences of activation of the two X chromosomes in somatic cells. In addition, this methodology may also assist in the analysis of autosomal chromosomes where only one allele is active. Thus, the imprinting of genes during differentiation and embryonic development of human cells may also be analyzed using this methodology.

The full potential of this methodology would be revealed when gene trapping will be exercised on a much larger scale with the aim of targeting all genes in the human genome. This analysis gives us new tools to study human genetics. As differentiation of ES cells recapitulates to some extent normal development, the effects of specific genetic mutations on differentiation in culture may be related to their gene function *in vivo*. In addition, we may also achieve specific differentiated cell lines mutated in certain genes from the gene trap clones. Although the gene trap analysis targets only one allele, homozygosity of the mutation may be achieved by selection

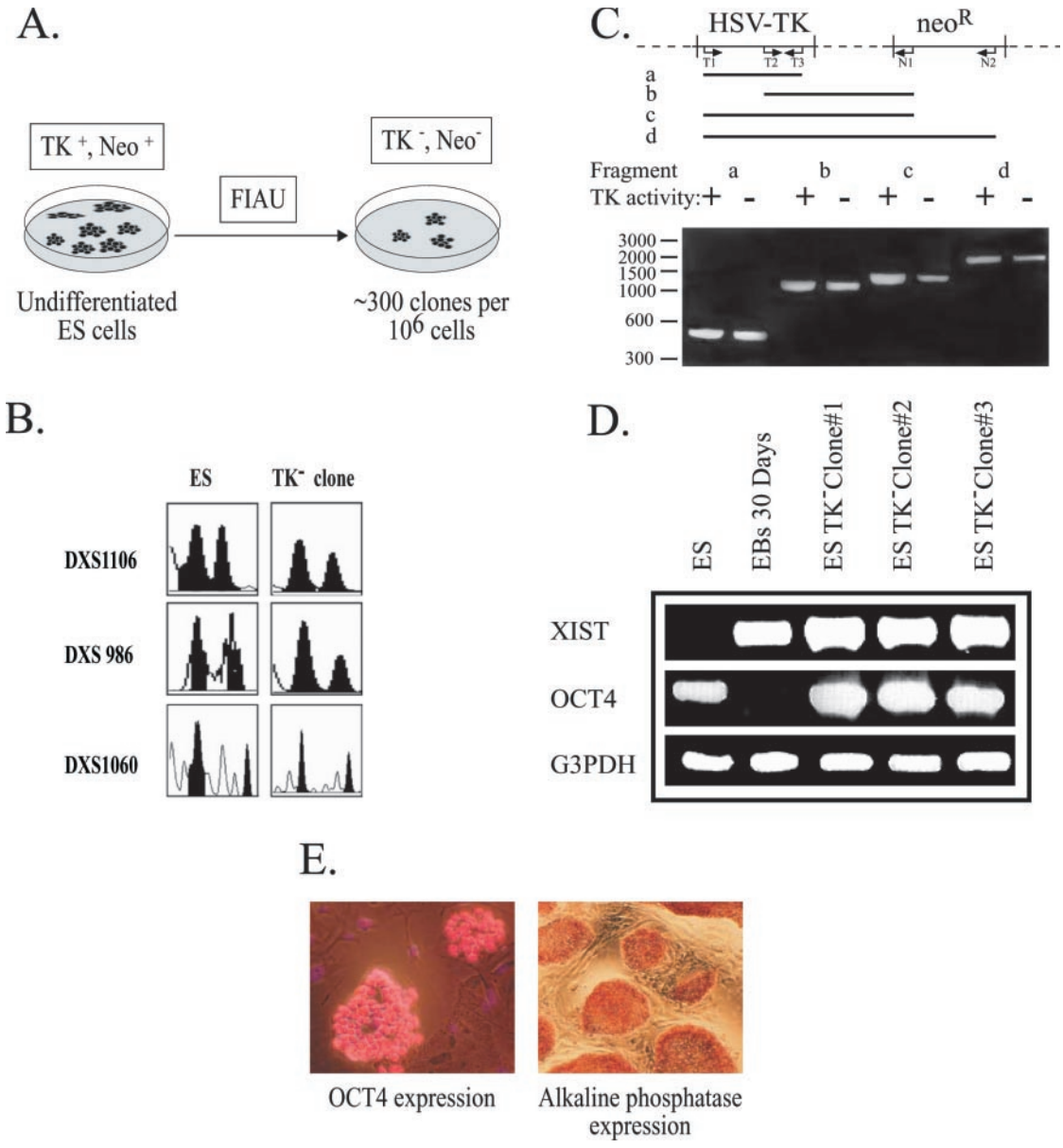

Figure 3. Isolation and characterization of transgene inactive clone by selection (**A**) Schematic diagram showing the strategy to isolate clones with inactivation of transgene on X chromosome. (**B**) Analysis of three representative polymorphic microsatellite markers on X chromosome (DX1106, DX986 and DX1060; see Materials and methods section). (**C**) PCR analysis of the HSV-TK and neo^R transgenes in the wild-type (+) and TK (−) clones. (**D**) RT–PCR analysis of expression of XIST, OCT4 and G3PDH (as control) in ES cells, 30 days EBs and in three transgene inactivated clones. (**E**) Oct4 antibody and alkaline phosphatase histochemical staining of transgene inactivated clones.

for gene conversion events (32). As the RET gene trap construct harbors a promoterless *GFP* gene, the expression pattern of the trapped gene can be easily monitored in living cells. In addition, this construct being a loxP-mediated removable exon trap, phenotypic reversion can be tested after provirus excision in the trap clones. Furthermore, since RET vector contains a negative selection marker, we may also induce deletions into specific regions in the genome and thus create a bank of chromosomal deletions. Since each clone could potentially yield deletions covering 1–20 cM (with an average of 10 cM) (19), deletions in our collection of mapped insertion clones may cover ~300 cM, or 10% of the human genome. In that case, the deletions would be primarily on single allele, our system may be optimal to study haploinsufficiency of the differentiated progenies of human ES cells, parental imprinting and X inactivation. We may also use it as a knock-in system to follow up the pattern of expression of specific genes and utilize it to isolate lineage-specific cells. Together with homologous recombination in human ES cells, this system should be the most important tool in studying human genetics *in vitro*. In conclusion, gene trap in human ES cells should be a useful tool not only in analyzing specific genes and in studying large genomic sequences but also in learning more about differentiation during development.

ACKNOWLEDGEMENTS

We are grateful to Dr Philip Leder (Harvard Medical School) for sharing with us the RET construct. We thank Dr Joseph Itskovitz-Eldor (Rambam Medical Center) for kindly providing us with the human ES cells as a collaborative research. S.K.D. is a Valazi-Pikovsky Fellow. This study was partially supported by funds from the Herbert Cohn Chair, by the Israel Science Foundation (grant no. 672/02-1), and by an NIH grant.

REFERENCES

1. Thomson,J.A., Itskovitz-Eldor,J., Shapiro,S.S., Waknitz,M.A., Swiergiel,J.J., Marshall,V.S. and Jones,J.M. (1998) Embryonic stem cell lines derived from human blastocysts. *Science*, **282**, 1145–1147.
2. Reubinoff,B.E., Pera,M.F., Fong,C.Y., Trounson,A. and Bongso,A. (2000) Embryonic stem cell lines from human blastocysts: somatic differentiation *in vitro*. *Nat. Biotechnol.*, **18**, 399–404.
3. Itskovitz-Eldor,J., Schuldiner,M., Karsenti,D., Eden,A., Yanuka,O., Amit,M., Soreq,H. and Benvenisty,N. (2000) Differentiation of human embryonic stem cells into embryoid bodies compromising the three embryonic germ layers. *Mol. Med.*, **6**, 88–95.
4. Schuldiner,M., Yanuka,O., Itskovitz-Eldor,J., Melton,D.A. and Benvenisty,N. (2000) Effects of eight growth factors on the differentiation of cells derived from human embryonic stem cells. *Proc. Natl Acad. Sci. USA*, **97**, 11307–11312.
5. Eiges,R., Schuldiner,M., Drukker,M., Yanuka,O., Itskovitz-Eldor,J. and Benvenisty,N. (2001) Establishment of human embryonic stem cell-transfected clones carrying a marker for undifferentiated cells. *Curr. Biol.*, **11**, 514–518.
6. Schuldiner,M., Itskovitz-Eldor,J. and Benvenisty,N. (2003) Selective ablation of human embryonic stem cells expressing a 'suicide' gene. *Stem Cells*, **21**, 257–265.
7. Henderson,J.K., Draper,J.S., Baillie,H.S., Fishel,S., Thomson,J.A., Moore,H. and Andrews,P.W. (2002) Preimplantation human embryos and embryonic stem cells show comparable expression of stage-specific embryonic antigens. *Stem Cells*, **20**, 329–337.
8. Evans,M.J. (1998) Gene trapping—a preface. *Dev. Dyn.*, **12**, 167–169.
9. Gossler,A., Joyner,A.L., Rossant,J. and Skarnes,W.C. (1989) Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes. *Science*, **244**, 463–465.
10. von Melchner,H. and Ruley,H.E. (1989) Identification of cellular promoters by using a retrovirus promoter trap. *J. Virol.*, **63**, 3227–3233.
11. Niwa,H., Araki,K., Kimura,S., Taniguchi,S., Wakasugi,S. and Yamamura,K. (1993) An efficient gene-trap method using poly A trap vectors and characterization of gene-trap events. *J. Biochem. (Tokyo)*, **113**, 343–349.
12. Salminen,M., Meyer,B.I. and Gruss,P. (1998) Efficient poly A trap approach allows the capture of genes specifically active in differentiated embryonic stem cells and in mouse embryos. *Dev. Dyn.*, **212**, 326–333.
13. Ishida,Y. and Leder,P. (1999) RET: a poly A-trap retrovirus vector for reversible disruption and expression monitoring of genes in living cells. *Nucleic Acids Res.*, **27**, e35.
14. Xu,R.H., Chen,X., Li,D.S., Li,R., Addicks,G.C., Glennon,C., Zwaka,T.P. and Thomson JA. (2002) BMP4 initiates human embryonic stem cell differentiation to trophoblast. *Nat. Biotechnol.*, **20**, 1261–64.
15. Berger,S.L. and Chirgwin,J.M. (1989) Isolation of RNA. *Methods Enzymol.*, **180**, 3–13.
16. Kent,W.J. (2002) BLAT—the BLAST-like alignment tool. *Genome Res.*, **12**, 656–664.
17. Benjamin,D., Bakel,I.V. and Craig I.W. (2000) A novel expression based approach for assessing the inactivation status of human X-linked genes. *Eur. J. Hum. Genet.*, **8**, 103–108.
18. Harlow,E. and Lane,D. (1999) *Using Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
19. Goodwin,N.C., Ishida,Y., Hartford,S., Wnek,C., Bergstrom,R.A., Leder,P. and Schimenti,J.C. (2001) DelBank: a mouse ES-cell resource for generating deletions. *Nature Genet.*, **28**, 310–311.
20. Frohman,M.A., Dush,M.K. and Martin,G.R. (1988) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl Acad. Sci. USA*, **85**, 8998–9002.
21. Dvash,T., Mayshar,Y., Darr,H., McElhaney,M., Barker,D., Yanuka,O., Kotkow,K.J., Rubin,L.L., Benvenisty,N. and Eiges,R. (2004) Temporal gene expression during differentiation of human embryonic stem cells and embryoid bodies. *Hum. Reprod.*, in press.
22. Kremer,E.J., Pritchard,M., Lynch,M., Yu,S., Holman,K., Baker,E., Warren,S.T., Schlessinger,D., Sutherland,G.R. and Richards,R.I. (1991) Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. *Science*, **252**,1711–1714.
23. Plath,K., Mlynarczyk-Evans,S., Nusinow,D.A. and Panning,B. (2002) Xist RNA and the mechanism of X chromosome inactivation. *Annu. Rev. Genet.*, **36**, 233–278.
24. Takagi,N. and Sasaki,M. (1975) Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature*, **256**, 640–642.
25. Ropers,B.J.,Wolff,G. and Hitzeroth,H.W. (1978) Preferential X inactivation in human placenta membranes:is the paternal X inactive in early embryonic development of female mammals? *Hum. Genet.*, **43**,265–272.
26. Migeon,B.R. and Do,T.T. (1979) In search of non-random X inactivation: studies of fetal membranes heterozygous for glucose-6-phosphate dehydrogenase. *Am. J. Hum. Genet.*, **31**, 581–585.
27. Zeng,S.M. and Yankowitz,J. (2003) X-inactivation patterns in human embryonic and extra-embryonic tissues. *Placenta*, **24**, 270–275.
28. Heard,E., Mongelard,F., Arnaud,D., Chureau,C., Vourc'h,C. and Avner,P. (1999) Human XIST yeast artificial chromosome transgenes show partial X inactivation center function in mouse embryonic stem cells. *Proc. Natl Acad. Sci. USA*, **96**, 6841–6846.
29. Migeon,B.R., Winter,H., Kazi,E., Chowdhury,A.K., Hughes,A., Haisley-Royster,C., Morrison,H. and Jeppesen,P. (2001) Low-copy-number human transgene is recognized as an X inactivation center in mouse ES cells, but fails to induce cis-inactivation in chimeric mice. *Genomics*, **71**, 156–162.
30. Disteche,C.M. (1995) Escape from X inactivation in human and mouse. *Trends Genet.*, **11**, 17–22.
31. Vasques,L.R., Klockner,M.N. and Pereira,L.V. (2002) X chromosome inactivation: how human are mice? *Cytogenet. Genome Res.*, **99**, 30–35.
32. Mortensen,R.M., Conner,D.A., Chao,S., Geisterfer- Lowrance,A.A. and Seidman,J.G. (1992) Production of homozygous mutant ES cells with a single targeting construct. *Mol. Cell. Biol.*, **12**, 2391–2395.