ARTICLE

A novel expression based approach for assessing the inactivation status of human X-linked genes

Don Benjamin[1], Inge Van Bakel[1] and Ian W Craig[2]

[1]Genetics Unit, University of Oxford, UK; [2]GDP Research Centre, Institute of Psychiatry, London, UK

We present a novel RT-PCR-based approach for determining the inactivation status of X-linked genes. Using cDNA from cloned female cell lines in which only the maternal or paternally derived X chromosome is active, we are able to demonstrate expression from only one allele in genes known to be inactivated. Following reverse transcription, amplification across a polymorphism will yield a product from a single allele if the gene of interest is inactivated, and products from both alleles in a gene escaping inactivation. We have verified this approach using the human androgen receptor and FMR1 loci which have been shown to be subjected to normal inactivation. The potential for widespread application of this approach was shown by the successful demonstration of inactivation at the MAOA and HPRT loci using intronic polymorphisms. *European Journal of Human Genetics* (2000) **8**, 103–108.

Introduction

Dosage compensation in mammalian females is achieved by the inactivation of one X chromosome early in development leading to the equality of X-linked gene products between male and female cells.[1] The inactive chromosome is not absolutely inert, this is demonstrated by the active expression of genes from the aptly named pseudoautosomal regions on the tips of the long and short arms. In addition, there are genes which are expressed from both the active and, at varying levels, the inactive chromosomes and are interspersed among other genes which are subject to regular dosage compensation.[2–4] At least 15 such genes have been described in humans, but of their mouse homologues, only the XE169 and UTX loci exhibit a similar escape from inactivation.[5,6] The majority of these genes are found in three main clusters on the short arm of the human X chromosome, at the pseudoautosomal boundary in Xp22.3, in a region centered around the UBE1 locus at Xp11.23, and around the DXS1272E locus at Xp11.21. In contrast, there are examples of genes such as the RPS4X and ZFX loci which escape

inactivation despite being closely flanked by inactive genes. Due to the paucity of current data, determination of the inactivation status of additional genes remains critical for the formulation and resolution of rival theories concerning the nature and spread of the inactivation process.

The most commonly employed means of determining the activity of an X-linked gene is via the construction of mouse–human somatic cell hybrids retaining only the inactive human X chromosome in a rodent background. As inactivation is effected at the level of transcription, detection of human specific transcripts by RT-PCR is suggestive of an escape from inactivation. This can then be confirmed by quantitation of the relevant transcripts from XO, XX, XXX or XXXX human cell lines.[7,8] We present an independent RT-PCR based approach using cDNA from cloned human cell lines. In a clonal female cell line, only one of either paternal or maternally derived X-chromosomes is active in every cell. Given a suitable polymorphism in a gene of interest, amplification of cDNA from clonal female cell lines across the polymorphism will amplify products from only one allele in a gene that is subject to inactivation. In the case of a gene escaping inactivation, however, both forms of the allele will be amplified with the ratio of their respective products indicating the degree of escape from inactivation. We have confirmed the validity of this approach by demonstrating the transcriptional inactivation of the human androgen receptor

(HUMARA), fragile X mental retardation 1 (FMR1), monoamine oxidase A (MAOA) and hypoxanthine phosphoribosyltransferase (HPRT) loci, all of which have been previously shown to be inactivated. Finally, we have applied the approach to the previously untested DXS977E locus and report preliminary evidence suggesting that transcription at this site may partially escape inactivation.

Materials and methods

Cell lines

Clonal cell lines were established from myoblasts obtained from muscle of consenting obligate carriers of DMD. Myoblast satellite cells were obtained from the biopsy sample by stirring in calcium and magnesium-free Hank's medium containing 0.25% trypsin (w/v) at 37°C for 30 min. Suspended cells were collected by decantation, concentrated by centrifugation and the entire harvest was plated to establish primary cultures. These were grown in Hams F10 containing BSA (0.5 g/L), fetuin (0.5 g/L), insulin (0.18 g/L), dexamethasone (0.9 mg/L), epidermal growth factor (10 mg/L), 15% (v/v) foetal calf serum and gentamycin (50 mg/L). At confluence, the cultures were cloned by serial dilution employing 96 well plates plated at the appropriate dilution and growth from a single cell confirmed by microscopic examination. The clonal lines were then expanded to 24-well plates and harvested at confluence for total RNA extraction. Genomic DNA from whole blood was also obtained from the same individuals and X-inactivation analysis using the previously described methylation-dependent X inactivation assays at the HUMARA[9] and MAOA[10] loci revealed non-skewed patterns of inactivation (D Benjamin, PhD thesis, University of Oxford, 1995). Cell lines were established from the following three individuals: ST (cell line Tu), CT (cell line Ti), Mo (cell lines 2C1, 2C5, 2E5, 3C8, 3C12, 3E4, 3G5, 12E5).

RNA extraction and cDNA synthesis

Total cellular RNA was prepared by the acid–phenol method.[11] cDNA synthesis was performed using the SuperScript Preamplification System (Gibco-BRL, Life Technologies, Inchinnon, UK) using random hexamers for first-strand synthesis.

PCR primers and conditions

Loci studied were the human androgen receptor (HUMARA), monoamine oxidase A (MAOA), fragile X mental retardation site (FMR1), hypoxanthine phosphoribosyl transferase (HPRT), and DXS977E which was identified as an EST[12] (Table 1). The following primers were employed: HUMARA primers P1 and P2,[9] MAOA primers 3-1 and 3-3,[10] FMR1 primers c and f,[13] HPRT primers a and b[14] and DXS977E primers 31A-L and 31A-R.[12] The annealing temperatures were 65°C, 54°C, 65°C, 63°C and 60°C respectively. Reaction conditions were 50 mM KCl, 10 mM TrisHCl (pH 8.0), 1.5 mM

Table 1 Loci and polymorphisms employed in this study. Primer sequences amplifying these polymorphisms are given in the Materials section. The location of the CA repeat in the DXS977E locus is not defined but is observed in the processed transcript[12]

Locus	Location	Polymorphism
HUMARA	Xq12	CAG_n (exon 1)
FMR1	Xq27.3	CGG_n (exon 1)
MAOA	Xp11.3	VNTR (23bp)$_n$, GT_n (intron 1)
HPRT	Xq26	$AGAT_n$ (intron 3)
DXS977E	Xp11.4	CA_n, (exon ?, 3′-UTR)

$MgCl_2$, 200 mM dNTPs, 1 mM primer and 2U Taq polymerase (Promega, Southampton, UK). For the FMR1 specific PCR, 10% (v/v) DMSO and a 3:1 mix of 7-deaza-dGTP:dGTP in place of dGTP was used. Thermal cycling parameters were: hot start at 94°C, 5 min, denaturation at 94°C, 1 min, annealing (1 min), extension at 72°C (30 s) for 32 cycles except for FMR1 which was performed for 41 cycles. HUMARA, FMR1 and DXS977E reactions were performed radioactively using one primer end-labelled with 32P-dATP and alleles were resolved on 6% denaturing polyacrylamide gels and visualised by autoradiography. Ratio of band intensities were calculated from densitometric scans using the IS-1000. Digital Imaging System. The MAOA and HPRT reactions were conducted non-radioactively and were resolved on 3% NuSieve: 1% normal agarose gels in 1X TBE (0.9 M Tris-Borate, 20 mM EDTA).

Results

Inactivation of the HUMARA and FMR1 loci

The polymorphisms used to discriminate between the different alleles at these loci are located within a single exon in these genes (Table 1). This feature may also prove to be a consideration in any future studies of a particular gene of interest. Therefore, it is vital to ensure the absence of any genomic DNA, as contamination will result in amplification from both alleles. Because both the *HUMARA* and *FMR1* genes have been previously shown to be inactive, we have used them to confirm the quality of the cDNA preparation and to demonstrate the absence of any genomic DNA. RT-PCR using HUMARA specific primers gave the following results (Figure 1a). All three individuals were polymorphic at his locus as seen by the two different alleles amplified from genomic DNA templates. However, only one of the two alleles was amplified from cDNA and from most of the putative clonal cell lines indicating complete inactivation of the gene at the transcriptional level, and the lack of genomic DNA in the cDNA preparation. In Ti, faint amplification of the lower (smaller) allele can be seen in the cDNA sample indicating trace amounts of genomic DNA. Essentially similar results were obtained for expression from the FMR1 locus in the informative cell lines from individuals Mo and Ti (Figure 1b). In sample 2C5, however, a faint product corresponding in size to the upper (larger) allele can be discerned.
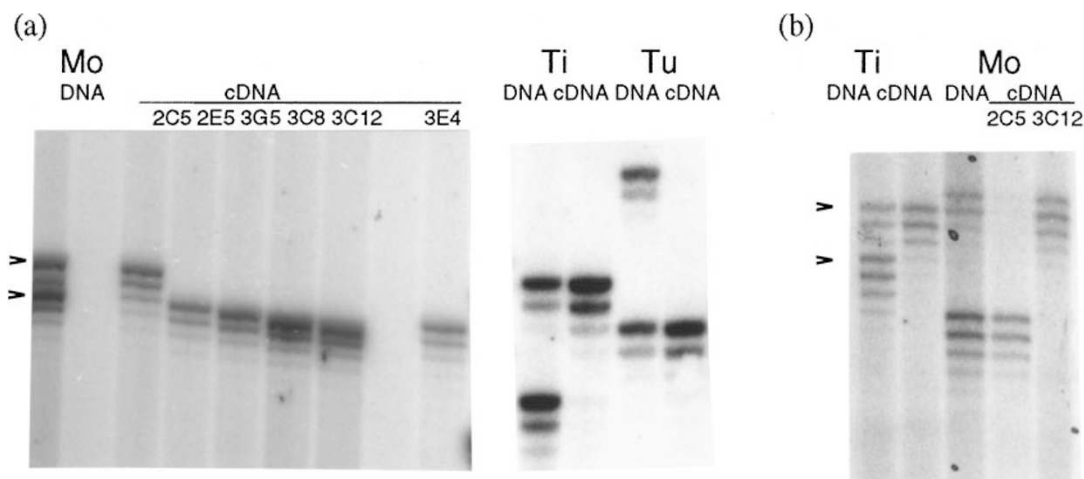
Figure 1 (**a**) RT-PCR primers for the HUMARA locus. (**b**) Similar analysis using FMR1 primers. Only Mo and Ti were polymorphic for FMR1. Amplification of only one allele from cDNA templates indicates X-inactivation at the level of transcription for both loci. A faint signal corresponding to the second signal can be observed in (**a**) Ti cDNA and (**b**) 2C5 cDNA (see text for comment). Arrows distinguish the main allele band from fainter shadow bands generated as a PCR artefact.

As only one allele is amplified from the other two samples, the signal presumably originates from trace amounts of genomic DNA.

Inactivation of the *MAOA* and *HPRT* genes

In practice, it may be rare to find a suitable polymorphism within the exon of a particular gene of interest. As a result, we attempted to demonstrate if a similar approach would be successful on intronic polymorphisms which, in contrast, are comparatively abundant. As cDNA synthesis was primed with random hexamers instead of polyA primers, reverse transcription of a sufficient number of precursor mRNA templates from total cellular RNA should, in principle, allow successful amplification from intronic sequences. The *MAOA* and *HPRT* genes were selected for study as they have been

shown to be inactivated and possess well characterised polymorphisms located in their introns. As seen in Figure 2, the inactivation of both the MAOA and HPRT loci could be demonstrated by this method. Success with this approach using the relatively low sensitivity approach of ethidium bromide visualisation on agarose gels required considerable care and the quality of the RNA preparations may be crucial. We observed that three out of six reactions for *MAOA* and four of seven for *HPRT* failed to give adequate amplification, even when repeated using twice the amount of template (not shown). The low signal intensities suggest that the assay may be working at the limit of detection for such polymorphisms. It should be noted that for the cell lines derived from the individual Mo, the results agree with that for the *HUMARA* and *FMR1* results (Figure 1) in that the same alleles are active
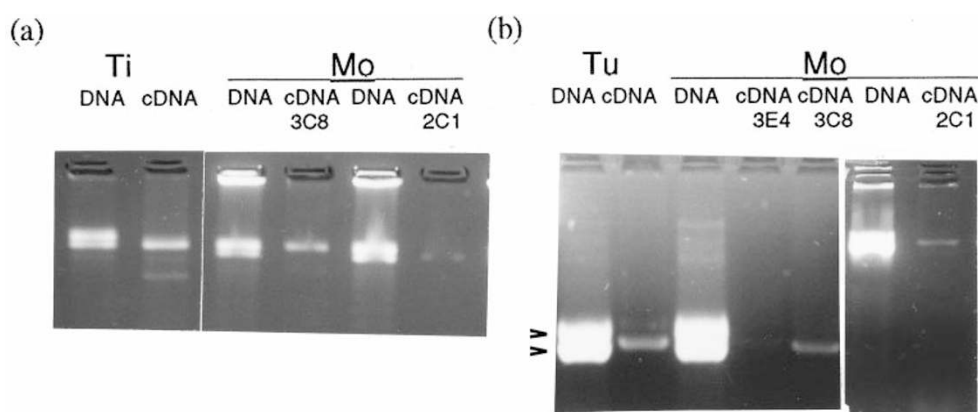

Figure 2 RT-PCR with (**a**) MAOA and (**b**) HPRT specific primers. Inactivation at both loci is demonstrated. cDNA from sample 3E4 in (**b**) is very weakly amplified. Arrows denote the position of both alleles amplified from genomic DNA. Weak amplification from cDNA suggests that the amount of nascent mRNA reverse-transcribed is low and is reflected in a high drop-out rate for these amplifications (see text).

for 3E4, 3C12, 3C8, 3G5, 2E5 and the other allele active in 2C5.

Determination of the activity status of the DXS977E locus

Only myoblasts from the individual Mo were found to be polymorphic at this locus. PCR amplification of cDNA from the various clones derived from the Mo myoblast lines resulted in significant products from both alleles (see Figure 3), suggesting that the locus escapes from inactivation. Although the result indicated that both alleles are transcribed, the signal from one is always weaker than that from the other, suggesting only a partial escape from inactivation. Low levels of DNA contamination could represent a possible explanation for the observed apparent transcription from both alleles; however, the correlation of the patterns observed for DXS977E allele expression in each of the clones is consistent with those deduced for the other X linked loci described in the previous sections. Analysis of densitometer scans of the autoradiographs suggested that its expression was at only about 30–40% of the presumptive, fully active allele.

Discussion

We have developed an RT-PCR based transcription assay to determine the activity of X-linked genes using cDNA from cloned cell lines. The primary requirement for the success of this approach is the existence of a polymorphism within the gene to be studied which thereby enables alleles from the active and inactive X chromosomes to be distinguished. Using primers specific for polymorphisms within the HUMARA and FMR1 loci, we have demonstrated the viability of this method. As these genes are known to be subject to

Figure 3 Results at the DXS977E locus. Only the individual Mo was polymorphic at this locus (**a**) initial results, (**b**) amplification was repeated and exposed for a longer period thereby unequivocally displaying amplification products from both alleles (arrows) in cell lines derived from Mo, indicating escape from inactivation. The ratio of the band intensities for the alleles suggests only a partial escape (expression at about 30–40% for the weaker allele).

inactivation, they also serve as preliminary tests to confirm the absence of any contaminating genomic DNA in the RNA preparation which will give rise to false positives. Because cloned myoblasts from primary cultures can only be passaged once or twice, this places a severe constraint on the number of cells that can ultimately be harvested for cDNA synthesis. As a consequence, we have omitted the standard negative control (reaction without reverse transcriptase) because the RT-PCR results from *HUMARA* and from *FMR1* represent a sufficient internal control for the detection of genomic DNA contamination. As a further control we attempted to amplify a portion of the *CD4OL* gene, which is not expressed in myoblasts from clonal myoblasts. No amplification was observed for any of the samples reported here (data not shown), indicating that any genomic DNA if present cannot support anything other than very weak amplification. As seen in the Results section, most of the cDNA samples gave only a single allele for HUMARA and FMR1 loci. The exceptions were Ti (HUMARA) and 2C5 (FMR1) where a faint product corresponding to the second allele was discernible. Using an alternative approach, Brown *et al*[15] have shown transcription from inactivated genes to be from 0.2 to 5% of that from the active allele. Hence the trace of DNA that may be present in some preparations achieved with the purification procedures employed, is below that required to detect significant escape from inactivation and is therefore acceptable for this type of analysis. Particularly, given that several clones are recommended for study and the evidence suggests that trace DNA contamination is a feature only of a minority. For this study, we have found that the unmodified acid phenol method for total RNA extraction[11] gave the best results. Any cDNA preparations which were shown to be significantly contaminated with genomic DNA by the above method were discarded.

Relatively few genes have easily recognized polymorphisms within the exons, therefore we also attempted to extend this approach to the intronic polymorphisms at the HPRT and MAOA loci as a test of the method's potential for wider application. Because the cDNA used for this study was reverse transcribed using random hexamers, intron-containing sequences from nascent RNA could reasonably be expected to be present as templates for PCR amplification. The results demonstrated the viability of this approach even with intronic sequences and confirm inactivation of expected patterns at these loci. Not all the samples, however, could be amplified, even by doubling the amount of template for amplification suggesting that the assay may be at the limits of its sensitivity in this particular formulation. The much higher sensitivity conferred by recent advances in fluorescent labelling and automated genotyping methodology may well provide the basis for extending the approach to any X-linked gene with intronic polymorphisms and whose unprocessed transcripts accumulate at a significant level.
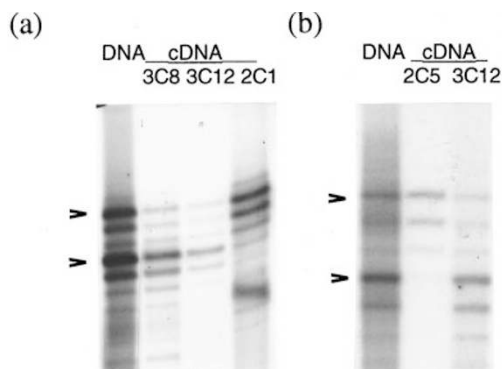
The amplification of both alleles of the DXS977E locus with these same cDNA preparations suggest that this locus may escape, albeit partially, X chromosome inactivation. The pattern of expression from the different cell lines (derived from the individual Mo) agrees with that previously determined at the HUMARA locus indicating that this is not an amplification artefact. The weaker amplified bands are also unlikely to be artefactual stutter bands which can be generated during amplification as they are larger than the main amplified band, clearly corresponding to one allele, in two samples (3C8 and 3C12). In contrast, stutter bands are invariably smaller than the major band (as in Figure 1).

It still may be that the more weakly amplified allele in each sample originates from trace amounts of genomic DNA. Against this interpretation, we have already shown the cDNA from 3C8 and 3C12 to be totally free of DNA in the HUMARA assay. Furthermore, the amount of template employed was less than that used for the MAOA and HPRT amplification which failed to amplify in some samples. We assume that this would not be the case if contaminating DNA were present in trace amounts as a template. Nevertheless, caution in the interpretation should be exercised and the result awaits further independent confirmation. DXS977E was originally isolated from a subtracted retinal cDNA library[16] and has been mapped to Xp11.4.[17] It is therefore of interest that two other genes from the same cytogenetic location have recently been shown to escape inactivation.[18,19] Unfortunately, these loci still await fine mapping to demonstrate if they are clustered or interspersed among normally inactivated genes. A similar approach to the one we describe, exploiting well characterised differences between laboratory strains, has been used in the mouse to demonstrate the inactivation of the *Rps4*, *Zfx* and *Ubel* genes.[20] The method outlined in our study represents a general approach, using clonal cell lines for the determination of inactivation in humans. Given the ubiquitous distribution of polymorphisms, the demonstration of the assays' applicability to intronic polymorphisms indicates that the technique can be extended to practically any X-linked locus provided a suitable polymorphism is available. While all the polymorphisms employed in this study were either mini or microsatellites, in principle any form of polymorphism can be used.

The most commonly employed method of determining the activity of X-linked genes at present is the detection of X chromosome specific transcripts from human–mouse somatic cell hybrids retaining an inactive human X chromosome as its only human component. This approach, however, suffers from the possible disadvantage that the inactivation status may be altered by prolonged propagation in tissue culture. An extensive, recent study of 33 X-linked genes revealed that a high proportion, perhaps as much as a quarter, may escape inactivation.[15] Another finding was the heterogeneity of expression of some genes which were only inactivated in a subset of cell lines. This phenomenon may reflect *in vivo* heterogeneity of inactivation between individ-

ual cells, of which very little is known. Despite the known instability of the inactive X chromosome in somatic cell hybrids,[21] the level of heterogeneity observed appears to be too high to be accounted for in this manner. The method developed in this study, based as it is on clonal cell lines, will provide a complementary approach towards resolving this issue as well as determining the inactivation status of more X-linked genes. Significantly, the inactivation status of ESTs chosen at random,[15,19] show a high percentage of escape from inactivation at levels similar to that which we have observed with the DXS977E locus. The actual number of genes escaping inactivation in the human X chromosome, at least to a partial extent, may therefore be substantial.

References

1 Lyon MF: Some milestones in the history of X chromosome inactivation. *Annu Rev Genet* 1992; **26**: 17–28.
2 Miller AP, Gustashaw K, Wolff DJ *et al*: Three genes that escape X chromosome inactivation are clustered within a 6Mb YAC contig and STS map in Xp11.21–p11.22. *Hum Mol Genet* 1995; **4**: 731–739.
3 Carrel L, Clemson CM, Dunn JM *et al*: X inactivation analysis and DNA methylation studies of the ubiquitin activating enzyme El and PCTAIRE-1 genes in human and mouse. *Hum Mol Genet* 1996; **5**: 391–401.
4 Disteche CM: Escape from X inactivation in humans and mouse. *Trends Genet* 1995; **11**: 17–22.
5 Wu J, Salido EC, Yen PH, Mohandas TK *et al*: The murine Xe169 gene escapes X inactivation like its human homologue. *Nat Genet* 1994; **7**: 491–496.
6 Greenfield A, Carrel L, Pennisi D *et al*: The UTX gene escapes X-inactivation in mice and humans. *Hum Mol Genet* 1998; **7**: 737–742.
7 Schneider-Gadicke A, Beer-Romero P, Brown LG, Nussbaum R, Page DC: The ZFX gene on the human X chromosome escapes X inactivation and is closely related to ZFY, the putative sex determinant on the Y chromosome. *Cell* 1989; **57**: 1247–1258.
8 Fisher EMC, Beer-Romero P, Brown LG *et al*: Homologous ribosomal protein genes on the human X and Y chromosomes: escape from X inactivation and possible implications for Turner Syndrome. *Cell* 1990; **63**: 1205–1218.
9 Allen RC, Zoghbi HY, Moseley AB, Rosenblatt HM, Belmont JW: Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen receptor gene correlates with X chromosome inactivation. *Am J Hum Genet* 1992; **51**: 1229–1239.
10 Hendriks RW, Chen Z-Y, Hinds H, Schuurman RKB, Craig IW: An X chromosome inactivation assay based on differential methylation of a CpG island coupled to a VNTR polymorphism at the 5' end of the monoamine oxidase A gene. *Hum Mol Genet* 1992; **1**: 87–94.
11 Chomczynski P, Sacchi N: Single step method of RNA isolation by acid guanidinium thiocyanate phenol-chloroform extraction. *Anal Biochem* 1987; **162**: 156–159.
12 Yan D, Wong D, Zheng K *et al*: Dinucleotide repeat polymorphism at the DXS977 locus. *Hum Mol Genet* 1994; **3**: 1030.
13 Fu YH, Kuhl DPA, Pizzuti A *et al*: Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* 1991; **67**: 1047–1058.

14 Edwards A, Hammond JA, Jin L, Caskey CT, Chakraborty R: Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* 1992; **12**: 241–253.

15 Brown CJ, Carrel L, Willard HF: Expression of genes from the human active and inactive X chromosomes. *Am J Hum Genet* 1997; **60**: 1333–1343.

16 Swaroop A, Xu J, Agarwal N, Weissman SM: A simple and efficient cDNA library subtraction procedure: isolation of human retina specific cDNA clones. *Nucleic Acids Res* 1991; **19**: 1954.

17 Nelson DL, Ballabio A, Cremers F, Monaco AP, Schlessinger D: Report of the Sixth International Workshop on X chromosome Mapping 1995. *Cytogenet Cell Genet* 1995; **71**: 307–342.

18 Jones MH, Furlong RA, Burkin H *et al*: The drosophila developmental gene fat facets has a human homologue in Xp11.4 which escapes X-inactivation and has related sequences on Yq11.2. *Hum Mol Genet* 1996; **5**: 1695–1701.

19 Esposito T, Gianfrancesco F, Ciccodicola A *et al*: Escape from X-inactivation of two new genes associated with DXS6974E and DXS7020E. *Genomics* 1997; **43**: 183–190.

20 Bressler SL, Lee KH, Adler DA, Chapman VM, Disteche CM: Maintenance of X inactivation of the Rps4, Zfx and Ube1 genes in a mouse in vitro system. *Somat Cell Molec Genet* 1993; **19**: 29–37.

21 Gartler SM, Goldman MA: Reactivation of inactive X linked genes. *Dev Genet* 1994; **15**: 504–514.