European Journal of Human Genetics (1998) 6, 552–562

http://www.stockton-press.co.uk/ejhg

ORIGINAL PAPER

Genetic mapping of X-linked loci involved in skewing of X chromosome inactivation in the human

Anna K Naumova[1], Leonard Olien[2], Lynne M Bird[3], Mark Smith[1], Andrei E Verner[1], Mark Leppert[4], Kenneth Morgan[2] and Carmen Sapienza[1,5]

[1]*Fels Institute for Cancer Research and Molecular Biology, Temple University School of Medicine, Philadelphia, USA*
[2]*Departments of Human Genetics and Medicine, McGill University, and Montreal General Hospital Research Institute, Montreal, Canada*
[3]*Division of Genetics, Children's Hospital and Health Center, San Diego, CA, USA*
[4]*Eccles Institute of Human Genetics, University of Utah, Salt Lake City, Utah, USA*
[5]*Department of Pathology and Laboratory Medicine, Temple University School of Medicine, Philadelphia, USA*

We have analyzed X-chromosome inactivation patterns in lymphocytes of 264 females from 38 families not known to have any genetic disease. Quantitative measures of X-inactivation showed strong sister–sister correlation in the degree of departure from equal numbers of cells having each X chromosome active, suggesting heritability of this phenotype. Strong sister–sister correlation was also observed for the fraction of cells having the same parent's X chromosome active, consistent with the possibility that this trait might be controlled by a *cis*-acting, X-linked gene. We used a sib-pair approach to determine whether X-inactivation phenotype was linked to loci in any region of the X chromosome. Both quantitative and discrete measures of X-inactivation phenotype showed evidence of linkage to markers in the region of the X inactivation center (*XIC*). The quantitative measure of X-inactivation phenotype used in our study also showed linkage to loci at Xq25–q26. This study provides the first evidence for X-linked inheritance of X chromosome inactivation phenotype derived from linkage analysis in phenotypically normal human families.

Keywords: skewed X-inactivation; mapping; sib-pair analysis

Correspondence: Carmen Sapienza, Fels Institute for Cancer Research and Molecular Biology, Temple University School of Medicine, 3307 N Broad Street, Philadelphia, PA 19140, USA. Tel: (215) 707-7373; Fax: (215) 707-1454; E-mail: CARMEN@SGI1.FELS.TEMPLE.EDU
Received 9 June 1998; revised 22 July 1998; accepted 5 August 1998

Introduction

One of the two X chromosomes in each somatic cell of female mammals is inactivated early in embryonic development.[1] The choice of which X chromosome to inactivate in any particular cell has a strong stochastic component but both epigenetic and genetic factors have been shown to influence the outcome of this process (see Belmont for review).[2] In both human[3–5] and mouse,[6,7] X chromosome inactivation is imprinted in early embryonic development such that the paternal X chromosome is inactivated in most cells of extra-embryonic lineages. In the mouse, an allelic series at the *Xce* locus has been demonstrated to affect the probability with which a particular X chromosome is inactivated in somatic cells of the embryo proper.[8] In contrast, the choice of which X chromosome is inactivated in the human is often said to be 'random'; meaning that on average, the maternal X chromosome is active in one half of the cells of an individual female and the paternal X chromosome is active in the other half.[1]

Individual females who deviate strongly from a 'random' pattern have been described, and familial clustering of females with highly skewed patterns of X-inactivation has also been observed (see reviews by Belmont[2] and Willard[9]). Plenge *et al*[10] used a candidate gene approach to identify two families in which skewed X-inactivation phenotype co-segregated with a specific allele at the *XIST* locus, which resides within the cytogenetic region defined as the X inactivation center,[11] or *XIC*. However, skewed X-inactivation phenotype cannot be demonstrated to map to the *XIC* in all families or individuals that show the trait,[12–16] despite evidence for the involvement of an X-linked gene in some of these instances.[16,17] In these cases, one cannot distinguish between models that invoke genetic control of a step in the X chromosome inactivation pathway and models that invoke selection for cells that have a particular X chromosome in the active or inactive state.[18,19]

Virtually all reports demonstrating 'non-random' X-inactivation presuppose that the cases described are likely to be exceptional and there are few indications that normal variation in X chromosome inactivation is heritable. However, failure to observe inheritance of this trait cannot be taken as strong evidence that it is not heritable because many factors might mask our ability to make this observation. Selection of the genetic model, unknown penetrance of the trait, the potential presence of phenocopies within families and the possibility of genetic heterogeneity[11,17,20–22] are all likely to be complicating factors. We have attempted to circumvent some of these difficulties by using a sib-pair approach to examine the heritability of X-inactivation phenotype in 264 females from 38 families and to search for linkage to loci on the X chromosome. Our analyses, which are largely independent of mode of inheritance, indicate that X chromosome inactivation phenotype is heritable and is linked to loci within or near the *XIC* at Xq13–q21 and to loci at Xq25–q26.

Materials and Methods

Quantitation of X-inactivation Skewing

The X-inactivation phenotype of individual females was scored using the PCR-based assay[23] for differential methylation of alleles at the androgen receptor locus (*AR*), as used in our previous studies.[17,24] Determination of X-inactivation by the *FRAXA* methylation PCR assay[25] was done, in addition, on one family, K1362, using primers 'c' – 5' GCTCAGCTCCGTTTCGGTTTCACTTCCGGT 3' and 'f' – 5' AGCCCCGCACTTCCACCACCAGCTCCTCCA 3'. Quantitation of the results of the *FRAXA* methylation assay was performed by laser-scanning densitometry or phosphorimaging analysis, as described for the *AR* methylation assay.[17,24]

We derived two measures of X-chromosome inactivation skewing:

1 The active proportion of the maternally inherited chromosome, denoted PA_{mat}

$$PA_{mat} = [m/(m + s)]^{Hha+}$$

where m is the densitometer quantitation or phosphorimage count corresponding to the paternal allele, and s is the value for the maternal allele, of the androgen receptor PCR product after the template has been digested with *Hha*I. For certain individuals whose *AR* alleles differed in length by only one trinucleotide repeat, the PA_{mat} score was adjusted by $[m/(m+s)]^{Hha-}$, as in Naumova *et al*.[24] The PA_{mat} score measures both the magnitude and direction of skewing. A PA_{mat} score of 0.5 indicates no skewing, a PA_{mat} score of zero indicates that all cells have an active paternal chromosome, whilst a PA_{mat} score of 1 indicates that all cells have an active maternal chromosome.

2 The degree of skewing score, denoted *DS*, is a measure of only the magnitude of skewing.[24] Zero indicates no skewing and 0.5 indicates complete skewing; that is, the active chromosome in every cell is of the same parental origin.

$$DS = |(PA_{mat} - 0.5)|$$

Correlation of Trait Values

We tested for sister–sister correlation and for mother–daughter correlation in both the *DS* score and the PA_{mat} score. The sister–sister correlation was estimated from sibships with at least two sisters using analysis of variance.[26] Significance was assessed using the permutation method[27]

because sibship sizes were small, and the trait is not normally distributed (see Figure 1 in Naumova *et al*[17]). Permutation replicates were generated by assigning individual females to sibships randomly, without replacement, from the pool of all females in the third generation. The distribution of the test statistic under the null hypothesis was found by calculating the test statistic for each of 2000 replicates. The empirical significance level (or achieved significance level,[27]) was defined as the proportion of permutation replicates with sister–sister correlation greater than the value calculated from the actual sample.

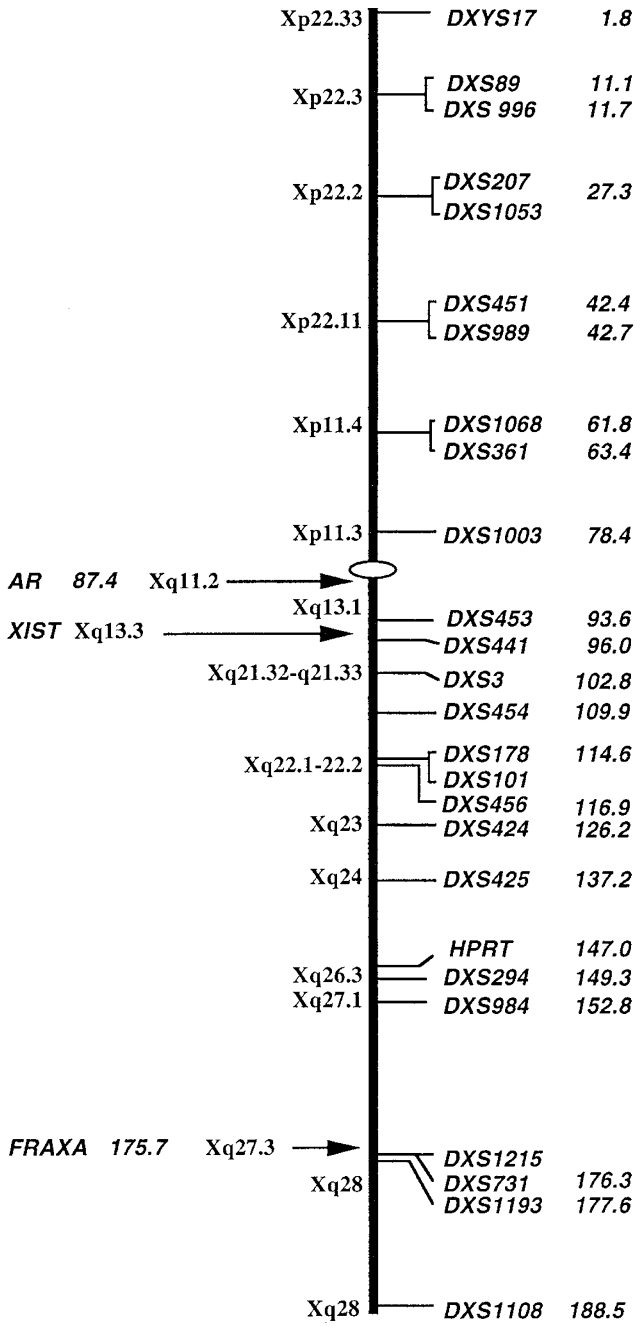

Figure 1 *Genetic map of the human X chromosome. Distance between markers and marker order, with the exception of* DXS101, *are taken from the linear version of the female X chromosome map*[31] *(see GenLink public database [http://www.genlink.wustl.edu]). Distances are displayed in Kosambi centiMorgans.* DXS101 *has been placed distal to* DXS178 *according to Vetrie et al.*[42]

Analysis of the mother–daughter correlation also utilized the permutation method. In this case, a permutation replicate was generated by randomly assigning mothers to families, sampling without replacement, from the pool of all mothers. In this way, the distribution of the mother–daughter correlation is conditioned on the correlation structure of the daughters. We used two estimators of the mother–daughter correlation: the pairwise estimator and the sib-mean estimator.[26,28] The pairwise estimator matched each daughter's score with the mother's score, whilst the sib-mean estimator matches the mean of a sibship with the score of the mother. The pairwise estimator performs well for low sibling correlation but gives greater weight to larger families.[26] The sib-mean estimator performs well for higher sibling correlation but has larger asymptotic standard error than the pairwise estimator.[28]

Linkage Analysis Using Sib-pairs

We used the Haseman-Elston sib-pair approach[29] to test for linkage of the PA_{mat} score to 15 markers on the X-chromosome. Define $Y_j = (y_{j1} - y_{j2})^2$, where y_{ji} is the PA_{mat} score of the ith member of the *j*th pair of sibs. In the standard approach, Y_j is regressed on p_j, the proportion of alleles at a given locus shared identical-by-descent (IBD) by the jth pair of sibs. In the present study, for each marker, individuals were included for whom the genotype was known or could be assigned with 95% probability (see below), therefore p_j was estimated by counting. Furthermore, p_j will equal either 1/2 or 1 for a pair of female sibs since they will have inherited either the same or different maternal alleles but will always inherit the same paternal allele. Therefore Haseman-Elston analysis is equivalent, in this case, to performing a test on $b = m_2 - m_1$, where m_k is the mean of the Y_j's for those sib-pairs that share *k* alleles IBD. Under the hypothesis of no linkage, the two groups will have identical means ($b = 0$). However, if the trait is linked to a particular marker, then *b* will be negative; that is, the alternative hypothesis is $b < 0$.

Significance of linkage to a given X chromosome locus was assessed using the parametric bootstrap method[27] on the maternally inherited allele for each daughter. Under the null hypothesis of no linkage, each daughter receives the grandpaternal allele from her mother with probability 1/2. We generated bootstrap replicates by randomly assigning, to each daughter, the grandparental origin of the maternally inherited allele. Two thousand replicates were generated and *b* was calculated for each replicate, which provided the distribution of *b* under the null hypothesis conditional on the sample structure. The empirical significance level was defined as the proportion of bootstrap replicates with *b* less than that of the actual sample. Because we bootstrapped marker alleles, the observed distribution of sibship sizes and sister–sister correlation of the trait were maintained for every replicate. Therefore, it was not necessary to adjust the statistic for non-independence of the sib-pairs since all replicates were generated under the same structure as the observed sample (see Hauser *et al*[30]).

Fifteen loci were chosen along the entire X-chromosome so that, where possible, adjacent markers were separated by not more than 15 cM. Where data were not available for a particular marker in a given individual, nearby markers were typed or retrieved from the GenLink database[31] (http://www.genlink.wustl.edu). The genotype at the marker of interest was then inferred according to the following criteria. The probability that the missing allele was of grandpaternal origin was calculated using flanker markers and joint recombination probabilities.[32] The missing allele was assigned to be of grandpaternal origin if the probability was greater than 0.95, whereas the allele was assigned to be of grandmaternal origin if the probability was less than 0.05. The data used in the linkage analysis comprised 15 markers and 76 individuals. Of the 1140 observations, 707 were typed by our lab or retrieved from the database, 328 were inferred as above, and 105 observations were uninformative for our analysis.

Permutation Test for Familial Aggregation of Skewing as a Discrete Trait

To test for familial aggregation of skewing, a contingency table of skewing status (the quantitative phenotype was transformed to the discrete trait 'skewed' or 'nonskewed' by designating individuals with at least 80% of their cells having the same X chromosome active as 'skewed') versus family was constructed using third-generation females. All families in which the skewing score was available for at least two sisters were included. Significance was estimated using the permutation method.[27] In this case, the null hypothesis is that there is no familial aggregation of skewing, that is skewing is independent of family. The test statistic is the probability of the contingency table, as calculated from the hypergeometric distribution.[33] Two thousand permutation replicates were generated at each threshold value. Each replicate was generated by randomly assigning individuals to families, while maintaining the distribution of sibship sizes. The empirical significance level is defined as the proportion of random permutation tables with probabilities less than or equal to the probability of the observed table. An empirical significance level of 0.05 was used to reject the null hypothesis of no familial aggregation of skewing.

Subjects Used in the Sibling Concordance Study

Five of the 38 families used in our initial screening were selected for the sibling concordance analysis on the basis of having two or more siblings who exhibited at least 80% of cells with the active X chromosome inherited from the same parent. Four of these five families came from the extended CEPH families collection and one family was collected by our laboratory. DNA from lymphocytes was used for X-inactivation assays, whereas DNA from lymphocytes and/or lymphoblastoid cell lines was used for determining the genotypes of the individuals at the X-chromosome loci tested (see below).

In our discrete trait mapping study, we considered the genotypes of only skewed females. Our rationale for this decision is that there are a number of ways, both genetic and non-genetic, in which a daughter who inherited an X-linked gene for skewing might *fail* to exhibit the skewing phenotype (inheritance of alleles at other loci that might counteract skewing, individual variation in the allocation of hematopoietic stem cells with one or the other X chromosome active, environmentally mediated counter selection, truncation selection imposed by transformation of the quantitative phenotype to a qualitative phenotype, etc.). However, given the results of our analyses of familial correlation and aggregation of skewing (Results and unpublished data), and our previous analysis of the frequency of skewing in unrelated females,[17] we regarded it as unlikely that skewing in multiple females within the same family as a result of non-genetic factors would be a serious concern.

Genotype Determination

Genotypes of all subjects at loci on the X chromosome were determined by PCR at markers spanning the length of the X chromosome. In all cases, we attempted to keep the distance between informative markers to less than 15 cM. Oligonucleotide PCR primers for the loci shown in Figure 1 were purchased from Research Genetics, Inc. (Huntsville, AL, USA), with the exception that primers for the *AR* locus were kindly provided by Dr Kim DeRiel, and primers for *DXS294*, *DXS984*, and *FRAXA* were purchased from Genosys Biotechnologies (The Woodlands, TX, USA). In cases where individuals were not informative for the markers listed above, we used data for closely linked markers from the public database, where it was available. Thus, the data for markers *DXS89*, *DXS342*, *DXS451*, *DXS1048*, *DXS319*, *DXS361*, *DXS453*, *DXS456*, *DXS425*, *HPRT*, and some of the data for *DXS989* and *DXS1108* for families K1332, K1341, K1347, K1375 and K1362 were taken from the GenLink database [http://www.genlink.wustl.edu, data release January 1996[31]]. The genotypes determined in our laboratory were compared, where possible, with those from the public database and no discrepancies were found.

Results

Skewed X-inactivation as a Quantitative Trait

We have used allele-specific methylation of CpG sites within the androgen receptor (*AR*) locus as a quantitative measure of X chromosome inactivation[23] (Materials and Methods). Two quantitative measures of X-inactivation were derived from these measurements. The 'degree of skewing' score (*DS*), measures only the absolute magnitude of 'skewing' and the PA_{mat} score takes into account both the magnitude and the direction (proportion of cells with the maternal X chromosome active) of skewing. The *DS* score is expected to be responsive to any genetic factor that affects skewing *per se*, whilst the PA_{mat} score is expected to be responsive to any X-linked gene that acts in *cis*, or any gene that exerts a parent-of-origin effect on X-inactivation.

All tests for sister–sister and mother–daughter correlations in *DS* score and PA_{mat} score were performed

Table 1 Sister–sister correlation in *DS* and PA_{mat} scores

Statistic	PA_{mat}	*DS*
Number of sisters	98	100
Number of sibships	26	27
Correlation	0.3791	0.2462
P-value[a]	0.0000	0.0025

[a]*P*-value determined by permutation test using 2000 replicates.

independently. Each test used 2000 random permutation replicates.

Sister–sister Correlation in X-inactivation Phenotype

We tested the hypothesis that there was no sister–sister correlation in the *DS* score or the PA_{mat} in 38 families (Table 1). For the *DS* score, only five of the 2000 permutation replicates (empirical significance level = 0.0025) had a sister–sister correlation greater than that of the observed families. For the PA_{mat} score, none of the permutation replicates had a sister–sister correlation greater than that found between sisters in the observed families. Thus, we reject the null hypothesis of no sister–sister correlation for both the *DS* score and the PA_{mat} score. The correlation in *DS* score indicates that the X-inactivation phenotype measured in siblings is not random, but that sisters are more similar to each other than are unrelated females. Because the PA_{mat} score also takes into account the parental origin of skewing, *i.e.* whether sisters have the same proportion of cells with the maternal X chromosome active, these data also suggest that if a genetic factor is influencing X-inactivation phenotype, that factor may lie on the X chromosome and act in *cis*.

Lack of Mother–daughter Correlation in X-inactivation Phenotype

We also tested for mother–daughter correlation in the *DS* and PA_{mat} (Table 2). None of the *P* values were less than 0.05, therefore we do not reject the null hypothesis of no mother–daughter correlation in *DS* or PA_{mat} score. These data indicate that, overall, mothers and daughters within the same family are no more similar in their X-inactivation phenotype than random permutations of mothers among the families. The interpretation of this observation with regard to genetic control of X-inactivation phenotype is, in large part, dependent on the genetic model by which X-inactivation phenotype is assumed to be inherited and expressed (see Discussion), but these data indicate that X-inactivation phenotype is not the result of a simple, X-linked dominant or autosomal dominant mode of inheritance.

Table 2 Mother–daughter correlation in *DS* and PA_{mat} scores

Statistic	PA_{mat}	*DS*
Number of daughters	75	82
Number of mothers	21	22
Pairwise estimator:		
Correlation	0.0720	–0.0599
P-value[a]	0.3800	0.6305
Sib-mean estimator:		
Correlation	0.0364	–0.0261
P-value[a]	0.4540	0.5490

[a]*P*-value determined by permutation test using 2000 replicates.

Mapping X-linked Genes that Influence X-inactivation by Haseman-Elston Sib-pair Analysis

Because of our analysis of the PA_{mat} score (Table 1) indicated that an X-linked gene might be involved in controlling this trait in *cis*, we modified the Haseman-Elston sib-pair approach of linkage analysis for this particular trait[29] (Materials and Methods). We determined the genotypes of each female at the X chromosome loci shown in Figure 1.[31] The number of sib-pairs, *b* statistic and empirical significance level are shown in Table 3. Four markers are linked to X-inactivation phenotype at a significance level of 0.05 or less. Two, *DXS453* and *DXS441* are closely linked to each other (approximately 3 cM apart,[31] see Figure 1) and are located at Xq13 in the region of the *XIC*. Two additional markers that are linked to X-inactivation

Table 3 Linkage analysis of X-inactivation skewing

Marker	*N*	*b*	*P*-value
DXS996	94	0.032	0.894
DXS207	104	0.003	0.558
DXS989	86	0.0076	0.627
DXS1068	100	–0.016	0.237
DXS1003	100	–0.0329	0.103
AR	120	–0.0299	0.0855
DXS453	116	–0.049	0.0175
DXS441	112	–0.0519	0.0195
DXS3	112	–0.0214	0.1715
DXS101	99	–0.0066	0.3585
DXS424	112	–0.0068	0.359
DXS425	101	–0.0466	0.035
DXS294	106	–0.06	0.0075
DXS731	111	–0.0387	0.0505
DXS1108	83	0.0343	0.881

N is the number of sib-pairs; $b=m_2-m_1$, where m_k is the mean of the squared difference in trait values of pairs of sibs that share *k* alleles IBD at a given locus (see Materials and Methods); *P*-value is the empirical significance level using the parametric bootstrap and 2000 replicates generated under the null hypothesis.

phenotype, *DXS425* and *DXS294*, are also linked to each other (12 cM apart,[3] see Figure 1) and are located at Xq25–q26. If we adjust the suggested guidelines for reporting the significance of linkage results proposed by Lander and Kruglyak[35] for the fact that we tested only loci on the X chromosome (the X chromosome contributes approximately 200 cM of the some 3300 cM that make up the human genome), both regions are 'suggestive' of linkage.

Skewed X-inactivation as a Discrete Trait

Familial Aggregation of 80% Skewing

In most studies seeking a genetic basis for variation of X chromosome inactivation in the human, the phenotype has been transformed from an inherently quantitative trait to a discrete trait by specifying a threshold value. Individuals having at least 80% of their cells with the same X chromosome active have been described as having non-random or skewed X-inactivation in a number of studies.[10,17,22,24,34] Although there is an important historical precedent for choosing the 80% threshold to designate an individual as having non-random X-inactivation ([7] and see discussion in Naumova *et al*[17]) it has not, to our knowledge, been demonstrated that individuals who are skewed to this level constitute a distinct phenotypic group. As one measure of this possibility, we tested whether females who were skewed to this degree or greater were distributed randomly among the 38 families or were clustered within families. Only 25 of the 2000 permutation replicates had a lower probability than the probability for the contingency table of observed families (empirical significance level = 0.0125). These results provide evidence for familial aggregation of skewed X-inactivation, as a discrete trait, when a threshold value of at least 80% of cells with the same X chromosome active is used.

Mapping of a Locus Involved in a Discrete Phenotype of Skewed X Chromosome Inactivation

We selected families containing two or more sisters skewed in favor of the same X chromosome active for use in an 'affected' sibling mapping study. Of the 38 three-generation families examined, only five contained two or more sisters who were concordantly skewed to the 80% level or greater (Figure 2). The skewed sisters in these five families were used to map an X-linked locus that may be responsible for the skewing phenotype. At each locus for which the sisters were informative, they were scored as concordant if they all inherited the same allele from their mother and were scored as discordant if *any* of them inherited a different allele. In our mapping strategy, X chromosome regions at which skewed sisters in all five families are concordant may contain genes involved in the skewed X-inactivation phenotype, whilst any locus that is discordant in one or more families is excluded. Because there may be females who are skewed for non-genetic reasons or because the trait may be genetically heterogeneous, this mapping strategy is likely to be conservative, in that loci are increasingly likely to be excluded as the number of skewed individuals considered becomes larger.

The regions of the X chromosome for which skewed females in each family may be concordant, assuming no recombination between consecutive concordant loci, is shown in Figure 3. Only a small region of the X chromosome is not excluded by one or more of the families. This region is limited on the proximal side by *DXS441*, within the *XIC* at Xq13, and on the distal side by *DXS3*, at Xq21. This interval is denoted as 'region of concordance' in Figure 3.

Discussion

We have examined X-chromosome inactivation phenotype for evidence of heritability in 38 'normal' families. We found a significant correlation between sisters in the degree to which they deviated from having equal proportions of lymphocytes with each X chromosome active, consistent with the possibility that this quantitative trait has a heritable component. Further evidence in favor of a heritable component of X-inactivation phenotype was obtained by comparing sisters for the fraction of their cells having an active maternal X chromosome. A significant sister–sister correlation was found, consistent not only with the heritability of this trait, but also with the possibility that this phenotype is controlled by a *cis*-acting, X-linked gene.

Although we found evidence for sister–sister correlation in the quantitative trait (Table 1), we found no evidence of mother–daughter correlation in either the *DS* score or the PA_{mat} score (Table 2). These results suggest that X-inactivation might be inherited in an unusual fashion. For this reason, linkage analysis was performed using a non-parametric method. Haseman-Elston sib-pair analysis provided evidence for linkage

of the X-inactivation trait to loci in two regions of the X chromosome, Xq13 and Xq25–q26 (Table 3).

Because there is considerable debate[35–37] over the appropriate nominal *P*-value required to report linkage, it is worth considering the possibility that one or both of the regions showing linkage to X-inactivation phenotype are false positive results due only to chance. In this regard, we may offer the following three counterpoints:

1 The sib-pair approach is robust. The method does not specify either a genetic model or a model for the distribution of the quantitative trait that is conditional on genotype, and the bootstrap method does not require assumptions on the distribution of the test statistic;

2 loci in the vicinity of both regions have been identified previously in conjunction with X-chromosome inactivation;[10,11,13–16,22,38,39]

3 we have not conducted a dense scan of the X-chromosome, nor have we tested for linkage at each point along the chromosome.

An additional consideration in interpreting these results is that, strictly speaking, we cannot distinguish whether any linkage identified reflects the activity of genes that might be involved directly in the process of X chromosome inactivation or whether some alleles at these loci might confer a proliferative advantage or disadvantage on the cells in which they are found (see reviews[2,9,19]). However, we note that one of the chromosome regions we have identified as showing

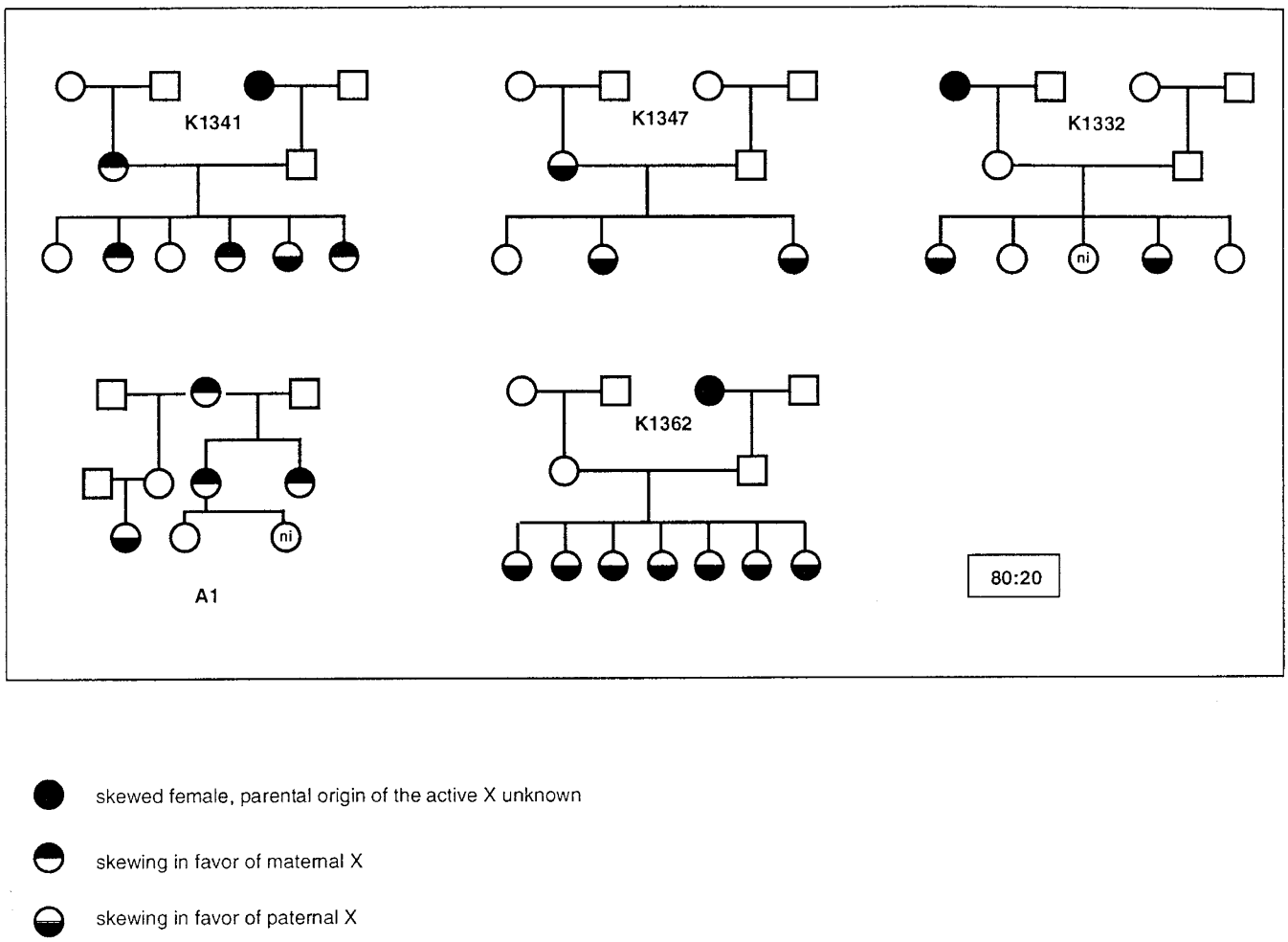

Figure 2 *Families selected for linkage analysis. Only the females of each family are shown in the bottom generation. Open circles – 'not skewed' females (see text), half-filled circles – 'skewed' females, ni – not informative for the* AR *locus or DNA not available.*

evidence for linkage to the quantitative trait (Xq13) contains the cytogenetically defined *XIC*,[38,39] as well as the *XIST* gene.[11].

We also transformed the X-inactivation phenotype from a quantitative trait to a discrete trait by designating those females with the same X chromosome active in at least 80% or their cells as skewed and selecting families for analysis on the basis of having two or more sisters skewed in favor of the same X chromosome active. Five of the 38 families fit these criteria and we used these families in a sibling concordance, exclusion mapping study. The only region of the X chromosome

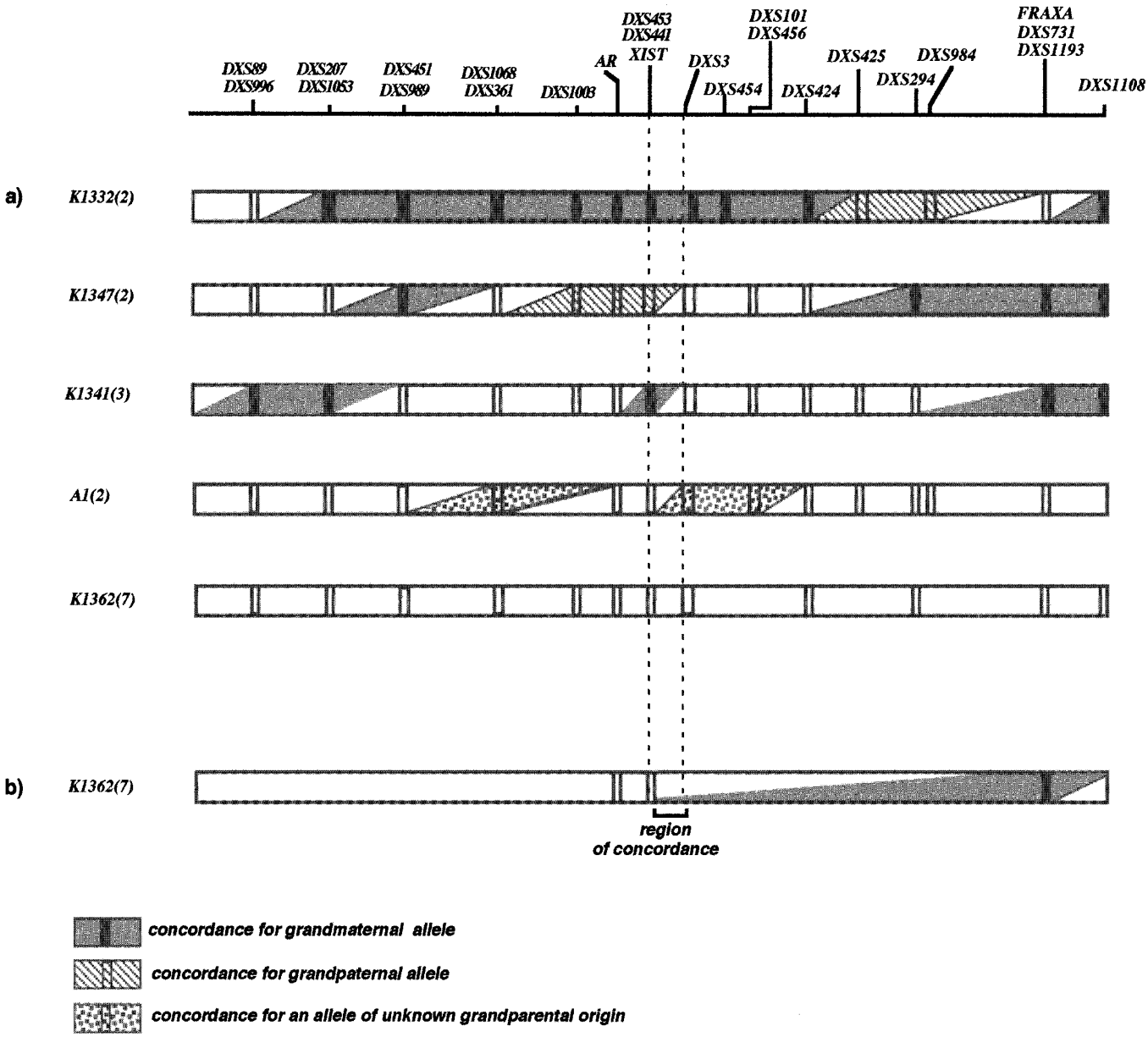

Figure 3 *Summarized haplotypes of skewed siblings. Family codes and number of informative siblings included in the analysis are displayed on the left, with number of informative siblings in parentheses. Each rectangle corresponds to a region of the X chromosome in which one or more loci were scored. At least one marker of those mapping to the corresponding region (Figure 1) was typed in each family. No rectangle appears when the family was not informative or no DNA was available for additional analysis. No information on the grandparental origin of the alleles was available for family A1. a) represents the maternal X chromosome (there is no locus at which all seven daughters in K1362 are concordant). b) shows the extent of the region for which all seven granddaughters in K1362 may be concordant for the X chromosome that was preferentially active in their paternal grandmother. The grandmother and granddaughters are concordant at* FRAXA *and are discordant at* AR *and* XIST *(see also Naumova et al[17]).*

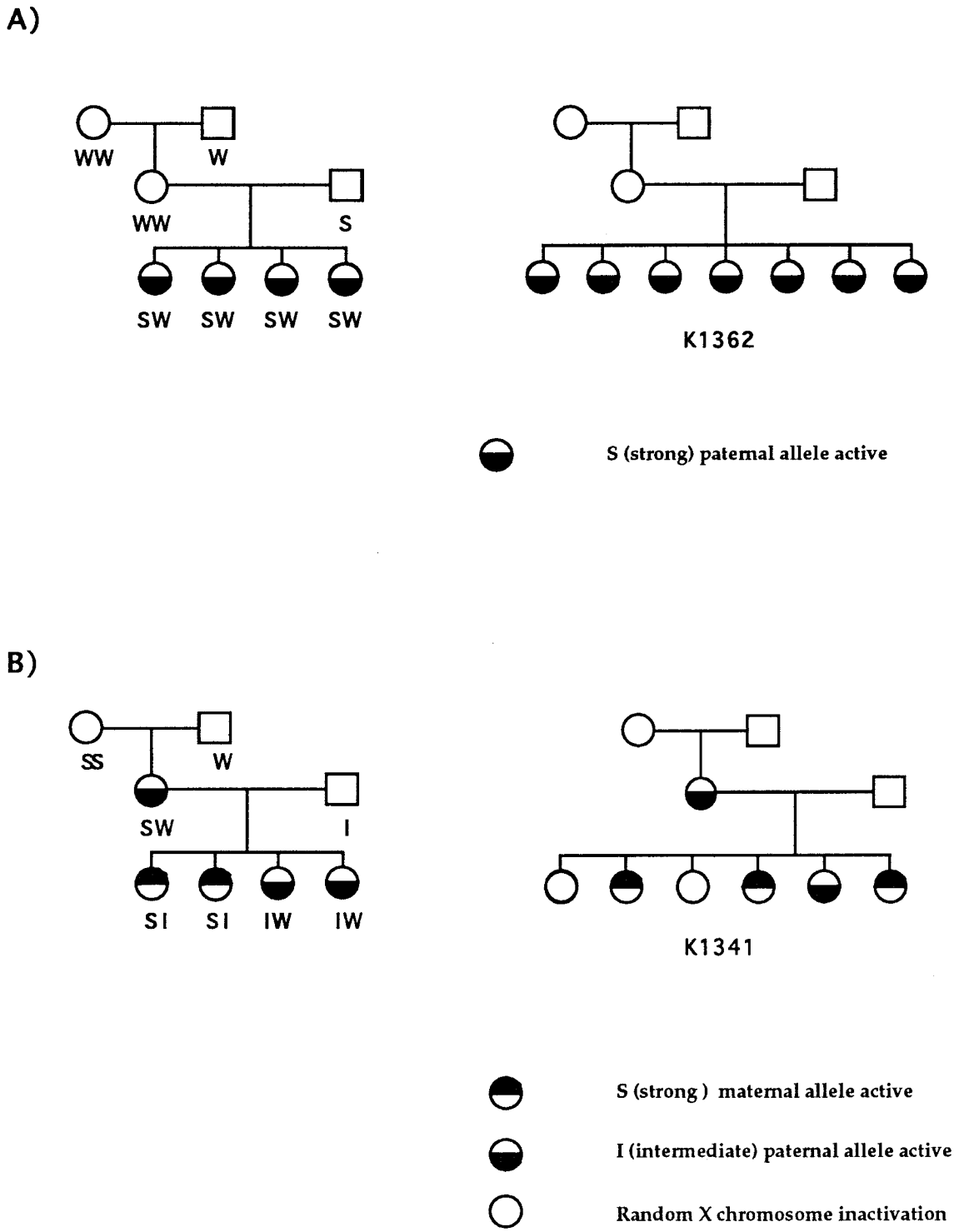

Figure 4 *Skewed X chromosome inactivation in human females as a result of interaction of* Xce*-like alleles of different strength. A) Model pedigree (left) with a mother homozygous for a weak allele and a father bearing a strong allele. All the daughters will have a skewed X-inactivation phenotype with their paternal X chromosome active. K1362 may be an example of such a family. B) Model pedigree with three* Xce*-like alleles of different strength. All the daughters are predicted to have a skewed X-inactivation phenotype, but some of the daughters will have the maternal X chromosome active, and some daughters will have the paternal X chromosome active. K1341 may be an example of such a family.*

not excluded by this analysis is defined on the proximal side by the marker *DXS441* (at Xq13) and on the distal side (at Xq21) by the marker *DXS3*. We note that *DXS441* is slightly distal to the *XIST* locus[40] but is within the cytogenetic region defined as the *XIC*. *DXS441* is also one of the loci that exhibited linkage to the quantitative trait in the Haseman-Elston sib-pair analysis (Table 3).

In terms of postulating a mode of inheritance for the discrete trait, we note that of the X chromosome-based genetic models that we considered, only one, the mouse *Xce* model,[8,41] appears capable of reconciling all of the observations. If alleles at an *Xce*-like locus in the human behave in the same way as those found in the mouse, then only individuals who are heterozygous for alleles of different 'strength' are predicted to exhibit non-random X-inactivation. X chromosomes bearing 'strong' *Xce* alleles have a high probability of remaining active, whilst X chromosomes bearing 'weak' *Xce* alleles have a low probability of remaining active.[41] This mode of inheritance is predicted to give rise to several unusual familial aggregations of X-inactivation phenotypes. For example, if a mother is homozygous for a weak *Xce* allele, she will exhibit random X-inactivation, but if the father of her daughters carries a strong *Xce* allele, all of their daughters will be skewed because they will be heterozygous for alleles of different strength. In such a family, of which K1362 may serve as an example (Figure 4), the daughters are predicted to exhibit the skewed phenotype regardless of which X chromosome they inherit from their mother. No mother–daughter correlation is expected because the mother will have random X-inactivation, whilst all of her daughters are predicted to be skewed. However, sisters in families such as K1362 are expected to show correlation in both the *DS* score (because they will all be skewed) and the PA_{mat} score (because daughters will tend to have an active paternal chromosome, regardless of which maternal X they have inherited.

Perhaps the most unusual type of family one might expect under this model is that in which some siblings are skewed in favor of an active maternal X whilst others are skewed in favor of an active paternal X. This circumstance is predicted to occur as a result of mating between a skewed mother, who was heterozygous for a strong allele and a weak allele, and a father who carried an allele of intermediate strength. Their daughters could be either of two genotypes: heterozygous for a strong allele and an intermediate allele, or heterozygous for a weak allele and an intermediate allele. Daughters of the first genotype would be skewed in favor of an active maternal X, whilst daughters of the second genotype would be skewed in favor of an active paternal X. Data gathered on the four skewed daughters of K1341 are consistent with such an inheritance pattern in the vicinity of the *XIC* (Figures 3 and 4 and data not shown).

Our genetic analysis of the X-inactivation phenotypes of females from 38 normal families disproves the hypothesis that X-inactivation in human females is 'random'. Although the precise mode by which this trait is inherited is uncertain, our data support a model in which X-inactivation phenotype is expressed as the result of the inheritance of specific alleles at one or two loci on the X-chromosome. In addition, our results are consistent with the existence of a human homologue of the mouse *Xce* locus.

Acknowledgements

This work was supported by grants from the American Cancer Society (to CS), the National Institutes of Health (GM52332, to CS) and the Canadian Genetic Diseases Network (NCE program, to KM). AKN and LO contributed equally to this work.

References

1 Lyon MF: Gene action in the X chromosome of the mouse (*Mus musculus L.*). *Nature* 1961; **190**: 372–373.
2 Belmont JW: Genetic control of X inactivation and processes leading to X-inactivation skewing. *Am J Hum Genet* 1996; **58**: 1101–1108.
3 Ropers H-H, Wolff G, Hitzeroth HW: Preferential X inactivation in human placenta membranes: Is the paternal X inactive in early embryonic development of female mammals. *Hum Genet* 1978; **43**: 265–273.
4 Harrison KB, Warburton D: Preferential X chromosome activity in human female placental tissues. *Cytogenet Cell Genet* 1986; **41**: 163–168.
5 Harrison KB: X chromosome inactivation in the human cytotrophoblast. *Cytogenet Cell Genet* 1989; **52**: 37–41.
6 West JD, Frels WI, Chapman VM, Papaionnou VE: Preferential expression of the maternally derived X chromosome in the mouse yolk sac. *Cell* 1977; **12**: 873–882.
7 Takagi N, Sasaki M: Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature* 1975; **256**: 640–641.
8 Cattanach BM, Perez JN, Pollard CE: Controlling elements in the mouse genome. *Genet Res* 1970; **15**: 189–195.
9 Willard HF: The sex chromosomes and X-chromosome inactivation. In: The metabolic and molecular bases of inherited disease. 7th edition, Scriver CR, Beaudet AL, Sly WS, Valle D. Eds. McGraw-Hill, Inc., 1995, 719–737.

10 Plenge RM, Hendrich BD, Schwartz C *et al*: A promoter mutation in the *XIST* gene in two unrelated families with skewed X-chromosome inactivation. *Nat Genetics* 1997; **17**: 353–356.
11 Brown CJ, Ballabio A, Rupert JL *et al*: A gene from the region of the human X-inactivation center is expressed exclusively from the inactive X chromosome. *Nature* 1991; **349**: 38–44.
12 Clarke JTR, Greer WL, Strasberg PM, Pearce RD, Skomorowski MA, Ray PN: Hunter disease (mucopolysaccaridosis type II) associated with unbalanced inactivation of the X chromosomes in a karyotypically normal girl. *Am J Hum Genet* 1991; **49**: 289–297.
13 Clarke JTR, Wilson PJ, Morris CP *et al*: Characterization of a deletion at Xq27-q28 associated with unbalanced inactivation of the nonmutant X chromosome. *Am J Hum Genet* 1992; **51**: 316–322.
14 Schmidt M, Certoma A, Du Sart D *et al*: Unusual X chromosome inactivation in a mentally retarded girl with an interstitial deletion Xq27: implications for the fragile X syndrome. *Hum Genet* 1990; **84**: 347–352.
15 Dahl N, Hu LJ, Chery M *et al*: Myotubular myopathy in a girl with a deletion at Xq27–q28 and unbalanced X-inactivation assigns the MTM1 gene to a 600-kb region. *Am J Hum Genet* 1995; **56**: 1108–1115.
16 Pegoraro E, Whitaker J, Mowery-Rushton P, Surti V, Lanasa M, Hoffman EP: Familial skewed X inactivation: a molecular trait associated with high spontaneous-abortion rate maps to Xq28. *Am J Hum Genet* 1997; **61**: 160–170.
17 Naumova AK, Plenge RM, Bird LM *et al*: Heritability of X-chromosome inactivation phenotype in a large family. *Am J Hum Genet* 1996; **58**: 1111–1119.
18 Puck JM, Stewart CC, Nussbaum RL: Maximum-likelihood analysis of human T-cell X-chromosome inactivation patterns: normal women versus carriers of X-linked severe combined immunodeficiency. *Am J Hum Genet* 1992; **50**: 742–748.
19 Migeon BR: The postulated X-inactivation center at Xq27 is most reasonably explained by ascertainment bias: heterozygous expression of recessive mutations is a powerful means of detecting unbalanced X inactivation. *Am J Hum Genet* 1993; **52**: 431–432.
20 Nance WE: Genetic tests with a sex-linked marker: glucose-6-phosphate dehydrogenase. *Cold Spring Harbor Symp Quant Biol* 1964; **29**: 415–424.
21 Pegoraro E, Schimke RN, Arahata K *et al*: Detection of new paternal gene mutations in isolated cases of dystrophinopathy in females. *Am J Hum Genet* 1994; **54**: 989–1003.
22 Ørstavik KH, Ørstavik RE, Eiklid K, Tranebjærg L: Inheritance of skewed X chromosome inactivation in a large family with an X-linked recessive deafness syndrome. *Am J Med Genet* 1996; **64**: 31–34.
23 Allen RC, Zoghbi HY, Moseley AB, Rosenblatt HM, Belmont JW: Methylation of *Hpa*II and *Hha*I 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.
24 Naumova AK, Olien L, Bird LM *et al*: Transmission-ratio distortion of X chromosomes among the male offspring of females with skewed X-inactivation. *Dev Genet* 1995; **17**: 198–205.
25 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.
26 Donner A, Eliasziw M: Methodology for inferences concerning familial correlations: a review. *J Clin Epidemiol* 1991; **44**: 449–455.
27 Efron B, Tibshirani R: *An Introduction to the Bootstrap*. Monographs on Statistics and Applied Probability No. 57. Chapman & Hall: New York, 1993.
28 Rosner B, Donner A, Hennekens CH: Estimation of interclass correlation from familial data. *J Roy Stat Soc, Series C (Applied Statistics)* 1977; **26**: 179–187.
29 Haseman JK, Elston RC: The investigation of linkage between a quantitative trait and a marker locus. *Behav Genet* 1972; **2**: 3–19.
30 Hauser ER, Boehnke M, Guo SW, Risch N: Affected-sib-pair interval mapping and exclusion for complex genetic traits: sampling considerations. *Genet Epidemiol* 1996; **13**: 117–137.
31 Fain PR, Kort EN, Chance PF *et al*: A 2D crossover-based map of the human X chromosome as a model for map integration. *Nat Genet* 1995; **9**: 261–266.
32 Ott J. *Analysis of Human Genetic Linkage*. John Hopkins University Press, Baltimore, 1991.
33 Guo SW, Thompson EA: *Analysis of sparse contingency tables: Monte Carlo estimation of exact* P-*values*, University of Washington, Department of Statistics, Technical Report NO. 187.
34 Busque L, Mio R, Mattioli J *et al*: Nonrandom X-inactivation patterns in normal females: lyonization ratios vary with age. *Blood* 1996; **88**: 59–65.
35 Lander E, Kruglyak L: Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995; **11**: 241–247.
36 Witte J, Elston R, Schork N: Genetic dissection of complex traits. *Nat Genet* 1996; **12**: 355–357.
37 Lander E, Kruglyak L: Genetic dissection of complex traits, Reply. *Nat Genet* 1996; **12**: 357–358.
38 Russel LB: Mammalian X-chromosome action: inactivation limited in spread and in region of origin. *Science* 1963; **140**: 976–978.
39 Mattei MG, Mattei JF, Vidal I, Girauld F: Structural anomalies of the X chromosome and inactivation center. *Hum Genet* 1981; **56**: 401–408.
40 Cuttichia AJ, Pearson PL (eds): *Human Gene Mapping 1993* John Hopkins University Press: Baltimore, 1994.
41 Cattanach BM, Beechey CV: Autosomal and X chromosome imprinting. *Development* (suppl) 1990; 63–72.
42 Vetrie D, Bobrow M, Harris A: Construction of a 5.2-megabase physical map of the human X chromosome at Xq22 using pulsed-field gel electrophoresis and yeast artificial chromosomes. *Genomics* 1993; **15**: 631–642.