ORIGINAL ARTICLE

Familial skewed X-chromosome inactivation linked to a component of the cohesin complex, *SA2*

Nisa KE Renault[1,6], Marc P Renault[2], Emily Copeland[3], Robin E Howell[4] and Wenda L Greer[5]

The gene dosage inequality between females with two X-chromosomes and males with one is compensated for by X-chromosome inactivation (XCI), which ensures the silencing of one X in every somatic cell of female mammals. XCI in humans results in a mosaic of two cell populations: those expressing the maternal X-chromosome and those expressing the paternal X-chromosome. We have previously shown that the degree of mosaicism (the X-inactivation pattern) in a Canadian family is directly related to disease severity in female carriers of the X-linked recessive bleeding disorder, haemophilia A. The distribution of X-inactivation patterns in this family was consistent with a genetic trait having a co-dominant mode of inheritance, suggesting that XCI choice may not be completely random. To identify genetic elements that could be responsible for biased XCI choice, a linkage analysis was undertaken using an approach tailored to accommodate the continuous nature of the X-inactivation pattern phenotype in the Canadian family. Several X-linked regions were identified, one of which overlaps with a region previously found to be linked to familial skewed XCI. *SA2*, a component of the cohesin complex is identified as a candidate gene that could participate in XCI through its association with CTCF.
Journal of Human Genetics (2011) 56, 390–397; doi:10.1038/jhg.2011.25; published online 17 March 2011

INTRODUCTION

X-chromosome inactivation (XCI) is a process whereby one of two X-chromosomes is inactivated early in mammalian female embryogenesis, to ensure that XX and XY individuals have a balanced expression of genes located on the X-chromosome.[1–3] Once the maternal or paternal X-chromosome has been transcriptionally silenced in a given somatic cell, the same X is inactivated in all progeny cells creating mosaic adult individuals. Despite the fact that XCI is essential and fundamental to normal female development, the mechanism is still largely unknown. It is well established that *XIST*, a non-coding RNA expressed exclusively from the inactive X, is responsible for initiating and maintaining silencing of that chromosome through CpG hypermethylation and chromatin remodelling.[4] However, the processes that differentiate an XX from XY cell and select one X for inactivation have yet to be elucidated.

The choice of which X is silenced has historically been considered a random process with the expectation that most females would have a ratio of active paternal X to active maternal X of ∼1:1 (balanced X inactivation pattern, XIP=0.5) and that significant variance from this pattern (skewed XIP) would be rare. There is growing evidence, however, that supports the concept that the choice step in XCI, and therefore the XIP, is genetically influenced. In mice, a polymorphic 'X-controlling element' (XCE), has been shown to influence the probability of XCI in *cis*, in a heritable fashion.[5,6] Although the XCE has yet to be isolated and characterised, it has been localised to the X-chromosome and four alleles with differing propensities for XCI have been identified. The XIPs in heterozygous mice are predictably skewed away from the balanced pattern that is observed in homozygous mice. Evidence in support of a human XCE is seen in reports of skewed XIPs clustering in families.[7,8] Naumova *et al.*[9] used an approach of classic linkage analysis, to show that markers DXS425 (Xq25) and DXS294 (Xq26) are likely linked to the human XCE. Cau *et al.* used a similar approach to identify a candidate region of 4.2 Mb cM ranging from DXS8067 (Xq24) to DXS8057 (Xq25), which partially overlaps the region identified by Naumova *et al.*[7]

In a previous report,[10] we described a family with three males and three females with clinical symptoms of haemophilia A, a bleeding disorder caused by low factor VIII activity. We showed that the three affected females had XCI ratios skewed toward activation of the mutated X-chromosome and that the degree of skewing correlated with FVIII activity and the severity of disease in all carriers in the family. Further, we found that more females showed a higher degree of skewing than would be predicted by a model of random XCI choice. Our data are consistent with a model of genetically influenced XCI

[1]Department of Pathology, Dalhousie University, Halifax, Nova Scotia, Canada; [2]UFR d'Informatique, Université Paris Diderot-Paris 7, Paris, France; [3]Department of Biology, Dalhousie University, Halifax, Nova Scotia, Canada; [4]Department of Oral and Maxillofacial Sciences, Dalhousie University, Halifax, Nova Scotia, Canada and [5]Department of Pathology, QEII Health Sciences Centre and Dalhousie University, Halifax, Nova Scotia, Canada
Correspondence: Dr WL Greer, Department of Pathology, Dalhousie University and QEII Health Sciences Centre, Department of Pathology, 5788 University Avenue, Mackenzie Building, Rm 223B, Halifax, Nova Scotia, Canada B3H 1V8.
E-mail: w.greer@dal.ca
[6]Current address: Unité Génétique Moléculaire Murine, Institut Pasteur, Paris, France.
Received 3 November 2010; revised 24 January 2011; accepted 31 January 2011; published online 17 March 2011

choice and support the possibility of a human XCE, similar to that described in mice.

Here we report further analyses of this family that have identified a critical region and *SA2* as a candidate gene for the human XCE.

MATERIALS AND METHODS

Ethics approval
This project was approved by the IWK Health Centre and the Capital District Health Authority Research Ethics Boards, Halifax, Nova Scotia, Canada (Project no. 2949). Informed consent/authorisation was obtained for all participants in this study.

Participants
The family studied here (Figure 1) is as we have previously described,[10] with the addition of seven new family members, II.4, II.10, III.4, III.6, III.7, IV.I and IV.2.

Factor VIII analyses
F8 intron 22 inversion analysis and functional factor VIII activity measurements were carried out as previously described.[10] In accordance with the IWK Health Sciences Centre guidelines on genetic testing of minors, the *F8* genotypes of young females, who were not obligate carriers, were not assessed. Factor VIII activity measurements were taken.

DNA extraction
DNA from peripheral blood samples was prepared as previously described.[10] Intraoral buccal mucosa cells were collected with cytology brushes (Orifice Medical AB, Ystad, Sweden). DNA was extracted using a QIAamp DNA Mini Kit (Qiagen, Germantown, MD, USA) according to the manufacturer's instructions.

XCI patterns in peripheral blood
XCI patterns (percentage of cells with Xa^p/Xi^m) in blood were determined by methylation sensitive endonuclease digestion (HhaI) followed by radioactive PCR amplification of the *HUMARA* microsatellite as described elsewhere.[10] At this locus, the inactive allele is hypermethylated,[11] therefore the peak with the greatest area is the least active. In some cases, the technique was adapted for analysis by fluorescent column electrophoresis (Supplementary Materials and methods). For participants who were informative at the *HUMARA* locus, but

whose alleles differed by only three base pairs in size, only the fluorescent column electrophoresis strategy could be used. In these instances, stutter peaks from the larger allele co-migrated with the smaller allele, falsely increasing the quantification of the latter. An analysis of allele stutter patterns (area of allele peak: area of stutter peak) in individuals whose alleles were well resolved revealed that stutter patterns were consistent in an allele-specific manner. Quantification of the smaller allele could therefore be calculated by subtracting the presumed area of the stutter peak from the total area quantified.

For individuals homozygous at the *HUMARA* locus, the *FMR-1* microsatellite was analysed[12] by the fluorescent CE strategy described (Supplementary Materials and methods) with the following modifications. After incubation with or without restriction endonuclease, the *FMR-1* microsatellite was amplified by PCR from aliquots of each solution (1.5 µl) using an FXS kit (Abbott Laboratories, Saint-Laurent Quebec, ON, Canada) according to the manufacturer's directions. For the visualisation of the PCR amplicons, 3.0 µl MM 1000 ROX Size Standard (BioVentures, Inc., Murfreesboro, TN, USA) was used. In contrast to the *HUMARA* locus, methylation of the *FMR-1* locus is associated with the active X-chromosome, therefore the peak with the greatest area corresponds to the most active allele. An established threshold[9] defines dramatically skewed XIPs as those > 0.80 or < 0.20. We considered an XIP somewhat skewed if it fell between 0.20–0.40 or 0.60–0.80, and balanced if it fell between 0.41–0.59. Values were calculated as an average of at least two independent determinations.

XIPs in buccal mucosa
XIPs in buccal mucosa were determined using a scaled-down version of the radioactive *HUMARA* protocol we have described previously,[10] where 123 ng DNA were digested in a total reaction volume of 3 µl. XIPs in buccal mucosa were an average of at least two determinations. When DNA was not limiting, three or more determinations were made.

Statistical calculations
The correlation between XIPs of blood and buccal mucosa was determined by linear regression analysis using Microsoft Excel (2003).

Candidate gene identification
Microsatellite analyses. Peripheral blood DNA was analysed at 48 X-linked microsatellite markers by 5-cM X-chromosome scan (Australia Genome Research Facility, Parkville, VIC, Australia, http://www.agrf.org.au) (Supplementary Table 1). Twenty
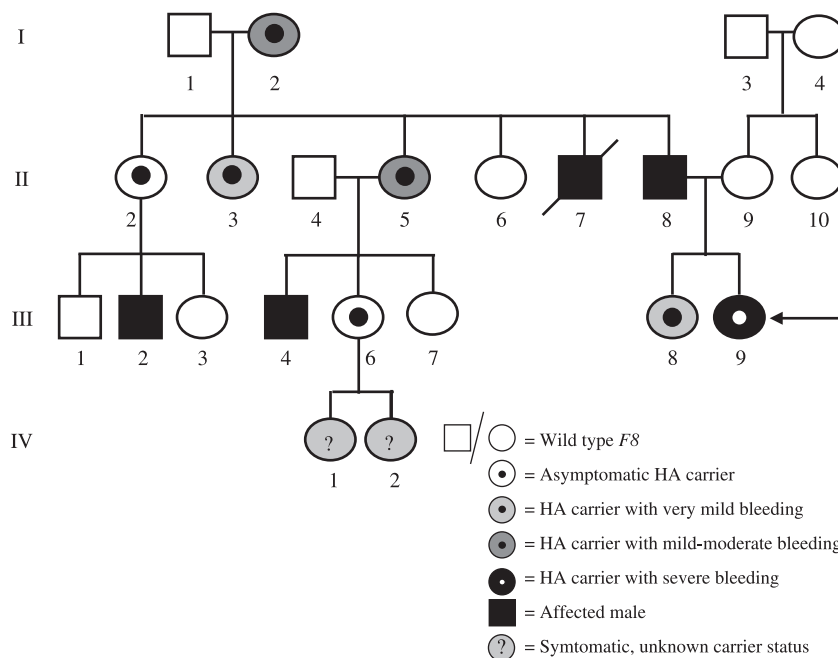

Figure 1 Atlantic Canadian family with unexpected haemophilia A expression in females. A four-generation family with haemophilia A (HA) expression in multiple females was ascertained when the proband (arrow) was first diagnosed with severe haemophilia A.

Table 1 Genotype sorting schemes for linkage analyses

#	Sorting schemes[a]
1	II.2=II.3=II.5=II.6
2	(II.2)≠(II.3=II.5=II.6)
3	(II.3)≠(II.2=II.5=II.6)
4	(II.5)≠(II.2=II.3=II.6)
5	(II.6)≠(II.2=II.3=II.5)
6	(II.2=II.3)≠(II.5=II.6)
7	(II.2=II.5)≠(II.3=II.6)
8	(II.2=II.6)≠(II.3=II.5)

[a]For all eight schemes, III.9≠III.8, IV.1≠IV.2 and III.6≠III.7.

additional X-linked markers were also analysed (Supplementary Materials and methods). The cytogenetic and physical locations of all markers were taken from the Ensembl Genome database (version 35) (http://www.ensembl.org), and the UCSC Golden Path database (http://genome.ucsc.edu/), respectively, and the primer sequences for in-house microsatellite analyses were taken from the Genome Database (http://www.gdb.org) (Supplementary Table 1).

Linkage analyses. To identify candidate *XCE* genes, a tailored linkage approach was taken. This was necessary because of both the continuous nature of the 'skewed XIP phenotype,' and the unknown genotype-phenotype correlation. Linkage analyses were first performed using data only from sister-pairs with clearly discordant XIP phenotypes. These were presumed to have differing XCE genotypes. For one sibship (generation II, Figure 1), it was unclear which sisters were likely to have the same genotypes. Assuming single-locus, X-linked inheritance, this sibship can have a maximum of two putative XCE genotypes. Sisters can therefore be sorted into a maximum of two genotype groups. For four sisters, there are a total of eight possible genotype 'sorting schemes' (Table 1). All sorting schemes were assessed for linkage to the X-chromosome. To facilitate these linkage analyses, we created a computer program, which sorts marker data by chromosomal location, phases haplotypes when possible, and identifies candidate regions fulfilling the genotype sorting scheme under assessment. The program was written in C++ using GNU Emacs (v. 21.3.1) and compiled with GCC (v. 4.3.2) or GCC MinGW (v. 3.4.5).

Candidate gene identification. Candidate regions identified through our linkage analyses were compared with XCE candidate regions described in the literature[7,9] Overlapping regions were recognised and candidate genes within the regions of overlap were identified using the Ensembl Genome Database. The Ensembl GNF development stage filter identified candidate genes expressed in the embryo.

SA2 sequencing

Exonic region identification. All *SA2* (ENSG00000101972) transcript variants, were identified in the Ensembl Genome Database. Regions of sequence coding for exons in any processed transcript variant (exonic regions) were identified and enumerated (Supplementary Table 2). To facilitate this process, we created a computer program, written in C++ and PERL (v. 5.10.0) using GNU Emacs (v. 21.3.1) and compiled with GCC (v. 4.3.2) or GCC MinGW (v. 3.4.5). Primers were designed to amplify and sequence each region using Primer3.[13] Amplifications were carried out in-house (Supplementary Materials and methods). Primarily, amplicons were spin-column purified using the illustra GFX PCR DNA and Gel Band Purification kit (GE Healthcare, Baie d'Urfe, Quebec, Canada) according to the manufacturer's directions, and sequenced at the Core Molecular Biology Facility, York University, Ontario, Canada (http://www.yorku.ca/biocore/) (indicated as 'York' sequencing protocol in Supplementary Table 2). Additional amplicons were sequenced in-house (Supplementary Materials and methods). All sequences were analysed with Mutation Surveyor v. 3.2.1 (SoftGenetics, Philadelphia, PA, USA).

RESULTS

Pedigree expansion and mutation analysis

A total of 15 females and seven males across four generations were included in this study (Figure 1). Factor VIII activity and *F8* genotype analyses were performed where appropriate. Heterozygosity for the intron 22 inversion (type II) was confirmed in seven females (Table 2). Two additional females (V.I and V.II) are considered carriers throughout the study due to their low factor VIII activities. Mutation analyses were omitted in both cases in accordance with IWK Health Center's policy on the genetic testing of minors.

Correlation of XIP in blood vs buccal indicates that selective pressures are unlikely to account for skewed XIP

XIPs can be skewed due to a bias at the onset of XCI (primary skewing) or can become skewed secondary to selective pressures (secondary skewing). Previously, we showed that the inheritance of the *F8* mutation is not affecting XIPs in this family.[10] To examine whether some unknown factor could be imparting a selective pressure leading to skewed XIPs, we assessed whether XIPs were tissue-specific, a hallmark of secondary skewing. XIPs in buccal mucosa were determined for 11 females, and ranged from balanced (III.7: 0.53 SD 0.04) to dramatically skewed (III.9: 0.84 SD 0.02) (Table 2). It was not possible to determine the XIPs in buccal mucosa from three females (I.4, IV.1, IV.2) due to poor yield and/or quality of extracted DNA, or two other females (II.10 and III.3) due to homozygosity at the HUMARA allele. An overall comparison of average blood and buccal XIPs reveals a good correlation ($R^2=0.8$; Figure 2). It has been reported that XIPs in blood become artificially skewed due to the progressive demethylation of the HUMARA locus in that tissue.[14] We do see that in 7/10 cases, the average XIPs in blood were more extreme (further from XIP=0.5) than buccal XIPs, however, this is not a significantly high proportion of cases ($P > 0.2$). Further, in 8/10 cases, any discrepancy between buccal and blood XIPs did not exceed the SD calculated from technical replicates. In III.9 and II.3, unusually small SDs (2 and 0, respectively) may account for the slight lack of fit between blood and buccal XIPs.

As the buccal mucosa and blood are derived from embryologically distinct germ layers, the correlation between XIPs in these tissues suggests that, for the most part, XIP skewing is occurring early in development in this family. This is consistent with a hypothesis of biased primary XCI, possibly at the level of XCI choice. Alternatively, there could be a selective pressure that is affecting all tissues. This is unlikely, however, as mutations dramatically affecting cell survival or proliferation probably would have been embryonically lethal if present in hemizygous males such as II.8 or I.3. Or, if the mutation were passed through maternal lineages only, we would expect a preponderance of paternally biased XIPs in females inheriting the mutation, where the deleterious maternal allele is preferentially inactive. A statistical analysis of the prevalence of paternally biased XIPs vs maternally biased XIPs among the female relatives of III.9 on the paternal side of her family finds that there is no significant preponderance of paternally biased XIPs ($P > 0.2$, data not shown).

X-chromosome-wide linkage analysis identifies an XCE candidate gene: *SA2*

Having ruled out common causes of genetically influenced skewed XIPs in a previous report,[10] we considered here, whether skewed XCI in this family might be caused by a mechanism analogous to that described in the mouse XCE hypothesis.[5,15] Neither the nature nor precise location of the putative human XCE are known, though Xq25 appears to be linked to familial skewed XCI.[7,9] To determine whether the skewed XCI

Table 2 Summary of molecular results for carrier and non-carrier females

Ppt:	I.2	I.4	II.2	II.3	II.5	II.6	II.9	II.10	III.3	III.6	III.7	III.8	III.9	IV.1	IV.2
Carrier	Y	N	Y	Y	Y	N	N	N	N	Y	N	Y	Y	-	-
FVIII (U ml^{-1})	0.39	1.61	0.82	0.84	0.28	1.61	1.33	—	1.35	1.93	1.12	0.52	0.02	0.55	0.29
Bld	0.84 (0.03)[a]	0.84 (0.12)	0.82 (0.13)	0.77 (0.09)	0.38 (0.03)	0.69 (0.03)	0.30 (0.04)	—	0.76 (0.10)[a]	0.77 (0.03)	0.56 (0.03)	0.53 (0.01)	0.92 (0.02)	0.49 (0.08)	0.18 (0.05)
Muc	0.78 (0.05)[a]	—	0.65 (0.09)	0.58 (0.00)	0.35 (0.03)	0.62 (0.05)	0.39 (0.07)	—	—	0.80 (0.03)	0.53 (0.04)	0.61 (0.09)	0.84 (0.02)	—	—

Abbreviations: Bld, XIP as measured in blood; FVIII, factor VIII activity levels; Muc, XIP as measured in buccal mucosa; Ppt, participant; '–', unknown.
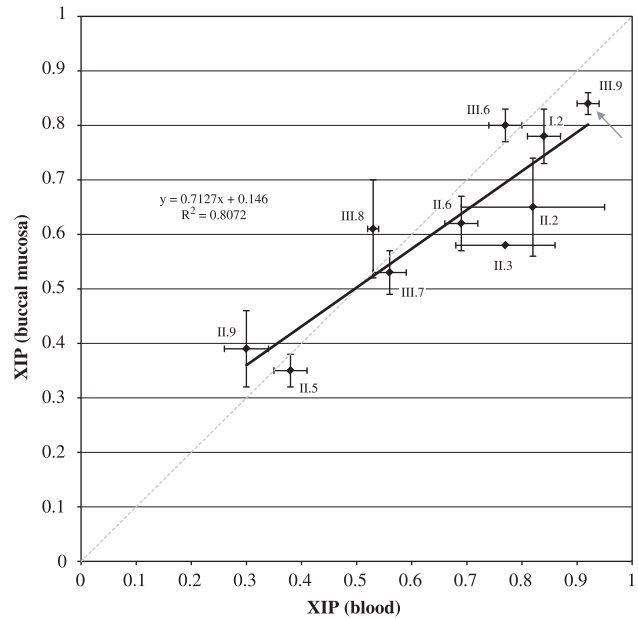s.d. is shown in parentheses
[a]Directionality of XIP unknown.

Figure 2 A good correlation between the XIPs in blood and buccal mucosa suggests that secondary skewing is not a major contributor to the XIPs observed. All females for whom we had both buccal and blood XIP data were included. Phase was randomly assigned to female I.2. The dotted grey line represents a perfect correlation between buccal and blood XIPs. The proband is marked by a grey arrow.

in the family studied here might also be linked to the X-chromosome, an X-chromosome-wide linkage analysis was undertaken.

As a skewed XIP phenotype is not precisely defined, and the XIP is a continuous trait, it is not possible to categorically assign phenotypes to each female. Therefore, we developed the following linkage approach. First, only sibships with clearly divergent phenotypes are considered (III.6-III.7, III.8-III.9 and IV.1-IV.2). Accordingly, in each case, sisters are assumed to have different genotypes. Our first linkage analysis considered only III.8 and III.9. We considered all markers to be a 'match' where III.8 and III.9 had inherited different alleles from their mother (Supplementary Table 3, step 1). All markers for which their mother was uninformative (homozygous or missing data) were also included. Next, we reduced the number of matching markers by considering IV.1 and IV.2. All markers at which their mother, III.6, was heterozygous but at which IV.1 and IV.2 had inherited the same maternal allele were eliminated. Similarly, we considered III.6 and III.7. Markers were eliminated if their mother, II.5, was heterozygous but sisters III.6 and III.7 had identical genotypes.

We then considered II.2, II.3, II.5, and II.6. Their intermediate XIPs precluded any obvious phenotype groupings. Assuming X-linked inheritance, a maximum of two XCE genotypes could be present in the sibship if their mother is heterozygous. This defines a total of eight genotype sorting schemes (Table 1). We tested each for linkage to the X-chromosome. Four sorting schemes (1, 3, 5, 6) were consistent with linkage to the X-chromosome for at least two consecutive markers (Supplementary Table 3). Three sorting schemes (3, 5, 6) reveal regions consistent with linkage for at least three consecutive markers (Table 3). The largest region identified, defined by sorting scheme 3, consists of 16.7 cM of Xq25-Xq27 (match region #1, Supplementary Figure 1). This overlaps by 0.7 cM with a region previously described in the literature as being linked to familial skewed XCI in humans[7,9] (match region 1.1, Supplementary Figure 1). Region 1.1, defined by markers DXS8098 and DXS8057, contains seven novel non-coding

Table 3 Tailored linkage approach reveals four linkage regions

Sorting scheme	Region	Size (Mb)
3	DXS8098 (Xq25)–DXS984 (Xq27.1)	16.7
5	DXS8091 (Xq28)–DXS1073 (Xq28)	6.1
6	DXS1238 (Xp21.1)–DXS8102 (Xp11.4)	6.9
6	DXS8102 (Xp11.4)–DXS1055 (Xp11.3)	7.3

RNA genes and four protein coding genes: *XIAP, SA2, SHZD1A* and *ODZ1*. Only *SA2* (*STAG2/Scc3* homologue), a component of the cohesin complex, is known to be expressed in the embryo.

The linkage approach used to identify these candidates considered each sibship in isolation. Haplotype analysis of the markers flanking region 1.1, which considers the entire pedigree, is also consistent with linkage to region 1.1. This analysis assumes a co-dominant model analogous to the mouse XCE model, where homozygous females, regardless of genotype, tend toward balanced XIPs, and heterozygocity results in skewed XCI. Figure 3 shows how the inheritance of three theoretical *SA2* alleles, designated A, B and C, could explain the XIP phenotypes in the family. In this model, A and B are strongly associated with the active X, and C is weakly associated with the active X as described by the relationship: $A > B > > C$. This model describes the degree of skewing for all females in the family, and the directionality of all but one, II.5. This discrepancy could indicate that we have not identified the correct gene. Alternatively, SA2 may indeed affect the degree of skewing but another element may affect the direction of skewing. It is also possible that the effect of SA2 on XCI choice may be influenced by chance, or that some unexplained secondary skewing is affecting the XIP of II.5 in her blood, buccal mucosa and, presumably her liver as well.

Sequencing of *SA2* reveals unusual intron-exon boundaries
Working under the assumption that III.8 and III.9 have discordant XCE genotypes, we sought any difference in their *SA2* gene sequences. *SA2* transcript analysis identified 45 unique exonic regions among 25 splice variants (Supplementary Table 2). No mutations were confirmed; however, seven poly-N tracks were present whose precise lengths could not be determined in III.8, II.9 or their parents (Supplementary Table 4). Six of these are intronic, lying adjacent to splice sites and one lies within the 3′ UTR of nine *SA2* transcript variants.

DISCUSSION
Evidence for primary skewing
Adult secondary X-inactivation patterns may differ from the primary XIP due to stochastic and selective events. We previously ruled out chromosomal abnormalities and selection against the factor VIII mutation[10] as possible causes of secondary skewing in this family and proposed that the skewing could principally be due to a primary bias in XCI choice. We have presented here further evidence that the skewing is of a primary nature. The good correlation between XIPs in blood and buccal mucosa argues against tissue specific secondary skewing due to selection, and the proportion of viable male descendents of I.2 (male:female=5:11, $P > 0.1$) makes selection against multiple or all tissues unlikely. Thus, primary skewing is most likely.

The cohesin component SA2, may be influencing XIPs in this family
Three mechanisms have been described in mammals, which affect primary XIPs by biasing XCI choice at the onset of XCI: (1) *XIC* mutations, (2) imprinting and (3) heterozygosity for a putative

X-controlling element, XCE. Only heterozygosity for an XCE has not been ruled out by our previous investigations.[10] There is evidence to support the existence of a human XCE,[7,9] although no gene has yet been identified. The XCE hypothesis suggests that the XCE phenotype depends on both the maternal and paternal alleles in a co-dominant fashion. This would explain how sib-pairs such as III.8 and III.9 could be discordant for the skewing phenotype (0.53 SD 0.01 and 0.92 SD 0.02, respectively), and how a child and her grandmother, but not her mother, could be dramatically skewed, as in the case of III.9, I.4, and II.9, respectively.

Using the XCE model, we designed a tailored linkage approach to identify loci influencing XIPs in this family. Our approach assumes that there is a single X-linked locus responsible for the XCE effect and that there is a good XCE genotype/XIP phenotype correlation. Autosomal factors are likely involved in XCI,[16–18] however, as the mouse XCE is X-linked,[19] and most of the key elements involved in XCI of humans and mice are X-linked, we reasonably began our search for the human XCE on the X-chromosome. It may be that several loci interact to determine XIPs. There are studies critical of the concept of a single locus having a large effect on XIPs,[20] however, these are predicated on a purely recessive or dominant mode of inheritance, and their data can be consistent with the XCE hypothesis if re-evaluated considering the XIP as a co-dominant trait.[21] The strength of the genotype/phenotype is unknown. Though our evidence suggests a low likelihood for significant secondary skewing effects, our data cannot rule out selective influences all together, nor can it illuminate the importance of stochastic events.

Given these assumptions, our linkage analysis reveals several match regions, one of which overlaps with XCE candidate regions described elsewhere.[8,9] Assuming a common aetiology, the XCE candidate gene region can be reduced to the interval ChrX: 122914524–123576797.

The putative XCE interval contains an embryologically expressed gene, SA2, whose known biological functions make it a strong candidate for XCE function. SA2 is a core component of the ring-like cohesin complex.[22–24] Cohesin is best known for its role in sister chromatid cohesion before and during mitosis,[23,24] however, other functions have recently been attributed to cohesin,[25] including nuclear re-organisation,[26] S-phase check point activation, DNA repair[27,28] and gene regulation.[29–31] SA2 itself has been shown to have transcriptional co-activator function.[32] Cohesin is recruited to chromatin through its interaction with DNA-binding proteins, including the CCCTCF binding factor (CTCF).[33] It co-occupies up to 90% of CTCF binding sites,[33,34] and some functions initially attributed to CTCF may depend on cohesin.[34] The cohesin-CTCF complex, which assembles in a cell-cycle-dependant[35] and developmentally-regulated[36] manner, can recruit RNA polymerase II to directly promote transcription,[37] and form DNA loop structures to mediate transcriptional insulation[31,34] or connect enhancers with core promoters.[25] Complexed with various other transcription factors, cohesin mediates tissue-specific[30,33] and hormone responsive transcription,[33] as well as the coordinated expression of genes with interrelated functions[38] and multiple genes in clusters.[39] The gene expression control functions of cohesin are molecularly separate from its role in chromatin cohesion such that mutations affecting gene expression do not necessarily affect genome integrity.[40]

Several lines of evidence suggest a potential role for cohesin in XCI. CTCF binding sites located within the *XIC*[41–43] likely recruit cohesin to the locus. CTCF binding is sensitive to methylation,[44] and XCI choice may be affected by differential methylation of CTCF binding sites within the *XIC*.[45] Mutations of the CTCF binding site within the *XIST* promoter result in skewed XCI choice.[46,47] XCI is initiated by
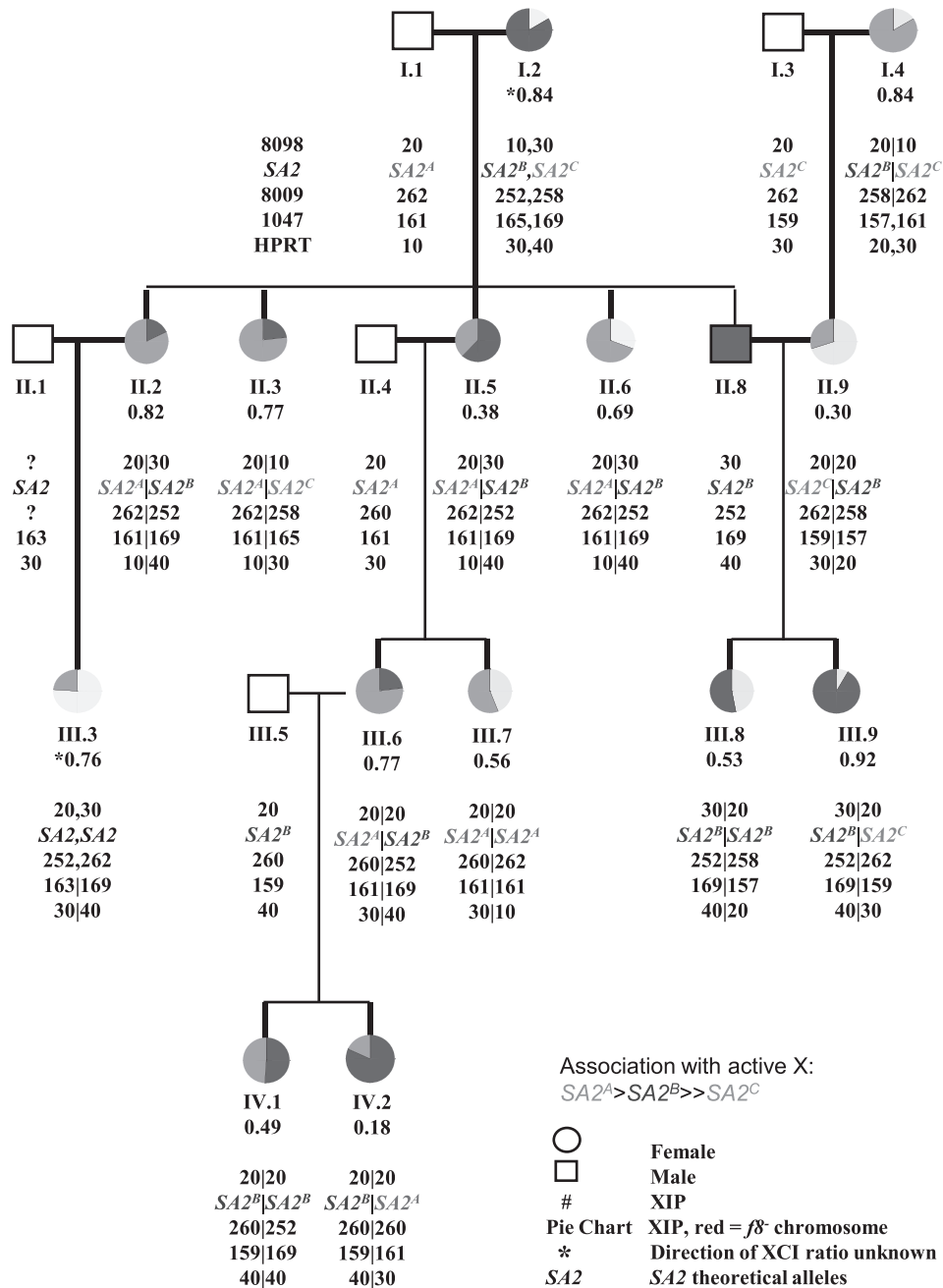
Figure 3 *SA2* and *Y-RNA* genes segregate with theoretical XCE alleles. Theoretical SA2 alleles are shown with the allelic designation of several flanking markers. X-chromosome inactivation (XCI) ratios are indicated below the pedigree number of each individual.

transient colocalisation of the X-chromosomes,[48] and CTCF is required for such pairing.[49] The enhancer-blocking function of CTCF has been implicated in XCI.[41] As a boundary element, cohesin-CTCF could establish the unique chromatin structure of the *XIST* gene, which is in a euchromatic state on the principally heterochromatic Xi.[50]

Our linkage data provides direct evidence linking the cohesin component SA2 to XCI. The specific mechanism by which SA2 could affect XCI choice is unknown. One possibility is that cohesin is involved in binding the two X-chromosomes together during their colocalisation before XCI initiation. As there is only one SA2 molecule in each cohesin ring complex, the orientation of the ring and location

of the SA2 molecule could favour transcriptional activation of one *XIC* versus the other. Biased XCI choice could thus result from genetic or epigenetic variation affecting the stability, level or orientation of cohesin binding. In the family studied here, no variations in the SA2 protein coding sequence were found, however, affects from different splice variants or protein expression levels remain reasonable possibilities. Variations in protein concentration of cohesin loading factors can affect the stability of cohesin binding and alter expression of cohesin-regulated genes.[51] RNAi mediated knockdown of Scc3/SA protein levels can also result in altered transcription of cohesin-regulated genes, perhaps by altering cohesin's insulator function.[52] Similarly to the finding that protein levels of X-linked genes can be

critical for the regulation of XCI,[53] precise SA2 expression levels could be important for XCI choice.

The work detailed here supports the hypothesis that a human XCE influences XCI choice. It identifies a region of the X-chromosome that is linked to familial skewed XCI. In conjunction with information presented by others,[7,9] a small region of linkage is determined, which contains a gene, *SA2*, with an inheritance pattern and known functions that make it a strong candidate for a role in XCI. This is the first time that the cohesin component SA2 or cohesin have been implicated in XCI.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

1 Lyon, M. F. Gene action in the X-chromosome of the mouse (Mus musculus L). *Nature* **190,** 372–373 (1961).
2 Lyon, M. F. Sex chromatin and gene action in the mammalian X-chromosome. *Am. J. Hum. Genet* **14,** 135–148 (1962).
3 Lyon, M. F. Lyonisation of the X chromosome. *Lancet* **2,** 1120–1121 (1963).
4 Lyon, M. F. & Rastan, S. Parental source of chromosome imprinting and its relevance for X chromosome inactivation. *Differentiation* **26,** 63–67 (1984).
5 Cattanach, B. M. & Isaacson, J. H. Controlling elements in the mouse X chromosome. *Genetics* **57,** 331–346 (1967).
6 Cattanach, B. M. & Rasberry, C. Enhanced specific-locus mutation response of 101/H male mice to single, acute X-irradiation. *Mutat. Res.* **311,** 77–84 (1994).
7 Cau, M., Addis, M., Congiu, R., Meloni, C., Cao, A., Santaniello, S. *et al.* A locus for familial skewed X chromosome inactivation maps to chromosome Xq25 in a family with a female manifesting Lowe syndrome. *J. Hum. Genet.* **51,** 1030–1036 (2006).
8 Naumova, A. K., Plenge, R. M., Bird, L. M., Leppert, M., Morgan, K., Willard, H. F. *et al.* Heritability of X chromosome—nactivation phenotype in a large family. *Am. J. Hum. Genet.* **58,** 1111–1119 (1996).
9 Naumova, A. K., Olien, L., Bird, L. M., Smith, M., Verner, A. E., Leppert, M. *et al.* Genetic mapping of X-linked loci involved in skewing of X chromosome inactivation in the human. *Eur. J. Hum. Genet.* **6,** 552–562 (1998).
10 Renault, N. K., Dyack, S., Dobson, M. J., Costa, T., Lam, W. L. & Greer, W. L. Heritable skewed X-chromosome inactivation leads to haemophilia A expression in heterozygous females. *Eur. J. Hum. Genet.* **15,** 628–637 (2007).
11 Allen, R. C., Zoghbi, H. Y., Moseley, A. B., Rosenblatt, H. M. & Belmont, J. W. 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.* **51,** 1229–1239 (1992).
12 Carrel, L. & Willard, H. F. An assay for X inactivation based on differential methylation at the fragile X locus, FMR1. *Am. J. Med. Genet.* **64,** 27–30 (1996).
13 Rozen, S. & Skaletsky, H. J. Primer3 on the WWW for general users and for biologist programmers. in: Krawetz S, Misener S (eds) *Bioinformatics. Methods and Protocols: Methods in Molecular in Molecular Biology*, Humana Press, Totowa, NJpp 365–386 (2000).
14 Swierczek, S. I., Agarwal, N., Nussenzveig, R. H., Rothstein, G., Wilson, A., Artz, A. *et al.* Hematopoiesis is not clonal in healthy elderly women. *Blood* **112,** 3186–3193 (2008).
15 Cattanach, B. M., Perez, J. N. & Pollard, C. E. Controlling elements in the mouse X-chromosome II Location in the linkage map. *Genet. Res* **15,** 183–195 (1970).
16 Silver, D. P., Dimitrov, S. D., Feunteun, J., Gelman, R., Drapkin, R., Lu, S. D. *et al.* Further evidence for BRCA1 communication with the inactive X chromosome. *Cell* **128,** 991–1002 (2007).
17 Percec, I., Plenge, R. M., Nadeau, J. H., Bartolomei, M. S. & Willard, H. F. Autosomal dominant mutations affecting X inactivation choice in the mouse. *Science* **296,** 1136–1139 (2002).
18 Percec, I., Thorvaldsen, J. L., Plenge, R. M., Krapp, C. J., Nadeau, J. H., Willard, H. F. *et al.* An N-ethyl-N-nitrosourea mutagenesis screen for epigenetic mutations in the mouse. *Genetics* **164,** 1481–1494 (2003).
19 Simmler, M. C., Cattanach, B. M., Rasberry, C., Rougeulle, C. & Avner, P. Mapping the murine Xce locus with (CA)n repeats. *Mamm. Genome* **4,** 523–530 (1993).
20 Bolduc, V., Chagnon, P., Provost, S., Dubé, M. P, Belisle, C., Gingras, M. *et al.* No evidence that skewing of X chromosome inactivation patterns is transmitted to offspring in humans. *J. Clin. Invest.* **118,** 333–341 (2008).
21 Renault, N. K. E. Examining the cause and effect of skewed X-chromosome inactivation patterns in humans, PhD thesis, Dalhousie University, Halifax, (2009).
22 Losada, A., Yokochi, T., Kobayashi, R. & Hirano, T. Identification and characterization of SA/Scc3p subunits in the Xenopus and human cohesin complexes. *J. Cell Biol.* **150,** 405–416 (2000).
23 Mannini, L., Menga, S. & Musio, A. The expanding universe of cohesin functions: a new genome stability caretaker involved in human disease and cancer. *Hum. Mutat.* **31,** 623–630 (2010).
24 Michaelis, C., Ciosk, R. & Nasmyth, K. Cohesins: chromosomal proteins that prevent premature separation of sister chromatids. *Cell* **91,** 35–45 (1997).
25 Wong, R. W. An update on cohesin function as a 'molecular glue' on chromosomes and spindles. *Cell Cycle* **9,** 1754–1758 (2010).
26 Gard, S., Light, W., Xiong, B., Bose, T., McNairn, A. J., Harris, B. *et al.* Cohesinopathy mutations disrupt the subnuclear organization of chromatin. *J. Cell Biol.* **187,** 455–462 (2009).
27 Kim, S. T., Xu, B. & Kastan, M. B. Involvement of the cohesin protein, Smc1, in Atm-dependent and independent responses to DNA damage. *Genes Dev.* **16,** 560–570 (2002).
28 Yazdi, P. T., Wang, Y., Zhao, S., Patel, N., Lee, E. Y. & Qin, J. SMC1 is a downstream effector in the ATM/NBS1 branch of the human S-phase checkpoint. *Genes Dev.* **16,** 571–582 (2002).
29 Dorsett, D. Cohesin, gene expression and development: lessons from Drosophila. *Chromosome Res.* **17,** 185–200 (2009).
30 Kagey, M. H., Newman, J. J., Bilodeau, S., Zhan, Y., Orlando, D. A., van Berkum, N. L *et al.* Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467,** 430–435 (2010).
31 Wendt, K. S. & Peters, J. M. How cohesin and CTCF cooperate in regulating gene expression. *Chromosome Res.* **17,** 201–214 (2009).
32 Lara-Pezzi, E., Pezzi, N., Prieto, I., Barthelemy, I., Carreiro, C., Martínez, A. *et al.* Evidence of a transcriptional co-activator function of cohesin STAG/SA/Scc3. *J. Biol. Chem.* **279,** 6553–6559 (2004).
33 Schmidt, D., Schwalie, P. C., Ross-Innes, C. S., Hurtado, A., Brown, G. D., Carroll, J. S. *et al.* A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Res.* **20,** 578–588 (2010).
34 Wendt, K. S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E. *et al.* Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451,** 796–801 (2008).
35 Kang, H. & Lieberman, P. M. Cell cycle control of Kaposi's sarcoma-associated herpesvirus latency transcription by CTCF-cohesin interactions. *J. Virol.* **83,** 6199–6210 (2009).
36 Degner, S. C., Wong, T. P., Jankevicius, G. & Feeney, A. J. Cutting edge: developmental stage-specific recruitment of cohesin to CTCF sites throughout immunoglobulin loci during B lymphocyte development. *J. Immunol.* **182,** 44–48 (2009).
37 Wada, Y., Ohta, Y., Xu, M., Tsutsumi, S., Minami, T., Inoue, K. *et al.* A wave of nascent transcription on activated human genes. *Proc. Natl Acad. Sci. USA* **106,** 18357–18361 (2009).
38 Skibbens, R. V., Marzillier, J. & Eastman, L. Cohesins coordinate gene transcriptions of related function within Saccharomyces cerevisiae. *Cell Cycle* **9,** 1601–1606 (2010).
39 Mishiro, T., Ishihara, K., Hino, S., Tsutsumi, S., Aburatani, H., Shirahige, K. *et al.* Architectural roles of multiple chromatin insulators at the human apolipoprotein gene cluster. *EMBO. J.* **28,** 1234–1245 (2009).
40 Horsfield, J. A., Anagnostou, S. H., Hu, J. K., Cho, K. H, Geisler, R., Lieschke, G. *et al.* Cohesin-dependent regulation of Runx genes. *Development* **134,** 2639–2649 (2007).
41 Chao, W., Huynh, K. D., Spencer, R. J., Davidow, L. S. & Lee, J. T. CTCF, a candidate trans-acting factor for X-inactivation choice. *Science* **295,** 345–347 (2002).
42 Donohoe, M. E., Zhang, L. F., Xu, N., Shi, Y. & Lee, J. T. Identification of a Ctcf cofactor, Yy1, for the X chromosome binary switch. *Mol. Cell* **25,** 43–56 (2007).
43 Navarro, P., Page, D. R., Avner, P. & Rougeulle, C. Tsix-mediated epigenetic switch of a CTCF-flanked region of the Xist promoter determines the Xist transcription program. *Genes Dev.* **20,** 2787–2792 (2006).
44 Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H. C. *et al.* Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132,** 422–433 (2008).
45 Boumil, R. M., Ogawa, Y., Sun, B. K., Huynh, K. D. & Lee, J. T. Differential methylation of Xite and CTCF sites in Tsix mirrors the pattern of X-inactivation choice in mice. *Mol. Cell Biol.* **26,** 2109–2117 (2006).
46 Pugacheva, E. M., Tiwari, V. K., Abdullaev, Z., Vostrov, A. A., Flanagan, P. T, Quitschke, W. W *et al.* Familial cases of point mutations in the XIST promoter reveal a correlation between CTCF binding and pre-emptive choices of X chromosome inactivation. *Hum. Mol. Genet.* **14,** 953–965 (2005).
47 Plenge, R. M., Hendrich, B. D., Schwartz, C., Arena, J. F, Naumova, A., Sapienza, C. *et al.* A promoter mutation in the XIST gene in two unrelated families with skewed X-chromosome inactivation. *Nat. Genet.* **17,** 353–356 (1997).
48 Bacher, C. P., Guggiari, M., Brors, B., Augui, S., Clerc, P., Avner, P. *et al.* Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. *Nat. Cell Biol.* **8,** 293–299 (2006).
49 Xu, N., Donohoe, M. E., Silva, S. S. & Lee, J. T. Evidence that homologous X-chromosome pairing requires transcription and Ctcf protein. *Nat. Genet.* **39,** 1390–1396 (2007).
50 Navarro, P. & Avner, P. An embryonic story: analysis of the gene regulative network controlling Xist expression in mouse embryonic stem cells. *Bioessays* **32,** 581–588 (2010).

51 Gause, M., Misulovin, Z., Bilyeu, A. & Dorsett, D. Dosage-sensitive regulation of cohesin chromosome binding and dynamics by Nipped-B, Pds5 and Wapl. _Mol. Cell Biol._ **30,** 4940–4951 (2010).

52 Rollins, R. A., Korom, M., Aulner, N., Martens, A. & Dorsett, D. Drosophila nipped-B protein supports sister chromatid cohesion and opposes the stromalin/Scc3 cohesion factor to facilitate long-range activation of the cut gene. _Mol. Cell Biol._ **24,** 3100–3111 (2004).

53 Jonkers, I., Barakat, T. S., Achame, E. M., Monkhorst, K., Kenter, A., Rentmeester, E. _et al._ RNF12 is an X-Encoded dose-dependent activator of X chromosome inactivation. _Cell_ **139,** 999–1011 (2009).

Supplementary Information accompanies the paper on Journal of Human Genetics website (http://www.nature.com/jhg)