UNIVERSITY OF SOUTHAMPTON

A Molecular Study of X Chromosome Inactivation in Humans

by Andrew James Sharp

Thesis submitted for the Degree of Doctor of Philosophy

Department of Human Genetics
Faculty of Medicine, Health and Biological Sciences

October 2002

The work presented in this thesis was conducted wholly whilst registered as a postgraduate student with the University of Southampton.

This thesis was submitted for examination in October 2002.

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF MEDICINE, HEALTH AND BIOLOGICAL SCIENCES

HUMAN GENETICS

Doctor of Philosophy

A MOLECULAR STUDY OF X CHROMOSOME INACTIVATION IN HUMANS

by Andrew James Sharp

Using DNA extracted from blood samples from 270 informative females, I determined that severely skewed X-inactivation in normal women is relatively common and increases with age ($P<0.05$). Samples of both buccal and urinary epithelia were also obtained from 88 of the females studied. Although there was a significant association of the X-inactivation ratios between different tissues in most individuals, wide variations were apparent in some cases, making accurate extrapolations between tissues impossible. The degree of correlation between tissues fell markedly with age. Overall, these data suggest that the major factors in the aetiology of skewed X-inactivation are secondary selection processes.

Previous studies in cases of trisomy rescue for a number of autosomes show a strong association with skewed X-inactivation. Data from the control group was used to test the hypothesis that trisomy 7 mosaicism causes Silver-Russell syndrome, a syndrome which has previously been attributed to imprinted genes. Consistent with the hypothesis, results showed a significant increase in the frequency of completely skewed X-inactivation in SRS patients (3 of 29) when compared to controls (3 of 270), suggesting the presence of undetected trisomy 7 in SRS patients.

Detailed studies of the spreading of X-inactivation into autosomal DNA in five unbalanced human X;autosome translocations were also performed. Using allele-specific RT-PCR long-range silencing of autosomal genes located up to 45Mb from the translocation breakpoint was observed, directly demonstrating the ability of X-inactivation to spread in cis through autosomal DNA. Spreading of gene silencing occurred in either a continuous or discontinuous fashion in different cases, suggesting that some autosomal DNA is resistant to the X-inactivation signal. Observations of late-replication, histone acetylation, histone methylation and XIST RNA show that X-inactivation can spread in the absence of cytogenetic features of the inactive X, although these histone modifications were found to be better cytogenetic correlates of the spread of X-inactivation than late-replication. Overall, there was good correlation between the pattern of gene silencing and the attenuation of clinical phenotype associated with each partial autosomal trisomy. Sequence analysis showed significant enrichment of LINE and LTR repeats within autosomal genes silenced by the spread of X-inactivation, providing compelling evidence that these are 'booster elements' which promote the spread of X-inactivation in cis, and explaining the variable spread of X-inactivation which is seen in X;autosome translocations. The association of LINEs and LTRs with X-inactivation fulfils Lyon's prediction that X-inactivation might represent a form of repeat-induced silencing, indicating that X-inactivation may have originally evolved from a genome defence mechanism against the invasion of transposable elements and ascribing functional significance to what is often called "junk DNA".

Contents

List of Figures

List of Tables

Acknowledgements

I am very grateful for the generous assistance of members of the Phlebotomy and Obstetrics and Gynaecology departments of Salisbury District Hospital and the patients and their families involved in this study, without whose co-operation this research would not have been possible. I would also like to thank Dr Trevor Cole (Clinical Genetics Unit, Birmingham Women's Hospital), Dr Evan Reid (East Anglian Medical Genetics Service, Cambridge) and especially Dr Nick Dennis (Wessex Clinical Genetics Service, Southampton) for their generous provision of clinical information and patient samples, Thomas Eggermann (Institute of Human Genetics, University of Aachen) and Gudrun Moore (Institute of Reproductive and Developmental Biology, Imperial College School of Medicine) for provision of SRS DNA samples, Jeffery Bailey (Case Western Reserve University, Ohio) and Will Tapper (Division of Human Genetics, University of Southampton) for performing repetitive sequence analysis, Hugh Spotswood and Bryan Turner (Chromatin and Gene Expression Group, University of Birmingham Medical School) for performing histone immunofluorescence, Lisa Hall and Jeanne Lawrence (Department of Cell Biology, University of Massachusetts) for performing *XIST* RNA FISH, Dr Andrew Barlow (Chromatin and Gene Expression Group, University of Birmingham Medical School) and Nicola Savage (Wessex Regional Genetics Laboratory) for advice with the replication timing assay, Paul Strike (Research and Development Support Unit, Salisbury District Hospital) for his expert help and advice with statistical analyses, and my supervisors David Robinson and Patricia Jacobs for helpful support, discussion and critical comment during the preparation of this manuscript. I would also like to thank Salisbury Hospitals Foundation for the provision of financial assistance towards the costs of presenting parts of this work at the 32nd meeting of the European Society of Human Genetics, Amsterdam 2000, the 51st meeting of the American Society of Human Genetics, San Diego 2001, and the 52nd meeting of the American Society of Human Genetics, Baltimore 2002, and, of course, The Wellcome Trust for their generous Prize Studentship (Ref. 058387) which funded this work.

Abbreviations

ABLIM1	Actin binding LIM protein 1
ANT2	Adenine nucleotide translocator 2
ANT3	Adenine nucleotide translocator 3
AR	Androgen receptor
AS	Angelman syndrome
BCS	Breast cancer screening
BRCA1	Breast cancer, type 1
BrdU	Bromodeoxyuridine
BRD2	Bromodomain containing 2
CCD	Charged-coupled device
cDNA	complementary-DNA
CF	Cystic fibrosis
cM	centiMorgan
CNTNAP2	Contactin associated protein-like 2
CpG	Cytidine-phosphate-guanosine
CPM	Confined placental mosaicism
CR1	Chicken Repeat 1
CSNK2B	Casein kinase 2, beta polypeptide
DAPI	Diamidinophenylindole
dATP	deoxyadenosine 5'-triphosphate
dCTP	deoxycytidine 5'-triphosphate
DEK	DEK oncogene (DNA binding)
dGTP	deoxyguanosine 5'-triphosphate
DMD	Duchenne muscular dystrophy
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
dTTP	deoxythymidine 5'-triphosphate
DUSP5	Dual specificity phosphatase 5
EBV	Epstein-Barr virus
EDTA	Ethylenediaminetetraacetate

Eed	Embryonic ectoderm development
EGTA	Ethylene glycol-bis(β-aminoethyl ether)
ES cell	Embryonic stem cell
EST	Expressed sequence tag
FAM	6-carboxyfluorescein
FITC	Fluorescein isothiocyanate
FISH	Fluorescent *in situ* hybridisation
FMR1	Fragile site, mental retardation 1
G-bands	Giesma-bands
GCLC	Glutamate-cysteine ligase, catalytic subunit
GFP	Green fluorescent protein
GMPR	Guanosine monophosphate reductase
GTL	G-bands by trypsin and Leishman's stain
G6PD	Glucose-6-phosphate dehydrogenase
HD	Huntington disease
HCR	Helix coiled-coil rod homologue
HLA-DRA	Major histocompatibility complex, class II, DR alpha
HLA-DRB5	Major histocompatibility complex, class II, DR beta 5
HLA-F	Major histocompatibility complex, class I, F
Homo	Homozygous
HPRT	Hypoxanthine guanine phosphoribosyltransferase
HPS	Hermansky-Pudlak syndrome
HRAS	v-Ha-ras Harvey rat sarcoma viral oncogene homolog
H3AcK14	Histone H3 acetylated at lysine 14
H4AcK8	Histone H4 acetylated at lysine 8
$H3Me_2K4$	Histone H3 dimethylated at lysine 4
H19	H19, imprinted maternally expressed untranslated mRNA
Igf2r	Insulin-like growth factor 2 receptor
IgG	Immunoglobulin G-group
IMS	Industrial methylated spirits
IRF4	Interferon regulatory factor 4
IUGR	Intra-uterine growth retardation
kb	Kilobase

KCNH2	Potassium voltage-gated channel, subfamily H (eag-related), member 2
lacZ	β-galactosidase
LINE-1/L1	Long interspersed nuclear element (type 1)
L1M1-4	LINE (type 1, mammalian specific)
L1P1-5	LINE (type 1, primate specific)
L1HS	LINE (type 1, Homo sapiens specific)
LINE-2/L2	Long interspersed nuclear element (type 2)
LMO2	LIM domain only 2 (rhombotin-like 1)
LTR	Long terminal repeat
mat	Maternal
Mb	Megabase
MCB	Macrochromatin body
MeCP1/MeCP2	Methyl-CpG binding protein 1/2
MICA	MHC class I polypeptide-related sequence A
MIR	Mammalian-wide interspersed repeat
M-MLV	Murine Moloney leukaemia virus
mRNA	Messenger RNA
MRPL23	Mitochondrial ribosomal protein L23
MXI1	MAX interacting protein 1
MZ	Monozygous
NG	Naso-gastric
NOTCH4	Notch homolog 4 (*Drosophila*)
ORF	Open reading frame
P	Probability
PAR	Pseudoautosomal region
pat	Paternal
PBS	Phosphate-buffered saline
PBT	Phosphate-buffered saline/Tween-20
PCTK1	Pctaire protein kinase 1
PCR	Polymerase chain reaction
PDE3B	Phosphodiesterase 3B, cGMP-inhibited

PDX1	Pyruvate dehydrogenase complex, lipoyl-containing component X1; E3-binding protein
PGK1	Phosphoglycerate kinase 1
PHA	Phytohaemagglutinin
PPC	Polyposis coli
PRDX3	Peroxiredoxin 3
PSMB9	Proteasome subunit, beta type, 9
PSMD13	Proteasome 26S subunit, non-ATPase, 13
PWS	Prader-Willi syndrome
r	Pearson correlation co-efficient
R	Co-efficient of determination
REP1	Repair of chromatin damage 1
RNA	Ribonucleic acid
rpm	Revolutions per minute
RPS4X	Ribosomal protein S4, X-linked
RRM1	Ribonucleotide reductase M1 polypeptide
RSA	Recurrent spontaneous abortion
RSB	Resuspension buffer
RT-PCR	Reverse-trancription PCR
SAA1	Serum amyloid A1
SCA1	Spinocerebellar ataxia 1
SDS	Sodium dodecyl sulphate
SINE	Short interspersed nuclear element
SLB	Sucrose lysis buffer
Smcx	Selected cDNA on X
SMPD1	Sphingomyelin phosphodiesterase 1, acid lysosomal
SNP	Single-nucleotide polymorphism
SRS	Silver-Russell syndrome
SSC	Saline sodium citrate
STS	Steroid sulphatase
SYBL1	Synaptobrevin-like 1
TAF2H	TAF10 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 30kDa
TAMRA	N, N, N',N'-tetramethyl-6-carboxyrhodamine

TE	Tris-EDTA
TEMED	N, N, N', N'-tetramethylethylenediamine
TES	Testis derived transcript (3 LIM domains)
TET	4, 7, 2', 7'-tetrachloro-6-carboxyfluorescein
TRITC	Texas Red isothiocyanate
TSIX/Tsix	X-inactive specific transcript - antisense
TSSC3	Tumor suppressing subtransferable candidate 3
U	Units
UBE1	Ubiquitin-activating enzyme 1
UPD	Uniparental disomy
XCE/Xce	X-controlling element
XIC/Xic	X-inactivation center
XIST/Xist	X-inactive specific transcript
X_a	Active X chromosome
X_i	Inactive X chromosome
YAC	Yeast artificial chromosome
ZFX	Zinc-finger protein, X-linked
ZNF76	Zinc finger protein 76 (expressed in testis)
3-D	Three-dimensional

Section 1 – Introduction

1.1 Introduction

The process of X-inactivation is a method of dosage compensation which occurs in females of all mammalian species. While the human X chromosome is large both in size and genetic content, the human Y chromosome is one of the smallest and contains only a handful of functional genes. Therefore there exists an inequality of gene content between males (46,XY) and females (46,XX). In order to compensate for this difference, one of the two X chromosomes in every somatic cell of female mammals is inactivated early in embryonic development. This inactivation causes the transcriptional silencing *in cis* of most of the genes on the inactive X chromosome (the X_i).

Analogous methods of dosage compensation have also been adopted in other species besides mammals. In the nematode *Caenorhabditis elegans* functional equivalence is achieved by halving the level of transcription from both X chromosomes in hermaphrodite XX individuals. Meanwhile in the fruit fly *Drosophila melanogaster* transcription from the single X chromosome of males is doubled in order to achieve equality with XX females. While there are obvious differences between these species and man, in each case there are close parallels in the molecular mechanisms by which transcriptional regulation is achieved. However, while all species achieve dosage compensation by influencing whole chromosomes as a single entity, the X-inactivation of mammals functions at an increased level of complexity by treating the two X chromosomes present in a single nucleus differently.

The process of X-inactivation was first proposed by Mary Lyon in 1961, based on the observation that while female mice that were heterozygous for X-linked coat-colour genes showed mottled patterns of expression, males were always uniform in colour (Lyon, 1961). Two years earlier, Ohno *et al.* (1959) had suggested that the Barr body represented a single heterochromatic X chromosome. Lyon proposed that this mosaic pattern of expression in females was a result of one or other of their X chromosomes being inactivated during embryonic development. Similar observations in humans were also made for females heterozygous for both ocular albinism and the

enzyme glucose-6-phosphate dehydrogenase (Beutler *et al.*, 1962), both of which are X-linked traits. As it is now commonly known, the 'Lyon hypothesis' was composed of four main proposals:

i) In normal females only one of the two X chromosomes present is genetically active, the other being inactivated.

ii) X-inactivation occurs early in development.

iii) The inactive X can be either maternal or paternal in origin, and the choice of which X to inactivate is random with respect to parental origin and independent of the choice in other cells of the embryo.

iv) X-inactivation is irreversible in somatic cells, such that the inactive X in a particular cell remains inactive in all descendants of that cell.

In the four decades since its inception, this theory has remained as the backbone of research into the process of X-inactivation. Recent work, in particular the identification of the gene *Xist* and its antisense partner *Tsix* as the likely mediators of the inactivation cascade, has shed light on how each of these processes occur.

It should be noted that much of the study of mammalian X-inactivation has been carried out in mice. Although the features of X-inactivation in mice mirror that in humans in many respects, several important differences are evident between the two species.

1.2 X-inactivation in the early embryo

During the earliest stages of female embryogenesis in placental mammals, both of the X chromosomes contributed to the zygote are active (Singer-Sam *et al.*, 1992). The earliest signs of X-inactivation are first seen in the extra-embryonic cells of the trophoectoderm and primitive endoderm, between 3.5 and 4.5 days post fertilisation in the mouse. In these first tissues to differentiate, X-inactivation - at least in the mouse - is imprinted such that the paternal X chromosome is preferentially inactive (Takagi *et al.*, 1982). In humans however it seems no such imprinting occurs. In the mouse, X-inactivation in the inner cell mass which is destined to form the embryo proper occurs at the late blastocyst stage, at 5.5-6.5 days post fertilisation (McMahon and Monk,

1983). In these tissues the paternal imprint no longer operates and the choice of which chromosome to inactivate is essentially random. On the basis of biochemical studies, X-inactivation in the mouse appears to be complete by the onset of gastrulation, at ~6.5 days post fertilisation (Monk and Harper, 1979). The presence of additional active X chromosomes is lethal beyond 10 days post fertilisation, presumably due to the increased levels of transcripts from dosage-sensitive X-linked genes (Takagi and Abe, 1990).

Once established the inactivation status remains stable throughout all subsequent somatic cell divisions. However, in the germline of female embryos the X_i is subsequently reactivated prior to meiotic division, at around 12.5-13.5 days post fertilisation (Kratzer and Chapman, 1981). It has been proposed that this reactivation may represent a requirement for a euchromatic state for normal meiotic chromosome pairing. Once reactivated the X chromosome then remains active in oocytes throughout ovulation and fertilisation until inactivation occurs in the developing zygote. Conversely in the male germline the single X chromosome is sequestered in the sex vesicle along with the Y chromosome and becomes transcriptionally inactive at the pachytene stage of meiosis (Lifschytz and Lindsley, 1972). It is unclear whether this inactivation during spermatogenesis is analogous to true X-inactivation seen in somatic cells, but such chromatin condensation is thought to prevent aberrant recombination events that might otherwise occur between the unpaired regions of the X and Y (McKee and Handel, 1993).

1.3 Features of the inactive X

Conformation - As identified by Barr and Bertram (1949), the Barr body is a condensed heterochromatic structure located at the nuclear periphery, characteristic of transcriptionally silent chromatin. This suggests that changes in the higher order packaging of the chromatin of the X_i occur which act to make it inaccessible to the transcriptional machinery of the cell. Indeed FISH studies have shown that sequences on the X_i are spatially compacted, while those on the active X chromosome (the X_a) tend to be decondensed (Dyer *et al.*, 1989). This is supported by experiments using DNA nucleases, which show that the X_i has a much higher resistance to DNAse

digestion than the X_a (Kerem *et al.*, 1983), and that DNA nuclease hypersensitive sites are present at the promoters of genes on the X_a, but not on the X_i (Yang and Caskey, 1987). While suggestive of changes in chromatin density, detailed studies using chromosome painting and 3-dimensional imaging have failed to find any significant differences in the volume occupied by the inactive X compared to its active counterpart (Eils *et al.*, 1996). Instead, it seems that changes in chromatin packaging and conformation occur, probably via alterations of nucleosome structure. This conclusion is supported by 3-D imaging studies of the X-linked genes *ANT2* and *ANT3*, which showed that the X-inactivated gene *ANT2* is positioned internally within the 3-D structure of the X_i but on the periphery of the X_a. In contrast *ANT3*, which is pseudoautosomal and therefore escapes X-inactivation, is located peripherally on both the X_i and X_a (Dietzel *et al.*, 1999).

The localisation of the X_i at the edge of the nucleus may also be an important component in its silencing. Recent reports have emphasised the significance of nuclear compartmentalisation in the regulation of transcription and replication, and in particular how the perinuclear localisation of chromatin can facilitate its transcriptional silencing (Andrulis *et al.*, 1998; Wei *et al.*, 1998). The telomeres of the X_i may function in this respect by anchoring it to the nuclear envelope, and perhaps account for its characteristic bend seen in metaphase spreads (Walker *et al.*, 1991).

Replication Timing - Studies of replication timing show that the X_i both initiates and completes its replication later in the cell cycle than the X_a (Willard and Latt, 1976). This late replication is the most precocious characteristic of the X_i to appear during embryonic development (Takagi *et al.*, 1982), preceding the conformational and biochemical modifications which all occur secondarily. Studies in female embryonic stem cells (ES cells), a good model system for studies of X-inactivation, have shown that a late replicating X is detectable immediately after the upregulation of *Xist* expression which initiates the X-inactivation process (Keohane *et al.*, 1996), and thus intimately linking transcriptional activity with changes in replication timing. This link is further strengthened by more detailed analyses of the replication timing of single X-linked genes. In the case of genes such as *FMR1*, *HPRT* and *G6PD*, all of which are subject to X-inactivation, the copy on the X_i was found to replicate later than the corresponding homologue on the X_a (Boggs and Chinault, 1994; Hansen *et al.*, 1996;

Torchia *et al.*, 1994). Conversely, analysis of the genes *RPS4X*, *ZFX*, *ANT3* and *STS*, all of which escape X-inactivation, revealed that the genes on the X_a and X_i both replicate early in the cell cycle (Boggs and Chinault, 1994). Therefore the replication timing of individual genes on the X appears to reflect their transcriptional activity, and is independent of the replication and transcriptional status of the chromosome on which they reside.

DNA Methylation - A number of studies have now shown that methylation of key sequences on the X_i plays an important role in the maintenance of its inactive status. Although the X_i and X_a have very similar global levels of methylation (Bernadino *et al.*, 1996), the CpG islands of genes subject to X-inactivation tend to be heavily methylated on the X_i and unmethylated on the X_a (Tribioli *et al.*, 1992). Meanwhile the CpG islands of genes that escape X-inactivation remain unmethylated on both the X_i and the X_a. CpG islands are regions of DNA with a high G and C content, particularly of GC doublets, which are associated with the 5' ends of genes. Once such sequences become methylated it is thought that they act to recruit the repressor proteins MeCP1 and MeCP2 (Boyes and Bird, 1992; Nan *et al.*, 1997), which bind solely to methylated DNA and inhibit the initiation of transcription.

The importance of methylation in the maintenance of X-inactivation has been demonstrated by experiments in which cells have been treated with the demethylating agent 5-azacytidine. This has been shown to cause the X_i to decondense with a concomitant increase in DNAse I sensitivity (Haaf, 1995), replicate earlier in the cell cycle (Jablonka *et al.*, 1985), and in one study using somatic cell hybrids also led to the reactivation of previously silenced genes (Hansen *et al.*, 1996). Additionally the relative instability of X-inactivation seen in marsupials and cultured human chorionic villi has been attributed to a lack of methylation (Cooper *et al.*, 1993; Migeon *et al.*, 1986).

However a number of observations now point to the fact that while methylation obviously plays an important role in the X-inactivation process, it is not in itself a primary cause of inactivation. Crucially, it is clear that the X chromosome is able to inactivate and silence transcription without methylation (Driscoll and Migeon, 1990). Studies into the chronological appearance of different features of the X_i also show that CpG island methylation often occurs late in embryonic development, well after X-

linked genes have inactivated (Lock *et al.*, 1987; Keohane *et al.*, 1998). Therefore, it seems that the primary function of methylation in X-inactivation is to stably maintain or 'lock' the X_i in its silent state. Importantly this methylation also acts as a mechanism for transmitting the inactivation signal through mitotic division, as any hemi-methylated CpG islands on newly replicated DNA are quickly methylated by DNA methyltransferases.

Histone Acetylation – Immunofluorescence studies have shown the X_i of both marsupial and eutherian mammals to be depleted of acetylated isoforms of histones H2A, H3 and H4 along most of its length (Wakefield *et al.*, 1997; Belyaev *et al.*, 1996). This same biochemical modification is associated with silent chromatin across many species, including the constitutive heterochromatin found in mammals, reflecting its fundamental role in chromatin dynamics. The correlation of transcriptional activity with histone acetylation status of the X_i is also maintained for those genes known to escape X-inactivation. Regions of residual histone H4 acetylation have been noted in the pseudoautosomal region and in Xp11.2-p11.3, both of which contain groups of genes which are transcribed from the X_i (Jeppesen and Turner, 1993). More detailed acetylation studies of the X chromosome show that differences between the X_i and X_a occur specifically at the promoters of individual genes. While genes that escape X-inactivation have highly acetylated promoters, the promoters of X-inactivated genes are markedly hypoacetylated in a way which mirrors the methylation of adjacent CpG dinucleotides (Gilbert and Sharp, 1999). However, as has been shown for methylation, hypoacetylation of the X_i does not appear to be an initiator of the X-inactivation process, but is more likely to be involved in the stable maintenance of inactive state. The chronological appearance of deacetylated histone H4 occurs several days after gene silencing is first detectable (Keohane *et al.*, 1996). Furthermore, this hypoacetylation of histones has recently been linked to the process of DNA methylation by the discovery that the methyl-CpG binding protein MeCP2 resides in a complex with histone deacetylase (Nan *et al.*, 1998; Jones *et al.*, 1998). Histone deacetylation has also been shown to act synergistically with DNA methylation in gene silencing (Cameron *et al.*, 1999) and therefore probably functions on the X_i to stably maintain its inactive state.

Histone Methylation – Immunofluorescence studies show that the active and inactive X chromosomes are differentially methylated at specific lysine residues of histone H3. While the X_i is enriched for H3 methylated at lysine 9, it is depleted of H3 dimethylated at lysine 4 along most of its length (Boggs *et al.*, 2001; Heard *et al.*, 2001), consistent with its heterochromatic nature (Jenuwein and Allis, 2001). Mirroring observations of histone acetylation on the X_i, residual H3 lysine 4 dimethylation is also apparent in Xp11 and at the Xp pseudoautosomal region, both of which are regions known to contain genes transcribed from the X_i. Punctate H3 lysine 4 dimethylation also occurs at approximately Xq25-26, suggesting this as a candidate region which may also contain genes escaping X-inactivation. Indeed, analysis by chromatin immunoprecipitation shows that H3 histones located at the promoters of genes subject to X-inactivation are methylated at lysine 9 and depleted of lysine 4 dimethylation on the X_i, whereas H3 methylation at the promoters of genes escaping X-inactivation shows the opposite distribution (Boggs *et al.*, 2001; Heard *et al.*, 2001). Temporal analysis of the enrichment of H3 lysine 9 methylation and depletion of H3 lysine 4 dimethylation on the X_i in differentiating ES cells show that these histone modifications are early events in the X-inactivation cascade and occur apparently co-ordinately. H3 lysine 9 methylation is visible on the X_i immediately after the initiation of the X-inactivation process, preceding the transcriptional silencing of X-linked genes, and is also a mitotically stable epigenetic mark (Heard *et al.*, 2001). Thus H3 methylation may play an important role in both the spread and maintenance of the X-inactivation signal.

1.4 Genes that escape X-inactivation

Some X-linked genes are expressed from both the active and the inactive X chromosome, and are said to 'escape' the inactivation process. Their existence was predicted on the basis that individuals with X chromosome aneuploidy show an abnormal phenotype, despite still having only a single active X chromosome. For unknown reasons, the number of genes that escape X-inactivation is much larger in man than in the mouse. This phenomenon may explain the relatively mild phenotype of the 45,X mouse compared to its human counterpart.

A gene is considered to escape X-inactivation if there is a detectable level of

expression from the X_i. The first such gene to be discovered was the XG blood group locus at the pseudo-autosomal region (PAR) in Xp22.3 (Fialkow, 1978). Lack of inactivation of such PAR genes was expected based on the fact that most have functional Y chromosome homologues that recombine during male meiosis. Their expression from both X chromosomes therefore maintains dosage equivalence between XY males and XX females. An interesting exception to this rule is the *SYBL1* gene located in the Xq PAR, which in contrast to most other pseudoautosomal genes is subject to X-inactivation and achieves dosage compensation by inactivation of the Y-linked copy (D'Esposito *et al.*, 1996).

However, not all genes that escape X-inactivation have corresponding homologues on the Y. Both *UBE1* and DXS423E located on the proximal short arm of the X lack Y-homologues (Brown and Willard, 1989; Brown *et al.*, 1995). The significance of the potentially increased level of transcription of such genes in females is unclear. Several studies in both man and mouse indicate that escape from inactivation may be a partial phenomenon in some cases. In the case of steroid sulphatase the actual level of expression may be lower from the copy on the X_i than from that on the X_a (Migeon *et al.*, 1982). Similarly for the *Smcx* gene in the mouse, transcription from the allele on the X_i is only 30-70% of that from the X_a (Carrel *et al.*, 1996a). Additionally expression of *Smcx* from the X_i was found to vary between both different individuals and different tissues.

Most assays of the inactivation status of X-linked genes have relied upon the use of mouse/human somatic cell hybrid systems which retain either an active or an inactive X chromosome, and have made little attempt to quantify relative levels of transcription. Preliminary results of such a survey identified 31 out of 125 genes along the X chromosome which are expressed from both the X_a and X_i (Carrel *et al.*, 1998). By extrapolation, this suggested that as many as one quarter of X-linked genes may escape inactivation. This systematic survey was later extended to include a total of 224 genes and expressed sequence tags (EST's), or approximately 10% of all X-linked transcripts (Carrel *et al.*, 1999). Of these, 28 (12.5%) were expressed in all X_i-containing hybrids tested, 147 (65.5%) were silenced, while 49 (22%) were heterogeneous in their expression, being silent in some hybrids but expressed in others. This heterogeneity of inactivation might reflect natural variation in X-inactivation, as found for *Smcx* and *REP1* (Carrel *et al.*, 1996a; Carrel and Willard, 1999), reactivation

in the mouse/human hybrid cells studied, or perhaps an innately unstable inactive state. Importantly, the genes that escaped X-inactivation were distributed non-randomly along the X chromosome, the majority being clustered in the short arm, particularly in distal Xp. By taking an unbiased sample of the genes analysed, Carrel *et al.* estimated that 19% of genes in Xp escape X-inactivation, compared to only 1% of those located in Xq.

Many of these genes known to escape inactivation are clustered within regions of the X chromosome, particularly in proximal Xp and in the PARs. This clustering has led to the suggestion that X-inactivation, and escape from it, may be a regional effect with entire blocks of genes being regulated in a co-ordinate fashion. Two genes which escape inactivation, *PCTK1* and *UBE1*, are separated by less than 5kb (Carrel *et al.*, 1996b), and lie within a larger 1Mb region which contains at least three other genes also known to escape inactivation (Brown *et al.*, 1995). This region however also contains a number of genes which are subject to X-inactivation (Miller *et al.*, 1995), and therefore it seems unlikely that any such co-ordinate control mechanism might operate at the megabase level.

Instead, this grouping of genes which escape inactivation probably reflects the evolutionary origins of the sex chromosomes. There is good evidence to suggest that most of the short arm of the human X chromosome resulted from a translocation of autosomal material to the ancestral sex chromosomes (Graves and Schmidt, 1992), and the PARs are similarly thought to have arisen by way of translocation (Graves *et al.*, 1998). By comparing the age and location of X-Y homologous gene pairs, Lahn and Page (1999) determined the human X chromosome to be composed of four distinct evolutionary strata. The oldest of these strata, the ancestral X chromosome, makes up the long arm of the X. As determined by Carrel *et al.* (1999), nearly all genes within this region are subject to X-inactivation. In contrast, the three more recent X chromosome strata define the proximal, medial, and distal regions of Xp respectively. These regions contain much higher numbers of genes escaping X-inactivation (Carrel *et al.*, 1999), and therefore, the density of genes escaping X-inactivation in each strata of the X chromosome inversely models its evolutionary age.

This distribution is consistent with prevailing evolutionary theory of how differentiation of the X and Y chromosomes occurred in the mammalian lineage. In this model, four separate chromosomal inversion events on the Y are thought to have

altered its gene order relative to the X, suppressing recombination between the two chromosomes (Lahn and Page, 1999). Lack of X-Y recombination within each stratum led to decay and eventual loss of genes in the affected region of the Y, providing a strong adaptive pressure for a mechanism of dosage compensation. Most likely, reduced or restricted expression from a Y-homologue caused a selective pressure in males for up-regulation of X-linked expression in order to maintain optimal gene activity. Subsequently, X-inactivation then acted as a counter-response to maintain optimal gene expression in females (Adler *et al.*, 1997; Jegalian and Page, 1998).

Recent evidence now points to a possible mechanism involved in propagating the spread of X-inactivation, and which forms further links between the evolution of the mammalian X chromosome and the distribution of genes upon it which escape inactivation (Bailey *et al.*, 2000). Gartler and Riggs (1983) first proposed that the X chromosome contains putative 'booster elements' or 'way stations' along its length that serve to promote the spread of the X-inactivation signal *in cis*. However, the exact nature of these 'booster elements' was then unknown. In 1998, Lyon put forward a "repeat hypothesis" in which she proposed the long interspersed repeat element (LINE) as a potential candidate. This hypothesis was based largely upon earlier FISH studies that had indicated that the X chromosomes of both mice and humans are greatly enriched for LINE elements when compared to the autosomes (Korenberg and Rykowski, 1988; Boyle *et al.*, 1990). In addition, Lyon had also noted that the extent of spread of X-inactivation in a number of murine X;autosome translocations seemed to correlate with cytogenetic bands that were rich in LINEs (Lyon, 1998). Lyon speculated that X-inactivation might therefore represent a form of repeat-induced silencing (a phenomenon whereby the expression of a serially arrayed DNA element is reduced as copy number increases, reviewed in Wolffe, 1997), mediated by LINE elements.

Using data published by the Human Genome Mapping Project, Bailey *et al.* (2000) directly compared the distribution of LINE elements on the X with that of the autosomes. As had been indicated by earlier FISH studies, their analyses found that the density of LINE elements on the X chromosome is overall some 2-fold higher than the rest of the genome. More specifically, classification of these repeats showed that almost all of this enrichment on the X is due to certain subfamilies of LINE-1 (L1)

elements that proliferated relatively recently in mammalian evolution. While older L1 forms showed only slight enrichment on the X, the younger L1M1 and L1P4 subclasses were between 3 and 5-fold more abundant on the X than on the autosomes. These subclasses are thought to have proliferated in the mammalian genome between 60 and 100 million years ago, corresponding with the radiation of the eutherian lineage and the evolution of their more complex X-inactivation cascade (Smit *et al.*, 1995; Wakefield *et al.*, 1997). In addition, Bailey *et al.* found that the distribution of these younger L1 elements showed strong clustering along the oldest stratum of the X chromosome. The cytogenetic band Xq13, containing the *XIC*, showed the highest density, with 45% of its sequence composed entirely of L1 repeats. Similar clusters of L1 elements were also located at bands Xq22, Xq24-25 and Xq27, thus satisfying the prediction of Gartler and Riggs' original model of a dense cluster at the site of initiation, with regularly spaced 'booster elements' to promote the spread of inactivation. In contrast, the band Xp22 was markedly reduced in L1 content (15.3%), and, consistent with its recent autosomal origin (Lahn and Page, 1999), very similar to the genomic average (15.6%). Further analyses of L1 composition along the X also demonstrated that the L1 content of segments that escape inactivation was significantly lower than those that contain genes which are subject to X-inactivation (10.3% vs. 20.5%). Taken together, Bailey *et al.* proposed that this non-random distribution of L1 elements on the X chromosome serves to propagate the X-inactivation signal *in cis*, perhaps by acting as binding sites for an *XIST* RNA/protein complex.

Despite these findings, the exact mechanism by which certain genes escape X-inactivation is currently unclear. It has been speculated that escape from inactivation may occur either primarily at the initiation of the X-inactivation process, such that certain genes are skipped over during the spread of inactivation, or secondarily through reactivation by a lack of appropriate maintenance. Studies into the patterns of expression of *Smcx* through embryonic development favour the latter explanation. Lingenfelter *et al.* (1998) found that *Smcx* was first inactivated in a proportion of cells in the early mouse embryo but progressively became expressed from every cell, implying that it becomes reactivated with time. The authors speculated that *Smcx* and other genes which escape X-inactivation are subject to fewer transcriptional controls than genes which are stably inactivated, thus allowing for their escape and expression.

1.5 Identification of the X-inactivation centre (*XIC*)

Early studies of sex chromosome aneuploidy showed a relationship between the number of X chromosomes in a cell and the presence of sex chromatin bodies. In 1959, Ford *et al.* found that the cells of a woman with Turner syndrome (45,X) contained no Barr bodies. In the same year, Jacobs *et al.* (1959) observed the presence of two Barr bodies in many cells of a 'superfemale' with a 47,XXX karyotype. These observations initially led Böök and Santesson (1960) to suggest that the autosomes might act to suppress the formation of a single sex chromatin body, with any additional X chromosomes being inactive. However, later studies showed that in cultured polyploid cells of an individual with Klinefelter syndrome multiple X chromosomes remained active (Harnden, 1962). As a result Böök and Santesson's 'n-1' rule was refined, invoking the existence of a counting mechanism which acts to somehow maintain one active X chromosome per diploid set of autosomes in every cell. Thus, Harnden put forward the following equation to express the relationship between the number of sex chromatin bodies (B) and the number of X chromosomes (x):

$$B = x-(p\div 2), \quad \text{where p is the ploidy of the cell.}$$

The presence of a specific locus required for X-inactivation was first proposed by Therman *et al.* (1974), who observed that when an X chromosome is involved in a translocation with an autosome, only one of the two remaining portions of the translocated X is able to inactivate. They noted that while inactivation occurred normally when many large segments of the X were deleted, the presence of the proximal long arm of the X was always required. This immediately suggested that X-inactivation requires the presence *in cis* of a central control region, termed the X-Inactivation Centre or *XIC*. Analysis of a number of such structurally abnormal X chromosomes in man resulted in the identification of a common minimal region of ~1Mb located in the proximal long arm of the X at Xq13.2 (Brown *et al.*, 1991a). At least two copies of this *XIC* are required for X-inactivation to occur, which must be located on different chromosomes (Muscatelli *et al.*, 1992), therefore implicating the *XIC* as a component of the counting mechanism invoked by the 'n-1 rule'.

1.6 The *XIST* Gene

Following the definition of the *XIC*, in 1991 Brown *et al.* localised a number of markers to the candidate minimal region in Xq13.2. One of these was the X-Inactive Specific Transcript or *XIST* gene. They found that an *XIST* cDNA probe hybridised to RNA prepared from females or somatic cell hybrids containing an inactive human X chromosome, but not to RNA from males or hybrid cells containing only an active X chromosome. Additionally, quantitative RNA blotting experiments showed that the level of *XIST* expression was proportional to the number of inactive X chromosomes present, thus demonstrating that *XIST* is expressed specifically from inactive, but not active, X chromosomes (Brown *et al.*, 1991b). *In situ* hybridisation further demonstrated that *XIST* is located at band Xq13. On the basis of its unique expression profile and chromosomal location, it was hypothesised that *XIST* is either primarily involved in, or uniquely influenced by, the X-inactivation process.

Analysis of the *XIST* gene showed that it is unusually large, comprising at least eight exons totalling 17kb of transcribed sequence. *XIST* is expressed ubiquitously in adult female somatic tissue, and has several alternatively spliced transcripts whose significance remains unknown. *XIST* also lacks any significant open reading frames (ORFs) and so does not appear to encode for any protein. Consistent with this, fluorescent *in situ* hybridisation (FISH) studies showed that the *XIST* RNA remains located within the nucleus in a position coincident with the Barr body (Brown *et al.*, 1992). This lack of translation and association of the *XIST* RNA with the X_i from which it is transcribed implicated *XIST* RNA as a possible structural component of the Barr body. The nuclear association of *XIST* RNA with the X_i was later confirmed by detailed three-dimensional FISH analysis (Clemson *et al.*, 1996). These studies showed that *XIST* RNA essentially paints the X_i during interphase in a particulate manner. Additionally Clemson *et al.* showed that *XIST* RNA is not bound to the DNA or chromatin of the X_i, but instead binds to part of the nuclear matrix, supporting the postulation that it may play a structural role in the architecture of the Barr body.

While observations of *XIST* RNA in human cells have shown that it remains associated with the X_i only during interphase (Clemson *et al.*, 1996), studies of murine *Xist* have found that this relationship is maintained until much later in the cell cycle

(Lee and Jaenisch, 1997). Duthie *et al.* (1999) utilised this dichotomy to study the localisation of *Xist* transcripts on rodent metaphase chromosomes. They found that *Xist* RNA exhibits a banded pattern on the X_i, binding preferentially to the gene-rich Giesma negative (G-light) bands in a number of different rodent species, and was excluded from regions of heterochromatin. Analysis of two murine X;autosome rearrangements also showed that *Xist* RNA associated with *cis*-linked autosomal chromatin much less readily, spreading some way into the attached autosomes in only a minority of cells. Similar observations were also made in two cases of human X;autosome rearrangements (Keohane *et al.*, 1999), leading Duthie *et al.* to postulate that factors required for the binding of *Xist/XIST* RNA and spread of inactivation are either absent or reduced on autosomes, relative to the X.

At about the same time as the discovery of *XIST* by Brown *et al.*, work in the mouse isolated the murine homologue, *Xist*. Like *XIST*, *Xist* is located in the murine *Xic* (Borsani *et al.*, 1991), is expressed exclusively from the X_i (Brockdorff *et al.*, 1991), lacks any significant ORFs, and remains within the nucleus associated specifically with the Barr body (Brockdorff *et al.*, 1992). Although *XIST* and *Xist* share a similar size and overall structure, comparisons between the human and murine homologues of *XIST/Xist* showed a relatively high level of sequence divergence between the two genes along most of their length, indicating no specific requirement of base composition for their function. However the 5' end of exon 1 was found to contain a stretch of tandem repeats which show strong homology and sequence conservation between the two species, suggestive of some functional significance (Brown *et al.*, 1992).

In 1993 Kay *et al.* provided further evidence in support of the involvement of *Xist* in the X-inactivation process. By studying the expression of *Xist* in the early mouse embryo, they found that *Xist* transcripts were first detectable at the 4-8 cell stage, at least one day prior to the earliest signs of an inactive X chromosome. Additional studies were also carried out using female mouse embryonic stem (ES) cells, a good model system for the study of X-inactivation. Prior to differentiating, female ES cells contain two active X chromosomes. However, upon differentiation they undergo X-inactivation in a way which is thought to mirror the developing

embryo (Martin *et al.*, 1978). Kay *et al.* found that *Xist* expression was up-regulated markedly in female ES cells upon their differentiation *in vitro*. Thus, by demonstrating that the onset of *Xist* expression precedes X chromosome inactivation, these observations indirectly suggested a causal role for *Xist* in the initiation of X-inactivation.

The first piece of evidence which directly proved the requirement of *Xist* for X-inactivation came in 1996, when a targeted deletion of 7kb of exon 1 and part of the promoter sequence was made on one of the two X chromosomes in female mouse ES cells (Penny *et al.*, 1996). When these cells were allowed to differentiate, X-inactivation occurred, indicating that the counting function of the deleted *Xic* was unaffected. Importantly however, only the normal X chromosome without the deletion was able to undergo inactivation, thus demonstrating that this knockout of *Xist* destroys its ability to cause X-inactivation *in cis* and proving the absolute requirement for *Xist* in X-inactivation. Furthermore, Penny *et al.* concluded that their targeted deletion did not affect the selection process of which X chromosome was to remain active. They suggested that elements that are required for counting the number of X chromosomes and choosing which to inactivate are located in regulatory elements either 5' or 3' to *Xist*. These *in vitro* results were later confirmed *in vivo* by Marahrens *et al.* (1997) who studied the phenotype of mice that carried an ~15kb deletion of exons 1 to 6 of *Xist*.

Further attempts to define the minimal region of the *XIC* and its functional elements were made by using a number of transgenic constructs. Lee *et al.* (1996) introduced 450kb of the murine *Xic*, including *Xist*, as a multicopy array into several different autosomes of male ES cells by transfection using Yeast Artificial Chromosomes (YACs). They observed that *Xist* was induced upon differentiation of the ES cells and could be expressed from either the endogenous or the ectopic loci, indicating that the 450kb transgene included the necessary elements for both the counting and choosing functions of the *Xic*. In addition, the ectopic *Xist* transcripts were found to co-localise to the autosome from which they were expressed and caused inactivation of a marker gene *in cis*, seemingly recapitulating X-inactivation process in its entirety. Thus, Lee *et al.* concluded that the *Xic* was contained entirely within their

450kb construct, and was sufficient to complete the counting, choosing and initiation of X-inactivation. Further studies using this same transgenic cell line showed that the association of the ectopic *Xist* RNA with its autosome was accompanied by a delay in replication timing and a marked hypoacetylation of histone H4, both characteristics of the X_i. Expression studies also showed a significant reduction in transcription from a number of housekeeping genes located up to 50 centiMorgans from the site of *Xic* insertion, although this appeared to vary from cell to cell. These latter findings suggested that the ability of *Xist* RNA to associate with and inactivate chromatin is not specific to the X chromosome, but can occur on any autosomal material that possesses an *Xic in cis* (Lee and Jaenisch, 1997).

Similar work published simultaneously studied the behaviour of male ES cells into which a multicopy cosmid containing the *Xist* gene and only 9kb of 5' sequence and 6kb of 3' sequence had been transfected (Herzing *et al.*, 1997). As was seen with the larger 450kb transgene generated by Lee *et al.*, this minimal *Xist* construct was also induced upon differentiation of the transgenic cell line - the *Xist* RNA associating with the autosome from which it originated accompanied by the silencing of a reporter gene *in cis*. In addition, Herzing *et al.* observed that some cells which contained the autosomal *Xist* insert appeared to inactivate the endogenous native *Xist* gene, suggesting that the transgene contained at least some of elements required for *Xic* counting. Therefore, based on its ability to seemingly carry out the processes of counting and *cis*-limited gene inactivation, it was suggested that the *Xic* is composed solely of the *Xist* gene alone.

However, it must be noted that the experiments of both Lee *et al.* (1996) and Herzing *et al.* (1997) generated ES cells which contained multiple copies of the *Xic* in tandem arrays, perhaps as many as 28 in one case, all located on a single autosome. While it is unclear exactly what effects this multicopy nature may have on the inactivation process, a number of observations made in these transgenic cell lines, such as a proportionate increase in *Xist* RNA levels with *Xic* copy number, indicates that their results should be interpreted with caution. Indeed, similar transgenic constructs containing almost identical fragments of the *Xic*, but present in only single or low copy number, have failed to express *Xist* or recapitulate any of its functions (Heard *et al.*,

1996; Matsuura *et al.*, 1996).

The continued expression of *Xist* and its association with the X_i in somatic cells implies that *Xist* plays a role in the maintenance of the inactive state. However, several observations now indicate that this is not the case. Brown and Willard (1994) found that portions of the X_i lacking *XIST* retained an inactive state in mouse/human somatic cell hybrids. Similarly, Rack *et al.* (1994) observed the continuing inactivation of rearranged inactive X chromosomes in leukaemic cells of two women. More recently, Csankovski *et al.* (1999) created murine fibroblasts heterozygous for a somatic deletion of *Xist*. Despite abolishing *Xist* transcription, the mutant X_i remained late replicating, hypoacetylated, and transcriptionally silent. Therefore, as had been indicated by earlier observations, Csankovski *et al.* concluded that the transcription and coating of the X_i by *Xist* is not required for the maintenance of X-inactivation.

1.7 The developmental expression of *XIST/Xist*

As demonstrated by Kay *et al.* (1994), *Xist* transcripts are first detectable by RT-PCR in the mouse embryo at the 4-cell stage, preceding the appearance of an inactive X. During the early cleavage stages, both X chromosomes of female embryos and the single X chromosome of male embryos are active (Epstein *et al.*, 1978), with X-inactivation occurring as cells differentiate from their totipotent state (Monk and Harper, 1978).

In 1995, studies by Beard *et al.* showed that prior to X-inactivation, *Xist* appeared to be expressed at very low levels from both alleles in undifferentiated female ES cells, and from the single X of undifferentiated male ES cells. This low level of expression was detectable only by the sensitive technique of RT-PCR, not by RNase protection, and could also be visualised directly by FISH as a small pinpoint signal over each *Xist* gene (Lee *et al.*, 1996). This low level of expression from all X chromosomes in undifferentiated ES cells contrasts dramatically with the situation following the onset of X-inactivation. Here *Xist* RNA is expressed only from the X_i, and is present in relatively large quantities coating the chromosome from which it is transcribed.

Panning *et al.* (1997) studied these patterns of *Xist* RNA accumulation in developing female mouse embryos. They observed that at the time of X-inactivation these cells exhibited a transient stage of expression with low levels of *Xist* RNA at one allele and high levels of *Xist* RNA at the other X chromosome. This suggested that X-inactivation is mediated by a shift from low-level to high-level *Xist* expression at the X_i, which occurs prior to a silencing of *Xist* expression from the X_a. However, analysis of the transcription rate of *Xist* showed no detectable differences between the two. Instead, Panning *et al.* found that the half-life of *Xist* RNA was significantly increased upon differentiation, indicating that the switch from low-level to high-level expression is not mediated by transcriptional activation but by a stabilisation of *Xist* transcripts at the X_i. This conclusion was further supported by Sheardown *et al.* (1997).

Johnston *et al.* (1998) apparently identified two novel promoters that control *Xist* expression. They interpreted their results to indicate that prior to the initiation of random X-inactivation, unstable *Xist* transcripts are produced from the upstream promoter, designated P_0, and that these fail to accumulate *in cis*. Johnston *et al.* suggested that at the onset of X-inactivation, transcription of *Xist* at the X_i switches to the downstream promoters P_1 and P_2 which produce stable transcripts and hence accumulate *in cis* at the X_i. Their results indicated that the accumulation of stable *Xist* transcripts correlates with, and is mediated by, a switch in promoter usage. Furthermore observations made in a number of *Xist* deletions led the authors to suggest that transcription from the putative P_0 promoter is required for an *Xist* allele to be registered by the counting mechanism at the onset of X-inactivation. They suggested that the putative switch from unstable P_0 to stable P_1/P_2 transcription represents the default situation occurring at *Xist* alleles during development, and that on n-1 alleles chosen to remain active this promoter switch is somehow blocked. As suggested by Panning *et al.* (1997), this block to the progress of X-inactivation could be achieved in two ways. Either by the presence of autosomally produced factors which are present in such limited quantities that they are able to occupy only a single *Xist* locus, or alternatively by the presence of a single privileged nuclear compartment per diploid genome which can be occupied by only one *Xist* allele.

However, more recent evidence now suggests that this promoter switch

hypothesis put forward by Johnston *et al.* is incorrect. Initially, studies using an 80kb transgene containing the *Xist* gene and only 730bp of its 5' sequence demonstrated that this construct was able to recapitulate X-inactivation, despite lacking the P_0 promoter (Lee *et al.*, 1999a). Subsequently, study of the nature of P_1 transcripts by Lee and colleagues indicated that they are neither intrinsically stable, nor accumulate *in cis*, as was suggested by Johnston *et al.* Further work also proved that the results obtained by Johnston *et al.* indicating the presence of an upstream P_0 promoter were actually due to non-specific PCR amplification of a ribosomal pseudogene located 5' to *Xist*. Finally, the supposed *Xist* transcription upstream of P_1 reported by Johnston *et al.* was instead found to originate from the opposing strand, representing the 3' end of the antisense transcript *Tsix* (Warshawsky *et al.*, 1999). As a result of this overwhelming evidence, Warshawsky *et al.* concluded that P_0 does not exist, and that the stabilisation and accumulation of *Xist* transcripts during differentiation does not occur as a result of a switch in promoter usage.

In contrast to the view of a block to inactivation occurring through a negative process which inhibits the progression of X-inactivation, Marahrens *et al.* (1998) hypothesised that the choice of which X chromosome to inactivate might instead be mediated by a positive mechanism which marks the future X_i. They found that in female mouse embryos that were heterozygous for a deletion of *Xist* exons 1-4 every cell had non-randomly inactivated the wild-type X chromosome. They suggested that this primary non-random inactivation indicates the presence of a positive element within their deletion that is required by an X chromosome for it to be chosen to inactivate. Comparing the phenotype of their *Xist* deletion mice to those of Penny *et al.* (1996) that retain the random choice mechanism defines a 6kb region of *Xist* that appears to be required for random choosing. Importantly, as the deletion of Penny *et al.* abolishes transcription of *Xist*, this 6kb region most likely functions as a DNA element rather than at the RNA level. The authors propose that this region may represent the binding site for specific proteins that initiate heterochromatin formation, and put forward a model in which the binding of a unique blocking factor and abundant initiation factors to X-inactivation occurs in a mutually exclusive fashion at each *Xist* allele, thus retaining a single active X per diploid cell.

A further targeted deletion of the *Xist* gene was created by Clerc and Avner (1998), extending for 65kb 3' to exon 6. In undifferentiated heterozygous ES cells *Xist* expression from the deleted allele was markedly reduced. Importantly though, upon differentiation X-inactivation occurred, but the normal X chromosome was never chosen to inactivate. Additionally, in cells which contained only a single deleted *Xist* allele, X-inactivation could occur normally. These latter observations indicated that the deleted 3' region of *Xist* contains a repressive regulatory element which prevents the default situation of X-inactivation, thus supporting the model proposed by Johnston *et al.* (1998).

Using strand-specific RNA FISH probes Lee *et al.* (1999b) identified a transcript anti-sense to *Xist* that was present in undifferentiated male and female ES cell nuclei. In reference to its reverse orientation, Lee *et al.* named this transcript *Tsix*. *Tsix* is apparently a single 40kb transcript which completely spans *Xist*, with a transcriptional start site beginning ~15kb 3' to *Xist*. It was therefore included in the deletion created by Clerc and Avner (1998), implicating it as a possible negative regulator of *Xist* expression. Consistent with this notion, Lee *et al.* found a dynamic association in the expression patterns of *Xist* and *Tsix* during the early stages of X-inactivation in ES cells. Prior to the onset of X-inactivation, both *Xist* and *Tsix* are bi-allelically expressed. At the onset of X-inactivation, *Tsix* expression becomes mono-allelic, preceding upregulation of *Xist* at the future X_i. At this stage, *Tsix* is expressed only from the future X_a, and once X-inactivation is established, *Tsix* is repressed. This expression profile suggests that *Tsix* regulates the early events of X-inactivation, and its anti-sense nature provides a mechanism through which it might block the function of *Xist* RNA. Like *Xist*, *Tsix* RNA localises within the nucleus specifically at the allele from which it is transcribed, but unlike *Xist* it never spreads beyond its transcription site. It also contains no conserved ORFs, and taken together with its nuclear localisation this strongly suggests it functions as an untranslated nuclear RNA. Lee *et al.* suggest a number of possible mechanisms by which *Tsix* may function to inhibit *Xist* RNA action - by base-pairing with *Xist* RNA and masking its functional sequences, alternatively the transcription of *Tsix* might in itself inhibit transcription of *Xist* on the opposing strand, or perhaps *Tsix* and *Xist* share a common enhancer for which they both compete.

Subsequent analysis of mice carrying a 3.7kb deletion encompassing the promoter region of *Tsix* further implicated it as a regulator of X-inactivation (Lee and Lu, 1999). While mice carrying this deletion failed to express *Tsix* from the mutant chromosome, levels of *Xist* expression were increased by some 50% from the mutant allele. Importantly though, these transcripts did not accumulate prior to differentiation. Furthermore, allele-specific RT-PCR and FISH demonstrated that at the onset of X-inactivation, *Xist* RNA was expressed exclusively from the deleted chromosome in heterozygous females, resulting in primary non-random inactivation. These observations indicate that *Tsix* functions as a repressor of *Xist in cis*, marking the future active X by blocking *Xist* accumulation in differentiating cells. Importantly however, the observation that deleting *Tsix* does not result in the accumulation of *Xist* RNA in undifferentiated cells implies that the involvement of additional factors, other than *Xist* and *Tsix,* are required for the initiation of X-inactivation. Similar to the model proposed by Marahrens *et al.* (1998), Lee and Lu hypothesised that *Xist* is regulated in a two-step pathway involving both positive and negative factors, as follows:

i) A unique 'blocking factor', present as a titrated complex of X-linked and autosomal factors, marks the future X_a by blocking *Xist* function, probably through *Tsix*.

ii) A 'competence factor' is also required to initiate X-inactivation on the future X_i, and probably acts to stabilise *Xist* transcripts. Composed of untitrated X-linked factors, this competence factor is therefore produced only in cells with more than one X per diploid genome.

In addition, while *Tsix* is required for the random choice aspect of X-inactivation, the deletion of *Tsix* created by Lee and Lu neither disrupted X chromosome counting, or affected proper silencing. Therefore chromosome counting, choice, and silencing are established as genetically distinct processes.

By searching for additional transcripts within the human *XIC*, Migeon *et al.* (2001) identified an overlapping transcript antisense to *XIST*, hence named *TSIX*. Strand-specific RT-PCR revealed *TSIX* is an apparently unprocessed, untranslated

transcript approximately 35kb in length, with a start site 27kb downstream of *XIST*. *TSIX* is expressed in cells of foetal origin from the inactive X chromosome, but not in adult somatic cells. Comparison with it's murine counterpart shows that *TSIX* and *Tsix* have little sequence homology. More importantly however, *TSIX* is truncated at it's 5' end and therefore is not linked to a CpG island that has been shown to be essential for normal *Tsix* function in the mouse (Lee and Lu, 1999; Lee, 2000; Sado *et al.*, 2001), and neither does its transcription overlap the promoter of *XIST*. Therefore, Migeon *et al.* concluded that the structural differences between *TSIX* and *Tsix* explain the lack of imprinting in human placental tissues, and further suggested that *TSIX* may merely be a functionless evolutionary relic with little or no role in humans.

Further studies of *TSIX* supported this hypothesis (Migeon *et al.*, 2002). *TSIX* and *XIST* are co-expressed in ES-derived cells with *TSIX* expression persisting in somatic cells for long periods, and do not show the dynamic relationship as seen in murine X-inactivation. Furthermore, *TSIX* expression is not imprinted, demonstrating that *TSIX* does not function as an antagonist to *XIST* and that the preferential silencing of the paternal X seen in the extra-embryonic tissues of the mouse does not have any human counterpart.

1.8 X-inactivation and histone macroH2A

The unique features of the X_i have led many authors to postulate the association of specific proteins with the chromatin of the Barr body (Willard and Salz, 1997). While studying the cellular distribution of a variant histone named macroH2A1.2, Costanzi and Pehrson (1998) consistently observed that the nuclei of female mammals contains a large, densely staining region, which they called the macrochromatin body (MCB). This MCB however was absent in the nuclei of normal males. Using chromosome-specific FISH probes, Costanzi and Pehrson demonstrated that the position of the MCB coincides with the X_i, leading to the suggestion that mH2A1.2 participates in the process of X-inactivation, perhaps by altering chromatin structure in association with *Xist* RNA.

Further studies of mH2A1.2 strengthened the case for its involvement in X-inactivation. Csankovszki *et al.* (1999) studied the effects of a somatic deletion of *Xist* in murine fibroblasts. Despite abolishing transcription of *Xist*, the mutant X_i remained

inactivated. However, MCBs never localised to this mutant chromosome. This observation indicated that *Xist* RNA mediates the preferential association of mH2A1.2 with the Xi, and that mH2A1.2 is not required for the maintenance of the inactive state. As mH2A1.2 is tightly bound to the nucleosome (Pehrson and Fried, 1992), while *Xist* RNA interacts with the nuclear matrix (Clemson *et al.*, 1996), Csankovski *et al.* suggested that an *Xist* RNA/mH2A complex might act as a mediator between the chromatin of the X_i and the nuclear matrix.

Mermoud *et al.* (1999) studied the distribution of macroH2A1.2 during the differentiation of murine ES cells. In contrast to somatic cells, they observed that an MCB is present prior to X-inactivation in both XX and XY ES cell lines, and did not associate with X chromatin. Following the initiation of X-inactivation at day 3 of differentiation, this pattern of mH2A1.2 staining remained unchanged. However, between days 7 to 14 of differentiation, the frequency of XY ES cells containing an MCB declined. In contrast, in XX cells there was a dramatic relocalisation of the MCB to the X_i between days 7 and 9. Furthermore, from days 12 to 16 following differentiation of these XX cells, additional MCBs were observed in DAPI (chromatin)-dense regions of each nucleus, with up to eight present in some cells. As the MCB localised to the X_i well after X-inactivation had occurred, Mermoud *et al.* concluded that macroH2A1.2 does not play a role in the initiation or propagation of the inactivation process, instead favouring its involvement in the maintenance of the inactive state. However, as indicated by the lack of colocalisation in a cancer cell line, this requirement is not absolute. Additionally, the presence of multiple non-X_i MCBs in chromatin-dense regions of many cells suggests that mH2A1.2 might play a more general role in heterochromatin formation (Costanzi and Pehrson, 1998). Further immunofluorescence studies demonstrated that in both male and female undifferentiated ES cells, the MCBs observed by Mermoud *et al.* actually represented concentration of mH2A1 at the centrosomes as a nonchromatin-associated pool (Rasmussen *et al.*, 2000). Rasmussen *et al.* showed that following the differentiation of female ES, accumulation of mH2A1 on the X_i corresponded with a loss of centrosomal mH2A1, suggesting that the centrosomes act as a store for mH2A1 prior to its incorporation into the nucleosomes.

By homology searching using the cDNA sequence of mH2A1, Chadwick and Willard (2001a) identified a second macroH2A variant named macroH2A2. mH2A2

shows 80% identity to mH2A1 and an almost identical genomic organisation, suggestive of a common evolutionary origin. Like mH2A1.2, mH2A2 also forms a macrochromatin body localised to the X_i during interphase, implying that the two may have overlapping and possibly redundant functions. The preferential association of these H2A variants with the X_i prompted Chadwick and Willard (2001a) to survey the distribution of other known histone H2 variants in relation to the Barr body. While H2A, H2A.X and H2B showed an apparently uniform distribution within the nuclei of female cells, H2A.Z was deficient on the X_i in a significant proportion of cells, suggestive of a role in the chromatin structure of the X_i.

Further investigations by Chadwick and Willard (2001b) identified a novel and more distantly related H2A variant which also showed a unique association with the X_i. Designated H2A-Bbd (Histone H2A variant, Barr-body deficient), this histone variant shows a diffuse localisation throughout interphase nuclei, but is excluded specifically from the region of the Barr body. Immunofluorescence investigations of metaphase spreads confirmed the exclusion of H2A-Bbd from most of the X_i, although some weak staining was apparent on regions of Xp and proximal Xq, similar to that seen for acetylated histone H4. Thus H2A-Bbd is apparently a counterpart to macroH2A with a mutually exclusive distribution, marking transcriptionally active chromatin.

1.9 X-inactivation and imprinting

Imprinting is a mechanism of gene regulation by which the expression of a gene is dependent upon the sex of the parent from which it is inherited. The first absolute demonstration of genomic imprinting in mammals was shown in 1991, when the selective maternal expression of *Igf2r* was shown in the mouse (Barlow *et al.*, 1991). As early as 1975, it was known that X-inactivation in the extra-embryonic tissues of the mouse showed a parental bias (Takagi and Sasaki, 1975). Similar preferential silencing of the paternal X chromosome is also seen in most, if not all tissues of marsupial mammals (Cooper, 1971), but evidence suggests that imprinted X-inactivation does not occur in humans.

The first demonstration of imprinted expression of *Xist* was made by Kay *et al.* (1993), who showed exclusive paternal *Xist* expression in very early pre-implantation

embryos. Exclusive expression of the paternal *Xist* allele continues in all cells of the extra-embryonic tissues. The deletion created by Marahrens *et al.* (1997) demonstrated that such preferential inactivation in the extra-embryonic tissues is a direct result of imprinted *Xist* expression. They observed that female mice that inherited the deleted *Xist* allele on the maternal X developed normally, whereas mice that inherited a paternal *Xist* deletion had poorly developed extra-embryonic membranes, and died soon after implantation. These results mirrored earlier observations made in gynogenetic and parthenogenetic mouse embryos, where both X chromosomes are of maternal origin. Here, X-inactivation either does not occur in the trophoblast, or does so only after a long delay (Shao and Takagi, 1990). In contrast, in androgenetic embryos both of the paternally derived X chromosomes are inactivated (Latham, 1996). Importantly, these observations suggest that imprinted X-inactivation differs from random inactivation, bypassing the counting and choice mechanisms and responding instead to parentally predetermined cues. They also indicate that it is the maternal *Xist* allele that carries some form of imprint to resist X-inactivation in the extra-embryonic tissues.

However, the fact that initial imprinted X-inactivation in the trophoectoderm is followed by random X-inactivation in the embryo proper indicates that any imprints on the maternal or paternal X chromosomes are developmentally labile, and are apparently lost specifically in the inner cell mass. So what might this imprint be? DNA methylation has been proposed by many as a likely candidate for imprinted X-inactivation, and is almost certainly involved in the imprinted expression of many autosomal genes (reviewed in Bartolomei and Tilghman, 1997). Additionally, the expression status of *Xist* in adult somatic tissues and in male and female germ cells correlates with its methylation status, implicating methylation as a controlling mechanism of *Xist* expression in these systems. In somatic cells the promoter of the expressed *Xist* allele on the X_i is hypomethylated, while the silent *Xist* allele on the X_a is hypermethylated (Norris *et al.*, 1994), and loss of this methylation reactivates *Xist* expression (Beard *et al.*, 1995). Also the *Xist* promoter of the inactivated X chromosome during spermatogenesis is demethylated, while the reactivated X of female germ cells has a methylated *Xist* allele (Ariel *et al.*, 1995). As a result the maternal *Xist* allele contributed to the zygote is fully methylated at its 5' end, while the

paternal *Xist* allele is hypomethylated and potentially open for transcription. Using restriction enzyme analysis, Zuccotti and Monk (1995) apparently demonstrated that this same pattern of methylation is retained in the extra-embryonic tissues that undergo imprinted X-inactivation in early embryogenesis. They concluded that this methylation difference is the imprint which dictates paternal X-inactivation in the trophoectoderm. However a re-examination of the methylation status of *Xist* in these tissues using an improved bisulphite-sequencing technique produced opposing results (McDonald *et al.*, 1998), and thus the involvement of methylation in imprinted *Xist* expression remains unclear. Monk and McLaren (1981) proposed the alternative hypothesis that the inactive, condensed state of the paternal X chromosome might be carried over from spermatogenesis, hence predisposing it to preferential inactivation in early differentiation.

A number of studies in humans have shown that, in contrast to the mouse, there does not appear to be exclusive inactivation of the paternal X chromosome in extra-embryonic tissues. By studying the allele-specific expression of an X-linked gene, Harrison (1989) demonstrated that the paternal X chromosome is often active in cells of human placental tissue, although an apparent excess of paternal inactivation was noted. *XIST* expression from the maternal X chromosome has also been demonstrated in early embryos and chorionic villi samples, apparently indicating that the X_i may be of maternal origin in these cells (Ray *et al.*, 1997). However, these studies utilised the highly sensitive techniques such as nested RT-PCR to assay for *XIST* transcription, and it remains possible that they are simply detecting the low levels of unstable *XIST* RNA produced prior to the initiation of X-inactivation (Beard *et al.*, 1995), rather than stable *XIST* transcripts which are associated with the X_i. Therefore, the results of these studies should be interpreted with caution. In contrast however, Goto *et al.* (1997) utilised the androgen receptor assay to show that the paternal X chromosome was preferentially, but not exclusively, inactivated in trophoblastic cells of the chorionic villi from two female conceptuses.

Further evidence of a lack of imprinting of X-inactivation in humans comes from studies of sex chromosome aneuploids. While the presence of additional maternal X chromosomes has profound effects in mouse embryos, a reported case of human maternal uniparental disomy of the X seemingly developed normally (Quan *et*

al., 1997). Additionally in cases of trisomy X and Klinefelter syndrome (47,XXY) in humans, the phenotype is identical irrespective of the parental origin of the supernumery X chromosome (MacDonald *et al.*, 1994). However in the mouse, where an additional maternal X chromosome is present, these same chromosome abnormalities lead to a complete failure in the development of the extra-embryonic membranes resulting in lethality, whereas additional paternal X's are tolerated (Tada *et al.*, 1993).

While studying the role of a gene named *Eed* in murine embryonic development, Wang *et al.* (2001) observed that while female *Eed* double-knockout embryos showed severe defects in trophoblast development, male embryos with the *Eed*-null mutation developed apparently normally. By cross-breeding a GFP (Green Fluorescent Protein) transgene onto the paternal X chromosome of *Eed*-null female embryos, Wang *et al.* observed that shortly after the initiation of X-inactivation, a progressive reactivation of the paternally-derived X occurred in the extra-embryonic tissues of these embryos. As Eed has been shown to reside in a complex with histone deacetylase (van der Vlag and Otte, 1999), Wang *et al.* therefore concluded that *Eed* acts to maintain imprinted X-inactivation in extra-embryonic tissues by mediating hypoacetylation of the paternal X chromosome.

1.10 Skewing of X-inactivation

Non-random or skewed X-inactivation is the predominant or exclusive silencing of the same X chromosome in all cells of a tissue, and can result from a number of different causes. Primary non-random inactivation is the preferential choice to inactivate the same X chromosome in all cells, and implies a distortion in the randomness of the X-inactivation process itself. In contrast, secondary non-random inactivation implies an initial random choice of inactivation, followed by post-inactivation selection for cells which carry either an active or inactive X chromosome. Both of these mechanisms operate in the mammalian embryo. An example of primary skewing is the imprinted nature of X-inactivation seen in somatic tissues of marsupials and the extra-embryonic cells of mice, while secondary selection effects have been shown to operate in cases of mutations and structural abnormalities of the X chromosome. Skewed X-inactivation is an important component in determining the clinical expression of X-linked disease in

heterozygous females. Partial or full expression of a disease allele is occasionally observed in such carriers, resulting from the majority of cells in an affected tissue having inactivated the wild-type allele. The effect of this skewed X-inactivation may therefore be to create functional nullisomy for certain X-linked genes in genotypically heterozygous individuals.

Skewed X-inactivation may also occur as a purely stochastic process. X-inactivation has been shown to initiate at about the 32-64 cell stage of embryogenesis (Lyon, 1972), and estimates of the actual number of progenitor cells destined to form any one tissue put this figure at between 10 and 20 (Fialkow, 1973). As the choice of which X chromosome to inactivate is thought to be a purely random event in humans, it follows that the number of cells that have inactivated the same X chromosome in any one embryo should follow a normal distribution. As a result of the relatively small number of progenitor cells for each tissue, this predicts that a proportion of females will have a skewed pattern of X-inactivation purely by chance.

This theoretical prediction has since been confirmed by a number of studies. Using the difference in the electrophoretic properties of isoforms of the X-linked enzyme G6PD in heterozygotes, Nance (1964) was the first to demonstrate that a significant proportion of normal women have skewed patterns of X-inactivation. More recently, a variety of techniques have been developed which rely on the differential sensitivity of the active and inactive X chromosomes to methylation-sensitive restriction enzymes. When combined with restriction-fragment length polymorphisms such as those found in the *PGK* and *HPRT* genes or variable repeat numbers such as the locus DXS255, this differential methylation allows the relative expression of the two X chromosomes in a sample to be quantified molecularly. However, data obtained by different investigators using these techniques has produced seemingly contradictory results. Vogelstein *et al.* (1987) studied a population of 81 normal females using methylation-sensitive assays at the *PGK* and *HPRT* loci, and found only 3 (4%) with significantly skewed X-inactivation, which he defined as being 'greater than 75% expression of one allele'. Despite its arbitrary nature, this definition of skewing has since found common use in many other studies of X-inactivation. Using the same techniques however, Gale *et al.* (1991) studied 65 apparently normal females and

found that 15 (23%) had skewed X-inactivation ratios in blood cells (as defined by Vogelstein *et al.*). And by these same criteria, Fey *et al.* (1992) concluded that 33% of their study population exhibited skewed X-inactivation patterns in blood, as determined by use of the probe M27β at the DXS255 locus. Further investigations have since concluded that the prevalence of skewed X-inactivation lies somewhere between these extremes (Harris *et al.*, 1992; Plenge *et al.*, 1997), but the reasons for such large discrepancies between studies remain unclear.

One possible explanation for these differences was first provided by Fey *et al.* (1994). During a survey of the X-inactivation clonality patterns in women with lymphoid disease, they observed that highly skewed X-inactivation was more frequent in the blood leukocytes taken from an elderly control population than in their young controls. This finding was apparently confirmed by Busque *et al.* (1996), who observed skewed X-inactivation patterns (ratios ≥75:25%) in 9% of neonates, 16% of 28-32 year olds, and 38% of women over 60 years of age. In contrast however, earlier studies had failed to find any evidence of an age-related effect on skewing (Gale *et al.*, 1994), although this may in part be due to the small sample sizes used. A further study by Gale *et al.* (1997) found that such age-related increases in skewing of X-inactivation were present in a myeloid cell line, but absent from lymphoid cells. The authors suggested that this difference in the degree of skewing seen in the two cell types was a reflection of their relative longevity – neutrophils being short-lived whereas T-lymphocytes are mainly long-lived. As a result, neutrophil stem cells must undergo much more rapid division, which Gale *et al.* suggested leads to the outgrowth of a small number of stem cell clones, resulting in skewed X-inactivation in many instances.

In addition to these apparent variations in the degree of skewing of X-inactivation between individuals of different ages, two studies have suggested that similar differences can occur among tissues within an individual. Gale *et al.* (1994) compared the X-inactivation patterns in blood, skin and muscle samples collected from twenty normal females, and found that while the results were similar for all tissues in eleven of their patients, in the remaining nine women they observed significant differences among tissues. While there was no consistent pattern to this in different

individuals, they most commonly observed the presence of skewed X-inactivation in peripheral blood, but not in skin or muscle tissue. A subsequent study by Azofeifa *et al.* (1996) similarly surveyed the X-inactivation ratios in blood leukocytes, muscle, thyroid gland and suprarenal gland taken from ten deceased women. These cells represent descendants from the three embryonic tissue layers - the ectoderm, mesoderm and endoderm, which are the first to differentiate in early embryonic development. As had been reported by Gale *et al.* (1994), Azofeifa *et al.* observed that skewed X-inactivation was more common in blood leukocytes than in other tissues. In contrast though, Azofeifa *et al.* found that X-inactivation ratios in the other tissues they examined were fairly well correlated in their study group.

It should be noted however that both of these studies utilised a number of different techniques to measure the X-inactivation ratios in their populations, which have been reported to give conflicting results in a proportion of cases (Fey *et al.*, 1994). Of the 20 individuals studied by Gale *et al.* (1994), 10 were analysed using the M27β probe, 9 at the *PGK* locus, and one at the *HPRT* locus. Similarly, Azofeifa *et al.* (1996) utilised the M27β and *PGK* probes in different individuals. In addition, a further potentially confounding element in the study by Gale *et al.* was the wide variation in age between individuals of their study population, which ranged from 18 to 96 years.

In addition to these observations of tissue-to-tissue variations of X-inactivation within individuals, anecdotal evidence suggests that X-inactivation ratios can vary even between different regions of the same organ in some females. During a study of the X-linked liver disease ornithine carbamoyltransferase deficiency, Reyal *et al.* (2000) noted marked variations in the X-inactivation ratios between 52 separate biopsies across different regions of the liver from one individual, ranging from 85:15 55:45.

1.11 Familial skewed X-inactivation

While the majority of cases of skewed X-inactivation are probably attributable to stochastic factors, several reports of multiple affected females in families in which X-linked syndromes are segregating suggests that, at least in some cases, skewed X-inactivation may be a heritable trait (Ropers *et al.*, 1977; Reddy *et al.*, 1984; Taylor *et al.*, 1991). It is known that in the mouse a locus called the X-controlling element

(*Xce*) influences the random choice of which X chromosome is inactivated to varying degrees (Cattanach and Williams, 1972). Three different *Xce* alleles have been identified to date in the laboratory mouse *Mus domesticus*, named Xce^a, Xce^b and Xce^c (Simmler *et al.*, 1993), and a fourth allele (Xce^d) in feral mouse species (Cattanach and Rasberry, 1994). Each of these alleles is of different 'strength' with regards to its probability of being inactivated, with Xce^a being the strongest and Xce^d the weakest. In mice heterozygous for Xce^a/Xce^b the X chromosome carrying the Xce^a allele is more likely to be inactivated than the X carrying the Xce^b allele, and similarly in Xce^b/Xce^c heterozygotes the X chromosome carrying the Xce^b allele is more likely to be inactive in the majority of cells. Most extreme non-random inactivation is seen in Xce^a/Xce^c heterozygotes.

By studying early mouse embryos which were initiating X-inactivation, Rastan (1982) demonstrated that the *Xce* operates through primary non-random inactivation, influencing the actual choice of which X chromosome is inactivated, rather than through secondary selection events. However, the actual mechanism by which it exerts its influence is unclear. Early mapping data localised the *Xce* locus to a region similar to that of the *Xic*, apparently in the locality of the *Xist* gene itself (Cattanach and Williams, 1972). More recently though it has been shown that the *Xce* and *Xist* are distinct entities, with the *Xce* lying 3' to *Xist* (Simmler *et al.*, 1993). Courtier *et al.* (1995) analysed this region and found that sites ~15kb distal to *Xist* show hypermethylation specific to the X_a, with the degree of methylation apparently correlating with the strength of the *Xce* allele on that chromosome. They proposed that such methylation might precede X-inactivation, thereby influencing the choice of which X chromosome to inactivate. Interestingly, this region of differential methylation identified by Courtier *et al.* has since been shown to coincide with the promoter of the anti-sense transcript *Tsix* (Lee *et al.*, 1999). While this provides an obvious mechanism by which the choice of X-inactivation could be influenced, it is also possible that such X_a-specific methylation may instead be imposed after the X-inactivation process, acting to repress *Tsix* transcription which is abolished on the X_a following the upregulation of *Xist* transcription. Quantitative studies of the levels of *Xist* transcript produced from different *Xce* alleles have also failed to find any significant or consistent differences which might contribute to its effects (Buzin *et al.*,

1994).

While there is no evidence for the existence of a homologue of the *Xce* in humans, several families have been identified in which skewed X-inactivation appears to be inherited between generations. In a screen for skewed X-inactivation in a large number of normal females using the PCR-based androgen receptor (*AR*) assay, Naumova *et al.* (1996) found one family in which all seven daughters and their paternal grandmother showed completely skewed X-inactivation patterns. Naumova *et al.* observed that while the same X chromosome was preferentially inactive in all seven of the daughters, the grandmothers X-inactivation profile was discordant with the skewing in the apparent reverse direction. This indicated that a recombination event had occurred between the *AR* locus used to assay the X-inactivation and the presumed *cis*-linked locus which was responsible for the skewing. On this basis Naumova *et al.* excluded linkage to the *XIST* gene as a candidate for the skewed X-inactivation in this pedigree, but the presence of only a single informative meiosis precluded further analysis.

In contrast Plenge *et al.* (1997) found two unrelated families in which all females showed preferential inactivation of the same X chromosome to varying degrees, and showed by haplotype analysis that this trait was linked to the *XIC*. By direct sequencing, Plenge *et al.* detected a single base-pair C to G mutation in the minimal promoter of the *XIST* gene in both families, and showed that this mutation was present on the preferentially inactive X chromosome in each case. This mutation was shown to be absent from over 1000 normal controls, and therefore does not represent a common polymorphism. Use of this mutant promoter to create a transgenic reporter construct demonstrated that it had only 20-50% the activity of the wild-type *XIST* promoter, establishing a functional mechanism through which this mutation might influence the random choice of X-inactivation.

A further case of familial skewed X-inactivation was detected by Pegoraro *et al.* (1997) in a large kindred with over 50 females. Sixteen of these women demonstrated complete skewing of X-inactivation, with preferential inactivation of the maternal X chromosome in every case. By linkage analysis this trait was localised to Xq28, where an ~800kb deletion of the Factor VIII gene was detected which co-

segregated perfectly with the skewed X-inactivation. Pegoraro *et al.* suggested that the Factor VIII mutation they detected is either lethal, or results in a severe growth disadvantage, to cells which carry it. As a result, heterozygous females show a loss of any cells which carry the mutation on their X_a, while in hemizygous males the mutation would be lethal. Consistent with this model Pegoraro *et al.* observed no males carrying the mutation, while all heterozygous females preferentially inactivated the deleted X chromosome and showed an increased rate of spontaneous abortion when compared to normal controls.

These studies serve to illustrate that the causes of heritable skewed X-inactivation are most likely heterogeneous in nature, with distinct causative loci and mechanisms of action occurring in separate pedigrees.

1.12 Skewed X-inactivation and secondary cell selection

Skewed X-inactivation often results due to selective pressures which tend to increase the relative abundance of cells which are expressing one or other X chromosome. This is most commonly seen in individuals with structurally abnormal X chromosomes, such as deletions, isochromosomes or ring chromosomes. In nearly all such cases, the structurally abnormal X chromosome is inactive in most or all cells, with the non-random inactivation thus acting to minimise the potential genetic imbalance which would otherwise be associated with the chromosome abnormality (Therman and Patau, 1974).

Similarly, non-random inactivation is also seen in the case of X;autosome translocations. Where such translocations are balanced, then the normal X chromosome is preferentially inactivated in most cases, with the two separate portions of the translocated X remaining active and thereby maintaining a single active X chromosome in each cell (Schmidt and DuSart, 1992). Conversely, in the case of unbalanced X;autosome translocations it is the translocation product that is invariably inactivated. This has been shown to result from a post-inactivation selection process against those cells which have retained an active abnormal X chromosome (McMahon and Monk, 1983). However, the degree of skewing may be influenced by both the nature of the translocated autosomal material, and the extent to which the inactivation signal may spread into it (Mattei *et al.*, 1982). Therefore, as with structurally

abnormal X chromosomes, the pattern of inactivation in X;autosome rearrangements also acts to minimise the consequences of the chromosome defect.

In addition to cell selection occurring where gross chromosomal defects are involved, secondary cell selection leading to skewed X-inactivation has also been suggested to occur for a number of X-linked single gene mutations. Many such instances of recessive X-linked disorders that give rise to skewed X-inactivation in the blood of asymptomatic carrier females now exist, such as Wiskott-Aldrich syndrome, dyskeratosis congenita, and focal dermal hypoplasia (Wengler *et al.*, 1995; Devriendt *et al.*, 1997; Gorski *et al.*, 1991). As with other structural abnormalities of the X chromosome, this selection acts to reduce the proportion of cells which possess a deleterious genotype.

In contrast to all other reports of skewing being the result of a negative selection process acting against certain clones, Migeon *et al.* (1981) proposed that positive selection favouring the mutant allele occurs in heterozygous carriers of adrenoleukodystrophy. However this conclusion was based largely on observations made in fibroblasts which had been cultured *in vitro*, a process which has since been demonstrated to artifactually influence X-inactivation ratios substantially (Hatchwell *et al.*, 1996). Furthermore an additional and much larger *in vivo* study of X-inactivation in adrenoleukodystrophy carriers found no evidence of positive selection for cells expressing the mutant allele (Watkiss *et al.*, 1993). Therefore, it seems unlikely that any such positive selection occurs.

1.13 Skewed X-inactivation due to reduction in precursor cell numbers

The limited number of cells present in the mammalian blastocyst at the time X-inactivation occurs (Lyon, 1972) means that any event leading to a reduction in the number of cells which go on to form the embryo proper will result in an increased probability of non-random X-inactivation. Thus, observations of skewed X-inactivation can be indicative of a "bottle-neck" in the embryonic precursor cell pool, and are associated with a number of developmental events.

Twinning – Female monozygous (MZ) twins discordant for Duchenne muscular dystrophy (DMD) were first described by Gomez *et al.* (1977), who hypothesised that the observed phenotypes were the result of differential X-inactivation of the normal and mutant alleles in each individual. Since then, there have been numerous other reports of female MZ twins discordant for X-linked disorders (Burn *et al.*, 1986; Richards *et al.*, 1990; Lupski *et al.*, 1991; Jorgensen *et al.*, 1992; Winchester *et al.*, 1992), suggesting a link between skewed X-inactivation and the formation of MZ twins. By re-evaluating observations of mosaicism of dystrophin-negative muscle fibres in MZ twins affected with DMD in the light of twin embryology, Nance (1990) suggested a mechanism that could explain the apparently high frequency of discordant skewed X-inactivation observed in MZ twins. Nance hypothesised that the X-inactivation process is initiated before the formation of MZ twins, which occurs by the division of a blastocyst into two separate embryos. As the commitment to inactivate the paternally- or maternally-derived X chromosome has already been made in each cell prior to this division, the splitting of an embryo in this way effectively reduces the number of progenitor cells contributing to each embryo, thereby increasing the probability that the resulting twins will have the same X chromosome inactive in an unequal proportion of cells. Thus, the process of MZ twinning will tend to produce more extreme X-inactivation ratios than were present in the blastocyst prior to the twinning event, especially where the splitting of the cell mass occurs unequally. As Nance pointed out, despite arising from approximately half the number of progenitor cells as singletons, MZ twins are not born at half the normal size, and thus show considerable "catch-up" growth. Singleton births heterozygous for an X-linked mutation often show a mosaic pattern of gene expression, with small areas of tissue derived from each progenitor cell with one or other X chromosome inactive. Nance reasoned that the reduced number of progenitors and subsequent catch-up growth of MZ twins should result in an increased size of each of these tissue patches. This is consistent with the histological studies of Burn *et al.* and Richards *et al.* who observed unusually large clumps of normal and diseased tissue in the muscle fibres of MZ twins discordantly affected with DMD.

Data to support the hypothesis that MZ twinning occurs after the commitment to X-inactivation came from two studies which examined X-inactivation ratios in a large number of twins (Bamforth *et al.*, 1996; Goodship *et al.*, 1996). While

discordant patterns of X-inactivation were not commonly observed in MZ twin pairs, extremely skewed X-inactivation ratios did occur significantly more frequently in MZ twins than dizygous twins or controls, consistent with the notion that MZ twinning is a relatively late embryological event.

Trisomy rescue – Lau *et al.* (1997) hypothesised that confined placental mosaicism (CPM) is a deleterious condition that leads to poor embryonic growth and a restriction in the number of foetal precursor cells. By surveying patterns of X-inactivation in foetuses and new-borns associated with CPM for an autosomal trisomy, they observed that 7 of 12 cases (58%) where the trisomy was of meiotic origin showed extremely skewed X-inactivation ratios. Similarly skewed X-inactivation was also observed in 3 of 7 cases (43%) of maternal uniparental disomy (matUPD) of chromosome 15 which also arose meiotically, consistent with the notion that a significant proportion of cases of trisomy rescue are associated with skewed X-inactivation. Lau *et al.* suggested that this resulted from either the trisomy rescue event occurring in a single or small number of cells after the commitment to X-inactivation had been made, or alternatively as a result of poor growth by and/or selection against trisomic cells in early embryogenesis causing a reduction in the precursor cell pool size. The association of skewed X-inactivation in cases of rescue from autosomal trisomy of meiotic origin was further strengthened by two additional studies which similarly found a significant association between skewed X-inactivation and CPM for trisomy 16 (Penaherrera *et al.*, 2000) and matUPD15 (Robinson *et al.*, 2000), suggesting that CPM may be a relatively common cause of skewed X-inactivation.

1.14 The spreading of X-inactivation in X;autosome rearrangements

By studying a number of murine X;autosome translocations, Russell (1963) observed phenotypic variegation that was coincident between the translocated autosomal chromosome and the coat colour genes located upon it. Russell hypothesised that the phenotype of these mice was caused by a variable spread of the X-inactivation signal into the translocated autosomal segment, suppressing expression of coat colour alleles upon it. Further phenotype studies suggested that the spread of X-inactivation into

autosomal DNA occurred in a graded fashion limited in distance. Additionally, Russell observed that this variegated phenotype only occurred in the presence of multiple X chromosomes with the translocation involving certain regions of the X chromosome, consistent with both the 'n-1' rule and the later discovery of the *XIC* (Therman *et al.*, 1974; Brown *et al.*, 1991a).

Shortly after the first confirmation of a human X;autosome translocation by banding techniques (Buckton *et al.*, 1971), Allerdice *et al.* (1971) utilised autoradiography to study the replication characteristics of an unbalanced X;14 translocation. Previous replication timing studies based on the differential incorporation of tritium-labelled thymidine had shown the X_i to replicate later in the cell cycle than either the X_a or the autosomes (German, 1964). Using this technique, Allerdice *et al.* observed that in their case the entire der(X), including the chromosome 14 segment, was apparently late-replicating, demonstrating cytologically the spreading of the X-inactivation signal into autosomal chromatin. In 1973, Latt developed an improved technique for studying chromosome replication timing based on the incorporation of BrdU, a thymidine analogue, which allowed the spreading of late-replication to be studied more easily and at much higher resolution than could be achieved using ^{3}H-thymidine. Based largely on the presumption that the spread of late-replication is a defining measure of the spread of X-inactivation, observations in over 60 independent cases of X;autosome translocation have since been reported (Leisti *et al.*, 1975 and references therein; Camargo and Cervenka, 1984 and references therein; Disteche *et al.*, 1984; Keitges and Palmer, 1986; Williams and Dear, 1987; Caiulo *et al.*, 1989; Schanz and Steinbach, 1989; Kulharya *et al.*, 1995; Preis *et al.*, 1996; Jones *et al.*, 1997; Canun *et al.*, 1998; Eggermann *et al.*, 1998). In reviews of these cases, Leisti *et al.* and Camargo and Cervenka noted the high degree of variation in the X-inactivation pattern, the extent of spread of late-replication, and the associated clinical phenotype of carrier individuals. However, as pointed out by Camargo and Cervenka, the majority of cases studied showed limited spreading of late-replication, usually extending only a short distance into the attached autosomal chromatin. Additionally in cases studied by Mohandas *et al.* (1981) and Keitges and Palmer (1986), spreading of late-replication occurred in a discontinuous fashion, skipping some autosomal bands and affecting others more distally, suggesting that autosomal chromatin does not either transmit or maintain the inactivation signal as efficiently as the X chromosome.

In 1979, Couturier *et al.* provided the first direct evidence to support the presumed relationship between the spread of late-replication into autosomal DNA and the inactivation of genes upon it. Their observations in a carrier of an unbalanced X;21 translocation showed a delayed replication timing of the translocated segment of chromosome 21 intermediate between that of the inactive X and the normal chromosome 21 homologues, accompanied by an attenuated phenotype. Consistent with this, biochemical studies of superoxide dismutase, a gene located upon the translocated region of chromosome 21, showed it to have only mildly increased activity that was much lower than that normally seen in individuals with trisomy 21. Three similar studies in other X;autosome translocation carriers also found corresponding correlations between the spread of late-replication, attenuation of clinical phenotype and the inactivation of single autosomal genes (Mohandas *et al.*, 1981; Mohandas *et al.*, 1982; Taysi *et al.*, 1982).

By isolating an X;4 translocation chromosome using somatic-cell hybrids, White *et al.* (1998) performed detailed studies of the spreading of inactivation, analysing 20 genes along the length of the translocated autosomal region by RT-PCR. Their results showed that while 14 of the genes analysed on the translocated 4q segment were inactivated, 6 remained expressed, apparently escaping the spread of X-inactivation. Although the X-inactivation signal had spread along virtually the entire 4q segment covering a distance of approximately 100Mb, these active genes were interspersed among groups of inactive genes, demonstrating the ability of X-inactivation to spread discontinuously as indicated by earlier replication timing studies (Mohandas *et al.*, 1981; Keitges and Palmer, 1986). Because of the relatively high proportion of autosomal genes escaping X-inactivation in their case (6 out of 20), White *et al.* concluded that autosomal DNA must in some way differ from the X chromosome in that it is relatively resistant to the X-inactivation signal.

Further evidence to support this view came from two cytogenetic studies of X;autosome rearrangements. Duthie *et al.* (1999) studied the spread of histone hypoacetylation and *Xist* RNA into autosomal chromatin in two murine X;autosome rearrangements. Both cases showed little or no spreading of these two features in the majority of cells analysed. Similar studies in two human X;autosome rearrangements (Keohane *et al.*, 1999) also found no spreading of histone hypoacetylation, late replication or *XIST* RNA into the attached autosomal chromatin, with all three

parameters defining the breakpoint between the X and the autosomal material.

1.15 Aims of the study

Variations of X-inactivation ratios in normal women – Previous studies have reported conflicting data on the prevalence of skewed X-inactivation in normal women. In particular, it has been suggested that X-inactivation ratios may change both with age (Fey *et al.*, 1994; Busque *et al.*, 1996) and between different tissues within an individual (Gale *et al.*, 1994; Azofeifa *et al.*, 1996). It was therefore intended to study X-inactivation ratios in a number of groups of normal women. Using DNA extracted from blood samples, both young (defined as ≤20 years) and elderly (defined as ≥60) populations were studied. Each was composed of approximately 100 individuals. Study of these would define both the distribution of X-inactivation patterns in normal women, and how this might change with age. These data would also provide a valuable control population to test the involvement of skewed X-inactivation in genetic disease.

Two previous studies have examined variation of X-inactivation ratios in different tissues, and reported contradictory results. Both of these studies utilised small numbers of individuals, some of whose ages varied widely, and used a number of different techniques within each population. In order to clarify this issue, investigation of tissue-specific X-inactivation ratios was also conducted using two further populations, each of approximately 50 normal females, aged ≤25 and ≥60. It was intended that from each individual samples of peripheral blood, buccal epithelia, and urine would be obtained. It was hoped that these three samples would contain cells descended from the three embryonic tissue layers - the mesoderm, ectoderm, and endoderm respectively, and thus provide representatives from all portions of the early embryo.

Previous studies have also indicated that skewed X-inactivation may be an inherited trait. In some instances, this may be due to secondary selection against X chromosomes which are carrying deleterious or lethal mutations. The data of Pegoraro *et al.* (1997) suggests that deleterious mutations that cause severely skewed X-inactivation may also cause embryonic lethality in hemizygous male offspring who

inherit them. It was therefore also intended to determine X-inactivation ratios in the relatives of females in the control populations. Correlation of X-inactivation ratios between first-degree relatives would indicate the existence of heritable loci that influence skewing, as is seen with the *Xce* in mouse. Furthermore, study of offspring ratios in women with severely skewed X-inactivation would determine if their skewing were the result of selection against lethal X-linked alleles.

Trisomy mosaicism in Silver-Russell syndrome – Silver-Russell syndrome (SRS) is a malformation syndrome characterised by severe pre- and postnatal growth retardation, asymmetry, craniofacial abnormalities, and other more variable features. Although the syndrome is probably heterogeneous in nature (Wakeling *et al.*, 1998), maternal uniparental disomy for chromosome 7 (matUPD7) is associated with approximately 7% of SRS patients (Kotzot *et al.*, 1995; Eggermann *et al.*, 1997; Preece *et al.*, 1997), leading to the assumption that one or more imprinted genes on chromosome 7 are responsible for at least a proportion of cases. More recently, the identification of duplications of proximal chromosome 7p in SRS patients (Joyce *et al.*, 1999; Monk *et al.*, 2000) suggests specifically the over-expression of gene(s) located within this region as the pathogenic mutation. However, observations in a familial case that both maternal and paternal inheritance of the region 7p12.1-p13 results in an SRS-like phenotype suggests that the causative genes may not be imprinted (Joyce *et al.*, 1999). Joyce *et al.* therefore proposed that SRS might result from the presence of an additional copy of proximal chromosome 7p genes, either as a result of sub-microscopic duplications of this region, or alternatively from undetected mosaic trisomy of chromosome 7.

Duplications involving a number of candidate genes within proximal 7p have been excluded from 87 SRS probands, (Joyce *et al.*, 1999; Monk *et al.*, 2000; Martinez *et al.*, 2001; S. Mergenthaler, personal communication) and thus do not represent a common cause of the disease. In addition, mutation and imprinting analyses of numerous candidate genes have also failed to uncover any significant defects in SRS (Abu-Amero *et al.*, 1997; Rieswijk *et al.*, 1998; Eggermann *et al.*, 1999; Mergenthaler *et al.*, 2000a; Mergenthaler *et al.*, 2001; Eggermann *et al.*, 2001) and thus the underlying genetic cause remains unknown. However, the alternative mechanism proposed by Joyce *et al.*, namely the possible involvement of mosaic

trisomy 7 as a cause of SRS, has not been adequately addressed. Every case of complete matUPD7 could have arisen by trisomy rescue (Mergenthaler *et al.*, 2000b), and trisomy 7 cells, apparently confined to the placenta, have been detected in five matUPD7 SRS cases (Mergenthaler *et al.*, 2000; Lau *et al.*, 1997; Kotzot *et al.*, 2000). Thus, it is possible that mosaicism for trisomy 7 contributes towards the phenotype in some cases of SRS, and that the occurrence of matUPD7 is an additional finding caused by the loss of the paternal chromosome 7 during trisomy rescue, which would be expected in one third of such cases (Spence *et al.*, 1988).

Previous reports have demonstrated that completely skewed X-inactivation is frequently observed in the diploid tissues of individuals with mosaic trisomy that has originated from a trisomy rescue event (Lau *et al.*, 1997; Penaherrera *et al.*, 2000; Robinson *et al.*, 2000). In these cases, the detected trisomic cells are often confined to the placenta, and skewed X-inactivation presumably occurs as a result of a reduction in the size of the disomic cell pool contributing to the developing foetus, because of either poor growth by, or selection against, the trisomic cells. Thus, skewed X-inactivation can act as a marker of clonality following a trisomy rescue event which has occurred, perhaps in only a single precursor cell, during foetal development. As most cases of matUPD7 probably originate from trisomy rescue, it was hypothesised that if trisomy mosaicism was involved in the aetiology of SRS, the frequency of skewed X-inactivation should be increased in a population of SRS patients when compared to controls. In order to test this hypothesis, an analysis of the X-inactivation patterns in a cohort of female SRS patients of unknown aetiology was performed, and the results compared to those obtained from normal females of similar age.

The spreading of X-inactivation in X;autosome translocations – Since Russell's (1963) observation that in cases of X;autosome translocation, X-inactivation is capable of spreading into attached autosomal DNA, numerous other studies have confirmed and extended these early observations (Reviewed in Camargo and Cervenka, 1984; Mattei *et al.*, 1982). However, in all but a few cases evidence for the extent of spread of X-inactivation into autosomal DNA has been based upon either replication timing studies or the associated clinical phenotype in carrier individuals. Direct evidence to support the presumption for an absolute correlation between the extent of spread of late-replication and transcriptional silencing in X;autosome translocation carriers is

weak. In particular, sporadic reports of individuals with unexpectedly mild phenotypes in which there is no detectable spread of late-replication suggests that at least in some cases late-replication may be a poor correlate of the spread of X-inactivation (Keitges and Palmer, 1986; Bettio *et al.*, 1994; Garcia-Heras *et al.*, 1997; Bacino *et al.*, 1999; Keohane *et al.*, 1999; Abuelo *et al.*, 2000).

It was therefore intended to perform detailed studies of the spreading of X-inactivation into autosomal DNA in a number of cases of X;autosome translocation. In each case, isolation of transcribed polymorphisms within the translocated autosomal segment would allow the expression status of individual genes along the translocated segment of autosome to be determined by allele-specific RT-PCR. A combination of immunofluorescence and FISH using whole chromosome paints would then allow the visualisation of any accompanying spread of cytogenetic hallmarks of the X_i, such as histone acetylation, histone methylation and late-replication, into the translocated autosomal chromatin. Similarly, it was also intended to study the distribution of *XIST* RNA on the der(X) in each case by a combination of RNA and DNA FISH. Correlation of these cytogenetic observations with the spread of X-inactivation as determined by gene expression studies would then allow an assessment of each parameter both against one another, and as potential predictors of the phenotype associated with each chromosome imbalance. Such studies might also provide additional insights into how the X-inactivation signal is spread along whole chromosomes.

Furthermore, previous observations of discontinuous spreading of late-replication (Mohandas *et al.*, 1981; Keitges and Palmer, 1986) and gene silencing (White *et al.*, 1998) in X;autosome rearrangements suggests that some autosomal chromatin lacks features important in the spread and/or maintenance of X-inactivation. Recently Mary Lyon proposed LINEs as candidates for these 'booster sequences', a hypothesis supported by their distribution on the X chromosome (Bailey *et al.*, 2000). It was therefore also intended to perform a sequence analysis using data from the Human Genome Mapping Project in an attempt to identify sequence features that might correlate with the spread of X-inactivation, with a particular emphasis on LINEs and other repeat sequences.

Section 2 - Subjects, Materials and Methods

2.1 X-inactivation ratios in normal women

2.1.1 Variations of skewed X-inactivation with age

Two cohorts of unrelated females were selected from patients referred to the Wessex Regional Genetics Laboratory for a variety of reasons between 1985 and 1997. All samples were anonymised prior to their use in this study. Population A consisted of 98 females aged less than 20 years old (Mean age 7.3 years, standard deviation 5.5 years), as listed in Appendix 1. Of these 43 had been referred with a diagnosis of Prader-Willi syndrome, 19 for Angelman syndrome, 29 for cystic fibrosis, and 7 for polyposis coli screening. Population B consisted of 115 women aged over 60 years old (Mean age 70.0 years, standard deviation 6.0 years), listed in Appendix 2. Of these 60 had been referred for screening of genes involved in the genesis of breast cancer and 55 were members of Huntington disease pedigrees. None of these conditions are known to have any influence on the X-inactivation ratio. Stored DNA samples from each individual had been extracted from peripheral blood collected with EDTA or lithium-heparin anticoagulants using the following salt-precipitation procedure:

1. In a fume cabinet the blood sample was poured into a 50ml centrifuge tube and 4 volumes of ice cold sterile distilled water added to lyse the red blood cells.
2. The sample was centrifuged at 3000rpm for 10 mins at 4°C (Sorvall RT6000B, DuPont) to precipitate the remaining intact white cells.
3. Following centrifugation the supernatant containing lysed cell debris was removed into a 2% solution of Virkon disinfectant (Antec International).
4. The cell pellet was re-suspended in 2 volumes of ice cold Sucrose Lysis Buffer (SLB) and centrifuged at 3000rpm for 10 mins at 4°C to collect the cell debris.
5. The supernatant was removed into a 2% solution of Virkon.
6. The pellet was re-suspended in 300µl of Resuspension Buffer (RSB) per ml of

original blood sample, and 15µl of 10% Sodium Dodecyl Sulphate (SDS) (Sigma) added per ml of blood to lyse the cells, and remove the lipid fraction and soluble proteins. 8µl of Proteinase K solution was also added to digest the remaining cell proteins.

7. The sample was mixed by agitation and incubated overnight at 37°C.
8. 150µl of 6M NaCl (Sigma) per ml of blood was added to precipitate the non-DNA components and tubes shaken for 20 seconds.
9. Samples were centrifuged at 3500rpm for 30 mins at room temperature to collect the precipitate.
10. Following centrifugation the supernatant was collected into a 15ml centrifuge tube, two volumes of 100% ethanol (Sigma) added to precipitate the DNA, and the tube inverted until DNA visibly collected as a 'hair ball'.
11. Using a sterile needle, the DNA pellet was removed, washed in 1ml of 70% ethanol to remove residual salts which might adversely affect future manipulations, and transferred to a DNA tube containing between 50-500µl of 1xTris-EDTA (TE) Buffer (Sigma) depending upon the yield of DNA, which improves the stability of stored DNA. If no visible precipitate was observed at this point, the sample was incubated at -70°C for 30 mins to encourage precipitation of nucleic acids, and then centrifuged at 14000rpm for 20 mins at 4°C. Supernatant was then removed and the remaining DNA pellet re-suspended in 50-500µl of 1xTE Buffer .
12. The DNA pellet was re-suspended by incubating at 37°C for 1 hour. DNA samples were then kept at -20°C for long-term storage, or temporarily at 4°C whilst being studied.

(Adapted from Miller *et al.*, 1988)

2.1.2 Tissue-specific variations of X-inactivation

Two further populations of females were recruited from individuals routinely referred to Salisbury District Hospital Pathology Department between April 1998 and May 1999. The first cohort (Population C) consisted of 50 women aged 25 or under (Mean age 21.5 years, standard deviation 3.3 years), listed in Appendix 3. The second (Population D) consisted of 51 women aged 60 years or over (Mean age 73.2 years,

standard deviation 8.5 years), listed in Appendix 4. The medical history of these patients was unknown. All individuals were volunteers to which the nature of the study was explained prior to their inclusion. The study was also approved by Salisbury Research Ethics Committee, application reference SA 03/98. From each patient three samples were obtained. Approximately 1-3ml of peripheral blood was taken by a trained phlebotomist into a sterile sample tube (Vacutainer) containing EDTA anticoagulant, two samples of buccal epithelia cells were taken using sterile mouthbrushes (Medical Wire and Equipment Co.), and a urine sample of between 5-80ml obtained in sterile sample jars. All samples were refrigerated at 4°C immediately after being taken and DNA extracted within 48 hours.

DNA was extracted from blood samples using the following procedure:

1. Approximately 1ml of blood was transferred to a 1.5ml microfuge tube and centrifuged at 13000rpm for 20 mins at 4°C (Sorvall RMC14, DuPont) to collect the cellular component.
2. The supernatant was removed into a 2% solution of Virkon, 1ml of ice cold distilled water added and the pellet gently re-suspended by pipetting.
3. The sample was centrifuged at 13000rpm for 20 mins at 4°C.
4. The supernatant was removed and the pellet re-suspended in 700µl of SLB.
5. The sample was centrifuged at 13000rpm for 10 mins at 4°C and the supernatant removed.
6. The cell pellet was re-suspended in 350µl of RSB by vortexing, and 40µl of 10% SDS and 3µl of Proteinase K solution added. The sample was then vortexed briefly and incubated at 37°C overnight.
7. 150µl of 6M NaCl was added, the sample was mixed by inversion, and then placed at -20°C for 5 mins until a white precipitate formed.
8. The sample was centrifuged at 13000rpm for 20 mins at 4°C.
9. The supernatant was removed into a clean 1.5ml microfuge tube, 2 volumes of ice cold 100% ethanol added, and the tube mixed by inversion until DNA visibly precipitated as a 'hairball'.
10. Samples were centrifuged at 13000rpm for 20 mins at room temperature to pellet the DNA, the supernatant removed, and 50-250µl of 1xTE buffer

added depending upon the yield of DNA.

11. DNA was re-suspended by incubation at 37°C for 1 hour, and samples stored at 4°C.

DNA was extracted from mouthbrush samples using the following procedure:

1. The two mouthbrushes from each individual were combined into a single tube.
2. 500µl of RSB was added to each sample and vortexed vigorously for 30 secs to suspend cells in the RSB.
3. The cell suspension was transferred to a 1.5ml tube using a Pasteur pipette, and 10µl of 10% SDS and 3µl of Proteinase K solution added.
4. The sample was vortexed briefly and incubated overnight at 37°C.
5. DNA was then precipitated as described in steps 7-11 above.

DNA was extracted from urine samples using the following procedure:

1. The volume of urine received was recorded, and if DNA was not to be extracted within 8 hours then a quantity equal to 10% of the sample volume of 95% methanol:5% glacial acetic acid (Sigma) was added as a preservative.
2. The sample was split equally into two 50ml centrifuge tubes. Samples were then centrifuged at 3000rpm for 10 mins at 4°C to fractionate the cellular component.
3. The supernatant was removed into a 2% solution of Virkon, 500µl of RSB added and the pellet resuspended by vortexing. At this point 50µl of the cell suspension was removed into a clean tube to which 500µl of 3:1 methanol:acetic acid added to fix the cells for microscopy (method detailed below).
4. The remainder of the sample was centrifuged at 13000 rpm for 5 mins at 4°C to pellet the cells, and the supernatant removed.
5. 10µl of 10% SDS and 3µl of Proteinase K solution were added, mixed by vortexing briefly, and incubated at 37°C overnight.
6. DNA was then precipitated as described in steps 7-11 above.

Fixed urine samples were prepared for microscopy using the following method:

1. 150µl of fixed cell suspension was spun onto a labelled microscope slide using a Cytospin II centrifuge (Shandon) at 3000rpm for 5 mins at room temperature.
2. Slides were allowed to air dry for 5 mins and then stained using an automated Varistain (Shandon) according to the following method:-

Step	Reagent	Incubation Time
1	100% ethanol	30 secs
2	70% ethanol	90 secs
3	dH_2O	30 secs
4	Haematoxylin Stain	150 secs
5	dH_2O	30 secs
6	0.5% HCl in 70% ethanol	60 secs
7	dH2O	120 secs
8	70% ethanol	30 secs
9	100% ethanol	30 secs
10	Orange G 6 Stain	30 secs
11	100% ethanol	30 secs
12	100% ethanol	30 secs
13	EA 50 Stain	60 secs
14	100% ethanol	60 secs
15	100% ethanol	60 secs
16	Xylene	60 secs
17	Xylene	60 secs

(Adapted from Papanicolaou, 1963)

3. After staining slides were placed in a fume cabinet to air dry briefly, a small quantity of Pertex Mounting Media (CellPath) pipetted over the stained cells and a coverslip placed on top. Particular care was taken to exclude air bubbles from the mounting media, which would hinder subsequent microscopy.

4. Slides were then left for approximately 1 week to allow the mounting media to set before being viewed and scored for approximate cell content using a light microscope (Zeiss).

2.1.3 Offspring ratios in women with skewed X-inactivation

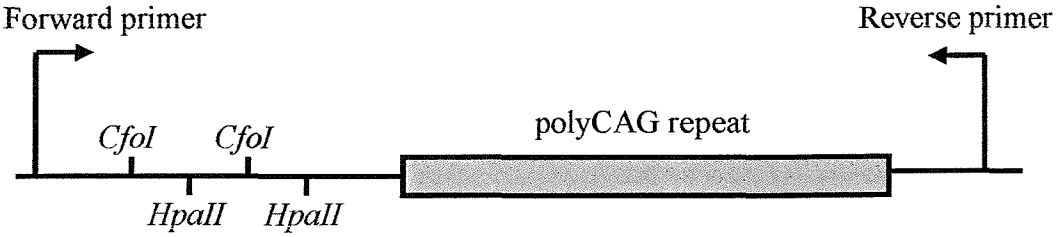
Seventeen further unrelated females were selected from patients referred to the Wessex Regional Genetics Laboratory between 1984 and 1995. Named Population E, (detailed in Appendix 5), each of these females was selected on the basis of having produced at least twice as many daughters as sons according to available pedigree data. The total numbers of offspring born to the females in Population E were 9 sons and 60 daughters. Fifteen of these females were members of pedigrees containing individuals diagnosed with Huntington disease, while the remaining two females were referred for polyposis coli screening. While the ages of six of the females in Population E were unknown, the ages of the remaining 11 ranged from 33 to 75 (Mean age 51 years, standard deviation 14.2 years). Stored samples of peripheral blood DNA were available from each of these individuals.

2.1.4 Determination of X-inactivation ratios

The assay used was a PCR-based test adapted from that described by Allen *et al.* 1992, and using the modifications described by Pegoraro *et al.* (1994). This procedure utilises the fact that the inactive X chromosome is methylated at many of its CpG dinucleotides at the 5' end of inactivated genes (Tribioli *et al.*, 1992). The PCR primers used amplify a region in exon 1 of the Androgen Receptor gene (*AR*) located at Xq12 (Shown in Figure 1). This amplified region contains a highly polymorphic CAG triplet repeat sequence, with approximately 20 different alleles seen in the normal population with a very high frequency of heterozygosity (~90% of females). Also contained within the amplicon are 4 restriction enzyme recognition sites, 2 for *HpaII* and 2 for *CfoI*. These restriction enzymes are methylation sensitive and will only cut unmethylated DNA. It has been shown that the recognition sequences for these restriction enzymes in this region are always methylated on the inactive X, and always unmethylated on the active X (Allen *et al.*, 1992). Therefore, *HpaII* and *CfoI* will only cleave DNA from the active unmethylated X and will leave DNA from the inactive methylated X chromosome intact. By first digesting DNA

with these methyl-sensitive enzymes, the quantity of product generated from each allele in the subsequent PCR will be proportional to its relative inactivation in the sample, as the PCR will only generate products from the uncut DNA from the X_i. By using PCR primers labelled with a fluorescent dye and resolving the reaction products on an ABI 377 automated sequencer, the amount of each allele produced could be accurately quantified, therefore allowing the relative activity of each X chromosome, the X-inactivation ratio, to be determined.

Figure 1 - Diagram of exon 1 of the *AR* gene at Xq12



As a result of the reaction kinetics of PCR, larger alleles tend to be less well amplified than shorter sequences of DNA. This preferential amplification becomes even more pronounced when the amplicon composition is GC-rich, as is the case here. The long tract of polyCAG's tends to form a hairpin-loop structure through complementary base-pairing between cytosine and guanine residues. The stability of this secondary structure increases with CAG repeat number, and inhibits the action of the DNA polymerase used in PCR. As a result, the degree of preferential amplification in each case depends upon the difference in repeat number between the two alleles. In order to account for this when using PCR to measure X-inactivation each sample has to be amplified in duplicate, once using untreated DNA, and once using DNA digested with *HpaII* and *CfoI*. The relative ratio of the alleles amplified from the untreated DNA can then be used to adjust the result generated from the restricted DNA, and thus account for the preferential amplification in each case.

It was found that the addition of dimethyl sulphoxide (DMSO) and 7' deaza-dGTP to the PCR reaction gave more uniform and reliable amplification. This was

particularly true for many old DNA samples that were otherwise difficult to amplify, presumably because of the poor quality of the DNA. DMSO is an inhibitor of the formation of molecular secondary structure, while 7' deaza-dGTP is unable to form complementary base pairs with cytosine. Therefore the inclusion of these two compounds into a PCR reaction inhibits the formation of the hairpin-loop structures which normally lead to long CAG repeat sequences being difficult to amplify.

Prior to PCR amplification the DNA sample to be tested was digested in a 20µl reaction mix containing 2µl DNA solution (approximately 500ng DNA), 1x Restriction Buffer L (10mM Tris-HCl, 10mM $MgCl_2$, 50mM NaCl, 1mM Dithioerythritol, pH 7.5) (Roche), 10 units *Hpa II* (Roche), and 10 units *CfoI* (Roche) in a 0.5ml reaction tube. A control digestion mix was also set up in parallel for each individual containing 2µl DNA solution and 1x Restriction Buffer L in a total volume of 20µl. Reactions were mixed by vortexing, the droplet collected by centrifugation, and then incubated overnight at 37°C. Tubes were then incubated at 94°C for 10 mins to inactivate any residual enzyme activity, and centrifuged briefly to collect the droplet. A positive control digest was also set up in each batch of reactions using DNA from a normal male. This single active X chromosome is unmethylated at the *AR* locus and so should be cleaved completely by *HpaII* and *CfoI*, thus generating no PCR product following complete restriction digestion. Visualisation of a product from this digested male DNA therefore indicates incomplete enzymatic cleavage.

Digested DNA was then amplified in duplicate using two sets of PCR primers, which both amplify the same region of the *AR* gene, but generate products of a different size. This generated two independent results for each individual from which a mean could be calculated. Primer sequences and modifications were as follows:

AR 1 - GCTGTGAAGGTTGCTGTTCCTCAT, labelled with 5' TET dye.
AR 2 - TCCAGAATCTGTTCCAGAGCGTGC, unlabelled.

AR 3 - GTGCGCGAAGTGATCCAGAA, labelled with 5' FAM dye.
AR 4 - CCAGGACCAGGTAGCCTGTG, unlabelled.

(Primers supplied by Oswel DNA Service, Southampton)

PCR reactions were performed in 0.2ml reaction tubes in an Applied

Biosystems 9600 Thermal Cycler in 20µl reaction volumes containing 4µl of DNA digestion reaction, 1x PCR Buffer II (10mM Tris-HCl, 50mM KCl, pH8.3)(Applied Biosystems), 200µM dATP, 200µM dCTP, 200µM dTTP, 125µM dGTP (Promega), 75µM 7' deaza-dGTP (Roche), 10% DMSO (Sigma), 50ng of *AR* 1/3, 50ng of *AR* 2/4, and 0.25 units of Ampli*Taq* Gold DNA polymerase (Roche).

PCR conditions were as follows:

1. An initial denaturation step at 94°C for 15 mins. This incubation period also functions to activate the Ampli*Taq* Gold DNA polymerase used in the PCR.
2. 30 amplification cycles of 94°C for 30 secs, 59°C for 30 secs, 72°C for 30 secs.
3. An extension step at 72°C for 7 mins
4. A final incubation at 60°C for 60 mins. Ampli*Taq* Gold also has a DNA terminal transferase activity, which occurs optimally at 60°C. As a result, during a PCR reaction it will sporadically add an extra adenine residue at the 3' end of some PCR products. This can result in the generation of 'split alleles', with a proportion of the allele 1bp larger than the remainder, leading to problems with subsequent quantification of the allele. With the inclusion of a final prolonged incubation at the optimal terminal transferase temperature, all PCR products should be of the same length.

Following PCR amplification, reactions were stored at 4°C in the dark. PCR products were then resolved using an ABI Prism 377 DNA sequencer (Applied Biosystems) using the following method:

1. 30cm gel plates were washed with a dilute detergent solution (Alconox), rinsed thoroughly under tap water, and then rinsed with distilled water before being left to air dry.
2. Gel plates were then assembled into a casting cassette, separated by 0.2mm spacers.
3. 300µl of 10% Ammonium Persulphate (Sigma) and 30µl of TEMED (Sigma) were added to a 4.25% denaturing polyacrylamide solution (19:1 bis

acrylamide, 7M urea, 1xTris-Borate EDTA) (Amresco) in a clean plastic bottle and mixed gently.

4. This gel mix was poured between the gel plates, a 36-well shark-tooth comb inserted in its reverse orientation and the gel left to polymerise for 1-2 hours.
5. Following polymerisation, the shark-tooth comb was removed, cleaned and re-inserted in its correct orientation into the top of the gel. The bottom section of the glass plates were thoroughly cleaned and the cassette placed into the sequencer.
6. After importing a pre-prepared sample sheet, a plate check was performed to ensure that the gel plates were thoroughly clean.
7. Buffer chambers were attached to the cassette, filled with 1xTris-Borate EDTA Buffer (Sigma), and a plate heater clipped on.
8. The gel was pre-run for 5-15 mins to remove impurities from the gel and ensure that the operating temperature of 50°C had been reached.
9. 0.5µl of each PCR product was aliquoted into numbered tubes containing 0.75µl of blue dextran (Sigma) in formamide (Amresco) and 0.25µl Genescan-500 TAMRA size standard (Applied Biosystems). This mixture was incubated at 94°C for 5 mins to denature the DNA to a single-stranded form.
10. The pre-run gel was paused, each well flushed using a hypodermic needle and syringe, and the denatured PCR products loaded into the odd-numbered wells only using 0.2mm Miniflex flat tips (Sorenson Bioscience).
11. The pre-run was resumed for 3 mins and subsequently paused again. The even numbered wells were then flushed and the remaining PCR products loaded into each of these.
12. After loading was complete, the pre-run was halted and the run initiated. Products were elctrophoresed until the 350 base-pair marker fragment was visible.
13. After electrophoresis was complete, the gel data was analysed using Genescan (Version 2.1, Applied Biosystems) and Genotyper (Version 2.0, Applied Biosystems) software. Peak heights and areas were quantified to give a measurement of the amount of PCR product from each allele.
14. The X-inactivation ratio was calculated in heterozygous individuals using

the following method:

i) In the undigested track the peak height of allele 1 was divided by the peak height of allele 2 to yield a 'Correction Factor'. This is needed to account for the preferential amplification of the smaller allele in the PCR reaction.

ii) In the corresponding track of restricted DNA, the peak height of allele 1 was divided by the peak height of allele 2. This ratio of peak heights was then divided by the 'Correction Factor' calculated in step (i) to yield an X-inactivation ratio.

iii) This figure was converted to a percentage by dividing the X-inactivation ratio by (itself + 1) and multiplying by 100. The resultant figure gives the X-inactivation ratio, expressed as the relative percentages of each peak height.

This can be expressed in the following equations:

$$\text{X-inactivation ratio} = \frac{(\text{Allele 1b} / \text{Allele 2b})}{(\text{Allele 1a} / \text{Allele 2a})}$$

$$\text{Percentage inactivation of allele 1} = \frac{\text{X-inactivation ratio}}{(\text{X-inactivation ratio} + 1)} \times 100$$

iv) In cases were the two alleles were closely spaced so that the stutter bands from the larger allele superimposed over the smaller allele a correction was applied. This correction was calculated from the height of the stutter band to the major band for each allele.

2.1.5 Statistical Analysis

χ^2 analysis, or alternatively Fisher's exact test (where expected or observed values were ≤ 5), were used to compare the proportion of females with severely skewed X-inactivation ratios in different populations. In order to compare the levels of correlation of X-inactivation ratios between different tissues within Populations C and

D, a number of analyses were performed. Pearson correlation co-efficients (r) were initially calculated to test for significant association between each tissue. The co-efficient of determination (R) yields a quantitative measure of the overall correlation between two parameters and was used to compare the degree of association of X-inactivation ratios in different tissues. In addition, regression analyses were performed and 95% prediction intervals generated for each tissue distribution. All statistical analyses were performed using Minitab, Statxact or customised software.

2.1.6 Determination of linearity of the *AR* assay

Samples of DNA from two normal males who possessed *AR* alleles which differed in size by 5 CAG repeats were selected and the concentration of the DNA samples determined by spectrophotometry and diluted to equilibrium. The two DNA samples were then mixed in serial dilutions ranging from 20:1 to 1:20 concentrations. These samples were then amplified by PCR using primer sets *AR1/2 and AR3/4* and the relative peak height of each allele produced in this reaction determined as described above. Results of relative input DNA dilution against relative peak height were then plotted graphically.

2.2 X-inactivation ratios in Silver-Russell Syndrome

The study population consisted of 34 unrelated female SRS probands of unknown aetiology, described previously (Preece *et al.*, 1997; Mergenthaler *et al.*, 2000b). The population had a mean age at sampling of 8.9 years, standard deviation 9.1, range 0.8-37, while the mean maternal age at birth was 28.6 years, standard deviation 5.5, range 19-39 (Appendix 7). A clinical diagnosis of SRS in each case was based on the following criteria: severe pre- and post-natal growth retardation, characteristic facial features, facial, trunk, or limb asymmetry, and a variety of other variable features (Wollmann *et al.*, 1995). High resolution cytogenetic examination of peripheral blood lymphocytes had previously revealed a normal karyotype in each case. Additionally, the analysis of 18-20 polymorphic markers along the length of chromosome 7 in each proband and their parents found no evidence of UPD7, indicating normal biparental inheritance of chromosome 7 in each case (Preece *et al.*, 1997; Mergenthaler *et al.*,

2000b). X-inactivation ratios in peripheral blood were determined using the *AR* gene PCR assay. Each result was generated from a mean of two independent PCR reactions.

X-inactivation ratios in two SRS patients with matUPD7 were also analysed. SR38 and SR57 were both heterodisomic, resulting from maternal meiotic non-disjunction of chromosome 7 followed by trisomy rescue (Eggermann *et al.*, 1997; Preece *et al.*, 1997).

Fisher's exact test (StatXact v.3) was used to compare the proportion of females with extreme or completely skewed X-inactivation ratios between test and control populations. As the hypothesis being tested predicted both the direction and magnitude (see discussion) of the increased frequency of skewed X-inactivation in SRS individuals compared to controls, a one-tailed test was used to calculate P-values, although the use of two-tailed tests gave similar results.

2.3 The spreading of X-inactivation in X;autosome translocations

2.3.1 Cell Culture

EBV-transformed lymphoblastoid cell lines from five patients with unbalanced X;autosome translocations (ID codes DD1289, DD0003, AL0044, BO0566, DD1550 for cases 1-5 respectively) were obtained from the European Collection of Animal Cell Cultures, Porton Down, Wiltshire, England, and grown in a 5% CO_2 atmosphere at 37°C in 30ml volumes of RPMI-1640 (Sigma) supplemented with 10% heat-inactivated foetal calf serum (Sigma), 2mM L-glutamine (Sigma), and 100U/ml each of penicillin and streptomycin (Sigma). Cultures were fed twice weekly and split at approximately one week intervals.

2.3.2 DNA and RNA extraction

DNA was extracted from lymphoblastoid cell lines and samples of peripheral blood using a salt-precipitation method, as detailed in Section 2.1.1.

RNA was extracted from lymphoblastoid cell lines and samples of peripheral

blood using Trizol reagent (Invitrogen). For samples of peripheral blood, T and B cells were separated as follows:

1. 10ml of Histopaque 1077 (Sigma) was placed in a 50ml screw cap tube.
2. The blood sample was gently layered on top of the Histopaque and centrifuged at 3000rpm for 20 mins.
3. The T and B cell layer was removed from the top of the interface layer and placed in a 15ml tube, and the red granulocyte pellet at the bottom removed for DNA extraction.
4. The T and B pellet was washed with PBS and centrifuged for 10 mins at 3000rpm.
5. The supernatant was removed into 2% Virkon, the cell pellet resuspended in 200µl PBS, transfered to an eppendorf tube and centrifuged for 2 mins at 13000rpm to pellet the cells.
6. The supernatant was removed into 2% Virkon. T and B cells, or lymphoblastoid cells, were then resuspend in 1ml of Trizol by gentle pipetting and stored at -70°C until ready for RNA extraction.

RNA was extracted as follows:

1. The frozen Trizol sample was defrosted for 1 hour at room temperature.
2. Samples were centrifuged at 11500rpm for 10 mins at 4°C.
3. The supernatant was transferred to a 2ml screw cap tube and 300µl of chloroform (Sigma) added.
4. This was mixed vigorously by shaking for 15 seconds, incubated for 2-3 mins at room temperature and centrifuged at 11500rpm for 15 mins at 2-8°C.
5. Following centrifugation the upper aqueous phase was transferred to a fresh autoclaved sterile 1.5ml eppendorf tube and 0.75ml of isopropanol (Sigma) added to precipitate the RNA.
6. This was incubated for 10 mins at room temperature and centrifuged at 2-8°C for 10 mins at 12500rpm.
7. The supernatant was removed, the RNA pellet washed with 1ml of 75% ethanol by vortexing, and centrifuged at 9200rpm for 5 mins at 2-8°C.
8. The supernatant was removed and the pellet allowed to briefly air dry before

dissolving in 50µl of RNase free water by vortexing and incubation at 55°C for 10 mins.

9. Samples were stored at -70°C until needed.

2.3.3 Determination of X-inactivation ratios

X-inactivation ratios were determined using the *AR* gene PCR assay as described in section 2.1.5.

2.3.4 Identification of transcribed polymorphisms

Putative transcribed single-nucleotide polymorphisms (SNPs) and sequence length polymorphisms contained within 10q23-qter, 11p12-pter, 7q22-qter and 6p12-pter were isolated from The Genome Database (http://gdbwww.gdb.org/), Locuslink (http://www.ncbi.nlm.nih.gov/LocusLink/), and HGBase (http://hgbase.interactiva.de/). Webcutter 2.0 (http://www.firstmarket.com/firstmarket/cutter/cut2.html) was used to identify those SNPs which either created or destroyed a restriction site, and primers to amplify these polymorphisms were designed from mRNA sequences in Genbank (http://www.ncbi.nlm.nih.gov/Web/Genbank) using Primer3 (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi). In Primer3, all primer picking settings were at default except for optimal T_m which was set at 60°C. In total, primers to amplify 216 putative polymorphisms in 135 genes were designed and tested for heterozygosity in the 5 probands AH, SP, SR, AL0044 and BO0566 by PCR, restriction digestion and agarose gel electrophoresis, as follows:

1. 50ng of genomic DNA was used as template in a 30µl PCR reaction containing 1x PCR Buffer (10mM Tris-HCl, 50mM KCl, 2mM $MgCl_2$, pH8.3) (Qiagen), 200µM each of dATP, dCTP, dTTP and dGTP, 33ng of each unlabelled primer, and 0.5U Hotstar *Taq* (Qiagen), using 35 cycles at 94°, 55-60°, and 72° for 30 secs each.
2. 15µl of this PCR reaction was digested for 2-16 hours in a 20µl reaction containing 5-20U of restriction enzyme (New England Biolabs/Roche) (Appendix 8) and 1x recommended buffer.
3. PCR products were resolved through 3% agarose containing 0.5ng/ml

ethidium bromide and visualised under ultraviolet light.

Polymorphisms, restriction enzymes, and primer details of the 32 used for allele-specific RT-PCR are detailed in Appendix 8.

2.3.5 Allele-specific quantitative RT-PCR

RNA samples (10µl) were incubated with 2U RQ1 RNAse-free DNAse (Promega) for 30 mins at 37°C in 18µl reaction volumes containing 1x First-Strand Buffer (50mM Tris-HCl pH8.3, 75mMKCl, 3mM $MgCl_2$) (Gibco BRL) and 8U RNAseOUT ribonuclease inhibitor (Gibco BRL). The DNAse was inactivated by the addition of 2µl Stop Buffer (20mM EGTA, pH8.0) and incubation at 65°C for 10 mins. 10µl of the sample was removed and diluted by the addition of 30µl of dH_2O and set aside as a control not subjected to reverse transcription.

66ng of each appropriate gene-specific reverse-strand primer were added to the remaining 10µl aliquot of RNA, incubated at 65°C for 5 mins under oil to denature the RNA, and placed on ice. This was used for cDNA synthesis in a 40µl reaction volume containing 1x First-Strand Buffer (Gibco BRL), 1mM each of dATP, dCTP, dTTP and dGTP, 20U of RNAse inhibitor (Gibco BRL), 10mM DTT and 500U of M-MLV reverse transcriptase (Gibco BRL) by incubation at 37°C for 60 mins. Finally the reaction was terminated by incubation at 95°C for 5 mins.

1µl (~50ng) of genomic DNA or cDNA were used as template in an initial 20µl PCR reaction containing 1x PCR Buffer (10mM Tris-HCl, 50mM KCl, 2mM $MgCl_2$, pH8.3) (Qiagen), 200µM each of dATP, dCTP, dTTP and dGTP, 33ng of each unlabelled primer, and 0.5U Hotstar *Taq* (Qiagen), using 30-38 cycles at 94°, 55-63°, and 72° for 30 secs each (detailed in Appendix 8). 3µl of this reaction was used as template in a second 30µl PCR reaction incorporating a fluorescently labelled primer, which was subjected to a single cycle of PCR. Use of this method avoided heteroduplex formation, thus facilitating accurate allele quantification by restriction digest (Uejima *et al.*, 2000). 15µl of this second PCR reaction was digested overnight in a 20µl reaction containing 10-20U of restriction enzyme (New England Biolabs) (Appendix 8) and 1x recommended buffer. Digests were run on an ABI377 sequencer and data analysed using Genescan and Genotyper software for quantification of alleles by peak height. The use of peak area for quantification gave similar results in

each case, but was found to be a less reliable measure of allele intensity than peak height in some instances, particularly where peaks were smeared or split.

For 26 of the 32 genes tested, comparison of results gained using proband DNA showed one allele to be approximately double the intensity of that in normal controls, demonstrating their inclusion in the translocated segment of autosome. For the remaining 6 genes (*RRM1*, *MRPL23*, *HCR*, *CSNK2B*, *HLA-DRA* and *PSMB9*) the primers used spanned large introns and were thus cDNA specific. Instead, their inclusion in the translocated segments of autosome was confirmed using physical and radiation-hybrid maps (http://genome.ucsc.edu/ and http://www.ncbi.nlm.nih.gov/genemap99/).

A gene was scored as inactive when the allele ratio gained using cDNA of the proband was significantly different from that obtained using DNA of the proband, and almost identical to those seen using DNA and cDNA from normal controls. Such results indicate that the gene is transcribed from only two alleles in the proband, implying that the copy on the der(X) is inactive (Fig. 16). In every case, analysis of parental DNA/RNA gave results that were consistent between the known origin of the der(X) and the transcriptionally silent allele. Similarly, a gene was scored as active when results obtained using cDNA of the proband were similar to those gained using DNA of the proband, and significantly different to those gained using both DNA and cDNA from heterozygous controls. Such results indicate that the gene is transcribed from all three alleles in the proband, implying that the copy on the der(X) remains active.

2.3.6 CpG island methylation analysis

NIX analysis (http://www.hgmp.mrc.ac.uk/Registered/Webapp/nix/) of genomic sequence clones deposited in Genbank was used to identify CpG islands located within the 5' region of translocated autosomal genes within 11p12-pter and 6p12-pter. Primers spanning sections of each CpG island were designed using Primer3, with primer picking settings set at default except for optimal T_m which was set at 62°C and optimal primer length at 22bp (Appendix 8). A further primer set to amplify the CpG island of the X-linked gene *PGK1*, which is known to be methylated on the inactive X and unmethylated on the active X (Gilbert and Sharp, 1999), was designed from representative genomic sequence (Appendix 8). Each amplicon was restriction

mapped using Webcutter 2.0 to determine the number and location of *HpaII*/*MspI* and *CfoI* recognition sites.

500ng DNA was digested overnight at 37°C in a 25µl reaction mix containing 1x Restriction Buffer L (10mM Tris-HCl, 10mM $MgCl_2$, 50mM NaCl, 1mM Dithioerythritol, pH 7.5) (Roche), 20U *DraI* (New England Biolabs) and either 10U *HpaII* (Roche), 10U *CfoI* (Roche) 10U *MspI* (Roche) or dH_2O. After overnight incubation a further 5U of *HpaII*, *CfoI* or *MspI* were added to the appropriate reactions and incubated for a further 5 hours at 37°C to ensure complete digestion. 4µl of this digest were then used as template to coamplify individual CpG islands and the control *PGK1* CpG island in a 20µl PCR reaction containing 1x PCR Buffer (Qiagen), 200µM each of dATP, dCTP and dTTP, 125µM dGTP (Promega), 75µM 7' deaza-dGTP (Roche), 10% DMSO (Sigma), 33ng of forward/reverse autosomal CpG island primer, 13-25ng of forward/reverse *PGK1* CpG island primer, and 0.5U Hotstar Taq (Qiagen) using 30-35 cycles at 94°, 60°, and 72° for 30 secs each (detailed in Appendix 8). PCR products were then resolved through 3% agarose containing 0.5ng/ml ethidium bromide and visualised under ultraviolet light.

2.3.7 Detection of late-replicating chromatin and *in situ* hybridisation

1. Lymphoblastoid cultures were resuspended by agitation, a 5ml aliquot removed into a 10ml flask containing 100µl of BrdU solution (final concentration 30µg/ml) and the culture mixed by inversion and incubated at 37°C.
2. After 3 hours, 100µl demecolcine (Sigma) was added (final concentration 200ng/ml), mixed by inversion and incubated at 37°C.
3. After a further 2 hours the culture was harvested and fixed as follows:
   - The culture was centrifuged for 5 mins at 1200rpm, the supernatant removed into 2% Virkon, and the cell pellet resuspended in 5ml of 75mM KCl by agitation.
   - The culture was centrifuged for 5 mins at 1200rpm, and the supernatant removed into 2% Virkon.
   - 5ml of freshly prepared 3:1 methanol:acetic acid was added dropwise with agitation to ensure resuspension of the cells.

- The culture was centrifuged for 5 mins at 1200rpm, the supernatant removed and the cell pellet resuspended in a further 5ml of 3:1 methanol:acetic acid. This step was then repeated.

4. Fixed cells were stored at –20°C until further use.
5. Slides of metaphase spreads were prepared in a controlled atmosphere (22°C and 60% humidity) by dropping one drop of fixed cell suspension onto ethanol-washed glass slides. Metaphases were then further spread by the addition of a further drop of freshly prepared 3:1 methanol:acetic acid. Slides were allowed to air dry.
6. Metaphases were hardened by exposure to ultraviolet light for 40 seconds.
7. 75µl of denaturing solution was added to the slide under a 22x50mm coverslip and incubated on a hotplate for 2 mins at 72°C.
8. The slide was rinsed briefly in 2xSSC.
9. The slide was dehydrated by serial incubations in ice-cold 70% IMS for 2 mins, 90% IMS for 2 mins, and 100% IMS for 2 mins, and allowed to air dry.
10. 75µl of PBT solution was added to the slide under a 22x50mm coverslip and incubated for 30 mins at room temperature.
11. The slide was incubated in 100ul of mouse monoclonal anti-BrdU (Sigma), diluted 1:50 in PBT, under a 22x50mm coverslip for 30 mins at room temperature.
12. The slide was washed twice for two minutes each in PBT at room temperature.
13. 100µl of anti-mouse IgG FITC conjugate (Sigma), diluted 1:500 in PBT, under a 22x50mm coverslip for 30 mins at room temperature.
14. The slide was washed twice for two minutes each in PBT at room temperature.
15. The slide was post-fixed in 4% paraformaldehyde for 5 mins at room temperature, and washed twice for two minutes each in PBT at room temperature.
17. 10µl aliquots of the appropriate TRITC-labelled whole chromosome paint (Appligene Oncor) was denatured at 72°C for 10 mins and pre-annealed at 37°C for 30 mins, as per manufacturers instructions, and applied to the slide under a 22x22mm coverslip, sealed with rubber cement, and hybridised overnight in a dark humid chamber at 37°C.
18. The coverslip was carefully removed and the slide washed for 5 mins in

0.5xSSC at 72°C and for a further 5 mins in 0.25xSSC at 72°C.

19. The slide was mounted in Vectorshield with DAPI (Vector) under a 22x50mm coverslip and viewed using Zeiss Axiophot microscope with a CCD camera. Images were captured and enhanced using Macprobe 4.1 software (PSI).

2.3.8 Histone immunofluorescence and *in situ* hybridisation

1. Antisera to histone H3 dimethylated at lysine 4, histone H3 acetylated at lysine 14, and histone H4 acetylated at lysine 8 were raised as described previously (Belyaev *et al.*, 1996). The antibody to H3 methylated at lysine 4 was raised against H3 peptides dimethylated at lysine 4 and does not recognise H3 trimethylated at lysine 4 (B. Turner and L. O'Neill, unpublished data).
2. Lymphoblastoid cultures were resuspended by agitation, a 5ml aliquot removed into a 10ml flask containing 100µl demecolcine (Sigma) (final concentration 200ng/ml), mixed by inversion and incubated at 37°C for 2 hours.
3. Cell cultures were centrifuged at 1200rpm for 5 mins to pellet cells. The supernatant was removed and the cell pellet initially resuspended in 1ml of hypotonic solution (75mM KCl). Cell density was determined using a haemocytometer and additional 75mM KCl solution added to adjust the cell density to $5x10^6$ cells/ml. Cells were then incubated at 37°C for 10 mins.
4. Glass slides were cleaned with ethanol and 100µl of cell suspension spun onto them using a Cytospin II (Shandon) at 1200rpm for 5 mins.
5. Slides were allowed to air-dry briefly and then incubated in KCM buffer for 10 mins at room temperature to permeabilise the cells.
6. Slides were blocked in KCB buffer for 10 mins at room temperature.
7. The slide was incubated with 100µl of antisera, diluted 1:200 in KCB buffer, under a 22x50mm coverslip for 30 mins at room temperature.
8. The slide was washed twice for two minutes each in KCB at room temperature.
9. The slide was incubated in 100µl of anti-rabbit IgG FITC conjugate (Vector), diluted 1:200 in KCB, under a 22x50mm coverslip for 30 mins at room temperature.
10. The slide was washed twice for two minutes each in KCB at room

temperature.

11. The slide was post-fixed in 4% formaldehyde for 5 mins at room temperature, and washed twice for two minutes each in 1xPBS at room temperature.
12. The slide was mounted in Vectorshield with DAPI (Vector) under a 22x50mm coverslip and viewed using Zeiss Axiophot microscope with a CCD camera. Immunofluorescence images were captured and enhanced using Macprobe 4.1 software (PSI) and cell co-ordinates taken using an Englandfinder slide.
13. The coverslip was carefully removed and the slide washed twice for two minutes each in 1xPBS at room temperature.
14. The slide was dehydrated by serial incubations in 70% IMS for 2 mins, 90% IMS for 2 mins, and 100% IMS for 2 mins, and allowed to air dry.
15. Slides were fixed by incubating in freshly-prepared 3:1 methanol:acetic acid for 5 mins at room temperature.
16. The slide was exposed to ultraviolet light for 40 secs to harden the cells.
17. The slide was denatured by incubating in 100µl of denaturing solution under a 22x50mm coverslip on a hotplate for 2 mins at 72°C.
18. The slide was washed for two minutes in 2xSSC.
19. The slide was dehydrated by serial incubations in 70% IMS for 2 mins, 90% IMS for 2 mins, and 100% IMS for 2 mins, and allowed to air dry.
20. 10µl aliquots of the appropriate TRITC-labelled whole chromosome paint (Appligene Oncor) was denatured at 72°C for 10 mins and pre-annealed at 37°C for 30 mins, as per manufacturers instructions, and applied to the slide under a 22x22mm coverslip, sealed with rubber cement, and hybridised overnight in a dark humid chamber at 37°C.
21. The coverslip was carefully removed and the slide washed for 5 mins in 0.5xSSC at 72°C and for a further 5 mins in 0.25xSSC at 72°C.
22. The slide was mounted in Vectorshield with DAPI (Vector) under a 22x50mm coverslip and viewed using Zeiss Axiophot microscope with a CCD camera. Previously captured cells were relocated using Englandfinder co-ordinates and *in situ* hybridisation images captured and enhanced using Macprobe 4.1 software (PSI).
23. Corresponding immunofluorescence and *in situ* hybridisation images were exported and overlaid in Adobe Photoshop.

2.3.9 LINE-1 repeat analysis of 10q24-25, 11p13-pter and 6p21.3-22.3

The Human Genome Browser, April 2001 assembly (http://genome.cse.ucsc.edu/), was used to determine the LINE-1 content of translocated autosomal regions and the locations of genes within them. Genomic sequence contigs were analysed for LINE-1 and L1 subclass content using RepeatMasker (Smit, A. F. A. & Green, P. at http://ftp.genome.washington.edu/RM/RepeatMasker.html) in sliding windows of 200kb at 20kb intervals, and the repeat content determined as a percentage of each window. Because of the draft nature of much of the sequence, anonymous bases (n), mostly representing gaps, were excluded from the analysis such that each window always contained 200kb of known sequence. Exceptions to this were made where anonymous base count exceeded 50kb of continuous sequence, as these imply contig gaps of unknown size and order. However, in order to represent the position of contig gaps, sequence windows whose start position was contained within anonymous sequence were scored as –1% repeat content.

2.3.10 Intragenic repetitive sequence analysis of translocated autosomal genes

The Human Genome Browser, December 2001 assembly, was used to determine the locations, orientation and sequence content of the majority of translocated autosomal genes investigated by RT-PCR and methylation analysis. The location of two genes (*HLA-DRB1* and *H19*) was not listed by the browser, and were instead assigned using representative sequences from GenBank (accession codes M26038 and BC006831 respectively). The browser listed 2 locations for a further five genes (*MICA*, *PSMB9*, *BRD2*, *TSSC3* and *RRM1*), probably representing artefactual duplication during sequence assembly as evidenced by perfect homology between each paired segment. For these genes the most probable order on the basis of radiation hybrid, genetic and cytogenetic mapping data was chosen. A total of 970,820bp inactive, 2,144,165bp partially active and 136,741bp fully active gene sequences were extracted and analysed to determine content of several types of repeat (Table 6). RepeatMasker analysis was performed at the sensitive setting to identify interspersed repeats (LINE,

L1, L2, CR1, SINE, Alu, MIR, and LTR) and low complexity DNA. RepeatMasker further classified L1 repeats into ten subfamilies (mammalian-wide L1M1 to L1M4, primate-specific L1P1 to L1P5 and the currently active L1HS). Tandem Repeats Finder (Benson, 1999) was used to identify tandem, di-, tri-, and tetra-nucleotide repeats. GC and anonymous base (n) content, mostly representing gaps, were also determined. Gene sequences were characterised by calculating the number of base pairs comprising each repeat type as a proportion of each gene sequence. This analysis was extended to include sequences extending 1kb, 2kb, 5kb and 10kb 5' from each gene. Despite the draft status of some of this dataset, only 3 and 6 contig gaps were encountered when 5kb and 10kb of additional 5' sequence were included, respectively. Sequence analysis was restricted to within relatively short distances of single genes as the spreading of X-inactivation was discontinuous in some cases, and the resolution of activation status was limited to those genes within which polymorphisms were identified. Each gene was coded by activation status from 1-10, with 1 being inactive and 2-10 representing an increasing percentage of activity, so that a score of 5 represents 50% activity.

2.3.11 Statistical analysis

Correlations between activation status and repeat density were evaluated using Spearman's correlation coefficients with mid-rank tie-corrections. Two-tailed P-values (Tables 6 and 7) were obtained as Monte Carlo estimates (20,000 samples) from permutation distributions, using a linear-by-linear association model (Sprent, 1995) (StatXact v.3). As a quality check, all rank correlations and associated tail-area probabilities were re-estimated (NPSTAT v3.8, May *et al.*, 1993). 95% confidence intervals on all correlations were obtained as bias-corrected Bootstrap estimates (Manly, 1998) (Simstat v. 2.01, Peladeau, 1996). No corrections for multiple testing error were employed because of the single primary hypothesis under investigation. Additionally the observed secondary correlations exhibit a structure that argues strongly against chance associations.

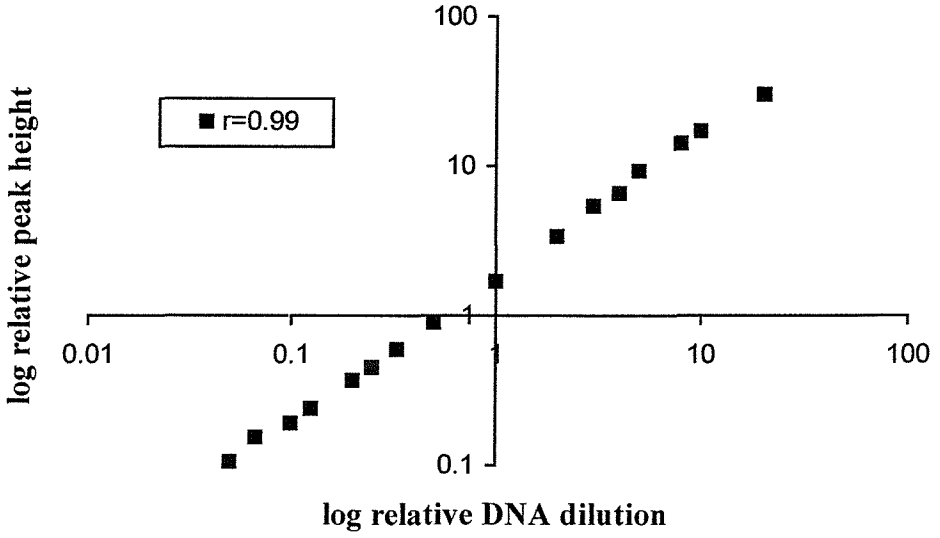
Section 3 – Results

3.1 X-inactivation ratios in normal women

3.1.1 Determination of linearity of the *AR* assay

The results of a 400-fold serial dilution and subsequent PCR of DNA from two normal males with different *AR* alleles are shown in Figure 2.

Figure 2 - Relative allele intensities produced following a serial dilution of DNA from two males and subsequent PCR at the AR locus



Results show that the peak height of each allele produced by the *AR* PCR assay is linearly related to the amount of input DNA (r=0.99), showing that the assay is quantitative over a wide range of input DNA dilutions.

3.1.2 Variations of skewed X-inactivation with age

Initial X-inactivation screening was of Population A, consisting of 98 females aged under 20, and Population B comprising 115 women aged 60 or over. X-inactivation ratios using DNA extracted from peripheral white blood cells from each individual were determined by amplification of the *AR* locus at Xq12. The relative amount of

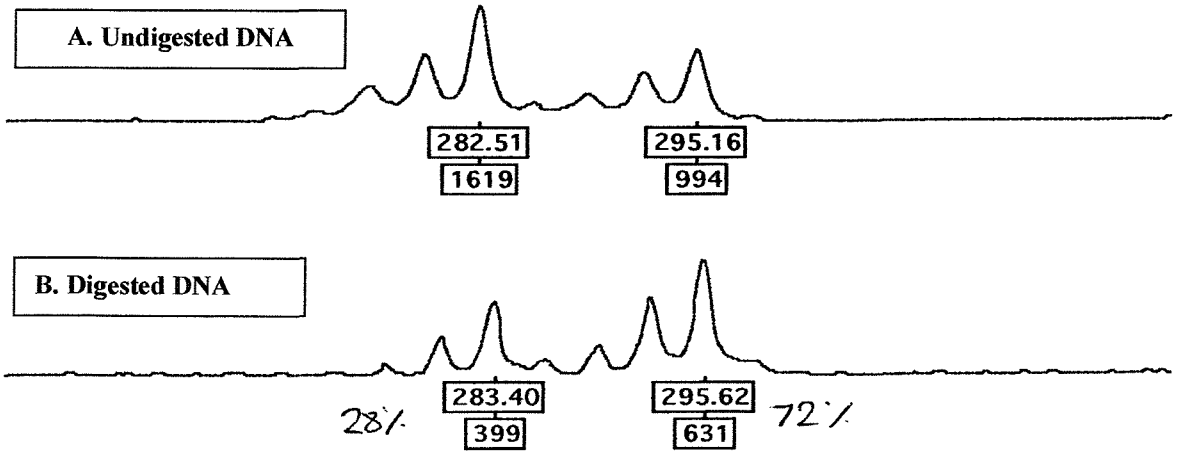
PCR product generated from each allele in heterozygous individuals was quantified using Genotyper software. Examples of results gained using this method are shown in Figure 3.

For each individual at least two independent PCR reactions were set up, generating multiple results for most subjects from which a mean X-inactivation ratio was calculated. In the vast majority of cases, these two independent PCR reactions generated X-inactivation ratios which differed by only a few percentage points (mean difference 4.8%, standard deviation 4.7% in Population A). This high degree of reproducibility suggests that the *AR* assay produced results which were both reliable and accurate. Results gained for Population A are listed in Appendix 1, while those for Population B are listed in Appendix 2.

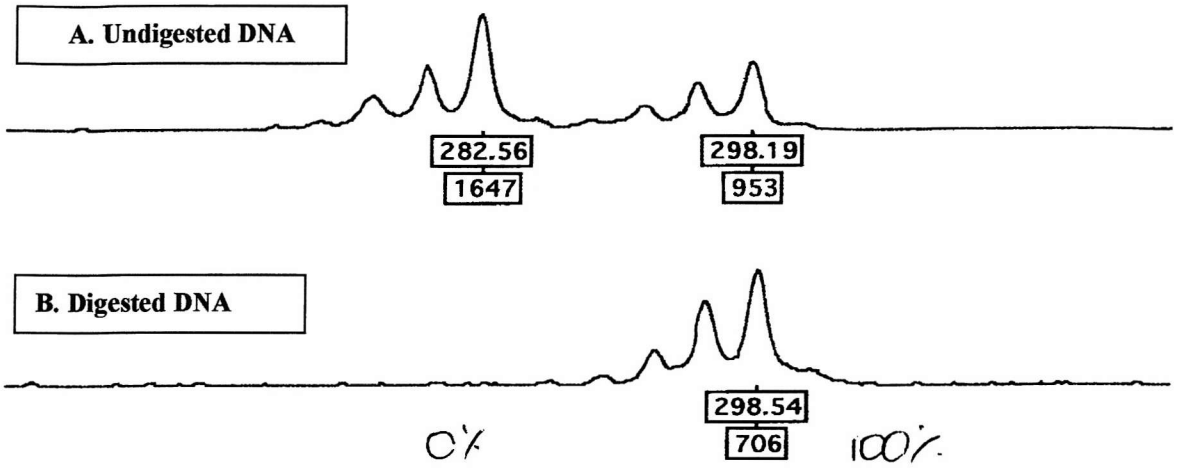
Initially allele intensity was quantified in two ways, by measuring both the peak height and the peak area defined by each allele. In the vast majority of cases, subsequent calculation of X-inactivation ratios using these two parameters yielded almost identical results. However, peak area was found to be unreliable as a measure of allele intensity in some instances. In particular, if the quality of the PCR products or gel resolution was poor, leading to smears or multiple 'split' alleles, then the use of peak height was much more satisfactory. Consequently, peak height was adopted as the preferred parameter to quantify allele intensity in all subsequent experiments.

Figure 3 – Examples of X-Inactivation Results at the *AR* Locus

1. Moderately Skewed X-inactivation

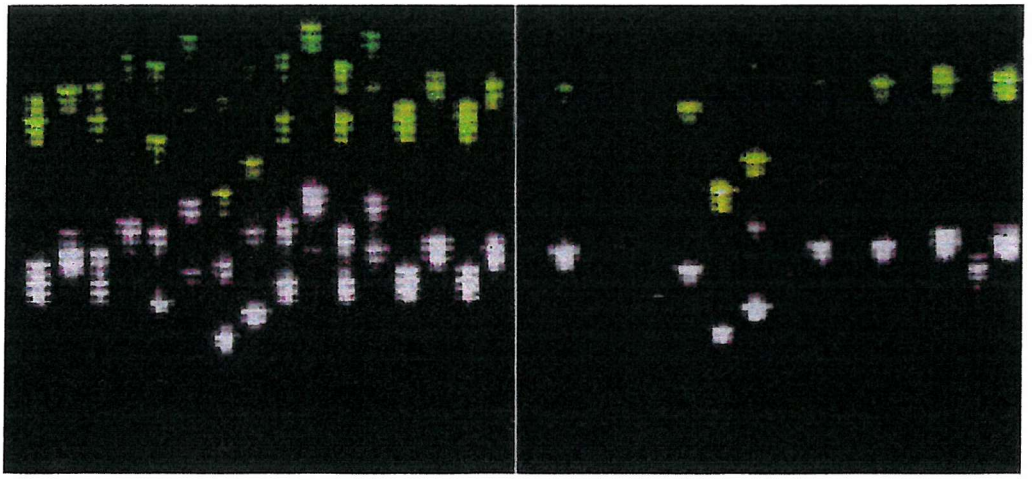


2. Completely Skewed X-inactivation



Reliable PCR amplification of some individuals DNA was found to be difficult, and in some cases there was extreme preferential amplification of the smaller alleles. This was a particular problem with many older samples, some of which had been in storage for over ten years and were most likely degraded and of poor quality. It was found that the addition of DMSO and 7' deaza-dGTP to the PCR reaction alleviated these problems of preferential amplification and produced more reliable and uniform results, as described by Mutter and Boynton (1995).

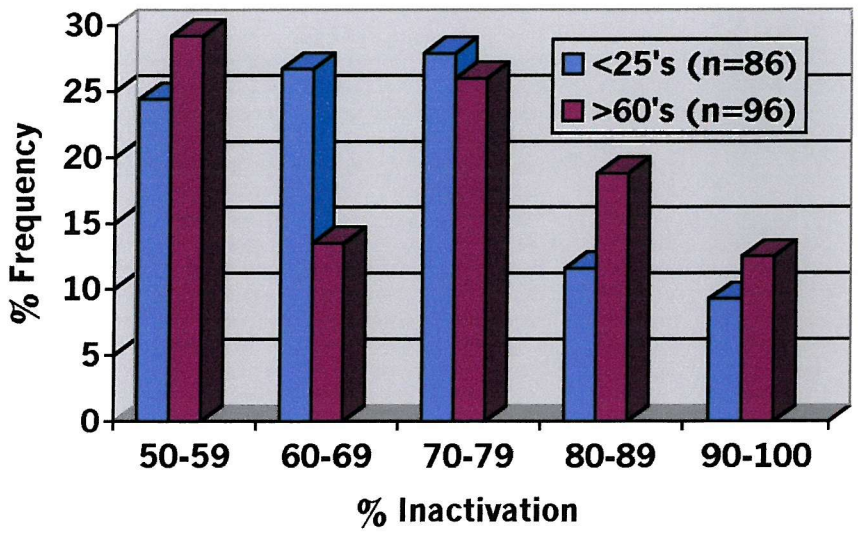
Figure 4 – The Influence of DMSO and 7' deaza-dGTP on PCR Results at the *AR* Locus



The improvement these chemicals make to the quality of PCR amplification is shown in Figure 4. Both sets of PCR products shown were amplified and resolved in parallel using identical DNA samples. However, those on the right were amplified using a standard PCR reaction mixture, while the image on the left shows the same DNA samples amplified using PCR reactions containing 10% DMSO and 75μM 7' deaza-dGTP. On the basis of these results, DMSO and 7' deaza-dGTP were added to all subsequent PCR amplifications of the *AR* locus.

Figure 5 shows the distribution of X-inactivation patterns obtained in nucleated blood cells from informative individuals in Populations A (Females ≤20 years) and B (Females ≥60 years). X-inactivation patterns are expressed as the percentage ratio of the predominantly inactive allele to the predominantly active allele and are displayed in 10% intervals. Using an X-inactivation ratio of ≥90:10 as a criterion for severe skewing, 9% of individuals under 20 had severely skewed X-inactivation compared to 13% of females >60.

Figure 5 - Skewing of X-Inactivation in Blood of Control Females of Different Ages



3.1.3 Tissue-specific variations of X-inactivation

X-inactivation ratios were next determined in cells from blood, mouthbrush and urine samples obtained from Populations C (women aged ≤25) and D (women aged ≥60). As with Populations A and B, at least two independent PCR reactions were set up for each sample, generating multiple results from which a mean X-inactivation ratio was calculated. In the vast majority of cases, these two independent PCR reactions generated X-inactivation ratios which differed by only a few percentage points in all three tissues (mean difference in blood 4.6%, standard deviation 4.2%; mean difference in buccal mucosa 4.7%, standard deviation 3.8%; mean difference in urinary epithelia 4.5%, standard deviation 5.1% in Population D). Results gained for Population C are listed in Appendix 3, while results from Population D are listed in Appendix 4. Of the 50 females recruited in Population C, seven were uninformative (homozygous) at the *AR* locus. Of the remaining 43 informative individuals, blood samples were obtained from 35, mouthbrushes from all 43, and urine samples from 41. However, PCR amplification consistently failed to produce any results using DNA extracted from three of the 41 urine samples. Six out of the 51 females in Population D (aged 60 years or over) were uninformative for the *AR* assay. Of the remaining 45 subjects, blood samples were obtained from 43 individuals, mouthbrushes from all 45, and urine samples from 40. As with Population C, DNA extracted from three of these urine specimens did not amplify by PCR analysis.

The cell content of urine samples obtained from 17 individuals in Population D was examined by light microscopy. Results from these samples showed that in most cases, squamous epithelia formed the bulk of the nucleated cells. Although small proportions (up to 20%) of urothelial cells were also present in some of the samples examined, these observations indicate that the cell content of urine samples is mainly of ectodermal origin.

The distribution of X-inactivation ratios in the three tissues obtained from Population C (females ≤25) is shown in Figure 6, while Figure 7 shows the equivalent data from Population D (females ≥60).

Severely skewed patterns of X-inactivation were more common in the elderly population in all tissues examined. Additionally, the incidence of severely skewed X-inactivation in these elderly women showed marked variation between different tissues. Of the females ≤25 years old in Population C, only 1 individual showed an

X-inactivation ratio ≥90:10, present only in her urinary epithelia. In contrast, in the elderly women of Population D, the incidence of severely skewed X-inactivation was 23% in white blood cells, 4% in buccal epithelia, and 11% in urinary epithelia.

Figure 6 - Tissue Specific Patterns of X-Inactivation in Women Under 25

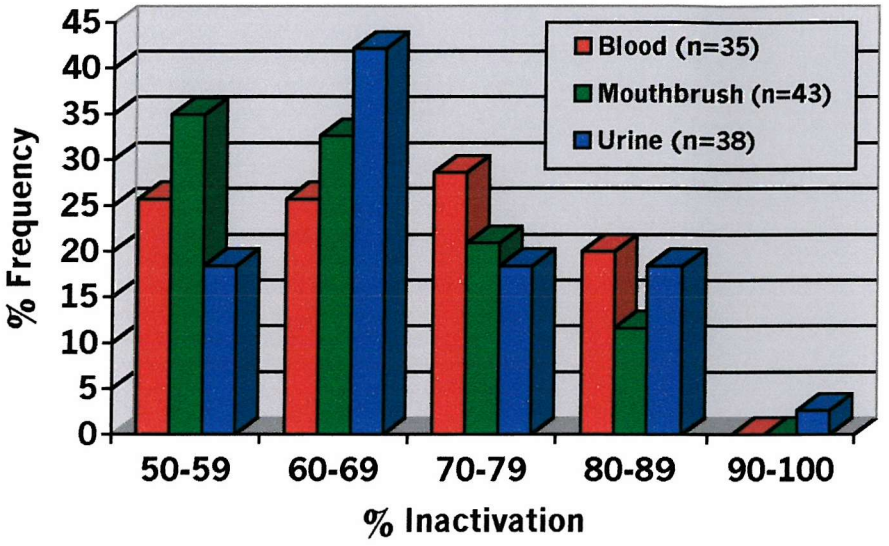
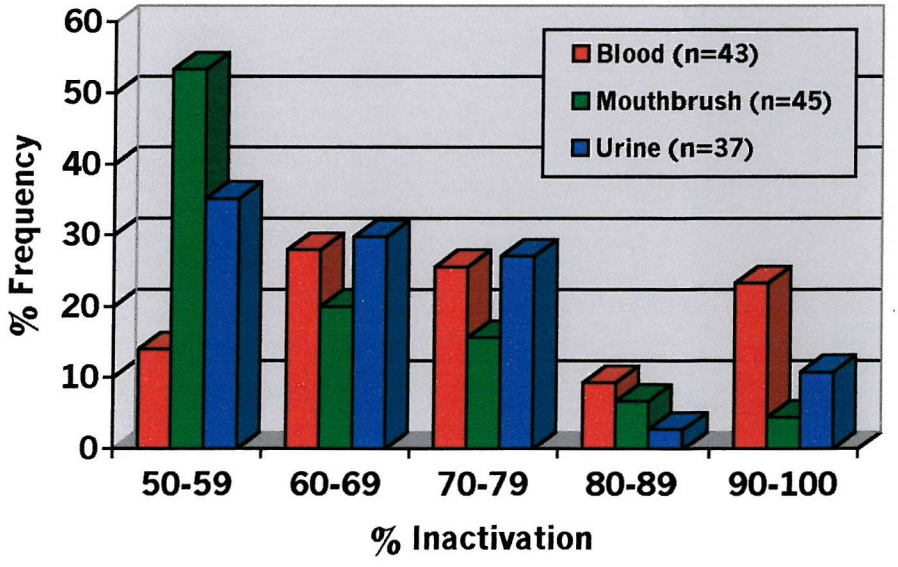


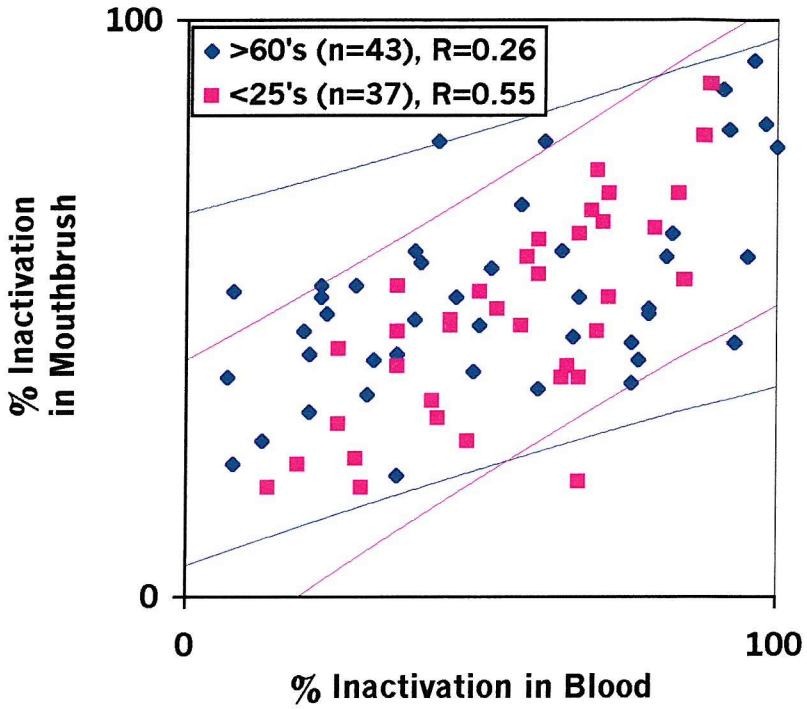
Figure 7 - Tissue Specific Patterns of X-Inactivation in Women Over 60



In order to study the correlation of X-inactivation ratios between different tissues of each individual, scatter plots were drawn to make pair-wise tissue comparisons in both the young and elderly populations. Calculation of correlation co-efficients showed that, on a population level, there was a statistically significant correlation of X-inactivation ratios between all of the tissues examined in both young and elderly women ($r \geq 0.51$, $P < 0.001$ in each case).

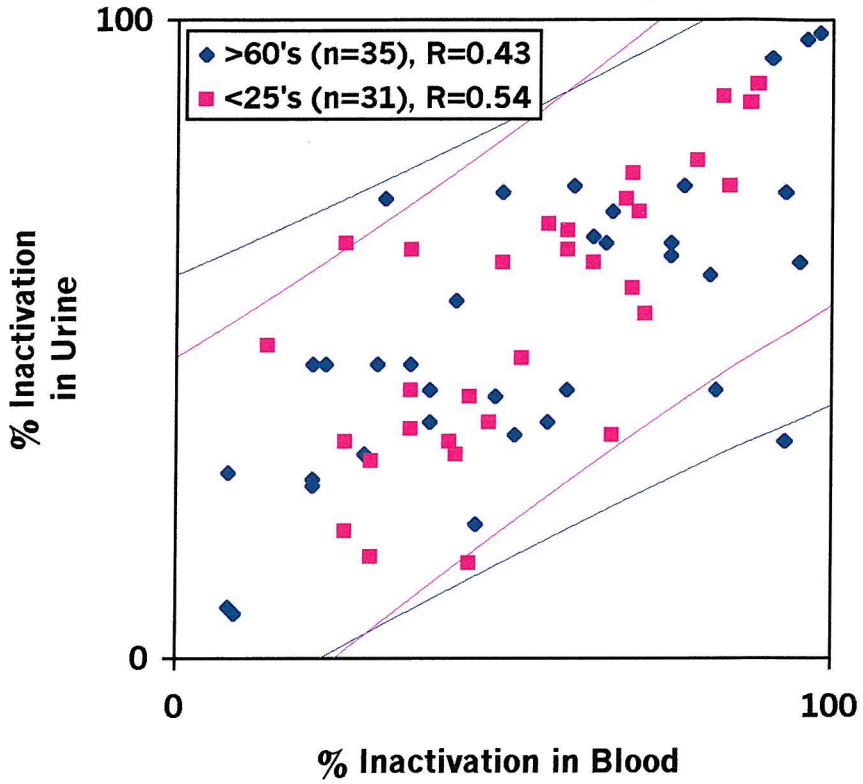
Figure 8 shows the correlation of X-inactivation ratios between white blood cells and buccal epithelia in Populations C and D. Inactivation values are expressed as the percentage ratio of the smaller *AR* allele to the larger *AR* allele, while coloured lines define 95% prediction intervals for the two age distributions. Results show that in the majority of individuals there was a significant correlation between the X-inactivation ratios in blood cells and buccal epithelia. Importantly though, in some individuals there was severe skewing present in blood, and a random inactivation pattern in the mouthbrush sample. This was much more prevalent in women ≥60 than in the younger age group. Calculation of co-efficients of determination (R) show that overall, there was a closer correlation between X-inactivation ratios in the two tissues in young individuals ($R = 0.55$, $P < 0.001$) than in older women ($R = 0.26$, $P < 0.001$).

Figure 8 - Correlation Between X-Inactivation in Blood and Mouthbrush Samples



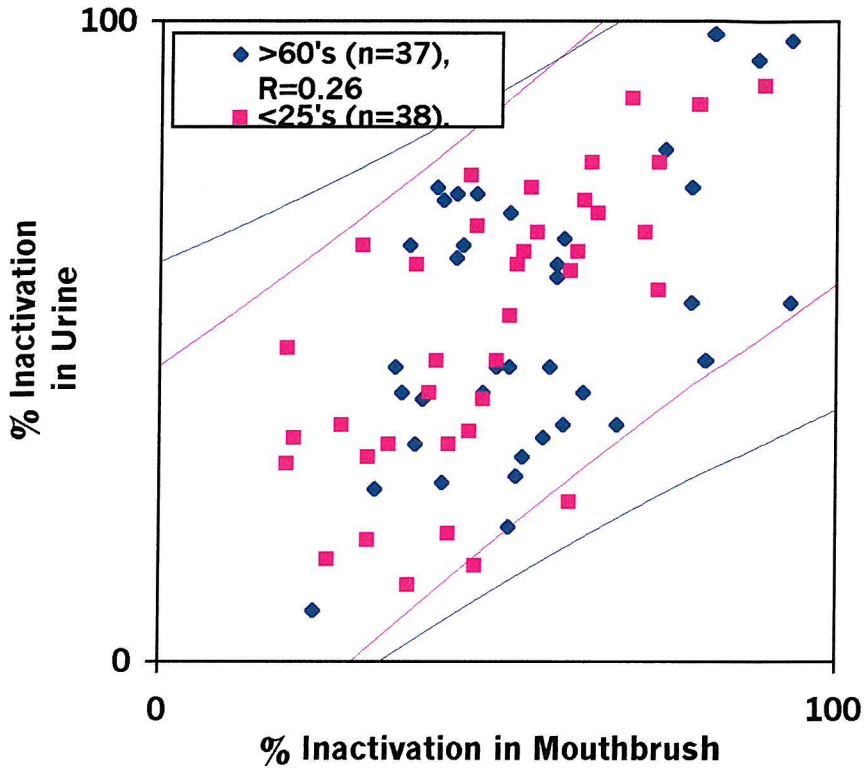
Further scatter plots correlating X-inactivation ratios between both blood and urine samples, and mouthbrush and urine samples, are shown in Figures 9 and 10 respectively. Correlations between blood and urine samples were very similar to those between blood and mouthbrush samples. Co-efficients of determination again indicated that X-inactivation ratios between blood and urine samples become more divergent with age ($R=0.53$, $P<0.001$ in females ≤ 25, $R=0.43$, $P<0.001$ in females ≥ 60). In a few individuals, extreme skewing was present in blood, while corresponding X-inactivation patterns in urinary epithelia were random.

Figure 9 - Correlation Between X-Inactivation in Blood and Urine Samples



Despite having a lower incidence of extreme skewing than in blood cells, correlation between mouthbrush and urine samples was the lowest of all the pair-wise comparisons made for both age groups ($R=0.47$, $P<0.001$ in females ≤ 25, and $R=0.26$, $P<0.001$ in women ≥ 60). Again however, discordant X-inactivation ratios were more common in elderly women.

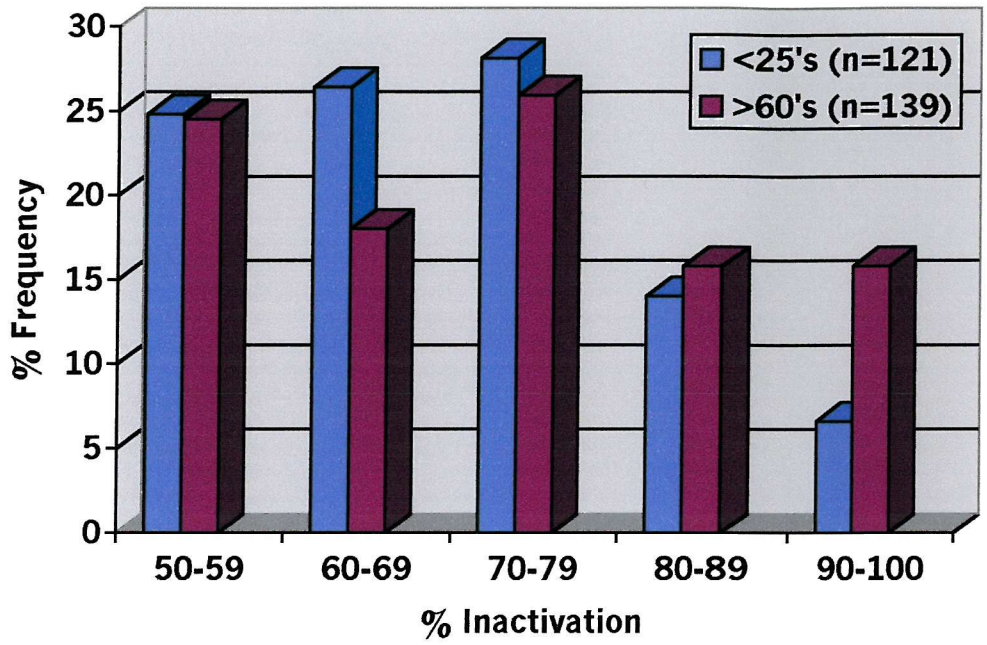
Figure 10 - Correlation Between X-Inactivation in Mouthbrush and Urine Samples



Overall, these correlations demonstrate that there was considerable variability in the different tissue patterns obtained between individuals. While some individuals showed random patterns in all tissues tested, others had severely skewed X-inactivation patterns in one, two, or all three cell-types. Interestingly none of the individuals tested had extreme skewing in opposite directions in different tissues; i.e. the same X chromosome was predominantly inactivated in all three tissues tested.

The results obtained for blood samples taken from Populations C and D were combined with those for Populations A and B to give an increased sample size. These combined data are shown in Figure 11, and give an incidence of extremely skewed X-inactivation (ratios ≥90:10) of 7% in females ≤25, and 16% in women aged ≥60 (statistically significant with χ^2=5.38, P<0.05). Using a less stringent criterion for skewing of ratios ≥75:25, 28% of females ≤25 had skewed X-inactivation compared with 48% of females aged ≥60 (statistically significant with χ^2=11.42, P<0.001).

Figure 11 - Skewing of X-Inactivation in Blood of Control Females of Different Ages



3.1.4 Offspring ratios in women with skewed X-inactivation

Pedigree information was available for only four of the women with X-inactivation ratios ≥90:10 in Population B. In addition, one further mother with extreme skewing was identified from Population A (detailed below). Figure 12 shows the number of offspring born to each of these five women.

Table 1 - Offspring Ratios in Mothers with Skewed X-Inactivation

No.	Age	X-inactivation ratio (in blood)	Sons	Daughters
1	67	96-4	1	2
2	73	94-6	0	3
3	66	94-6	0	2
4	63	91-9	0	1
5	35	100-0	3	1*
		TOTALS	**4**	**9**

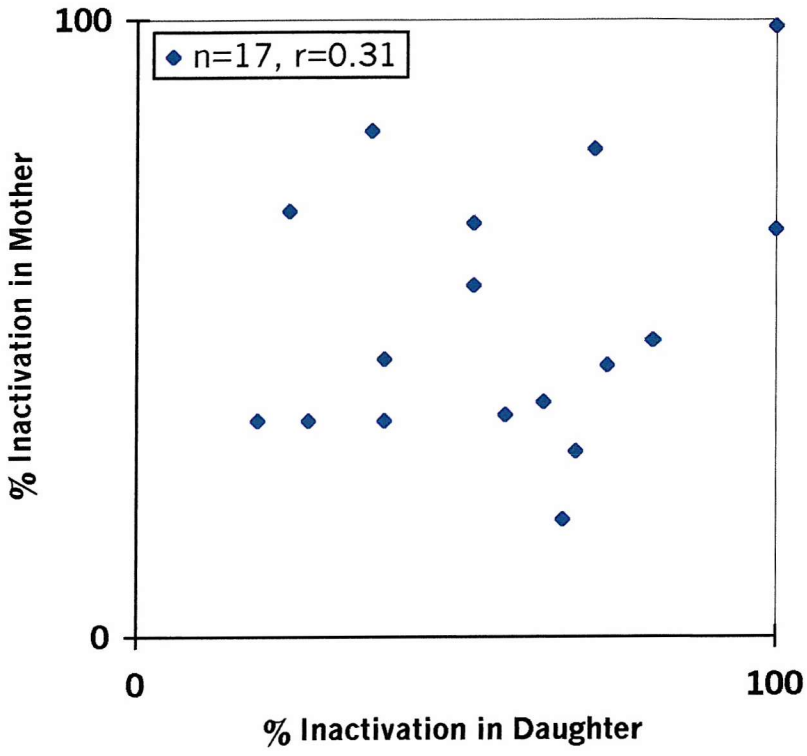
Although there was an apparent excess of daughters born to these five women with skewed X-inactivation ratios, this difference was not statistically significant (P=0.52, Fisher's exact test). No data were available on the frequency of miscarriages. Four of the five women produced more daughters than sons, in contrast to mother 5 who had only one daughter and three sons. However, examination of the X-inactivation pattern in this single daughter (asterisked) showed that, like her mother, she too had a completely skewed X-inactivation ratio in white blood cells, suggesting the possible inheritance of a locus influencing skewing in this case. Unfortunately, no DNA samples were available from any of the male sibs, or from any of the offspring of mothers 1-4.

In order to further test the hypothesis that women with skewed X-inactivation produce more daughters than sons, an additional cohort of 17 women was studied, named Population E. Detailed in Appendix 5, each of these females was selected on the basis of having produced at least twice as many daughters as sons. Results showed that only one of the 16 informative females studied (6%) in Population E had X-inactivation ratios ≥90:10, compared to 7-16% of controls in Populations A-D.

3.1.5 Heritability of X-inactivation ratios

Maternal blood DNA samples were available for 29 of the females in Population A. X-inactivation ratios were determined in each individual, and the results for the 17 informative mother-daughter pairs plotted on the scatter diagram shown in Figure 12. Inactivation values are expressed as the percentage ratio of the smaller *AR* allele to the larger *AR* allele for mothers, and the percentage inactivation of the maternally-derived allele for daughters. Results are listed in Appendix 6. As indicated by the low correlation co-efficient (r=0.31, P=0.23), the X-inactivation ratios in the mother-daughter pairs tested were apparently unrelated in most cases.

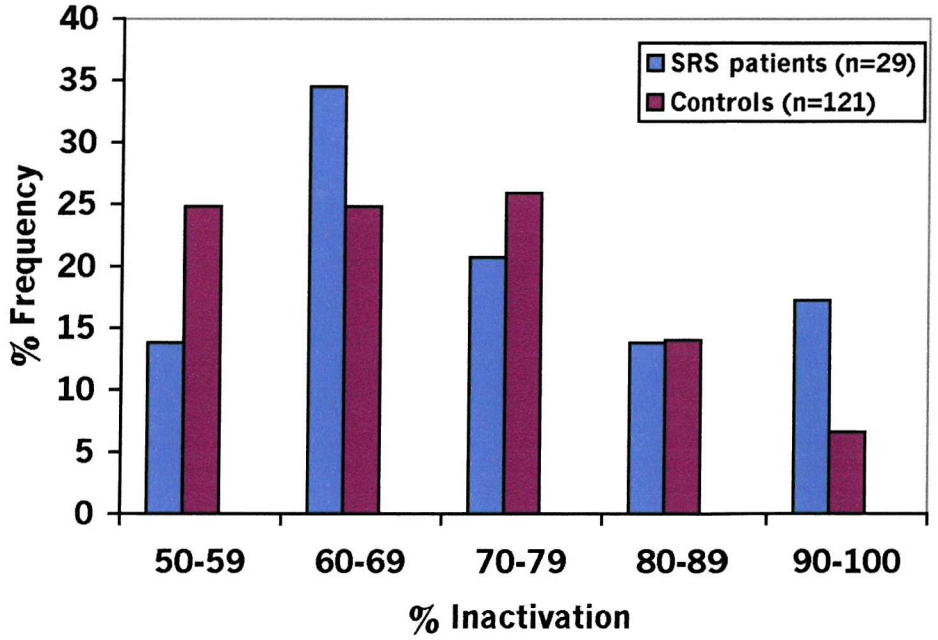
Figure 12 - Correlation of X-Inactivation in Mother-Daughter Pairs



3.2 X-inactivation ratios in Silver-Russell syndrome

29 of the 34 SRS probands studied were informative for the *AR* assay (Appendix 7). Extremely skewed X-inactivation (ratios ≥90:10) was observed in 5 of these 29 (17.2%), compared to 8 of 121 (6.6%) controls of similar age (not significant, P=0.079). Additionally, completely skewed X-inactivation (ratios 100:0) was observed in 3 of these 5 SRS probands (representing 10.3% of the cohort), compared to only 3 of 270 (1.1%) normal controls of all ages (statistically significant, P=0.014). Figure 13 shows the relative distributions of X-inactivation ratios obtained in non-UPD7 SRS patients and controls.

Figure 13 - X-inactivation ratios in SRS patients and controls



Analysis in the two SRS cases with matUPD7 showed X-inactivation ratios of 86:14 and 85:15.

3.3 The spreading of X-inactivation in X;autosome translocations

3.3.1 Case 1 - AH, 46,X,der(X)t(X;10)(q26.3;q23.3) mat

3.3.1.1 Clinical details

The pedigree is shown in Figure 14. The proband, AH, was referred to a clinical geneticist with secondary amenorrhea, aged 14 years. She began menstruation aged 10, but periods were always irregular, finally ceasing altogether. All other secondary sexual characteristics were normal. Her height was 157cm. She was healthy as a child, and showed no abnormal external physical features. Her maternal aunt (ED) only ever had one period, and is childless. Aside from this, no obvious abnormalities were present in any other family members.

Figure 14 – Pedigree of AH

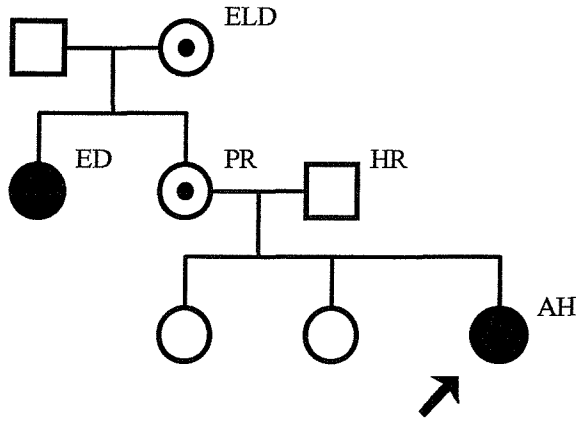


Fig. 14. Balanced carriers of the translocation are represented by dotted circles, unbalanced carriers by filled circles. The proband AH is indicated by an arrow.

3.3.1.2 Cytogenetic analysis

GTL-banding of PHA-stimulated peripheral blood lymphocytes revealed an unbalanced X;10 translocation in the proband AH, karyotype 46,X,der(X)t(X;10)(q26.3;q23.3). She is therefore trisomic for 10q23.3-qter, and monosomic for Xq26.3-qter. Her maternal aunt (ED) has the same unbalanced karyotype, while her mother (PR) and maternal grandmother (ELD) were also found to carry the same translocation, but in a balanced form, karyotype 46,X,t(X;10)(q26.3;q23.3).

3.3.1.3 X-inactivation, parental origin, and breakpoint analysis

Methylation analysis at the *AR* locus demonstrated a completely skewed X-inactivation pattern in both peripheral blood and EBV-transformed lymphocytes of AH. A completely skewed X-inactivation ratio was also observed in her mother PR, who carries the balanced form of the translocation. Although results (Figure 15) appear to indicate that the same X chromosome is inactive in both mother and daughter, PCR analyses of microsatellite markers in Xq demonstrated the presence of a recombination event between the *AR* locus (Xq12) and the translocation breakpoint in Xq26.3 (data not shown). Therefore, the maternally-inherited der(X;10) is exclusively inactive in AH, while the normal X chromosome is silenced in PR. These X-inactivation patterns are consistent with those seen in the majority of carriers of unbalanced and balanced X;autosome translocation carriers, respectively (Mattei *et al.*, 1982). Consistent with this finding, PCR analysis of polymorphic markers demonstrated the der(X;10) to be of maternal origin, with breakpoints lying between D10S583 and HPS, and DXS1187 and DXS1062 respectively.

Figure 15 – X-inactivation ratio analysis in the pedigree of AH using the *AR* assay

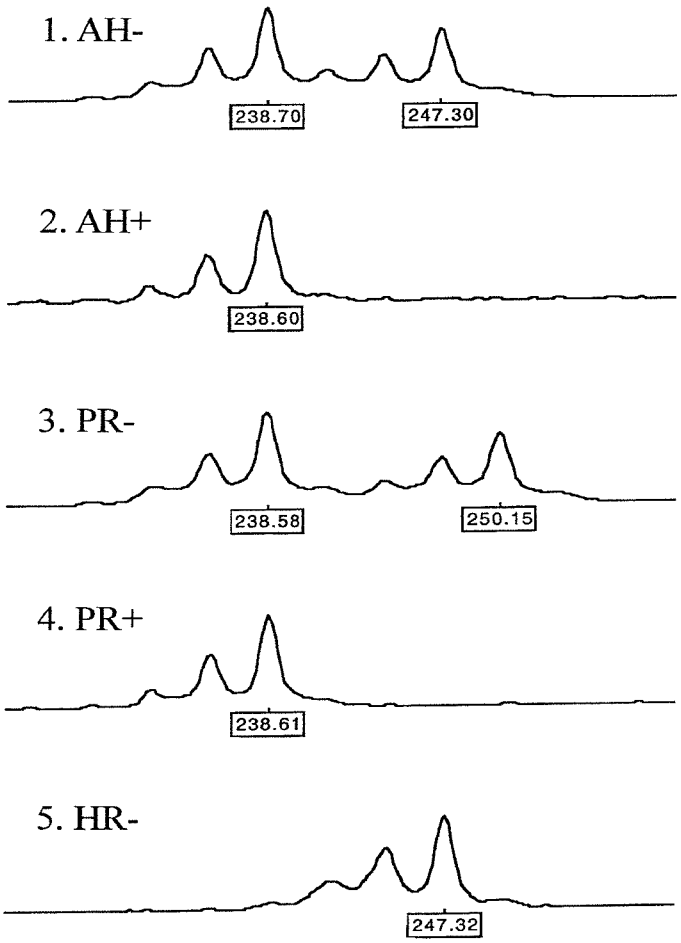


Fig. 15. (-) denotes PCR of undigested DNA, (+) denotes PCR using DNA digested with *HpaII/CfoI*. Figures below each allele represent size in base pairs.

3.3.1.4 Gene expression analysis

Heterozygous polymorphisms were identified in five genes within the translocated segment 10q23.3-qter in AH, and were used for allele-specific RT-PCR of RNA extracted from EBV-transformed lymphoblasts. Results for three of these genes (*HPS, MXI1* and *PRDX3*) are shown in Figure 16, and demonstrate an apparently continuous spread of gene silencing across nearly the entire translocated 10q segment (Table 2). The transcriptional status of three genes (*HPS, MXI1* and *ABLIM*) was also

analysed using RNA extracted from peripheral blood of AH, with concordant results.

Figure 16 – Allele-specific PCR/RT-PCR of three genes within the translocated segment 10q23-qter in AH

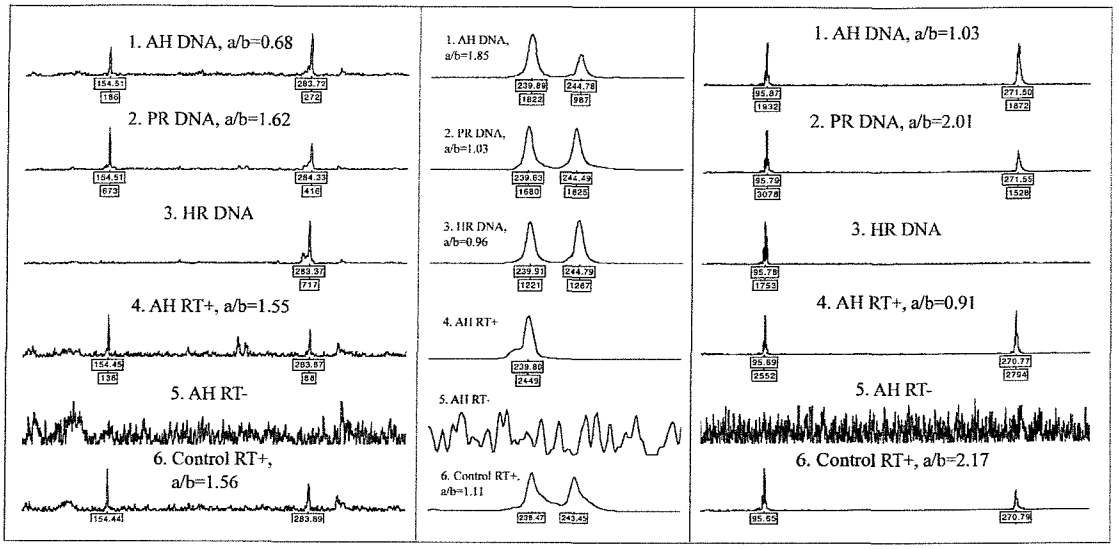


Fig. 16. Results using DNA of AH (track 1) clearly show one allele to be doubled in intensity compared to parental DNA (tracks 2 and 3), indicating the inclusion of the locus in the trisomic region 10q23.3-qter. For *HPS* and *MXI1*, analysis of cDNA from AH (track 4) is consistent with the disappearance of one maternally-inherited allele, indicating that the copy on the der(X;10) is inactive. In contrast, for *PRDX3*, analysis of cDNA from AH (track 4) gives results similar to those from DNA of AH, and significantly different to those from control cDNA (track 6) and DNA, indicating that the copy on the der(X) remains active. PCR using RNA not subjected to reverse transcription (track 5) shows this amplification is cDNA-specific. Amplification of control cDNA (track 6) is also shown. Figures below each allele represent size in base pairs and peak height respectively. The ratio 'a/b' represents the relative intensity of the smaller allele 'a' to the larger allele 'b' by peak height in each case.

Table 2 – Summary of gene locations and their transcription status on the der(X;10) in AH

Marker	Cytogenetic Band*	Distance from 10pter, Mb*	Transcriptional status on der(X;10) in AH
10q breakpoint	10q23.32	98.5-105.4	---
HPS	10q24.2	105.4	Inactive
MXI1	10q25.2	117.7	Inactive
DUSP5	10q25.2	118.0	Inactive
ABLIM	10q25.3	122.2	Inactive
PRDX3	10q25.3	128.5	Active
10qter	10qter	142.1	---

*Mapping data obtained from Human Genome Browser Gateway, December 2001 assembly (http://genome.ucsc.edu/).

3.3.1.5 Histone immunofluorescence

In AH, results gained by combined *in situ* hybridisation and immunolabelling using antisera against H3AcK14, H4AcK8 and H3Me_2K4 were concordant and demonstrated the der(X;10) to be pale staining in every cell examined, consistent with previous results. In each of 32 lymphoblastoid cells examined there was a continuous and almost complete spread of depletion of histone acetylation and H3 lysine 4 dimethylation across the translocated 10q segment (Figures 17a-i). However, in every cell examined a small region of H3/H4 acetylation and H3 lysine 4 dimethylation was clearly visible at the distal end of the translocated segment of 10q. In a proportion of cells, similar punctate staining was also apparent at the distal tip of Xp, in Xp11.2, and on the long arm of the X at approximately Xq22-25, consistent with published observations (Belyaev *et al.*, 1996; Boggs *et al.*, 2001).

3.3.1.6 Replication timing analysis

A fluorescent late-pulse BrdU assay combined with *in situ* hybridisation using whole chromosome 10 paint was performed to determine the extent of spread of late-replication in the derivative X;10 chromosome in AH. Separate lymphoblast cultures were exposed to BrdU for 3, 4, 5 and 6 hours. Incubation with BrdU for the last 5 hours of cell culture was found to give optimal incorporation, but all cultures gave identical results. In each of 50 cells examined, the derivative chromosome was late-replicating, concordant with the results of the *AR* assay. However, the late-replicating region extended only as far as the X;autosome boundary, and did not visibly spread

into the translocated segment of 10q in any of the cells examined (Figure 17j).

Figure 17 - Combined immunofluorescence and *in situ* hybridisation in AH

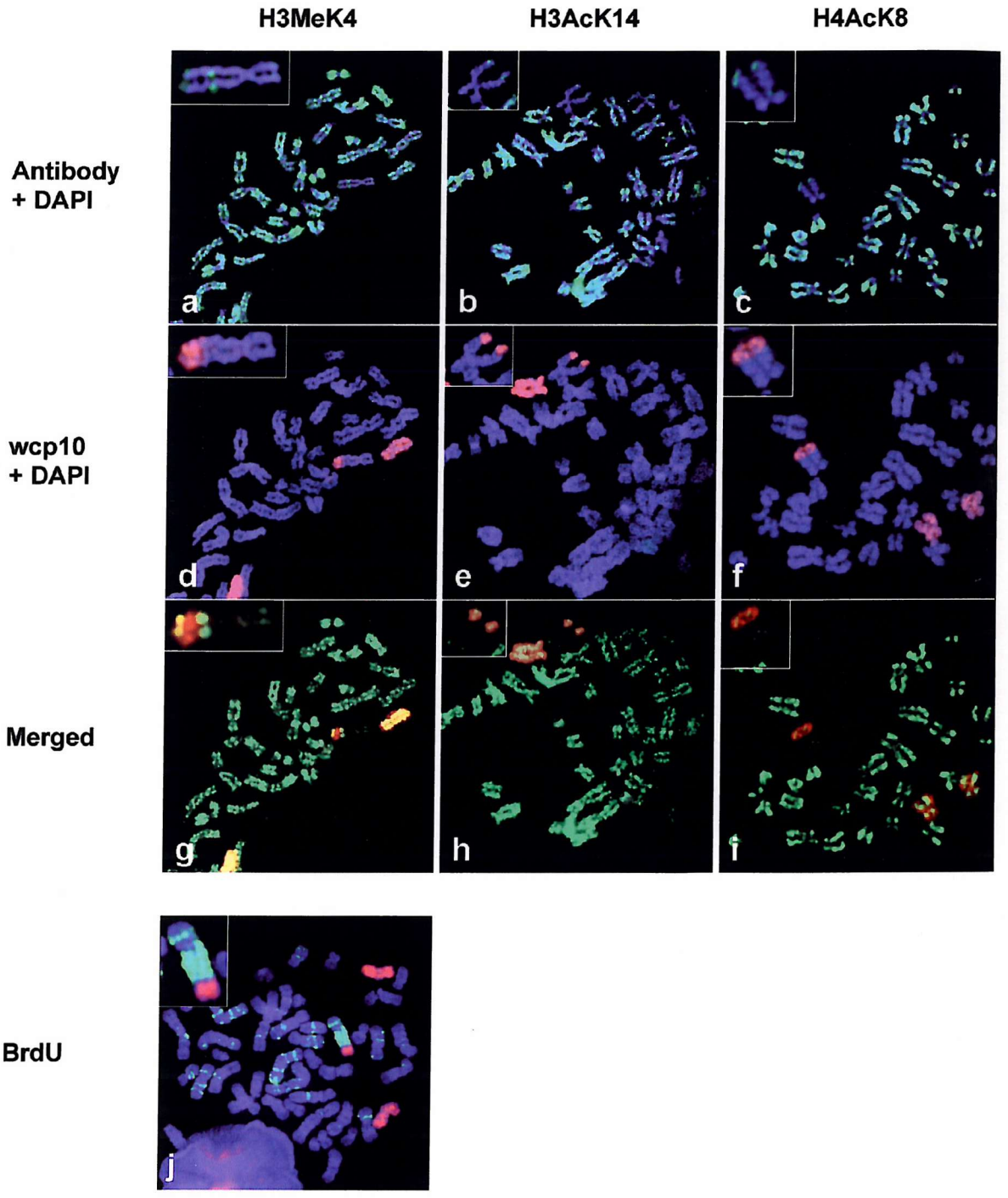


Fig. 17. (Previous page) Results of immunolabelling of histones **(a)** H3 dimethylated at lysine 4 ($H3Me_2K4$), **(b)** H3 acetylated at lysine 14 (H3AcK14), and **(c)** H4 acetylated at lysine 8 (H4AcK8) (labelled green in each case), and **(d-f)** subsequent *in situ* hybridisation using chromosome 10 paint (red). In panels **(g-i)** immunofluorescence and FISH images of each metaphase have been merged and DAPI counterstain removed to allow visualisation of overlap between the two signals on the der(X;10). An enlargement of the der(X) is shown in the top-left corner of each panel. **(a-c)** Results using antisera specific to each histone epitope are concordant and show the der(X;10) is clearly pale-staining along almost its entire length. **(g-i)** However, a small region of H3 lysine 4 dimethylation and H3/H4 acetylation is clearly visible at the distal tip of the translocated segment of 10q. Punctate staining is also apparent at the distal tip of Xp, at Xp11.2, and on the long arm of the X at approximately Xq22-25. **(j)** Results of replication timing analysis, using anti-BrdU (green), combined with *in situ* hybridisation using chromosome 10 paint (red). The X portion of the der(X;10) chromosome is late replicating, indicated by its predominant labelling with BrdU. However, each of 50 lymphoblasts examined shows a complete absence of spreading of late replication into the translocated segment of 10q, with the late replicating region appearing to define the boundary between the autosomal and X chromatin.

3.3.2 Case 2 – SP, 46,X,der(X)t(X;11)(q26.3;p12) *de novo* (pat)

3.3.2.1 Clinical details

Aged 5 years, upon clinical examination SP was noted to have mild developmental delay, large stature with growth on the 97th centile, a very long tongue although this was not considered typical macroglossia, short neck, and minor facial dysmorphisms including mild hypertelorism, mild epicanthus and slightly downward slanting eyes.

3.3.2.2 Cytogenetic analysis

GTL-banding of PHA-stimulated peripheral blood lymphocytes revealed an unbalanced X;11 translocation in the proband SP, karyotype 46,X,der(X)t(X;11)(q26.3;p12). She is therefore trisomic for 11p12-pter, and monosomic for Xq26.3-qter. Parental karyotypes were normal.

3.3.2.3 X-inactivation, parental origin, and breakpoint analysis

Methylation analysis at the androgen receptor (*AR*) locus demonstrated a completely skewed X-inactivation pattern in both peripheral blood and EBV-transformed lymphocytes of SP, with the paternally derived chromosome exclusively inactive. Consistent with this finding, PCR analysis of microsatellite markers demonstrated the

der(X;11) to be of paternal origin, with breakpoints lying between DXS1187 and DXS1062, and D11S4102 and D11S1355 respectively.

3.3.2.4 Gene expression analysis

Heterozygous polymorphisms were identified in eleven genes within the translocated segment 11p12-pter in SP, and were used for allele-specific RT-PCR of RNA extracted from EBV-transformed lymphoblasts. Results demonstrate an apparently continuous spread of gene silencing across nearly the entire translocated 11p segment (Table 3). The transcriptional status of five of these genes (*HRAS*, *TSSC3*, *TAF2H*, *LMO2* and *PDX1*) was also analysed using RNA extracted from peripheral blood of SP, with concordant results.

For *HRAS*, in cDNA of SP the relative intensity of one paternally-derived allele was approximately 50% that seen in the DNA of SP, and in the DNA and cDNA of normal controls. This indicates either an approximate halving of the level of transcription of *HRAS* on the der(X;11) in every cell, or alternatively mosaic silencing of this gene, with inactivation occurring in some 50% of cells. As previous studies of the nearby imprinted locus *H19* in SP have shown it to be methylated at levels intermediate between normal controls and individuals with paternal uniparental disomy of chromosome 11 (Reik *et al.*, 1994), it seems likely that both *HRAS* and *H19* display a mosaic pattern of inactivation.

Table 3 – Summary of gene locations and their transcription and CpG island methylation status on the der(X;11) in SP

Marker	Cytogenetic Band*	Physical Distance from 11pter, Mb*	Transcriptional / methylation status on the der(X;11) in SP
11p telomere	11pter	0	---
MRPL23	11p15.5	0.05	Active
PSMD13	11p15.5	0.2	Active (not highly methylated)
H19	11p15.5	0.3	50% methylated#
HRAS	11p15.5	1.0	50% Inactive (not highly methylated)
TSSC3	11p15.5	1.9	Inactive
RRM1	11p15.4	3.3	Inactive
SMPD1	11p15.4	5.7	Inactive
TAF2H	11p15.4	5.9	Inactive (highly methylated)
PDE3B	11p15.2	12.6	Inactive
SAA1	11p15.1	17.6	Inactive
LMO2	11p13	34.6	Inactive
PDX1	11p13	35.7	Inactive (highly methylated)
11p breakpoint	11p12	37.7-46.3	---

*Mapping data obtained from Human Genome Browser Gateway, December 2001 assembly (http://genome.cse.ucsc.edu/). # (Reik *et al.*, 1994).

3.3.2.5 CpG island methylation analysis

CpG islands were identified in the 5' regions of *PDX1*, *TAF2H, HRAS,* and *PSMD13,* and their methylation status investigated by PCR following *HpaII* or *CfoI* digestion of genomic DNA. Analysis of control individuals showed that the CpG island of each gene was unmethylated on normal chromosome 11. In contrast, analysis of DNA extracted from both peripheral blood and lymphoblasts of SP showed the presence of high levels of methylation at the CpG islands of *PDX1* (inactive) and *TAF2H* (inactive) (Fig. 18). However, no methylation was detected at the CpG islands of *HRAS* (approximately 50% inactive) or *PSMD13* (active) by either *HpaII* or *CfoI* analysis although, because of the presence of multiple recognition sites for each of these enzymes within both amplicons, these results do not exclude the presence of partial methylation (Appendix 8).

Figure 18 - Methylation analysis of four CpG islands within 11p12-pter in SP

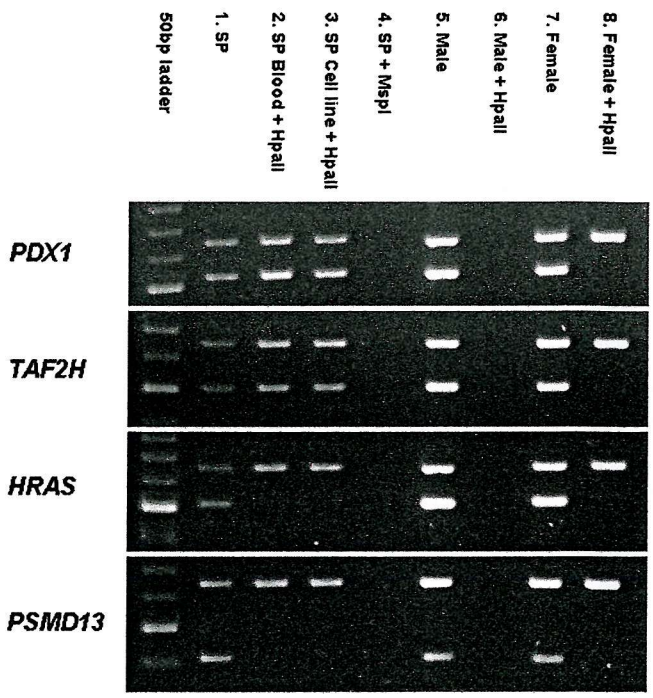


Fig. 18. Genomic DNA was either undigested, digested with the methylation-sensitive restriction enzyme *HpaII*, or with its methylation-insensitive isoschizomer *MspI*. Digests were then coamplified using primers spanning the CpG island of each gene (lower band in each case) and control primers spanning the CpG island of *PGK1* (upper band), an X-linked gene which is known to be methylated on the inactive X and unmethylated on the active X (Gilbert and Sharp, 1999). Following digestion with *HpaII* only methylated DNA remains intact and available as template in the subsequent PCR reaction. Analysis of control male and female DNA (tracks 5-8) shows the CpG island of each gene to be unmethylated on normal chromosome 11, represented by their failure to amplify following *HpaII* digestion. As the single active X chromosome of males is also unmethylated the control *PGK1* amplicon does not amplify following *HpaII* digestion either (track 6). In contrast, PCR of DNA extracted from both peripheral blood (track 2) and lymphoblasts (track 3) of SP which has been digested with *HpaII* still amplifies the CpG islands of the inactivated 11p genes *PDX1* and *TAF2H*, but does not following digestion with *MspI* (track 4), demonstrating the presence of high levels of methylation at these loci. However, following *HpaII* digestion (tracks 2 and 3), no amplification of the CpG islands of *HRAS* or *PSMD13* occurs using DNA of SP, indicating that some or all of the recognition sites for *HpaII* within both of these amplicons are unmethylated. Identical results at each of these four loci were also obtained when *HpaII* was substituted with *CfoI*. This methylation sensitive restriction enzyme has a different recognition sequence to *HpaII* and thus examines the methylation status of a different subset of CpG dinucleotides.

3.3.2.6 Histone immunofluorescence

Immunolabelling of histone H3 acetylated at lysine 14 (H3AcK14), H4 acetylated at lysine 8 (H4AcK8), or H3 dimethylated at lysine 4 ($H3Me_2K4$), was combined with *in situ* hybridisation using whole chromosome 11 paint to determine the extent of spread of each histone modification in the der(X;11) chromosome in SP. Results using antisera specific to each histone epitope were concordant, and showed that in the 36 lymphoblasts examined, the der(X;11) was clearly depleted of histone acetylation and H3 lysine 4 dimethylation along almost its entire length (Figs. 19a-i). However, in every cell examined a small punctate region of H3/H4 acetylation and H3 lysine 4 dimethylation was clearly visible at the distal tip of the translocated segment of 11p. In a proportion of cells, similar punctate staining was also apparent at the distal tip of Xp, in Xp11.2, and on the long arm of the X at approximately Xq22-25, consistent with previous observations (Belyaev *et al.*, 1996; Boggs *et al.*, 2001).

3.3.2.7 Replication timing analysis

A fluorescent late-pulse BrdU assay combined with *in situ* hybridisation using whole chromosome 11 paint was performed to determine the extent of spread of late-replication in the der(X;11) chromosome in SP. Results are shown in Figure 19j,k. In each of 50 cells examined in both peripheral blood and lymphoblasts, the derivative chromosome was late replicating, concordant with the results of the *AR* assay. In all cells a partial spreading of late-replication into the translocated 11p segment was observed, although some cell-to-cell variation was apparent. In most cases the late replicating region extended in a continuous fashion from the X chromatin across approximately half to two-thirds of the autosomal segment. However, variations in the extent of spread of late-replication occurred. In some cells the late-replicating region extended over almost the entire 11p segment, while in others BrdU staining was evident only on the most proximal portion of the translocated 11p. However, overall there was good concordance in the extent of spread of late-replication observed between both peripheral blood and lymphoblasts (Figs. 19j,k).

Figure 19 - Combined immunofluorescence and *in situ* hybridisation in SP

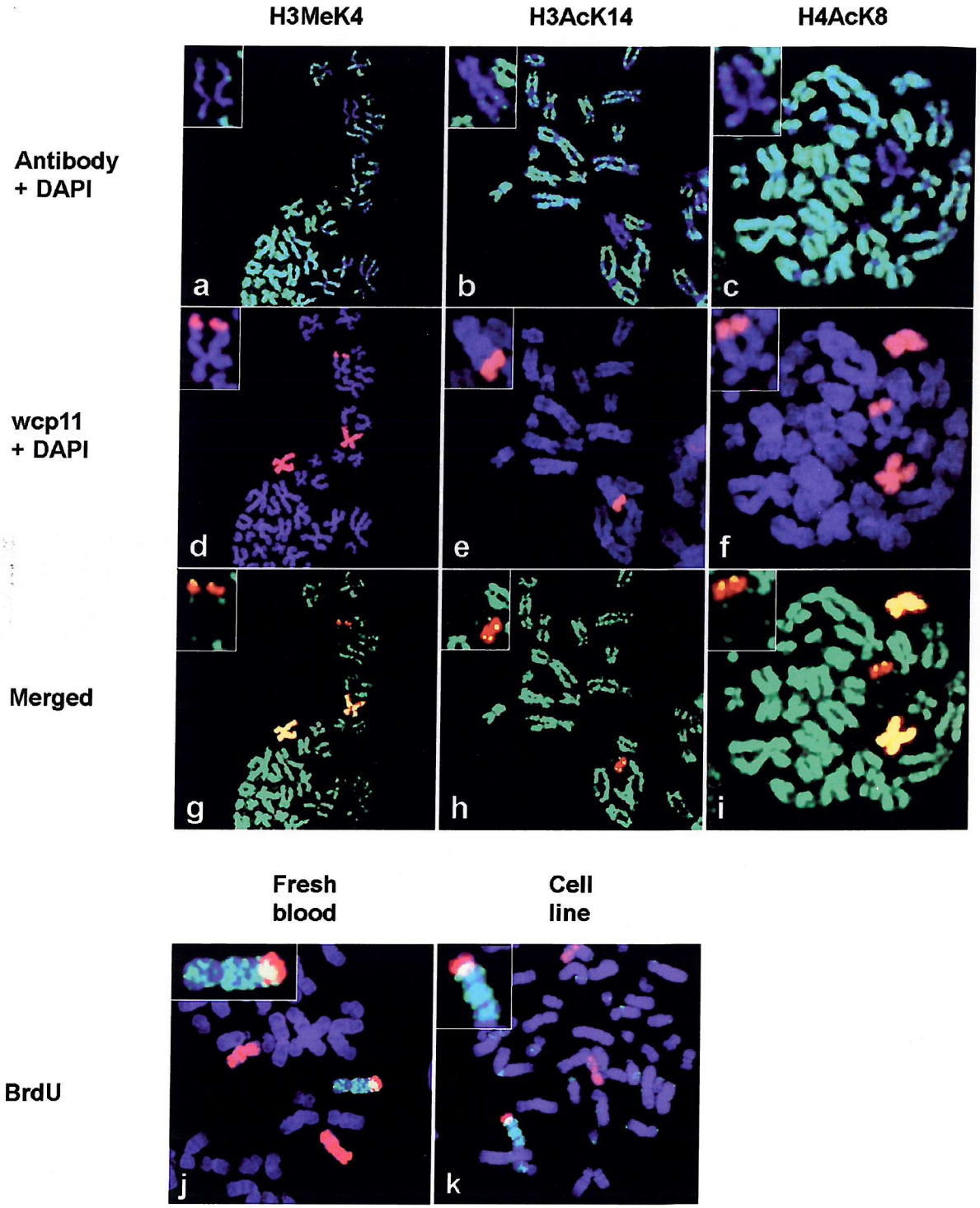


Fig. 19. (Previous page) Results of immunolabelling of histones **(a)** H3 dimethylated at lysine 4 ($H3Me_2K4$), **(b)** H3 acetylated at lysine 14 (H3AcK14), and **(c)** H4 acetylated at lysine 8 (H4AcK8) (labelled green in each case), and **(d-f)** subsequent *in situ* hybridisation using chromosome 11 paint (red). In panels **(g-i)** immunofluorescence and FISH images of each metaphase have been merged and DAPI counterstain removed to allow visualisation of overlap between the two signals on the der(X;11). An enlargement of the der(X) is shown in the top-left corner of each panel. **(a-c)** Results using antisera specific to each histone epitope are concordant and show the der(X;11) is clearly pale-staining along almost its entire length. **(g-i)** However, a small punctate region of H3 lysine 4 dimethylation and H3/H4 acetylation is clearly visible at the distal tip of the translocated segment of 11p. Punctate staining is also apparent at the distal tip of Xp and on the long arm of the X at approximately Xq22-25. **(j,k)** Results of replication timing analysis, using anti-BrdU (green), combined with *in situ* hybridisation using chromosome 11 paint (red). The der(X;11) chromosome is late replicating, indicated by its predominant labelling with BrdU. In both **(j)** peripheral blood, and **(k)** lymphoblasts, results are concordant and show a partial spreading of late-replication into approximately half of the translocated segment of 11p, visualised as the overlap between anti-BrdU (green), and chromosome 11 paint (red), pseudocoloured yellow.

3.3.3 Case 3 - SR, 46,X,der(X)t(X;7)(q27.3;q22.3) mat

3.3.3.1 Clinical details

At the age of 16, SR displayed a number of severe phenotypic abnormalities with profound motor and developmental delay and severe mental retardation. She has no speech, and shows a variety of orthopaedic disorders with scoliosis and dislocated hips.

3.3.3.2 Cytogenetic analysis

GTL-banding of PHA-stimulated peripheral blood lymphocytes revealed an unbalanced X;7 translocation in the proband SR, karyotype 46,X,der(X)t(X;7)(q27.3;q22.3). She is therefore trisomic for 7q22.3-qter, and monosomic for Xq27.3-qter. Her mother was also found to carry the same translocation, but in a balanced form, karyotype 46,X,t(X;7)(q27.3;q22.3).

3.3.3.3 X-inactivation, parental origin, and breakpoint analysis

Methylation analysis at the *AR* locus demonstrated a completely skewed X-inactivation pattern in both peripheral blood and lymphoblasts of SR, with the maternally derived chromosome exclusively inactive. PCR analysis of microsatellite

markers also demonstrated the der(X;7) to be of maternal origin, with breakpoints lying between DXS998 and DXS1684, and D7S2420 and D7S523 respectively.

3.3.3.4 Gene expression analysis

Heterozygous polymorphisms were identified in three genes within the translocated segment 7q22.3-qter in SR, and were used for allele-specific RT-PCR. Results are summarised in Table 4.

For *CNTNAP2*, in cDNA of SR the relative intensity of one maternally derived allele was approximately 30% that seen in the DNA of SR, and in the DNA and cDNA of normal controls, consistent with a mosaic pattern of inactivation.

Table 4 – Summary of gene locations and their transcription status on the der(X;7) in SR

Marker	Cytogenetic Band*	Physical Distance from 7pter, Mb*	Transcriptional status on the der(X;7) in SR
7q breakpoint	7q22.3	110.3-114.1	---
TES	7q31.2	116.2	Inactive
CNTNAP2	7q35	151.6	~70% Inactive
KCNH2	7q36.1	154.1	Active
7q telomere	7qter	163.8	---

*Mapping data obtained from Human Genome Browser Gateway, December 2001 assembly (http://genome.cse.ucsc.edu/).

3.3.3.5 Histone immunofluorescence

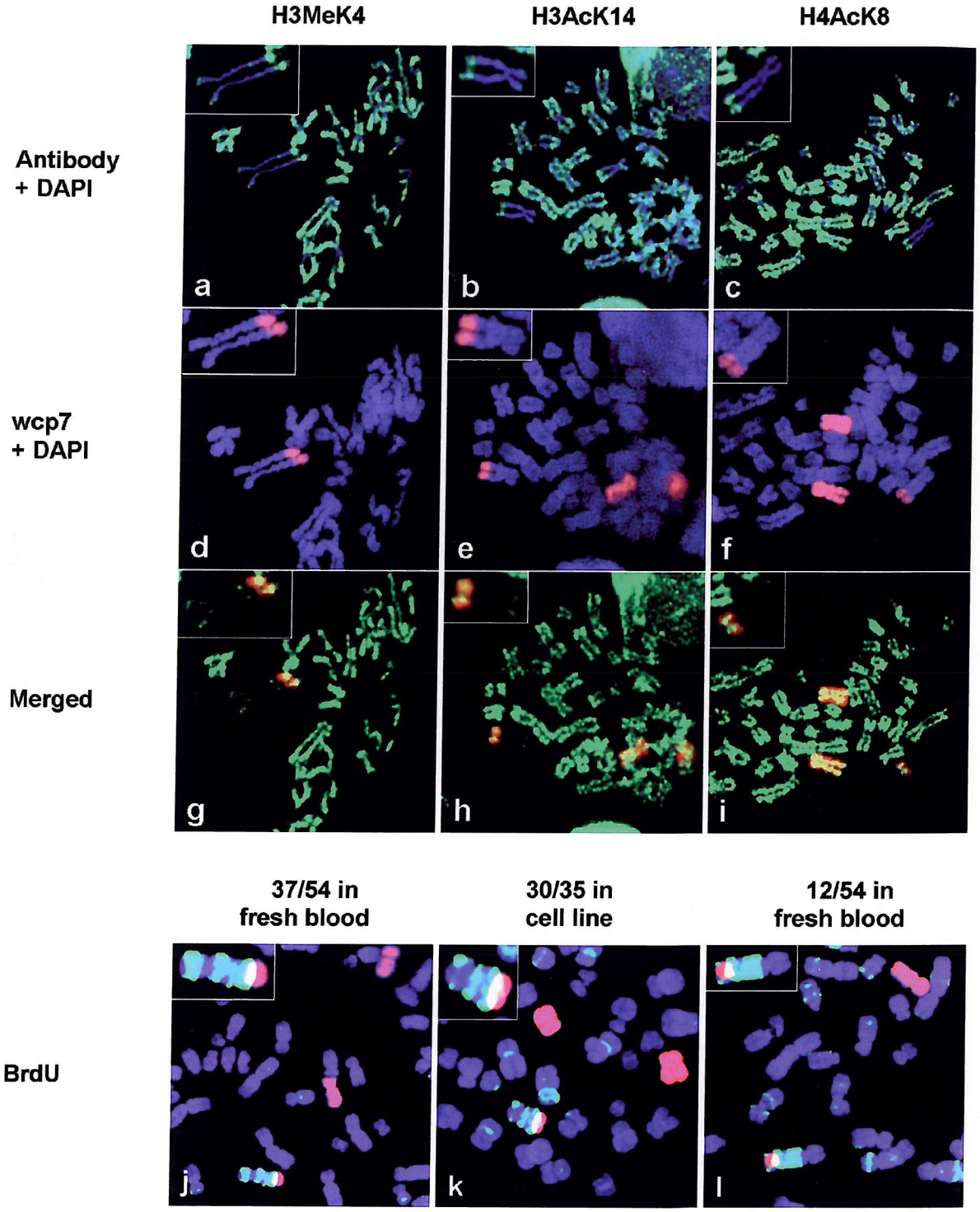
Results gained by immunolabelling using antisera against H3AcK14, H4AcK8 and $H3Me_2K4$ combined with *in situ* hybridisation were concordant and demonstrated the der(X;7) to be exclusively inactive. In each of 27 lymphoblastoid cells examined there was a partial and continuous spread of histone hypoacetylation and depletion of H3 lysine 4 dimethylation across approximately one third of the 7q segment (Figs. 20a-i). However, the remaining distal portion of the translocated 7q segment was indistinguishable from the corresponding regions on the two normal chromosome 7 homologues. Residual antibody staining was also apparent at the distal tip of Xp, in Xp11.2, and on the long arm of the X at approximately Xq22-25 in a proportion of cells.

3.3.3.6 Replication timing analysis

Replication timing analysis demonstrated that in each of 54 peripheral blood and 35 lymphoblastoid cells examined, the der(X;7) was late replicating, concordant with the results of the *AR* assay. However, the extent of spread of late replication into the translocated segment of 7q varied among cells (Figs. 20j-l). 49/54 peripheral blood cells and 30/35 lymphoblastoid cells showed a continuous spread of late-replication across approximately one third of the 7q segment. Among these 49 peripheral blood cells there was also evidence for a discontinuous spread of late-replication where the resolution of the metaphases was higher. In addition to the late replicating region on the proximal portion of 7q, 12 cells also showed an isolated focus of BrdU staining at or just below the telomere (Fig. 20l). In the remaining 5/54 blood cells and 5/35 lymphoblasts the late-replicating region extended only as far as the X;autosome boundary, and did not visibly spread into the translocated segment of 7q.

Fig. 20. (Following page) Results of immunolabelling of histones **(a)** H3 dimethylated at lysine 4 ($H3Me_2K4$), **(b)** H3 acetylated at lysine 14 (H3AcK14), and **(c)** H4 acetylated at lysine 8 (H4AcK8) (labelled green in each case), and **(d-f)** subsequent *in situ* hybridisation using chromosome 7 paint (red). In panels **(g-i)** immunofluorescence and FISH images of each metaphase have been merged and DAPI counterstain removed to allow visualisation of overlap between the two signals on the der(X;7). An enlargement of the der(X) is shown in the top-left corner of each panel. **(a-c)** Results using antisera specific to each histone epitope are concordant and show the der(X;7) is predominantly pale-staining, with **(g-i)** a partial and continuous spread of histones depleted of acetylation and dimethylation across approximately one third of the 7q segment. However, the remaining distal portion of the translocated 7q segment appears indistinguishable from the corresponding regions on the two normal chromosome 7 homologues. Punctate staining is also apparent at the distal tip of Xp, in Xp11.2, and on the long arm of the X at approximately Xq22-25. **(j-l)** Results of replication timing analysis, using anti-BrdU (green), combined with *in situ* hybridisation using chromosome 7 paint (red). The der(X;7) chromosome is late replicating, indicated by its predominant labelling with BrdU. In both **(j)** 37 of 54 peripheral blood cells examined, and **(k)** 30 of 35 lymphoblasts, results are concordant and show a partial spreading of late-replication into approximately one third to one half of the translocated segment of 7q, visualised as the overlap between anti-BrdU (green), and chromosome 7 paint (red), pseudocoloured yellow. **(l)** In some peripheral blood cells where the resolution of the metaphases is higher a discontinuous spread of late-replication is apparent. In addition to the late replicating region on the proximal portion of 7q, 12 of 54 blood cells also show an isolated focus of BrdU staining at or just below the translocated 7q telomere.

Figure 20 - Combined immunofluorescence and *in situ* hybridisation in SR



3.3.4 Case 4 – AL0044, 46,X,der(X)t(X;6)(p11.2;p21.1) mat

Studies in AL0044 have been reported previously (Keohane *et al.*, 1999). Briefly, these found that hypoacetylation of histone H4, late-replication and *XIST* RNA were coincident and apparently excluded from the translocated segment of 6p on the der(X;6).

3.3.4.1 Clinical details

AL0044 has mild developmental delay, learning difficulties (IQ=75), and short stature.

3.3.4.2 Cytogenetic analysis

GTL-banding of PHA-stimulated peripheral blood lymphocytes revealed an unbalanced X;6 translocation in the proband AL0044, karyotype 46,X,der(X)t(X;6)(p11.2;p21.1). She is therefore trisomic for 6p21.1-pter, and monosomic for Xp11.2-pter. Her mother was also found to carry the same translocation, but in a balanced form, karyotype 46,X,t(X;6)(p11.2;p21.1).

3.3.4.3 X-inactivation, parental origin, and breakpoint analysis

Methylation analysis at the *AR* locus demonstrated a completely skewed X-inactivation pattern in lymphoblasts of AL0044, with the maternally derived chromosome exclusively inactive. PCR analysis of polymorphic markers also demonstrated the der(X;6) to be of maternal origin, with breakpoints lying between DXS1058 and DXS8083, and *ZNF76* and D6S269 respectively.

3.3.4.4 Gene expression analysis

Heterozygous polymorphisms were identified in nine genes within the translocated segment 6p21.1-pter in AL0044, and were used for allele-specific RT-PCR. Results show a discontinuous spreading of gene silencing across the entire translocated segment of 6p, with active genes interspersed among inactive genes, and are summarised in Table 5.

For *SCA1*, in cDNA of AL0044 the relative intensity of one maternally derived allele was approximately 30% that seen in the DNA of AL0044, and in the DNA and cDNA of normal controls, consistent with a mosaic pattern of inactivation.

Table 5 – Summary of gene locations and their transcription and CpG island methylation status on the der(X;6) in AL0044 and BO0566

Marker	Cytogenetic Band*	Physical Distance from 6pter, Mb*	Transcriptional / methylation status on the der(X;6) in AL0044	Transcriptional / methylation status on the der(X;6) in BO0566
6p telomere	6pter	0	---	---
IRF4	6p25.3	0.6	Inactive	n/i
SCA1	6p25.2	5.0	~70% Inactive	n/i
GMPR	6p25.2	5.3	n/i	Inactive
DEK	6p22.3	19.5	Active	n/i
HLA-F	6p21.1	39.6	Inactive (highly methylated)	n/i (highly methylated)
HCR	6p21.1	41.0	Inactive (not highly methylated)	Active (not highly methylated)
MICA	6p21.1	41.3	Inactive (highly methylated)	n/i (highly methylated)
CSNK2B	6p21.1	41.6	Active	n/i
NOTCH4	6p21.1	42.1	n/i	Inactive
HLA-DRA	6p21.1	42.3	n/i	Active
HLA-DRB5	6p21.1	42.4	Inactive	Inactive
PSMB9	6p21.1	42.7	n/i	Active
BRD2	6p21.1	42.8	Active	Active
6p breakpoint in AL0044	6p21.1	45.1-52.5	---	---
6p breakpoint in BO0566	6p12	52.5-56.6	---	---

n/i denotes marker not informative. *Mapping data obtained from Human Genome Browser Gateway, December 2001 assembly (http://genome.cse.ucsc.edu/).

3.3.4.5 CpG island methylation analysis

CpG islands were identified in the 5' regions of *HLA-F*, *HCR*, and *MICA*, and their methylation status investigated by PCR following *HpaII* or *CfoI* digestion of genomic DNA. Analysis of control individuals showed that the CpG island of each gene was unmethylated on their normal chromosome 6. In contrast, analysis of DNA extracted from lymphoblasts of AL0044 showed the presence of high levels of methylation at the CpG islands of *HLA-F* (inactive) and *MICA* (inactive) (Fig. 21). However, no methylation was detected at the CpG island of *HCR* (inactive) by either *HpaII* or *CfoI* analysis although, because of the presence of multiple recognition sites for each of

these enzymes within the amplicon, this does not exclude the presence of partial methylation.

Figure 21 - Methylation analysis of three CpG islands within 6p21-pter in AL0044 and BO0566

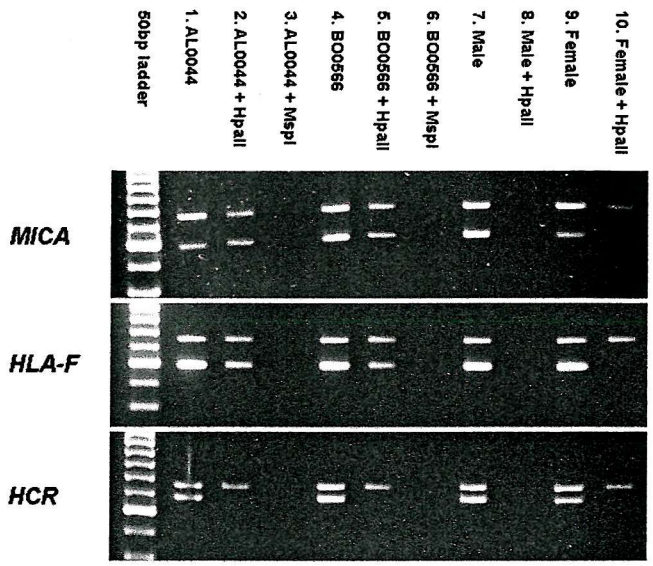


Fig. 21. Genomic DNA was either undigested, digested with the methylation-sensitive restriction enzyme *HpaII*, or with its methylation-insensitive isoschizomer *MspI*. Digests were then coamplified using primers spanning the CpG island of each gene (lower band in each case) and control primers spanning the CpG island of *PGK1* (upper band), an X-linked gene which is known to be methylated on the inactive X and unmethylated on the active X (Gilbert and Sharp, 1999). Following digestion with *HpaII* only methylated DNA remains intact and available as template in the subsequent PCR reaction. Analysis of control male and female DNA (tracks 7-10) shows the CpG island of each gene to be unmethylated on normal chromosome 6, represented by their failure to amplify following *HpaII* digestion. As the single active X chromosome of males is also unmethylated the control *PGK1* amplicon does not amplify following *HpaII* digestion either (track 10). In contrast, PCR of DNA from both AL0044 (track 2) and BO0566 (track 5) which has been digested with *HpaII* still amplifies the CpG islands of the inactivated 6p genes *MICA* and *HLA-F*, but does not following digestion with *MspI* (tracks 3 and 6), demonstrating the presence of high levels of methylation at these loci. However, following *HpaII* digestion (tracks 2 and 5), no amplification of the CpG islands of *HCR* occurs using DNA of AL0044 or BO0566, indicating that some or all of the recognition sites for *HpaII* within both of these amplicons are unmethylated. Identical results at each of these three loci were also obtained when *HpaII* was substituted with *CfoI*. This methylation sensitive restriction enzyme has a different recognition sequence to *HpaII* and thus examines the methylation status of a different subset of CpG dinucleotides.

3.3.4.6 Histone immunofluorescence

Results gained by combined *in situ* hybridisation and immunolabelling using antisera against H3AcK14, H4AcK8 and H3Me_2K4 were concordant and demonstrated the der(X;6) to be exclusively inactive. In each of 23 lymphoblastoid cells examined there was a discontinuous spread of histone hypoacetylation and depletion of H3 lysine 4 dimethylation across the translocated 6p segment (Figs. 22a-i). This region depleted of histone acetylation and H3 lysine 4 dimethylation varied in size, in some instances covering up to approximately one third of the translocated autosome, but was always confined to the distal end of the translocated 6p chromatin. In every cell the more proximal translocated portion of 6p appeared indistinguishable from the corresponding regions on the two normal chromosome 6 homologues. In a proportion of cells punctate antibody staining was also apparent on the long arm of the X at approximately Xq22-25 with all three antisera.

3.3.4.7 Replication timing analysis

Replication timing analysis in lymphoblasts of AL0044 demonstrated that in each of 50 cells examined the der(X;6) was late replicating, concordant with the results of the *AR* assay. Of these, 20/50 showed a complete absence of spreading of late replication into the translocated segment of 6p, with the late replicating region appearing to define the boundary between the autosomal and X chromatin. However, in the remaining 30/50 cells there was a discontinuous spread of late replication into the translocated 6p segment, with a variably sized region of BrdU staining visible on the translocated 6p telomere, with no staining on the intervening segment (Figs. 22j,k).

Figure 22 - Combined immunofluorescence and *in situ* hybridisation in AL0044

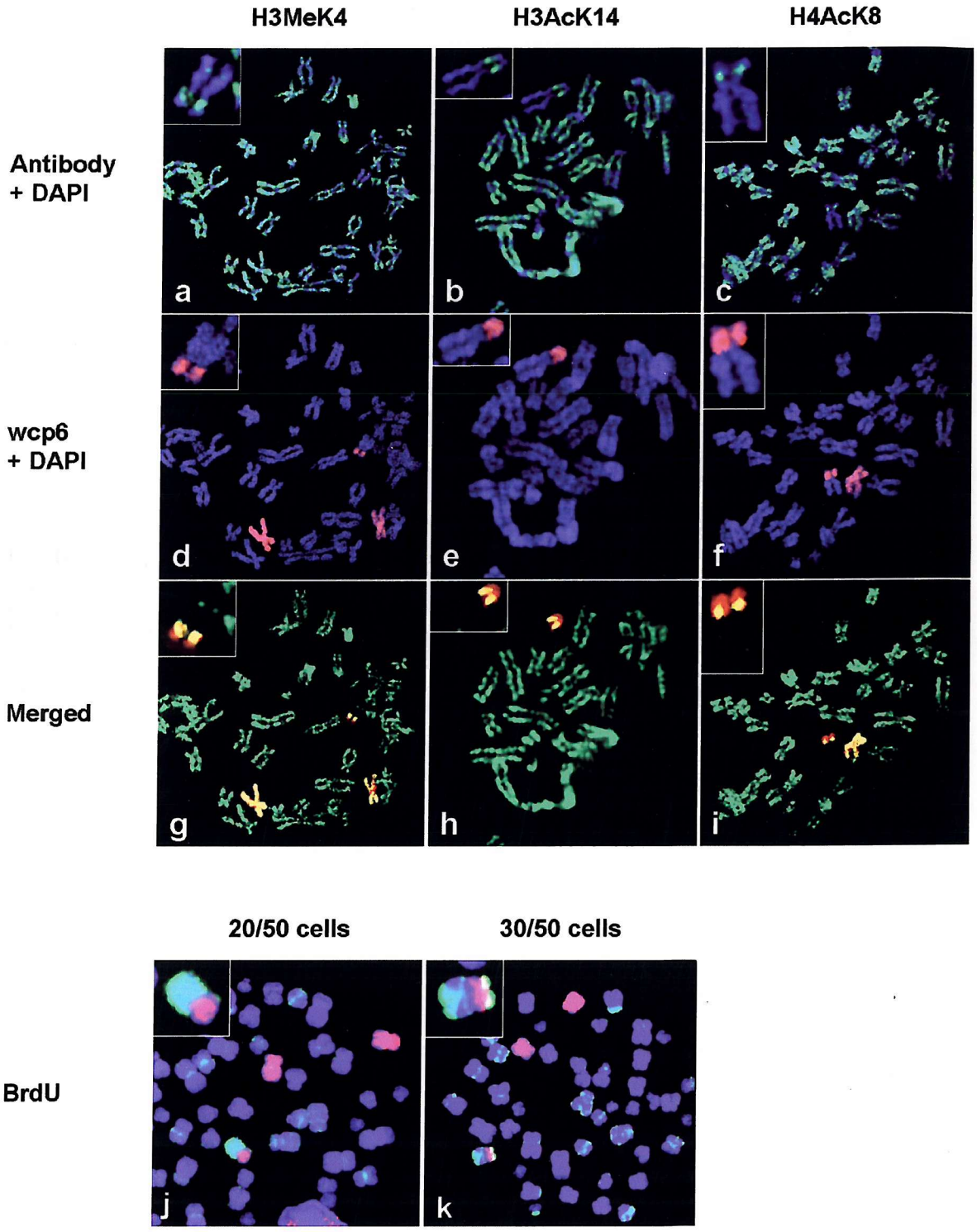


Fig. 22. (Previous page) Results of immunolabelling of histones **(a)** H3 dimethylated at lysine 4 ($H3Me_2K4$), **(b)** H3 acetylated at lysine 14 (H3AcK14), and **(c)** H4 acetylated at lysine 8 (H4AcK8) (labelled green in each case), and **(d-f)** subsequent *in situ* hybridisation using chromosome 6 paint (red). In panels **(g-i)** immunofluorescence and FISH images of each metaphase have been merged and DAPI counterstain removed to allow visualisation of overlap between the two signals on the der(X;6). An enlargement of the der(X) is shown in the top-left corner of each panel. **(a-c)** Results using antisera specific to each histone epitope are concordant and show the der(X;6) is predominantly pale-staining with a discontinuous spread of histones depleted of acetylation and dimethylation covering only the distal region of the 6p segment. **(g-i)** A variably sized pale-staining region is visible at the translocated 6p telomeric region, with no staining on the intervening segment. The remaining proximal portion of the translocated 6p segment appears indistinguishable from the corresponding regions on the two normal chromosome 6 homologues, with the proximal transition in antibody staining appearing to define the X;autosome boundary. Punctate staining is also apparent at the distal tip of Xp and on the long arm of the X at approximately Xq22-25. **(j-k)** Results of replication timing analysis, using anti-BrdU (green), combined with *in situ* hybridisation using chromosome 6 paint (red). The der(X;6) chromosome is late replicating, indicated by its predominant labelling with BrdU. **(j)** 20 of 50 lymphoblasts examined show a complete absence of spreading of late replication into the translocated segment of 6p, with the late replicating region appearing to define the boundary between the autosomal and X chromatin. **(k)** However, in 30 of 50 cells examined there was a discontinuous spread of late replication into the translocated 6p segment, with a variably sized region of BrdU staining visible at the translocated 6p telomeric region, with no staining on the intervening segment.

3.3.5 Case 5 – BO0566, 46,X,der(X)t(X;6)(q28;p12) *de novo* (pat)

3.3.5.1 Clinical details

BO0566 failed to thrive as a baby, with poor feeding and diarrhoea. At 16 months of age, her head circumference was on the 3rd centile, with a small anterior fontanelle, mild hypertelorism, thin lips, low set ears and a left ear lobe crease. At 3 years of age, she was able to sit unaided, had babbling speech and had to be fed liquidised food. At 4½ years of age she was able to stand, but had no speech and was being fed via an NG tube. She developed epilepsy at 8 years of age. Aged 10, she has reasonable motor skills, has no speech or sign language, and is able to take a few steps but mostly 'bottom shuffles'. She has some behavioural problems, nipping and scratching, with inappropriate laughter. Generally her development is similar to a 9-12 month old. She also has a heart murmur, severe swallowing difficulties and experiences recurrent chest infections.

3.3.5.2 Cytogenetic analysis

GTL-banding of PHA-stimulated peripheral blood lymphocytes revealed an unbalanced X;6 translocation in the proband BO0566, karyotype 46,X,der(X)t(X;6)(q28;p12). She is therefore trisomic for 6p12-pter, and monosomic for Xq28-qter. Parental karyotypes were normal.

3.3.5.3 X-inactivation, parental origin, and breakpoint analysis

Methylation analysis at the *AR* locus demonstrated a completely skewed X-inactivation pattern in lymphoblasts of BO0566, with the paternally derived chromosome exclusively inactive. Consistent with this finding, PCR analysis of polymorphic markers demonstrated the der(X;6) to be of paternal origin, with breakpoints lying between DXS1073 and DXS1108, and D6S269 and *GCLC* respectively.

3.3.5.4 Gene expression analysis

Heterozygous polymorphisms were identified in seven genes within the translocated segment 6p12-pter in BO0566, and were used for allele-specific RT-PCR. As in AL0044, results show a discontinuous spreading of gene silencing across the translocated segment of 6p, with active genes interspersed among inactive genes, and are summarised in Table 5.

3.3.5.5 CpG island methylation analysis

Analysis of DNA extracted from lymphoblasts of BO0566 showed the presence of high levels of methylation at the CpG islands of *HLA-F* and *MICA* (transcriptional status unknown). No methylation was detected at the CpG island of *HCR* (active) by either *HpaII* or *CfoI* analysis (Fig. 21).

3.3.5.6 Histone immunofluorescence

Results gained by combined *in situ* hybridisation and immunolabelling using antisera against H3AcK14, H4AcK8 and H3Me_2K4 were concordant and demonstrated the der(X;6) to be exclusively inactive. In each of 24 lymphoblastoid cells examined there was a complete absence of spreading of histone hypoacetylation and depletion of H3 lysine 4 dimethylation into the translocated segment of 6p, with the transition in antibody staining appearing to define the X;autosome boundary (Figs. 23a-i). In a

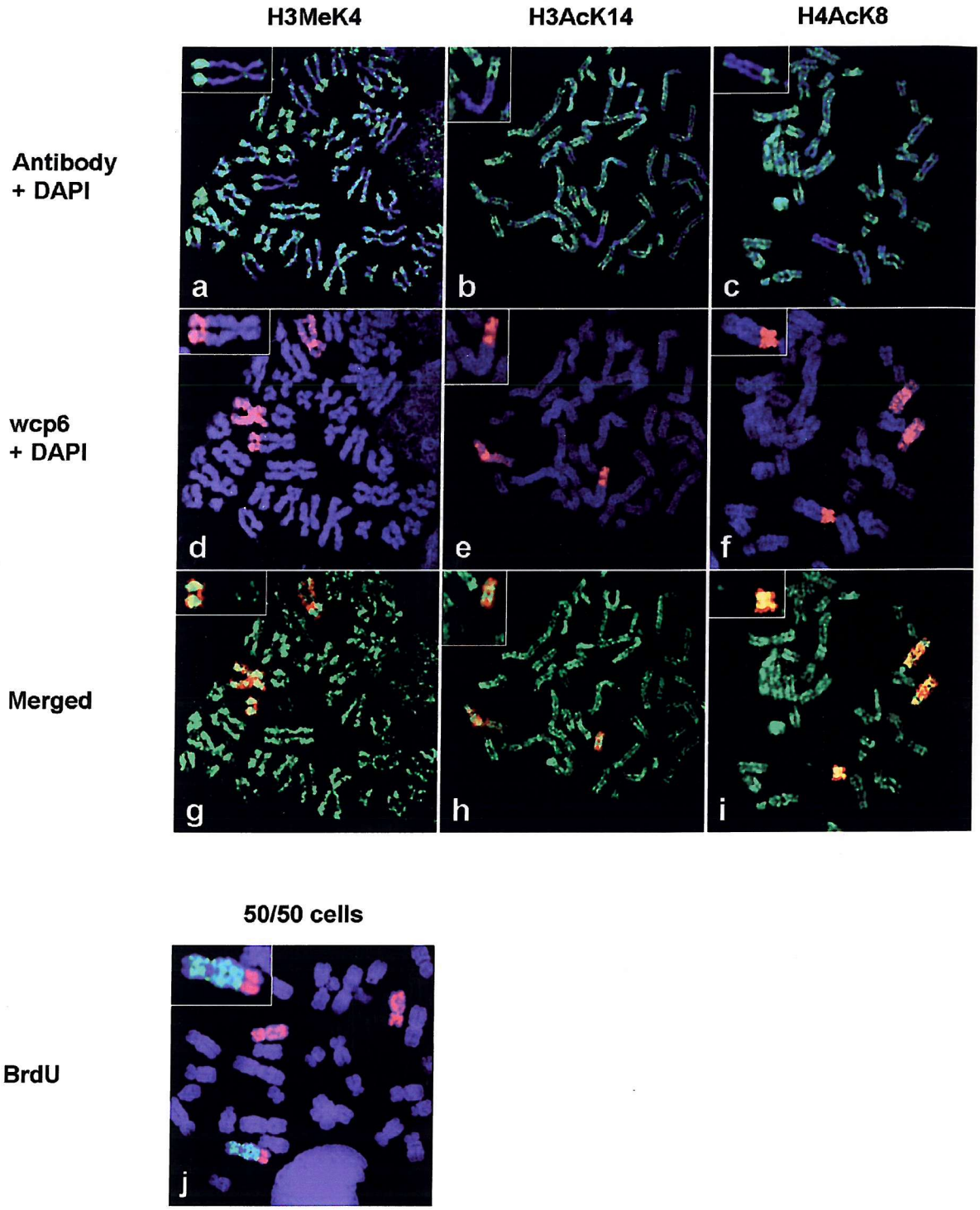
proportion of cells punctate antibody staining was also apparent at the distal tip of Xp, in Xp11.2, and on the long arm of the X at approximately Xq22-25 with all three antisera.

3.3.5.7 Replication timing analysis

Replication timing analysis in lymphoblasts of BO0566 demonstrated that in each of 50 cells examined the der(X;6) was late replicating. In every cell there was a complete absence of spreading of late replication into the translocated segment of 6p, with the late replicating region appearing to define the X;autosome boundary (Fig. 23j).

Fig. 23. (Following page) Results of immunolabelling of histones **(a)** H3 dimethylated at lysine 4 ($H3Me_2K4$), **(b)** H3 acetylated at lysine 14 (H3AcK14), and **(c)** H4 acetylated at lysine 8 (H4AcK8) (labelled green in each case), and **(d-f)** subsequent *in situ* hybridisation using chromosome 6 paint (red). In panels **(g-i)** immunofluorescence and FISH images of each metaphase have been merged and DAPI counterstain removed to allow visualisation of overlap between the two signals on the der(X;6). An enlargement of the der(X) is shown in the top-left corner of each panel. **(a-c)** Results using antisera specific to each histone epitope are concordant and show the der(X;6) is predominantly pale-staining. **(g-i)** However, there is a complete absence of spreading of histones depleted of acetylation and dimethylation into the translocated segment of 6p, with the transition in antibody staining appearing to define the X;autosome boundary. In some cells punctate antibody staining is also apparent at the distal tip of Xp, in Xp11.2, and on the long arm of the X at approximately Xq22-25 with all three antisera. **(j)** Results of replication timing analysis, using anti-BrdU (green), combined with *in situ* hybridisation using chromosome 6 paint (red). The X portion of the der(X;6) chromosome is late replicating, indicated by its predominant labelling with BrdU. However, each of 50 lymphoblasts examined shows a complete absence of spreading of late replication into the translocated segment of 6p, with the late replicating region appearing to define the boundary between the autosomal and X chromatin.

Figure 23 - Combined immunofluorescence and *in situ* hybridisation in BO0566



3.3.6 *XIST* RNA *in situ* hybridisation

Results of RNA *in situ* hybridisation in lymphoblasts of AL0044, a normal control, SP and SR using a probe specific for *XIST* are shown in Figs. 24a-d, respectively. Because *XIST* RNA dissociates from the human X chromosome in prophase, analysis is of its nuclear distribution is limited to interphase.

Studies in AL0044 were performed previously (Keohane *et al.*, 1999), combining RNA FISH for *XIST* transcripts and subsequent DNA FISH using chromosome 6 paint to localise both the two normal chromosomes 6 and the translocated segment of 6p. Results showed that in 41 of 67 cells examined there was no overlap between the *XIST* RNA and chromosome 6 signals, indicating a complete lack of spread of *XIST* RNA into the translocated segment of 6p on the der(X;6) (Fig 24a). In the remaining 26/67 cells there was some overlap of the two signals. However, because the technique employs two-dimensional microscopy to view a three-dimensional cell, depending upon the orientation of each nucleus, some apparent overlap in signals would be expected in a proportion of cells. Thus within technical limits, these observations are consistent with no spreading of *XIST* RNA into the translocated 6p segment in most or all cells in AL0044.

Studies of the distribution of *XIST* RNA in lymphoblasts of SP and SR (Figs. 26c and 26d, respectively) showed that, where detectable, it had a diffuse signal compared to normal controls (Fig. 24b), indicating poor localisation to the X_i. In addition, in both SP and SR, *XIST* RNA was apparently not expressed in many cells.

Figure 24 – Results of *XIST* RNA *in situ* hybridisation

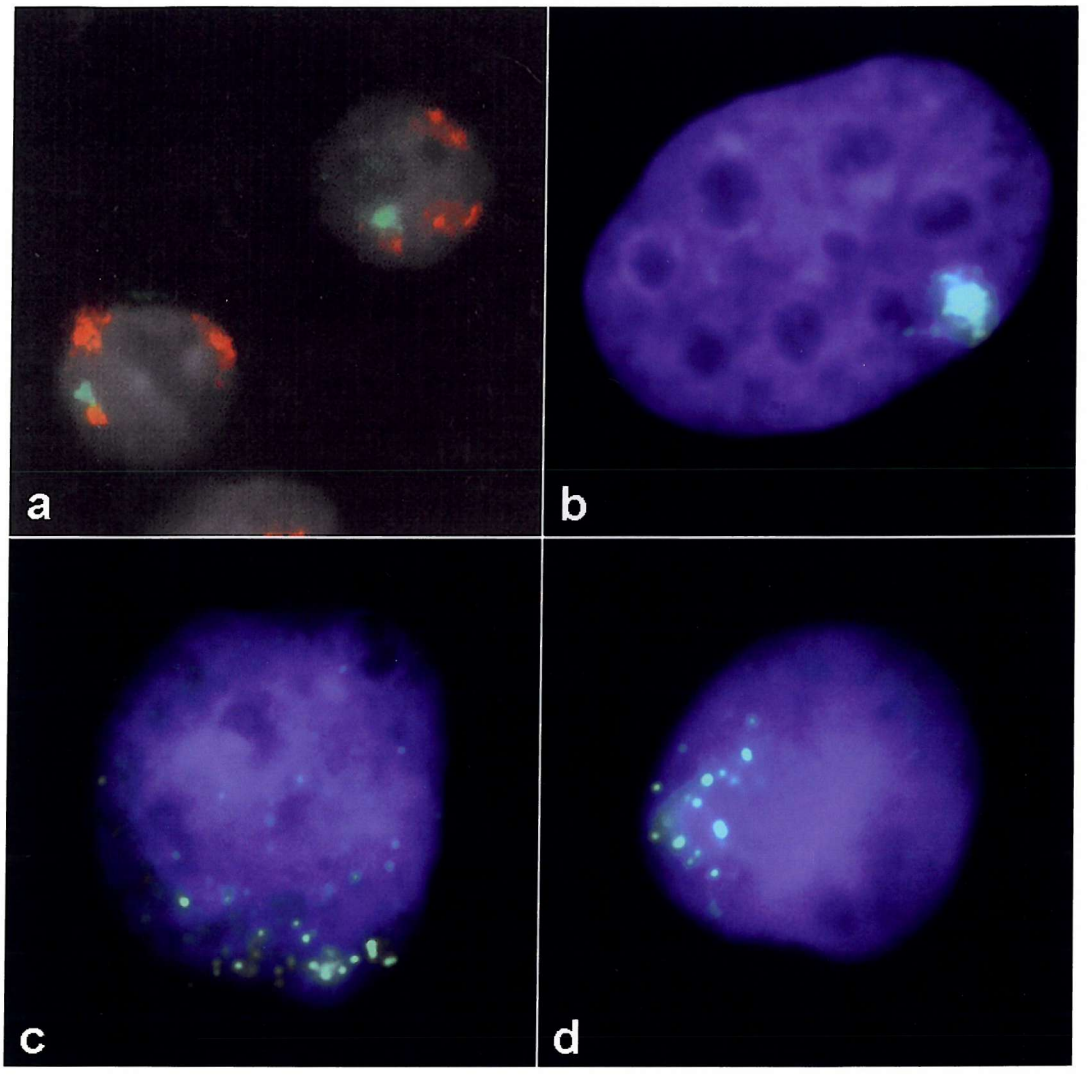


Fig. 24. Results of RNA *in situ* hybridisation using a probe specific for *XIST* (green) in **(a)** AL0044 (taken from Keohane *et al.*, 1999), **(b)** normal control, **(c)** Case 2, SP and **(d)** Case 3, SR. In **(a)** subsequent *in situ* hybridisation using a chromosome 6 paint (red) has also been performed showing the relative locations of both the two normal chromosome 6 homologues (large signals) and the translocated segment 6p21-pter (small signal adjacent to *XIST* RNA). In **(a)** AL0044 there is little or no overlap between the *XIST* RNA and chromosome 6 signals, indicating a lack of spread of *XIST* RNA into the translocated segment of 6p on the der(X;6). In both **(c)** SP and **(d)** SR *XIST* RNA has a diffuse distribution compared to **(b)** the normal control, indicating poor localisation to the X_i. In addition, in both SP and SR, *XIST* RNA was apparently not expressed in many cells.

3.3.7 LINE-1 repeat analysis of 10q24-25, 11p13-pter and 6p21.3-22.3

Results of the analysis of LINE-1 content in 10q24-25, 11p13-pter and 6p21.3-22.3 are shown in Figs. 25, 26, 27 respectively. The relative locations of autosomal genes, as listed in the Human Genome Browser, April 2001 assembly, whose transcriptional status had been determined by RT-PCR and/or methylation analysis is indicated, with inactive genes shown in red, partially active genes shown in yellow, and fully active genes shown in green. Although subjective in nature, overall there was no apparent relationship between L1 density and the spread of inactivation. Inactive genes (eg. *ABLIM*) were in some instances located in L1-poor regions, and conversely active genes (eg. *PSMB9* and *HLA-DRA*) were found within L1-rich regions. In addition, *MICA* and *CSNK2B*, which are differentially inactivated, are located adjacently within a L1-poor region.

A similar analysis of the primate-specific L1P1 to L1P5 LINE subclasses, which are particularly enriched on the X chromosome (Bailey *et al.*, 2000), showed that these have an almost identical distribution to L1 repeats as a whole.

Figure 25 – LINE-1 content in 10q24-25

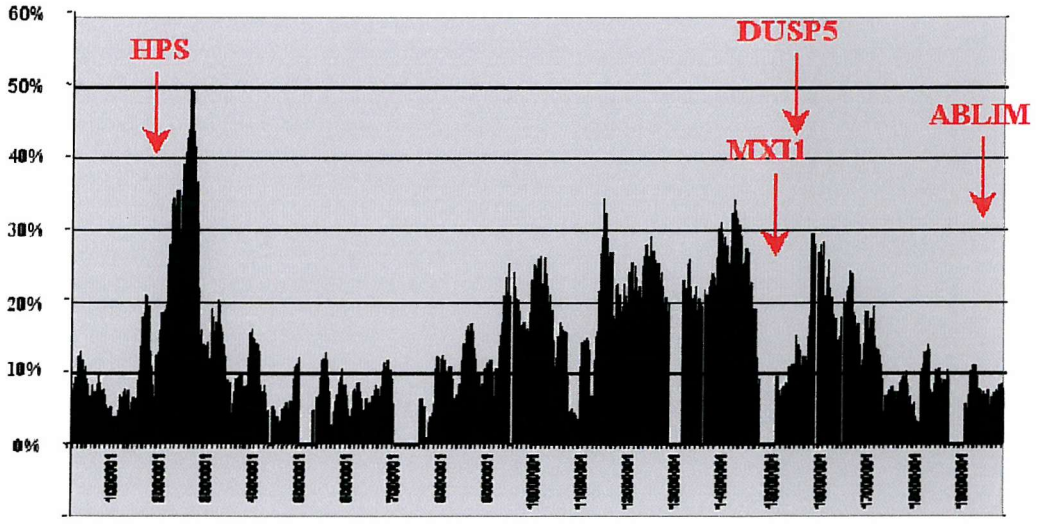


Figure 26 – LINE-1 content in 11p13-pter

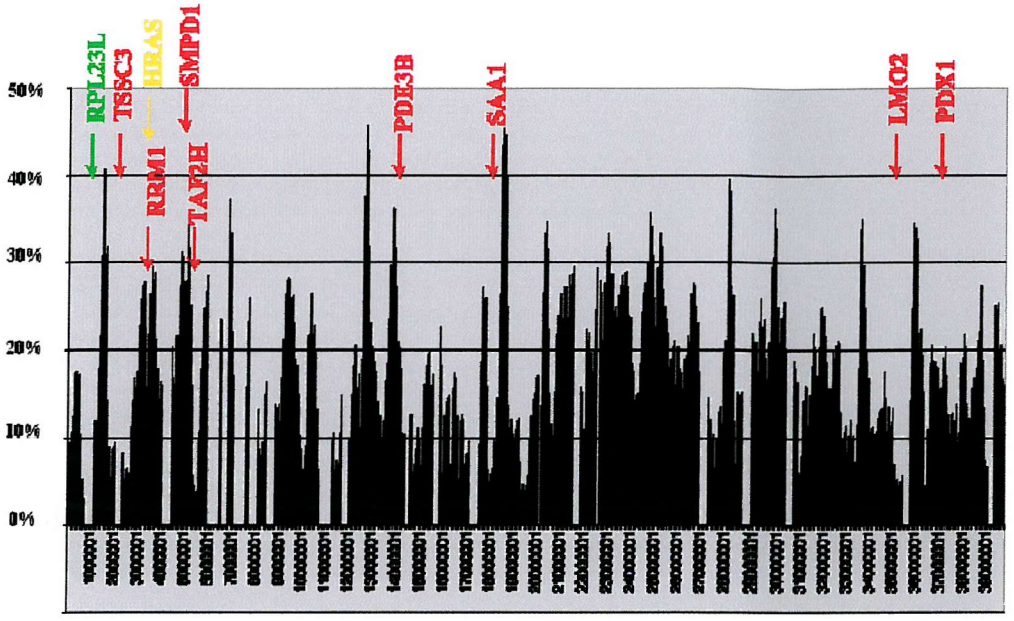
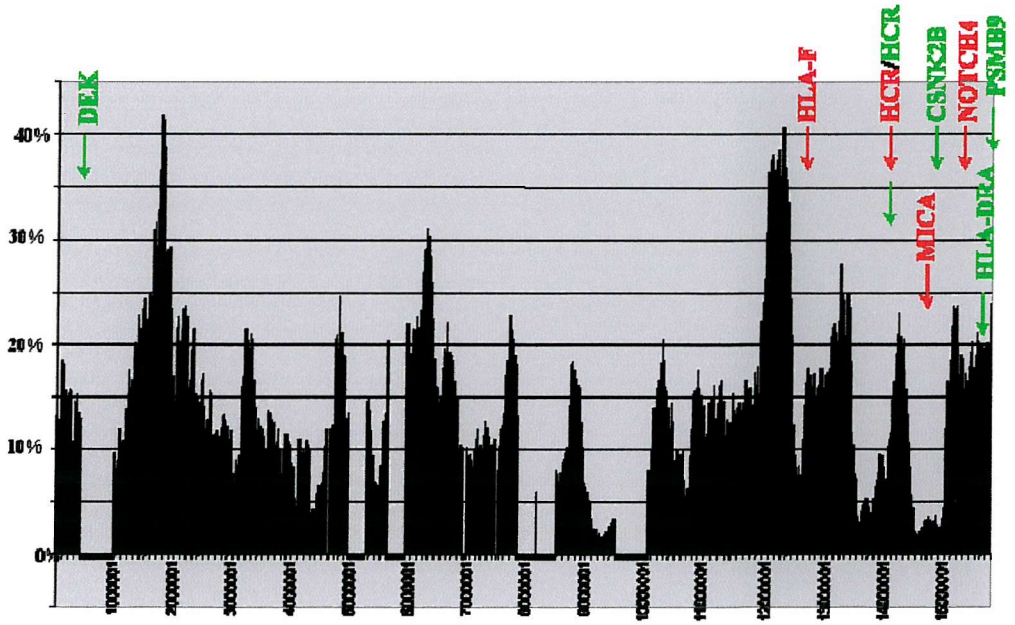


Figure 27 – LINE-1 content in 6p21.31-p22.3



3.3.8 Intragenic repetitive sequence analysis of translocated autosomal genes

There was a significant correlation between total intragenic repeat density and inactivation. The mean total repeat content of autosomal genes silenced by the spread of X-inactivation was 30.03%. This compared to 25.12% in partially active genes and 16.02% in fully active genes (P=0.048, r=0.295). LINE and LTR interspersed repeats accounted for over 90% of this variation in repeat content (Table 6, Appendix 9).

There was a significant correlation between intragenic LINE density and inactivation (P=0.004, r=0.46, Table 6). Long Terminal Repeats (LTRs) were also significantly associated with inactivated genes (P=0.005, r=0.45, Table 6). However, activation status was not associated with SINEs, simple repeats, low complexity DNA, tandem repeats, or G+C content (P>0.1 in each case, Table 6).

I next examined the association of inactivation with the three LINE subclasses: L1s, L2s and CR1s. Individually, L1 repeats were more strongly associated with inactivation than L2 repeats (r=0.45, P=0.003 for intragenic L1s; r=0.35, P=0.021 for intragenic L2s, Table 6) whilst CR1 was not significant (r=0.19, P=0.146). I explored the relative contribution of L1 and L2 repeats to the association of LINEs with inactivation using partial correlations. When the effect of L1s was removed partial correlations were non-significant (P>0.2 for gene only in each case). In contrast, after removal of L2 repeats partial correlations with inactivation remained significant (P=0.03 for gene only), suggesting that inactivation is specifically associated with L1 repeats. When the ten L1 subtypes (L1M1 to L1M4, L1P1 to L1P5, L1HS) were examined individually there was no significant difference in concentration between active and inactive genes (data not shown), although this may simply reflect the relatively small proportion of total sequence comprised by each subtype in the dataset. However, when old (L1M2 to L1M4) and young (L1M1, L1P1 to L1P5, L1HS) L1 subtypes were combined, inactivated genes were found to contain significantly greater concentrations of both old and young subtypes (P=0.018 for intragenic old L1s; P=0.009 for intragenic young L1s, Table 7).

Table 6 - Mean intragenic sequence content of autosomal genes in five unbalanced X;autosome translocations

Sequence type	% of inactive genes, n=19	% of partially active genes, n=5	% of fully active genes, n=9	Correlation, r	P-value
G+C content	47.7	52.2	49.2	0.16	0.167
LINE	11.00	8.09	1.53	0.46	0.004
L1	8.70	5.19	1.23	0.45	0.003
L2	2.21	2.74	0.31	0.35	0.021
CR1	0.09	0.17	0	-0.19	0.146
SINE	11.70	10.86	11.05	-0.07	0.337
Alu	8.88	9.72	8.04	-0.08	0.320
MIR	2.82	1.14	3.01	-0.06	0.382
LTR	4.08	2.80	0	0.45	0.005
Simple repeats	0.28	0.39	0.29	-0.08	0.322
Low complexity DNA	0.75	0.24	0.37	-0.23	0.110
Tandem repeats	1.98	2.69	2.45	0.12	0.257
Di-, tri- and tetranucleotide repeats	0.24	0.05	0.33	0.03	0.436
Total repeat content	**30.03%**	**25.12%**	**16.02%**	**0.29**	**0.0485**

Repeat element density is expressed as the mean percentage of intragenic sequence comprising each repeat type, with Spearman's correlation coefficients and associated probabilities obtained as Monte Carlo estimates from sample permutation distributions.

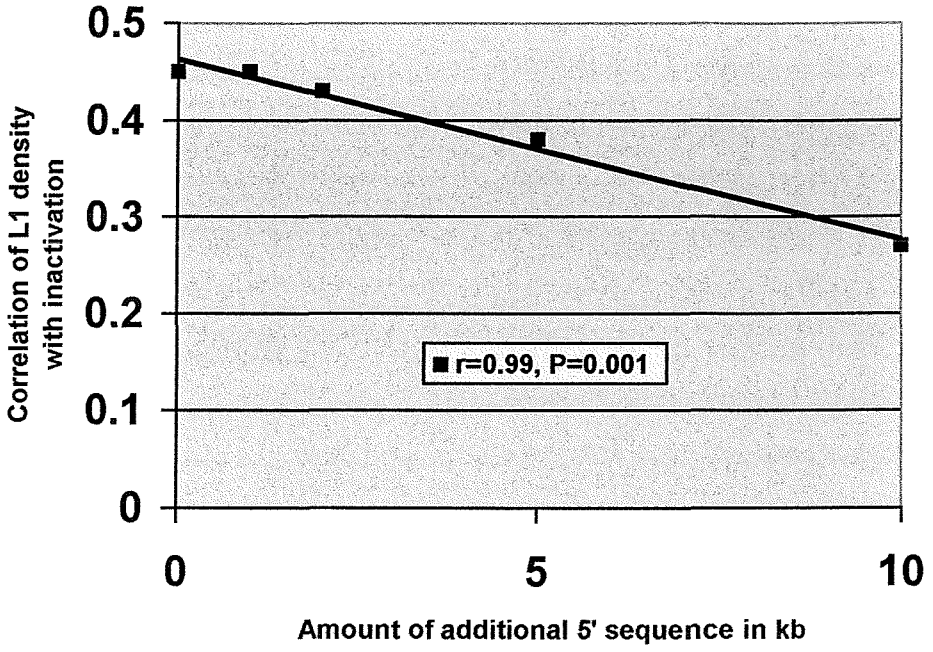
Table 7 - Correlations between inactivation and repeat content with increasing 5' sequence

Dataset	Total LINEs	L1s	L2s	Old LINEs	Young LINEs	LTRs
Gene only	r=0.46, P=0.004	r=0.45, P=0.003	r=0.35, P=0.021	r=0.39, P=0.018	r=0.41, P=0.009	r=0.45, P=0.005
Gene+1kb 5'	r=0.48, P=0.002	r=0.45, P=0.005	r=0.39, P=0.011	r=0.39, P=0.016	r=0.41, P=0.008	r=0.48, P=0.002
Gene+2kb 5`	r=0.46, P=0.004	r=0.43, P=0.008	r=0.32, P=0.037	r=0.38, P=0.020	r=0.41, P=0.007	r=0.49, P=0.002
Gene+5kb 5`	r=0.40, P=0.017	r=0.38, P=0.022	r=0.31, P=0.053	r=0.33, P=0.059	r=0.44, P=0.005	r=0.40, P=0.018
Gene+10kb 5`	r=0.35, P=0.045	r=0.27, P=0.127	r=0.43, P=0.010	r=0.29, P=0.110	r=0.22, P=0.168	r=0.38, P=0.018

Spearman's correlation coefficients with associated probabilities obtained as Monte Carlo estimates from sample permutation distributions. 95% confidence intervals were obtained for every correlation estimate using a bias corrected Bootstrap estimator. The median interval width spanned r±0.27 for gene only, rising to r±0.32 for gene + 10kb 5'.

As elements located in the upstream regions of genes are often involved in transcriptional regulation, we extended our analysis to include additional sequence 1kb, 2kb, 5kb and 10kb 5' of each gene (Table 7). However, the correlation between L1 density and inactivation diminished linearly with additional upstream sequence (linear regression R^2=0.98, P=0.001, Figure 28), suggesting that sequences located 5' of autosomal genes are not associated with the spread of X-inactivation.

Figure 28 - Correlation of L1 density with inactivation diminishes with additional 5' sequence



Section 4 – Discussion

4.1 X-inactivation ratios in normal women

4.1.1 Introduction

Numerous investigations have attempted to define the prevalence of skewed X-inactivation in normal females. Despite this, reports in the literature vary widely, with up to ten-fold differences in the estimates of women with 'non-random' patterns of X-inactivation quoted by different authors (e.g. Vogelstein *et al.*, 1987; Fey *et al.*, 1992). Comparisons between previous studies which might account for these discrepancies are hampered by several factors.

Firstly, definitions of 'skewed X-inactivation' vary widely. Based on the assumption that X-inactivation is a truly random event which takes place in a population of twenty independent cells, Vogelstein *et al.* (1987) defined skewed X-inactivation as ratios ≥75:25, on the basis that this represented a statistically significant departure from the mean (with $P<0.05$). However, other authors have since defined skewed X-inactivation as ratios of 80:20, 85:15, 90:10, or even 95:5.

Secondly, numerous techniques now also exist for the determination of X-inactivation ratios. Many earlier studies have utilised Southern blotting-based assays which rely on restriction-fragment length polymorphisms at either the *PGK*, *HPRT* or DXS255 loci. However, more recent studies have instead used PCR-based techniques, in particular the *AR* assay. While these different techniques generally produce concordant results, discrepancies have been noted in some cases (Gale *et al.*, 1991).

Finally, recent evidence suggests that X-inactivation ratios might vary both with age, and between different tissues (Busque *et al.*, 1996; Gale *et al.*, 1994). These latter findings indicate that meaningful comparisons between populations of females must use both age- and tissue-matched controls.

In an attempt to try and unify these observations, this study has set out to screen a large number of normal females using a single, reliable technique. Blood samples from females of different age groups were studied in order to determine possible changes in the prevalence of skewing with age. Additional samples of blood,

mouthbrush and urine were collected from further females of different ages in order to determine tissue-specific patterns of X-inactivation. Further studies were also carried out in an attempt to determine possible factors which might influence X-inactivation ratios in these apparently normal women.

4.1.2 Variations of skewed X-inactivation with age

The incidence of skewed X-inactivation is controversial, with previous studies reporting that between 4% and 33% of normal females have ratios ≥75:25. Using a more stringent criterion (ratios ≥90:10), this study has found the incidence of severe skewing in females aged ≤25 to be 7%. In addition, characterisation of more elderly females suggests that this incidence increases with age, with 16% of women ≥60 found to have similarly severe skewing in blood cells. Therefore, these results demonstrate that severely skewed X-inactivation, at least in blood cells, is a relatively common phenomenon. This is in general agreement with some previous surveys of X-inactivation patterns in normal females which have also used the *AR* assay (Naumova *et al.*, 1996).

The observation that severe skewing occurs more frequently in elderly females suggests that skewed X-inactivation is a trait that can be acquired with age. Because the process of X-inactivation occurs in early embryogenesis and is then stably maintained in all progeny cells, theoretically the proportion of cells which have inactivated each X chromosome should remain constant from birth to death. However, this assumes that all cells have both similar life spans, and divide at equivalent rates. Therefore, any factor which might alter either a cells' longevity or its rate of division compared to another clone could conceivably influence the relative preponderance of that cell lineage within a tissue. This mechanism may well account for the increase in the frequency of skewed X-inactivation in blood cells of elderly females observed in this study.

Blood cells are one of the most mitotically active tissues of the body, many of the constituent cells dividing rapidly with a short life span. This rapid and continuous turnover means even mild selective pressures would be able to exert an influence during the lifetime of an individual. X-inactivation results in the formation of two populations of cells which differ only in the choice of which X chromosome they are

expressing. Due to the many polymorphisms which are present between chromosomes, this dichotomy immediately provides the potential for one cell lineage to be at a selective disadvantage to the other. Any variations of X-linked genes which result in even a tiny metabolic or growth difference between these two clones might result in the selective growth of cells expressing a particular X chromosome. The net result of this would be the development of skewed X-inactivation ratios with time.

Similar secondary cell selection phenomena are well documented for a variety of X-linked syndromes, such as Wiskott-Aldrich syndrome, X-linked recessive deafness, and X-linked α-thalessaemia (Wengler *et al.*, 1995; Orstavik *et al.*, 1996; Gibbons *et al.*, 1992). In particular, in female carriers of Fragile X, an inverse correlation has been noted between age and the proportion of blood cells which retain an active X chromosome carrying a full mutation (Rousseau *et al.*, 1991). This latter observation is entirely consistent with the notion that certain X-linked genotypes may be at a selective disadvantage and are subsequently out-competed, resulting in progressively skewed X-inactivation patterns with age. However, over the lifetime of an individual, cell selection might not require the presence of specific gene mutations *per se*. Instead, any polymorphism which slightly alters the expression or function of a vital gene might provide sufficient bias on which selection could act. Longitudinal studies of X-inactivation ratios in a number of normal females over an extended time period would be needed to definitively confirm this hypothesis.

One such study was recently reported by Abkowitz *et al.* (1998). Here, longitudinal analysis of G6PD variants in the blood cells of Safari cats showed a slow change towards the preferential expression of one isoenzyme in seven of the eleven females analysed. The rate of change of X-inactivation ratios, as measured by G6PD polymorphism, varied between individuals, taking between one and eight years to show a significant deviation from the mean in different animals. In addition, Abkowitz *et al.* found that while this change occurred in short-lived erythrocytes and granulocytes, the X-inactivation ratio in long-lived T-lymphocytes remained unchanged, even after 13 years in one animal. These results indicate that a selection process favouring cells expressing a particular X chromosome occurs in many normal females, resulting in the generation of skewed X-inactivation with time.

Heritable variations that occur between X chromosomes and on which cell

selection can act could account for familial cases of skewed X-inactivation such as those reported by Naumova *et al.* (1996 and 1998). As discussed by Migeon (1993), this hypothesis provides a simple explanation for familial skewing and, consistent with the linkage data of Naumova *et al.* (1998), causative variations would be expected to map to different X-linked loci in separate pedigrees.

It follows from this that the frequency of skewed X-inactivation seen in young females is probably a better representation of the incidence of truly stochastic skewing, i.e. that which occurs purely by chance. However, it is probable that a certain amount of cell selection occurs during the first few years of life, and so the true incidence of stochastic skewed X-inactivation (ratios ≥90:10) is probably lower than the 7% found in females ≤25 in this study. Studies of X-inactivation patterns in the blood of neonates would confirm this, as determined by Busque *et al.* (1996). In this study the authors found severe skewing (ratios ≥90:10) in 2% of neonates, 5% of females aged 28-32, and 23% of females aged ≥60. These results indicate that the vast majority of extreme skewing seen in the blood of normal adult females is acquired with age.

It is unclear however why the results of Busque *et al.*, and those of other authors using the *AR* assay, differ appreciably from this study. One possibility is the different methods used to quantify the intensity of the alleles produced in the PCR reaction. Busque *et al.* utilised radioisotope labelling of their PCR products, and the gel images were then scanned by densitometry. In contrast, this study has utilised fluorescently labelled PCR products, and the allele intensity determined automatically by gel analysis software. Accurate quantification using densitometry of gel images can be problematic, especially where PCR products are poorly defined or there is a high degree of background signal which must be compensated for. This is especially true where alleles are of weak intensity, as is often the case when using the *AR* assay. Inspection of the raw data of Busque *et al.* shows that less than 10% of their subjects in each age category were scored as having X-inactivation ratios in the 'moderate skewing' interval of 75:25 to 95:5. This suggests that their densitometric method for quantification of alleles may have exaggerated the degree of skewing in many cases, resulting in individuals with moderate skewing being misclassified as having severely skewed X-inactivation. This may well account for the high incidence of severe skewing reported by Busque *et al.* in their elderly age group. Conversely, the semi-

automated fluorescent PCR assay utilised here was shown to be highly reproducible in all individuals, suggesting the results gained were both accurate and reliable.

The inclusion of DMSO and 7' deaza-dGTP in the PCR used to amplify the *AR* repeat also made a significant improvement in assay reliability, and allowed uniform and reproducible amplification of otherwise problematic samples (Mutter and Boynton, 1995).

An additional technical factor may also account for some of the discrepancies in the frequency of skewing reported by different authors. While this study has utilised a double methyl-sensitive restriction digestion of patient DNA using both *Hpa*II and *Cfo*I, other studies have often used only single digestion with *Hpa*II alone. While this should theoretically make no difference, single digestion with *Hpa*II alone can fail in some instances (authors observations). Despite the use of control DNA to check for apparent complete digestion, a failure of *Hpa*II to completely cleave all unmethylated cut-sites in some samples would go unnoticed, resulting in an under-estimate of the incidence of severe skewing. Use of only a single restriction enzyme is potentially unreliable for a number of reasons. Firstly, a reduction in the activity of a specific batch of enzyme would yield false results, although this should be detected if appropriate controls are used. This is far less likely to be a problem if two independent enzymes are used. Secondly, a proportion of the target DNA sequences may retain secondary structure that could render restriction cut-sites inaccessible to a single enzyme digest. However, a double restriction digest will generate DNA fragments of much smaller size, which are far less likely to retain any inhibitory secondary structure. Thirdly, the use of *Hpa*II alone in the *AR* assay means that the ability to distinguish the active and inactive X chromosomes is entirely reliant upon the differential methylation status of only two CpG dinucleotides. Any aberrant methylation, or indeed incomplete cleavage, at these two sites will result in the generation of inaccurate results. Again, the combined use of two enzymes increases the number of methylation sites under investigation to four, and therefore making the assay less liable to inaccuracies due to incomplete methylation of the X_i. In summary, the use of only one methylation-sensitive restriction enzyme for any assay of X-inactivation is much more prone to error than a combined digest, and may tend to underestimate the true degree of skewing in some individuals. This fact may explain the comparatively low incidences

of severe skewing (ratios ≥90:10) of 3.5% and 1.5% reported by Plenge *et al.* (1997) and Lanasa *et al.* (1999), both of whose studies utilised digestion of patient DNA with *Hpa*II alone. In contrast the large study of Naumova *et al.* (1996) in which both *Hpa*II and *Cfo*I were used found an incidence of 9%, similar to that reported here for females ≤25.

Finally, it is possible that sampling differences may have contributed to variations between this study and others, although there was no evidence for this. Populations A and B were recruited from individuals referred to the laboratory as part of pedigrees diagnosed with a variety of autosomal genetic disorders (see Materials and Methods). While none of these conditions are known to influence X-inactivation ratios, there is some very limited evidence suggesting that skewed X-inactivation is more common in carriers of *BRCA1* mutations (Buller *et al.*, 1999). However, no significant differences in the proportion of individuals from each ascertainment group with skewed X-inactivation were observed, indicating that sampling differences did not influence results.

4.1.3 Tissue-specific variations of X-inactivation

Despite previous investigations into the variation of X chromosome inactivation patterns in different tissues (Gale *et al.*, 1994; Azofeifa *et al.*, 1996), the degree of correlation between two tissue lineages within an individual remains unclear. In order to address this issue, this study has determined the X-inactivation ratios in 3 easily accessible tissues, namely blood, buccal epithelia, and urinary cells, taken from 88 informative females of different ages.

Results from these 3 cell types indicate that while the degree of severe skewing was similar in each tissue of young females (aged ≤25 years). In elderly women (aged ≥60) skewed patterns of X-inactivation were significantly more common in blood than in other tissues. There are several possible explanations that might account for this observation. Firstly, this high degree of skewing may be due to somatic cell selection, as detailed earlier. As blood cells have a very high rate of turnover compared to other tissues, there is a greater opportunity for selective pressures to be exerted against particular clones. This mechanism could easily result in skewing being relatively more common in blood than in other, less mitotically active, tissues (Abkowitz *et al.*, 1998).

Alternatively, the different frequencies of skewed X-inactivation seen between tissues may be due to the relative size of the cell pool contributing to each tissue at the time that X-inactivation occurs in the developing embryo. Following random inactivation, the proportionate expression of each X chromosome should follow a normal distribution about a mean of 50:50. The breadth of this distribution though is dependent upon the number of cells present when X-inactivation occurs. If the number of cells is large, then extreme X-inactivation ratios will be uncommon. However, if the number of cells contributing to any one tissue is small, then skewing will be a relatively common event.

By studying developing mouse embryos carrying an X-linked *lacZ* transgene, Tan *et al.* (1993) concluded that X-inactivation in the mouse occurs progressively in different tissues, with some cell lineages apparently completing X-inactivation much earlier than others. This would suggest that the number of cells contributing to each tissue differs significantly, potentially explaining why severe skewing might be more common in some tissues than in others. However, it is possible that the results reported by Tan *et al.* do not reflect differential timing of X-inactivation, but are merely an artifact of tissue variation in the stability of the transgenic *lacZ* mRNA/protein that was used as a reporter in their assay. In addition, no tissue differences in the preponderance of skewing in our younger age group were observed, which would be expected if differential cell numbers contributed to each tissue.

Finally, it is also possible that the increased amount of skewing seen in Population D is a result of sampling bias. While the medical history of each individual in Populations C and D was unknown, patients in these two age categories were recruited from different sources. The majority of females in the younger Population C were pregnant mothers attending obstetric clinic, and could be expected to represent a reasonable population sample. Conversely, individuals in Population D were all referrals made to a pathology department. It is possible that a small proportion of these patients were referred due to haematological malignancies, which in some cases may have influenced their blood X-inactivation status. However, previous studies of both post-chemotherapy and haematologically abnormal women have found no significant differences in blood X-inactivation status when compared with healthy controls (Gale *et al.*, 1991; Fey *et al.*, 1994). Therefore, it seems unlikely that this sampling difference would have had any major influence on results.

Although correlation analysis demonstrated that there was an overall significant association of X-inactivation ratios between each tissue in both young and elderly women, this degree of correlation fell markedly with age. Importantly, in some individuals, extreme skewing was present in one or two tissues, while corresponding X-inactivation ratios in other tissues were random. This was observed more frequently in elderly women than in young females. These results demonstrate that X-inactivation patterns can vary widely between different tissues within normal females. In addition, the consistent observation of reduced levels of correlation of X-inactivation between different tissues in elderly women further supports the hypothesis that secondary cell selection processes are the major contributing factor towards the development of skewing in clinically normal females.

Regression analysis of each tissue distribution was performed, from which 95% prediction intervals were generated (shown in Figures 8, 9 and 10). Given that the X-inactivation ratio in a certain tissue of an individual is known, these allow the corresponding range of X-inactivation ratios in other tissues to be predicted with 95% certainty. For example, if the X-inactivation ratio in the blood of a young woman has been determined to be 90:10, then the corresponding ratio in her buccal epithelia will lie within the range 45:55 to 97:3. The fact that this interval is so broad illustrates that extrapolations of X-inactivation ratios from one cell type to another within an individual should be treated with caution.

4.1.4 Offspring ratios in women with skewed X-inactivation

The data of Pegoraro *et al.* (1997) suggests that deleterious X-linked mutations that cause severely skewed X-inactivation may also cause embryonic lethality in hemizygous male offspring who inherit the abnormal chromosome. If such mutations are a common cause of severe skewing, then this predicts that a population of women with skewed X-inactivation will give birth to an excess of daughters to sons. The study of offspring ratios in women with severely skewed X-inactivation therefore provides a potential means of determining mechanisms that lead to skewed X-inactivation.

Unfortunately, pedigree data was only available for 5 of the women identified from Population A with blood X-inactivation ratios ≥90:10, precluding any meaningful

analyses. However, these 5 women produced a total of 9 daughters and 4 sons, very close to the 2:1 ratio predicted by the presence of a male-lethal trait on one of the maternal X chromosomes. Therefore, although far from conclusive, these data are at least consistent with the hypothesis that women who present with extremely skewed X-inactivation ratios may be heterozygous carriers of deleterious X-linked mutations, and that this may manifest as a lethal trait in half of their male pregnancies. However, the study of a much larger population of women presenting with skewed X-chromosome inactivation patterns is needed to confirm this.

Two recent studies (Lanasa *et al.*, 1999; Sangha *et al.*, 1999) have investigated the possible association of skewed X-inactivation with recurrent spontaneous abortion (RSA). Both found an excess of severe skewing in populations of women who suffer from RSA when compared to normal controls, further supporting the idea that X-linked lethal traits may be relatively common, and a significant cause of pregnancy loss.

Due to the small number of mothers identified with skewed X-inactivation, a further attempt was made to test the hypothesis that women with skewed X-inactivation are carriers of deleterious X-linked variations. In order to do this, an additional cohort of 17 women was studied, named Population E. Each of these females was selected on the basis of having produced at least twice as many daughters as sons. Due to the large amount of effort involved in recruiting sufficient numbers of women with skewed X-inactivation, it was hoped that studying this population would provide a more effective solution than further screening for highly skewed females. If the hypothesis was correct, then this population should have an increased prevalence of women with highly skewed X-inactivation patterns compared to controls.

Results showed that only one of the 16 informative females studied (6%) in Population E had X-inactivation ratios ≥90:10 compared to 7-16% of controls in Populations A-D, indicating that the hypothesis is incorrect. However, a number of factors meant that firm conclusions are difficult to draw from this comparison. Most significantly, as a result of the small size of Population E, only very large enrichments in the frequency of skewed X-inactivation could be detected by comparison with the controls. There are numerous reasons why a woman might conceive more daughters than sons, including random probability and paternal contribution, and these undoubtedly account for a significant proportion of such families. As no attempts

were made to exclude individuals from Population E with a known aetiology for their skewed offspring ratio, the power of this already small survey was likely further reduced. Similarly, the offspring ratios of many individuals in the 'control' Populations A-D were unknown, again diluting the power of this comparison to detect any differences. An ideal control population should consist of age-matched women with the same total number of children of evenly mixed sex. These shortcomings make it unlikely that this survey would detect any real differences in X-inactivation ratios that may exist between test and control populations.

4.1.5 Heritability of X-inactivation ratios

In an attempt to determine if X-inactivation ratios can be inherited as a genetic trait, patterns of blood X-inactivation were determined in 29 mother-daughter pairs. Of these, 17 were informative for both the degree and direction (parental origin) of skewing. Analysis of the correlation between the maternal and daughters' X-inactivation ratios showed no significant association ($r=0.31$, $P=0.23$), indicating that, in most cases, mother-daughter X-inactivation ratios were unrelated. However, the small size of the sample population again means that firm conclusions on this issue cannot be drawn.

Results of two previous familial studies of X-inactivation have both indicated the existence of a heritable component. Orstavik *et al.* (1999) examined blood X-inactivation ratios in 171 elderly female twin pairs. They found a significantly higher correlation between monozygous ($r=0.57$) than dizygous ($r=0.22$) twins, providing strong evidence that a significant proportion of skewed X-inactivation is attributable to genetic factors. Naumova *et al.* (1998) similarly studied X-inactivation ratios in 38 multiplex families, and noted a significant sister-sister correlation in both the degree and direction of skewing. However, as was found here, no such correlation existed between mother-daughter pairs in their study population. Naumova *et al.* proposed that these results could be explained if X-inactivation ratios in humans were influenced primarily, by an X-linked locus equivalent to the murine *Xce*. They postulated that the inheritance of '*XCE*' alleles of different 'strength' within families would lead to aggregations of skewed X-inactivation, with strong sister-sister correlation but a lack of mother-daughter correlation. While potentially explaining the observations made by

Naumova *et al.*, attempts to map this putative '*XCE*' by linkage showed no single region of concordance in one family, and linkage to different X-linked loci in four other pedigrees. These results argue against the existence of a single *Xce*-like control locus in humans, indicating instead that any inheritance of X-inactivation is influenced by different loci in separate pedigrees.

As discussed earlier, X-linked variations upon which secondary cell selection can act provide a simple alternative explanation that could account for inherited X-inactivation ratios, without the need to invoke the existence of multiple X-linked loci with a primary influence on the degree of skewing. Interaction between maternally and paternally derived variations could result in a correlation of X-inactivation ratios between siblings, but not necessarily between mother and daughter. In addition, monozygous twin pairs would show a closer correlation of X-inactivation ratios than would dizygous twins (Orstavik *et al.*, 1999). As observed by Naumova *et al.*, these variations would also map to different regions of the X chromosome in different families. Finally, the longitudinal study of Abkowitz *et al.* (1998), and the data presented here, are inconsistent with the hypothesis that skewed X-inactivation is a primary trait. Instead, they strongly suggest that skewed X-inactivation is acquired over time as a result of secondary selection processes. Taken together, these data indicate that X-inactivation ratios themselves may be heritable, but only as a secondary consequence of the inheritance of X-linked variations upon which cell selection can act.

4.1.6 Summary

The incidence of skewed X-inactivation in normal women is controversial, with up to 10-fold differences reported by different authors. In order to clarify this issue, I have conducted the most comprehensive survey of X-inactivation patterns to date in 270 informative females from different age groups, using the Androgen Receptor gene PCR assay. Results using DNA extracted from blood samples show that the incidence of severe skewing (defined here as ratios ≥90:10) is relatively common and increases with age ($P<0.05$), occurring in 7% of normal women under 25 years of age, and 16% of normal women over 60.

In order to study tissue specific patterns of X-inactivation, samples of both

buccal and urinary epithelia were also obtained from 88 of the females studied. Although there was a significant association of the X-inactivation ratios between each tissue in most individuals, wide variations were apparent in some cases, making accurate extrapolations between tissues impossible. The degree of correlation between each tissue also fell markedly with age.

Attempts were also made to determine the potential mechanisms that might cause skewed X-inactivation in normal women. However, while the study of offspring ratios in women with severely skewed X-inactivation suggested a possible excess of daughters born to these women, results of this study, and of females with a sex-biased offspring ratio, were inconclusive due to small sample size.

Examination of the correlation of X-inactivation between mother-daughter pairs found no significant association. However, small sample size again meant that firm conclusions on the heritability of X-inactivation could not be drawn.

Overall, the data presented here are consistent with the hypothesis that the major factors in the aetiology of skewed X-inactivation are secondary selection processes.

4.2 X-inactivation ratios in Silver-Russell syndrome

Maternal uniparental disomy for chromosome 7 is associated with approximately 7% of SRS cases (Kotzot *et al.*, 1995; Eggermann *et al.*, 1997; Preece *et al.*, 1997), leading to the suggestion that one or more imprinted genes on chromosome 7 are responsible for the disease. However, as it is likely that most cases of matUPD7 arise by trisomy 7 rescue (Mergenthaler *et al.*, 2000b), it is unclear whether the SRS phenotype is due to the presence of matUPD7 alone.

Based on previous reports demonstrating an increased frequency of skewed X-inactivation in cases of trisomy rescue (Lau *et al.*, 1997; Penherrera *et al.*, 2000), this study has set out to determine the possible role of trisomy 7 in SRS by analysing the frequency of skewed X-inactivation in a cohort of 29 non-UPD7 SRS patients. Consistent with the hypothesis, results showed an increased frequency of extremely skewed X-inactivation in SRS (17.2%) compared to age-matched controls (6.6%), although this difference does not reach statistical significance (P=0.079). However,

this is probably a result of the relatively small number of SRS cases analysed, and the low P-value is strongly suggestive that a significant effect would be observed had a larger cohort been studied. More specifically however, the frequency of completely skewed X-inactivation in the SRS population is some ten-fold higher than that found in controls (P=0.014), consistent with the possibility that in some of the cases analysed, trisomy rescue has occurred. Complete skewing in the diploid tissues of these individuals suggests that they are derived from a single, or very small number of, progenitor cells in which trisomy rescue occurred in the developing embryo, and predicts the presence of undetected trisomic cell lines, either confined to the placenta or within other somatic tissues (Lau *et al.*, 1997; Penherrera *et al.*, 2000). These data are therefore consistent with the hypothesis that mosaicism for trisomy 7 exists in some cases of SRS, and suggests that the occurrence of matUPD7 seen in a proportion of cases is a consequence of trisomy rescue. However, these results do not exclude a role for imprinting in SRS, as indicated by a proband with segmental matUPD7 confined to 7q31-qter which is unlikely to be associated with trisomy 7 (Hannula *et al.*, 2001), and by several other SRS patients with abnormalities of chromosome 7 (D. Monk, personal communication).

The proportion of SRS cases in which both extreme and complete skewing was observed is similar to that predicted from the 7% incidence of matUPD7 in SRS, as follows. As one third of cases of trisomy rescue will result in UPD, it can therefore be estimated that trisomy 7 occurs in some 21% of SRS patients. This study has examined only those individuals with biparental inheritance of chromosome 7, representing 14% of total cases, and previous studies of trisomy rescue have found that approximately two-thirds of such cases where the trisomy is of meiotic origin show completely skewed X-inactivation (Lau *et al.*, 1997; Penherrera *et al.*, 2000). Approximately half of SRS cases with matUPD7 have arisen from a trisomy of meiotic origin (Kotzot *et al.*, 1995; Eggermann *et al.*, 1997; Preece *et al.*, 1997). Thus, some 5% of non-UPD SRS females should exhibit skewing as a result of trisomy rescue, in addition to the 'background' incidence of 7% found in age-matched controls, totalling 12%, statistically similar to that observed here (17%). However, there did not appear to be an association of increased maternal age with SRS, which would be expected in those cases where the trisomy was of meiotic origin (Risch *et al.*, 1986). In particular, the mean maternal age for four of the five individuals with extreme skewing for whom

data was available was 28 (range 21-34), although the small sample size made firm conclusions impossible. In addition, skewed X-inactivation was not observed in three cases of maternal heterodisomy 7 (Lau *et al.*, 1997). As these most likely arose by trisomy rescue complete skewing might be expected in these cases, but once again firm conclusions cannot be drawn from such small numbers.

The possible association of trisomy 7 with SRS is similar to that seen in trisomy 16. All cases of trisomy 16 are of maternal meiotic origin (Hassold *et al.*, 1995), and mosaicism for trisomy 16, often confined to the placenta, is associated with matUPD16 in many instances. However, in some cases there is poor correlation between matUPD16 and intrauterine growth retardation (IUGR) or abnormal pregnancy outcome (Penaherrera *et al.*, 2000). Instead, it is likely that the IUGR and foetal abnormalities associated with matUPD16 are partly due to the effects of high levels of trisomy 16 in the placenta, and their possible persistence in some somatic tissues (Robinson *et al.*, 1997).

Although the observed matUPD and structural abnormalities of chromosome 7 in SRS strongly suggest the specific involvement of chromosome 7, the finding of skewed X-inactivation in SRS is also compatible with alternative explanations other than trisomy 7 rescue. Many X-linked syndromes are associated with skewed X-inactivation, and X-linked inheritance has been suggested for some familial SRS cases (Duncan *et al.*, 1990). While there is currently no direct evidence to support the involvement of the X chromosome in SRS, as the syndrome is undoubtedly genetically heterogeneous our observations may also indicate an X-linked cause in a proportion of cases. Alternatively, the increased frequency of skewing in SRS could indicate the involvement of trisomy rescue for other autosomes, besides chromosome 7. A previous case of confined placental mosaicism for trisomy 16 showed asymmetric IUGR with a relatively large head, similar to that seen in many SRS cases (Kalousek *et al.*, 1993), and trisomy 18 mosaicism has also been reported in a child with features of SRS (Chauvel *et al.*, 1975). However, UPD for many autosomes, particularly those that are frequently observed as aneuploids, has previously been excluded from 70 SRS probands (Eggermann *et al.*, 1997; Preece *et al.*, 1997). Thus, while trisomy mosaicism for chromosomes other than 7 may be responsible for a small number of cases, it seems unlikely to play any significant role in the disease.

Previous studies have failed to detect the presence of trisomy 7 cells in the

blood or fibroblasts of SRS patients (Monk *et al.*, 2001), although the techniques used in some cases have been limited by low sensitivity (Eggermann *et al.*, 1997; Preece *et al.*, 1997). In addition, high-level mosaicism or complete trisomy 7 is a lethal condition, with different phenotypic features from those seen in SRS (Fryns, 1990c). Therefore, SRS might be characterised by distinct patterns of mosaicism for trisomy 7 cells, confined to certain tissues. Somatic mosaicism for trisomic cells might account for the asymmetrical growth and patchy skin pigmentation that is observed in some individuals with SRS, and mosaicism for a ring chromosome 7 has been reported in one case of SRS (Miyoshi *et al.*, 1999). While such asymmetry might be predicted in individuals with mosaicism, only one of the three SRS individuals in whom we found completely skewed X-inactivation showed asymmetrical growth. In contrast to the formation of matUPD7, patUPD7, which is compatible with normal growth, probably results from monosomy rescue (Hoglund *et al.*, 1994; Pan *et al.*, 1998), and would not be associated with trisomy.

In summary, the increased frequency of completely skewed X-inactivation found in SRS patients with biparental inheritance of chromosome 7 is consistent with the possible presence of trisomy 7 confined to certain tissues in some SRS patients. Whether this contributes to the phenotype of SRS requires further experimentation.

4.3 The spreading of X-inactivation in X;autosome translocations

4.3.1 Introduction

Since the ability of X-inactivation to spread variably into autosomal DNA was first recognised (Russell, 1963), numerous other studies of X;autosome rearrangements have confirmed and extended these early observations (Reviewed in Mattei *et al.*, 1982; Camargo and Cervenka, 1984). However, in all but a few cases evidence for the extent of spread of X-inactivation into autosomal DNA has been based upon either replication timing studies or the associated clinical phenotype in carrier individuals. The former approach is problematic, as direct evidence to support the presumption for an absolute correlation between the extent of spread of late-replication and

transcriptional silencing in X;autosome translocation carriers is weak. While in many cases there is an apparently good correlation between the chromosome imbalance, spreading of late-replication, and clinical phenotype, sporadic reports of individuals with unexpectedly mild phenotypes in which there is no detectable spread of late-replication suggests that in some cases this relationship does not hold true (Keitges and Palmer, 1986; Bettio *et al.*, 1994; Garcia-Heras *et al.*, 1997; Bacino *et al.*, 1999; Abuelo *et al.*, 2000).

While observations in four X;autosome rearrangements have noted the inactivation of single autosomal genes (Couturier *et al.*, 1979; Mohandas *et al.*, 1981; Mohandas *et al.*, 1982; Taysi *et al.*, 1982), only one other previous study has directly measured the spread of X-inactivation through autosomal DNA by gene expression analysis. White *et al.* (1998) found that the spread of inactivation in an X;4 translocation occurred in a discontinuous fashion, with active genes interspersed among inactive regions of chromatin. This mirrors similar observations of discontinuous spreading of late-replication (Mohandas *et al.*, 1981; Keitges and Palmer, 1986), suggesting that some autosomal chromatin lacks features important in the spread and/or maintenance of X-inactivation.

Recently Mary Lyon proposed LINE repeats as candidates for these *cis*-acting elements which promote the spread of X-inactivation along a chromosome (Lyon, 1998), a hypothesis supported by the distribution of LINEs on the X chromosome (Bailey *et al.*, 2000). As the density and distribution of LINEs is a constant property of each chromosome, Lyon's repeat hypothesis predicts that in X;autosome rearrangements, the spread of X-inactivation into autosomal chromatin should be an intrinsic property dependent upon LINE content.

This study has set out to perform detailed studies of the spreading of X-inactivation in unbalanced X;autosome translocations. Five such cases were studied, each associated with varying severity of phenotype. In each case the spread of X-inactivation into the translocated autosomal material was directly determined on a gene-by-gene basis using allele-specific RT-PCR and CpG island methylation analysis. In an attempt to identify cytogenetic features of the inactive X that might correlate with this spread of gene silencing, a combination of immunofluorescence and FISH was performed to localise the distribution of acetylated and methylated histones, late-replicating chromatin, and *XIST* RNA on each der(X). Finally, in order to test Lyon's

repeat hypothesis, a sequence analysis using data from The Human Genome Project was performed to identify repetitive DNA elements, in particular LINEs, that might correlate with the spread of X-inactivation observed in each case.

4.3.2 Analysis of five unbalanced X;autosome translocations

I have directly studied the spreading of X-inactivation in five X;autosome translocations and in each case demonstrated the long-range silencing of autosomal genes located up to 45Mb from the translocation breakpoint. Thus, whatever factors are responsible for the spread of X-inactivation are not unique to the X chromosome. However, in each case this spreading has occurred in either an incomplete or discontinuous fashion, suggesting that autosomal chromatin does not either transmit or maintain the inactivation signal as efficiently as the X chromosome. Furthermore, these results show that the spreading of X-inactivation through autosomal chromatin can occur in the absence of cytogenetic features normally associated with the inactive X, such as late-replication, histone hypoacetylation and coating with *XIST* RNA.

Many previous studies have utilised techniques to determine the spreading of late-replication into autosomal chromatin in X;autosome rearrangements, and inferred a direct relationship between delayed replication timing and genetic silencing in such cases (Keitges and Palmer, 1986; Schanz and Steinbach, 1989; Surralles and Natarajan, 1998). Indeed, there are many cases in which the attenuated phenotype of individuals trisomic for a segment of autosome, translocated onto an inactive X, appears to correlate well with the spreading of late-replication into it (Mohandas *et al.*, 1982; Disteche *et al.*, 1984; Keitges and Palmer, 1986; Wiliams and Dear, 1987; Canun *et al.*, 1998). Such cases imply that the inactivated region of autosome can be defined by the extent of spread of late-replication. However, my observations indicate that, at least in some instances, silencing of autosomal genes by X-inactivation can occur without any apparent global delay in the replication timing of the surrounding chromatin. In both Case 1 (AH) and Case 5 (BO0566), there was no evidence to suggest that the translocated autosomal segment was delayed in its replication compared to the normal homologues. Instead, the late-replicating domain of both these der(X) chromosomes appeared to define the boundary between the autosomal and X chromatin. A review of published cases of unbalanced X;autosome

translocation suggests that this phenomenon is not restricted to these two patients. A number of individuals have been previously described where no spreading of late replication into the attached autosome was observed, but whose unusually mild clinical phenotype strongly suggests inactivation of translocated autosomal genes (Keitges and Palmer, 1986; Bettio *et al.*, 1994; Garcia-Heras *et al.*, 1997; Bacino *et al.*, 1999; Abuelo *et al.*, 2000). Thus, these data show that the use of replication timing studies to characterise the extent of spread of X-inactivation in X;autosome rearrangements can be misleading.

In contrast, while there was a complete lack of spread of late-replication in Case 1 (AH), observations of depletion of histone acetylation and H3 lysine 4 dimethylation in this individual show a good correlation with the pattern of gene silencing. Similarly in Case 2 (SP), the almost complete spread of gene silencing is accurately reflected by the pattern of histone acetylation and H3 lysine 4 dimethylation, but less so by the relatively limited spread of late-replication. Therefore, while observations in Case 3 (SR) do show an apparent association between the spread of late-replication and gene silencing, these data demonstrate that the distribution of histone modifications which distinguish the inactive X, such as H3AcK14, H4AcK8 and $H3Me_2K4$, are superior to late-replication as cytogenetic measures of the spread of X-inactivation. Additionally, they also show that these histone modifications are distinct from and independent of late-replication (Keohane *et al.*, 1996). However, where the spread of gene silencing occurs discontinuously, as observed in the two X;6 translocations studied, there was no corresponding distribution of depletion of histone acetylation and H3 lysine dimethylation. This may simply reflect the low resolution of immunofluorescence, and histone states at the gene level may differ from that of the surrounding chromatin, as is observed for X-linked genes which escape X-inactivation (Gilbert and Sharp, 1999; Boggs *et al.*, 2001).

While late-replication is a poor cytogenetic correlate of the spread of X-inactivation, it is interesting to note that in every case studied, autosomal genes which were located within cytogenetically late-replicating regions were inactive. Therefore this late-replicating chromatin may represent domains in which the spread of X-inactivation is maintained in a more stable fashion. All late-replicating autosomal regions were both hypoacetylated and depleted of H3 dimethylated at lysine 4, and furthermore the location of autosomal genes in SP, SR and AL0044 which showed

partial inactivation corresponded with observations of mosaicism in the spread of late-replication.

These data show that the CpG islands of autosomal genes silenced by the spread of X-inactivation become highly methylated, as occurs on the inactive X (Tribioli *et al.*, 1992). CpG island methylation probably represents an important component of the spread and/or maintenance of inactivation through autosomal DNA, as in both AL0044 and BO0566 methylation at *MICA* and *HLA-F* represents the only difference detected between the inactive translocated copy of these genes and their active normal homologues. However, sequence analysis did not detect CpG islands adjacent to or within several other inactivated autosomal genes, suggesting that possession of a CpG island is not a prerequisite for silencing by the spread of X-inactivation.

Perhaps the most surprising observation to emerge from this study is that the spread of X-inactivation in each of the five cases studied showed different characteristics, reflecting the complexity of the X-inactivation cascade. In Case 1 (AH) there was an apparently continuous spread of gene silencing covering the majority of the translocated 10q segment. This spread of inactivation correlated with the pattern of depletion of histone acetylation and H3 lysine 4 dimethylation but not with the spread of late-replication. In Case 2 (SP), there was a continuous spread of gene silencing over almost the entire translocated segment of 11p, accompanied by a partial but variable spread of late-replication and an almost complete spread of depletion of histone acetylation and H3 lysine 4 dimethylation. In Case 3 (SR) in the great majority of cells studied there was a partial and apparently continuous spread of gene silencing which correlated well with both the spread of late replication and histone modifications. However, there was also a discontinuous spread of late-replication in a small proportion of cells. In contrast, both of the X;6 translocations studied (Cases 4 and 5, AL0044 and BO0566) showed a discontinuous spread of gene silencing, but cytogenetic observations were discordant between the two. In AL0044, there was a discontinuous and variable spread of late-replication and depletion of histone acetylation and H3 lysine 4 dimethylation which were confined to the distal portion of the translocated 6p segment. The apparent absence or minimal spread of these cytogenetic features in many cells may explain why they were not reported in a previous study of this same case (Keohane *et al.*, 1999). Observations of *XIST* RNA

distribution in this case showed little or no spread into translocated 6p segment (Keohane *et al.*, 1999). In BO0566 no spread of late-replication or histone modifications were observed.

Despite these variations, in every case there was a good correlation between the pattern of gene silencing and the attenuation of clinical phenotype associated with each partial autosomal trisomy. Despite being trisomic for the region 10q23.3-qter, Case 1 (AH) presented only with secondary amenorrhea, as did her aunt with the same unbalanced rearrangement, a finding consistent with rearrangements of Xq (Mattei *et al.*, 1982). She lacked any clinical features normally seen in patients with distal 10q trisomy, which include growth and mental retardation, facial and skeletal anomalies, and a variety of internal malformations (Fryns, 1990a). Consistent with this phenotype, results indicated an apparently continuous spread of inactivation from the adjacent X chromatin, silencing multiple genes over some 25Mb of DNA in the translocated chromosome 10. Surprisingly, the most distal gene tested, *PRDX3*, was shown to remain active, indicating that the spread of inactivation into the 10q material was incomplete. Although the possibility that this gene was initially silenced and has later reactivated cannot be excluded (Schanz and Steinbach, 1989), this indicates that over-expression of *PRDX3*, an antioxidant protein that localises to the mitochondrion (Tsuji *et al.*, 1995), does not contribute significantly to the phenotype associated with trisomy 10q.

Case 2 (SP) showed mild features of Beckwith-Wiedemann syndrome, consistent with the paternal origin of her additional copy of 11p, a region known to be imprinted. However, trisomy for the segment 11p12-pter is normally associated with much more severe abnormalities than present in SP, and her phenotype is instead similar to that seen in cases with smaller paternal duplications of 11p15-pter (Fryns and Chrzanowska, 1990). This correlates well with the pattern of gene expression observed on the der(X;11), with an apparently continuous spread of X-inactivation silencing all but the distal tip of the translocated 11p segment. Similarly, trisomy for the region 7q22-qter is normally associated with profound mental retardation and multiple congenital abnormalities, and most cases are either stillborn or die shortly after birth (Schinzel, 1990). The phenotype seen in Case 3 (SR) is considerably milder than this, and instead her abnormalities are similar to those expected with trisomy for a more distal region of 7q, consistent with the partial spread of X-inactivation observed

in this case. Finally, both Cases 4 (AL0044) and 5 (BO0566) also show milder phenotypes than would normally be predicted simply from their chromosome abnormalities. Trisomy for 6p21-pter normally causes multiple congenital anomalies, with severe growth and mental retardation, and internal abnormalities are such that most cases die during the neonatal period (Fryns, 1990b). However, AL0044 presented only with mild learning difficulties and developmental delay as a 7 year old child, consistent with inactivation of the majority of genes tested within the translocated segment of 6p. In contrast, BO0566 displays much more severe features than AL0044. This may be partly attributed to her larger translocated 6p segment, but seems more likely to be a result of the less efficient spread of inactivation observed, by comparison with that in AL0044.

This concordance between the spread of gene silencing and the associated clinical phenotype suggests that our *in vitro* observations in lymphoblastoid cells are a good reflection of that *in vivo*, although we cannot exclude the possibility that the pattern of inactivation observed is a result of selection removing those cells in which dosage-sensitive autosomal genes remained active. Additionally, it is possible that in each of the cases studied the spread of gene inactivation, late-replication and/or histone modifications we observed is a result of inefficient maintenance of the X-inactivation signal in autosomal chromatin. Immediately after the onset of X-inactivation in the developing embryo, the spread of inactivation through each translocated autosome may have been more complete but subsequently receded due to a lack of appropriate maintenance. A failure of maintenance could be characterised by the progressive or differential loss of features of X-inactivation such as coating with *XIST* RNA, late-replication and histone hypoacetylation prior to the reactivation of autosomal genes, which could account for our observations of the distribution of inactive autosomal genes in relation to cytogenetic features of X-inactivation. Although I cannot discount this possibility, the study of gene expression, CpG island methylation and late-replication in AH, SP and SR using both EBV-transformed lymphoblasts and peripheral blood did not detect any differences between the two cell types. This suggests that the spread of X-inactivation at the DNA, RNA and chromatin level is maintained in a relatively stable fashion in both systems and is not influenced by EBV-transformation.

Previous observations of the distribution of *XIST* RNA in interphase

lymphoblast nuclei of AL0044 showed it to be localised specifically to the inactive X with little or no spread into the adjacent 6p chromatin (Keohane *et al.*, 1999). These observations are consistent with those made in murine X;autosome rearrangements (Duthie *et al.*, 1999), and suggest that *XIST/Xist* RNA has a reduced affinity for autosomal chromatin. Our observations in AL0044 clearly demonstrate that silencing of autosomal genes by a spread of X-inactivation can occur in the absence of coating by *XIST* RNA. Unfortunately it was not possible to study the distribution of *XIST* RNA on the der(X) in other cases as FISH studies showed an aberrant distribution of *XIST* transcripts in the cell lines available. In lymphoblast nuclei of both SP and SR, where *XIST* expression was detectable it showed a diffuse distribution that was poorly localised compared to controls, while in many cells *XIST* was apparently not expressed. Given that *XIST* transcripts were apparently normally expressed and localised in EBV-transformed lymphoblasts of AL0044, it is unclear why they showed an aberrant distribution in similarly transformed cell lines of both SP and SR. However, this sporadic failure of expression and localisation of *XIST* transcripts in some cell lines might be an artefactual consequence of EBV-transformation, similar to that seen in transgenic lines (Clemson *et al.*, 1998). Although it was not possible to perform *XIST* RNA FISH studies in fresh samples, as we were unable to detect any appreciable differences in the spread of X-inactivation in SP or SR between fresh and EBV-transformed samples it seems likely that this aberrant *XIST* expression did not have any appreciable influence on the spread of inactivation observed, nor represents the situation *in vivo*.

There was no apparent relationship between the spread of inactivation and the location of translocation breakpoints or G-bands, as has been suggested (Camargo and Cervenka, 1984; Duthie *et al.*, 1999). Inactive genes were contained within both G-light and G-dark bands at similar frequency. Neither did the extent of spread of inactivation appear to be related to absolute distance from the *XIC*, although in AH, SP and SR there were clear 'gradient effects', with inactivation diminishing in effect with increasing distance from X chromatin. It has also been suggested that on the human X the separation of the short arm from the *XIC* by the centomere may be responsible for the high density of genes escaping X-inactivation in Xp (Carrel *et al.*, 1999; Disteche, 1999). However, comparison of AL0044 and BO0566 suggests that centromeric heterochromatin does not represent a barrier to the spread of X-

inactivation.

In three of the five cases studied we observed that late-replication and/or histone hypoacetylation and H3 lysine 4 dimethylation sharply define the boundary between the X and autosomal chromatin, strongly suggesting that autosomal DNA has distinct properties from the X and is resistant to the X-inactivation signal. The discontinuous spread of gene silencing observed in AL0044 and BO0566 suggests that certain regions of chromosome 6p in particular lack features important for the spread and/or maintenance of X-inactivation. However, the differences in the spread of late-replication, histone modification and the discordant inactivation of *HCR* between the two cases indicates that sequence-specific factors are not the sole determinants of X-inactivation, and that the relative position of translocation breakpoints also influences the extent of its spread in X;autosome rearrangements.

It remains unclear though what dictates whether the spread of X-inactivation through autosomal DNA is accompanied by global changes to the surrounding chromatin, such as histone modifications and delayed replication timing. The absence of these features on much of the translocated 6p segment in both AL0044 and BO0566 where gene silencing occurred discontinuously, suggests that these hallmarks of the inactive X may be limited to chromatin regions that are globally inactive. However, as evidenced by Case 1 (AH), a delay in replication timing is clearly not a direct consequence of a continuous spread of either gene silencing or histone hypoacetylation. Instead, the association of these features with the spread of X-inactivation may be related to some intrinsic property of chromatin that is currently unrecognised.

The inactive X chromosome is distinguished by multiple mechanisms of transcriptional control. Methylation, hypoacetylation, conformational and replication changes are all undoubtedly important in the establishment or maintenance of X-inactivation. Removal of these control mechanisms can destabilise the inactive state (Haaf, 1995; Migeon *et al.*, 1986), leading to the reactivation of previously silenced genes (Hansen *et al.*, 1996). The limited spreading of features such as late-replication and histone hypoacetylation suggest that at least some autosomal chromatin is resistant to these characteristic changes that distinguish the inactive X chromosome from transcriptionally active chromatin, and may contribute to the relatively limited ability of X-inactivation to spread and be maintained *in cis* through autosomal DNA.

4.3.3 Sequence analysis

The mechanism by which X-inactivation spreads along a chromosome is poorly understood. Lyon (1998) recently proposed that Long Interspersed Nuclear Elements (LINEs) might function to facilitate this spread, a hypothesis supported by the distribution of LINEs on the X chromosome (Bailey *et al.*, 2000). Lyon's repeat hypothesis predicts that autosomal genes that are contained within LINE-rich regions should be silenced by the spread of X-inactivation, while LINE-poor autosomal regions should escape the X-inactivation signal. This was tested in two ways. Initially a study of the distribution of L1s at the cytogenetic band level in three of the translocated regions of autosome studied was performed. In apparent disagreement with Lyon's hypothesis, when genomic sequence was analysed in windows of 200kb, there was no obvious relationship between L1 density and the spread of inactivation. However, because of the subjective nature and incomplete coverage of the analysis it was difficult to draw firm conclusions based on these data. Therefore a second analysis which focused specifically on repetitive sequences within and immediately upstream of each of the 33 autosomal genes examined by RT-PCR and/or methylation analysis was performed. Consistent with Lyon's hypothesis (1998) and the data of Bailey *et al.* (2000), results of this analysis established that autosomal genes susceptible to the spread of X-inactivation have a different repetitive sequence content from those that resist this signal. Inactivated autosomal genes contained significantly higher densities of both LINE-1 and LTR interspersed repeats, suggesting that the spread of X-inactivation is sequence-specific. These results support Lyon's repeat hypothesis (Lyon, 1998) and, in the context of their distribution on the X chromosome (Bailey *et al.*, 2000), provide compelling evidence that LINE-1s are 'booster elements', first proposed by Gartler and Riggs (1983), which mediate the spread of X-inactivation *in cis*.

LINEs clearly satisfy the requirements of these *cis*-acting elements, being distributed throughout the genome but enriched specifically on the X (Korenberg and Rykowski, 1988), where their organization is suggestive of a role in the X-inactivation process (Bailey *et al.*, 2000). They also occur throughout the mammalian lineage. Although a recent comparative sequence study of mouse and human X chromosomes concluded that the murine X is not significantly enriched for LINEs, the authors did

not conduct any formal analysis (Chureau *et al.*, 2002). While the absolute frequency of LINEs on the murine X (16.2%) is not as high as in humans (30.2%), as demonstrated by early FISH studies (Boyle *et al.*, 1990), the data of Chureau *et al.* (2002) show that the murine X is significantly enriched for LINEs by comparison with the murine autosomes ($P<0.0001$, chi-square test). Indeed, the magnitude of this enrichment of LINEs on the X when compared to the autosomes is similar in both mouse (1.54-fold) and human (1.78-fold), consistent with Lyon's repeat hypothesis.

Bailey *et al.* observed that X-linked genes escaping X-inactivation have significantly lower L1 content than genes subject to X-inactivation (Bailey *et al.*, 2000). However, the human X chromosome is thought to have evolved by way of successive translocation of autosomal material onto its short arm (Lahn and Page, 1999). As each of these more recent additions to the X contains progressively higher densities of genes escaping X-inactivation (Carrel *et al.*, 1999; Disteche, 1999), it was unclear whether the association of reduced L1 density with escape from inactivation was of functional significance, or instead simply a reflection of the recent autosomal origins of Xp. The data presented here show that this correlation between L1 density and the spread of X-inactivation exists not only for the X but also on the autosomes, strongly suggesting that L1 density is an important factor in determining susceptibility to X-inactivation.

These data also indicate that the association between repeat density and X-inactivation is specifically limited to intragenic repeat sequences, i.e. those within introns and untranslated regions. When the analysis was extended to include additional 5' non-coding sequence the correlation between L1 density and inactivation diminished linearly. This may explain why the initial investigation of the relationship between the spread of inactivation and LINE density at the cytogenetic band level found no apparent correlation. However, this was surprising given that Lyon's repeat hypothesis was based partly on her observations of an apparent correlation between the spread of X-inactivation and LINE density at the cytogenetic level. The failure of this preliminary sequence study to show any such correlation may also be due in part to the draft and incomplete nature of much of the dataset in the April 2001 Genome Assembly which was used in the analysis. At this time large regions of chromosomes 11p and 6p in particular, which contain the majority of the translocated autosomal genes analysed by RT-PCR/methylation analysis, were relatively poorly sequenced. In

contrast, in the latter more focused analysis which used the improved December 2001 Genome Assembly, a total of only 3 and 6 contig gaps were encountered when 5kb and 10kb of additional 5' sequence were included, respectively. Thus, it is possible that the failure of the initial study to identify any large-scale correlation between LINE density and the spread of X-inactivation was due simply to the poor quality of the sequence data upon which it was based.

The distribution of L1s and LTRs within the genome also explains the variable and discontinuous spread of X-inactivation which is seen in X;autosome translocations (Mattei *et al.*, 1982; White *et al.*, 1998). As LINE elements occur throughout the mammalian genome at varying density (Korenberg and Rykowski, 1988; Boyle *et al.*, 1990), X-inactivation may spread to different extents through autosomal chromatin, jumping regions of low LINE density and silencing LINE-rich genes located more distally.

The enrichment of LINEs on the X chromosome is mostly attributable to younger L1 elements (Bailey *et al.*, 2000). Bailey *et al.* hypothesised that this occurred by selective retrotransposition of young L1 subtypes onto the X and demonstrated that autosomes were excluded from this process. This may explain why we observed that both old and young L1 subtypes were associated with inactivated autosomal genes and also partly why the spread of inactivation into autosomal DNA is less efficient than on the X (White *et al.*, 1998). However, the modest correlation ($r=0.45$) of L1s with inactivation and the obvious overlap in L1 content between some active and inactive genes (Appendix 9) indicates that the spread of X-inactivation is not dictated by LINE density alone. LTRs, which are enriched on the human X chromosome (Bailey *et al.*, 2000), were also correlated with inactivation independently of LINEs. These data suggest that in addition to LINEs, LTRs may also function to mediate the spread of X-inactivation, although this possibility will require further investigation. Additionally, in X;autosome translocations factors such as the position of translocation breakpoints and the relative distance of a particular gene from *cis*-linked X chromatin also influence the spread of inactivation, and other as yet unidentified factors may also play a role. Because of these confounding effects, the relatively small number of genes examined, and the imperfect nature of the current Human Genome sequence assembly, it is likely that the degree of correlations reported here of LINEs and LTRs with inactivation are conservative.

Do LINE elements represent binding sites for an *XIST* RNA silencing complex? FISH studies show that *XIST/Xist* RNA localises to G-light bands on the X chromosome (L. Hall and J. Lawrence, personal communication; Duthie *et al.*, 1999) while LINEs show the opposite distribution, being preferentially associated with G-dark bands on the autosomes (Korenberg and Rykowski, 1988; Boyle *et al.*, 1990). Although there is no such distinct relationship between LINEs and G-bands on the X chromosome (Korenberg and Rykowski, 1988; Boyle *et al.*, 1990; Bailey *et al.*, 2000), the available evidence indicates that LINE elements alone are not good candidates, and thus this is one aspect of Lyon's repeat hypothesis that is not supported by these data. However, previous studies have demonstrated that *XIST* RNA binds to the chromatin scaffold and not directly with DNA (Clemson *et al.*, 1996). In addition, the relatively low nuclear abundance of *XIST* RNA means that it is only able to bind to the inactive X at a mean density of approximately one molecule per 100kb of DNA (Buzin *et al.*, 1994). Given these factors, it seems perhaps unlikely that *XIST* binds simply to LINEs.

In her hypothesis, Lyon speculated that X-inactivation might represent a form of repeat-induced silencing facilitated by dense arrays of LINEs being recognized and converted to heterochromatin (Lyon, 1998). Repeat-induced silencing is a mechanism which operates in eukaryotes whereby the transcription of a repeated element reduces with increasing copy number, and is thought to act as a genome defence mechanism to protect against the invasion and proliferation of viruses and other retrotransposons (Wolffe, 1997). Both LINEs and LTRs were associated with the spread of X-inactivation. Not only are these repeats retrotransposons, but also LTRs are almost exclusively derived from retroviruses (Smit, 1999). Thus, the association of these elements with the spread of X-inactivation fulfils Lyon's prediction, suggesting that the X-inactivation process may have evolved from a genome defence mechanism against the invasion of transposable elements and ascribing functional significance to what is often referred to as "junk DNA". It is interesting to note that repetitive DNA has also been recently implicated with both imprinting (Greally, 2002) and in the germ-line inactivation of the X chromosome in *C. elegans* (Fong *et al.*, 2002), and it is therefore tempting to speculate that repeat sequences may have acquired a more widespread role in epigenetic regulation than is currently appreciated.

4.3.4 Summary

In summary I have characterised the spread of X-inactivation through autosomal chromatin in five unbalanced X;autosome translocations. This study demonstrates that transcriptional silencing by a spread of X-inactivation can occur in the apparent absence of features normally associated with the inactive X, raising further questions about the mechanism by which the X-inactivation signal is transmitted and maintained in these cases. These data emphasise that transcription studies are necessary to accurately determine the spread of X-inactivation into autosomal chromatin, and cytogenetic characteristics such as late-replication and histone hypoacetylation are not necessarily associated with inactivation of autosomal genes. In particular, I have shown that the use of replication timing studies to characterise the extent of spread of inactivation in X;autosome rearrangements can be misleading, forcing a reassessment of previous studies which have relied upon this parameter to determine the spread of X-inactivation. However, where X-inactivation spreads in a continuous fashion, cytogenetic features such as depletion of histone acetylation and H3 lysine 4 dimethylation provide more reliable indicators of the extent of spread of X-inactivation than replication timing studies.

Investigation of the repetitive sequence content of translocated autosomal genes that are either subject to or resist the X-inactivation signal provides compelling evidence to support Lyon's hypothesis that LINE-1 repeats are 'booster elements' which promote the spread of X-inactivation *in cis*. LTRs are also associated with inactivation suggesting that these elements may also function in this regard, although confirmation of the role of these repeats in the X-inactivation process will require further experimentation. These results explain the variable and often discontinuous spread of X-inactivation which is seen in X;autosome translocations and indicate the underlying mechanism from which the X-inactivation process may have evolved, suggesting a functional role for "junk DNA" within our genome.

Section 5 - Appendices

5.1 Appendix 1. Population A - Females aged ≤20

NO	SPEC ID	CLASS	AGE		X-INACTIVATION RESULT A		X-INACTIVATION RESULT B		MEAN RATIO	
			Years	Months	Height	Area	Height	Area	Height	Area
1	92/0005	PWS[i]	0	0	75-25		70-30		73-27	
2	94/552	PWS	16	3	81-19		79-21		80-20	
3	942002	PWS	0	1	80-20		65-35		73-27	
4	952264	PWS	0	0	69-31	69-31	63-37	65-35	66-34	67-33
5	9604489	PWS	0	3	63-37	61-39				
6	9703641	PWS	0	1	55-45	58-42	55-45	52-48	55-45	55-45
7	901785	PWS	1	9	67-33		60-40		64-36	
8	91/1471	PWS	1	8	72-28	75-25	74-26	75-25	73-27	75-25
9	952097	PWS	1	8	100-0		100-0		100-0	
10	950499	PWS	1	5	63-37					
11	901540	PWS	2	1	87-13					
12	91/157	PWS	2	4	80-20					
13	93/1184	PWS	2	2	68-32		66-34		67-33	
14	92/368	PWS	2	7	77-23		67-33		72-28	
15	92/583	PWS	2	4	62-38		56-44		59-41	
16	93/1539	PWS	2	2	Homo	Homo	Homo	Homo	Homo	Homo
17	93/1565	PWS	3	0	56-44	56-44	79-21	86-14	68-32	71-29
18	951962	PWS	3	9	52-48		53-47		53-47	
19	942229	PWS	4	11	Homo	Homo	Homo	Homo	Homo	Homo
20	941685	PWS	4	9	84-16		77-23		81-19	
21	89/843	PWS	5	9	Homo	Homo	Homo	Homo	Homo	Homo
22	91/105	PWS	5	6	58-42					
23	93/226	PWS	5	11	67-33	63-37				
24	942199	PWS	16	8	65-35	65-35	69-31	69-31	67-33	67-33
25	92/1417	PWS	5	6	77-23		77-23	78-22	77-23	73-27
26	951078	PWS	5	6	71-29		76-24		74-26	
27	94/709	PWS	6	11	51-49	50-50	52-48	51-49	52-48	51-49
28	9608804	PWS	7	3	54-46		69-31	76-24	62-38	
29	90/1865	PWS	8	1	81-19		81-19		81-19	
30	92/226	PWS	14	8	93-7		90-10		92-8	
31	9607104	PWS	8	8	Homo	Homo				
32	93/1249	PWS	9	3	62-38					
33	93/1328	PWS	9	3	75-25	71-29	75-25	76-24	75-25	74-26
34	950621	PWS	9	6	91-9		82-18		87-13	
35	9605051	PWS	9	4	69-31	70-30	52-48	52-48	61-39	61-39
36	941620	PWS	10	0	Homo	Homo				
37	950550	PWS	10	0	91-9	91-9	94-6		93-7	
38	94/212	PWS	11	7	Homo	Homo				
39	941142	PWS	11	1	50-50					
40	950050	PWS	11	7	78-22		74-26		76-24	
41	9605806	PWS	11	9	69-31	72-28				
42	94/629	PWS	13	7	53-47					
43	91/1456	PWS	13	10	59-41		53-47		56-44	
44	92/226	PWS	14	8	93-7		90-10		92-8	
45	951503	AS[ii]	15	1	53-47	53-47				
46	951501	AS	5	5	73-27		61-39		67-33	
47	941985	AS	8	3	67-33		65-35		66-34	
48	90/1606	AS	2	6	70-30					
49	91/196	AS	2	2	59-41		57-43		58-42	

NO	SPEC ID	CLASS	AGE		X-INACTIVATION RESULT A		X-INACTIVATION RESULT B		MEAN RATIO	
			Years	Months	Height	Area	Height	Area	Height	Area
50	91/1126	AS	2	5	53-47					
51	950303	AS	17	11	81-19	82-18	81-19	83-17	81-19	83-17
52	94/154	AS	18	10	57-43	55-45	65-35	69-31	61-39	62-38
53	91/1229	AS	12	5	71-29					
54	92/0075	AS	0	10	75-25		73-27		74-26	
55	92/0101	AS	0	11	84-16	86-14	86-14	87-13	85-15	87-13
56	92/1249	AS	2	3	55-45		54-46		55-45	
57	950059	AS	1	4	63-37		76-24		70-30	
58	93/0641	AS	0	2	51-49		48-52		50-50	
59	93/1086	AS	0	7	72-28					
60	94/442	AS	2	7	66-34		61-39		63-37	
61	941434	AS	3	6	70-30		74-26	75-25	72-28	77-23
62	941072	AS	4	3	93-7	97-3	100-0	100-0	97-3	99-1
63	950488	AS	2	7	Homo	Homo	Homo	Homo	Homo	Homo
64	9608096	CF[iii]	4	2	68-32	73-27	57-43		63-37	
65	9601728	CF	3	7	77-23					
66	9703914	CF	3	3	61-39		52-48		57-43	
67	9700661	CF	17	4	Homo	Homo				
68	9701653	CF	6	4	50-50	51-49	52-48	52-48	51-49	52-48
69	9703119	CF	9	3	60-40	61-39				
70	9705243	CF	18	6	94-6		94-6		94-6	
71	9603890	CF	10	0	Homo	Homo	Homo	Homo	Homo	Homo
72	9600467	CF	1	11	Homo	Homo	Homo	Homo	Homo	Homo
73	9701789	CF	3	6	75-25	80-20	77-23	76-24	76-24	78-22
74	9608738	CF	6	8	65-35		77-23		71-29	
75	9704965	CF	3	2	58-42	59-41				
76	9609345	CF	5	8	93-7	93-7	91-9		92-8	
77	9705540	CF	6	8	52-48	57-43				
78	9702319	CF	14	7	50-50	51-49	51-49	49-51	51-49	50-50
79	9703316	CF	6	6	Homo	Homo	Homo	Homo	Homo	Homo
80	9701625	CF	6	1	60-40					
81	9701464	CF	8	4	56-44	63-37				
82	9701162	CF	15	11	67-33					
83	9608878	CF	8	1	59-41		64-36		62-38	
84	9704409	CF	7	4	Homo	Homo				
85	9701743	CF	2	9	88-12	89-11	85-15		87-13	
86	9605805	CF	7	8	74-26	71-29				
87	9606143	CF	6	1	66-34		63-37		65-35	
88	9600816	CF	6	0	78-22	79-21				
89	960029	CF	12	0	73-27	75-25				
90	952802	CF	11	4	58-42	57-43				
91	9700360	CF	4	6	95-5	96-4	93-7		94-6	
92	9603815	CF	3	9	63-37	64-36	64-36	65-35	64-36	65-35
93	93/1615	PPC[iv]	19	1	78-22	82-18	88-12		83-17	
94	9602562	PPC	10	3	100-0	100-0	100-0	100-0	100-0	100-0
95	9705228	PPC	16	3	78-22	82-18				
96	8922	PPC	10	1	76-24	68-32	71-29		74-26	
97	941724	PPC	17	2	73-27	72-28	68-32		71-29	
98	93/1134	PPC	15	8	78-22	75-25				
99	92/888	PPC	16	1	51-49	58-42				

[i] Prader-Willi syndrome
[ii] Angelman syndrome
[iii] Cystic fibrosis
[iv] Polyposis coli

5.2 Appendix 2. Population B - Females Aged ≥60

NO	SPEC ID	CLASS	AGE		X-INACTIVATION RESULT A		X-INACTIVATION RESULT B		MEAN RATIO		CHILDREN		GRAND CHILDREN	
			Years	Month	Height	Area	Height	Area	Height	Area	MALE	FEMALE	MALE	FEMALE
1	921859	BCS[i]	76	7	70-30		56-44		63-37					
2	92/974	BCS	64	0	63-37	53-47	55-45	60-40	57-43	57-43				
3	92/1101	BCS	73	0	85-15	87-15								
4	92/1256	BCS	73	1	76-24		40-60		58-42					
5	92/1375	BCS	62	8	Homo	Homo	Homo	Homo	Homo	Homo				
6	92/1376	BCS	62	7	91-9		89-11		90-11					
7	93/0896	BCS	69	2	60-40	58-42								
8	93/0897	BCS	65	1	73-27	71-29								
9	93/980	BCS	65	4	74-26	77-23								
10	93/1135	BCS	68	4	73-27		74-26		74-26					
11	93/1329	BCS	78	2	67-33		74-26		71-29					
12	93/1441	BCS	65	6	57-43	56-44								
13	93/1527	BCS	80	5	61-39	64-36								
14	93/1566	BCS	61	9	Homo	Homo								
15	93/1582	BCS	69	8	60-40		56-44		58-42					
16	93/1614	BCS	68	9	87-13	88-12								
17	93/1631	BCS	79	1	57-43		56-44		57-43					
18	93/1653	BCS	69	2	94-6		96-4		95-5					
19	93/1806	BCS	67	9	78-22		73-27		76-24					
20	94/454	BCS	84	4	80-20		87-13		84-16					
21	94/747	BCS	72	1	Homo	Homo								
22	94/748	BCS	64	10	86-14		87-13		87-13					
23	94/917	BCS	68	5	Homo	Homo	Homo	Homo	Homo	Homo				
24	941352	BCS	64	7	92-8	88-12	91-9	95-5	92-8	92-8				
25	941374	BCS	80	10	Homo	Homo	Homo	Homo	Homo	Homo				
26	942279	BCS	70	6	88-12									
27	942334	BCS	65	8	81-19		88-12		85-15					
28	952008	BCS	71	11	69-31		66-34		68-32					
29	952676	BCS	66	7	80-20		84-16		82-18					
30	950064	BCS	73	4	66-34		73-27		70-30					
31	951691	BCS	73	11	83-17	79-21								
32	951714	BCS	81	7	75-25	75-5								

NO	SPEC ID	CLASS	AGE		X-INACTIVATION RESULT A		X-INACTIVATION RESULT B		MEAN RATIO		CHILDREN		GRAND CHILDREN	
			Years	Month	Height	Area	Height	Area	Height	Area	MALE	FEMALE	MALE	FEMALE
33	951716	BCS	84	7	81-19	80-20								
34	9600729	BCS	81	1	74-26	76-24								
35	9601938	BCS	63	10	81-19	82-18								
36	9602697	BCS	62	7	79-21	82-18								
37	9603896	BCS	61	10	60-40									
38	9603964	BCS	65	9	71-29	74-26								
39	9604831	BCS	72	1	52-48	54-46	55-45	55-45	54-46	55-45				
40	9606079	BCS	62	11	75-25	75-25	73-27	74-26	74-26	75-25				
41	9608098	BCS	64	5	55-45		55-45		55-45					
42	9609001	BCS	67	2	57-43	58-42								
43	9609152	BCS	77	0	Homo	Homo								
44	9609414	BCS	62	0	64-36	63-37								
45	9700062	BCS	73	4	65-35	67-33								
46	9700205	BCS	76	4	97-3	98-2								
47	9700844	BCS	67	2	56-44	55-45								
48	9701166	BCS	61	6	51-49		50-50		51-49					
49	9701165	BCS	65	5	85-15	87-13								
50	9701955	BCS	67	11	73-27	72-28								
51	9702315	BCS	64	5	73-27	72-28								
52	9703764	BCS	77	4	Homo	Homo								
53	9704548	BCS	67	1	53-47	52-48								
54	9704810	BCS	64	1	57-43	57-43								
55	9705481	BCS	69	2	80-20	79-21								
56	9705632	BCS	61	2	81-19	80-20								
57	9706173	BCS	68	2	79-21	77-23								
58	9706291	BCS	83	8	56-44	58-42								
59	9706490	BCS	61	0	84-16	84-16								
60	9706696	BCS	71	0	70-30	72-28								
61	85/1280	HD[ii]	71	5	Homo	Homo	Homo	Homo	Homo	Homo	2	6	11	11
62	85/1337	HD	63	10	66-34						2	0	3	2
63	86/510	HD	77	5	Homo	Homo	Homo	Homo	Homo	Homo	3	2	7	8
64	86/2132	HD	68	6	58-42						0	3	3	5
65	86/3225	HD	71	6	92-8	91-9					1	0		
66	86/5799	HD	61	7	53-47	66-34					1	2	0	1
67	87/2296	HD	72	4	54-46		51-49		53-47		1	4	8	1
68	87/3066	HD	70	8	65-35						2	2	2	2

NO	SPEC ID	CLASS	AGE		X-INACTIVATION RESULT A		X-INACTIVATION RESULT B		MEAN RATIO		CHILDREN		GRAND CHILDREN	
			Years	Month	Height	Area	Height	Area	Height	Area	MALE	FEMALE	MALE	FEMALE
69	960062	HD	63	10	64-36									
70	8800308	HD	77	5	76-24		79-21		78-22		1	1	3	3
71	88/2555	HD	67	10	58-42		54-46		56-44		0	0		
72	88/2841	HD	67	0	95-5		97-3		96-4		1	2	3	1
73	88/2975	HD	76	11	Homo	Homo					1	2	5	1
74	88/2976	HD	73	5	96-4	97-3	94-6	96-4	95-5	97-3	0	3		
75	88/3191	HD	66	3	93-7	95-5	94-6	96-4	94-6	96-4	0	2	2	2
76	88/3300	HD	62	11	72-28	72-28	63-37	62-38	68-32	67-33	3	1		
77	88/4053	HD	73	8	55-45	56-44	53-47	54-46	54-46	55-45	1	1	0	2
78	88/6000	HD	65	10	50-50						2	1	0	4
79	89/325	HD	74	3	73-27	74-26					3	4	1	1
80	89/487	HD	66	0	Homo	Homo	Homo	Homo	Homo	Homo	2	0	1	1
81	89/575	HD	68	7	84-16	85-15					0	2	1	3
82	89/621	HD	70	9	88-12		87-13		88-12		1	1	2	0
83	91/667	HD	86	11							2	0	4	4
84	91/1129	HD	61	5	Homo	Homo	Homo	Homo	Homo	Homo				
85	91/1218	HD	63	1	92-8						0	1		
86	92/181	HD	72	10	73-27						2	1		
87	921778	HD	65	10	90-10	76-24					4	4	6	3
88	93/224	HD	61	4	67-33	67-33					0	1		
89	92/915	HD	67	11	93-7	94-6								
90	92/1078	HD	74	2	74-26	76-24								
91	93/0340	HD	63	11	Homo	Homo	Homo	Homo	Homo	Homo		1		
92	93/904	HD	64	9	53-47	44-56	61-39	60-40	57-43	52-48				
93	93/1285	HD	71	4	58-42	71-29	58-42	75-25	58-42	73-27	2	0		
94	93/1634	HD	65	7	84-16	85-15	85-15	85-15	85-15	85-15				
95	93/1802	HD	72	7	Homo	Homo	Homo	Homo	Homo	Homo				
96	94/0043	HD	64	2	Homo	Homo	Homo	Homo	Homo	Homo				
97	94/221	HD	72	10	86-14	86-14	70-30	74-26	78-22	80-20				
98	94/516	HD	69	8	87-13	89-11	97-3	98-2	92-8	94-6				
99	94/690	HD	69	0	86-14	80-20	85-15	86-14	86-14	83-17				
100	941511	HD	75	0	81-19	73-27	52-48	55-45	67-33	64-36				
101	941572	HD	67	9	53-47	56-44								
102	941701	HD	81	10	88-12	95-5	98-2	98-2	93-7	97-3				
103	94/2210	HD	70	0	52-48	51-49								

NO	SPEC ID	CLASS	AGE		X-INACTIVATION RESULT A		X-INACTIVATION RESULT B		MEAN RATIO		CHILDREN		GRAND CHILDREN	
			Years	Month	Height	Area	Height	Area	Height	Area	MALE	FEMALE	MALE	FEMALE
104	942221	HD	68	3	96-4	97-3	97-3	97-3	97-3	97-3				
105	942269	HD	74	8	69-31	67-33	55-45	64-36	62-38	66-34	2	0		
106	950090	HD	69	2	75-25	78-22								
107	950992	HD	70	0	78-22						0	1	0	1
108	952456	HD	74	9	82-18	85-15								
109	952568	HD	72	4	Homo	Homo	Homo	Homo	Homo	Homo				
110	9603887	HD	64	5	Homo	Homo								
111	9608688	HD	77	0	55-45						1	1		
112	9704168	HD	66	3	54-46	52-48								
113	9700380	HD	74	5	88-12	89-11								
114	9602706	HD	75	0	79-21	79-21								
115	9701196	HD	77	3	91-9	92-8	92-8	91-9	92-8	92-8				

[i] Breast cancer screening
[ii] Huntington disease

5.3 Appendix 3. Population C - Females aged ≤25

No.	DOB	AGE	BLOOD			MOUTHBR USH			URINE		
		Y+M	ResultA	Result B	Mean	Result A	Result B	Mean	Result A	Result B	Mean
1	17.10.79	18 + 7	37-63	35-65	36-64	39-61	42-58	40-60	46-54	38-62	42-58
2	21.9.73	24 + 8	62-38	57-43	60-40	59-41	64-36	62-38	63-37	65-35	64-36
3	5.4.79	19 + 0	72-28	69-31	71-29	65-35			70-30		
4	30.4.74	24 + 1	84-16			67-33	72-28	70-30	87-13	88-12	88-12
5	18.11.78	19 + 7	32-68	29-71	30-70	19-81	20-80	19-81	27-83	34-66	31-69
6	21.1.80	18 + 5	Homo	Homo	Homo	Homo	Homo	Homo	Homo	Homo	Homo
7	11.5.74	24 + 3	36-64			54-46			64-36		
8	10.3.74	24 + 5	50-50			53-47			62-38		
9	5.9.88	10 + 0				43-57	32-68	37-63	13-87	12-.88	12-.88
10	25.12.87	10 + 9	40-60	44-56	42-58	35-65	33-67	34-66	31-69	38-62	34-66
11	13.11.83	14 + 10	55-45	58-42	57-43	41-59	53-47	47-53	68-32	67-33	68-32
12	20.10.72	25 + 11	85-15	85-15	85-15	54-46	55-45	55-45	75-25	73-27	74-26
13	19.1.77	21 + 9	64-36	52-48	58-42	62-38	56-44	59-41			
14	23.2.75	23 + 8	65-35			38-62	42-58	40-60			
15	3.9.77	21 + 1	45-55			48-52	48-52	48-52	41-59		
16	28.6.79	19 + 8	20-80	19-81	19-81	22-78	25-75	23-77			
17	5.5.79	19 + 7	26-74			43-57			34-66		
18	27.2.76	22 + 9	70-30			46-54			76-24		
19	12.6.79	19 + 6	Homo	Homo	Homo	Homo	Homo	Homo	Homo	Homo	Homo
20	10.5.77	21 + 7	70-30			74-26			58-42		
21	14.3.77	21 + 11	44-56	46-54	45-55	47-53			17-83	12.-82	15-85
22	24.2.76	22 + 11	Homo	Homo	Homo	Homo	Homo	Homo	Homo	Homo	Homo
23	18.10.75	23 + 3	60-40			56-44			67-33		
24	8.1.74	25 + 1	21-79	38-62	30-70	22-78	27-73	25-75	14-86	17-83	16-84
25	29.6.78	20 + 8	36-64			46-54			26-74	47-53	36-64
26	11.2.76	23 + 0	67-33	61-39	64-36	38-62			58-42	65-35	62-38
27	19.11.81	17 + 3				74-26			73-27	83-17	78-22
28	27.7.78	20 + 7	14-86			19-81			49-51		
29	15.5.80	18 + 8	85-15	92-8	89-11	89-11			91-9	89-11	90-10
30	29.10.74	24 + 5				72-28			67-33		
31	24.8.76	22 + 7	53-47			50-50			47-53		
32	12.7.74	24 + 8				61-39			55-45	67-33	61-39
33	3.10.78	20 + 5	Homo	Homo	Homo	Homo	Homo	Homo	Homo	Homo	Homo
34	7.4.73	25 + 11				56-44	65-35	61-39	26-74	24-76	25-75
35	19.4.77	21 + 11	Homo	Homo	Homo	Homo	Homo	Homo	Homo	Homo	Homo
36	4.10.77	21 + 4				66-34					
37	2.11.73	25 + 4	65-35	78-22	72-28	55-45	48-52	52-48	56-44	51-49	54-46
38	13.6.76	22 + 9	67-33	67-33	67-33	20-80			32-68	38-62	35-65
39	14.6.76	22 + 9	90-10	85-15	88-12	80-20			87-13		
40	8.4.75	23 + 11	34-66	17-83	26-74	30-70	30-70	30-70	68-32	62-38	65-35
41	1.9.73	25 + 6	41-59	46-54	43-57	31-69			26-74	39-61	32-68
42	25.4.78	20 + 11	45-55	52-48	48-52	30-70	25-75	27-73	37-63		
43	6.4.76	23 + 0	Homo	Homo	Homo	Homo	Homo	Homo	Homo	Homo	Homo
44	5.12.80	18 + 3	67-33	71-29	69-31	63-37			72-28		
45	15.2.77	22 + 1	Homo	Homo	Homo	Homo	Homo	Homo	Homo	Homo	Homo
46	2.9.76	22 + 8	28-72	24-76	26-74	40-60	46-54	43-57	20-80		
47	11.10.75	23 + 5				41-59			47-53		
48	1.6.77	21 + 9	72-28			69-31	70-30	70-30			
49	15.4.74	25 + 0				32-68	29-71	31-69	19-81		
50	30.11.77	21 + 5	78-22	82-18	80-20	64-63			75-25	81-19	78-22

5.4 Appendix 4. Population D - Females aged ≥60

NO	DOB	AGE	BLOOD			MOUTHBRUSH			URINE			URINE VOLUME	URINE CYTOLOGY
		Years+Months	Result A	Result B	Mean	Result A	Result B	Mean	Result A	Result B	Mean		
1	20.5.24	73 + 11	31-69	28-72	29-71	54-46	54-46	54-46	35-65	29-71	32-68	40ml	100% squamous
2	27.7.37	60 + 9	76-24			37-63	37-63	37-63	65-35	64-36	65-35	20ml	
3	4.8.24	74 + 6	39-61	33-67	36-64	26-74	17-83	21-79	3 alleles	3 alleles	3 alleles	60ml	100% squamous
4	3.8.21	76 + 8				91-9	94-6	93-7	57-43	55-45	56-44	30ml	
5	28.12.30	67 + 4	99-1	97-3	98-2	78-22	85-15	82-18	98-2	97-3	98-2	30ml	
6	25.9.16	81+ 7	Homo	Homo	Homo	Homo	Homo	Homo					
7	17.4.22	76 + 0	78-22	80-20	79-21	49-51	49-51	49-51					
8	11.3.27	71 + 1	65-35	66-34	66-34	54-46	36-64	45-55	63-37	66-34	65-35	50ml	90% squamous + 10% urothelium
9	18.7.21	76 + 9	74-26	81-19	78-22	41-59			74-26			70ml	70% squamous + 30% urothelium
10	20.4.38	60 + 0	35-65	14-86	24-76	45-55	54-46	49-51				30ml	100% squamous
11	5.12.22	75 + 4	92-8	94-6	93-7	35-65	41-59	38-62	39-61	29-71	34-66	2ml	
12	10.7.21	76 + 10	82-18	81-19	82-18	61-39	57-43	59-41	71-29	48-52	60-40	20ml	100% squamous
13	18.1.16	82 + 3	34-66	39-61	36-64	38-62	46-54	42-58					
14	8.10.34	63 + 7	46-54	53-47	49-51	34-66	44-56	39-61	45-55	38-62	41-59	70ml	
15	7.7.32	65 + 10	23-77	20-80	21-79	48-52	53-47	50-50	45-55	48-52	46-54	50ml	95% squamous+ 5% urothelium
16	25.10.30	67 + 7	Homo	Homo	Homo	Homo	Homo	Homo					
17	17.11.08	97 + 6	52-48	52-48	52-48	55-45	59-41	57-43	35-65			40ml	100% squamous
18	19.8.22	75 + 9	100-0	100-0	100-0	79-21	76-24	78-22					
19	6.2.21	77 + 3	59-41	60-40	60-40	37-63	34-66	36-64	46-54	38-62	42-58	60ml	
20	11.1.34	64 + 4	43-57	49-51	46-54	50-50	55-45	52-48	21-79			70ml	100% squamous
21	21.8.34	63 + 9	15-85	88-12	13-87	30-70	24-76	27-73					
22	14.10.24	73 + 7	94-6	95-5	95-5	60-40	57-43	59-41	65-35	58-42	62-38		100% squamous
23	23.8.24	73 + 9	59-41	54-46	57-43	68-32	67-33	68-32	29-71	45-55	37-63	2ml	
24	30.10.22	75 + 7	23-77	18-82	20-80	41-59	51-49	46-54				20ml	80% squamous + 20% urothelium
25	22.1.32	66 + 4	34-66	44-56	39-61	49-51	48-52	48-52	42-58	42-58	42-58	10ml	80% squamous + 20% urothelium
26	3.2.24	74 + 3	41-59	38-62	39-61	58-42	62-38	60-40	35-65	39-61	37-63		100% squamous
27	26.6.32	65 + 11	Homo	Homo	Homo	Homo	Homo	Homo	Homo	Homo	Homo		100% squamous
28	20.12.22	75 + 6	27-73	20-80	23-77	52-48	52-48	52-48	45-55	28-52	46-54		95% squamous + 5% urothelium
29	11.6.29	68 + 11	35-65	28-72	31-69	38-62	32-68	35-65	46-54	46-54	46-54		90% squamous+10% urothelium
30	7.2.22	76 + 3	37-63	49-51	43-57	79-21	79-21	79-21	57-43	55-45	56-44		90% urothelium+10% squamous

NO	DOB	AGE	BLOOD			MOUTHBRUSH			URINE			URINE VOLUME	URINE CYTOLOGY
		Years+Months	Result A	Result B	Mean	Result A	Result B	Mean	Result A	Result B	Mean		
31	25.7.37	60 + 11	92-8	91-9	92-8	80-20	81-19	81-19	47-53				
32	18.11.21	76 + 7	92-8	93-7	93-7	44-56	45-55	44-56	71-29	75-25	73-27		
33	15.10.21	76 + 8	Homo	Homo	Homo	Homo	Homo	Homo	Homo	Homo	Homo		
34	26.3.29	69 + 3	17-83	26-74	21-79	43-57	41-59	42-58	26-74	30-70	28-72		
35	10.5.15	83 + 1	Homo	Homo	Homo	Homo	Homo	Homo	Homo	Homo	Homo		
36	4.1.18	80 + 5	26-74	16-84	21-79	29-71	34-66	32-68	22-78	32-68	27-73		
37	28.1.20	78 + 5				70-30	80-20	75-25	80-20	80-20	80-20		
38	14.3.38	60 + 3	62-38	60-40	61-39	75-25	83-17	79-21	73-27	75-25	74-26		
39	15.2.19	79 + 5	25-75	21-79	23-77	51-49	56-44	54-46					
40	21.3.26	72 + 3	8.-92	9.-91	8.-92	53-47	52-48	53-47	29-71				
41	3.11.37	60 + 8	61-39	66-34	64-36	57-43	62-38	60-40	66-34	65-35	66-34		
42	28.2.15	83 + 4	51-49	49-51	50-50	49-51	46-54	47-53	73-27				
43	13.11.28	69 + 8	95-5	96-4	96-4	92-8	94-6	93-7	97-3	96-4	97-3		
44	10.4.25	73 + 3	68-32	66-34	67-33	61-39	43-57	52-48	72-28	67-33	70-30		
45	29.1.35	63 + 6	40-60	24-76	32-68	41-59	43-57	42-58	71-29	72-28	72-28		
46	29.2.08	90 + 6	Homo	Homo	Homo	Homo	Homo	Homo	Homo	Homo	Homo		
47	7.7.35	63 + 0	77-23	74-26	76-24	42-58	46-54	44-56	63-37				
48	8.4.28	70 + 3	11.-89	5.-95	8.-92	27-73	20-80	23-77	7.-93	10.-90	8.-92		
49	7.7.38	60 + 0	40-60	32-68	36-64	59-41	56-44	58-42	46-54	46-54	46-54		
50	12.7.12	86 + 0	81-19	85-15	83-17	66-34	60-40	63-37	40-60	44-56	42-58		
51	20.11.10	87 + 8	93-7	89-11	91-9	85-15	91-9	88-12	94-6	93-7	94-6		

5.5 Appendix 5. Population E - Mothers With a High Ratio of Female:Male Offspring

No.	Spec. Id	DOB	Age Years	Category	No. of Sons	No. of Daughters	X-Inactivation Result A	X-Inactivation Result B	Mean X-Inactivation
1	951225	4.2.48	47	HD 15[i]	0	2	Homo	Homo	Homo
2	86/2138	25.10.17	69	HD 16	0	3	51-49	55-45	53-47
3	84/3870	30.9.39	45	HD 31	0	3	43-57		
4	87/3302			HD 59	0	2	47-53	50-50	48-52
5	88/2124			HD 65	1	3	53-47		
6	89/525	12.3.40	49	HD 81	0	3	26-74		
7	1753			HD 94	0	3	27-73	32-68	29-71
8	89/676	27.9.56	33	HD 1	0	4	76-24	75-25	76-24
9	88/1700	27.12.36	52	HD 1	1	6	65-35	67-33	66-34
10	87/3468			HD 4	1	4	72-28	68-32	70-30
11	89/104	5.9.34	65	HD 13	3	6	24-76	21-79	22-78
12	91/0042	30.4.40	51	HD 36	0	4	36-64	38-62	37-63
13	88/2957			HD 47	1	6	54-46	50-50	52-48
14	90/1961	20.8.15	75	HD 56	1	4	46-54		
15	88/2127			HD 65	1	3	36-64		
16	94/232	20.2.58	36	PPC 18[ii]	0	2	55-45	61-39	58-42
17	94/530	1.7.61	33	PPC 20	0	2	51-49	46-54	48-52

[i] Huntington disease
[ii] Polyposis coli

5.6 Appendix 6. X-Inactivation Ratios in Mother-Daughter Pairs

FAMILY NO.	NO.	SPEC. ID	CLASS	AGE		X-Inactivation Ratio
				Years	Months	
1	7	901785	PWS[i]	1	9	64-36
1M	7b	902107	PWS			38-62
2	9	952097	PWS	1	8	100-0
2M	9b	952098	PWS			66-34
3	10	950499	PWS	1	5	63-37
3M	10b	9301579	PWS			Homo
4	11	901540	PWS	2	1	87-13
4M	11b	902742	PWS			N/I
5	14	92/368	PWS	2	7	72-28
5M	14b	9201077	PWS			79-21
6	19	942229	PWS	4	11	Homo
6M	19b	950540	PWS			32-68
7	20	941685	PWS	4	9	81-19
7M	20b	9504391	PWS			48-52
8	21	89/843	PWS	5	9	Homo
8M	21b	894573	PWS			31-69
9	22	91/105	PWS	5	6	58-42
9M	22b	865863	PWS			36-64
10	23	93/226	PWS	5	11	67-33
10M	23b	9300680	PWS			19-81
11	26	951078	PWS	5	6	74-26
11M	26b	9504181	PWS			44-56
12	29	90/1865	PWS	8	1	19-81
12M	29b	90/4093	PWS			N/I
13	34	950621	PWS	9	6	87-13
13M	34b	950622	PWS			N/I
14	35	9605051	PWS	9	4	39-61
14M	35b	9605083	PWS			45-55
15	36	941620	PWS	10	0	Homo
15M	36b	941621	PWS			78-22
16	40	950050	PWS	11	7	76-24
16M	40b	950051	PWS			31-69
17	41	9605806	PWS	11	9	69-31
17M	41b	9605801	PWS			30-70

FAMILY NO.	NO.	SPEC. ID	CLASS	AGE		X-Inactivation
				Years	Months	
17	41	9605806	PWS	11	9	69-31
17M	41b	9605801	PWS			30-70
18	42	94/629	PWS	13	7	53-47
18M	42b	941121	PWS			57-43
19	46	951501	AS[ii]	5	5	67-33
19M	46b	951502	AS			Homo
20	50	91/1126	AS	2	5	53-47
20M	50b	914372	AS			40-60
21	51	950303	AS	17	11	19-81
21M	51b	950990	AS			35-65
22	52	94/154	AS	18	10	39-61
22M	52b	940519	AS			35-65
23	60	94/442	AS	2	7	37-63
23M	60b	950443	AS			82-18
24	61	941434	AS	3	6	72-28
24M	61b	943786	AS			Homo
25	62	941072	AS	4	3	97-3
25M	62b	941726	AS			Homo
26	63	950488	AS	2	7	Homo
26M	63b	942941	AS			36-64
27	64	950059	AS	1	4	70-30
27M	64b	950130	AS			Homo
28	65	951503	AS	15	1	53-47
28M	65b	951505	AS			67-33
29	102	9602562	PPC[iii]	10	3	100-0
29M	102b	9602957	PPC			99-1

[i] Prader-Willi syndrome
[ii] Angelman syndrome
[iii] Polyposis coli

5.7 Appendix 7. SRS females with biparental inheritance of chromosome 7.

Case No.	Age at sampling	Maternal Age at Birth	*AR*1	*AR*2	Mean
1	8	?	72;18	80;20	76;24
2	3	24	59;41	54;46	57;43
3	7	28	Homo	Homo	Homo
4	13	39	64;36	61;39	63;37
5	6	26	78;22	77;23	78;22
6	4	32	51;49	58;42	55;45
7	3	26	53;47	51;49	52;48
8	4	24	63;37	Fail	
9	20	?	56;44	65;35	61;39
10	4	31	65;35	66;34	66;34
11	37	20	83;17	78;22	81;19
12	10	?	80;20	77;23	79;21
13	2.5	39	62;38	60;40	61;39
14	4	19	71;29	57;43	64;36
15	11	29	62;38	66;34	64;36
16	4	31	Homo	Homo	Homo
17	2	24	72;28	71;29	72;28
18	1.5	28	96;4	96;4	96;4
19	12	?	Homo	Homo	Homo
20	2.5	32	Homo	Homo	Homo
21	4	?	100;0	100;0	100;0
22	1.5	34	83;17	83;17	83;17
23	5	26	Homo	Homo	Homo
24	1.5	30	66;34	62;38	64;36
25	8	33	Homo	Homo	Homo
26	14.5	34	91;9	91;9	91;9
27	5.7	?	63;37	56;44	60;40
28	21.8	27	100;0	100;0	100;0
29	35.5	35	78;22	65;35	72;28
30	9.8	23	80;20	84;16	82;16
31	8.2	21	100;0	100;0	100;0
32	0.9	?	70;30	63;37	67;33
33	8.3	?	60;40	58;42	59;41
34	5.8	?	75;25	77;23	76;24

5.8 Appendix 8 – Polymorphisms, restriction enzymes, and primer details used for allele-specific RT-PCR and CpG island methylation analysis

5.8.1 – Polymorphisms within 10q23-qter

Gene	Polymorphism	Primer sequences and labels	Restriction enzyme	Product sizes in bp	Amplification conditions
HPS	c>a	F: cacgcagcccttctccagctt TET R: attccatctgcctcccaaag	*Pst*I	284>284+154	58°, 30 cycles
MXI1	pentamer repeat	F: ttgggagtgacgagggttac TET R: ctctgccccacttttgcta	---	239-244	55°, 30 cycles
DUSP5	c>a	F: tgtctgtctctgttgcttcg R: tggacccaagagggtgtcta HEX	*Hph*I	134>108+87	58°, 35 cycles
ABLIM1	t>a	F: ggagctgcgttgttttcttc FAM R: cctgttccctgatggagttc	*Mse*I	285>285+149	58°, 30 cycles
PRDX3	a>c	F: ggtggctggctagctagttt R: tgatcagttaagtcactctgc FAM	*Nla*III	329>271+96	58°, 35 cycles

5.8.2 – Polymorphisms within 11p12-pter

Gene	Polymorphism	Primer sequences and labels	Restriction enzyme	Product sizes in bp	Amplification conditions
MRPL23	g>a	F: aagccggactacaaggtcg HEX R: ctctgcctctgctgcctct	*HaeII*	153>153+99	55°, 35 cycles
PSMD13	c>t	F: ggagagaaagaacgtgccag FAM R: gttaacaaagcagaggggttt	*AciI*	138>138+98	55°, 35 cycles
HRAS	hexamer repeat	F: ctgtgggtttgcccttcaga FAM R: ctcctacagggtctcctgcc	---	106-118	58°, 38 cycles, 7% DMSO, 1M Betaine, 75µM deaza-dGTP
TSSC3	g>a	F: cagagacagagccccacatt TET R: tcagcgtggatttattttatttca	*PstI*	154>154+95	55°, 32 cycles
RRM1	c>a	F: ggaggaattggtgttgctgt TET R: gtctaaatgccaaggctcca	*TaqI*	187>187+115	55°, 35 cycles
SMPD1	g>a	F: ctcccagaggcccagagcct FAM R: agaaggtcctgtttccccggc	*HphI*	197>148+100	58°, 35 cycles
TAF2H	g>?	F: tctccttagctgcccagaaa R: cacattgatgccatactcgc FAM	*HaeIII*	171>171+95	58°, 32 cycles
PDE3B	g>a	F: tagtttgccaactccacagc TET R: atgccaaattcttggtgagg	*HaeIII*	116>116+97	58°, 35 cycles
SAA1	?	F: ggcacttgaagaagccaactag R: cctggaaaggaaaggaaaggcag TET	*HindIII*	188>188+105	62°, 36 cycles
LMO2	a>g	F: cctggctttttaactcccct R: gtggtttccattctcaaccg FAM	*MscI*	164>164+95	58°, 32 cycles
PDX1	t>g	F: gaagtgatcttcaaagagatggc HEX R: caataatgttgccataacccc	*BsrGI*	161>161+94	58°, 32 cycles

5.8.3 – Polymorphisms within 7q22-qter

Gene	Polymorphism	Primer sequences and labels	Restriction enzyme	Product sizes in bp	Amplification conditions
TES	c>g	F: catgaacttctgcccaatgaggc TET R: tgtgtcaaggatgccacaatgcc	*HphI*	134>134+76	65°, 32 cycles
CNTNAP2	c>g	F: aatcaaactaccttccctggttc R: ttgccagtttttaagcagagc FAM	*BsmAI*	166>166+128	58°, 35 cycles
KCNH2	a>g	F: cgtgctgttcttgctcatgt R: gctgctgttgtagggtttgc TET	*HaeIII*	159>159+108	58°, 35 cycles

5.8.4 – Polymorphisms within 6p12-pter

Gene	Polymorphism	Primer sequences and labels	Restriction enzyme	Product sizes in bp	Amplification conditions
IRF4	c>t	F: tgtgttctgtagactgccatca TET R: ggcaagttttctctgtcttcca	*CfoI*	129>129+76	58°, 32 cycles
SCA1	triplet repeat	F: aactggaaatgtggacgtac FAM R: caacatgggcagtctgag	---	210-240	58°, 32 cycles, 25% Solution Q, 75µM deaza-dGTP
GMPR	c>g	F: ctgctcctgaatggtggaat FAM R: ggtgcaagtgttggtttgaa	*AvaI*	174>174+86	58°, 35 cycles
DEK	g>a	F: cagttaccctgctttgcctc TET R: cgtgaagctggctaggtttc	*TaqI*	254>254+235	58°, 32 cycles
HLA-F	triplet repeat	F: ctgtcctatttcatatgctcagg HEX R: atgaacttgtcctgagaatgaag	---	250-330	62°, 32 cycles
HCR	t>a	F: ttacaagcagcagcgactgt HEX R: tcatcgagcaggacagagag	*AluI*	166>166+91	58°, 35 cycles
MICA	triplet repeat	F: ccttttttcagggaaagtgc TET R: ccttaccatctccagaaactgc	---	182-191	58°, 28 cycles
CSNK2B	t>c	F: tggcaatgaattcttctgtgaa FAM R: gcctgctcaatcaggtcact	*TaqI*	174>174+89	58°, 32 cycles
NOTCH4	triplet repeat	F: atgcagccccttcactgct TET R: tgggtctgaccactgagaca	---	67-76	58°, 32 cycles, 25% Solution Q, 75µM deaza-dGTP
HLA-DRA	c>a	F: cgatcaccaatgtacctcca TET R: cctgtggtgacaggttttcc	*TaqI*	147>125+87	58°, 32 cycles
HLA-DRB5	t>c	F: cacagcaggagtacagatgca FAM R: ttcagcagctctgcagaaaa	*MseI*	159>159+83	58°, 35 cycles
PSMB9	a>g	F: gcgttgtgatgggttctgat TET R: gcagctgaaccagagagtgc	*CfoI*	119>119+88	58°, 35 cycles
BRD2	t>c	F: gccaagaaccaaagactgct HEX R: acttggcctgcatgactacc	*DdeI*	166>118+75	58°, 35 cycles

5.8.5 – CpG islands

Gene	Location	Primer sequences	*MspI/HpaII* sites	*CfoI* sites	Product size, bp	PCR conditions
PGK1	Xq21.1	F: ctgggtctcgcacattcttcac R: aggaacagggcccacactac	4	5	325	60°, 32-35 cycles
PSMD13	11p15.5	F: ccaggtagacgtagaggcgagt R: cccagtagtgttttgcgaggat	2	2	206	60°, 35 cycles, 0.4x *PGK1*
HRAS	11p15.5	F: ccgattcagcatcacaggtc R: acaagtatttgctgagcgccta	5	4	255	60°, 35 cycles, 0.33x *PGK1*
TAF2H	11p15.4	F: cataaacgtcgaagccaggtct R: gcctgtctctgttagtgctctcc	1	1	252	60°, 35 cycles, 0.75x *PGK1*
PDX1	11p13	F: cggtattcccaggaggagagg R: ggtgaggtcacctaaacgctct	3	4	267	60°, 35 cycles, 0.5x *PGK1*
HLA-F	6p21.1	F: ggggattttggcctaaactga R: caacgcacagactgaccgagt	5	2	249	60°, 32 cycles, 0.66x *PGK1*
HCR	6p21.1	F: tctccgagtagcctcctcttca R: aaaaagtggaaggcgggagat	3	7	290	60°, 35 cycles, 0.4x *PGK1*
MICA	6p21.1	F: ggaggtgcaaaagggaagatg R: tctccagcccactggaatttt	4	1	243	60°, 35 cycles, 0.35x *PGK1*

5.9 Appendix 9 – Intragenic repetitive element content of autosomal genes in AH, SP, SR, AL0044 and BO0566

Locus	ACT	GC	LINE	L1	L2	CR1	SINE	*ALU*	MIR	LTR	SIMPL	LOW	TAN	MONO	DI	TRI	TET
IRF4	1	48.50	1.50	1.30	0.21	0	6.04	5.33	0.71	0	0.26	1.55	4.48	0	0.48	0	0
GMPR	1	47.80	3.87	3.19	0.50	0.19	19.04	17.06	1.97	20.67	0.24	0.20	1.93	0	0.37	0.20	0
HLA-F	1	58.30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MICA	1	49.70	32.49	32.49	0	0	6.66	5.11	1.55	0	0.29	0	1.40	0	0	0	0
NOTCH4	1	51.80	12.66	8.31	4.34	0	23.47	18.81	4.66	0	0.31	0.86	2.50	0	0.20	0	0
HLA-DRB1	1	41.70	22.00	15.17	6.83	0	11.40	9.29	1.95	20.75	0.50	0.32	1.38	0	0.14	0	0
TES	1	37.70	9.23	5.12	4.10	0	7.08	3.65	3.43	10.48	0.64	0.80	0	0	0	0	0
HPS	1	49.70	13.55	9.94	3.60	0	9.02	4.95	4.07	7.70	0.27	0.60	0	0	0	0	0
MXI1	1	37.30	14.01	11.42	2.59	0	11.59	9.83	1.75	1.13	0.31	1.77	0	0	0	0	0
DUSP5	1	49.00	1.24	1.24	0	0	6.54	2.23	4.31	0	0	0.51	0	0	0	0	0
ABLIM	1	42.70	8.54	5.78	2.55	0.20	18.11	14.48	3.63	2.59	0.64	0.46	0	0	0	0	0
TSSC3	1	64.90	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RRM1	1	40.30	10.09	5.82	4.27	0	26.34	23.11	3.23	3.98	0.10	1.09	2.13	0	0.22	0	0
SMPD1	1	56.20	0	0	0	0	11.40	11.40	0	0	0	0	0	0	0	0	0
TAF2H	1	59.90	0	0	0	0	7.03	0	7.03	0	0	4.62	0	0	0	0	0
PDE3B	1	37.00	40.74	36.17	3.76	0.81	7.94	5.90	2.04	7.97	0.33	0.72	4.81	0	0.07	0	0
SAA1	1	50.80	0	0	0	0	23.83	19.99	3.84	0	0	0	4.19	0	0	0	0
LMO2	1	46.40	11.17	5.84	5.33	0	17.61	10.38	7.22	0.48	1.01	0.20	5.38	0	2.80	0	0
PDX1	1	35.90	27.80	23.53	3.81	0.47	9.36	7.09	2.27	1.69	0.51	0.52	5.59	0	0.10	0	0
SCA1	3	41.90	10.04	7.06	2.65	0.33	15.50	12.61	2.89	4.96	0.81	0.44	3.81	0	0.21	0.03	0
CNTNAP2	3	38.30	19.38	14.85	4.03	0.50	9.98	7.95	2.03	9.05	1.14	0.61	0	0	0	0	0
HCR	5	51.70	11.03	4.03	7.00	0	28.83	28.04	0.78	0	0	0.14	0.98	0	0	0	0
H19	5	63.60	0	0	0	0	0	0	0	0	0	0	7.10	0	0	0	0
HRAS	5	65.90	0	0	0	0	0	0	0	0	0	0	1.54	0	0	0	0
DEK	10	39.10	0	0	0	0	0	0	0	0	0	0	2.60	0	0.28	0	0
CSNK2B	10	51.90	0	0	0	0	10.14	7.27	2.87	0	0	0	2.01	0	2.01	0	0
HLA-DRA	10	43.00	0	0	0	0	8.71	6.04	2.67	0	0.47	0	1.12	0	0	0	0
PSMB9	10	47.70	0	0	0	0	19.33	11.75	7.58	0	0	0	0	0	0	0	0
BRD2	10	51.80	0	0	0	0	2.54	2.54	0	0	0.57	0.55	0.78	0	0.32	0	0

Locus	ACT	GC	LINE	L1	L2	CR1	SINE	*ALU*	MIR	LTR	SIMPL	LOW	TAN	MONO	DI	TRI	TET
KCNH2	10	59.30	6.42	5.42	1.00	0	6.72	1.79	4.93	0	1.04	1.32	0	0	0	0	0
PRDX3	10	45.90	0	0	0	0	32.64	28.00	4.63	0	0	0.48	0	0	0	0	0
PSMD13	10	43.90	7.38	5.61	1.77	0	18.35	14.96	3.39	0	0.58	0.54	6.87	0	0.38	0	0
MRPL23	10	60.50	0	0	0	0	1.05	0	1.05	0	0	0.48	19.43	0	0	0	0

Repeat density is expressed as a percentage of the total sequence. Activation status (ACT) is coded from 1-10, with 1 being inactive and 2-10 representing an increasing percentage of activity, so that a score of 5 represents 50% activity.

5.10 Appendix 10. Buffers and Reagents

5.10.1 Solutions for DNA extraction

Sucrose Lysis Buffer (SLB) - Stored at 4°C

Sucrose	0.32 M (Sigma)
Triton X-100	1% (Sigma)
$MgCl_2$	1 M (Sigma)
Tris	10 mM (Sigma)
pH	7.5

Resuspension Buffer (RSB) - Stored at 4°C

NaCl	0.075 M (Sigma)
EDTA	0.5 M (Sigma)
pH	8.0

Proteinase K Solution - Stored at -20°C

Proteinase K	50 mg/ml (Roche)
Tris	50 mM (Sigma)
EDTA	10 mM (Sigma)
NaCl	0.2 mM (Sigma)

Tris-EDTA (TE) Buffer - Stored at room temperature

Tris-HCl	1 M (Sigma)
EDTA	0. 1M (Sigma)
pH	8.0

5.10.2 Solutions for staining of urine samples

Haematoxylin Solution - Filtered and stored for not more than 2 weeks in the dark at room temperature

Haematoxylin (Shandon)	5 grams
Ammonium aluminium sulphate (Sigma)	100 grams
Ethyl alcohol (Sigma)	50 ml
dH_2O	1000 ml
Mercuric oxide (Sigma)	2.5 grams
Glacial acetic acid (Sigma)	42 ml

Orange G 6 Solution - Stored at room temperature

95% ethyl alcohol (Sigma)	950 ml
Phosphotungstic acid (Sigma)	0.15 grams
10% aqueous Orange G (Sigma)	50 ml

EA 65 Solution - Filtered and stored in the dark at room temperature

This solution comprises three different stains, namely Light Green SF yellowish, Bismarck Brown and Eosin yellowish. Initially 10% aqueous stock solutions of each of these three stains were made. From these diluted aqueous stocks three alcoholic stock solutions were made as follows:-

Alcoholic stock solution A	0.05% Light Green SF yellowish solution in 95% ethanol
Alcoholic stock solution B	0.5% Bismarck Brown solution in 95% ethanol
Alcoholic stock solution C	0.5% Eosin yellowish solution in 95% ethanol

Finally the EA 65 stain solution was made as follows:-

Alcoholic stock solution A	180 ml
Alcoholic stock solution B	40 ml
Alcoholic stock solution C	180 ml
Phosphotungstic acid	2.4 grams

5.10.3 Solutions for detection of late-replicating chromatin

Bromodeoxyuridine (BrdU) solution - Stored at -20°C

Bromodeoxyuridine (Sigma)	1.5mg/ml
Phosphate-buffered saline (Sigma)	1x solution

Denaturing solution - Stored at 4°C

Formamide (Sigma)	70%
2xSSC	30%

PBS/Block/Tween-20 (PBT) - Stored at 4°C for up to 1 month

Na_2HPO_4 (Sigma)	1mM
K_2HPO_4 (Sigma)	0.3mM
NaCl (Sigma)	12.5mM
Tween-20 (Sigma)	0.1%
Bovine Serum Albumin (Sigma)	1.5%

4% Paraformaldehyde – Filtered and stored at 4°C for up to 3 months.

Paraformaldehyde (Sigma)	4%
Phosphate-Buffered Saline (Sigma)	1x concentrate
pH	7.2

5.10.4 Solutions for histone immunofluorescence

KCM Buffer - Stored at 4°C for up to 1 month

KCl (Sigma)	120mM
NaCl (Sigma)	20mM
Tris-HCl, pH8.0 (Sigma)	10mM
EDTA (Sigma)	0.5mM
Triton X-100 (Sigma)	0.1%

KCB Buffer (Blocking Reagent) - Stored at 4°C for up to 1 month

KCM Buffer	1x
Bovine serum albumin (Sigma)	1%

5.10.5 Solutions for *in situ* hybridisation

Saline Sodium Citrate (SSC) Buffer (1x concentrate) - Stored at room temperature

NaCl (Sigma)	3M
Sodium citrate (Sigma)	0.3M

pH7.0

Denaturing solution - Stored at 4°C

Formamide (Sigma)	70%
2xSSC	30%

Section 6 - References

ABKOWITZ JL, TABOADA M, SHELTON GH, CATLIN SN, GUTTORP P, KIKLEVICH JV (1998) An X chromosome gene regulates haematopoietic stem cell kinetics. *Proc Natl Acad Sci USA* **95**(7):3862-3866

ABU-AMERO S, PRICE S, WAKELING E, STANIER P, TREMBATH R, PREECE MA, MOORE GE (1997) Lack of hemizygosity for the insulin-like growth factor I receptor gene in a quantitative study of 33 Silver Russell syndrome probands and their families. *Eur J Hum Genet* **5**:235-241

ABUELO DN, AHSANUDDHIN AN, MARK HFL (2000) Distal 5q trisomy resulting from an X;5 translocation detected by chromosome painting. *Am J Med Genet* **94**:392-399

ADLER SA, RUGARLI EI, LINGENFELTER PA, TSUCHIYA K, POLINSKI D, LIGGIT HD, CHAPMAN VM, ELLIOT RW, BALLABIO A, DISTECHE CM (1997) Evidence of evolutionary up-regulation of the single active X-chromosome in mammals based on *Clc4* expression levels in *Mus spretus* and *Mus musculus*. *Proc Natl Acad Sci USA* **94**:9244-9248

ALLEN RC, ZOGHBI HY, MOSELEY HB, ROSENBLATT HM, BELMONT JW (1992) 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**:1229-239

ALLERDICE PW, MILLER OJ, KLINGER HP, PALLISTER PD, OPITZ JM (1971) Demonstration of a spreading effect in an X-autosome translocation by combined autoradiographic and quinacrine-fluorescence studies. *Excerpta Medica Int Cong Ser* **233**:14-15

ANDRULIS ED, NEIMAN AM, ZAPPULLA DC, STERNGLANZ R (1998) Perinuclear localization of chromatin facilitates transcriptional silencing. *Nature* **394**:592-595

ARIEL M, ROBINSON E, MCCARREY JR, CEDAR H (1995) Gamete-specific methylation correlates with imprinting of the murine *Xist* gene. *Nature Genet* **9**:312-330

AZOFEIFA J, WALDHERR R, CREMER M (1996) X-chromosome methylation

ratios as indicators of chromosomal activity: evidence of intraindividual divergencies among tissues of different embryonal origin. *Hum Genet* **97**:330-333

BACINO CA, LEE B, SPIKES AS, SHAFFER LG (1999) Trisomy 16q in a female newborn with a *de novo* X;16 translocation and hypoplastic left heart. *Am J Med Genet* **82**:128-131

BAILEY JA, CARREL L, CHAKRAVARTI A, EICHLER EE (2000) Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: The Lyon repeat hypothesis. *Proc Natl Acad Sci USA* **97**(12):6634-6639

BAMFORTH F, MACHIN G, INNES M (1996) X-chromosome inactivation is mostly random in placental tissues of female monozygotic twins and triplets. *Am J Med Genet* **61**:209-215

BARLOW DP, STOGER R, HERMANN BG, SAITO K, SCHWEIFER N (1991) The mouse insulin-like-2 receptor is imprinted and closely linked to the *Tme* locus. *Nature* **349**:84-87

BARR ML, BERTRAM EG (1949) A morphological distinction between neurones of male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. *Nature* **163**:676-677

BARTOLOMEI MS, TILGHMAN SM (1997) Genomic imprinting in mammals. *Ann Rev Genet* **31**:493-525

BEARD C, LI E, JAENISCH R (1995) Loss of methylation activates *Xist* in somatic but not embryonic cells. *Genes Dev* **9**:2325-2334

BELYAEV N, KEOHANE AM, TURNER BM (1996) Differential underacetylation of histones H2A, H3 and H4 on the inactive X chromosome in human female cells. *Hum Genet* **97**:573-578

BENSON G (1999) Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res* **27**:573-580.

BERNADINO J, LAMOLIATTE E, LOMBARD M, NIVELEAU A, MALFOY B (1996) DNA methylation of the X chromosomes of the human female: an *in situ* semi-quantitative analysis. *Chromosoma* **104**:528-535

BETTIO D, RIZZI N, GIARDINO D (1994) Familial translocation (X;3) (p22.3;p23): Chromosomal *in situ* suppression (CISS) hybridization and inactivation pattern study. *Clin Genet* **46**:360-363

BEUTLER E, YEH M, FAIRBANKS VF (1962) The normal human female as a mosaic for X-chromosome activity: Studies using the gene for G-6-PD deficiency as a marker. *Proc Natl Acad Sci USA* **48**:9

BOGGS BA, CHINAULT AC (1994) Analysis of replication timing properties of human X-chromosomal loci by fluorescence *in situ* hybridization. *Proc Natl Acad Sci USA* **91**:6083-6087

BOGGS BA, CHEUNG P, HEARD E, SPECTOR DL, CHINAULT AC, ALLIS D (2001) Differentially methylated forms of histone H3 show unique association patterns with inactive human X chromosomes. *Nature Genet* **30**:73-76

BÖÖK JA, SANTESSON B (1960) More chromosome anomalies. *Lancet* **ii**:191

BORSANI G, TONLORENZI R, SIMMLER MC, DANDOLO L, ARNAUD D, CAPRA V, GROMPE M, PIZZUTI A, MUZNY D, LAWRENCE C, WILLARD HF, AVNER P, BALLABIO A (1991) Characterization of a murine gene expressed from the inactive X chromosome. *Nature* **351**:325-329

BOYES J, BIRD A (1992) Repression of genes by DNA methylation depends on CpG density and promotor strength: evidence for involvement of a methyl-CpG binding protein. *EMBO J* **11**:327-333

BOYLE AL, BALLARD SG, WARD DC (1990) Differential distribution of long and short interspersed element sequences in the mouse genome: chromosome karyotyping by fluorescence in situ hybridization. *Proc Natl Acad Sci USA* **87**:7757-7761

BROCKDORFF N, ASHWORTH A, KAY GF, MCCABE VM, NORRIS DP,PENNY GD, PATEL D, RASTAN S (1991) Conservation of position and exclusive expression of mouse *Xist* from the inactive X chromosome. *Nature* **351**:329-331

BROCKDORFF N, ASHWORTH A, KAY GF, MCCABE VM, NORRIS DP, COOPER PA, SWIFT S, RASTAN S (1992) The product of the mouse *Xist* gene is a 15kb inactive-X specific transcript containing no conserved ORF and located within the nucleus. *Cell* **71**:515-526

BROWN CJ, WILLARD HF (1989) Noninactivation of a selectable human X-linked gene that complements a murine temperature-sensitive cell-cycle defect. *Am J Hum Genet* **45**:592-598

BROWN CJ, LAFRENIERE RG, POWERS VE, SEBASTIO G, BALLABIO A,

PETTIGREW AJ, LEDBETTER DH, LEVY E, CRAIG IW, WILLARD HF (1991a) Localization of the X inactivation centre on the human X chromosome in Xq13. *Nature* **349**:82-84

BROWN CJ, BALLABIO A, RUPERT JL, LAFRENIERE RG, GROMPE M, TONLORENZI R, WILLARD HF (1991b) A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* **349**:38-44

BROWN CJ, HENDRICH BD, RUPERT JL, LAFRENIERE RG, XING Y, LAWRENCE J, WILLARD HF (1992) The human *XIST* gene: analysis of a 17kb inactive-X specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell* **71**:527-542

BROWN CJ, WILLARD HF (1994) The human X-inactivation centre is not required for maintenance of X-chromosome inactivation. *Nature* **368**:154-156

BROWN CJ, MILLER AP, CARREL L, RUPERT JL, DAVIES KE, WILLARD HF (1995) The DXS423E gene in Xp11.21 escapes X chromosome inactivation. *Hum Mol Genet* **4**:251-255

BUCKTON KE, JACOBS PA, RAE LA, NEWTON MS, SANGER R (1971) An inherited X-autosome translocation in man. *Ann Hum Genet* **35**:171-178

BULLER RE, SOOD AK, LALLAS T, BUEKERS T, SKILLING JS (1999) Association between nonrandom X-inactivation and BRCA1 mutation in germline DNA of patients with ovarian cancer. *J Natl Cancer Inst* **91**(4):339-346

BURN J, POVEY S, BOYD Y, MUNRO EA, WEST L, HARPER K, THOMAS D (1986) Duchenne muscular dystrophy in one of monozygotic twin girls. *J Med Genet* **23**:494-500

BUSQUE L, MIO R, MATTIOLI J, BRAIS E, BLAIS N, LALONDE Y, MARAGH M, GILLILAND DG (1996) Non-random X-inactivation patterns in normal females: Lyonization ratios vary with age. *Blood* **88**:59-65

BUZIN C, MANN JB, SINGER-SAM J (1994) Quantitative RT-PCR assays show *Xist* RNA levels are low in mouse female adult tissue, embryos and embryoid bodies. *Development* **120**:3529-3536

CAIULO A, BARDONI B, CAMERINO G, GUIOLI S, MINELLI A, PIANTANIDA M, CROSATO F, DALLA FIOR T, MARASCHIO P (1989) Cytogenetic and

molecular analysis of an unbalanced translocation (X;7)(q28;p15) in a dysmorphic girl. *Hum Genet* **84**:51-54

CAMARGO M, CERVENKA J (1984) DNA replication and inactivation patterns in structural abnormality of sex chromosomes. I. X-A translocations, rings, fragments, isochromosomes, and pseudo-isodicentrics. *Hum Genet* **67**:37-47

CAMERON EE, BACHMAN KE, MYOHANNEN S, HERMAN JG, BAYLIN SB (1999) Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. *Nature Genet* **21**:103-105

CANUN S, MUTCHINICK O, SHAFFER LG, FERNANDEZ C (1998) Combined trisomy 9 and Ullrich-Turner syndrome in a girl with a 46,X,der(9)t(X;9)(q12;q32) karyotype. *Am J Med Genet* **80**:199-203

CARREL L, HUNT PA, WILLARD HF (1996a) Tissue and lineage specific variation in inactive X chromosome expression of the murine *Smcx* gene. *Hum Mol Genet* **5**:1361-1366

CARREL L, CLEMSON CM, DUN JM, MILLER AP, HUNT (1996b) X inactivation analysis and DNA methylation studies of the ubiquitin activating enzyme Ube1 and Pctaire-1 genes in human and mouse. *Hum Mol Genet* **5**:391-401

CARREL L, COTTLE A, WILLARD HF (1998) X inactivation profile of the X chromosome: survey of the first 125 genes (ASHG abstract). *Am J Hum Genet* **63;4**:A4

CARREL L, COTTLE AA, GOGLIN KC, WILLARD HF (1999) A first generation X-inactivation profile of the human X chromosome. *Proc Natl Acad Sci USA* **96**(25):14440-14444

CARREL L, WILLARD HF (1999) Heterogeneous gene expression from the inactive X chromosome: An X-linked gene that escapes X-inactivation in some human cells but is inactivated in others. *Proc Natl Acad Sci USA* **96**(13):7364-7369

CATTANACH BM, WILLIAMS CE (1972) Evidence of non-random X chromosome activity in the mouse. *Genet Res* **19**:229-240

CATTANACH BM, RASBERRY C (1994) Identification of the *Mus cutaneous Xce* allele. *Mouse Genome* **92**:114

CHADWICK BP, WILLARD HF (2001a) Histone H2A variants and the inactive X chromosome: identification of a second macroH2A variant. *Hum Mol Genet* **10**:1101-

1113
CHADWICK BP, WILLARD HF (2001b) A novel chromatin protein, distantly related to H2A, is largely excluded from the inactive X chromosome. *J Cell Biol* **152**:375-384
CHAUVEL PJ, MOORE CM, HASLAM RH (1975) Trisomy 18 mosaicism with features of Russell-Silver syndrome. *Dev Med Child Neurol* **17**:220-224
CHUREAU C, PRISSETTE M, BOURDET A, BARBE V, CATTOLICO L, JONES L, EGGEN A, AVNER P, DURET L (2002) Comparative sequence analysis of the X-inactivation center region in mouse, human, and bovine. *Genome Res* **12**:894-908
CLEMSON CM, MCNEIL JA, WILLARD HF, LAWRENCE JB (1996) *XIST* RNA paints the inactive X-chromosome at interphase: Evidence for a novel RNA involved in nuclear/chromosome structure. *J Cell Biol* **132**(3):259-275
CLEMSON CM, CHOW JC, BROWN CJ, LAWRENCE JB (1998) Stabilization and localization of *Xist* RNA are controlled by separate mechanisms and are not sufficient for X inactivation. *J Cell Biol* **142**:13-23
CLERC P, AVNER P (1998) Role of the region 3' to *Xist* exon 6 in the counting process of X-chromosome inactivation. *Nature Genet* **19**:249-253
COOPER DW (1971) Directed genetic change model for X chromosome inactivation in eutherian mammals. *Nature* **230**:292-294
COOPER DW, JOHNSTON PG, WATSON JM, GRAVES JAM (1993) X-inactivation in marsupials and monotremes. *Semin Dev Biol* **4**:117-128
COSTANZI C, PEHRSON JR (1998) Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature* **393**:599-601
COURTIER B, HEARD E, AVNER P (1995) *Xce* haplotypes show modified methylation in a region of the active X-chromosome lying 3' to *Xist*. *Proc Natl Acad Sci USA* **92**:3531-3535
COUTURIER J, DUTRILLAUX B, GARBER P, RAOUL O, CROQUETTE M-F, FOURLINNIE JC, MAILLARD E (1979) Evidence for a correlation between late replication and autosomal gene inactivation in a familial translocation t(X;21). *Hum Genet* **49**:319-326
CSANKOVSZKI G, PANNING B, BATES B, PEHRSON JR, JAENISCH R (1999) Conditional deletion of *Xist* disrupts histone macroH2A localization but not maintenance of X inactivation. *Nat Genet* **22**:323-324

D'ESPOSITO M, CICCODICOLA A, GIANFRANCESCO F, ESPOSITO T, FLAGIELLO L (1996) A synaptobrevin-like gene in the Xq28 pseudoautosomal region undergoes X-inactivation. *Nature Genet* **13**:227-229

DEVRIENDT K, MATTHIJS G, LEGIUS G, SCHOLLEN E, BLOCKMANS D, VAN GEET C, DEGREEF H, CASSIMAN JJ, FRYNS JP (1997) Skewed X-chromosome inactivation in female carriers of dyskeratosis congenita. *Am J Hum Genet* **60**:581-587

DIETZEL S, SCHIEBEL K, LITTLE G, EDELMANN P, RAPPOLD GA, EILS R, CREMER T (1999) The 3D positioning of ANT2 and ANT3 genes within female X-chromosome territories correlates with gene activity. *Exp Cell Res* **252**(2):363-375

DISTECHE CM, SWISSHELM K, FORBES S, PAGON RA (1984) X-inactivation patterns in lymphocytes and skin fibroblasts of three cases of X-autosome translocations with abnormal phenotypes. *Hum Genet* **66**:71-76

DISTECHE CM (1999) Escapees on the X chromosome. *Proc Natl Acad Sci USA* **96**:14180-14182

DRISCOLL DJ, MIGEON BR (1990) Sex difference in methylation of single copy genes in human meiotic germ cells: implications for X chromosome inactivation, imprinting, and origin of CpG mutations. *Somat Cell Mol Genet* **16**:267-282

DUNCAN PA, HALL JG, SHAPIRO LR, VIBERT BK (1990) Three generation dominant transmission of the Silver-Russell syndrome. *Am J Med Genet* **35**:245-250

DUTHIE SM, NESTEROVA TB, FORMSTONE EJ, KEOHANE AM, TURNER BM, ZAKIAN SM, BROCKDORFF N (1999) *Xist* RNA exhibits a banded localization on the inactive X chromosome and is excluded from autosomal material in *cis*. *Hum Mol Genet* **8**(2):195-204

DYER KA, CRANFIELD TK, GARTLER SM (1989) Molecular cytological differentiation of active from inactive X domains in interphase: Implications for X-chromosome inactivation. *Cytogenet Cell Genet* **50**:116-120

EGGERMANN T, WOLLMANN HA, KUNER R, EGGERMANN K, ENDERS H, KAISER P, RANKE MB (1997) Molecular studies in 37 Silver-Russell syndrome patients: frequency and etiology of uniparental disomy. *Hum Genet* **100**:415-419

EGGERMANN T, MAU U, KLEIN-VOGLER U, KENDZIORRA H, MACKENSEN-HAEN S, SIEVERDING L, ENDERS H, KAISER P (1998)

Molecular and cytogenetic analysis of an X/autosomal translocation: 45,X,dic. *Clin Genet* **53**:293-297

EGGERMANN K, WOLLMANN HA, TOMIUK J, RANKE MB, KAISER P, EGGERMANN T (1999) Screening for mutations in the promotor and the coding regions of the *IGFBP1* and *IGFBP3* genes in Silver-Russell syndrome patients. *Hum Hered* **49**:123-128

EGGERMANN T, KLOOS P, MERGENTHALER S, EGGERMANN K, DOBOS M, RANKE M, WOLLMANN H (2001) *IRS1* and *GRB2* as members of the IGF signal transduction pathway are not associated with intrauterine growth retardation and Silver-Russell syndrome. *Clin Genet* **59**:371-373

EILS R, DIETZEL S, BERTIN E, SCHROCK E, SPEICHER MR (1996) Three-dimensional reconstruction of painted human interphase chromosomes: active and inactive X chromosome territories have similar volumes but differ in shape and surface structure. *J Cell Biol* **135**:1427-1440

EPSTEIN CJ, SMITH S, TRAVIS B, TUCKER G (1978) Both X chromosomes function before visible X chromosome inactivation in female mouse embryos. *Nature* **274**:500-503

FEY MF, PETER HJ, HINDS HL, ZIMMERMANN A, LIECHTI-GALLATI S, GERBER H, SUDER H, TOBLER A (1992) Clonal analysis of human tumours with M27β, a highly informative polymorphic X chromosome probe. *J Clin Invest* **89**:1438-1444

FEY MF, LIECHTI-GALLATI S, VON ROHR A, BORISCH B, THEILKAS L, SCHNEIDER V, OESTREICHER M, NAGEL S, ZIEMIECKI A, TOBLER A (1994) Clonality and X-inactivation patterns in hematopoietic cell populations detected by the highly informative M27β DNA probe. *Blood* **83**:931-938

FIALKOW PJ (1973) Primordial cell pool size and lineage relationships of five human cell types. *Ann Hum Genet* **37**:39-48

FIALKOW PJ (1978) X-chromosome inactivation and the Xg locus. *Am J Hum Genet* **22**:460-463

FONG Y, BENDER L, WANG W, STROME S (2002) Regulation of the different chromatin states of autosomes and X chromosomes in the germ line of *C. elegans*. *Science* **296**:2235-2238

FORD CE, JONES KW, POLANI PE, DE ALMEIDA J, BRIGGS JH (1959) A sex chromosome anomaly in a case of gonadal dysgenesis (Turner's syndrome). *Lancet* **i**:711-713

FRYNS JP (1990a) Chromosome 10, Trisomy 10q2; in BUYSE ML (ed): *Birth Defects Encyclopaedia*, Vol 1, pp 359-360. Cambridge, MA

FRYNS JP (1990b) Chromosome 6, trisomy 6p2; in BUYSE ML (ed), *Birth Defects Encyclopaedia*, Vol 1, pp 343-344. Cambridge, MA

FRYNS JP (1990c) Chromosome 7, mosaic trisomy 7; in BUYSE ML (ed): *Birth Defects Encyclopaedia*, Vol 1 pp 348. Cambridge, MA

FRYNS JP, CHRZANOWSKA K (1990) Chromosome 11, trisomy 11p; in BUYSE ML (ed): *Birth Defects Encyclopaedia*, Vol 1 pp 364. Cambridge, MA

GALE RE, WHEADON H, LINCH DC (1991) X-chromosome inactivation patterns using HPRT and PGK polymorphisms in haematologically normal and post-chemotherapy females. *Br J Haematol* **79**:193-197

GALE RE, WHEADON H, BOULOS P, LINCH DC (1994) Tissue specificity of X-chromosome inactivation patterns. *Blood* **83**:2899-2905

GALE RE, FIELDING AK, HARRISON CN, LINCH DC (1997) Acquired skewing of X-inactivation patterns in myeloid cells of the elderly suggests stochastic clonal loss with age. *Br J Haematol* **98**:512-519

GARCIA-HERAS J, MARTIN JA, WITCHEL SF, SCACHERI P (1997) De novo der(X)t(X;10)(q26;q21) with features of distal trisomy 10q: Case report of paternal origin identified by late replication with BrdU and the human androgen receptor assay (HAR). *J Med Genet* **34**:242-245

GARTLER SM, RIGGS AD (1983) Mammalian X-chromosome inactivation. *Annu Rev Genet* **17**:155-190

GERMAN J (1964) The pattern of DNA synthesis in the chromosomes of human blood cells. *J Cell Biol* **20**:37

GIBBONS RJ, SUTHERS GK, WILKIE AOM, BUCKLE VJ, HIGGS DR (1992) X-linked α-thalessemia/mental retardation (ATR-X) syndrome: localization to Xq12-q21.31 by X-inactivation and linkage analysis. *Am J Hum Genet* **51**:1136-1149

GILBERT SL, SHARP PA (1999) Promoter-specific hypoacetylation of X-inactivated genes. *Proc Natl Acad Sci USA* **96**(24):13825-13830

GOMEZ MR, ENGEL AG, DEWALD G, PETERSON HA (1977) Failure of inactivation of Duchenne dystrophy X-chromosome in one of female identical twins. *Neurology* **27**:537-541

GOODSHIP J, CARTER J, BURN J (1996) X-inactivation patterns in monozygotic and dizygotic female twins. *Am J Med Genet* **61**:205-208

GORSKI JL (1991) Father-to-daughter transmission of focal dermal hypoplasia associated with non-random X chromosome inactivation: support for X-linked inheritance and paternal X chromosome mosaicism. *Am J Med Genet* **40**:332-337

GOTO T, WRIGHT E, MONK M (1997) Paternal X-chromosome inactivation in human trophoblast cells. *Mol Hum Reprod* **3**:77-80

GRAVES JAM, SCHMIDT MM (1992) Mammalian sex chromosomes: design or accident? *Curr Opin Genet Dev* **2**:890-901

GRAVES JAM, WAKEFIELD MJ, TODER R (1998) The origin and the evolution of the pseudoautosomal regions of the human sex chromosomes. *Hum Mol Genet* 7:1991-1996

GREALLY JM (2002) Short interspersed transposable elements (SINEs) are excluded from imprinted regions in the human genome. *Proc Natl Acad Sci USA* **99**:327-332

HAAF T (1995) The effects of 5-azacytidine and 5-azadeoxycytidine on chromosome structure and function: implications for methylation-associated cellular processes. *Pharmacol Ther* **65**:19-46

HANNULA K, LIPSANEN-NYMAN M, KONTIOKARI T, KERE J (2001) A narrow segment of maternal uniparental disomy of chromosome 7q31-qter in Silver-Russell syndrome delimits a candidate gene region. *Am J Hum Genet* **68**:247-253

HANSEN RS, CRANFIELD TK, FJELD AD, GARTLER SM (1996) Role of late replication timing in the silencing of X-linked genes. *Hum Mol Genet* **5**:1345-1353

HARNDEN DG (1962) Nuclear sex in triploid XXY human cells. *Lancet* **ii**:488

HARRIS A, COLLINS J, VETRIE D, COLE C, BOBROW M (1992) X inactivation as a mechanism of selection against lethal alleles: further investigation of incontinentia pigmenti and X linked lymphoproliferative disease. *J Med Genet* **29**:608-614

HARRISON KB (1989) X-chromosome inactivation in the human cytotrophoblast. *Cytogenet Cell Genet* **52**:37-41

HASSOLD T, MERRILL M, ADKINS K, FREEMAN S, SHERMAN S (1995)

Recombination and maternal-age dependent nondisjunction: molecular studies of trisomy 16. *Am J Hum Genet* **57**:867-874

HATCHWELL E, ROBINSON D, CROLLA JA, COCKWELL AE (1996) X inactivation analysis in a female with Hypomelanosis of Ito associated with a balanced X;17 translocation: evidence for functional disomy of Xp. *J Med Genet* **33**:216-220

HEARD E, KRESS C, MONELARD F, COURTIER B, ROUGEULLE C, AVNER P (1996) Transgenic mice carrying an *Xist*-containing YAC. *Hum Mol Genet* **5**:441-450

HEARD E, ROUGELLE C, ARNAUD D, AVNER P, ALLIS CD, SPECTOR DL (2001) Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during inactivation. *Cell* **107**:727-738

HERZING LB, ROMER JT, HORN JM, ASHWORTH A (1997) *Xist* has properties of the X-chromosome inactivation centre. *Nature* **386**:272-275

HOGLUND P, HOLMBERG C, DE LA CHAPELLE A, KERE J (1994) Paternal isodisomy for chromosome 7 is compatible with normal growth and development in a child with congenital chloride diarrhea. *Am J Hum Genet* **55**:747-752

JABLONKA E, GOITEIN R, MARCUS M, CEDAR H (1985) DNA hypomethylation causes an increase in DNase-I sensitivity and an advance in the time of replication of the entire inactive X chromosome. *Chromosoma* **93**:152-6

JACOBS PA, BAIKIE AG, COURT BROWN WM, MACGREGOR TN, MACLEAN N, HARNDEN DG (1959) Evidence for the existence of the 'superfemale'. *Lancet* **ii**:423-425

JEGALIAN K, PAGE DC (1998) A proposed path by which genes common to mammalian X and Y chromosomes evolve to become inactivated. *Nature* **394**:776-780

JENUWEIN T, ALLIS CD (2001) Translating the histone code. *Science* **293**:1074-1080

JEPPESEN P, TURNER BM (1993) The inactive X-chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. *Cell* **74**:281-289

JOHNSTON CM, NESTEROVA TB, FORMSTONE EJ, NEWALL AET, DUTHIE SM, SHEARDOWN SA, BROCKDORFF N (1998) Developmentally regulated *Xist* promotor switch mediates initiation of X inactivation. *Cell* **94**:809-817

JONES C, BOOTH C, RITA D, JAZMINES L, BRANDT B, NEWLAN A,

HORSTHEMKE B (1997) Bilateral retinoblastoma in a male patient with an X;13 translocation: Evidence for silencing of the *RB1* gene by the spreading of X inactivation. *Am J Hum Genet* **60**:1558-1562

JONES PL, VEENSTRA GJC, WADE PA, VERMAARK D, KASS SU, LANDSBERGER N, STROUBOULIS A, WOLFFE AP (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nature Genet* **19**:187-191

JORGENSEN AL, PHILIP J, RASKIND WH, MATSUSHITA M, CHRISTENSEN B, DREYER V, MOTULSKY AG (1992) Different patterns of X inactivation in MZ twins discordant for red-green color-vision deficiency. *Am J Hum Genet* **51**:291-298

JOYCE CA, SHARP A, WALKER JM, BULLMAN H, TEMPLE IK (1999) Duplication of 7p12.1-p13, including *GRB10* and *IGFBP1*, in a mother and daughter with features of Silver-Russell syndrome. *Hum Genet* **105**:273-280

KALOUSEK DK, LANGLOIS S, BARRETT I, YAM I, WILSON DR, HOWARD-PEEBLES PN, JOHNSON MP, GIORGIUTTI E (1993) Uniparental disomy for chromosome 16 in humans. *Am J Hum Genet* **52**:8-16

KAY GF, PENNY GD, PATEL D, ASHWORTH A, BROCKDORFF N, RASTAN S (1993) Expression of *Xist* during mouse development suggests a role in the initiation of X chromosome inactivation. *Cell* **72**:171-182

KAY GF, BARTON SC, SURANI A, RASTAN S (1994) Imprinting and X-chromosome counting mechanisms determine *Xist* expression in early mouse development. *Cell* 77:639-650

KEITGES EA, PALMER CG (1986) Analysis of the spreading of inactivation in eight X autosome translocations utilizing the high resolution RBG technique. *Hum Genet* **72**:231-236

KEOHANE AM, O'NEIL LP, BELYAEV ND, LAVENDER JS, TURNER BM (1996) X-inactivation and histone H4 acetylation in embryonic stem cells. *Dev Biol* **180**:618-630

KEOHANE AM, LAVENDER JS, O'NEILL LP, TURNER BM (1998) Histone acetylation and X-inactivation. *Dev Genet* **22**:65-73

KEOHANE AM, BARLOW AL, WATERS J, BOURN D, TURNER BM (1999) H4 acetylation, *XIST* RNA and replication timing are coincident and define X;autosome boundaries in two abnormal X chromosomes. *Hum Mol Genet* **8**(2):377-383

KEREM B-S, GOTEN R, RICHLER C, MARCUS M, CEDAR H (1983) In situ nick translation distinguishes between active and inactive X chromosomes. *Nature* **304**:88-90

KORENBERG JR, RYKOWSKI MC (1988) Human genome organization: *Alu*, lines, and the molecular structure of metaphase chromosome bands. *Cell* **53**:391-400

KOTZOT D, SCHMITT S, BERNASCONI F, ROBINSON WP, LURIE IW, ILYINA H, MEHES K, HAMEL BC, OTTEN BJ, HERGERSBERG M (1995) Uniparental disomy 7 in Silver-Russell syndrome and primordial growth retardation. *Hum Mol Genet* **4**:583-587

KOTZOT D, BALMER D, BAUMER A, CHRZANOWSKA K, HAMEL BC, ILYINA H, KRAJEWSKA-WALASEK M, LURIE IW, OTTEN BJ, SCHOENLE E, TARIVERDIAN G, SCHINZEL A (2000) Maternal uniparental disomy 7 – review and further delineation of the phenotype. *Eur J Paediatr* **159**:247-256

KRATZER PG, CHAPMAN VM (1981) X chromosome reactivation in oocytes of *Mus caroli*. *Proc Natl Acad Sci USA* **78**:3093-3097

KULHARYA AS, ROOP H, KUKOLICH MK, NACHTMAN RG, BELMONT JW, GARCIA-HERAS J (1995) Mild phenotypic effects of a de novo deletion Xpter-Xp22.3 and duplication 3pter-3p23. *Am J Med Genet* **56**:16-21

LAHN BT, PAGE DC (1999) Four evolutionary strata on the human X chromosome. *Science* **286**:964-967

LANASA CM, HOGGE WA, KUBIK C, BLANCATO J, HOFFMAN EP (1999) Highly skewed X chromosome inactivation is associated with idiopathic recurrent spontaneous abortion. *Am J Hum Genet* **65**:252-254

LATHAM KE (1996) X chromosome imprinting and inactivation in the early mammalian embryo. *Trends Genet* **12**:134-138

LATT SA (1973) Microfluorometric detection of deoxyribonucleic acid replication in human metaphase chromosomes. *Proc Natl Acad Sci USA* **70**:3395-3399

LAU AW, BROWN CJ, PENAHERRERA M, LANGLOIS S, KALOUSEK DK, ROBINSON WP (1997) Skewed X-chromosome inactivation is common in fetuses or newborns associated with confined placental mosaicim. *Am J Hum Genet* **61**:1353-1361

LEE JT, STRAUSS WM, DAUSMAN JA, JAENISCH R (1996) A 450kb transgene

displays properties of the mammalian X-inactivation centre. *Cell* **86**:83-94

LEE JT, JAENISCH R (1997) Long-range *cis* effects of ectopic X-inactivation centres on a mouse autosome. *Nature* **386**:275-279

LEE JT, LU N, HAN Y (1999a) Genetic analysis of the mouse X-inactivation center defines an 80-kb multifunction domain. *Proc Natl Acad Sci USA* **96**:3836-3841

LEE JT, DAVIDOW JS, WARSHAWSKY D (1999b) *Tsix*, a gene antisense to *Xist* at the X-inactivation centre. *Nature Genet* **21**:400-404

LEE JT, LU N (1999) Targeted mutagenesis of *Tsix* leads to nonrandom X-inactivation. *Cell* **99**:47-57

LEE J (2000) Disruption of imprinted X inactivation by parent-of-origin effects at *Tsix*. *Cell* **103**:17-27

LEISTI JT, KABACK MM, RIMOIN DL (1975) Human X-autosome translocations: differential inactivation of the X chromosome in a kindred with an X-9 translocation. *Am J Hum Genet* **27**:441-453

LIFSCHYTZ E, LINDSLEY DL (1972) The role of X inactivation during spermatogenesis. *Proc Natl Acad Sci USA* **69**:182-186

LINGENFELTER PA, ADLER DA, POSLINSKI D, THOMAS S, ELLIOTT RW, CHAPMAN VM, DISTECHE CM (1998) Escape from X inactivation of *Smcx* is preceeded by silencing during mouse development. *Nature Genet* **18**:212-213

LOCK LF, TAKAGI N MARTIN GR (1987) Methylation of the *Hprt* gene on the inactive X occurs after X chromosome inactivation. *Cell* **48**:39-46

LUPSKI JR, GARCIA CA, ZOGHBI HY, HOFFMAN EP, FENWICK RG (1991) Discordance of muscular dystrophy in monozygotic female twins: evidence supporting asymmetric splitting of the inner cell mass in a manifesting carrier of Duchenne dystrophy. *Am J Med Genet* **40**:354-364

LYON MF (1961) Gene action in the X-chromosome of the mouse (Mus musculus L.). *Nature* **190**:372-373

LYON MF (1972) X-chromosome inactivation and developmental progression in mammals. *Biol Rev* **47**:1-35

LYON MF (1998) X-chromosome inactivation: a repeat hypothesis. *Cytogenet Cell Genet* **80**:133-137

MACDONALD M, HASSOLD T, HARVEY J, WANG LH, MORTON NE, JACOBS

PA (1994) The origin of 47,XXY and 47,XXX aneuploidy: heterogeneous mechanisms and role of aberrant recombination. *Hum Mol Genet* **3**:1365-1371

MANLY BFJ (1998) *Randomization, Bootstrap and Monte Carlo Methods in Biology.* Chapman and Hall, London.

MARAHRENS Y, PANNING B, DAUSMAN J, STRAUSS W, JAENISCH R (1997) *Xist* deficient mice are defective in dosage compensation but not in spermatogenesis. *Genes Dev* **11**:156-166

MARAHRENS Y, LORING J, JAENISCH R (1998) Role of the *Xist* gene in X chromosome choosing. *Cell* **92**:657-664

MARTIN GR, EPSTEIN CJ, TRAVIS B, TUCKER, YATZIV (1978) X-chromosome inactivation during differentiation of female teratocarcinoma stem cells in vitro. *Nature* **271**:329-333

MARTINEZ MJ, BINKERT F, SCHINZEL A, KOTZOT D (2001) No evidence of dup(7)(p11.2p13) in Silver-Russell syndrome. *Am J Med Genet* **99**:335-337

MATSUURA S, EPISKOPOU C, HAMVAS R, BROWN SDM (1996) *Xist* expression from an *Xist* YAC transgene carried on the mouse Y chromosome. *Hum Mol Genet* **5**:451-459

MATTEI MG, MATTEI JF, AYME S, GIRAUD F (1982) X;autosome translocations: Cytogenetic characteristics and their consequences. *Hum Genet* **61**:295-309

MAY RH, HUNTER M, MASSON M (1993) *NPSTAT v 3.8.* Dept of Psychology, University of Victoria, P.O. Box 3050, Victoria, British Columbia, Canada V8W 3P5

MCDONALD LE, PATERSON CA, KAY GF (1998) Bisulfite genomic sequencing-derived methylation profile of the *Xist* gene throughout early mouse development. *Genomics* **54**:379-386

MCKEE BD, HANDEL MA (1993) Sex chromosomes, recombination, and chromatin conformation. *Chromosoma* **102**:71-80

MCMAHON A, MONK M (1983) X-chromosome activity in female mouse embryos heterozygous for *Pgk1* and Searles translocation. *Genet Res Camb* **41**:69-83

MERGENTHALER S, EGGERMANN K, TOMIUK J, RANKE MB, WOLLMANN HA, EGGERMANN T (2000a) Exclusion of a disease relevant role of *PAX4* in the aetiology of Silver-Russell syndrome: screening for mutations and determination of imprinting status. *J Med Genet* **37**:E44

MERGENTHALER S, WOLLMANN HA, BURGER B, EGGERMANN K, KAISER P, RANKE MB, SCHWANITZ G, EGGERMANN T (2000b) Formation of uniparental disomy 7 delineated from new case and a UPD7 case after trisomy rescue. Presentation of own results and a review of the literature. *Ann Genet* **43**:15-21

MERGENTHALER S, HITCHINS MP, BLAGITKO-DORFS N, MONK D, WOLLMANN HA, RANKE MB, ROPERS HH, APOSTOLIDOU S, STANIER P, PREECE MA, EGGERMANN T, KALSCHEUER VM, MOORE GE (2001) Conflicting reports of imprinting status of human *GRB10* in developing brain: how reliable are somatic cell hybrids for predicting allelic origin of expression? *Am J Hum Genet* **68**:543-545

MERMOUD JE, COSTANZI C, PEHRSON JR, BROCKDORFF N (1999) Histone macroH2A1.2 relocates to the inactive X chromosome after initiation and propogation of X-inactivation. *J Cel Biol* **147**(7):1399-1408

MIGEON BR, MOSER HW, MOSER AB, AXELMAN J, SILLENCE D, NORUM RA (1981) Adrenoleukodystrophy: Evidence for X-linkage, inactivation, and selection favouring the mutant allele in heterozygous cells. *Proc Natl Acad Sci USA* **78**:5066-5070

MIGEON BR, SHAPIRO LJ, NORUM RA, MOHANDAS T, AXELMAN J, DABORA RL (1982) Differential expression of steroid sulphatase locus on active and inactive human X chromosome. *Nature* **299**:838-840

MIGEON BR, SCHMIDT M, AXELMAN J, CULLEN CR (1986) Complete reactivation of X chromosomes from human chorionic villi with a switch to early DNA replication. *Proc Natl Acad Sci USA* **83**:2182-2186

MIGEON BR (1993) The postulated X inactivation centre 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* **52**:431-432

MIGEON BR, CHOWDHURY AK, DUNSTON JA, MCINTOSH I (2001) Identification of *TSIX*, encoding an RNA antisense to human *XIST*, reveals differences from its murine counterpart: Implications for X inactivation. *Am J Hum Genet* **69**:951-960

MIGEON BR, LEE CH, CHOWDHURY AK, CARPENTER H (2002) Species differences in *TSIX/Tsix* reveal the roles of these genes in X-chromosome inactivation.

Am J Hum Genet **71**:286-293

MILLER AP, GUSTASHAW K, WOLFF DJ, RIDER SH, MONACO AP (1995) Three genes that escape X chromosome inactivation are clustered within a 6Mb contig and STS map in Xp11.21-11.22. *Hum Mol Genet* **4**:731-739

MILLER SA, DYKES DD, POLESKY HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* **16**:1215

MIYOSHI O, KONDOH T, TANEDA H, OTSUKA K, MATSUMOTO T, NIIKAWA N (1999) 47,XX,UPD(7)mat,+r(7)pat/46,XX,UPD(7)mat mosaicism in a girl with Silver-Russell syndrome (SRS): possible exclusion of the putative gene from a 7p13-q11 region. *J Med Genet* **36**:326-329

MOHANDAS T, CRANDALL BF, SPARKES RS, PASSAGE MB, SPARKES MC (1981) Late replication studies in a human X/13 translocation: correlation with autosomal gene expression. *Cytogenet Cell Genet* **29**:215-220

MOHANDAS T, SPARKES RS, SHAPIRO LJ (1982) Genetic evidence for the inactivation of a human autosomal locus attached to an inactive X chromosome. *Am J Hum Genet* **34**:811-817

MONK D, WAKELING EL, PROUD V, HITCHINS M, ABU-AMERO SN, STANIER P, PREECE MA, MOORE GE (2000) Duplication of 7p11.2-p13, including *GRB10*, in Silver-Russell syndrome. *Am J Hum Genet* **66**:36-46

MONK D, HITCHINS M, RUSSO S, PREECE M, STANIER P, MOORE GE (2001) No evidence for mosaicism in Silver-Russell syndrome. *J Med Genet* **38**:E11

MONK M, HARPER MI (1978) X chromosome activity in preimplantation embryos from XX and XO mothers. *J Embryol Exp Morphol* **46**:53-64

MONK M, HARPER MI (1979) Sequential XCI coupled with cellular differentiation in early mouse embryos. *Nature* **281**:311-313

MONK M, MCLAREN A (1981) X chromosome activity in foetal germ cells of the mouse. *J Embryol Exp Morphol* **63**:75-84

MUSCATELLI F, LENA D, METTEI MG, FONTES M (1992) A male with two contiguous inactivation centers on a single X chromosome: study of X inactivation and XIST expression. *Hum Mol Genet* **1**:115-9

MUTTER GL, BOYNTON KA (1995) PCR bias in amplification of androgen receptor alleles, a trinucleotide repeat marker used in clonality studies. *Nucleic Acids Res*

23:1411-1418

NAN X, CAMPOY J, BIRD A (1997) MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* **88**:471-481

NAN X, HG H, JOHNSON CA, LAHERTY CD, TURNER BM, EISENMAN RN, BIRD A (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393**:386-389

NANCE WE (1964) Genetic tests with a sex-linked marker, glucose-6-phosphate dehydrogenase. *Cold Spring Harbor Syp Quant Biol* **29**:415-424

NANCE WE (1990) Do twin Lyons have larger spots? *Am J Hum Genet* **46**:646-648

NAUMOVA AK, PLENGE RM, BIRD LM, LEPPERT M, MORGAN K, WILLARD HF, SAPIENZA C (1996) Heritability of X-chromosome inactivation phenotype in a large family. *Am J Hum Genet* **58**:1111-1119

NAUMOVA AK, OLIEN L, BIRD LM, SMITH M, VERNER AE, LEPPERT M, MORGAN K, SAPIENZA C (1998) Genetic mapping of X-linked loci involved in skewing of X chromosome inactivation in the human. *Eur J Hum Genet* **6**:552-562

NORRIS DP, PATEL D, KAY GF, PENNY GD, BROCKDORFF HN (1994) Evidence that random and imprinted *Xist* expression is controlled by pre-emptive methylation. *Cell* **77**:41-51

OHNO S, KAPLAN WD, KINOSITA R (1959) Formation of the sex chromatin by a single X chromosome in liver cells of *Rattus norvegicus*. *Exp Cell Res* **18**:415-418

ORSTAVIK KH, ORSTAVIK RE, EIKLID K, TRANEBJAERG L (1996) Inheritance of skewed X-chromosome inactivation in a large family with an X-linked recessive deafness syndrome. *Am J Med Genet* **64**:31-34

ORSTAVIK KH, KRISTIANSEN M, HAGEN-LARSEN H, SKYTTHE A, BATHUM L, VAUPEL JW, CHRISTENSEN K (1999) Evidence of a genetic influence on the skewed X chromosome inactivation in elderly females (Abstract). *Eur J Hum Genet* **7**:P170

PAN Y, MCCASKILL CD, THOMPSON KH, HICKS J, CASEY B, SHAFFER LG, CRAIGEN WJ (1998) Paternal isodisomy of chromosome 7 associated with complete situs inversus and immotile cilia. *Am J Hum Genet* **62**:1551-1555

PANNING B, DAUSMAN J, JAENISCH R (1997) X chromosome inactivation is mediated by *Xist* RNA stabilization. *Cell* **90**:907-916

PAPANICOLAOU GN (1963) Memorandum on staining. In *The Atlas of Exfoliative Cytology* pp12-13. Harvard University Press, Cambridge, USA

PEGORARO E, SCHIMKE RN, ARAHATA K, HAYASHI Y, STERN H, ARKS H, GLASBERG MR, CARROLL JE, TABER JW, WESSEL HB, BAUSERMANN SC, MARKS WA, TORIELLO HV, HIGGINS JV, APPLETON S, SCHWARTZ S, GARCIA CA, HOFFMAN EP (1994) Detection of new paternal dystrophin gene mutations in isolated cases of dystrophinopathy in females. *Am J Hum Genet* **54**:989-1003

PEGORARO E, WHITAKER J, MOWERY-RUSHTON P, SURTI U, LANASA M, HOFFMAN EP (1997) Familial skewed X-inactivation: A molecular trait associated with high spontaneous-abortion rate maps to Xq28. *Am J Hum Genet* **61**:160-170

PEHRSON JR, FRIED VA (1992) MacroH2A, a core histone containing a large nonhistone region. *Science* **257**:1398-1400

PELADEAU N (1996) *Simstat v.2.01.* Provalis Research, 2414 Bennett Street, Quebec, Canada, H1V 3S4

PENAHERRERA MS, BARRETT IJ, BROWN CJ, LANGLOIS S, YONG SL, LEWIS S, BRUYERE H, HOWARD-PEEBLES PN, KALOUSEK DK, ROBINSON WP (2000) An association between skewed X-chromosome inactivation and abnormal outcome in mosaic trisomy 16 confined predominantly to the placenta. *Clin Genet* **58**:436-446

PENNY GD, KAY GF, SHEARDOWN SA, RASTAN S, BROCKDORFF N (1996) Requirement for *Xist* in X chromosome inactivation. *Nature* **379**:131-137

PLENGE RM, HENDRICH BD, SCHWARTZ C, FERNANDO ARENA J, NAUMOVA A, SAPIENZA C, WINTER RM, WILLARD HF (1997) A promoter mutation in the *XIST* gene in two unrelated families with skewed X-chromosome inactivation. *Nature Genet* **17**:353-356

PREECE MA, PRICE SM, DAVIES V, CLOUGH L, STANIER P, TREMBATH RC, MOORE GE (1997) Maternal uniparental disomy 7 in Silver-Russell syndrome. *J Med Genet* **34**:6-9

PREIS W, BARBI G, LIPTAY S, KENNERKNECHT I, SCHWEMMLE S, POHLANDT F (1996) X/autosome translocation in three generations ascertained through an infant with trisomy 16p due to failure of spreading of X-inactivation. *Am J*

Med Genet **61**:117-121

QUAN F, JANAS J, TOTH-FEJEL SE, JOHNSON DB, WOLFORD JK (1997) Uniparental disomy of the entire X chromosome in a female with Duchenne muscular dystrophy. *Am J Hum Genet* **60**:160-165

RACK KA, CHELLY J, GIBBONS RJ, RIDER S, BENJAMIN D, LAFRENIERE RG, OSCIER D, HENDRIKS RW, CRAIG IW, WILLARD HF, MONACO AP, BUCKLE VJ (1994) Absence of the XIST gene from late-replicating isodicentric X chromosomes in leukaemia. *Hum Mol Genet* 3:1053-1059

RASMUSSEN TP, MASTRANGELO MA, EDEN A, PEHRSON JR, JAENISCH R (2000) Dynamic relocalization of histone macroH2A1 from centrosomes to inactive X chromosomes during X inactivation. *Journal Cell Biol* **150**:1189-1198.

RASTAN S (1982) Primary non-random X-inactivation caused by controlling elements in the mouse demonstrated at the cellular level. *Genet Res* **40**:139-147

RAY PF, WINSTON RML, HANDYSIDE AH (1997) *XIST* expression from the maternal X chromosome in human male preimplantation embryos at the blastocyst stage. *Hum Mol Genet* **6**:1323-1327

REDDY BK, ANANDAVALLI TE, REDDI OS (1984) X-linked Duchenne muscular dystrophy in an unusual family with manifesting carriers. *Hum Genet* **67**:460-462

REIK W, BROWN KW, SLATTER RE, SARTORI P, ELLIOTT M, MAHER ER (1994) Allelic methylation of H19 and IGF2 in the Beckwith-Wiedemann syndrome. *Hum Mol Genet* **3**:1297-1301

REYAL F, GIGAREL N, RABIER D, GILBERT B, BARDET J, COLLEAUX L, SAUDUBRAY JM, AMIAL J, MUNNICH A, BONNEFONT JP (2000) Relationship between clinical phenotype, residual OTC activity and pattern of liver X-inactivation in symptomatic OTC-deficient females. *Am J Hum Genet* (abstract) **67**:A1601

RICHARDS CS, WATKINS SC, HOFFMANN EP, SCHNEIDER NR, MILSARK IW, KATZ KS, COOK JD, KUNKEL LM, CORTADA JM (1990) Skewed X inactivation in a female MZ twin results in Duchenne muscular dystrophy. *Am J Hum Genet* **46**:672-681

RIESEWIJK AM, BLAGITKO N, SCHINZEL AA, HU L, SCHULZ U, HAMEL BC, ROPERS HH, KALSCHEUER VM (1998) Evidence against a major role of *Peg1/MEST* in Silver-Russell syndrome. *Eur J Hum Genet* **6**:114-120

RISCH N, STEIN Z, KLINE J, WARBURTON D (1986) The relationship between maternal age and chromosome size in autosomal trisomy. *Am J Hum Genet* **39**:68-78

ROBINSON WP, BARRETT IJ, BERNARD L, TELENIUS A, BERNASCONI F, WILSON RD, BEST RG, HOWARD-PEEBLES PN, LANGLOIS S, KALOUSEK DK (1997) Meiotic origin of trisomy in confined placental mosaicism is correlated with presence of fetal uniparental disomy, high levels of trisomy in trophoblast, and increased risk of fetal intrauterine growth restriction. *Am J Hum Genet* **60**:917-927

ROBINSON WP, CHRISTIAN SL, KUCHINKA BD, PENAHERRERA MS, DAS S, SCHUFFENHAUER S, MALCOLM S, SCHINZEL AA, HASSOLD TJ, LEDBETTER DH (2000) Somatic segregation errors predominantly contribute to the gain or loss of a paternal chromosome leading to uniparental disomy for chromosome 15. *Clin Genet* **57**:349-358

ROPERS HH, WIENKER TF, GRIMM T, SCHROETTER K, BENDER K (1977) Evidence for preferential X-chromosome inactivation in a family with Fabry disease. *Am J Hum Genet* **29**:361-370

ROUSSEAU F, HEITZ D, OBERLE I, MANDEL JL (1991) Selection in blood cells from female carriers of the fragile X syndrome: inverse correlation between age and proportion of active X-chromosomes carrying the full mutation. *J Med Genet* **28**:830-836

RUSSELL LB (1963) Mammalian X-chromosome action: inactivation limited in spread and region of origin. *Science* **140**:976-978

SADO T, WANG Z, SASAKI H, LI E (2001) Regulation of imprinted X-chromosome inactivation in mice by *Tsix*. *Development* **128**:1275-1286

SANGHA KK, STEPHENSON MD, BROWN CJ, ROBINSON WP (1999) Extremely skewed X-inactivation is increased in women with recurrent spontaneous abortion. *Am J Hum Genet* **65**:913-917

SCHANZ S, STEINBACH P (1989) Investigation of the "variable spreading" of X inactivation into a translocated autosome. *Hum Genet* **82**:244-248

SCHINZEL AAGL (1990) Chromosome 7, trisomy 7q2-3; in BUYSE ML (ed): *Birth Defects Encyclopaedia*, Vol 1 pp 349-350. Cambridge, MA

SCHMIDT M, DUSART D (1992) Functional disomies of the X-chromosome influence the cell selection and hence the X-inactivation patterns in females with

balanced X;autosome translocations: A review of 122 cases. *Am J Med Genet* **42**:161-169

SHAO C, TAKAGI N (1990) An extra maternally derived X chromosome is deleterious to early mouse development. *Development* **110**:969-975

SHEARDOWN SA, DUTHIE SM, JOHNSTONE CM, NEWALL AET, FORMSTONE EJ, ARKELL RM, NESTEROVA TB, ALGHISI GC, RASTAN S, BROCKDORFF N (1997) Stabilisation of *Xist* RNA mediates initiation of X chromosome inactivation. *Cell* **91**:99-107

SIMMLER MC, CATTANACH BM, RASBERRY C, ROUGUELLE C (1993) Mapping the murine *Xce* locus with $(CA)_n$ repeats. *Mamm Genome* **4**:523-530

SINGER-SAM J, CHAPMAN V, LEBON JM, RIGGS AD (1992) Parental imprinting studied by allele specific primer extension after PCR: paternal X-chromosome linked genes are transcribed prior to preferential paternal X chromosome inactivation. *Proc Natl Acad Sci USA* **89**:10469-10473

SMIT AF, TOTH G, RIGGS AD, JURKA J (1995) Ancestral, mammalian-wide subfamilies of LINE-1 repetitive sequences. *J Mol Biol* **246**:401-417

SMIT AF (1999) Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Curr Opin Genet Dev* **9**:657-663

SPENCE JE, PERCIACCANTE RG, GREIG GM, WILLARD HF, LEDBETTER DH, HEJTMANCIK JF, POLLACK MS, O'BRIEN WE, BEAUDET AL (1988) Uniparental disomy as a mechanism for human genetic disease. *Am J Hum Genet* **42**:217-226

SPRENT P (1995) *Applied Nonparametric Statistical Methods.* Chapman and Hall, London.

SURRALLES J, NATARAJAN AT (1998) Position effect of translocations involving the inactive X chromosome: physical linkage to *XIC/XIST* does not lead to long-range *de novo* inactivation in human differentiated cells. *Cytogenet Cell Genet* **82**:58-66

TADA T, TAKAGI N, ADLER ID (1993) Parental imprinting on the mouse X chromosome: effects on the early development of XO, XXY, and XXX embryos. *Genet Res* **62**:139-148

TAKAGI N, SASAKI M (1975) Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature* **256**:640-642

TAKAGI N, SUGAWARA O, SASAKI M (1982) Regional and temporal changes in the pattern of X-chromosome replication during the early post-implantation development of the female mouse. *Chromosoma* **85**:275-286

TAKAGI N, ABE K (1990) Detrimental effects of two active X chromosomes on early mouse development. *Development* **109**:189-201

TAN SS, WILLIAMS EA, TAM PPL (1993) X-chromosome inactivation occurs at different times in different tissues of the post-implantation mouse. *Nature Genet* **3**:170-174

TAYLOR SA, DEUGAU KV, LILLICRAP DP (1991) Somatic mosaicism and female-to-female transmission in a kindred with Haemophilia B (Factor IX deficiency). *Proc Natl Acad Sci USA* **88**:39-42

TAYSI K, SPARKES RS, O'BRIEN TJ, DENGLER DR (1982) Down's syndrome phenotype and autosomal gene inactivation in a child with presumed (X;21) *de novo* translocation. *J Med Genet* **19**:144-148

THERMAN E, PATAU K (1974) Abnormal X chromosomes in man: origin, behavior and effects. *Humangenetik* **25**:1-16

THERMAN E, SARTO GE, PATAU K (1974) Center for Barr body condensation on the proximal part of the human Xq: a hypothesis. *Chromosoma* **44**:361-366

TORCHIA BS, CAL LMI, MIGEON BR (1994) DNA replication analysis of *FMR1*, *XIST*, and Factor 8C loci by FISH shows non-transcribed X-linked genes replicate late. *Am J Hum Genet* **55**:96-104

TRIBIOLI C, TAMANINI F, PATROSSO C, MINALESI L, VILLA A (1992) Methylation and sequence analysis around *EagI* sites: identification of 28 new CpG islands in Xq24-q28. *Nucleic Acids Res* **20**:727-733

TSUJI K, COPELAND NG, JENKINS NA, OBINATA M (1995) Mammalian antioxidant protein complements alkylhydroperoxide reductase (ahpC) mutation in *Escherichia coli*. *Biochem J* **307**: 377-381

UEJIMA H, LEE MP, CUI H, FEINBERG AP (2000) Hot-stop PCR: a simple and general assay for linear quantification of allele ratios. *Nature Genet* **25**:375-376

VAN DER VLAG J, OTTE AP (1999) Transcriptional repression mediated by the human polycomb-group protein EED involves histone deacetylation. *Nature Genet* **23**:474-478

VOGELSTEIN B, FEARON ER, HAMILTON SR, PREISINGER AC, WILLARD HF, MICHELSON AM, RIGGS AD, ORKIN SH (1987) Clonal analysis using recombinant DNA probes from the X-chromosome. *Cancer Res* **47**:4806-4813

WAKEFIELD MJ, KEOHANE AM, TURNER BM, GRAVES JAM (1997) Histone underacetylation is an ancient component of mammalian X-chromosome inactivation. *Proc Natl Acad Sci USA* **94**:9665-9668

WAKELING EL, ABU-AMERO S, PRICE SM, STANIER P, TREMBATH RC, MOORE GE, PREECE MA (1998) Genetics of Silver-Russell syndrome. *Horm Res* **49**:32-36

WALKER CL, CARGILE CB, FLOY KM, DELANNOY M, MIGEON BR (1991) The Barr body is a looped X-chromosome formed by telomere association. *Proc Natl Acad Sci USA* **88**:6191-6195

WANG J, MAGER J, CHEN Y, SCHNEIDER E, CROSS JC, NAGY A, MAGNUSON T (2001) Imprinted X inactivation maintained by a mouse *Polycomb* group protein. *Nature Genet* **28**:371-375

WARSHAWSKY D, STAVROPOULOS N, LEE JT (1999) Further examination of the *Xist* promoter-switch hypothesis in X inactivation: Evidence against the existence and function of a P_0 promoter. *Proc Natl Acad Sci USA* **96**(25):14424-14429

WATKISS E, WEBB T, BUNDEY S (1993) Is skewed X inactivation responsible for symptoms in female carriers of adrenoleukodystrophy? *J Med Genet* **30**:651-654

WEI X, SAMARABANDU J, DEVDHAR RS, SIEGEL AJ, ACHARYA R, BEREZNEY R (1998) Segregation of transcription and replication sites into higher order domains. *Science* **281**:1502-1505

WENGLER G, GORLIN JB, WILLIAMSON JM, ROSEN FS, BING DH (1995) Non-random inactivation of the X chromosome in early lineage haematopoietic cells in carriers of Wiskott-Aldrich syndrome. *Blood* **85**:2471-2477

WHITE WM, WILLARD HF, VAN DYKE DL, WOLFF DJ (1998) The spreading of X inactivation into autosomal material of an X;autosome translocation: Evidence for a difference between autosomal and X-chromosomal DNA. *Am J Hum Genet* **63**:20-28

WILLARD HF, LATT SA (1976) Analysis of deoxyribonucleic acid replication in human X chromosomes by fluorescence microscopy. *Am J Hum Genet* **28**:213-227

WILLARD HF, SALZ HK (1997) Remodelling chromatin with RNA. *Nature*

386:228-229

WILLIAMS J, DEAR PRF (1987) An unbalanced t(X;10) mat translocation in a child with congenital abnormalities. *J Med Genet* **24**:633

WINCHESTER B, YOUNG E, GEDDES S, GENET S, HURST J, MIDDLETON-PRICE H, WILLIAMS N, WEBB M, HABEL A, MALCOLM S (1992) Female twin with Hunter disease due to nonrandom inactivation of the X-chromosome: a consequence of twinning. *Am J Med Genet* **44**:834-838

WOLFFE AP (1997) Transcription control: repressed repeats express themselves. *Current Biology* **7**:R796-R798

WOLLMANN HA, KIRCHNER T, ENDERS H, PREECE MA (1995) Growth and symptoms in Silver-Russell syndrome: review on the basis of 386 patients. *Eur J Paediatr* **154**:958-968

YANG TP, CASKEY T (1987) Nuclease sensitivity of the mouse *Hprt* gene promotor region: Differential sensitivity on the active and inactive X chromosomes. *Mol Cell Biol* **7**:92-94

ZUCCOTTI M, MONK M (1995) Methylation of the mouse *Xist* gene in sperm and eggs correlates with imprinted *Xist* expression and paternal X-inactivation. *Nature Genet* **9**:316-320