Université de Montréal

GENETIC-EPIDEMIOLOGIC ANALYSIS OF X-INACTIVATION SKEWING IN HUMAN FEMALES: SUGGESTIVE EVIDENCE FOR TWO DISTINCT TRAITS

par

Robert Mio

Département de médecine

Faculté de médecine

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de
Philosophiae Doctor (Ph.D.)
en sciences biomédicales

novembre 2003

Faculté des Études Supérieures
Grade octroyé
à compter du
0 3 FEV 2005
Université de Montréal

© Robert Mio, 2003

W
4
U58
2005
V.012

Université de Montréal

Direction des bibliothèques

AVIS

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal
Faculté des études supérieures

Cette thèse intitulée:

GENETIC-EPIDEMIOLOGIC ANALYSIS OF X-INACTIVATION SKEWING IN HUMAN FEMALES: SUGGESTIVE EVIDENCE FOR TWO DISTINCT TRAITS

présenté par:
Robert Mio

A été évalué par un jury composé des personnes suivantes:

Présidente-rapporteur	Manjunath Puttaswamy, PhD
Directeur de recherche	Lambert Busque, MD (FRCPC)
Membre du jury	Eric Milot, PhD
Examinateur externe	Anna Naumova, PhD (McGill University)
	Gary Gilliland, MD PhD (Harvard University)
Représentante du doyen	Dennis Claude Roy, MD (FRCPC)

Thèse acceptée le

SUMMARY

Introduction: X inactivation (Xi) randomly inactivates a single X chromosome (maternal or paternal) in female somatic cells. In a significant proportion of females however, Xi is not random, a phenomenon termed skewed Xi [arbitrarily defined as preferential inactivation (≥75%) of the maternal (Xm) or paternal X (Xp)]. The proportion of Xp inactivated relative to Xm is termed the X inactivation ratio (XIR). The study of Xi in humans is hampered by two unrelated phenotypes. i) The *primary Xi* trait. Initiated during early embryogenesis, in a developmental context, it induces a similar Xi pattern, which can vary from random to skewed, among various tissues. A resulting skewed Xi pattern, termed primary skewing (PS), may result from a small number of stem cells present when Xi is initiated However, other possibilities include genetic influences, such as heterozygosity for the X-linked *Xce* locus, as observed in certain mice hybrids. That 9% of human female neonates demonstrate a skewed Xi pattern in cord blood supports a PS trait in humans. ii) the *secondary skewing* trait, usually associated with a skewed Xi pattern in a tissue-specific manner, occurs after the initiation of Xi. Secondary skewing often results from a growth competition between X-linked alleles, such as in female carriers of various X-linked immunodeficiency disease alleles. However, X-linked disease alleles are rare and do not explain the high prevalence of skewing (38%) as observed in peripheral blood (PB) of 'healthy' females 60 years of age and older. The latter trait, termed acquired skewing (AS), has been assigned various etiologies. Recent data support an X-linked genetic component influencing hematopoietic stem cell (HSC) growth/survival kinetics. Clinically, skewed Xi has been associated with various biomarkers (breast cancer and recurrent spontaneous abortion for example). In light of this data, a study was undertaken to resolve the etiologies and biological / clinical associations of Xi traits in human females. Methods: French-Canadian nuclear families (females only) were recruited for study analysis. Two biological tissues were obtained: PB for analysis of the AS trait and buccal cells (BC) for analysis of the primary Xi trait. PB was fractionated and cell-sorted to obtain pure-cell populations. The XIR was determined by the HUMARA clonality assay. Xi phenotypes included. i) The XIR derived from BC for analysis of the primary Xi trait. ii) The XIR of PB for analysis of the AS trait and iii) the relative value of AS obtained by quantitating deviation from the primary XIR (i.e., difference between the BC and PB XIR). Qualitative analyses included a skewed Xi pattern (≥25% deviation from random Xi for the XIR and an AS value ≥0,25 for the relative AS trait). Subject data included a medical questionnaire and blood counts, which were analyzed as a function of Xi phenotypes. Genetic effects were determined by heritability studies. Results and Conclusion: 1144 females derived from 193 nuclear families were recruited. Age ranged from 38-96 years, with a mean of 63,3 years. 90,8% of females were informative for the HUMARA assay. Findings related to the BC XIR were four-fold: i) the incidence of Xi skewing in BC was low (12,4% - similar to that reported in

neonatal cord blood, i.e. 8,6%); ii) the incidence of Xi skewing was relatively stable with advancing age ($p=0{,}21$); iii) the BC XIR was modestly correlated with that of PB lineages ($0{,}46<r<0{,}56$) and iv) heritability analysis revealed a genetic (plausibly X-linked) component ($p<0{,}0001$; $h^2=0{,}30$). These findings are similar to the Xce-influenced primary Xi trait in mice, supporting evidence for an XCE-like primary Xi trait in humans. Further, this finding suggests that the primary Xi trait in humans is not strictly a stochastic process as previously suggested. Clinically, the BC and blood XIR was associated with asthma. Among hematopoietic lineages, the incidence of skewed Xi was higher versus BC: granulocytes (PMN) 36%; monocytes 36,6%; T cells 20,1% and B cells 26,5%. The incidence of relative AS was 22,7; 27,2; 11,4 and 16,3%, respectively. With the exception of T cells, significant correlation of AS values among B cells and myeloid lineages ($0{,}73<r<0{,}85$) was consistent with a HSC origin of AS. The incidence of AS (relative and absolute) increased significantly with age, particularly for myeloid lineages. Lack of T cells contribution may be attributed to long-lived memory T cells. Heritability estimates attributed 20-39% of the variance of AS values to genetic effects (plausibly X-linked), supporting linkage studies to map the trait(s). Clinically, increasing AS values were associated with a decreased eosinophil count. Because eosinophil count is a predictor of all-cause mortality, the finding suggests AS may be associated with increased longevity. In effect, that the AS trait demonstrated a different biological profile versus the PS trait (i.e., late versus early-onset, increased incidence of skewing with age versus stable, different incidence of skewing, and a different clinical profile) is convincing evidence for two distinct traits.

Key words: X chromosome, acquired skewing, primary skewing, HUMARA, heritability, hematopoietic stem cells, buccal cells, human, females, Québec, hematopoietic lineages, association studies, clinical data, complete blood counts, age

RÉSUMÉ

Introduction : Dans les cellules somatiques de femmes, l'inactivation du chromosome X (iX) inactive, de façon aléatoire, l'un des deux chromosome X (maternel ou paternel). Cependant, dans une certaine proportion de femmes, l'inactivation n'est pas aléatoire, un phénomène appelé iX biaisé [arbitrairement défini comme étant une inactivation préférentielle (≥75%) du X maternel (Xm) ou paternel (Xp)]. La proportion de Xp inactivé relatif au Xm est appelé le ratio d'inactivation du X (RIX). Chez l'humain, l'étude de iX est entravée par deux phénotypes non reliés: i) Le *iX primaire*. Amorcé tôt durant l'embryogenèse, dans un contexte de développement donné, induit un RIX similaire dans divers tissus, soit aléatoire ou biaisé, ce dernier est appelé le biaisé primaire (BP). Le BP peut provenir d'un petit nombre de cellules souches présentent lors de l'initiation du iX. D'autres causes incluant les influences génétiques, comme l'hétérozygocité du locus *Xce* lié au X, observée dans certaines hybrides de souris. Le fait que chez 9% des nouveau-nés de sexe féminin on observe l'iX biaisé dans le sang du cordon suggère un phénotype BP chez l'humain. ii) Le phénotype *biaisé secondaire*, habituellement associé avec l'iX biaisé dans un tissu spécifique, a lieu après l'initiation de iX. Le biais est le résultat d'un désavantage de croissance conféré par des allèles mutant liés au X. Les exemples incluent des porteuses de différents allèles de maladies immunodéficientes liées au X. Malgré cela, ces allèles sont rares et n'expliquent pas la haute fréquence (38%) de ce biais observé dans le sang périphérique (SP) chez des femmes âgées de 60 (au plus) et en santé. Ce phénotype, appelé biais acquis (BA), a été associé à diverses causes. De récentes données indiquent une composante génétique liée au X qui influence la cinétique des cellules souches hématopoïétiques (CSH). Cliniquement, le phénotype iX biaisé a été associé à différents marqueurs biologiques comme le cancer du sein et l'avortement spontané répétitif. À la lumière de ces données, une étude a été entreprise pour élucider les causes et les associations biologiques des deux phénotypes iX chez les femmes. Méthodes: Des familles Canadiennes Françaises (femmes seulement) ont été recrutées pour participer à l'étude. Deux tissus biologiques ont été recueillis pour analyse : le sang périphérique (SP) pour l'analyse du phénotype BA et des cellules buccales (CB) pour l'analyse du phénotype iX primaire. Le SP a été fractionné, et ses cellules classées pour obtenir des populations de cellules pures. Le RIX a été déterminé par la méthode de HUMARA. Les phénotypes de iX inclus sont: i) Le RIX obtenu des CB, représentatif du phénotype iX primaire. ii) La valeur absolue du RIX pour l'analyze du phénotype BA et iii) la valeur BA relatif, calculé selon la différence entre les RIX primaires des CB et du SP. Pour les analyses qualitatives, le RIX biaisé (≥25%) et la valeur BA relatif (≥0,25) ont été utilisé. Les données cliniques incluant un questionnaire médicale et les analyse sanguines, ont été analysées en relation avec le RIX. Une composante génétique/familiale des phénotypes a été évaluée par une

étude familiale. Résultats et conclusion : Parmi 193 familles, 1144 femmes ont été recrutées. Ces dernières étaient âgées entre 38 et 96 ans, pour une moyenne de 63,3 ans. Les résultats obtenus pour le phénotype iX primaire sont de quatre ordres : i) la fréquence du iX biaisé dans les CB était faible (12,4%-similaire au résultat obtenu dans les cordons de nouveau-nés (8,6%)), ii) le RIX des CB était relativement stable avec l'âge ($p=0.21$), iii) le RIX des CB corrélait signifitivement avec celui des types cellulaires du SP ($0,46<r<0,56$) et iv) l'héritabilité des RIX des CB suggère une composante génétique ($p<0,0001$; $h^2=0,30$). Ces propriétés sont en lien direct avec un phénotype iX primaire, possiblement en relation avec un locus *Xce* lié au X. Ceci suggère que le iX primaire chez l'humain n'est pas un processus strictement stochastique comme suggéré précédemment. Cliniquement, les RIX ont été associé à une augmentation des cas d'asthme. Parmi les types hématopoïétiques, l'incidence du iX biaisé était plus élevée par rapport au CB: granulocytes (PMN) 36%, monocytes 36,6%; lymphocytes T 20,1% et lymphocytes B 26,5%. En utilisant la valeur BA relatif, l'incidence du BA (valeur $\geq 0,25$) était 22,7; 27,2; 11,4 et 16%, respectivement. À l'exception des cellules T, chez un individu les RIX entre les différentes cellules hématopoïétiques corrèlent d'un façon significative ($0,73<r<0,85$), ce qui supporte une d'origine CSH. L'incidence de BA augmentent significativement avec l'âge, particulièrement pour les types myéloïdes. Une incidence du BA moins fréquente chez les lymphocytes T pourrait être attribuée à une plus grande longévité. Les études d'héritabilité ont attribué 20 à 39% de la variance des valeurs de BA aux effets génétiques, ce qui supportent les études de liaisons génétiques pour localiser les phénotypes. Cliniquement, l'augmentation des valeurs de BA était associée à une diminution du nombre d'éosinophiles. Puisque le compte d'éosinophiles est associé à la mortalité, une association négative suggère que le BA est associé à la longévité. En conséquence, le fait que le phénotype BA ait démontré un profil biologique différent de celui du phénotype BP (i.e. : l'apparition acquise versus primaire, l'incidence de iX biaisé differentes et les profils cliniques differentes) suggère fortement deux phénotypes distincts.

Mots clés: chromosome X, biaisé acquise, biaisé primaire, HUMARA, héritabilité, cellules souches hématopoïétiques, cellules buccales, humain, femme, Québec, lignées hématopoïétiques, études d'association, donnée clinique, analyse sanguines, l'âge

TABLE OF CONTENTS

SUMMARY iii

RÉSUMÉ v

TABLE OF CONTENTS vii

LIST OF FIGURES xv

LIST OF TABLES xvi

LIST OF ABBREVIATIONS xviii

ACKNOWLEDGEMENTS xxi

Chapter I 1

INTRODUCTION AND LITERATURE REVIEW 1

1 THE X CHROMOSOME 2

1.1 Basic principles 2

1.2 Lyon's Hypothesis 2

1.3 Anatomical human X-mosaicism 3

1.4 Timing of X inactivation: *concurrent with tissue differentiation* 4

2 MECHANISM OF X INACTIVATION 5

2.1 The X inactivation center (Xic/XIC) 5

Abbreviations: X^m, maternal X; X^p, paternal X; a, active; i, inactive 6

2.2 Genetic elements of the Xic: *Xist* gene 9

2.2.1 Developmental regulation of Xist 10

2.3 Characteristics acquired with *Xist* upregulation 10

Methylation of CpG islands of Xi-linked genes 10

Asynchronous replication of the inactive X 10

Covalent modifications of X chromosome associated histones 12

Enrichment of histone variants 13

Enrichment of mRNA species 13

2.4 Stability of X inactivation 13

3 REGULATORY ELEMENTS OF THE XIC 13

3.1 Further characterization of the *Tsix* gene 14

4 IMPRINTED X INACTIVATION 15

4.1 Preferential inactivation of the paternal X in extraembryonic tissues 15

4.2 X-imprinting in humans 16

5 ESCAPE FROM X INACTIVATION 16

6 CLINICAL APPLICATION OF THE LYON'S HYPOTHESIS 17

6.1 X inactivation-based clonality assays 17

6.2 Principles of X-inactivation based clonality assays 17
6.2.1 Distinguish Xp from Xm 18
6.2.2 Distinguish active X (aX) from inactive X (iX) 18
6.3 Utility of Xi based clonality assays 18
6.3.1 HUMARA assay 18
6.4 Hematological application of clonality assays 27
6.4.1 Clonal origin of hematologic malignancies 27
6.4.2 Stem cell origin of hematologic malignancies 27
6.4.3 Carrier detection of X-linked disorders 27
6.5 Limitation of X-inactivation based clonality assay 28
7 NONRANDOM X-INACTIVATION 28
7.1 Binomial distribution of X-inactivation ratios 28
7.2 Skewed X-inactivation as a discrete trait 28
7.3 Etiologies of skewed X inactivation 29
7.3.1 Etiology of primary skewing: Stochastic event 31
7.3.2 Etiology of primary skewing: a heritable trait 31
7.3.2.1 Xce allele heterozygosity: a skewed Xi pattern in multiple tissues 32
7.3.2.2 XIST/TSIX aberrations as a cause of primary skewing 33
7.3.3 Etiology of Secondary Skewing: selection against X-linked disease allele(s) 33
7.3.3.1 Suggestive evidence that secondary skewing may be incompletely penetrant 34
7.3.3.2 Nonhematopoietic tissues are generally robust to secondary skewing 35
7.4 Familial clustering of skewed Xi: Primary or secondary etiology? 35
7.5 Randomly ascertained families to identify the etiology of skewed Xi 37
8 INCIDENCE OF SKEWED Xi IN THE GENERAL FEMALE POPULATION: *variable* 37
8.1 Higher incidence of skewed Xi in peripheral blood versus NHTs 38
8.2 Correlation of XIRs among various tissues 38
8.2.1 Correlation of XIRs among blood cell lineages: *strong* 38
8.2.2 Correlation of XIRs among NHTs: *moderate, evidence for a body-wide XIR* 39
8.2.3 Correlation of XIRs between blood cells and NHTs: *low* 39
Concordance of Xi patterns between blood and NHTs: evidence of a body-wide Xi pattern 41
8.3 Suggestive evidence for an *XCE*-like X chromosome skewing trait 41
9 ANALYSIS OF BLOOD Xi PATTERNS AS A FUNCTION OF AGE 41
9.1 AS: *confirmation studies* 42
9.2 Clarifications in light of the AS phenomenon 43
10 ETIOLOGY OF SKEWED Xi IN CORD BLOOD: *primary skewing trait* 43
11 PLAUSIBLE ETIOLOGIES OF ACQUIRED SKEWING 44

11.1 Stochastic event: *clonal dominance secondary to stem cell depletion* 44
11.2 Stochastic event: *clonal "pre-neoplasia" hematopoiesis* 45
11.3 Continuation of the PS trait 46
11.4 Genetic: *X-linked hemizygous cell selection* 46
12 X-LINKED HEMIZYGOUS CELL SELECTION ETIOLOGY OF AS 46
12.1 AS in felines 46
12.1.1 Hematopoietic stem cell origin of AS 47
12.2 X-linked genetic basis of the AS trait extended to humans 48
12.3 Twin studies support an X-linked genetic basis to AS in humans 48
12.3.1 Correlation of XIRs in elderly MZ twin pairs 48
12.3.2 Genetic contribution to the AS trait 50
12.4 Intrinsic versus extrinsic factors in the etiology of AS 50
12.5 Molecular characteristics of the AS trait: *Genetic model* 51
13 CLINICAL ASSOCIATIONS OF SKEWED Xi 52
13.1 Skewing and expression of X-linked disease allele(s) 52
13.2 Skewing and RSA 53
13.3 Skewing and susceptibility to ovarian/breast cancer 53
13.4 Skewing and predisposition to autoimmunity 54
13.5 Skewing and longevity 55
13.6 Therapeutic applications 55
14 HYPOTHESIS and GOALS 55
15 STRATEGIES, RATIONALE AND EXPERIMENTAL DESIGN 56
15.1 Recruitment of elderly females 57
15.2 Genetic analyses: sib-pair approach to dissect a complex genetic trait 57
15.3 Collection of families 57
15.3.1 Benefit of a relatively isolated population 57
15.4 Sample size (λ estimation): power studies 58
15.5 Definition of the phenotype 59
Chapter II 60
MATERIALS AND METHODS 60
2.1 Subjects 61
2.2 Medical Questionnaire 61
2.2.1 Age 61
2.2.2 Smoking habits: environmental stimulus 61
2.2.3 Family data: skewing and X-linked mutant allele(s) 62
2.2.4 Clinical data 62

2.3 Biological samples 63
2.3.1 Primary Xi pattern / PS trait 63
2.3.2 AS trait: peripheral blood 63
2.4 Complete blood counts 64
2.5 Cellular fractionation 64
2.6 Immunophenotyping and cell sorting 64
2.7 DNA isolation from blood cells 65
2.8 DNA isolation from buccal cells 65
2.9 HUMARA clonality assay 66
2.9.1 HUMARA allelic ladder 67
2.9.2 Quantitation of HUMARA alleles 67
2.9.3 Reliability of quantification: accuracy 68
2.10 Calculation of the XIR 68
2.10.1 DS score (degree of skewing): 68
2.10.2 PAmat score (Proportion Active of the Maternally inherited X chromosome): 68
2.10.3 AS_{PAmat} score: 69
2.10.4 AS_{DS} score 69
2.11 Discrete criteria to delineate a skewed Xi pattern 69
2.11.1 Skewing as a discrete trait 69
2.11.2 Acquired skewing as a discrete trait 70
2.12 Familial resemblance of Xi pattern traits 70
2.12.1 Familial aggregation of skewed Xi: recurrence risk ratio (RRR) 70
2.12.2 Sibling correlation of XIRs by ANOVA 71
2.12.3 Heritability analyses 71
2.13 Intra-individual correlation of XIRs: 72
Body-wide Xi pattern: *correlation of XIRs between BC and hematopoietic lineages* 72
HSC origin of AS: *correlation of XIRs among hematopoietic lineages* 72
2.14 Statistical analyses 72
Chapter III 73
RESULTS (PART I) 73
3.1 Population data 74
3.2 Specimen data 74
3.3 Incidence of skewed Xi 74
3.3.1 Skewed X inactivation: DS score ≥ 0,25 74
BC skewing: primary skewed Xi pattern (PS) / body-wide skewing trait 74
Peripheral blood skewing: higher incidence of skewed Xi versus BC 75

3.3.2 Incidence of the AS trait: AS_{QL} and AS_{QT} analyses 75

Incidence of acquired skewing - qualitative analyses: AS_{QL} 75

Incidence of acquired skewing - quantitative analyses: AS_{QT} 75

3.4 Distribution and mean DS and AS_{DS} scores 76

DS score 76

AS_{DS} score 76

3.5 XIRs as a function of age 77

3.6 Evidence supporting a primary Xi (PS) trait: intraindividual correlation of XIRs between BC and hematopoietic lineages 78

3.6.1 Correlation of PAmat scores between BC and leukocytes 79

3.6.2 Concordance for a skewed Xi pattern among various tissues 79

3.7 Evidence for a HSC origin of AS: intraindividual correlation of XIRs among hematopoietic lineages 80

3.7.1 Correlation of PAmat scores among blood lineages: HSC origin 80

3.7.2 Correlation of AS_{PAmat} scores among blood lineages: HSC origin of AS 80

3.8 Suggestive evidence for distinct Xi-skewing traits: PS versus AS 81

DISCUSSION FOR PART I 82

Population and specimen data 82

Suggestive evidence of a primary Xi trait: Primary skewing trait 82

Further characterization of the AS trait 84

Hematopoietic lineages implicated in the AS trait 85

AS represents departure from the primary XIR 86

HSC origin of AS 86

Suggestive evidence that primary skewing and AS are distinct traits 87

Clinical implications 87

Interpretation of Xi patterns for myeloproliferative disorders 87

Control tissue for X-linked disorders: 88

Chapter IV 108

RESULTS (PART II) 108

4.1 Characteristics of the sample population 110

4.1.1 Hematopoietic indices 110

4.1.2 Clinical characteristics 111

4.1.2.1 Prevalence of clinical traits versus reference populations: 111

4.1.3 Smoking characteristics 112

4.1.4 Family / parity data 112

4.2 Confounding variables 112

4.2.1 Age as a confounding variable 112
Age-hematopoietic indices 113
Age-clinical data 113
Age-smoking habits 113
Age-parity 114
4.2.2 Clinical data as a confounding variable 114
4.2.2.1 Clinical data-blood counts 114
4.2.2.2 Clinical data – parity data analysis 116
4.2.3 Smoke as a confounding variable 117
4.2.3.1 Smoke-clinical data 117
4.2.3.2 Smoke-blood counts 119
4.2.3.3 Smoke-parity 119
4.2.4 Parity data as a confounding variable 120
Parity and blood counts 120
4.2.5 Covariates associated with blood counts: multivariate analyses 120
4.2.6 Maximum likelihood heritability estimates of confounding variables 120
Blood counts 120
Parity data 121
Smoking habits 121
4.3 Familial aggregation of skewed Xi patterns 121
4.3.1 Primary skewing trait (DS score ≥ 0,25) 121
4.3.2 AS trait (AS_{DS} score ≥ 0,25) 122
4.4 Physiological relevance of skewed Xi patterns: association with biological variables 122
4.4.1 Hematopoietic indices and Xi skewing 123
4.4.1.1 DS score 123
4.4.1.2 AS_{DS} score 124
4.4.2 Clinical data and Xi patterns 125
4.4.2.1 DS score 125
4.4.2.2 AS_{DS} score 126
4.4.3 Parity data and Xi patterns 127
4.4.3.1 DS score 127
4.4.3.2 AS_{DS} score 128
4.5 Etiologies of skewed Xi patterns 128
4.5.1 Environmental etiology: cigarette smoke 128
4.5.1.1 DS score 128

4.5.1.2 AS_{DS} scores ... 129
4.5.2 Familial/genetic etiology: sibling correlation of XIRs and heritability coefficients ... 129
4.5.2.1 Xi Skewing trait: deviation from random Xi ... 129
4.5.2.2 AS trait: Deviation from the BC XIR ... 130
4.5.3 Segregation analysis of skewed Xi patterns (preliminary findings) ... 130
4.5.3.1 Segregation analysis of skewed Xi patterns (DS score ≥ 0,25) ... 130
4.5.3.2 Segregation analysis of the AS trait (AS_{DS} score ≥ 0,25) ... 131
DISCUSSION FOR PART II ... 132
Descriptive statistics and confounding variables ... 132
Genetic component to variability of hematopoietic cell numbers ... 133
Primary Xi pattern: BC XIR ... 133
Evidence of an X-linked genetic component: familial aggregation / correlation analyses ... 133
Candidate gene: ... 134
Grounds for a genetically complex trait ... 134
Number of progenitor cells present when Xi is initiated ... 134
Linkage analysis of a genetically complex trait ... 135
Parametric analysis ... 135
Nonparametric analysis ... 136
Allelic association ... 136
AS trait: hematopoietic XIRs ... 137
Evidence of an X-linked genetic component: familial aggregation / correlation analyses ... 137
T lymphocyte XIRs: strong evidence of a familial / X-linked genetic component ... 138
AS trait: grounds for a genetically complex trait ... 138
Candidate gene(s) for AS: modulation of HSC kinetics ... 138
Role of environmental factor(s) in variance of the AS trait: cigarette smoke ... 139
Hematologic associations / consequence of AS ... 139
Blood Xi-skewing (DS score) and hemoglobin concentration ... 140
AS trait (DS and AS_{DS} scores) and eosinophil counts ... 140
Blood Xi-skewing (DS score) and platelet count ... 141
Clinical consequence of the primary Xi pattern ... 141
Primary Xi pattern and asthma ... 141
XCE-like locus may influence expression of X-linked genes ... 142
Preferential inactivation of Xm versus Xp ... 143
Clinical associations of the AS trait ... 144
AS and susceptibility to RA ... 144

AS and immune-modulation 144
Chapter V 184
CONCLUSION, PERSPECTIVES AND FUTURE DIRECTIONS 184
5.1 Conclusion 185
Primary Xi trait 185
Genetic (X-linked) component to variability of BC XIRs 185
Clinical findings associated with the BC XIR 186
AS trait 186
Hematopoietic stem cell origin of AS 186
Genetic (X-linked) component to variability of AS values 186
Hematologic associations of the AS trait 186
Clinical associations of AS 187
5.2 Perspectives and Future Directions 188
Reference List 191
APPENDIXES i
Annex 1 ii
HERITABILITY OF NONMRANDOM X-INACTIVATION IN HEALTHY FEMALES ii
Annex 2 xviii
SKEWED Xi AND LYMPHOCYTE CLONALITY xviii
Annex 3 xxii
EVIDENCE OF A HSC ORIGIN OF SKEWING IN HUMANS xxii
Annex 4 xxiv
PILOT STUDY: EVIDENCE FOR TWO Xi TRAITS IN ELDERLY HUMAN FEMALES xxiv
Annex 5 xxxi
FORMULAIRE DE CONSENTEMENT xxxi
Annex 6 xxxv
QUESTIONNAIRE MÉDICAL xxxv

LIST OF FIGURES

Chapter 1

Figure 1 Timing of X inactivation in the mouse model....6
Figure 2 The Xic and it's elements....7
Figure 3 *Xist*, *Tsix* and X inactivation....11
Figure 4 HUMARA clonality assay....22-23
Figure 5 X-linked hemizygous theory of AS....49

Chapter 3

Figure 6 Distribution of family size (female siblings)....89
Figure 7 Age distribution of female participants....89
Figure 8 Incidence of skewed X inactivation (DS score $\geq$ 0.25)....90
Figure 9 Incidence of AS in hematopoietic lineages....91
Figure 10 Frequency distribution of DS scores....92
Figure 11 Frequency distribution of AS_{DS} scores....93
Figure 12 Age-PAmat scatter plots....94
Figure 13 Age-DS score scatter plots....95
Figure 14 Age-AS_{PAmat} score scatter plots....96
Figure 15 Age-AS_{DS} score scatter plots....97
Figure 16 Body-wide XIR: intra-individual correlation of PAmat scores....98
Figure 17 Blood skewing: intra-individual correlation of PAmat scores....99
Figure 18 Acquired skewing: intra-individual correlation of AS_{PAmat} score. 100
Figure 19 Direction of AS relative to body-wide XIR....101

Chapter 4

Figure 20 Frequency distribution of blood counts....146-147
Figure 21 Frequency distribution of blood counts without outliers....148-149

LIST OF TABLES

Chapter 1

Table I The Xic and its elements....7
Table II Characteristics acquired with *Xist* upregulation....12
Table III X inactivation based clonality assays....19
Table IV Utility of the HUMARA assay....24-25
Table V Primary versus secondary etiologies of skewing....30
Table VI Intra-individual correlation of XIR among various tissues....40

Chapter 2

Table VII Cell-sorted populations and purity....65

Chapter 3

Table VII Number of informative females (DNA amplified....102
Table IX Incidence of skewing and AS (n=3)....102
Table X Mean X inactivation ratios....103
Table XI Skewing as a function of age....104
Table XII Intra-individual correlation of XIRs....105
Table XIII Evidence supporting a body-wide skewing trait....106
Table XIV Direction of AS relative to the body-wide XIR....107

Chapter 4

Table XV Descriptive statistics for hematopoietic indices....150
Table XVI Outliers removed from blood counts....151
Table XVII Descriptive statistics of clinical data....152
Table XVIII Descriptive statistics of smoking habits....153
Table XIX Descriptive statistics of family / parity data....154
Table XX Association between age and hematopoietic indices....155
Table XXI Association between age and clinical data....156
Table XXII Association between age and smoking characteristics....157
Table XXIII Association between age and parity....158
Table XXIV Association between clinical data and hematopoietic 159-160
Table XXV Association between clinical data and parity....161
Table XXVI Association between smoking habits and clinical data....162
Table XXVII Association between smoking habits and blood counts....163
Table XXVIII Association between smoking habits and parity....164
Table XXIX Association between parity and blood counts....165

Table XXX Covariates associated with hematopoietic indices....................166

Table XXXI Heritability estimates of confounding variables......................167

Table XXXII Familial aggregation of skewing and AS...............................168

Table XXXIII Association between skewing and blood counts............... 169-170

Table XXXIV Association between skewing and clinical data.................171-172

Table XXXV Parity data and skewing..173

Table XXXVI Association between smoking habits and skewing............174-175

Table XXXVII Familial resemblance of XIRs..176

Table XXXVIII Segregation analysis of the skewing trait........................177-183

LIST OF ABBREVIATIONS

λ	recurrence risk ratio
ar	allele ratio
APOE-4	apolipoprotein E - 4
AS	acquired skewing
BM	bone-marrow
BMT	bone marrow transplant
bp	base-pair
BRCA1	breast cancer 1 gene
CBC	complete blood count
CPM	confined placental mosaicism
COPD	chronic obstructive pulmonary disease
$CysLT_1$	cysteinyl leukotriene receptor
DMD	Duchenne muscular dystrophy
DNA	deoxyribonucleic acid
dpc	days post-coitum
DS	degree of skewing
DZ	dizygous
ES	embryonic stem
ET	essential thrombocythemia
FITC	fluorescent isothiocyanate
G6PD	glucose-6-phosphate dehydrogenase
GFP	green fluorescent protein
GLM	general linear model
HPRT	hypoxanthine phosphoribosyltransferase
HSC	hematopoietic stem cell
HUMARA	human androgen receptor
ICF	immunodeficiency, centromeric instability, facial dysmorphism
Kb	kilo base
LA	linkage analysis
L1	long interspersed repeat elements-1
LD	linkage disequilibrium
LOH	loss of heterozygosity
MCV	mean corpuscular volume

MDS	myelodysplastic syndrome
MLL	mononuclear layer
MPD	myeloproliferative disorder
mRNA	messenger RNA
MS	multiple sclerosis
MZ	monozygous
ng	nanogram
NHT	nonhematopoietic tissue
OA	osteoarthritis
PALA	proportion active of the larger HUMARA allele
PAmat	proportion active of the maternal allele
PAR	pseudoautosomal region
PB	peripheral blood
PCR	polymerase chain reaction
PCV	polycythemia vera
PGK	phosphoglycerate kinase
PE	phycoerythrin
PMN	polymorphonuclear cells
PS	primary skewing
RA	rheumatoid arthritis
RBC	red blood cells
RE	restriction enzyme
RFLP	restriction fragment polymorphism
RNA	ribonucleic acid
RR	relative risk
RSA	recurrent spontaneous abortion
SA	spontaneous abortion
SLE	systemic lupus erythematosus
SNP	single nucleotide polymorphism
SOLAR	Sequential Oligogenic Linkage Analysis Routines
SRS	simple repeat sequence
TE	Tris-EDTA
TSG	tumor suppressor gene
TDT	transmission disequilibrium test
TRD	transmission ratio distortion
Tsix/TSIX	'Xist spelled right-to-left'

VNTR	variable number tandem repeat
WAS	Wiskott-Aldrich syndrome
X-SCID	X-linked severely combined immunodeficiency
Xa	active X chromosome
Xce	X-chromosome controlling element
XCI	X chromosome inactivation
Xi	inactive X chromosome
Xic/XIC	X inactivation center
XIP	X inactivation pattern
XIR	X inactivation ratio
Xist/XIST	X-inactivation specific transcript
XLA	X-linked agammaglobulinemia
Xm	maternal X chromosome
Xp	paternal X chromosome

ACKNOWLEDGEMENTS

I am very grateful to my director and mentor Lambert Busque for timely discussions, intellectual stimulation, and for allowing me the opportunity and time allotted to advance the chromosome X project. His insight and encouragement were equally appreciated.

I thank the team members I've been collaborating with over the duration of the project, without whom the project would not have been feasible: in particular Sylvie Provost, Marianne Gingras and Sylvie Fafard for technical experience, encouragement and timely discussions; Linda Lizotte and Nathalie Rathé for specimen and data collection; Marie-Pierre Dubé, Andrew Paterson and Dushanthi Pinnaduwage for statistical insight and analyses; and to Marc Lussier and Pierre Chagnon for project management, genetic expertise and encouragement. I am also grateful to Nathalie Labreque and Martin Guimond for technical assistance and/or immunological insight. Technical assistance and discussion of provoking genetical issues with Patrick Scott is gratefully acknowledged. My indebtedness goes to Alain Bonnardeaux, Denis Cournoyer, Tran Hoang and Claude Perreault for being members of my Masters thesis review board and/or pre-doctoral committee.

To my family and friends, thanks for your support and encouragement.

To my Parents, Lisa, Steven and Sylvia

Through error you come to the truth!

I am a man because I err!

Fyodor Dostoevsky

Necessity is the mother of invention.

Chapter I

INTRODUCTION AND LITERATURE REVIEW

1 THE X CHROMOSOME

1.1 Basic principles

The term "X chromosome", used to denote the sex chromosome that determines the development of the homogametic female sex, has been adopted in honor of Henking's observation of a densely stained body that appeared in half of secondary spermatocytes of the heteropteran insect, Pyrrhocoris apterus. He was uncertain of the clear nature of the body so he labeled it "X" for unknown, years later identified as the sex chromosome. The X is an exceptional chromosome as it can undergo inactivation or reactivation - dependant on chromosome company, developmental pathway (soma versus germline) and stage of development. Among eutherian mammals, with few exceptions, it is conserved in size and compromises roughly 5% of the haploid genome (Ohno, 1967). Moreover, as the X chromosome lacks a pairing partner in males (aside from the PAR regions that recombine with the Y), Ohno hypothesized that it is relatively protected from rearrangements thus conserving gene linkage to the X chromosome across various eutherian species.

Among mammalian species, X-linked gene dosage equivalence between XX females and XY males is accomplished by a fundamental mechanism of gene regulation: X chromosome inactivation. By the same mechanism, aneuploidy of the X chromosome (extra copies or monosomy) is well tolerated. X inactivation (Xi) is initiated during early female development and is a fundamental requirement for normal development. Failure to do so has been associated with severe developmental defects and embryonic lethality (Migeon et al., 1993). In particular, ectodermal cell death and absence of mesodermal formation has been observed in mouse embryos bearing two active Xs (Takagi and Abe, 1990). To achieve gene dosage equivalence with the autosomes, gene expression from the single active X is upregulated approximately 2-fold, as demonstrated in different mice strains where a particular locus was X-linked in one but autosomal in another (Adler et al., 1997), consistent with Ohno's hypothesis for an evolutionary requirement for high level gene expression from the single active X chromosome (Ohno, 1967).

1.2 Lyon's Hypothesis

The X chromosome inactivation hypothesis, first proposed by Mary Lyon (Lyon, 1961), states that: (i) one of the two X chromosomes in mammalian female cells is genetically inactivated; (ii) the inactive X could be maternal or paternal in origin; (iii) inactivation occurs early in embryonic development and remains fixed in progeny cells. The latter is consistent with clonal inheritance of

X inactivation patterns (Davidson et al., 1963). In other words, "X chromosome inactivation is the transcriptional silencing of a randomly selected X chromosome initiated in early female development". As a result, females are functional mosaics for X-linked polymorphisms, with two distinct cell lines/populations. Lyon based her hypothesis on summation of the following observations: i) the Barr body (then referred to as the X-chromatin body) was formed by condensation of a single X chromosome (Ohno et al., 1959); ii) the asynchronous labeling of the X chromosomes (Gilbert et al., 1962); iii) mice with an XO genotype were normal fertile females, suggesting a single X is required for development (Welshons and Russell, 1959); and iv) mice heterozygous for an X-linked coat color gene (mottled, brindled, tortoise-shell, or tabby - most of which are lethal in the hemizygous state) present a variegated coat phenotype (Welshons and Russell, 1959) (Lyon, 1960). In fact, in the heterozygous state, these genes give rise to a random patchy, somewhat linear, distribution of abnormal and wild type coat colors (Lyon, 1962). Dosage compensation by X inactivation has been tested and extended to humans (Gartler et al., 1992). Indeed, very similar patterns can be seen in female carriers of various X-linked skin disease alleles.

1.3 Anatomical human X-mosaicism

In humans, the lines of Blashko, a nonrandom developmental pattern of the skin, manifest in the heterozygous state of various X-linked gene defects, with mutant gene expression covering affected areas and wild-type gene expression constituting normal skin. Examples of X-linked skin disorders with clinical manifestations following the Blashko lines are focal dermal hypoplasia, chondrodysplasia punctata, and hypohidrotic ectodermal dysplasia (Happle, 1985). The nature and origins of the lines of Blashko may be best explained by the visible consequences of Lyonization (Happle, 1985), possibly reflecting the stream or trend of embryonic tissue growth. They describe a V-shape over the spine, on the abdomen they frequently form whorls and on the limbs they run in a more or less perpendicular linear direction. When compared to mosaic defects of human skin, the banding observed in mice is usually much more coarse, vaguely resembling the distribution of Blashko's lines.

Although the linear skin lesions are likely to reflect clonal proliferation of two functionally different cell populations, analysis of small skin specimens (3-6mm) from normal individuals has found similar X inactivation patterns (XIP) in different regions of the same individual, suggesting skin growth is characterized by considerable cell mixing (Fialkow, 1973). Similar findings were reported for uterine tissue (Linder and Gartler, 1965) and hair roots (Gartler et al., 1969). Moreover, in human skin samples composed of approximately 35 basal keratinocytes, a fine mosaic of tiles was observed, with the maternal or paternal X chromosome inactivated in each tile (Asplund et al.,

2001). It appears therefore that the distribution pattern of skin cells may be influenced by X-linked genetic factors: in healthy females, a fine mosaic of tiles; in carrier's of certain X-linked skin disease alleles on the other hand, a linear pattern may be observed.

1.4 Timing of X inactivation: *concurrent with tissue differentiation*

Although the process and mechanism of Xi has been investigated in several species, a prime model is the mouse as it is experimentally amenable and a plethora of information is available. The following is an outline.

Pre-fertilization

During early female development, oogenesis is characterized by reactivation of the inactive X, occurring prior to the leptotene stage (Kratzer and Chapman, 1981), (Gartler et al., 1980), ensuing in two active X chromosomes (Epstein, 1972). During male gametogenesis on the other hand, the active X chromosome is progressively inactivated around first meiotic prophase (Lifschytz and Lindsley, 1972), ensuing in formation of the XY body. Akin to the female soma, the inactive X is relatively condensed/heterochromatic. In contrast however, although transcriptionally inactive (Richler et al., 1992), the CpG islands of house keeping genes are relatively unmethylated (Driscoll and Migeon, 1990).

Post-fertilization

Key observations of Xi occurring post-fertilization have been documented *in vivo* using harvested mouse preimplantation embryos and in embryonic stem (ES) cells. In both systems, initiation of Xi has been tightly regulated to tissue differentiation - see Figure 1 (page 6) . Upon fertilization, the zygote and early blastocyst (prior to implantation) embodies two active X chromosomes (Gartler et al., 1972). Reactivation of the sperm-derived inactive X chromosome however has not been well characterized. Upon differentiation, one X is inevitably selected for inactivation (Epstein et al., 1978), (Kratzer and Gartler, 1978), (Monk and Kathuria, 1977), (Monk and Harper, 1979), (Penny et al., 1996). In cells destined to form extraembryonic lineages (trophectoderm and primitive endoderm), the paternal X bears a putative imprint marking the chromosome for early inactivation, wherein Xi is initiated 4,0-5,0 days post-coitum (dpc), shortly around or after the time of implantation (Takagi and Sasaki, 1975), (West et al., 1977), (Takagi et al., 1978), (Monk and Harper, 1979), (Costanzi et al., 2000). After implantation (5,5 – 6,5 dpc), a genome-wide demethylation event is believed responsible for erasure of the parental imprint(s) (Rastan, 1982a) (Monk et al., 1987), ensuing in random inactivation of parental X chromosomes in cells of the inner cell mass (Gardner and Lyon, 1971).

Furthermore, although preliminary data found that Xi occurs concurrently with tissue differentiation (Monk and Harper, 1979), unequivocal evidence was derived from a mouse line transgenic for an X-linked *Lac Z* transgene. The transgene, which was subject to inactivation, underwent inactivation in a developmental context, gradually proceeding in sub-populations and lineages, and was virtually completed by 11,5 dpc (Tan et al., 1993; Tan et al., 2000). Thereafter, the inactive state of the chromosome was clonally inherited in daughter cells. In humans, Xi is speculated to occur approximately 16 dpc (Park, 1957).

2 MECHANISM OF X INACTIVATION

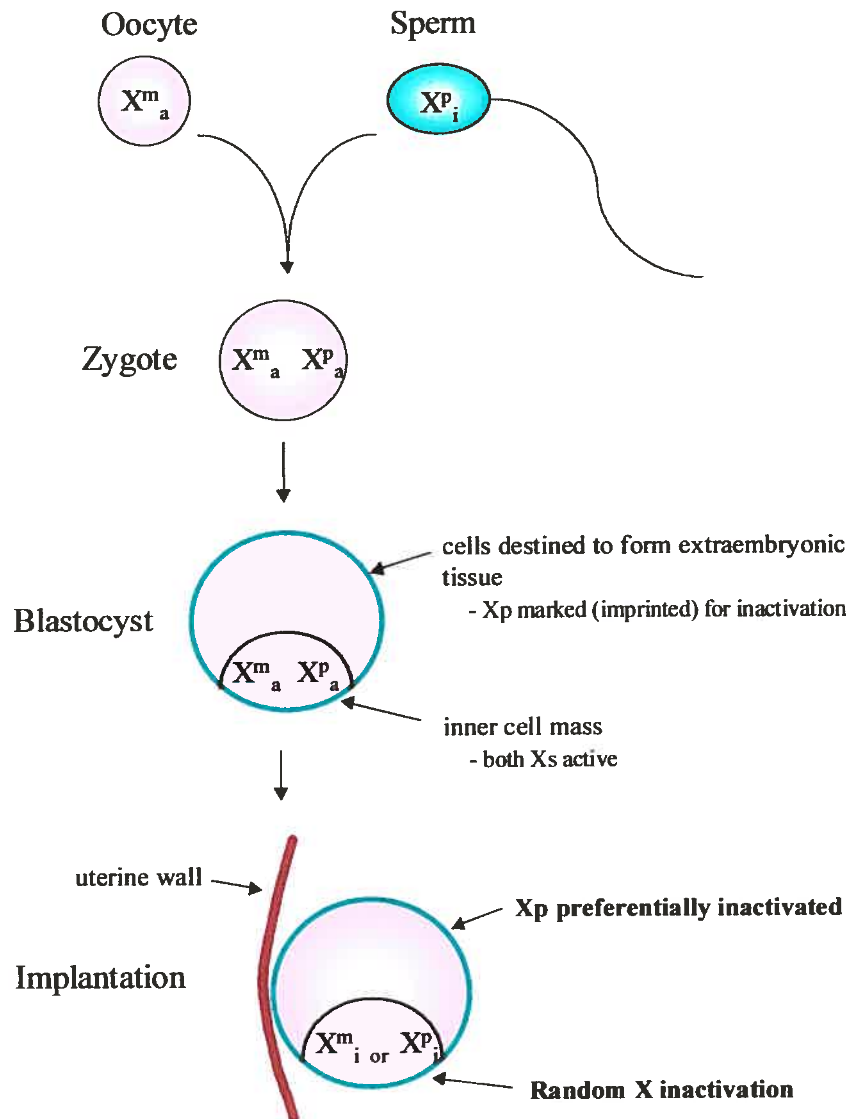
Xi is an intensely studied subject that involves several molecular mechanisms: counting and choosing, initiation/propagation and maintenance. An introduction to the underlying principles, properties and mechanisms of Xi will be provided (for a comprehensive review see (Willard, 1996a; Avner and Heard, 2001)). The X-inactivation center (Xic), spanning approximately 1MB, is a key *cis*-regulator for the initiation of Xi. Several elements involved in the Xi process reside within the Xic (Figure 2, page 7) (Table I, page 7).

2.1 The X inactivation center (Xic/XIC)

The first clues for an "initiator of X chromosome inactivation" came from genetic studies wherein mutant X chromosomes lacking a particular region were unable to undergo inactivation. This deleted region was thereafter called the X chromosome inactivation center (Xic). Evidence that the Xic functions *in cis* has been progressively defined by the partial spreading of inactivation onto autosomal segments in X-autosome translocations containing the Xic/XIC (Russell, 1963), (Rastan, 1983), (Cattanach et al., 1991), (Jeppesen and Turner, 1993; Jeppesen and Turner, 1993). Localization of the human XIC has been assigned to band Xq13 (Brown et al., 1991a), (Leppig et al., 1993), (Lafreniere et al., 1993). There is no convincing evidence for more then a single XIC (Willard, 1996b). The corresponding syntenic region in mouse is somewhat smaller and the overall organization of the Xic/XIC is poorly conserved between human and mouse (Debrand et al., 1998).

Three general processes have been attributed to the Xic/XIC: counting, choice and initiation/propagation.

Figure 1. Timing and appearance of X inactivation during female mouse development. After formation of the zygote, both Xs are relatively active. An imprinting pathway, where Xp (paternal X chromosome) is preferentially inactivated, is operative in tissues destined to form extraembryonic structures. At around the time of implantation, random X inactivation begins in cells destined to form epiblast-derived tissues.



Abbreviations: **X^m, maternal X; X^p, paternal X; a, active; i, inactive**

Figure 2. The Xic region and its defined elements. The Xic, located at Xq13, has been implicated in three defined processes of X inactivation: i) counting, ii) choice, and iii) initiation/propagation. Genetic elements identified within the Xic that play a role in the X inactivation process are *Xist*, *Tsix* and the *Xce* element. Distances between elements and the size of the genetic elements are approximate.

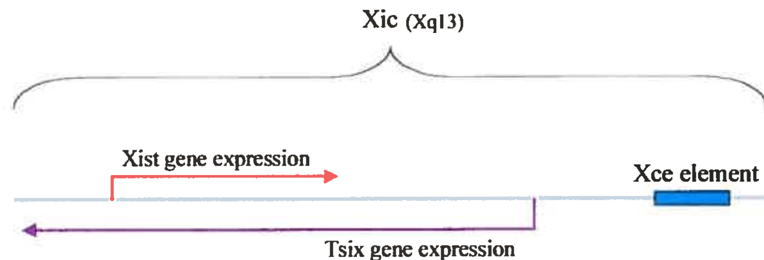


Table I. The Xic and its elements.

<u>Xic</u>: A master control region required for the initiation of X inactivation. As the Xic functions *in cis,* X chromosomes void of an Xic fail to undergo inactivation. A counting, choosing and spreading role has been attributed to this locus. Effector elements residing within the Xic include *Xist*, *Xce* and *Tsix*.
<u>*Xist*</u>: Localized to the Xic region, the X inactive specific transcript (*Xist*) is a non-protein coding mRNA essential for the initiation of X inactivation *in cis* and appears to be the primary signal for propagation of inactivation along the chromosome.
<u>*Tsix*</u>: *Tsix* is a mRNA species transcribed antisense to and approximately 15 Kb downstream of the *Xist* gene. In mice, the transcript spans the whole of the *Xist* gene while in humans a shorter transcript has been identified. Targeted deletions in the 5' region of *Tsix* leads to skewed X inactivation and disrupts imprinted X inactivation in extraembryonic tissues, suggesting *Tsix* influences X chromosome choice. A possible role of *Tsix* may be regulating *Xist* activity *in cis*.
<u>*Xce*</u>: The apparent function of the *cis*-acting X-chromosome controlling element (*Xce*) is to modulate the probability upon which a particular X chromosome is inactivated. In the mouse, a gradient of alleles has been demonstrated. In the heterozygous state, the X chromosome bearing a stronger allele is preferentially selected to remain active versus the X with the weaker allele, resulting in primary nonrandom X inactivation. *Xce* activity resides 3' to *Xist*.

i) **Counting**: This step senses the number of X chromosomes in the cell and ensures that only a single X chromosome remains active per diploid cell - all other X(s) are inactivated (Therman and Patau, 1974) (Rastan, 1982a). The absolute requirement of at least two Xics is required for initiation of X inactivation (Rastan and Robertson, 1985), (Rastan, 1983). How a single X is selected is not presently clear. It is speculated that a blocking factor produced in limited amounts binds to a single Xic (reviewed in (Migeon, 1994). Alternatively, allelic methylation differences within a proposed differential methylation region may be another (Chao et al., 2002). Since the number of active X chromosomes is dependant on the number of chromosomal sets (ploidy number of autosomes), as there are one to two active X's in XXX triploids (69 chromosomes) and two active X's in XXXX tetraploids (92 chromosomes), an autosomal origin for the blocking factor has been supported (Carr, 1971), (Jacobs and Migeon, 1989). In fact, recent findings identified the CTCF insulator / transcription factor as 'the' or 'one of' several putative trans-acting "blocking factor". As several CTCF binding sites were identified in the mouse Tsix/DXPas34 region, the authors postulate that Tsix and CTCF work together to designate the future active X by inhibiting Xist activity *in cis* (Chao et al., 2002). For a review of plausible mechanisms, see (Percec and Bartolomei, 2002). The role for CTCF as the human "blocking factor" is presently unclear as the number of CTCF binding sites in the syntenic human region has significantly less CTCF binding sites.

ii) **Choice**. During the choice process, one X is selected to remain active and the other to be inactivated. Although believed to be a random process, with maternal and paternal X boasting equal inactivation probability, deviation from random inactivation has been noted. For example, in extra-embryonic lineages of mouse, the paternal X is marked for preferential inactivation. In the embryo proper alternatively, the imprint is believed erased soon after implantation, ensuing in random X inactivation. Nonetheless, in mice, allelic variants of the *Xce* locus (the nature of which is presently unclear) can compromise choice, ensuing in primary nonrandom X inactivation patterns. A gradient of *Xce* alleles has been identified, each influencing the probability of inducing cis-inactivation (Cattanach and Isaacson, 1967), (Cattanach et al., 1969), (Cattanach and Perez, 1970). It is speculated that the Xce affects affinity for the trans-blocking factor. Additional elements recently implicated in the choice step are *Tsix* (Lee and Lu, 1999), *Xist* (Marahrens et al., 1997), (Boumil and Lee, 2001), (Newall et al., 2001) and *Xite* (Ogawa and Lee, 2003). Further, as CTCF may act in concert with Tsix in designating the future active X, it may also play a role in the choice step.

iii) **Initiation / propagation**: This step involves initiation and spreading of the inactivation signal onto neighboring regions in a *cis*-mediated manner (Rastan, 1983), (Lee et al., 1996), (Willard, 1996a). Given the cis-limited function of Xist, binding sites or nucleation centers (booster elements) have been speculated (Gartler and Riggs, 1983), (Riggs, 1990). As a candidate booster element,

Lyon (Lyon, 1998) proposed long interspersed repeat elements-1 (L1), based on two observations. i) The X chromosome is enriched for such elements, approximately two-fold versus the autosomes (Bailey et al., 2000). ii) In the event of X-autosome translocations, the extent of inactivation spreading onto autosomal regions correlated with levels of L1 elements. Moreover, X-linked regions that contain genes that escape inactivation are significantly reduced in L1 content versus chromosomal segments containing genes subject to inactivation (Bailey et al., 2000).

Although Xist RNA coating of the inactive X chromosome may have a chromatin-remodeling role, (Clemson et al., 1998), Xist expression per se may not be required for 'maintenance' of inactivation (Csankovszki et al., 1999) (Wutz and Jaenisch, 2000) as loss of the XIC in humans has not been associated with instability of the inactive state (Brown and Willard, 1994), thus suggesting that once silencing has been achieved, it is irreversible and thereof independent of Xist. Moreover, when Xist expression was reactivated from the active X or when ectopic Xist expression was induced, inactivation could not be initiated (Wutz and Jaenisch, 2000) (Clemson et al., 1998), consistent with a developmental context of Xist expression.

2.2 Genetic elements of the Xic: *Xist* gene

The *Xist* gene, for X inactivation specific transcript, is believed to play a primary role in the X inactivation process: i) it is located within the Xic (Xq13.2), ii) is transcribed specifically from the inactive X chromosome in both human (Brown et al., 1991b) and mouse (Borsani et al., 1991), (Brockdorff et al., 1991), iii) is transcribed prior to Xi (Kay et al., 1993), and iv) has been shown to be essential for *cis*-mediated Xi by loss-of-function experiments (Penny et al., 1996), (Marahrens et al., 1997), (Lee et al., 1996). Although necessary for Xi, the *Xist* gene is not required for male development since male mice bearing a Xist deletion are fertile and physiologically normal (Marahrens et al., 1997).

The *Xist/XIST* gene encodes a large untranslated RNA (Brown et al., 1992; Lee et al., 1996), (Brockdorff et al., 1992) localized to the nucleus (Brown et al., 1992; Clemson et al., 1996; Lee et al., 1996). Although both the mouse and human gene contain at least 8 exons, the XIST cDNA spans a maximal 19.3 KB (Hong et al., 2000) while Xist cDNA spans at least 17.4 KB, however smaller isoforms have been produced (Hong et al., 1999). The mouse and human *Xist/XIST* gene exhibit moderate but significant conservation of sequence and gene structure (Nesterova et al., 2001).

2.2.1 Developmental regulation of Xist

Dynamic changes in Xist expression mark one of the earliest events of the Xi process (Figure 1.3a). In undifferentiated murine embryonic stem (ES) cells (both male and female), the X-chromosomes are relatively active and marked by basal Xist expression (Hong et al., 1999; Tai et al., 1994), (Beard et al., 1995), (Daniels et al., 1997a), (Panning and Jaenisch, 1996) (Panning et al., 1997), (Panning et al., 1997), (Johnston et al., 1998). This RNA however accumulates at the site of transcription and does not localize across the chromosome. Upon differentiation however, *Xist* RNA expression is upregulated (>15-fold) from the X destined for inactivation and is downregulated on the future active X (Panning and Jaenisch, 1996) (Panning and Jaenisch, 1996). Xist upregulation is thought to result from an increased RNA half-life (Panning et al., 1997), (Sheardown et al., 1997), possibly owing to a switch in Xist promoter usage (Johnston et al., 1998). However, recent data suggest an alternate mechanism: Tsix gene interference (vide infra) (Warshawsky et al., 1999). Nonetheless, Xist accumulation is rapid, coating the inactive X within one cell cycle (Wutz and Jaenisch, 2000), (Panning et al., 1997). Several thousand X-linked genes are transcriptionally silenced shortly after the accumulation of Xist RNA (Keohane et al., 1996).

2.3 Characteristics acquired with *Xist* upregulation

Acquisition of Xi involves several progressively acquired changes in structure and function, including recruitment of regulatory / chromatin proteins, many of which have been identified. Both genetic and epigenetic modifications have been characterized, see Table II (page 12) for an outline. Cytologically, the inactive X assumes a condensed, compact appearance coined the Barr body (Barr and Bertram, 1949).

Methylation of CpG islands of Xi-linked genes

During the course of Xi, the CpG islands of Xi-linked genes are generally hypermethylated (Wolf and Migeon, 1982; Wolf et al., 1984) (Keith et al., 1986) (Bird, 1986), ensuing in transcriptional silencing.

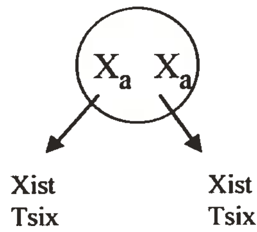
Asynchronous replication of the inactive X

With respect to its' active counterpart, the inactive X chromosome (and residing genes) replicate

Figure 3. *Xist*, *Tsix* and X inactivation. a) *Xist* and *Tsix* expression in undifferentiated versus differentiated ES cells. b) *Xist* and *Tsix* expression in extraembryonic (pre-implantation) versus embryonic tissues (post-implantation).

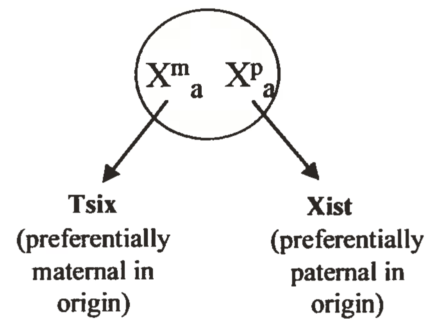
a) ES Cells

undifferentiated



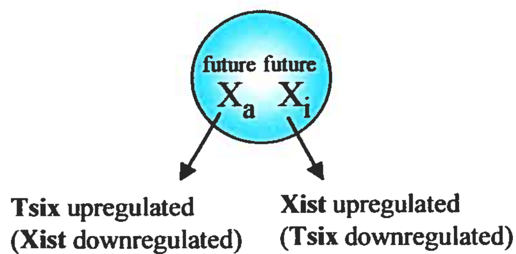
b) Embryo

pre-implantation

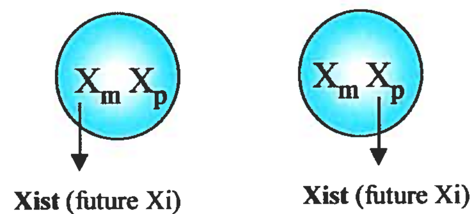


- preferential (imprinted) inactivation of Xp in extraembryonic tissues

differentiated



post-implantation



- erasure of imprint
- random X inactivation

Table II. Genetic/epigenetic and morphological characteristics acquired with inactivation of the X chromosome.

Genetic modifications

- *Xist* gene upregulation
- *Tsix* gene downregulation
- General shutdown of most X-linked genes

Epigenetic modifications

- Xist mRNA coating
- hypermethylation of CpG islands
- asynchronous replication (usually late-replicating)
- hypoacetylation of histones H3 and H4
- hypermethylation of histone H3
- high concentration of core histone macroH2A1.2
- enrichment for BRCA1 mRNA
- condensed chromatin structure, coined the Barr body

asynchronously (reviewed in (Heard et al., 1997)), namely in late S phase, (Boggs and Chinault, 1994), (Schmidt and Migeon, 1990). Genes that escape X inactivation on the other hand, replicate in synchronicity with its' active X homologue (Boggs and Chinault, 1994).

Covalent modifications of X chromosome associated histones

Histone acetylation

The transcriptionally active X is characterized by hyper-acetylation of NH_2-terminal lysine residues of histone H3 and H4 (Jeppesen and Turner, 1993), (Gilbert and Sharp, 1999), (Boggs et al., 1996), while those associated with promoters of transcriptionally inactive genes of the inactive X chromosome are essentially hypoacetylated.

Histone methylation

During X chromosome inactivation, methylation of lysine 9 of histone H3 (H3-K9 methylation) occurs within or shortly after Xist accumulation (Heard et al., 2001), (Mermoud et al., 2002).

Enrichment of histone variants

The inactive X is further characterized by a high concentration of the core histone macroH2A1.2 (Costanzi and Pehrson, 1998). Localization of macroH2A1.2 to the inactive X may be mediated by interaction with Xist mRNA, possibly by formation of a ribonucleoprotein complex (Csankovszki et al., 1999).

Enrichment of mRNA species

Recent studies have demonstrated that BRCA1 mRNA associates (possibly essential) with XIST mRNA in female somatic cells, mediating chromatin stability of the inactive X (Ganesan et al., 2002). Thereof, in absence of functional BRCA1 (either by germline mutations, LOH or RNA interference), i) the chromatin state of the inactive X was destabilized, ii) enrichment for macroH2A1.2 was suppressed and iii) particular X inactivated genes (a GFP transgene for example) were partially upregulated, thus suggesting BRCA1 plays a pivotal role in mediating Xi. Thus, the gender-specificity upon which BRCA1 mutations transcend to ovarian and breast cancer may be explained by genomic instability of the inactive X chromosome. Supportive evidence was derived from various $BRCA1^{-/-}$ tumors, of which demonstrated defects in chromatin structure and over-expression of several genes linked to the inactive X (Jazaeri et al., 2002). One hypothesis may be that derepression of Xi leads to a dosage imbalance of genes required for normal function/development of breast and ovaries.

2.4 Stability of X inactivation

Once Xi has been achieved, the chromatin state is clonally inherited (Davidson et al., 1963) and the inactive state believed stably maintained for duration of the cell lifespan (Migeon, 1972). However, exceptions have been noted. i) The inactive X is reactivated as a normal part of oogenesis. ii) Albeit at a low frequency (10^{-5} to 10^{-4}), in vitro cultivation has been associated with reactivation of various Xi-linked genes (Dyer et al., 1989) (Mohandas et al., 1981), (Graves, 1982). In addition, there is convincing evidence in mice that particular Xi-linked genes are partially reactivated with advancing age (Wareham et al., 1987) (Brown and Rastan, 1988).

3 REGULATORY ELEMENTS OF THE XIC

To identify regulatory regions within the Xic, deletion and transgenetic analyses in and adjacent to the *Xist* gene and their consequential effects on cis-mediated Xi were investigated.

Initiation element

Targeted deletion of the 5' region of *Xist* (7-Kb, including 30 bp of the promoter region) abolished initiation/propagation (Penny et al., 1996), delimiting initiation of Xi to this region (Lee et al., 1999a).

Counting elements

The counting process was delimited to a 37-Kb region lying 3' to *Xist*, but genetically separable from *Tsix* promoter and *Xite*. Moreover, as the aberrant X was also inactivated in differentiating XY cells, sex-specific factors in the initiation of Xi are putatively excluded.

Choice elements

The choice process appears influenced by various elements within the Xic. i) The elusive *Xce* locus, reputably associated with nonrandom Xi patterns (Cattanach and Papworth, 1981), has been delimited to a rather large region 3' to the *Tsix* locus (reviewed in section 7.3.2.1). ii) Targeted deletion of the Xist antisense transcript, i.e., *Tsix*, induces constitutive Xist expression and nonrandom Xi of the aberrant X (Lee and Lu, 1999). iii) The recently identified *Xite* gene (downstream from *Tsix*) also mediates choice as targeted deletion of the gene downregulated Tsix activity *in cis*, resulting in Xist upregulation and preferential inactivation of the aberrant X (Ogawa and Lee, 2003). iv) In humans, a rare base-pair mutation in the *XIST* promoter region has been associated with preferential inactivation of the mutant X (Plenge et al., 1997), suggesting mutation/polymorphisms in the *Xist* gene may influence Xi patterns. A plausible mechanism may be increased Xist transcription, thus increased probability of undergoing Xi. v) Induced chemical mutagenesis of the mouse genome identified two autosomal loci (yet to be clearly identified) which when mutated altered X chromosome choice (Percec et al., 2002).

3.1 Further characterization of the *Tsix* gene

The *Tsix* gene, so dubbed as it is transcribed antisense to *Xist*, encodes a 40-Kb transcript originating some 28 Kb downstream of *Xist* in both human and mouse (Lee et al., 1999b), (Sado et al., 2001), (Migeon et al., 2001). Although originally defined as having no conserved ORF, recent studies suggest it is partially processed, giving rise to a 2,7 and 4,3 Kb transcript (Sado et al., 2001). It is found exclusively in the nucleus, but unlike *Xist*, remains localized to the Xic (Lee et al., 1999b).

Role of Tsix in regulation of Xist

Because *Tsix* is i) transcribed antisense to *Xist*, ii) is concurrent with early Xist expression and iii) Tsix upregulation from the future Xa is associated with Xist downregulation (Panning and Jaenisch, 1996), (Lee et al., 1999b), a role in regulating Xist activity has been speculated. Available evidence suggests high level Tsix expression represses Xist likely though a transcriptional mechanism (Lee and Lu, 1999), however other mechanisms are plausible (see (Avner and Heard, 2001) for review). Once Xi is established, Tsix expression from Xa is downregulated to undetectable levels.

Similar to the murine homologue, human *TSIX* is expressed exclusively in epiblast derived cells. However, it is truncated at its 3' end, ending approximately at exon 5 of the *XIST* gene, deriving a transcript of approximately 35 Kb (Migeon et al., 2001). Moreover, human *TSIX* lacks the 5' CpG island putatively required for murine Tsix imprinting function (Sado et al., 2001), suggesting *TSIX* may not function like its murine counterpart (Migeon et al., 2001).

4 IMPRINTED X INACTIVATION

Genomic imprinting is a process whereby gene function is affected in a parental origin-specific manner and is manifested as a difference in expression of parental alleles. The effects of genomic imprinting arise from differential epigenetic modification of parental alleles during gametogenesis, followed by additional epigenetic processes that may occur after fertilization (Latham, 1999). Genomic imprinting may limit the effects of growth factors in the embryo and/or in extra-embryonic tissues, in which there is differential parental investment (Haig, 1993). Strong epigenetic effect of parental imprinting on the X inactivation process has been observed in mouse and weaker evidence has been found in human.

4.1 Preferential inactivation of the paternal X in extraembryonic tissues

In mice, unlike embryonic tissues that exhibit random Xi, the paternal X is preferentially inactivated in the first cells to differentiate, the physiological consequence of which is presently not clear, eventually giving rise to tissues destined for extra-embryonic development (Takagi and Sasaki, 1975), (West et al., 1977), (Harper et al., 1982), with 87-88% of trophoblasts possibly showing Xp inactivation (Takagi and Sasaki, 1975), (West et al., 1977). Resistance of Xm from undergoing inactivation has been ascribed to acquisition of an imprint during oogenesis (Lyon and Rastan, 1984), (Tada et al., 2000). The *Xist* gene is thought to play a vital role in the imprinted pathway since female mice inheriting a paternal *Xist* deletion bear 2 active Xs and die early in embryogenesis. Mice inheriting a maternal *Xist* deletion on the other hand, are normal and exhibit

exclusive paternal Xi (Marahrens et al., 1997). In humans however, there is no evidence for parental-specific preference of Xist expression (Daniels et al., 1997b) (Ray et al., 1997). Interestingly however, preferential inactivation of Xp in extra-embryonic tissues may not be related to imprinting at all. Alternatively, relative to the active maternal X, the inactive X derived from sperm may be more receptive to the inactivation signal (Monk and McLaren, 1981). Supporting evidence was derived from elegant cloning experiments where the inactive X (maternal or paternal in origin) of the donor-derived nucleus was preferentially inactivated in extra-embryonic tissues of the cloned embryo (Eggan et al., 2000), thus suggesting that the donor-derived inactive X, like in sperm, is poised for early inactivation.

4.2 X-imprinting in humans

Unlike the mouse, evidence for preferential inactivation of Xp in extraembryonic tissues in humans remains controversial. Although random Xi was observed in chorionic villi of early gestation (Migeon and Do, 1979), (Migeon et al., 1985), (Mohandas et al., 1989), (Bamforth et al., 1996), preferential inactivation of Xp was observed by others (Harrison and Warburton, 1986), (Harrison, 1989), (Goto et al., 1997) and partially supported by (Uehara et al., 2000). Furthermore, whereas trophoblasts from full-term placentas showed lack of preferential inactivation of Xp (Looijenga et al., 1999), Xp was preferentially inactivated in another study (Ropers et al., 1978). Discrepancies may be explained by methodology used, tissue type and/or maternal cell contamination.

5 ESCAPE FROM X INACTIVATION

An intriguing aspect of Xi is that certain genes escape inactivation, resulting in expression from both X chromosomes. Approximately 30% of genes on the short arm of X and less than 5% of genes on the long arm escape Xi, averaging over 10% of X-linked genes (Carrel et al., 1999). Relative to mice, humans are believed to have a larger number of genes that escape Xi (Disteche, 1999). These genes may be independent or clustered together, the latter found in or near the pseudoautosomal region (PAR) (Lyon, 1962), (Miller and Willard, 1998), suggesting a chromosomal domain model of regulation. It is unclear whether genes are initially inactivated and then re-activated, or simply remain active. It was hypothesized that genes escape Xi to compensate for functional Y homologues in males. However, many genes that escape Xi do not have an Y homologue, suggesting dosage imbalance between the sexes. Interestingly, several genes display variable escape from Xi, being expressed from the inactive X in some females but subject to inactivation in others (Carrel and Willard, 1999), (Anderson and Brown, 1999). Escape

from Xi may be a significant factor for physiological and disease susceptibility differences between the sexes (see (Brown and Robinson, 2000) for review).

6 CLINICAL APPLICATION OF THE LYON'S HYPOTHESIS

Several methods have been developed to determine the clonal nature of a pathological tissue. Originally these methods relied on specific gene defects associated with particular cancers (reviewed in (Gilliland et al., 1991a)). For example (i) immunoglobulin and T cell receptor gene rearrangements in lymphomas, (ii) cytogenetic and fluorescence in situ hybridization analysis of genetic aberrations (translocations, deletions, duplications and inversions) in acute myeloid leukemia, chronic myeloid leukemia and myeloproliferative disorders, (iii) point mutations in critical genes such as proto-oncogenes and tumor suppressor genes in hematological and gastrointestinal cancers, and (iv) viral integration in viral-associated lymphomas. However, not all tumors exhibit particular tumor-specific markers or the specific abnormalities are not known.

6.1 X inactivation-based clonality assays

Xi based clonality assays were developed on precepts of the Xi hypothesis (Linder and Gartler, 1965). These assays, which do not rely on the presence of tumor-specific markers, have the advantage of detecting clonal derivation of cells in any informative female. The finding that a pathological tissue is monoclonal is consistent with a neoplastic process (i.e., that a tumor has arisen from acquired somatic mutation(s) in a single progenitor cell (Nowell and Hungerford, 1960)) whereas normal and reactive cell populations (inflammatory tissues) are typically polyclonal in origin. Application of these assays is central to many aspects of our understanding of the normal biology of hematopoiesis as well as the pathogenesis of hematologic malignancies ((Busque and Gilliland, 1998) for review). For example, the stem cell origin of differentiated hematopoietic lineages (Gartler et al., 1969), the clonal origin of numerous hematological and non-hematological malignancies (Linder and Gartler, 1965), and the stem cell origin of various hematologic malignancies. In addition, these assays can be utilized for detection status of female carriers of various X-linked disease alleles (reviewed in (Puck and Willard, 1998)).

6.2 Principles of X-inactivation based clonality assays

Xi-based clonality assays, rely on two basic prerequisites. The first is to differentiate parental origin of the chromosomes: Xp from Xm. The second is to distinguish the active X (aX) from the inactive

X (iX). A limitation to these assays is exclusion of male (XY) samples. Table III (page 18) provides a list of informative Xi-based assays.

6.2.1 Distinguish Xp from Xm

The ability to distinguish Xp from Xm is based on polymorphisms that exist in the general population at different loci on the X chromosome (reviewed in (Busque and Gilliland, 1998)). These DNA polymorphisms are typically of a single base-pair variation or variation in length of a simple repeat sequence (SRS). Both PCR and Southern-blot based techniques are utilized, but PCR is more efficient and does not require large quantities of DNA starting material (pico- to nanogram quantities acceptable). Expression-based assays have been developed, which may be transcript- or protein-based. The informativeness of a polymorphism is directly related to the frequency and number of alleles present in the population, which may vary among different ethnic groups.

6.2.2 Distinguish active X (aX) from inactive X (iX)

The iX chromosome, characterized by hypermethylation of CpG dinucleotides in promoter regions and repression of gene activity, is amenable to both epigenetic and expression-based assays respectively (see (Busque and Gilliland, 1998) and (Gale and Linch, 1998) for review). Absolute requirements of the epigenetic-based assay are i) one allele be invariably methylated and the other invariably unmethylated, and ii) DNA methylation patterns be stable with advancing age. Expression-based assays rely on the principle that transcript/protein expression occur exclusively from the active X chromosome. Moreover, expression-based assays have the advantage of analyzing non-nucleated cells such as platelets and RBC. A comprehensive list of informative (> 30% heterozygosity) X-linked clonality assays is provided in Table III (page 18).

6.3 Utility of Xi based clonality assays

To be useful, a clonality assay should be highly informative and exhibit strict differential methylation patterns. Although not a single X-linked clonality assay is informative in 100% of females, the HUMARA assay was utilized for our analyses as it exhibits a high degree of heterozygosity, results are reproducible, has been thoroughly validated, and is the most frequently used assay to infer Xi status of normal and neoplastic specimens.

6.3.1 HUMARA assay

The HUMARA locus (DXS1213; Xq11-12) consists of eight exons spanning more than 90 Kb of DNA (Kuiper et al., 1989), (Sleddens et al., 1992), (Lubahn et al., 1988), (Chang et al., 1988). The

Table III. X inactivation based clonality assays.

gene / locus	location	polymorphism	% heterozygosity	assay	PCR	reference
AR (HUMARA)	Xq11-12	SRS	80-90	methylation	+	(Allen et al., 1992)
				expression	+	(Busque et al., 1994)
BGN	Xq28	SNP	~50	expression	+	(Kutsche and Brown, 2000)
DXS15-134	Xq28	SRS	~46	methylation	+	(Okamoto et al., 1998)
DXS16	Xq22	RFLP	~33	methylation	-	(Khalifa et al., 2001)
DXS225 (M27b)	Xp11.22	SRS	>80	methylation	-	(Boyd and Fraser, 1990)
FMR1	Xq27.3	SRS	45-65	methylation	+	(Lee et al., 1994), (Carrel and Willard, 1996)
G6PD	Xq28	RFLP/SNP	0-30*	expression°	-	(Beutler et al., 1962), (Linder and Gartler, 1965)
				expression	+	(Beutler and Kuhl, 1990), (Curnutte et al., 1992), (Prchal and Guan, 1993), (Liu et al., 1997)
HPRT	Xq26-27.2	RFLP	~28	methylation	-	(Vogelstein et al., 1985)
IDS	Xq28	RFLP/SNP	0-51*	expression	+	(el Kassar et al., 1997), (Harrison et al., 1998), (Gregg et al., 2000)
MAOA	Xp11	SRS	~75	methylation	+	(Hendriks et al., 1992)
p55	Xq28	RFLP/SNP	37-48	expression	+	(Belickova et al., 1994), (Luhovy et al., 1995), (el Kassar et al., 1997), (Liu et al., 1997)
PGK	Xq13	RFLP	33-52	methylation	-	(Vogelstein et al., 1987), (Tsukamoto et al., 1993)
				methylation	+	(Gilliland et al., 1991b)
TM4SF2	Xp	SNP	~40	expression	+	(Kutsche and Brown, 2000)
VBP1	Xq28	SNP	~40	expression	+	(Kutsche and Brown, 2000)
XIST	Xq13	SNP	~50	expression	+	(Rupert et al., 1995)

* Heterozygosity dependant on ethnic group analyzed
° Protein expression assay, otherwise mRNA

first exon contains a highly polymorphic in-frame CAG trimeric repeat. In the general population, 8 to 36 CAG repeat units have been observed, corresponding to a 28 allele-spectrum (personal observations). The alleles follow a Gaussian distribution which approaches a constitutional heterozygosity value of 90%, as determined in several racial groups (Sleddens et al., 1992), (Allen et al., 1992), (Edwards et al., 1992), (Busque et al., 1996), (Desmarais et al., 1998). Significant repeat expansion (38-66 repeats) has been associated with spinal and bulbar muscular atrophy (Kennedy disease), an X-linked motor neuron disease (La Spada et al., 1991). Proximal to the repeats are four particular methylation sensitive restriction enzyme (RE) sites that meet obligate methylation patterns. The cytosine residues within these sites are strictly methylated on the iX but unmethylated on the aX. Utilizing these key features, the PCR-based HUMARA clonality assay was developed (Allen et al., 1992). Since the polymorphism is in the coding region, an expression-based assay (mRNA) has also been developed (Busque et al., 1994). Figure 4 (page 22-23) depicts the principles of the HUMARA clonality assay and sample results are shown.

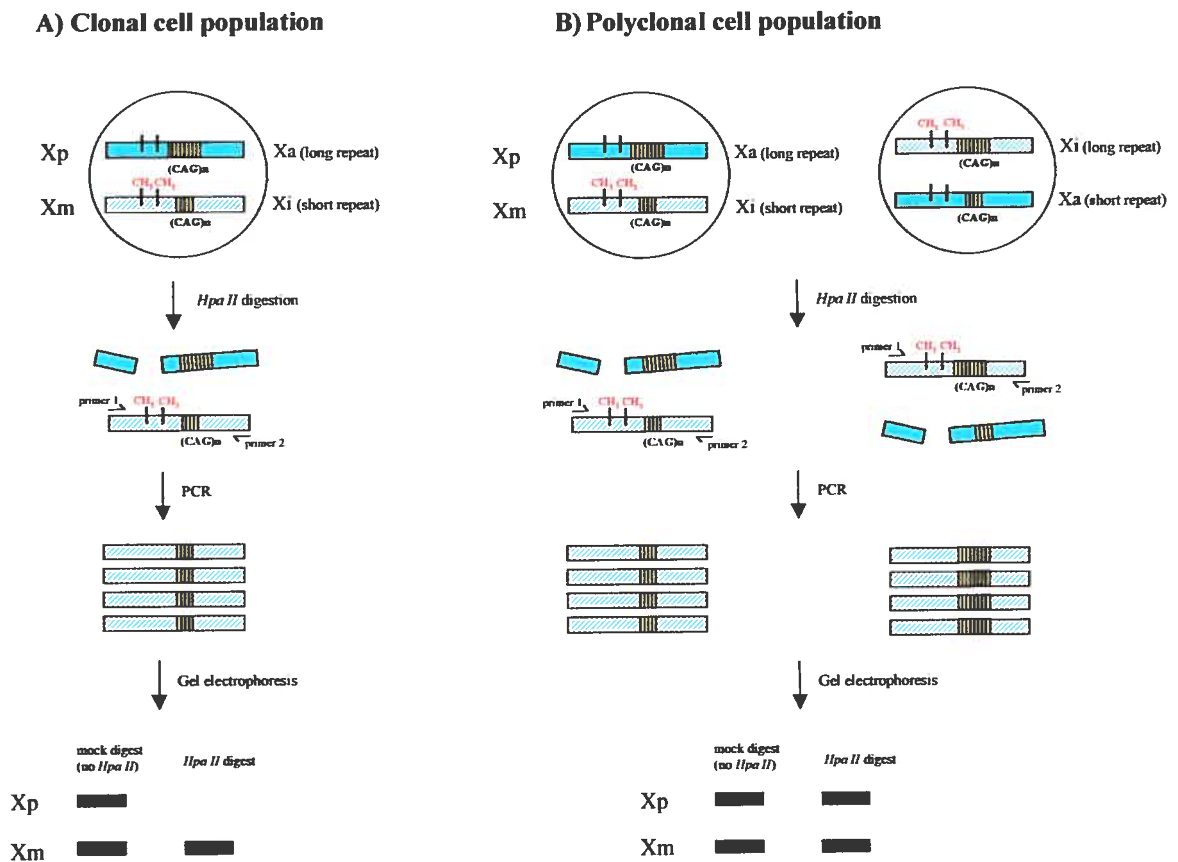
Utility of the HUMARA assay

Advantages of the HUMARA assay include PCR-based requiring little DNA as starting material, a high rate of heterozygosity, and precise/accurate quantitation of alleles. Commonly used techniques in the assessment of band intensities are visual evaluation, laser densitometry, and more recently, fluorochrome and phosphor imaging technology. The technique has been validated by several investigators, establishing a reliable and reproducible assay to study the clonal origin of sample specimens. In fact, it has confirmed or refuted the monoclonal and/or clonal origin (consistent with a neoplastic growth) of several diseases/disorders (see Table IV, page 24-25, for a comprehensive list). However, clonality does not necessitate expansion of a mutant clone, since in certain instances, large X inactivation patches, as observed on diffuse intimal thickening of coronaries, may mimic clonality (Murry et al., 1997). Further, in presence of a large polyclonal background, the assay is unable of detecting a clonal cell population (Allen et al., 1992), (Gilliland et al., 1991b). Because of the high heterozygosity rate, the HUMARA locus is frequently used as a genotype marker applicable for forensic/paternity testing (Desmarais et al., 1998) and linkage analyses. Moreover, as length of the repeat has been correlated with functional activity of the receptor, role of the AR gene in prostate, breast cancer (Giguere et al., 2001) and a wide variety of other diseases and conditions has been investigated (reviewed in (Yong et al., 2000)).

Figure 4. HUMARA clonality assay (figures A - C).

A) Clonal cell population. In this hypothetical example, the paternal X chromosome (bearing the smaller HUMARA allele) is inactive in all cells. The DNA is extracted and subjected to methylation sensitive restriction enzyme digestion (i.e., *Hpa II*). As the paternal allele is methylated, it is resistant to enzyme digestion, thus amenable to PCR amplification. Upon gel electrophoresis, a single band is detected on autoradiographic film, reflecting amplification of the inactive allele, signifying a monoclonal cell population. A mock digest (without addition of *Hpa II*) is also included, serving as an internal control (determines zygosity status and quantifies potential preferential amplification of the smaller allele). In a non-pathological setting, as in normal tissue samples, such a pattern is referred to as a skewed or a nonrandom X inactivation pattern.

B) Polyclonal cell population. A mixed population of cells where Xp and Xm are equally inactivated. After DNA extraction and RE digestion with *Hpa II*, only the intact/methylated alleles are amenable to PCR amplification, in this case, both alleles. Upon gel electrophoresis, two bands of similar intensity are detected on autoradiographic film, reflecting a polyclonal cell population. In a non-pathological setting, as in normal tissue samples, such a pattern is referred to as a random X inactivation pattern.



C) Sample HUMARA results of a clonal and polyclonal cell population. *Clonal cell population.* As two alleles are detected in the mock-digest lane, the sample is heterozygous thus informative for clonality analysis. In the *Hpa II* digest lane, a single predominant band is observed, consistent with clonal derivation (or a skewed X inactivation pattern if tissue was derived from a healthy female). *Polyclonal cell population.* Two alleles are detected in the mock-digest lane, consistent with an informative sample. In the *Hpa II* digest lane, two alleles of approximate equal intensity are detected, consistent with a polyclonal cell population (known as a random X inactivation pattern if tissue was derived from a healthy female).

Clonal cell population
(skewed X inactivation)

Polyclonal cell population
(random X inactivation)

Table IV. Utility of the HUMARA assay to confirm or refute the clonal origin of diseases/disorders.

Disease or disorder	Clonal	Reference
Renal system		
tuberous sclerosis harmatomas (sporadic and familial)	+	(Green et al., 1996)
renal sporadic angiomyolipoma	+	(Paradis et al., 1998a)
- smooth muscle/blood vessel component	+	(Saxena et al., 1999)
- adipose tissue component	-	(Saxena et al., 1999)
Skin and soft tissue		
melanoma (malignant)	+	(Harada et al., 1997)
nevacellular nevus	-	(Harada et al., 1997)
acquired navi	+	(Robinson et al., 1998)
malignant fibrous histiocytoma	+	(Fujita et al., 1998)
benign fibrous histiocytoma	+	(Chen et al., 2000)
giant cell tumor of tendon sheath	-	(Vogrincic et al., 1997)
sporadic basal cell carcinomas	+/- (60%)	(Walsh et al., 1998)
desmoid tumor	+	(Li et al., 1996), (Lucas et al., 1997)
Liver		
hepatocellular carcinoma	+	(Piao et al., 1997)
macronodules of liver cirrhosis	+/- (54%)	(Paradis et al., 1998b)
micronodules of liver cirrhosis	+/- (51%)	(Paradis et al., 2000)
large regenerative nodules of liver cirrhosis	-	(Piao et al., 1997)
focal nodular hyperplasia of the liver	-	(Paradis et al., 1997)
Lung and arteries		
sclerosing hemangioma of the lung	+	(Niho et al., 1998)
atypical adenomatous hyperplasia of the lung	+	(Niho et al., 1999)
endothelial cell proliferation in plexiform lesions in primary pulmonary hypertension	+/- (77%)	(Lee et al., 1998)
atherosclerotic vessels	-	(Murry et al., 1997)
plaques smooth muscle cells	+	(Murry et al., 1997)
Brain and CNS		
malignant gliomas	+	(Kattar et al., 1997)
low-grade gliomas	+	(Kattar et al., 1997)
gangliogliomas	+	(Zhu et al., 1997)
meningiomas	+/- (9/15)	(Zhu et al., 1995)

24

Hematological		
Reed-Sternberg cells in nodular sclerosing Hodgkin's disease	+	(Chang et al., 1999)
"aggressive" mastocytosis	?	(Horny et al., 1997)
Langerhans'- cell histiocytosis	+	(Willman et al., 1994)
essential thrombocythemia	+/- (69%)	(el Kassar et al., 1995)
idiopathic thrombocytosis	+/- (66%)	(Shih et al., 2001)
Rosai-Dorfman disease	-	(Paulli et al., 1995)
Chester-Erdheim disease	+/- (3/4)	(Chetritt et al., 1999)
Reproductive tract and sexual organs		
noninvasive papillary transitional cell carcinomas of the urinary tract	+	(Li and Cannizzaro, 1999)
endometrial adenocarcinoma	+	(Mutter et al., 1995)
uterine leiomyomata	+	(Mashal et al., 1994)
high-grade squamous intra-epithelial lesions of the uterine cervix	+	(Enomoto et al., 1997)
ovarian endometrial cysts	+	(Jimbo et al., 1997), (Yano et al., 1999)
vulvar intraepithelial neoplasia	+	(Tate et al., 1997)
disseminated peritoneal leiomyomatosis	+	(Quade et al., 1997)
carcinosarcoma of the breast		
- carcinomatous component	+	(Wada et al., 1998)
- sarcomatous component	+	(Wada et al., 1998)
Digestive system		
colonic crypts	+ ?	(Endo et al., 1995)
superficial depressed-type gastric carcinoma	+	(Bamba et al., 1998)
Type I gastric carcinoid tumors	+	(D'Adda et al., 1999)
Miscellaneous		
Kaposi's sarcoma	-	(Delabesse et al., 1997)
Kaposi's sarcoma	+	(Rabkin et al., 1997)
salivary gland pleomorphic adenomas	+	(Lee et al., 2000)
epithelial cells of Wartin's tumor	-	(Honda et al., 2000)

In certain instances, <100% clonality is consistent with a multi-hit hypothesis of neoplasia

6.4 Hematological application of clonality assays

6.4.1 Clonal origin of hematologic malignancies

Several precepts of modern hematologic oncology rely on data obtained by Xi-based clonality assays. The documentation of monoclonal derivation of cells in acute myeloid leukemia (Wiggans et al., 1978), acute lymphocytic leukemia (Dow et al., 1985), chronic myeloid leukemia (Fialkow et al., 1967), myelodysplastic syndromes (MDS) (Raskind et al., 1984), myelofibrosis with myeloid metaplasia (Kreipe et al., 1991), and in various other myeloproliferative disorders (MPDs) (Anger et al., 1990), (Fialkow et al., 1981) has provided a monoclonal/neoplastic origin for most hematologic malignancies.

6.4.2 Stem cell origin of hematologic malignancies

The demonstration of a skewed Xi pattern in multiple hematopoietic lineages has invoked a multipotent stem cell origin for various MPDs (Adamson et al., 1976), (Gaetani et al., 1982), (Raskind et al., 1985), (Fialkow et al., 1981), (Jacobson et al., 1978). Other studies however have cited heterogeneity in lineage involvement (Anger et al., 1990), (Asimakopoulos et al., 1996), (el Kassar et al., 1995), (Tsukamoto et al., 1994), (Gilliland et al., 1991b), (Janssen et al., 1990). In MDS subjects furthermore, although some scientists have reported the involvement of T lymphocytes in disease pathogenesis thus suggesting a pluripotent stem cell involvement (Tsukamoto et al., 1993), (Janssen et al., 1989), (Tefferi et al., 1990), others have excluded the uniform involvement of T lymphocytes (Raskind et al., 1984), (Abrahamson et al., 1991), (Lawrence et al., 1987), consistent with a more differentiated stem cell origin. Disparity in findings remains enigmatic.

6.4.3 Carrier detection of X-linked disorders

As a consequence of random Xi, female carriers of X-linked mutant alleles are functional mosaics, with one population of cells expressing the wild type allele and the other the mutant. Although the wild type cells are generally sufficient to spare females from the clinical defects of recessive X-linked mutant alleles, in the hematopoietic department, a unique capacity for cellular compensation exists. For certain X-linked disease alleles, cells expressing the mutant allele fail to mature along a specific developmental pathway, rendering lineage development to be derived from precursor cells expressing the wild type allele. Examples include female carriers of various X-linked immunodeficiencies such as X-linked severely combined immunodeficiency (X-SCID), X-linked agammaglobulinemia (XLA), and Wiskott-Aldrich syndrome (WAS) (see section 7.3.3 for further analysis). As such, identification of a clonal (or skewed) Xi pattern in the pertinent cell lineage(s)

may be used as a diagnostic marker to identify female carriers of X-linked disease alleles (reviewed in (Belmont, 1995)).

6.5 Limitation of X-inactivation based clonality assay

Aside from pathological conditions such as hematologic malignancies and female carriers' of X-linked disease alleles, clonal origin of cells has been observed at a phenomenally high frequency in the general female population. Clinically, it poses a potential limitation to Xi-based clonality assays since it mimics clonal derivation of cells but does not equate with malignancy. For fundamental purposes, recent efforts have attempted to identify and characterize (incidence, implicated cell types, etiologies, biological associations and consequences) of this "frequent/non-pathological" form of clonality.

7 NONRANDOM X-INACTIVATION

7.1 Binomial distribution of X-inactivation ratios

Because of the random nature of Xi, the maternal or paternal X chromosome may be inactivated in any given cell (Gartler and Riggs, 1983). In any given female tissue, the proportion of cells inactivating Xp relative to Xm can be quantitated, deriving the X inactivation ratio (XIR). Upon examination of a large number of healthy females, XIRs tend to follow a binomial or bell-shaped distribution (Fialkow, 1973), (Gale et al., 1991; Gale et al., 1992; Gale et al., 1994) (Naumova et al., 1996a) (Busque et al., 1996). XIRs about the mean represent symmetrical or equal inactivation of the maternal and paternal X chromosome, defined as a random Xi pattern. Extremes of the distribution represent significant departure from equality of inactivation, which can occur in favor of the maternal or paternal X chromosome, defined as preferential Xi, skewed X inactivation or nonrandom X inactivation (a clonal Xi pattern is used in a pathological setting). As skewed Xi was observed in 'healthy' females, a benign/non-pathological origin was speculated.

7.2 Skewed X-inactivation as a discrete trait

The threshold value commonly used to delineate a skewed Xi pattern was arbitrarily assigned to a disproportion of 75%:25% (3:1) or greater between the number of cells that inactivate Xp versus Xm, or vice-versa. The term skewed Xi simply denotes a bias in choice of parental X selected for inactivation (no association with the preferentially active or inactive X is given, unless otherwise stated). This ratio is widely accepted in the literature to delineate clonal derivation of cells (Gale et

al., 1991; Gale et al., 1992), (Gale et al., 1994); (Vogelstein et al., 1987), (Allen et al., 1994), (Busque et al., 1996). Using this threshold criterion, samples can be arbitrarily categorized as skewed or non-skewed. A more stringent criterion for a skewed Xi pattern, i.e., a ratio of ≥10:1, which corresponds to greater than 90% of cells inactivating the same X chromosome, has also been utilized (Gale et al., 1992), (Fey et al., 1994), (Busque et al., 1996), (Plenge et al., 1997). Using the 3:1 criterion, the incidence of skewed X inactivation in the general female population was found to vary from 3.7 to 23% (Fialkow, 1973), (Vogelstein et al., 1987), (Gale et al., 1991), (Gale et al., 1992).

In the absence of methods to precisely quantify the Xi pattern (or XIR), the trait has been traditionally limited to qualitative analyses, where the Xi pattern was dichotomized as skewed or non-skewed. However, with the advancement of reliable quantitative techniques, quantitative trait analyses are amenable.

7.3 Etiologies of skewed X inactivation

The normal distribution of XIRs observed in the general population implies a highly variable trait. Two general and unrelated phenotypes of Xi have been reported. i) The *primary Xi* trait. Initiated during early embryogenesis, it induces Xi in a developmental context, concurrent with tissue differentiation. Anecdotal evidence suggests it induces a relatively similar Xi pattern among various tissues within the individual (i.e., a body-wide Xi pattern). Skewed Xi as a result of the primary Xi trait has been aptly termed primary skewing (PS). Stochastic processes such as a small number of stem cells present when Xi is initiated are believed responsible for this skewed Xi pattern. However, more compatible with the developmental context of initiation of Xi are other etiologies. For instance, genetic influences, such as heterozygosity for the X-linked *Xce* locus, as observed in certain strains of mice, can induce a similar phenotype. ii) The second phenotype, termed secondary skewing, usually associated with a skewed Xi pattern in a cell lineage or tissue-specific manner, occurs after the initiation of Xi. The most widely accepted etiology is presence of an X-linked mutant allele. Examples include female carriers of various X-linked immunodeficiency disease alleles. In particular, skewing is the result of a growth competition between X-linked wild-type versus mutant allele(s). However, another etiology includes clonal derivation of cells [see Table V (page 30) and (Belmont, 1996) and (Migeon, 1998) for review].

Table V. Primary versus secondary etiologies of skewed Xi patterns. Primary skewing has been fundamentally associated with a biased event of X inactivation occurring during early embryological development and associated multiple tissue skewing. In contrast, secondary skewing ensues after the initiation of X inactivation, is normally tissue-specific and may reflect clonal derivation of cells or reflect a growth competition between cells expressing genetic variants of an X-linked gene.

Etiologies of Primary skewing: embryological event	
Stochastic (based on a binomial probability of X inactivation)	- Small number of precursor cells: unequal inactivation of parental X chromosomes - Twinning: reduced number of precursor cells - CPM: reduced number of precursor cells (trisomy rescue) - Early X inactivation: reduced number of precursor cells
Genetic	- *Xce* allele heterozygosity - *XIST*, *TSIX* or possibly *XITE* gene mutations - Imprinting of extraembryonic tissues - Timing of X inactivation gene, may influence number of cell present at initiation of X inactivation event

Etiologies of secondary skewing: occurs after the initiation of random X inactivation	
Stochastic	- clonal hematopoiesis (pre-neoplasia) - clonal dominance secondary to stem cell depletion
Genetic	- growth advantage conferred by an X-linked allele (hemizygous cell selection). Examples include female carriers of X-linked disease alleles.

7.3.1 Etiology of primary skewing: Stochastic event

As the choice of X selected for inactivation at the outset of Xi is speculated to be a random event, the number of primordial cells present at such time will be a strong determinant of Xi ratios (McLaren, 1972), (Fialkow, 1973), (Luzzatto et al., 1979). As such, the variance in the distribution of XIRs will follow a binomial distribution where deviation from the mean is inversely proportional to the number of cells present. It follows that if the number of cells is small, the probability of unequal Xi is statistically large (for example, 37,5% if 4 cells are present). Likewise, if the number of cells is large (e.g., n=16), the probability of nonrandom Xi is small (4,90%). Therefore, by assessing the incidence of skewing in the general female population, several studies have deduced the number of stem cells present at the onset of Xi. Assuming that i) Xi is random, ii) there is no selection of X-linked alleles, and iii) that the variation XIRs is not the consequence of early versus late Xi, estimates for the progenitor cell pool size, at least for the hematopoietic department, have ranged from 4-20 cells (Gandini et al., 1968), (Fialkow, 1973), (Gale et al., 1991), (Puck et al., 1992), (Prchal et al., 1996), (Gale et al., 1997), (Tonon et al., 1998). A progenitor pool size of approximately n=16 cells has been deduced in felines (Abkowitz et al., 1998). Hypothetical mechanisms that can possibly increase this variance, some of which may be genetically influenced, are monozygous (MZ) twinning (Goodship et al., 1996), (Bamforth et al., 1996), confined placenta mosaicism (CPM) (Lau et al., 1997), and earlier X chromosome inactivation (Ingerslev et al., 1989). A plausible mechanism may be a decrease in stem cell precursor size, further increasing the variance and probability of a skewed Xi pattern.

Originally, Xi was speculated to precede tissue differentiation (Nesbitt, 1971), (Nesbitt and Gartler, 1971), implicating a common precursor cell pool from which all tissues are derived, ensuing in common XIRs among multiple tissues within the individual. That Xi patterns are relatively consistent among various tissues within the individual, in both mice and human (vide infra: section 8.2) supports this model. However, more recent data suggests Xi occurs in a developmental context, gradually proceeding in sub-populations and lineages (Monk and Harper, 1979), (Tan et al., 1993). Consequently, according to this model, Xi patterns would be tissue-specific. However, that Xi patterns are body-wide argues in favor of alternative etiologies. For instance a genetic component to acquisition of the primary Xi pattern.

7.3.2 Etiology of primary skewing: a heritable trait

Evidence for a genetic component to derivation of the primary Xi pattern is based on the following: i) the familial clustering of skewed Xi patterns in healthy families (Luzzatto et al., 1979) (Naumova et al., 1996b), and ii) the frequent concordance of a skewed Xi pattern in mother-daughter pairs

(Harris et al., 1992), (Tihy et al., 1994), (Rupert et al., 1995), (Azofeifa et al., 1995), (Parolini et al., 1998), (Orstavik et al., 1999) (and personal observations see Annex 1, page ii).

Genetic elements hitherto identified that may play a role in PS reside within the Xic region. This includes allelic variants of the *Xce* locus and genetic perturbations of the *Xist/XIST* gene. Imprinted inactivation of Xp in extraembryonic tissues has been omitted from this section as it was previously described (section 4.1).

7.3.2.1 Xce allele heterozygosity: a skewed Xi pattern in multiple tissues

In mice, the X chromosome-controlling element (*Xce*), which co-localizes to the *Xic* but a separate genetic element from *Xist* (Simmler et al., 1993), biases the choice of X chromosome selected for inactivation (Cattanach and Isaacson, 1967; Cattanach and Perez, 1970). Although, the molecular correlates responsible for *Xce* allelism have not been fully elucidated (Avner and Heard, 2001), *Xce* allele strength has been correlated with the methylation status of the DXPas34 locus (located 15 Kb downstream of *Xist)*. In fact, DXPas34 is hypermethylated on *Xce* alleles that are more likely to remain active (Courtier et al., 1995). However, methylation analysis of this locus at different stages of embryogenesis has determined differential methylation to be a late event, making it an unlikely candidate for *Xce* allelism (Prissette et al., 2001). More recently however, the *Xite* gene, approximately 40-Kb downstream from *Xist*, was found to regulate Tsix activity *in cis* (Ogawa and Lee, 2003). In fact, deletion of *Xite* downregulated Tsix activity, resulting in a skewed Xi pattern. Further, as molecular / allelic variants of *Xite* were noted, a candidate element for Xce-allelism was speculated. Regardless, at least three *Xce* alleles (*Xce*^{a}, *Xce*^{b}, *Xce*^{c}) and possibly a fourth (*Xce*^{d}) have been identified. Mice heterozygous for the locus are susceptible to skewing, the degree of which dependent on the strength and combination of alleles (XIRs are essentially random in homozygotes). The alleles are classified on a gradient of increasing tendency to remain active: Xce^{d}> *Xce*^{c}> *Xce*^{b}> *Xce*^{a} (Cattanach and Williams, 1972), (West and Chapman, 1978), (Johnston and Cattanach, 1981). For example, in the *Xce*^{a}/ *Xce*^{c} heterozygote, the X chromosome bearing the *Xce*^{c} allele is more likely to remain activate versus the X bearing the Xce^{a} allele, resulting in a skewed Xi pattern. Albeit a Gaussian distribution about the mean (Plenge et al., 2000), thus implying variability in phenotypic expression, the distortion induced by *Xce* allele skewing normally results in an average skew of 20-30%. *Xce* allele strength was originally defined by counting vibrissae number in mice heterozygous for the X-linked mutation Tabby (Cattanach et al., 1969), but more recent assays have examined protein polymorphisms (West and Chapman, 1978), gene expression profiles (Singer-Sam et al., 1992), (Buzin et al., 1994), (Penny et al., 1996), (Plenge et al., 2000) and DNA methylation patterns (Courtier et al., 1995), (Avner et al., 1998).

Cytological staining studies found that Xce X chromosome skewing results from a primary bias in choice of X selected for inactivation rather then a consequence of secondary cell selection (Rastan, 1982b). The XIR induced by Xce allele skewing is relatively consistent among multiple tissues (Johnston and Cattanach, 1981), (Nesbitt, 1971) (Krietsch et al., 1986), (Plenge et al., 2000). In fact, when six different tissues of 18 mice were analyzed, there was ten times more variation between animals then among tissues within any given mouse (Plenge et al., 2000). Further, correlation analyses of XIRs from two tissues in 74 *Xce* heterozygotes derived a correlation coefficient of 0,62. These findings suggest that the XIR, as a result of *Xce* allele heterozygosity, is generally body-wide. Consequently, due to the synteny between the mouse and human X chromosome and correlates of X-linked loci implicated in the Xi process, we speculate an *XCE*-like locus may be operative in humans, although none has yet been clearly identified.

7.3.2.2 XIST/TSIX aberrations as a cause of primary skewing

Mutations in and around the *Xist/Tsix* gene, as discussed in section 3 (choice element), have been associated with a skewed Xi pattern in multiple tissues. Of special interest, a more in-depth analysis of the *Xist* promoter mutation will be discussed here. In two unrelated families, female carrier's (n=9) of a *XIST* promoter mutation (C43G) were associated with preferential inactivation of the mutated X in multiple tissues (Plenge et al., 1997). Paradoxically however, promoter construct analysis found that the mutant promoter was consistently less active versus the wild type promoter. Further, that the *XIST* mutation was detected in only one of 1166 additional X chromosomes analyzed implies a rare mutation. In addition, the mutation was not detected in 32 Duchenne/Becker muscular dystrophy carriers nor in 34 normal control females exhibiting a skewed Xi pattern, suggesting the mutation is not a common cause of skewing and/or the mutation by itself is insufficient to induce skewing (Pereira and Zatz, 1999). Furthermore, that i) carriers of the mutation exhibited variable skewing, ii) non-carrier family members also exhibited a skewed Xi pattern, and iii) that males acquiring the mutation were healthy, suggests that the *XIST* promoter mutation in of itself is insufficient to induce skewing. If XIST polymorphisms are indeed implicated as a primary cause of skewing, they have yet to be identified.

7.3.3 Etiology of Secondary Skewing: selection against X-linked disease allele(s)

After random Xi is initiated, there is potential for natural selection pressures to influence the balance of the two cell populations (reviewed in (Belmont, 1995), (Belmont, 1996), (Puck and Willard, 1998), (Migeon, 1998)). In the case of large X chromosome rearrangement, selection favors the 'least disruptive' gene-dosage event. Thus, cells bearing the normal X in the active state are selectively eliminated (Disteche et al., 1981), (Russell and Cacheiro, 1978). Since negative

selection occurs during embryogenesis, it is associated with multiple tissue skewing. In the event of a deletion on the other hand, cells with the deleted X in the active state are selectively eliminated. Mutations in single genes have also been associated with a skewed Xi pattern, mostly in a lineage/tissue-specific manner, with cells expressing the mutant allele failing to thrive along a specific developmental pathway. Examples include female carriers of X-linked immunodeficiency alleles such as female carriers of the *WASP* gene mutation whom exhibit a skewed XIP among hematopoietic lineages (Prchal et al., 1980), (Fearon et al., 1988) but a random Xi pattern among fibroblasts and oral epithelial cells (Fearon et al., 1988), (Wengler et al., 1995); female carriers of the X-linked agammaglobulinemia gene defect whom exhibit a skewed Xi pattern limited to B lymphocyte (Conley et al., 1986); and carriers of the *X-SCID* gene mutation whom exhibit a skewed Xi pattern limited to lymphocyte (Puck et al., 1987). Other diseases include female carriers of X-linked thrombocytopenia, whom have a skewed Xi pattern in lymphocytes (Saint-Basile et al., 1991) and female carriers of the Lesch-Nyhan gene defect whom have a skewed Xi pattern in erythroid cells (Nyhan et al., 1970), (Albertini and DeMars, 1974). Less frequently however, single-gene disorders may also result in a skewed XIP among multiple tissues, as seen in females with X-linked Dyskeratosis Congenita. These subjects demonstrate a skewed Xi pattern in peripheral blood and possibly in fibroblasts and buccal mucosa (Devriendt et al., 1997), (Vulliamy et al., 1997). Nonetheless, gross X chromosome aberrations and X-linked disease alleles are relatively rare genetic events, therefore a negligible cause for the high incidence of skewed Xi patterns as observed in the general female population.

7.3.3.1 Suggestive evidence that secondary skewing may be incompletely penetrant

Selection against X-linked mutant alleles in female carriers of X-linked immunodeficiency syndromes is usually greater than 95%. However, this may not be the case in "non-immunodeficiency" syndromes, as speculated by (Plenge et al., 1999). In a family originally described for segregating a *DDP* gene mutation (Orstavik et al., 1996), of 8 siblings, all of whom carried the mutation, 7 demonstrated a skewed Xi pattern. However, that one of the carriers exhibited a random Xi pattern argues against secondary cell selection as a cause of skewing. Alternatively, the authors speculated a primary etiology of skewing. They hypothesize that the *DPP* mutant allele was in linkage disequilibrium with an X chromosome skewing allele in 7/8 siblings. In the non-skewed sibling however, a recombination event occurred between the two loci. Alternatively, (Plenge et al., 1999) argues in favor of a secondary etiology of skewing, wherein selection against the mutant allele is not fully penetrant in some subjects. However, that the tissue expression pattern of the *DDP* gene-product has not been fully evaluated, aside from high level expression in fetal and adult brain (Jin et al., 1996), selection against *DDP* mutant allele as a cause of secondary skewing in leukocytes warrants further investigation.

7.3.3.2 Nonhematopoietic tissues are generally robust to secondary skewing

Although the hematopoietic department demonstrates a unique compensatory mechanism for dealing with cells expressing X-linked mutant alleles, anecdotal evidence suggests that nonhematopoietic tissues (NHTs) are less subject to intrinsic factors of skewing such as X-linked mutant allele(s). For example, females with Lesch-Nyhan syndrome, whom exhibit deficiency for the X-linked enzyme hypoxanthine phosphoribosyl transferase (HPRT), exhibit normal HPRT activity in both red blood cells (RBC) and leukocytes but intermediate activity in skin specimens (namely fibroblasts), suggesting that the mutant clone is selectively eliminated in the hematopoietic department (i.e., negative selection occurring at the HSC level) but not in skin (Nyhan et al., 1970). In addition, selection in blood appears to be age-dependant, as the HPRT-deficient clone was detected at low levels (5-10%) in young female heterozygotes (7-17 yo) (Albertini and DeMars, 1974) but undetectable in adult lymphocyte and erythrocyte lineages (McDonald and Kelley, 1972). A similar finding was made in mice: negative selection of the Hprt deficient clone was age-dependant and limited to the hematopoietic department as selection was not seen in NHTs such as skeletal muscle, kidney, liver, lung and brain (Ansell et al., 1991), (Samuel et al., 1993). Further findings were made in female carriers of class I glucose-6-phosphate dehydrogenase (*G6PD)* mutations (Filosa et al., 1996), whom (n=4) demonstrated a skewed Xi pattern in several hematopoietic lineages but a relatively random Xi pattern in NHTs (buccal mucosa cells, urinary tract cells, and hair bulbs).

Lack of a secondary mechanism of skewing operating in NHTs may be explained by various factors. i) NHTs, in contrast to hematopoietic lineages, may reflect lower metabolic/gene activity of the X-linked gene of interest, therefore less prone to selection pressures. ii) Tissue-specific differences (differentiated versus less differentiated) in timing of gene expression. For instance, expression in NHTs may be occurring in more differentiated cell types. As such, the stem cell pool would be unaffected, therefore a continuous source of mutant and wild-type cell types would be provided.

Inasmuch as X-linked mutant alleles are indeed capable of inducing skewed Xi in NHTs, they are exceedingly rare in the population, therefore a negligible cause of NHT skewing in the general female population.

7.4 Familial clustering of skewed Xi: Primary or secondary etiology?

Familial clustering of skewed Xi was reported by several scientists (Ropers et al., 1977), (Orstavik et al., 1996), (Marcus et al., 1992), (Tihy et al., 1994), thus inferring a familial/genetic component. Various examples are given below.

i) In a family segregating a paternally derived Fabry gene, (Ropers et al., 1977), leukocyte samples from 4 siblings carrying the mutation exhibited alpha-galactosidase activity 50% of normal (thus speculative of a random Xi pattern). Another four siblings also carrying the mutation exhibited enzyme activity levels ≤ 20% of normal (thus speculative of a skewed Xi pattern), suggesting 2 genetically heterogeneous subgroups. Since four heterozygotes had 50% of normal enzyme activity and that the vast majority of Fabry heterozygotes have enzyme activity ≥ 30% of normal is supportive evidence against negative selection as a general feature of Fabry carriers. Further, skewing as a result of a stochastic processes of Xi was argued against as this would service a Gaussian distribution of enzyme activity among the siblings. Alternatively, that enzyme activity was positively correlated in different tissues of the same individual is consistent with a primary etiology of Xi pattern derivation. Heterogeneity of enzyme activity among siblings (50% versus 20%) may be best explained by heterogeneity for a primary skewing gene (possibly *XCE*-like). Accordingly, since the disease gene was paternally derived and applying the rules of Xce skewing, pairing of the paternal X (assuming bearer of a strong *XCE*-like allele) with a maternal X of weaker *XCE*-like activity, preferential inactivation of the maternal X is speculated, resulting in low enzyme activity, as seen among four siblings. On the other hand, if paired with a maternal allele of equal *XCE*-like activity, a random Xi pattern is speculated, as seen with the remaining siblings. Of interest however, that there was a high female-to-male ratio among the siblings, one can speculate that the mother may be transmitting a 'more' detrimental X-linked disease allele (Migeon, 1993), causing lethality in males and a skewed Xi pattern in favor of the 'least' disruptive allele in daughters (i.e., the paternal X).

ii) (Orstavik et al., 1996) studied the Xi patterns of a family segregating X-linked recessive deafness syndrome (Mohr-Tranebjaerg syndrome). Observed was a high incidence of skewed Xi in leukocytes of obligate female carriers. One however exhibited a random Xi pattern, a finding arguing against a secondary etiology of skewing. Alternatively, as described above, a primary etiology of skewing was speculated.

iii) (Parolini et al., 1998) identified a female carrier of a paternally derived denovo *WASP* gene mutation whom had the majority of blood cells and oral epithelial cells expressing the mutant allele. Normally however, cells expressing a mutant *WASP* allele are selectively eliminated from

circulation. As her mother also exhibited a skewed Xi pattern, with the maternal X preferentially inactivated in both mother and daughter, two etiologies of skewing were speculated: i) a primary etiology of skewing genetically transmitted from mother to daughter or ii) selection against a 'more' detrimental X-linked mutant allele (Migeon, 1993).

iv) Among a mother-daughter pair carrying of a dystrophin gene deletion, the mother was asymptotic while the daughter was affected with severe Becker muscular dystrophy (Tihy et al., 1994). A probable cause for discordant phenotype was attributed to a discordant Xi pattern: mutant X preferentially inactivated in the mother, while preferentially active in the daughter. Since the DMD mutation is not normally associated with negative cell selection however, a primary etiology of skewing was speculated.

The clustering of skewed Xi patterns within these families and mother-daughter pairs is strong evidence in favor of a familial/genetic component to derivation of Xi patterns. However, because these families also segregate X-linked mutant alleles, the "true" cause of skewing (primary versus secondary) was obscured (reviewed in (Migeon, 1998)). Therefore, to properly identify the formal genetics and the putative loci that influence X chromosome choice, families that do not segregate X-linked mutant alleles should be evaluated.

7.5 Randomly ascertained families to identify the etiology of skewed Xi

In a preliminary study attempting to identify genetic control of X inactivation, Xi patterns of 36 randomly ascertained families were analyzed (Naumova et al., 1996a). One family demonstrated significant aggregation of skewing, where all 7 daughters and the mother of the father exhibited a skewed Xi pattern in lymphocyte samples. Preferential and consistent inactivation of the maternal X among the daughters argues against a normally distributed trait, but compliant with an X-linked pattern of inheritance, plausibly transmitted from paternal grandmother to grandchildren. An XCE-like etiology of skewing was speculated. Small family size however limited genetic linkage studies. Segregation analysis of HUMARA and XIST alleles has excluded linkage to this region. Secondary selection against a maternally derived X-linked mutant allele was excluded on basis of a random Xi pattern in both mother and maternal grandmother. This study has provided preliminary evidence in favor of a genetic component to derivation of Xi patterns in humans.

8 INCIDENCE OF SKEWED Xi IN THE GENERAL FEMALE POPULATION: *variable*

In the past 30 years, the incidence of skewed Xi in the human female population, as estimated by different investigators, has been variable. For example, reanalysis of normal tissue samples (granulocytes, skin, and muscle), as first described by (Fialkow, 1973), but using a discrete criteria of skewed Xi (>75:25), found 13% of samples to demonstrate a skewed Xi pattern. In a later study (Vogelstein et al., 1987), a lower incidence of skewing (3,7%) was reported for normal tissue samples (colonic mucosa and lymphocytes). More recently, the prevalence of skewing in blood and bone marrow samples of hematologically normal females was figured at 22-23% (Gale et al., 1991) (Gale et al., 1992). Similar findings have been reported in other studies (Brown and Brown, 1993), (Harris et al., 1992), (Puck et al., 1992). The disparity in the reported incidence of skewing was initially explained by the diversity of assays used, criteria used to delineate skewed Xi, and small population sizes. However, a more plausible explanation came from studies that analyzed multiple tissues from the same individual.

8.1 Higher incidence of skewed Xi in peripheral blood versus NHTs

In contrast to NHTs such as gastrointestinal mucosa and thyroid which demonstrated a low incidence of skewed Xi (Fey et al., 1992), an incidence of 33% was detected in blood leukocytes (Fey et al., 1994). Similar findings were reported by (Gale et al., 1994), who observed a significantly higher incidence of skewing in blood-derived cells (40%) versus NHTs such as skin (20%), muscle (25%), and colon mucosa (12%). When tissues of different embryonal origin were analyzed (Azofeifa et al., 1996), a higher incidence of skewing was observed in leukocytes (70%) versus NHTs such as muscle (20%), thyroid gland (30%) and suprarenal gland (40%). The tissue-specific incidence of skewing was thought to reflect differences in the timing of Xi (Tan et al., 1993), together with tissue-specific differences in the number of progenitor cells present at the time when Xi is initiated. The higher variance of skewing in blood implied a smaller progenitor pool and/or earlier time of Xi.

8.2 Correlation of XIRs among various tissues

A limited number of studies have investigated the correlation of XIRs among various tissues and cell types within an individual, which includes i) between the various hematopoietic cell lineages, ii) between leukocytes and NHTs, and iii) among various NHTs.

8.2.1 Correlation of XIRs among blood cell lineages: *strong*

Within the hematopoietic department, with the exception of red blood cells (RBC), XIRs are well correlated and relatively similar among the various hematopoietic lineages within the individual (Table VI, page 40), with a correlation coefficient r ranging from 0,97 to 0,99 (Gale et al., 1994), similar to previous findings (Gandini and Gartler, 1969). Strong correlation was an expected finding since these cells are derived from a common progenitor cell (i.e., HSC). Correlation of RBC with leukocytes was modestly weaker (r=0,77 with granulocytes; r=0,66 with lymphocytes), possibly reflecting the influence of extrinsic factors [dietary factors (e.g., iron, fava beans) and altitude] on polymorphic X-linked genetic determinants (G6PD for example) governing RBC kinetics.

8.2.2 Correlation of XIRs among NHTs: *moderate, evidence for a body-wide XIR*

Correlation of XIRs among NHTs, with the exception of muscle-thyroid comparison ($0,46<r<0,64$), revealed a modest correlation coefficient (r), varying from 0,71 to 0,99 (see Table VI, page 40). With few exceptions, these findings suggest that within an individual, XIRs are generally similar from one tissue to the next (i.e., body-wide XIR), consistent with properties of the primary Xi pattern. Similar findings were observed in mice (Nesbitt, 1971), (Johnston and Cattanach, 1981), (Krietsch et al., 1986), (Plenge et al., 2000).

8.2.3 Correlation of XIRs between blood cells and NHTs: *low*

Among the various NHT-blood cell lineage comparisons, extreme variation in correlation was observed: $0,19 < r < 0,99$. Weakest correlation was observed between thyroid tissue and leukocytes ($0,19<r<0,24$) and between skin and RBC (r=0,34) (Table VI, page 40). Lack of significant correlation between RBCs and a NHT (hair roots) was previously demonstrated (Gartler et al., 1969), as well as weak correlation between polymorphonuclear cells (PMN) and hair bulbs, r=0,31 (Tonon et al., 1998). High variability in correlation may be explained by the higher incidence of skewed Xi in PB (section 8.1).

Table VI. Correlation of XIR of different tissues within the same individual

Comparison	Investigator			Summary
	Gale (94)	*Fialkow (73)*	*Azofeifa (96)*	
Blood - Blood	gran - (mono/B cells) 0,99 (n=26) gran - T cells 0,97 (n=25)	gran - lympho 0,97 (n=20) gran - RBC 0,77 (n=42) RBC - lympho 0,66 (n=20)		0.66 - 0.99
NHT - NHT	skin - muscle 0,79 (n=20) skin - colon 0,71 (n=9) muscle - colon 0,94 (n=9)		muscle-muscle 0,88 muscle - thyr 0,46 - 0,64 muscle - supra 0,85 - 0,95 thyr - thyr 0,89 thyr - supra 0,88 - 0,99 supra - supra 0,94	0,46 - 0,99
NHT - Blood	skin - gran 0,74 (n=20) muscle - gran 0,63 (n=20) colon - gran 0,69 (n=9)	skin - gran 0,88 (n=18) muscle - gran 0,99 (n=5) skin - lympho 0,89 (n=11) skin - RBC 0,34 (n=18) muscle - RBC 0,99 (n=5)	muscle - leuko 0,61- 0,73 thyr - leuko 0,19 - 0,24 supra - leuko 0,46 - 0,51	0,19 - 0,99

Note. Abbreviations: gran, granulocytes; mono, monocytes; lympho, lymphocytes; RBC, red blood cells; thyr, thyroid gland; supra, suprarenal gland; leuko, leukocytes

Concordance of Xi patterns between blood and NHTs: evidence of a body-wide Xi pattern

Skewing in a NHT was generally associated with a skewed Xi pattern in blood (i.e., 100% concordance for muscle-granulocyte comparison) (Gale et al., 1994). A skewed Xi pattern in blood on the other hand was less often associated with skewing in a NHTs (i.e., 45% concordance for blood-muscle comparison). Similar results were observed with skin-blood and colon-blood comparisons (Gale et al., 1994) and by other scientists (Fialkow, 1973), (Azofeifa et al., 1996), (Buller et al., 1999). These findings suggest Xi patterns are generally body-wide, consistent with the properties of the primary Xi pattern. The higher incidence of skewing may be unrelated to skewing in NHTs.

8.3 Suggestive evidence for an *XCE*-like X chromosome skewing trait

Findings of a body-wide XIR/Xi pattern imply a common mechanism/etiology of Xi pattern derivation among various tissues within an individual. And although recent data suggests Xi proceeds with different schedules in different tissues (Tan et al., 1993), strong correlation of XIRs among multiple NHTs argues against a stochastic etiology of Xi pattern derivation as this would foster tissue-specific XIRs. A secondary etiology of skewing is also unlikely as X-linked mutant alleles are exceedingly rare and do not generally induce skewing in NHTs (section 7.3.3.2). Rather, i) that strong correlation of XIRs among multiple tissues in mice corresponds to a primary etiology of XIR derivation, i.e., *Xce* allelism, and ii) due to synteny of Xi-related elements between human and mouse (Xic, Xist, Tsix), we speculate that the Xi pattern of a human NHT reflects zygosity for an *XCE*-like locus. No conclusive evidence, albeit suggestive (Naumova et al., 1998), for a human *XCE*-correlate has been identified.

Identification of the genetic element(s) implicated in derivation of the primary Xi pattern may shed light on the molecular elements implicated in the Xi pathway and/or the choice step of Xi. Further, as PS results in preferential expression of a parental X chromosome (similar to imprinting), the clinical consequences (fecundity, disease susceptibility) warrants investigation.

9 ANALYSIS OF BLOOD Xi PATTERNS AS A FUNCTION OF AGE

Unequivocal insight into the higher incidence of skewed Xi patterns in blood versus NHTs was provided by the examination of XIRs as a function of age. Preliminary data suggested a significantly higher incidence of extremely skewed Xi patterns in leukocytes of elderly females (aged 75-96 years) versus healthy children (aged 2-8 years) and younger females (aged 20-58 years)

(Fey et al., 1994). However, given the lack of reliability and ambiguity in interpreting results based on the M27β clonality assay (Fey et al., 1994), we conducted a similar study based on a larger sample size utilizing a more reliable clonality assay, i.e., HUMARA. This study comprised the core of my Master's research project entitled: Clonality Analysis of Blood-Derived Cells in Healthy Females: Evidence for Age-Dependent Increase in Clonal Hematopoiesis. In total, 295 normal females cross-sectioned into three different age groups were recruited. We found that the incidence of skewed Xi increased significantly with advancing age: 8,6% in cord blood of neonates, 16,4% in peripheral blood (PB) of 28-32 year olds, and 37,9% in PB of females 60 years of age and older (Busque et al., 1996). The phenomenon (trait) was termed acquired skewing (AS) and represents significant departure from random Xi in blood cells with advancing age. Although the increased incidence of skewing in the 28-32 year old (yo) group was marked, it was not significantly higher versus neonates. The incidence of skewing in the elderly group (>60 yo) was significantly higher versus neonates ($p<0,0001$) and the 28-32 yo group ($p=0,0064$).

9.1 AS: *confirmation studies*

The increased incidence of skewed Xi with advancing age has recently been confirmed. In one study the incidence of skewed Xi in the polymorphonuclear (PMN) fraction of cord blood was 8.3% (Tonon et al., 1998), while in a second study 26% (Gale et al., 1997). The discrepancy may be due to the small sample size analyzed by Gale (n=23) versus n=36 (Tonon et al., 1998) and n=162 (Busque et al., 1996). In middle age females (17-50 yo), the incidence of skewed Xi in the PMN fraction of PB was 16.7% in one study (Tonon et al., 1998) and 22% when total white blood cell/neutrophil fractions were analyzed (Gale et al., 1997). Both are similar to the 16.4% Xi skewing incidence reported by (Busque et al., 1996) for females aged 28-32 yo. When elderly females were analyzed (75 yo and older), results have been relatively consistent, with the incidence of skewed Xi ranging from 35-56% (Gale et al., 1997), (Tonon et al., 1998), (Christensen et al., 2000). Remarkably, when centenarians were analyzed (Christensen et al., 2000), 67% of samples demonstrated a skewed Xi pattern.

That the incidence of skewed Xi reached 67% in centenarians ($p<0,01$ versus 73-93 yo), implies that the AS trait gradually increases with advancing age. However, that the variation in XIRs in sequential analyses, 24 months apart (Tonon et al., 1998) or longer (Prchal et al., 1996), were not significantly different, suggests that XIRs vary very slowly. In a further study, XIR analysis during an 18-month period of three hematopoietic lineages (granulocytes, monocytes and T cells) derived from healthy females of varying age groups (young, middle-aged and elderly) detected no significant fluctuation of XIRs (van Dijk et al., 2002). These results argue against the hypothesis

that AS reflects the stochastic contribution of a small number of emerging stem cells from the inactive HSC pool.

9.2 Clarifications in light of the AS phenomenon

In light of the AS trait, ambiguities raised in prior studies may be better clarified.

i) AS plausibly represents the major confounding variable for discrepancies in the incidence of skewed Xi observed in the human female population (section 7.2), thus invoking age as an important variable in skewed Xi studies.

ii) The large range in estimates of stem cell precursor size (section 7.3.1) may be invariably explained by the AS trait. In effect, estimate of primordial pool size should be mathematically dependent on age: as the incidence of skewed Xi increases with advancing age, size estimation of the primordial pool should decrease. As such, estimates should be based on neonatal samples (Tonon et al., 1998).

iii) Since AS mimics clonal derivation of cells, ambiguity in the stem cell origin (undifferentiated versus more differentiated) for various MPDs (section 6.4.2) may be best explained by the AS trait. For instance, in absence of AS, a skewed Xi pattern may be limited to pathological lineage(s) only. However, in the event of AS, a skewed Xi pattern may appear in additional lineages, thus implying a more undifferentiated stem cell origin. Thus, age should be considered a confounding variable in clonality studies.

iv) A smaller progenitor pool size ascribed to the hematopoietic department (HD) versus NHTs (section 8.1) may be best explained by the AS trait. Consequently, and confounding for age, the precursor pool size for the HD may not differ significantly from NHTs.

v) The low correlation of XIRs observed between blood and NHTs (section 8.2.3) may be reasonably explained by AS trait. In the absence of AS, as in the case of young females, XIRs among various tissues should be similar. In the event of AS however, as in older females, XIR correlation between blood and NHTs should be lower.

10 ETIOLOGY OF SKEWED Xi IN CORD BLOOD: *primary skewing trait*

To speculate on the cause of skewed Xi patterns in cord blood samples, the following was considered. Foremost, it is unlikely to reflect a secondary etiology of skewing since X-linked disease alleles are relatively rare. Moreover, it is unlikely to reflect AS since it is a late-onset trait. Rather, that XIRs and Xi patterns are generally body-wide, a primary etiology of Xi pattern derivation (*XCE*-like) was speculated. Consequently, we hypothesize that in neonates, XIRs of various tissues be well correlated, warranting tissue-correlation studies for confirmation.

11 PLAUSIBLE ETIOLOGIES OF ACQUIRED SKEWING

Plausible etiologies of AS are discussed below, these include both stochastic (clonal derivation, stem cell depletion) and genetic (X-linked polymorphisms). AS is unlikely to reflect selection against X-linked disease alleles as these are relatively rare.

11.1 Stochastic event: *clonal dominance secondary to stem cell depletion*

Various estimates on the number of HSC contributing to hematopoiesis have been made, the majority suggesting that the hematopoietic department has an enormous reserve potential, with HSC depletion plausibly prevented by relatively few stem cells being used at anyone time, i.e., clonal succession model (Kay H.E.M, 1965). That HSCs are generally quiescent supports this model. In fact, in mice, approximately 5% of HSC are believed to contribute to the maintenance of hematopoiesis at any one point in time, with 99% estimated to enter cell cycle once every 57 days (Cheshier et al., 1999). Alternatively, others theorize that the HSC pool may be small but capable of continuous self-renewal. In effect, regulation of the stem cell pool requires further analysis and is of major therapeutic importance if manipulation of HSCs is warranted.

Preliminary insight into HSC behavior was provided in female Safari cats undergoing bone marrow transplantation with a limited number of autologous marrow cells. As these cats were heterozygous for the G6PD locus, Xi pattern analysis was possible. While untransplanted cats retained polyclonal hematopoiesis, transplanted cats developed a skewed Xi pattern, demonstrating that small doses of marrow cells can induce a skewed Xi pattern (Abkowitz et al., 1995) (Abkowitz et al., 1998). Moreover, these studies were the basis for mathematical modeling showing that clonal dominance can occur simply by chance when the number of hematopoietic stem cells (HSC) is small (Abkowitz et al., 1996). Consequently, assuming HSC decisions are stochastic and implementing predefined rates of HSC self-renewal, differentiation and apoptosis, simulation studies predicted

that small differences in HSC kinetics can generate the observed drift in G6PD phenotype (Catlin et al., 1997).

As a human correlate, stochastic depletion of HSC with advancing age and random differentiation of the few residual HSC has been a plausible mechanism of AS (Gale et al., 1997). In support of this view, by age 70, hematopoietic cellularity of the iliac crest in bone marrow is reduced to approximately 30% relative to that of young adults (Gilleece and Dexter, 1993). Similar findings were recently obtained, but only in much older females (80-90 years old) (Ogawa et al., 2000). A trend towards a decreased absolute number of hematopoietic progenitors and/or CD34+ stem cells in PB and bone marrow of elderly people has also been observed (Egusa et al., 1998), (Bagnara et al., 2000), (Hirota et al., 1988), (Nilsson-Ehle et al., 1995). However, the ability of elderly mouse marrow to support serial transplantation is the same as that of young mice (Harrison and Astle, 1982), and HSC from bone marrow of elderly mice are five times more frequent than those from young mice, albeit a reduced homing capacity (Morrison et al., 1996). Furthermore, in human recipients of allograft bone marrow transplantation, whom receive a relatively small number of the donor's HSC, hematopoiesis is stable and polyclonal (Nash et al., 1988), (Turhan et al., 1989), (Saunders et al., 1995), (Tonon et al., 1998), (Mathioudakis et al., 2000), suggesting that severe HSC depletion, well above that which occurs in transplanted patients, is an unfounded cause of AS. To ultimately rule out stem cell depletion as a putative cause of AS, the number of HSCs present in skewed versus non-skewed females should be investigated. However, since HSCs are infrequent (<1 in 10^6 nucleated marrow cells) and difficult to isolate by physical or immunological methods (Wang et al., 1997), analysis may prove difficult.

11.2 Stochastic event: *clonal "pre-neoplasia" hematopoiesis*

According to the multi-hit model of neoplasia, a number of random genetic hits are required for neoplastic transformation. In theory, a limited number of genetic mutations (autosomal and/or X-linked) acquired by a HSC may provide the cell with a non-pathological but measurable growth advantage, allowing it to dominate over polyclonal hemopoiesis, thus resulting in a clonal/skewed Xi pattern. An argument in favor of this hypothesis is the increased incidence of hematological malignancies with age, particularly for diseases such as MDS and AML (Oscier, 1997), (Sandler, 1987). However, the rarity of hematologic malignancies, incidences of less than 0,01% for even the most common, makes this explanation favorably unlikely. As indirect support, women with a skewed Xi pattern have normal blood counts (Busque et al., 1996). Moreover, when a group of women exhibiting AS were prospectively examined for several years, no evidence for increased occurrence of hematologic malignancy was observed (Tonon et al., 1998). Although the frequency

of clonal populations of lymphocytes in healthy females increases with advancing age (Posnett et al., 1994), clonal populations of T and B lymphocytes did not differ between elderly skewed and elderly non-skewed females (Annex 2, page xviii), ruling out lymphocyte clonality as a probable cause of AS. Nonetheless, to directly implicate clonal hematopoiesis as a probable cause of AS, one must identify genetic mutations/aberrations directly responsible for pre-neoplastic growth.

11.3 Continuation of the PS trait

Although the AS trait is late-onset and that of PS early-onset, thus implying distinct traits, that both are potentially X-linked theoretically implies that they may be one and the same. Because PS normally results in a maximal skewing of 20-30%, it is possible that AS represents a continuation (or completion) of the PS trait. If true, we postulate that direction of AS (in favor of Xp or Xm) should occur in favor of the same parental X preferentially inactivated in the primary Xi trait. Analysis of the direction of AS as a function of the primary X pattern should provide insight.

11.4 Genetic: *X-linked hemizygous cell selection*

Since the hematopoietic system is mitotically active (producing 10^{11} cells on a daily basis), thus particularly sensitive to genetic polymorphisms affecting cell growth, we speculate that AS may result from a relative growth competition between cells expressing genetic variants of an X-linked gene.

12 X-LINKED HEMIZYGOUS CELL SELECTION ETIOLOGY OF AS

12.1 AS in felines

Insight into the genetic basis of AS was provided by Xi analysis of aging female Safari cats - the F1 generation derived from crossing the domestic cat with the Geoffroy cat. Since the parental strains have evolved independently for 12 million years, it is highly probable that the F1 generation will be heterozygous for several genetic loci. In particular, the female Safari cats were heterozygous for the *G6PD* locus (with the G allele derived from the Geoffroy cat and the d allele from the domestic cat), thus informative for Xi-based assays. In young cats (2-3 months), balanced hematopoiesis (i.e., random Xi) (equal number of cells expressing the G and d allele) was observed. With advancing age however, and in absence of any pathological consequence, there was an increasing percentage of blood cells expressing the G allele. By 4-6 years of age, 67% of cats exhibited skewed Xi ($\geq$ 75% expression of a parental-specific X chromosome), with the X derived from the Geoffroy cat

preferentially active. A primary etiology of skewing was rejected as hematopoiesis was initially polyclonal. Alternatively, a relative growth advantage conferred by an X-linked allele was speculated (Abkowitz et al., 1998). These findings argue against both the stem cell depletion and clonal "pre-neoplasia" etiologies of AS since skewed Xi would have occurred in favor of either parental X chromosome when F1 felines were examined. An imprinted autosomal basis is also unlikely as an imprinted gene can be paired with either parental X chromosome. An imprinted X-linked genetic basis however, could not be ruled out as parental derivation of G6PD alleles (i.e., Xp versus Xm) was not investigated.

12.1.1 Hematopoietic stem cell origin of AS

Since hematopoiesis can be partially defined as differentiation of pluripotent stem cells into various differentiated hematopoietic lineages (Till and McCulloch, 1980), identifying the hematopoietic lineages implicated in AS may provide insight into the cell origin of AS.

The contribution of lymphocytes to the AS trait has been contradictory. In our initial study (Busque et al., 1996), as a majority of females were extremely skewed (XIR≥10:1), we speculated that both myeloid and lymphoid lineages contributed to the AS trait, consistent with a HSC origin of AS. However, when cell lineages were analyzed independently, significantly more Xi skewing was observed in granulocytes (46%) versus T lymphocytes (19%) of elderly females (63-95 years), suggesting AS differently affects the granulocyte and T lymphocyte lineage ($p<0,01$) (Champion et al., 1997). Similar findings were reported by (Gale et al., 1997) who observed a skewed Xi incidence of 56% in neutrophils of elderly females (aged 75 years and older) versus 40% in T lymphocytes. When a more stringent criterion of skewed Xi was used (> 90% expression of one allele), 33% of females exhibited a skewed Xi pattern in neutrophils versus 9% in T lymphocytes. In a further study, although the overall correlation of XIRs between PMN and T cells was significant ($r=0,66$), correlation was stronger in younger females (25-32) ($r=0,77$) versus older females (>75 years) ($r=0,57$) (Tonon et al., 1998). These findings suggest that relative to T lymphocytes, granulocytes/PMN are more affected by AS. Lack of T cell contribution may reflect longevity of lymphocytes (Gale et al., 1997). Alternatively, it may reflect restriction of the AS process to the granulocyte/myeloid lineage.

Preliminary insight into the plausible HSC origin of AS was derived by analysis of autologous BM transplantation studies in female Safari cats (Abkowitz et al., 1998). When individual hematopoietic lineages were analyzed, Xi skewing was detected in progenitor cells, RBC, and granulocytes. T lymphocytes, demonstrated a lack of Xi skewing, postulated to reflect either longevity of T cells

and/or restriction of the AS phenomenon to the myeloid lineage. For clarification, autologous transplantations were performed. Among elderly cats established as having AS, although sample size was small (n=2), BM was harvested and reserved for autologous infusion. The cats subsequently received myeloablative therapy and re-infused with the harvested BM. Upon Xi analysis 3-20 months post-BMT, a skewed Xi pattern was detected in all lineages analyzed: progenitor cells, granulocytes and T lymphocytes. The latter was consistent with derivation from a common precursor, an argument in favor of a HSC origin of AS. These results corroborate that T cells are less likely to contribute to the AS trait due to their long circulatory-half lives. Anecdotal evidence from a pilot study also demonstrated a HSC origin of skewing (Annex 3, page xxii).

12.2 X-linked genetic basis of the AS trait extended to humans

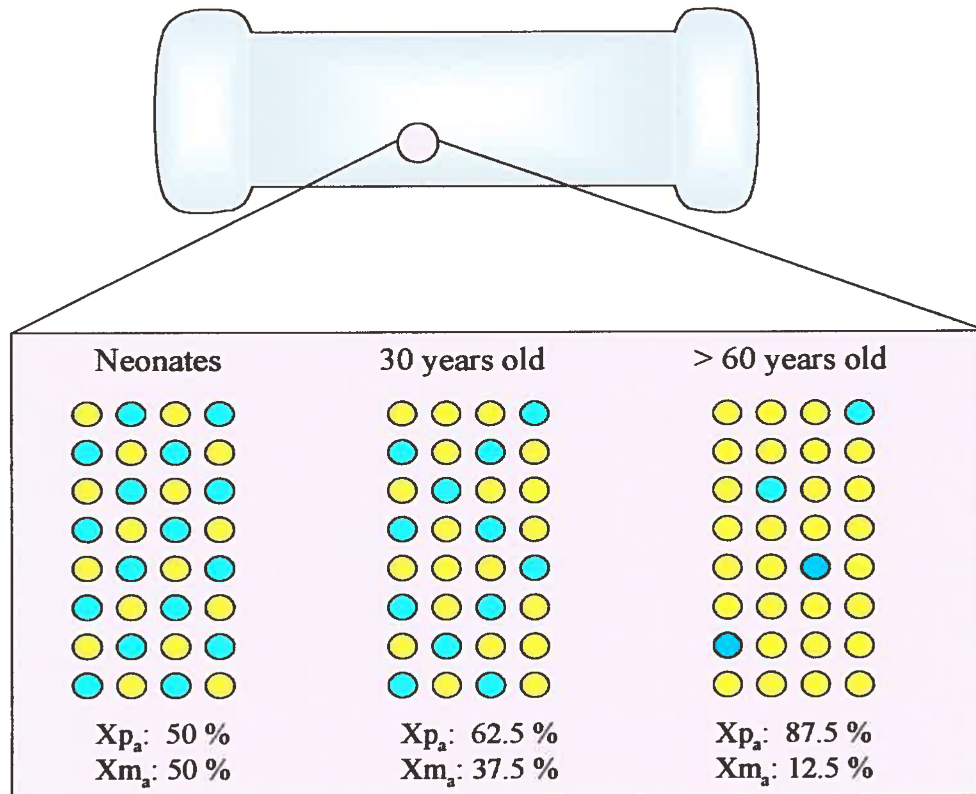
The finding of an AS-like trait in felines has been instrumental is assigning an X-linked genetic basis to the AS trait in humans. Due to the resemblance of various Xi skewing features between felines and humans (e.g., age-related penetrance and exclusion of T lymphocytes), we speculate an X-linked genetic basis to the AS trait in humans. We speculate that in female carrier's of the X-linked AS allele (i.e., heterozygous for the AS gene), HSCs bearing a strong AS allele (on the active X) will out-compete HSCs bearing a weak AS allele (on the active X), manifesting as AS. In theory, the X-linked gene can modify HSC kinetics such as replication, differentiation and/or apoptosis. A model of the AS trait is presented in Figure 5 (page 49). As observed, XIRs in neonates are initially random (reflective of the primary Xi pattern). With advancing age however, HSC bearing the stronger X-linked AS allele outgrow cells expressing the weaker allele, significantly dominating in cell number by age 60 and older.

12.3 Twin studies support an X-linked genetic basis to AS in humans

12.3.1 Correlation of XIRs in elderly MZ twin pairs

Since MZ twins are essentially genetically identical, a genetically influenced trait should be well correlated within the pair. As such, PB was collected from 71 MZ twin pairs aged 73 to 93 yo and XIRs determined. 35% of the subjects exhibited a skewed Xi pattern. Moreover, twin pairs demonstrated preferential inactivation of the same parental X chromosome (r=0,57), evidence in favor of an X-linked genetic component (Christensen et al., 2000). If stochastic or autosomal etiologies were implicated, little or no correlation in preferential inactivation of a particular parental

Figure 5. X-linked genetic hemizygous theory of AS. Diagrammatic representation of a typical bone marrow where hematopoiesis occurs. Blow-up of the circular object depicts HSC expressing either the maternal or paternal X chromosome. Yellow cells - Xp active, blue cells - Xm active. In neonates, X inactivation is initially random. With advancing age however, dominance of one cell type occurs. By age 30 and 60 yo and older, HSCs expressing the "stronger" X-linked AS allele (Xp in this example) significantly outnumber HSCs expressing the "weaker" AS allele. For this hypothetical example, XIRs and Xi-skewing in favor of the paternal X were arbitrarily assigned.



X chromosome would have been observed. A limitation however, is that a proportion MZ twins may share a common blood supply during intra-uterine growth, thus possibly inflating the correlation coefficient.

12.3.2 Genetic contribution to the AS trait

Classical twin studies offer a dynamic model in which to quantify the contribution of genetic factors to human trait variation. Correlation of phenotype is compared among MZ twins, whom are essentially genetically identical, versus dizygous (DZ) twins whom on average share 50% of their autosomal material and 75% of their X-linked genetic material. Environmental influences are not considered since MZ and DZ twin pairs hypothetically share similar environments. Consequently, if a trait is genetically influenced, correlation should be stronger among MZ versus DZ twin pairs. To quantify the contribution of genetic factors in the etiology of AS, the XIR (measured as a function of Xp inactivated relative to Xm) of 29 MZ and 18 DZ twin pairs were studied (Vickers et al., 2001). Mean ages of twins were 61 and 63 years, respectively. Among the MZ twin pairs, the intraclass correlation of allele ratios was 0,53 and 0,26 for granulocytes and T cells respectively. Among the DZ twin pairs, the intraclass correlation of allele ratios was 0,19 and 0,36 for granulocytes and T cells respectively, deriving a heritability score [$h^2 = 2(r_{MZ} - r_{DZ})$] of 0,68 for granulocytes and zero for T cells, supporting a modest genetic (X-linked) component to derivation of XIRs in granulocyte. However, results may be artifactual since a proportion of MZ twins share a common blood supply during intra-uterine growth, thus artifactually increasing correlation of XIRs. It is not clear why T cell XIR correlation was stronger in DZ versus MZ twin pairs, but significant variability in correlation coefficients due to smaller sample size is speculated. That genetic factor(s) explain only 68% of the variance of XIRs in granulocytes suggests additional factors may play a role in derivation of blood XIR: clonal disorders, environmental influences, stochastic events and experimental error.

12.4 Intrinsic versus extrinsic factors in the etiology of AS

That environmental factors possibly influence XIRs has been previously documented. In South African females heterozygous for X-linked G6PD deficiency, found were random XIRs in youths (thus half the cells deficient for G6PD activity and the other half full enzyme activity). Despite the hematopathology associated with G6PD deficient cells, they are relatively resistant to malaria (probably owing to impaired entry and growth of the parasites during erythropoiesis), therefore selectively retained in circulation (Luzzatto et al., 1969). After puberty however, as females acquire a stronger immune system, immune-related anti-malarial activity is upregulated, thus eliminating malarial parasites from circulation. Consequently, in absence of the stimulus (malaria), cells

expressing the G6PD (-) allele, whom are at a relative growth disadvantage versus G6PD (+) cells, are effectively eliminated from circulation (Hitzeroth and Bender, 1981). The study suggests women fully exploit polymorphic X-linked genes. For instance, in the presence of a particular external influence(s), cells bearing an X-linked allele which haphazardly provides it with a growth disadvantage is counter-balanced by preferential destruction (by malarial parasites) of cells bearing the stronger allele. Alternatively, in absence of the extrinsic factor, cell survival kinetics are based solely on intrinsic factor(s): X-linked allele(s) which can provide the cell with a relative growth advantage.

In a similar study, the G6PD phenotype was analyzed in a group of 77 Sardinian females heterozygous for G6PD deficiency. It was found that the majority of females had an excess of G6PD (+) cells in their blood (Rinaldi et al., 1976). High susceptibility of G6PD (-) RBCs to hemolytic noxae was the most likely explanation. This was further supported by the presence of mild noncompensated anemia in heterozygotes. However, as not all females were skewed in favor of the normal (+) allele, but rather some in favor of the deficient (-) allele, it was speculated that in some instances, the G6PD (-) allele may have a relative growth advantage. As mentioned above, G6PD (-) cells are retained/selected in young females with active malarial parasitemia. However, other causative agents are possible.

Thus, the G6PD locus offers a dynamic model to illustrate the intricacies of the AS trait. It introduces the mechanism by which both intrinsic and extrinsic factors influence preferential dominance of X-linked alleles, thus implying a multifactorial trait where genetic, environment and gene-environment interactions are plausible. Thus we predict that both intrinsic (X-linked polymorphisms) and extrinsic factors may play a role in the etiology of AS. Possible external/environmental factor(s) includes infectious diseases, the BM microenvironment, medicinal products, allergens, pollutants, altitude, diet, and certain lifestyle factors (smoking habits, psychological state, physical activity).

12.5 Molecular characteristics of the AS trait: *Genetic model*

Assuming i) that AS occurs in females who are heterozygous for an AS gene, ii) that females heterozygous for the AS gene are not at a viable disadvantage and iii) that the AS gene is fully penetrant, a two-allele model for the AS gene would predict a maximum incidence of skewed Xi of 50%. That 67% of centenarians demonstrated a skewed Xi pattern argues against such a model. This implies i) if limited to a single locus, there are more then two alleles; ii) genetic heterogeneity; and/or iii) a polygenic trait (additive or multiplicative epistasis model).

The X-linked genetic variation / polymorphism may consist of a single base-pair substitution or may vary in length of a simple sequence repeat. It is however unlikely to be a gross genetic alteration or a deleterious X-linked mutation since these are i) relatively rare, ii) normally manifest as a clinical condition in female carriers and iii) lethal in males, inducing a high female-to-male ratio distortion among offspring. The polymorphism may reside within coding regions and/or regulatory elements. The genetic variant may not be limited to the DNA sequence as epigenetic modification(s) are also plausible.

Aside from clonal derivation, the AS trait does not preclude the involvement of autosomal factors, the limiting factor however, would have to be heterozygosity for an X-linked genetic polymorphism as homozygosity would equally favor the parental X's, resulting in a random Xi pattern. X-linked genetic candidate elements include: a gene promoter region(s), transcription factor, enzyme, signal transduction molecule, cell surface receptor, cell membrane channel, and/or structural protein. Linkage analysis of families segregating the trait is one method of identifying potential candidate gene(s) (Naumova et al., 1998). However, a candidate gene approach is also feasible (Plenge et al., 1997).

13 CLINICAL ASSOCIATIONS OF SKEWED Xi

Although the vast majority of females with skewed Xi appear physiologically normal, various pathophysiological associations have been cited. Whether the clinical association is a consequence or a cause of skewing, or associated with a primary or secondary etiology of skewing is sometimes discussed.

13.1 Skewing and expression of X-linked disease allele(s)

In the "unfortunate" event of a skewed Xi pattern in favor of the X bearing the mutant allele preferentially active, female carriers of recessive X-linked disease alleles can manifest X-linked recessive traits (Tihy et al., 1994), (Aral et al., 1996), (Favier et al., 2000), (Orstavik et al., 1999), (Azofeifa et al., 1995). As secondary skewing normally reflects negative selection of the mutant clone, these females are likely to reflect a primary etiology of Xi skewing, or selection against a 'more' detrimental X-linked disease allele (Migeon, 1993). Moreover, based on observation that some obligate carriers of various X-linked hematological disease alleles exhibit late-onset disease symptoms, a role for AS in decreasing expression is speculated. This list includes X-linked sideroblastic anemia (Cazzola et al., 2000), G6PD deficiency (Beutler et al., 1996), X-linked

agammaglobulinemia (Aruffo et al., 1994), X-linked SCID (Smith and Notarangelo, 1997), and X-linked chronic granulomatous disease (Cazzola et al., 1985), (Rosen-Wolff et al., 2001). Surprisingly, hematopoietic cells with the X chromosome expressing the mutant allele were preferentially selected over their wild type counter-part, suggesting i) that the AS gene is a stronger determinant of Xi patterns over X-linked mutant alleles or ii) expression of the AS allele occurs in a less differentiated HSC versus the an X-linked mutant allele.

13.2 Skewing and RSA

Versus controls, a high frequency of skewed Xi was observed among females' experiencing recurrent spontaneous abortions (RSA). RSA has been defined as 2-3 or more consecutive losses under 20 weeks gestation, and estimated to affect 1-2% of couples wishing to bear children (Pegoraro et al., 1997), (Sangha et al., 1999), (Lanasa et al., 1999). A plausible etiology is transmission of a lethal X-linked allele inducing both skewed Xi in the female carrier (wild-type X preferentially active) and the spontaneous abortion of a male conceptus (Lanasa and Hogge, 2000).

13.3 Skewing and susceptibility to ovarian/breast cancer

Klinefelter males (XXY genotype) comprise a unique set of individuals in which to analyze the role of the extra X chromosome and its associated gene products in promoting or protecting disease pathogenesis. An increased incidence of autoimmune diseases (rheumatoid arthritis and systemic lupus erythematosus) and breast carcinoma were noted among Klinefelter subjects, suggesting that a gene derived from the extra X chromosome (possibly escaping Xi) may increase disease susceptibility (Bandmann et al., 1984). Moreover, versus controls, a higher incidence of skewed Xi has been observed in leukocytes of females experiencing invasive ovarian cancer and in ovarian cancer patients carrying a germline BRCA1 mutation (Buller et al., 1999), however a cause and effect relationship could not be determined. Moreover, because of the AS trait, lack of proper age-matched controls may have confounded results. Nonetheless, that BRCA1 associates with the inactive X chromosome (Ganesan et al., 2002) supports a relationship between Xi and breast cancer susceptibility. Moreover, frequent loss of heterozygosity of the X chromosome in ovarian cancer patients (Osborne and Leech, 1994) suggests that an unidentified tumor suppressor gene (TSG) may reside on the X chromosome (Cheng et al., 1996). The AR gene is a candidate locus since mutations have been associated with male breast cancer (Lobaccaro et al., 1993). Moreover, length of the CAG polymorphism in exon 1 of the *AR* gene has been controversially associated with female breast cancer development. Although some studies associate aggressive forms of breast cancer with shorter CAG repeat lengths (Yu et al., 2000), other studies, versus age-matched controls, found no

association between CAG repeat length and early-onset breast cancer (Spurdle et al., 1999), (Kristiansen et al., 2002). However, among BRCA1 mutation carriers, longer repeat lengths were associated with increased risk of developing early-age breast cancer (Rebbeck et al., 1999). In spite of these conflicting results, the relationship between skewing (whether primary or secondary) and ovarian/breast cancer susceptibility merits further investigation.

13.4 Skewing and predisposition to autoimmunity

Compared to males, females have enhanced immuno-reactivity, which although protects them from many types of infections, has also been associated with increased susceptibility to autoimmune disorders. For example, females demonstrate an increased prevalence of systemic lupus erythematosus (SLE), multiple sclerosis (MS), and rheumatoid arthritis (RA), suggesting disease susceptibility alleles are sex-influenced, developmentally and/or hormonally regulated. Estrogens for example, which possibly elicit an immuno-stimulatory effect, have been associated with disease exacerbation (Cooper et al., 1998). Androgens on the other hand, have been linked with a decrease in autoimmune activity (Cutolo and Masi, 1998). Alternatively, the XX genotype of females may pose a disease risk factor. Since several genes escape Xi, dosage imbalance for any one of these genes may increase disease susceptibility. Similarly, a protective role for Y-linked genes in males, *SRY* for example, is another possibility.

Moreover, an association between Xi patterns and autoimmunity was speculated. In one study, the authors, in combination with their own data, conducted a survey of the literature and found that female carriers of X-linked immunodeficiency diseases normally associated with random Xi, i.e., X-linked lymphoproliferative syndrome (XLP), X-linked Hyper-IgM (XLHM), and X-linked granulomatous disease (XCGD), had clear features of autoimmunity versus immunodeficiency diseases normally associated with skewed Xi patterns, i.e., agammaglobulinemia (XLA) and X-linked severe combined immunodeficiency (X-SCID). These results suggest that a skewed Xi pattern protects females from an autoimmune-related phenotype (Martin-Villa et al., 1999), possibly owing to negative selection of the mutant allele. Alternatively, theoretical data suggests that a skewed Xi pattern may pose a risk factor for susceptibility to autoimmune disorders. As first speculated by (Stewart, 1998), the Xi pattern of the thymus may constitute a risk factor for loss of T cell tolerance to self-antigens. Theoretically, a skewed pattern of Xi in thymic tissue may lead to inadequate elimination of potentially self-reactive T cells. Consequently, in visceral tissues, antigen expression from the X chromosome preferentially inactivated in the thymus (but preferentially active in visceral tissue) may be recognized by T cells as non-self, possibly eliciting an autoimmune response. Recently, the incidence of skewed Xi in leukocytes in a population of females with

putative autoimmune disorders (MS, SLE, juvenile RA, and type I diabetes mellitus) was compared with control females. A major role for skewed Xi in female predisposition to autoimmune disorders however could not be supported nor disproved. One possibility may be lack of appropriate tissue analysis, as thymus was unavailable (Chitnis et al., 2000). Nonetheless, the relationship between skewing (whether primary or secondary) and autoimmune function merits further investigation.

13.5 Skewing and longevity

Since i) AS occurs in a vast majority of elderly females and ii) females live relatively longer than males, an X-linked genetic basis to longevity was postulated (Christensen et al., 2000). AS reflecting the selection of X-linked 'longevity allele(s)' may be one possibility.

13.6 Therapeutic applications

Since the AS gene plausibly modulates HSC kinetics (replication, differentiation, survival), identification of the AS gene(s) may be of clinical/therapeutic relevance. For example, assuming that the AS gene can be identified and manipulated, the following are theoretically possible: i) stimulation of HSC to replicate, thus assisting horizontal DNA integration into normally quiescent cells. ii) Stimulation of HSC to differentiate, thus inducing differentiation of hematopoietic malignancies characterized by inhibition of differentiation. iii) Regulation of apoptosis, ideal for inducing apoptosis in malignancies.

Consequently, identifying the AS gene may promote novel therapies in treating medical disorders and provide a clearer understanding of the aging hematopoietic system.

14 HYPOTHESIS and GOALS

We hypothesize that Xi pattern analysis in humans is hampered by two unrelated events. One which corresponds to the primary Xi trait, and the other to the AS trait. By proper tissue, age and clinical / biological analyses, we speculate that both traits can be properly distinguished, quantitated and etiologies determined.

Our goals therefore were to distinguish the two traits.

1) Characterization of the two traits

- Determine incidence of skewed Xi as a function of tissue type
- Analyze XIRs as a function of age
- Analyze intraindividual correlation of XIRs between different tissues
- Analyze intraindividual correlation of XIRs among various hematopoietic lineages

2) Determine etiologies of skewed Xi patterns

- Determine heritability coefficient of XIRs
- Analyze segregation of the skewed Xi trait
- Analyze familial aggregation of skewed Xi
- Determine role of environmental factors

3) Clinical relevance of Xi patterns

- Analyze association between XIRs and family data
- Determine impact of AS on hematopoietic indexes
- Determine whether skewed Xi is associated with disease pathogenesis

15 STRATEGIES, RATIONALE AND EXPERIMENTAL DESIGN

We speculate that prior attempts to map the XCE-like gene, although putatively mapped to two regions of the X chromosome, one consistent with the XIC region (Naumova et al., 1998), was primarily confounded by etiologic heterogeneity. For instance, since their sample population did not consist of neonates and choice of study tissue was lymphocytes, we speculate that their skewing trait was confounded by both the PS and AS traits. Moreover, since selection criterion was not based on ethnicity, thus increasing overall probability of genetic and environmental variability, increased genetic variability is speculated. Thus, to circumvent these confounding variables, prudent choices such as appropriate tissue controls, age and ethnicity were considered in our study design.

Our approach therefore was to analyze skewed Xi in a population of females where i) the skewed Xi trait would be relatively high. ii) Etiologic heterogeneity would be relatively reduced. iii) Appropriate tissue controls would be available. iv) Large family size would be available, permitting trait segregation and potential genetic linkage analyses.

15.1 Recruitment of elderly females

Since AS is a late-onset trait, analysis of elderly females was required. In addition, a family versus a population-based approach was elected, thus permitting genetic linkage analysis in the event of significant evidence of heritability.

15.2 Genetic analyses: sib-pair approach to dissect a complex genetic trait

Notwithstanding an X-linked genetic component of inheritance, on basis of variable expressivity, impossibility of assigning Xi phenotype in males, age-dependant phenotype, tissue-specificity of phenotype, plausible role of environmental factors and speculation of genetic heterogeneity, a complex genetic basis to skewed Xi was speculated. As specification of mode of inheritance and large multigeneration families are not required, the most robust approach to dissect the genetics of a complex trait would be by non-parametric sib-pair analysis. Parametric/large pedigree analyses would not be ideal since it requires a specific genetic model and large multigeneration families. In addition, since the AS trait is late onset (≈ 60 yo), the likelihood of obtaining living parents is markedly reduced, and children moreover would not be informative for AS, thus limiting analyses to a single generation.

15.3 Collection of families

To reduce or eliminate ascertainment biases, families were ascertained without regard to trait status and family history. However, relatively large families were selected as this would increase i) the number of affectids for qualitative analyses and ii) the number of informative meioses in which to study genetic transmission of the skewing trait. An arbitrarily assigned criterion for family size was 4 or more female siblings. Further, since AS is a late-onset trait, a criterion for study enrollment was to have at least one female sibling 60 years of age or older. To be eligible for study enrollment, subjects had to be relatively healthy with no active cancer; and with the exception of self-reported anemia, no active blood disorder. Upon recruitment of a subject, she was responsible for recruitment of her sisters.

15.3.1 Benefit of a relatively isolated population

Ideally, when conducting a genetic-epidemiologic and/or linkage study, minimal population stratification, population admixture and environmental influences are warranted. As such, study of an isolated population offers several key advantages [see (Peltonen et al., 2000) for review] as background noise is reduced thus increasing statistical power. For one, isolated populations are

founded by a relatively small founder population, reducing overall genetic heterogeneity, thus families in the sample should have a similar underlying genetic predisposition. By studying isolated populations in which genetic variability is relatively reduced, a complex pattern of inheritance can be converted to one resembling a simple Mendelian trait (Jorde, 1995), thus improving the ability to detect the effects of a particular gene. Due to linguistic and geographical considerations, French-Canadian families residing in the province of Québec, Canada, founded by a relatively small population (estimated to range from 2500 to 8500 individuals) between the years 1609 - 1759, has been relatively isolated until the early 1900's; thus providing a relatively homogeneous population in which to study complex genetic traits. Noting that people in isolated populations typically share a common environment and culture (diet, exercise, seasonal influences and exposure to infectious diseases) reduced variability by non-genetic factors (phenocopies) to trait variance is speculated. In addition, well maintained genealogical records, minimal ethnic outbreeding and relatively large family sizes are ideal criteria for genetic-epidemiologic analyses. Due to the lifestyle during the early 1900's, e.g., Catholic upbringing thus the banning of contraception by the church, it was not uncommon to find families with a large number of children. In addition to being remotely available, it was on basis of these criteria that families of French-Canadian origin were accrued for study analysis.

15.4 Sample size (λ estimation): power studies

For complex diseases, power to detect linkage is dependant on the following factors: unit of relationship, level of recombination between marker and trait, marker heterozygosity, mode of inheritance, genetic heterogeneity and the relative risk (λ value) (Risch, 1990). The λ value, a measure of familial aggregation possibly due to genetic effects, is generally obtained by comparing the trait incidence in affected family members to that observed in the general population. For common trait alleles, the preferred model to estimate λ value is among sibs. However, since frequency of the skewing trait is both age and tissue dependent, thus prevalence of the trait in the general population an inappropriate measure, a modification was made to calculation of the λ value as follows. The recurrence risk ratio (or λr) of the skewed Xi phenotype was calculated by using a bank of families (n=98) recruited on basis of other diseases, namely HLA typing and calcium stone formation. Two-generation families where at least one daughter and mother were both informative for the HUMARA clonality assay (i.e., XIR available) were included in the analysis (n=53 families). XIR determination was performed on DNA extracted from PB. Briefly, one daughter was randomly selected from each family and dichotomized as skewed (case) or non-skewed (control). The Xi pattern of the corresponding mother was then analyzed. Of n=16 cases (skewed Xi pattern),

62,5% (10/16) had mothers who also demonstrated a skewed Xi pattern. On the other hand, of n=37 controls (random Xi pattern); 37,8% (14/37) had a mother with a skewed Xi pattern, deriving a λ value of 1,65. Although small, the value is above 1, supporting a familial / genetic component. Since calculation of λ is influence by trait prevalence in the general population, a low value is not surprising for a common trait such as skewed Xi. Other diseases with a low λ value included Alzheimer disease (λ =3,5). Nonetheless, assuming a trait frequency of approximately 40% in a 60 yo population, and an average sib-ship size of 6, we estimate 2 affected sib pairs per family. Following published tables, a λ of 1,6 requires 400 affected sib pairs to detect linkage (lod score > 3,0) with 59% power. This translates to an accrual of approximately 1200 females (~200 families).

15.5 Definition of the phenotype

An obstacle frequently encountered in complex trait analyses is uncertainty of phenotype definition. For skewed Xi, most studies have traditionally quantitated significant departure from random Xi (i.e., XIR $\geq$ 3:1) in PB. However, in light that both the PS and AS traits influence Xi patterns in PB (i.e., a combination of both the primary Xi pattern and AS trait), re-evaluation of methods to quantitate the Xi pattern in PB is warranted. Thus a model to quantify the AS trait in absence of the primary Xi pattern (i.e., relative AS) was developed. Although there is as yet no "correct" model from which to quantify relative AS, the utilization of T cells as a control tissue from which to quantify AS has been previously examined (Vickers et al., 2001). However, since i) T cells clonally expand in response to antigenic exposure, ii) clonal populations of T lymphocytes increase with advancing age and iii) as there is no evidence which unequivocally precludes T cell involvement in the AS trait, it appears that T cells are not an ideal control tissue from which to quantify AS. Alternatively, to properly quantify relative AS, the primary XIR (derived from a NHT) should be subtracted from the blood XIR, deriving a putative true AS value (relative AS).

Chapter II

MATERIALS AND METHODS

2.1 Subjects

'Healthy' elderly female subjects and her 'healthy' female siblings were recruited through multiple sources, including advertisements in senior citizen newsletters and through word of mouth. When available, mothers (n=37) and her female siblings were also recruited. All procedures were undertaken with ethical approval from the Hopital Maisonneuve-Rosemont Ethics Committee (see Annex 5 for consent form) and with the subjects' written informed consent. For confidentiality, family pedigrees were diagrammed and each family and individual received an identifier code number.

2.2 Medical Questionnaire

Each participant responded to a medical questionnaire (Annex 6) administered by trained health-care professionals. Data was collected on age, smoking habits, family data, and clinical data (self-reported medical conditions and medicinal use). Some of these variables were selected a priori and others as a result of exploratory analyses. All data was collected in two installments, by interview on the day of biological sample collection and by phone interview approximately 6-12 months after sample collection. Information was collected about themselves, parents, siblings, and offspring. It is not clear whether self-reported conditions were self-diagnosed or derived from a medical clinician.

2.2.1 Age

Age of participant was derived directly by interview and verified by medical ID cards.

2.2.2 Smoking habits: environmental stimulus

The role of environmental factor(s) as a cause of Xi pattern skewing was investigated. Rationale for selecting cigarette smoke as a possible stimulus is i) an affect on hematopoiesis has been documented, ii) smoking habits are quantifiable (discrete and continuous), and iii) cigarette smoke, with many of the undesirable effects attributed to nicotine, benzene and/or carbon monoxide, causes hemato-toxicity (leading to an imbalance in hematopoietic homeostasis), leukocytosis and altered immunological function.

Discrete variables of smoking habits included current smoker and ex-smoker status. Current smokers were defined as those currently exposed to direct/voluntary cigarette consumption and ex-smokers as those who had ever regularly smoked cigarettes but have now quit. For the ex-smoker variable, in our data sheets, a null-value was given if the subject was never a smoker, a zero value if a current smoker, and a true value (i.e., 1) if an ex-smoker. Current smoker was given a zero value

if never smoked or ex-smoker. Quantitative variables of smoke included number of years exposed to cigarette smoke (years-smoke), number of years stop smoking (years-stop) and the quantity of smoke exposure (pack-years), calculated as follows: packs of cigarettes consumed per day x number of years smoked. The variables years-smoke and pack-years are the cumulative of both current and ex-smoker. A zero value was assigned to years-smoke and pack-years if the subject was a non-smoker. As information on smoking habits was obtained by telephone interview, a certain rate of false-positive/false-negative is speculated but undetermined. Lack of honesty may be attributed to subjects attempting to paint themselves in a better light.

2.2.3 Family data: skewing and X-linked mutant allele(s)

Assuming X-linked lethal alleles can induce both a skewed Xi pattern in carrier females and the spontaneous abortion / miscarriage of a male conceptus, one hypothesis is that a proportion of skewing is caused by X-linked mutant allele(s). As such, we investigated the relationship between a skewed Xi pattern and family data, in particular offspring sex-ratio, occurrence of a miscarriage/spontaneous abortion, and the number of miscarriages. No data on the recurrence frequency of miscarriages was collected. Other family data collected was number and sex of siblings, whether parents of index cases were living and sex/number of offspring. Interestingly, since we recruited families with a high number of female siblings, we may have incidentally selected families transmitting X-linked mutant alleles. However, due to the lifestyle of the time (early 1990), it was not uncommon for families to have large numbers of children, therefore an unlikely bias.

2.2.4 Clinical data

Insight into the biological associations of Xi skewing was investigated by analyzing clinical data (complete blood counts, self-reported medical conditions and medicinal use) as a function of Xi patterns. Self-reported medical conditions included chronic diseases with a plausible immune/autoimmune component: asthma, systemic lupus erythematosus (SLE), rheumatoid arthritis (RA), and anemia. Reasons for selecting these disorders was based on a plausible association with a skewed Xi pattern (section 13.4 of introduction) and a female preponderance. Information on cancer history (type of cancer) and cancer related therapy (surgery, radiotherapy, chemotherapy) was also investigated as an association between Xi skewing and cancer (breast and ovarian) has been suggested (section 13.3). Of interest, although chemotherapy (busulfan administration) was suggested to accelerate the Xi skewing trait (Abkowitz et al., 1988), more recent studies do not support such a claim: i) no difference in Xi patterns was found between females who had received

chemotherapy versus controls (Gale et al., 1991) and ii) in a cohort of breast cancer patients who had received chemotherapy, no change in Xi patterns was observed a year later (Buller et al., 1999).

As medical records were not evaluated a limitation to these analyses could be a certain rate of false-positives. For example, in self-reported RA, a false-positive rate of 55% has been reported (Lynn et al., 1995). Individuals falsely claiming RA in fact had osteoarthritis (OA), an inflammatory/atrophy joint disorder. False negatives are unlikely as a 0% false-negative rate was also reported.

Self-reported medicinal use was categorized into 9 sub-types, as determined by the principal investigator: i) use of hormone replacement therapy (HRT), ii) medication for allergies (med-allergy), iii) anti-convulsants (anti-conv), iv) anti-inflammatory (anti-inflam), v) medication for asthma (med-asthma), vi) miscellaneous medication (med-other), vii) heart medication (med-heart), viii) hematological medication (med-hemato), and ix) vitamins, supplements and herbal remedies (med-vit-other). Medicinal-use was used as an indirect method to infer a medical condition. However, whether medicinal products have a direct affect on Xi skewing is not known.

2.3 Biological samples

To characterize the primary Xi and AS trait(s), appropriate tissues were required.

2.3.1 Primary Xi pattern / PS trait

The primary XIR was derived from buccal cell tissue. It was easily obtained by a gentle 10-second stroking (performed by the health care professional) of the two inner cheeks (mucous membrane lining the mouth) using a CEP buccal swab (Gibco, BRL). Samples obtained outside the Montreal region were transported with over-night delivery.

2.3.2 AS trait: peripheral blood

The AS trait was analyzed in independent hematopoietic lineages. Three tubes (20-25 ml) of heperanized peripheral blood (PB) were obtained by venipuncture from each participant. Samples obtained outside the Montreal region were transported with over-night delivery. PB was fractionated and cell-sorted to obtain pure cell populations: PMN, monocytes, T and B lymphocytes. To our knowledge, this is the first investigation analyzing the AS trait as a function of various hematopoietic lineages.

2.4 Complete blood counts

To determine whether AS modulates hematopoietic indexes, XIRs were analyzed as a function of blood parameters such as blood counts, hemoglobin concentration and MCV. A complete blood count (CBC) and white cell differential was performed on blood samples using a GenS automated cell counter (Coulter). No samples were excluded because of abnormal parameters.

2.5 Cellular fractionation

PB was separated into polymorphonuclear (PMN) and mononuclear (MLL) cellular fractions using standard density gradient centrifugation (Ficoll-Paque, Pharmacia). The granulocyte fraction (PMN) was assessed morphologically and was typically greater then 98% pure. An aliquot of the MLL fraction was subjected to immunophenotyping to obtain pure populations of CD14+ cells (monocytes), CD3+ cells (T lymphocytes), and CD19+ cells (B lymphocytes) (see Chapter 2.6).

2.6 Immunophenotyping and cell sorting

Isolation of T lymphocytes, B lymphocytes, and monocytes were obtained by one of two methods. Originally, the isolation of T lymphocytes (CD3+), B lymphocytes (CD19+) and monocytes (CD14+) was carried out solely by fluorescence-activated cell sorting (method 1). However, to increase the number of samples analyzable on a daily basis, a reduction in manipulating time was required. As such, a step prior to cell sorting was the isolation of T cells using CD3-coated magnetic beads (method 2), followed by cell sorting of the effluent enriched for CD14 and CD19 positive cells.

Method 1: T lymphocytes, B lymphocytes and monocytes were directly isolated using a fluorescence-activated cell sorter (FACStar-plus, Becton Dickinson) after staining with monoclonal antibodies conjugated to fluorescent isothiocyanate (FITC) or phycoerythrin (PE). The monoclonal antibodies used in this study consisted of anti-CD3/FITC (IgG1), anti-CD19/PE (IgG1), and anti-CD14/FITC (IgG2a), as well as their isotypic msIgG1 and msIgG2a /FITC or /PE controls (Coulter Immunology, ID Labs).

Method 2: Prior to anti-CD14 and anti-CD19 labeling as outlined in method 1, CD3-positive cells were depleted using CD3-coated magnetic beads (Dynal) according to manufacturer's instructions. Number of cells collected and cell purity are provided in Table VII (page 65). DNA from the cell-sorted populations was isolated by standard methods (Chapter 2.7).

Table VII. Cell-sorted populations and purity based on light-scattering properties.

Cell type	Number of cells collected (mean)	Cell-sorting purity (mean %)
CD3 (T lymphocytes)	$1,3x10^5$	99,7 ± 0,25
CD3-coated magnetic beads	$1,2x10^5$	97,9 ± 0,29
CD14 (monocytes)	$7,3x10^4$	99,5 ± 0,33
CD19 (B lymphocytes)	$1,0x10^5$	99,4 ± 0,48

2.7 DNA isolation from blood cells

Prior to DNA isolation from the PMN fraction, nuclei were isolated by lysing cells with Triton X-100 (Sigma, St-Louis) solution (0,32 M sucrose; 10mM Tris pH 7,5; 5 mM $MgCl_2$; 1% Triton X-100). To isolate the DNA, nuclei of the PMN fraction and the cell pellet of cell-sorted populations were subjected to conventional cell lysis and proteinase K/SDS digestion. Briefly, the cell and nuclei DNA pellets were digested in NaCl/EDTA (75nm/25mM) buffer supplemented with proteinase-K (2 mg/ml) / SDS (5%) at 37°C for 24-48 hours. DNA was extracted and rid of proteins with a phenol, phenol/chloroform and chloroform extraction procedure; precipitated with 2 volumes of 95% ethanol and 1/10 volume of 3 M Sodium Acetate; washed with 70% ethanol, and resuspended in an appropriate volume of Tris-EDTA (TE) buffer. DNA derived from the PMN fraction had a mean concentration of 275 ng/ul (20-1900 ng/ul). DNA samples were conserved at 4°C.

2.8 DNA isolation from buccal cells

DNA from buccal cells was isolated according to manufacturer's protocol (Gibco, BRL). Briefly, to the 500 ul ependorff tube containing the buccal swab was added 250 ul NaCl/EDTA (75nm/25mM), 50 ul proteinase-K (2 mg/ml) and incubated for 2 hrs at 55°C. After digestion, the tube was vortexed for several seconds. On a sterile bench, the bottom of the tube was perforated with a sterile needle and the tube was then placed into a larger 2-ml ependorff tube. The supernatant was collected in the 2 ml tube by centrifugation (30 sec at 10 000 rpm) and then subjected to DNA extraction by the lithium chloride method. Briefly, to each tube was added ½ volume (150 ul) of lithium chloride (American Chemicals LTD) (7.5 M), vortexed for several seconds and placed at –

80°C for 10 minutes, followed by a centrifugation at 14000 rpm for 15 min. Carefully, the supernatant was transferred to a new tube and mixed with an equal volume of 95% ethanol. The tube was gently shaken for several seconds and DNA precipitation carried out at room temperature for 30 min. The tube was centrifuged for 30 minutes at 14000 rpm, the supernatant was decanted and the pellet subjected to one cycle of washing with 70% ethanol. After centrifugation (15 min, 140000 rpm) and decantation of the supernatant, the tube was placed under the chemical hood to evaporate residual ethanol. The pellet was resuspended in 50 ul TE and incubated at 50°C for 15 min. DNA samples were conserved at 4°C.

2.9 HUMARA clonality assay

The HUMARA assay was carried out as previously described (Busque et al., 1996). Briefly:

Kinasing primer protocol: 6 µl (5 pmol/µl) of primer HUMARA I was added to 10X kinase buffer (2 µl), γ-^{32}P dATP (6 µl, 3000 µCi/mmol), polynucleotide kinase (PK; 1,2 µl) and H_2O (4,8 µl), followed by incubation at 37°C for 35 min and inactivation of PK at 94°C for 4 min.

Pre-cutting of genomic DNA: genomic DNA was pre-cut by mixing sample DNA (50ng-500 ng in 4-5 µl) with *Hpa* II (24 units), *Rsa* I (12 units), L buffer (2,5 µl,) and H_2O (16,6 µl). Buffers and enzymes were supplied by Boehringer Mannheim. An auto-control (mock-digest) was prepared in the same way except that *Hpa* II was omitted from the mix. As *Rsa* I RE sites are outside of the HUMARA PCR target, it was included to prevent preferential amplification of the inactive allele. Samples were incubated at 37°C overnight, and heat inactivated at 80°C for 15 min prior to PCR amplification.

PCR amplification of the HUMARA locus: 7 µl of digested DNA was added to 20 µl of a PCR mix containing buffer (10X: 500 mM NaCl, 100 mM Tris-HCl, pH 8,2; 15 mM $MgCl_2$, 0,1 % gelatin); dNTPs (200 µM each); primer HUMARA I: 5'-GCTGTGAAGGTTGCTGTTCCTCAT-3', and primer HUMARA II: 5'-TCCAGAATCTGTTCCAGAGCGTGC-3' (10,0 pmol each); 1 ul DMSO (Sigma); γ-^{32}P end-labeled HUMARA I primer (1,0 pmol); Taq polymerase (0,5 units); and H_2O to final volume of 20 µl. Samples were amplified on a programmable thermal cycler (MJ Research, Inc.): initial DNA denaturation at 94°C for 3 min, then 28 cycles of amplification starting with denaturation at 94°C for 45 sec, annealing at 60°C for 30 sec, and elongation at 72°C for 30 sec.

Gel electrophoresis: Upon completion of the PCR reaction, 10 µl of formamide loading buffer (95% formamide, 20 mM EDTA, 0,05% bromophenol blue, 0.05% xylene cyanol) was added to each sample, samples were denatured at 95°C for 3 minutes, followed with rapid chilling to 4°C. Amplified PCR product (8-10 µl) were electrophoresed on a 4% polyacrylamide denaturing gel at 60 watts for approximately 3 hours.

2.9.1 HUMARA allelic ladder

For identification of the HUMARA allele size, a referencing ladder was constructed (Desmarais et al., 1998) with the following modifications. A total of 12 individuals, 3 males and 9 females, spanning the observed HUMARA allele spectrum (alleles 8-36) were selected. Individual samples were subjected to PCR amplification using the above protocol with the exception that the HUMARA primer 1 was not end-labeled with γ-^{32}P dATP. PCR products were pooled together and DNA extracted by phenol, phenol/chloroform and chloroform extraction, precipitated with 2 volumes of 95% ethanol and 1/10 volume of 3 M Sodium Acetate, washed with 70% ethanol, and resuspended in Tris-EDTA (TE) buffer. A small aliquot (4 ul) was included in every PCR plaque subjected to the HUMARA clonality assay.

2.9.2 Quantitation of HUMARA alleles

Gels were vacuum-dried and exposed overnight to a Bio-Rad CS imaging screen (phosphor-imaging technology). The latent image was recovered by infrared light scanning of the screen in a GS-525 Molecular Imager (Bio-Rad) and recorded as a digitized image. Band quantitation analysis was processed with the Molecular Analyst software (Bio-Rad). The following protocol was strictly followed. Individuals homozygous for the HUMARA locus were excluded from band quantitation analysis. In heterozygous individuals, a mock digest and an enzymatic/*Hpa II* digest were available. The mock digest was included to determine zygosity and to quantitate potential preferential amplification of the smaller allele. In each lane, band intensity of each allele was quantitated. A box of equal size was utilized to determine the pixel intensity of each allele. Local background, as determined by the software, was subtracted from each sample. The allele ratio (ar) was determined by dividing the band intensity of the larger allele by the band intensity of the smaller allele. An ar was derived for both the mock and *Hpa II* digest lane.

ar = (band intensity of the larger allele) / (band intensity of the smaller allele)

In order to compensate for potential preferential amplification of the smaller allele, the ar of the mock digest was included in the calculation of the XIR, deriving the Proportion active of the larger HUMARA allele (PALA) score.

PALA = (ar mock digest) / [(ar mock digest) + (ar *Hpa II* digest)]

2.9.3 Reliability of quantification: accuracy

Reliability of quantification was determined by comparing the XIR of observed versus theoretically expected ratios. In brief, two male samples, each bearing a different HUMARA allele length and of equal concentration (40ng/ul), were mixed to the following ratios: 10:0, 9:1, 8:2, 7:3, 6:4, 5:5, 4:6, 3:7, 2:8, 1:9, 0:10 (thus comprising the theoretical spectrum of allele ratios). The observed XIR was determined for each sample and compared to the theoretical XIR. As ratios were not significantly different between the two, a reliable and accurate method for quantitating XIRs was ensured.

2.10 Calculation of the XIR

The XIR was reported by two methods: the DS and PAmat scores (Naumova et al., 1998). In addition, for blood cells, a novel measure of XIRs included the AS_{DS} and AS_{PAmat} scores.

2.10.1 DS score (degree of skewing):

A quantitative measure of deviation from random Xi; values ranged from 0 to 0,5. A score of zero indicates a random Xi pattern while a score of 0,5 signifies a complete skewed Xi pattern. The DS score was calculated by subtracting the PALA score from 0,5, followed by taking the absolute of this value.

DS = | 0,5 - PALA |

2.10.2 PAmat score (Proportion Active of the Maternally inherited X chromosome):

Individuals informative for phase of the X chromosome, the Proportion Active of the Maternally inherited chromosome (PAmat score) was calculated, a measure of both the magnitude and direction of skewing relative to the maternal X chromosome active; values ranged from 0 to 1. Parental derivation of HUMARA alleles was possible only when the mother was genotyped and/or when the paternal allele was indirectly identified. A PAmat score of 0,5 indicated random Xi, a score above 0,5 indicated preferential inactivation of the paternal X chromosome and a score below 0,5 indicated preferential inactivation of the maternal X chromosome. The PAmat score was calculated by two methods, pending on parental origin of the larger HUMARA allele.

i) When the larger HUMARA allele was paternal in origin:

PAmat = PALA

ii) when the smaller HUMARA allele was paternal in origin:

PAmat = (1 - PALA)

2.10.3 AS_{PAmat} score:

To derive a quantitative value of AS in absence of influence from the primary Xi patter, the AS_{PAmat} score was determined. It was obtained by subtracting the PAmat score of BC from the PAmat score of blood. Theoretically, the score could range from -1 to +1. A negative score signifies preferential inactivation of the maternal X while a positive score represents preferential inactivation of the paternal X.

$$AS_{PAmat} = (\text{PAmat blood}) - (\text{PAmat BC})$$

2.10.4 AS_{DS} score

To derive a quantitative measure of AS in absence of parental direction (Xp versus Xm), the AS_{DS} score was determined, calculated by taking the absolute of the AS_{PAmat} score. Theoretically, the score should range from 0 to 1.

$$AS_{DS} = | AS_{PAmat} |$$

2.11 Discrete criteria to delineate a skewed Xi pattern

2.11.1 Skewing as a discrete trait

To analyze Xi skewing as a discrete trait, an arbitrarily defined DS score $\geq 0{,}25$ was utilized. It corresponds to at least 75% of cells inactivating the same parental X chromosome. This value is widely accepted in the literature to delineate clonal derivation of cells.

Skewed X inactivation = (DS score $\geq 0{,}25$)

2.11.2 Acquired skewing as a discrete trait

Two methods were utilized to measure the incidence of the AS trait. The first was a qualitative method (AS_{QL}), determined as follows: a random Xi pattern in BC (0,75 > PAmat > 0,25) but a skewed Xi pattern in blood (0,75 ≤ PAmat ≤ 0,25). A limitation to this method however, was lack of stringency. For example, a PAmat score of 0,74 in BC and 0,76 in PMN, a difference of only 0,02, would signify AS.

$$AS_{QL} = \text{(skewed Xi pattern in blood, random Xi pattern in BC)}$$

The second was a quantitative method (AS_{QT}), where an arbitrarily defined AS_{DS} score ≥ 0,25 signified AS as a discrete trait. Unlike the AS_{QL} method, this method was more stringent as a deviation of at least 0,25 was required to delineate a skewed Xi pattern. A limitation however occurred when the PAmat score of BC was greater than 0,75 or less than 0,25. Under these circumstances, if skewing in blood occurred in the same direction as in BC, it was impossible to obtain an AS_{PAmat} score ≥ 0,25, thus possibly underestimating the actual incidence of AS.

$$AS_{QT} = (AS_{DS} \text{ score} \geq 0{,}25)$$

2.12 Familial resemblance of Xi pattern traits

Several approaches were taken to analyze the familial resemblance of Xi patterns, a positive finding possibly suggestive of a genetic component: i) familial aggregation of skewed Xi, ii) sibling correlation of XIRs analyzed by ANOVA, and iii) heritability analyzes.

2.12.1 Familial aggregation of skewed Xi: recurrence risk ratio (RRR)

To determine whether skewing was aggregated within families, the recurrence risk ratio (RRR) was determined. It was calculated by first determining the recurrence risk (RR) in relatives of the proband. The RR is simply the proportion of affected relatives. It was calculated for relatives of the affected case (skewed Xi) and for relatives of the unaffected control (random Xi), thus a case-control method. In this study, relatives included all female siblings of the proband. From each family, one individual was randomly chosen and categorized as affected (case) or unaffected (control). The affection status in all remaining siblings was then determined, deriving the RR. The RRR was derived by dividing the RR of affectids by the RR of controls. A RRR > 1,0 normally suggests familial aggregation of a trait.

2.12.2 Sibling correlation of XIRs by ANOVA

To derive evidence of a familial (possibly genetic) component to skewing, quantitative analyses were conducted. In particular, correlation of XIRs within versus among families, by means of analysis of variance (ANOVA) (Donner and Eliasziw, 1991). In an outbred population, one expects a genetically influenced trait be more correlated within biologically related individuals than among non-biological relatives. Families with at least two siblings whom had a PAmat or AS_{PAmat} score were included for study analysis. We hypothesize that if skewing is X-linked, siblings should demonstrate stronger correlation of XIRs versus unrelated individuals, i.e., stronger correlation within versus among families. However, correlation of trait values among siblings is only suggestive evidence of a genetic component since shared environmental factors may induce similar patterns.

2.12.3 Heritability analyses

To acquire knowledge of the genetic architecture of quantitative traits, estimates of heritability (fraction of the total phenotypic variance attributable to additive genetic effects) are routinely performed. In principle, these analyses quantify the degree of familial resemblance due to genetic effects. And although heritability estimates provide presumptive evidence of genetic determinants, other explanations (e.g., unidentified environmental influences) are possible. In this study, the testing of various variables for contribution to trait variance and their corresponding p-values were derived using the SOLAR (Sequential Oligogenic Linkage Analysis Routines) statistical software package. Various variables of skewing included DS, PAmat, AS_{DS} and AS_{PAmat} scores. Covariates taken into consideration were: age, pack-years of cigarette smoke, current smoker and ever-smoker. In the variance component approach, the likelihood of the pedigrees for the trait of interest is assumed to follow a multivariate normal distribution (Almasy and Blangero, 1998). Estimates of the means and variances of components of the models are obtained by maximum likelihood method (Lange et al., 1976). Maximum likelihood methods are used to estimate simultaneously the genetic and covariate effects. The effect of a covariate on each trait is determined by comparing a model with covariates to a model with a specific covariate removed. The resulting likelihood ratio statistic is used to test the significance of the effects, comparing the full model (i.e., genes + covariates) with the reduced model. The trait was modeled against a linear combination of effects caused by covariates, genetic effects, and residual variability. The effect caused by genes was assumed to be independent from other effects and normally distributed with mean of zero and variance of s^2g. Heritability (h^2) was defined as the variance of the trait caused by genes (s^2g) divided by the total phenotypic variance (s^2). This analysis was performed with the SOLAR v1.5.7 software program (Almasy and Blangero, 1998).

In addition to XIR scores, other traits considered for heritability estimates were blood counts, number of spontaneous abortions, some clinical conditions (RA and asthma) and smoking habits.

2.13 Intra-individual correlation of XIRs:

Body-wide Xi pattern: *correlation of XIRs between BC and hematopoietic lineages*

To derive evidence of a primary Xi trait (i.e., body-wide Xi pattern), correlation of PAmat scores between BC and hematopoietic lineages were analyzed. We speculate that in the event of a primary Xi trait, PAmat scores should be strongly correlated between the two tissues. Alternatively, if skewed Xi were tissue-specific, low or no correlation is postulated.

HSC origin of AS: *correlation of XIRs among hematopoietic lineages*

To derive evidence in favor or against a HSC origin of AS, correlation of XIRs among the various hematopoietic lineages was investigated. Strong correlation would argue in favor of a common mechanism of skewing, possibly consistent with a HSC origin. Alternatively, weak correlation would imply lineage-specific etiologies of AS.

2.14 Statistical analyses

Linear Regression (LR), Multiple linear regression (General Linear Model procedure – GLM), tests of correlation, t-tests and multivariate Ftests were carried out using the SAS 8.0 and SPSS 10.1 softwares. Categorical variables included clinical conditions, medicinal products, various smoking variables and various skewed Xi pattern phenotypes. Other variables were treated as continuous variables. Family, when necessary, was treated as a class variable – General Estimating Equations (GEE) method. In fact, GEE, a marginal regression model, is an extension of the usual GLM (Liang et al., 1986). It is designed for correlated data, such as observations within a family. It is a semi-parametric approach which uses correlation as a measure of association (Tregouet and Tiret, 2000). It is robust to mis-specifications of the familial correlation and results are not influenced by family size.

Chapter III

RESULTS (PART I)

3.1 Population data

A total of 1144 female subjects, derived from 193 nuclear families were recruited for study enrollment. No males were recruited. The number of female sibling participants per family ranged from 3-12, with a mean of 5,6 and a median of 5 (Figure 6, page 89). In 19,7 % of these families (n=38), the mother and in some instances her female siblings were also recruited, deriving a mean of 5,9 females per family. The age range of female participants varied from 38-96 years of age, and demonstrated a normal distribution with a mean of 63,3 ± 10,0 (Figure 7, page 89).

3.2 Specimen data

From each participant the following cell types were obtained for Xi analysis: buccal cells (BC), polymorphonuclear cells (PMN or granulocytes), CD14+ sorted cells (monocytes), CD3+ sorted cells (T lymphocytes), and CD19+ sorted cells (B-lymphocytes). Of 1144 females, 1039 were informative for the HUMARA clonality assay, deriving a heterozygosity rate of 90,8%, in agreement with previous findings (Busque et al., 1996). Table VIII (page 102) provides data on the number of informative females (DNA amplified and heterozygous for the HUMARA locus) for each cell type. These samples were amenable for Xi analyses.

To analyze the degree to which PCR contributed to the variation in the data, we tested for reproducibility of the PCR assay. For BC, the HUMARA assay was repeated on 48 randomly selected samples. The absolute mean difference of PAmat scores reported at trial 1 versus trial 2 was 0,048 ± 0,046. To assess reproducibility for blood samples, we estimated the variability of our positive control (specimen 808 a and h), which was included in every PCR assay. For 175 replicates, the mean DS score was 0,374 ± 0,051.

3.3 Incidence of skewed Xi

3.3.1 Skewed X inactivation: DS score ≥ 0,25

The incidence of skewed Xi (DS score ≥ 0,25) was determined for each cell type.

BC skewing: primary skewed Xi pattern (PS) / body-wide skewing trait

As shown in Figure 8 (page 90), the incidence of skewed Xi in BC was estimated at 12,4% (129/1039). We speculate that this incidence of skewed Xi reflects the frequency of the PS trait. If indeed reflective of an *XCE*-like etiology, as a DS score of ≥25% was utilized would suggest strong

acting *XCE*-like alleles in humans, as *Xce* allele Xi skewing in mice induces a maximal skewing bias of 20-30%.

Peripheral blood skewing: higher incidence of skewed Xi versus BC

Versus BC, an elevated incidence of skewed Xi pattern was observed in hematopoietic lineages. PMN and monocytes had similar (but not significantly different) incidences of skewing: 36,0% (374/1039) and 36,6% (366/1000), respectively. The incidence of Xi skewing in T lymphocytes (20,1%; 205/1021) was significantly less versus B-lymphocytes (26,8%; 271/1017) ($p<0{,}001$). A higher incidence of Xi skewing in PMN versus T cells ($p<0{,}001$) was previously documented (Gale et al., 1997), (Champion et al., 1997), (Tonon et al., 1998), and speculated to reflect longevity of T cells. We speculate that the increased incidence of Xi skewing in hematopoietic lineages versus BC is due to the AS trait.

3.3.2 Incidence of the AS trait: AS_{QL} and AS_{QT} analyses

Incidence of acquired skewing - qualitative analyses: AS_{QL}

For this method, a random Xi pattern in BC but a skewed Xi pattern in blood was used to delineate AS as a discrete trait. As depicted in Figure 9a (page 91) there was a high incidence of AS_{QL} in PMN (32,2%, 293/910) and monocytes (34,6%; 303/875). However, this difference was not significant. The incidence of AS_{QL} in B lymphocytes (22,5%; 200/890) was significantly lower versus PMN ($p<0{,}001$) and monocytes ($p<0{,}001$). The lowest incidence of AS_{QL} was found in T lymphocytes (16,3%; 146/894); which was significantly lower versus PMN ($p<0{,}001$), monocytes ($p<0{,}001$) and B lymphocytes ($p<0{,}001$).

Incidence of acquired skewing - quantitative analyses: AS_{QT}

The incidences of AS_{QT} (measured as significant deviation from the body-wide XIR, i.e., AS_{DS} score $\geq 0{,}25$) are provided in Figure 9b (page 91). The AS_{QT} trait was most prevalent in PMN (22,7%; 203/896) and monocytes (27,2%; 235/863), with a significantly higher incidence in monocytes versus PMN ($0{,}025<p<0{,}05$). The incidence of AS_{QT} was 11,4% (100/879) and 16,3% (143/877) in T and B-lymphocytes respectively. The incidence of AS_{QT} in T cells was significantly lower versus B cells ($0{,}001<p<0{,}005$), PMN ($p<0{,}001$) and monocytes ($p<0{,}001$).

When comparing the two analyses (AS_{QL} versus AS_{QT}), a higher incidence of AS was reported with the AS_{QL} criterion for all four lineages. A likely explanation is the lack of stringency the latter method employed. Nonetheless, based on the following explanation, the AS_{QT} model may be a more

accurate estimate. As shown in Table IX (page 102) if the frequency of primary skewing (12,4%) is subtracted from the frequency of blood skewing (e.g., 36,0% for PMN), an incidence of AS similar to the AS_{QT} method (22,7%) is derived, i.e., 36,0 – 12,4 = 23,6%, a difference of only 0,9%.

3.4 Distribution and mean DS and AS_{DS} scores

DS score

For buccal cells, setting a DS score of 0,00 as "absolute" random Xi, i.e., 50% of cells inactivating Xp and 50% Xm, the proportion of individuals deviating from this value were determined. The distribution of these scores is given in Figure 10 (page 92). Individual DS scores ranged from 0 (no skewing) to 0,5 (complete skewing). All samples were skewed to the right, consistent with the majority of females exhibiting random XIPs. For PMN and monocytes however, a higher frequency of DS scores in the range of 0,25 to 0,50 can be observed, consistent with the higher incidence of skewing for these cell types.

As shown in Table X a) (page 103), the mean DS score for BC was 0,123 ±0,094, in agreement with previous findings (Monteiro et al., 1998). The mean DS score for PMN (0,204 ±0,129) was not significantly different from monocytes (0,207 ±0,136) ($p=0,634$). The mean DS score of T lymphocytes (0,153 ±0,111) was significantly lower versus B-lymphocytes (0,174 ±0,112) ($p<0,0001$). The mean DS score of buccal cells was lower versus all blood cell lineages ($p<0,0001$). Overall, mean DS scores among the various cell types were significantly different ($F=91,3$; $p<0,0001$).

AS_{DS} score

For blood cells, the distributions of AS_{DS} scores are given in Figure 11 (page 93). AS_{DS} scores ranged from 0 (no AS) to 0,75 (maximum AS_{DS} score observed). For all lineages, the distribution of scores was skewed to the right, consistent with most females exhibiting a small degree of AS. For PMN and monocytes, versus T and B-lymphocytes, a higher frequency of females with AS_{DS} score ≥ 0,25 can be observed, consistent with the higher incidence of AS for these cell types.

As shown in Table X b) (page 103), mean AS_{DS} scores were highest for myeloid-derived lineages: a mean AS_{DS} score of 0,159 ±0,125 was observed for PMN, which was not significantly different from that of monocytes, mean AS_{DS} = 0,173 ±0,133 ($p=0,057$). For T lymphocytes, the mean AS_{DS} score was 0,127 ±0,099, which was significantly lower versus B-lymphocytes, mean AS_{DS} = 0,142 ±0,110 ($p=0,046$). The mean AS_{DS} scores of both PMN and monocytes were significantly higher

versus T and B-lymphocytes ($p<0{,}001$). Overall, the difference in mean AS_{DS} scores among the various blood lineages was significantly different ($p<0{,}001$). Assuming that the AS trait represents departure from the body-wide XIR, these results imply that deviation occurs primarily in myeloid-derived lineages. Lower mean AS_{DS} scores for B and T lymphocytes imply less deviation from the body-wide XIR, particularly for T lymphocytes.

3.5 XIRs as a function of age

To identify phenotypic differences between the primary Xi and AS traits, XIRs were analyzed as a function of aging.

PAmat score

As shown in Figure 12 (page 94) and Table XI a) (page 104), that PAmat scores were not significantly associated with age for all tissues suggests that derivation of XIRs is indifferent to source of parental alleles, thus arguing against an imprinted basis for skewing (primary and acquired forms).

DS score

As presented in Figure 13 (page 95) and Table XI b) (104), the DS score of BC was relatively stable with advancing age. In fact, it slightly decreased but was not significant (GEE results: $p= 0{,}2069$; $b= -0{,}0004$). This is a novel finding, suggesting that the observed incidence of Xi skewing in BC is acquired prior to the age range 38-96 (mean 63,3) and is relatively stable thereafter. This finding supports a primary etiology of XIR derivation in BC. As such, it supports our model to quantitate AS by measuring departure from the primary Xi pattern.

In both PMN ($p= 0{,}0004$; $b= 0{,}0016$) and monocytes ($p=0{,}0056$; $b= 0{,}0013$), the DS scores increased significantly with advancing age, consistent with the properties of the AS trait. Similar p-values were demonstrated by the SOLAR statistical method. However, as the beta estimates were small, the data implies that age is a moderate predictor of XIRs in our elderly female population. This finding is consistent with prior data that AS occurs primarily prior to age 60 but increases steadily thereafter (Christensen et al., 2000). Among T ($p= 0{,}1341$, $b= 0{,}0006$) and B-lymphocytes ($p= 0{,}3481$; $b= 0{,}0004$), the DS score slightly but insignificantly increased with advancing age (as depicted in Figure 13, page 95), consistent with the lower incidence of skewing for these lineages. Similar results were obtained by the SOLAR statistical method (Table XI b, page 104). These results suggest that the incidence of skewed Xi ($DS\geq0{,}25$) observed in T and B cells does not

significantly vary with age, consistent with acquisition of skewed Xi patterns prior to age range 38-96.

In light that BC DS scores were relatively stable while those of hematopoietic lineages increased with advancing age is consistent with different phenotypes, thus suggesting disparate traits.

AS_{PAmat} score

As shown in Figure 14 (page 96) and Table c) XI (page 104), AS_{PAmat} scores were not significantly associated with age for all lineages, suggesting that derivation of AS is indifferent to source of parental alleles, arguing against an imprinted basis.

AS_{DS} score

As a function of aging (Figure 15, page 97), the AS_{DS} scores increased significantly for all hematopoietic lineages, as demonstrated by results of the LR and GEE methods [Table d) XI (page 104)]: PMN ($p<0,0001$; $b=0,0024$), monocytes ($p=0,0004$; $b=0,0019$), T cells ($p=0,0378$; $b=0,0009$), and B cells ($p=0,0021$; $b=0,0013$), consistent with the properties of the AS trait. Like DS score results, the beta-estimates were relatively small, suggesting that age is a moderate predictor of AS scores in an elderly population of females. Results derived from the SOLAR statistical method were slightly different. Although the AS_{DS} scores of PMN and monocytes were significantly associated with age ($p=9x10^{-6}$ and $p=0,0035$, respectively), consistent with results derived from LR and GEE methods, the AS_{DS} scores of T and B lymphocytes unlike LR and GEE results were not associated with age ($p=0,13$ and $p=0,14$, respectively). Discrepancies may invariably be explained by the method of statistical analyses.

That the AS_{DS} but not the DS score of T and B lymphocytes slightly but significantly increased with age suggests that as a function of aging, the Xi pattern of lymphocytes deviates more from the primary Xi pattern versus from random Xi. In other words, it appears that the observed incidence of Xi skewing in lymphocytes has not increased drastically between the ages 38-96. One possibility as for why AS_{DS} but not DS scores of lymphocytes increased significantly with age, as calculation of the AS_{DS} score involved subtraction of the BC PAmat score from hematopoietic lineages, may be attributed to a slight decrease of DS scores in BCs with advancing age.

3.6 Evidence supporting a primary Xi (PS) trait: intraindividual correlation of XIRs between BC and hematopoietic lineages

Evidence for a primary Xi (PS) trait was investigated by analyzing correlation of PAmat scores and concordance of skewed Xi patterns between BC and hematopoietic lineages.

3.6.1 Correlation of PAmat scores between BC and leukocytes

As shown in Table XII a) (page 105) and Figure 16 (page 98), the PAmat score of buccal cells correlated equally well with T lymphocytes (r=0,563), B-lymphocytes (r=0,546), and PMN (r=0,535). Correlation with monocytes was slightly weaker (r=0,462). All correlation coefficients were significant ($p<10^{-17}$). These results suggest a common mechanism of XIR derivation between the two tissues, evidence in favor of a body-wide / primary Xi trait. Lack of complete correlation between the two tissues may plausibly be explained by the AS trait. That correlation was weakest with the cell type exhibiting the highest incidence of AS, i.e., monocytes (p=.462), is consistent with the latter. However, that correlation of BC with PMN (r=0,535) was similar to that of BC with T cells (r=0,563), suggests that in addition to AS, other etiologies are possible. For instance, that the incidence of skewed Xi in lymphocytes was acquired prior to age range 38-96 and the DS score relatively stable within this age range suggests that in addition to a body-wide Xi pattern, Xi patterns may also be partially tissue-specific.

3.6.2 Concordance for a skewed Xi pattern among various tissues

Evidence supporting a primary (body-wide) Xi skewing trait (PS). *Methods*: Females with a skewed Xi pattern in BC were examined for Xi pattern analysis in hematopoietic lineages. We hypothesized that in the event of a primary Xi skewing trait, BC skewing should be highly concordant with leukocyte skewing (i.e., body-wide skewing). On the other hand, if Xi-skewing were tissue-specific, the incidence of leukocyte skewing should be unrelated to BC skewing, thus expected to reflect the incidence of skewing observed in our cohort population of females (column 3 of Table XIII, page 106). As observed in column 4, concordance for a skewed Xi pattern between the two tissues varied from 53,6 to 64,3%, an incidence higher than expected if Xi skewing was tissue-specific, supporting evidence for a body-wide Xi skewing trait. Lack of complete concordance can be explained by lineage specific etiologies such as the AS trait, clonal derivation of cells, stochastic processes, methodologies used and/or technical variability. Moreover, as direction of skewing (preferential inactivation of a parental-specific X chromosome) was highly concordant between the two tissues (94-98,6%; column 5) suggests a common mechanism in derivation of skewed Xi patterns between the two tissues, congruent with an X-linked etiology or derivation from a common stem cell pool.

3.7 Evidence for a HSC origin of AS: intraindividual correlation of XIRs among hematopoietic lineages

3.7.1 Correlation of PAmat scores among blood lineages: HSC origin

As shown in Table XII b) (page 105) and Figure 17 (page 99), the PAmat score of PMN, a cell type that demonstrated a high incidence of skewing, exhibited strong correlation with monocytes (r=0,881), followed with B lymphocytes (r=0,824). Strong correlation of PAmat scores between PMN and monocytes is consistent with a common ontology, as both are of myeloid origin. Correlation of PAmat scores between PMN and T lymphocytes was slightly weaker (r=0,695), consistent less contribution of T cell involvement in Xi skewing. This finding is in agreement with published results where correlation of XIRs between PMN and T cells was stronger in younger (r=0,77) versus older females (r=0,57) (Tonon et al., 1998). Monocytes, a cell type which also demonstrated a high incidence of skewing, also exhibited stronger correlation with B-lymphocytes (r=0,787) versus T lymphocytes (r=0,649). These findings suggest that relative to T lymphocytes, PMN, monocytes and B-lymphocytes share a common mechanism of skewing. Strong but incomplete correlation of PAmat scores between T and B-lymphocytes (r=0,712), suggests that in addition to a common mechanism of XIR derivation (such as derivation from a common precursor) XIRs to some extent are lineage-specific in lymphocytes.

In general, significant correlation of PAmat scores among the various hematopoietic lineages is consistent with a common mechanism of XIR derivation, namely derivation from a common hematopoietic precursor. Variability in correlation may reflect longevity of lymphocytes, lineage-specific etiologies of Xi-skewing (such as clonal hematopoiesis, stochastic processes) and technical variability.

3.7.2 Correlation of AS_{PAmat} scores among blood lineages: HSC origin of AS

To determine whether departure from the primary Xi pattern (i.e., the BC XIR) occurred in favor of the same parental X among the various hematopoietic lineages, i.e., evidence in favor of a common etiology of AS (HSC origin of AS), correlation of AS_{PAmat} scores were investigated. Albeit a slight reduction in strength, results were similar to PAmat correlation results (Table XII c); page 105) and Figure 18). As shown, correlation was strongest between cell types that demonstrated the highest incidence of AS, i.e., PMN and monocytes (r=0,846). Correlation of PMN was stronger with B cells (r=0,749) versus T cells (r=0,562), reflective of the higher incidence of AS in B versus T cells. Monocytes produced similar correlation results as PMN, exhibiting stronger correlation with B cells (r=0,728) versus T cells (r=0,550). Correlation between T and B cells was estimated at r=0,600.

Lack of stronger correlation between T and B lymphocyte AS_{PAmat} scores plausibly reflects the lower incidence of skewing in T versus B cells and lineage-specific etiologies of AS. And as observed in the T cell – B cell scatter plot of AS_{PAmat} scores, aggregation of scores within the center may further add lack of a strong correlation coefficient. The latter is consistent with minimal deviation from the primary Xi pattern.

In general, significant correlation of AS_{PAmat} scores implies that the AS trait is occurring in favor of the same parental X among the various hematopoietic lineages, consistent with a HSC origin of AS. Lack of stronger correlation may reflect lineage-specific etiologies of skewing (clonal hematopoiesis, lineage-specific selection of X-linked allele(s)), stochastic processes and technical variability.

3.8 Suggestive evidence for distinct Xi-skewing traits: PS versus AS

Although preliminary findings suggests that PS and AS are independent traits, as one is reflective of a primary event and the other acquired with advancing age, that both are putatively X-linked however, suggests that they may be one and the same, thus sharing a common mechanism of skewing. Given that the hematopoietic department is the most mitotically active tissue in the body, and provided sufficient time has elapsed, we hypothesized that AS may reflect completion of the initiated (but unfinished) primary Xi skewing trait. If true, AS should reflect selection of cells bearing the preferentially inactive X in BC tissue. Alternatively, if PS and AS are independent traits, the direction of AS should occur independent to the primary (BC) Xi pattern. For clarification, direction of skewing refers to preferential inactivation of a parental-specific X chromosome.

As shown in Figure 19 (page 101), when direction of AS (AS_{PAmat} score) was plotted relative to that of the primary XIR (BC PAmat score), direction of AS appeared to occur in either direction of the BC PAmat score, this was particularly true for PMN and monocytes. For instance, taking a relatively skewed BC PAmat score of 0,60 (i.e., preferential inactivation of Xp) and analyzing the number of individuals skewed towards Xp versus those skewed in favor of Xm (AS_{PAmat} score), it appears that AS occurred in either direction. In fact, as seen in Table XIV (page 107), when direction of AS relative to the BC PAmat score was counted and tallied in each informative subject, direction of AS was normally distributed, i.e., direction of AS was independent to the BC Xi pattern, thus suggesting AS does not reflect completion of the primary Xi trait. Rather, the data suggests that if both traits are indeed X-linked, they are genetically distinct. Alternatively, if both

traits are dictated by one locus, the data suggests selection of X-linked alleles in AS occur by a stochastic mechanism (however the latter is inconsistent with findings from twin studies).

In T and B-lymphocytes however, it appears that direction of AS was not normally distributed relative to the BC PAmat score. In fact, as seen in Table XIV (page 107), AS in lymphocytes appears to reflect selection of cells bearing the alternate inactive X chromosome ($p<0,0001$ for T lymphocytes and $p=0,02$ for B lymphocytes). These findings suggest that the etiology of AS in lymphocyte (particularly T cells) may be different to that occurring in myeloid lineages.

A limitation to the scatter-plot analyses however were the blank areas in the upper right and lower left quadrants. This limitation occurs as a result of the AS_{PAmat} calculation. Since the PAmat score of BC is subtracted from the blood PAmat score, a score difference greater than 1,0 is not possible, thus blank areas in the scatter-plots.

DISCUSSION FOR PART I

Prior to identifying and defining the etiologies and biological associations of skewing (Part II), findings supporting the existence of two distinct forms of skewing, i.e. primary versus acquired, will be discussed.

Population and specimen data

A cross-section of elderly French-Canadian females and her female siblings (and sometimes mother if living) residing within the province of Québec were accrued for study enrollment. Mean age of our population was 63,3 +/- 10 years, which was normally distributed. A mean of 5,9 females participated per family. A total of 193 families were accrued, comprising a total of 1144 females. To date, this is the largest data set comprising analysis of Xi pattern in healthy elderly females. Two biological tissues were collected for X inactivation analysis: BC for analysis of the primary Xi trait and PB for analysis of the AS trait. Since AS was speculated to reflect deviation from the body-wide XIR, we developed a model to quantitate AS by subtracting the primary XIR from the blood XIR. To our knowledge, this is the first study to perform such an analysis.

Suggestive evidence of a primary Xi trait: Primary skewing trait

In BC tissue, the incidence of skewed Xi was estimated at 12,4%, and was speculated to reflect a primary / body-wide skewing trait, supported by the following observations.

1. The DS score of BC was relatively stable with advancing age. This finding is consistent with a primary or early-onset etiology of XIR derivation (at least prior to age range 38-96) and consistent with published data where the XIR of a NHT was relatively resistant to factors (X-linked disease alleles) which can induce skewed Xi patterns (Nyhan et al., 1970), (Filosa et al., 1996).

2. The incidence of skewing in BC (12,4%) was low and similar to that observed in cord blood of neonates - previously estimated at 8,6% (Busque et al., 1996). As such, we speculate that similar rates of skewing infer a common trait, i.e., a primary etiology of XIR derivation. Nonetheless, formal evidence merits the examination of multiple tissues in neonates (constrained by practical and ethical issues however). Of special interest, that the incidence of skewing was slightly higher in BC (12,4%) versus cord blood (8,6%), raises speculation to the theory of a primary Xi pattern. Possible confounding variables include: i) that in addition to a body-wide skewing trait, skewing may be partially tissue-specific, with a higher incidence of skewed Xi in BC tissue versus cord blood. ii) Neutrophil contamination of BC scrapings (as previously cited - data not presented by the scientist however (Gale et al., 1994)). However, a) that our buccal scrapings consisted of a gentle stroking of the inner cheeks, thus unlikely to disrupt the basal membrane where neutrophils may reside; b) that leukocyte staining of a small sample of BC scrapings (n=10) did not detect leukocyte contamination (data not shown) and c) that the DS score of BC actually slightly decreased with advancing age unlike that of PMN which increased, are supportive evidence against neutrophil contamination of BC scrapings. iii) Technical variability reflecting differences in technical assessment of band quantitation. While our previous study utilized laser densitometry, the current project employed phosphor-imaging technology, which may be more sensitive and accurate, thus artifactually increasing the incidence of skewed Xi patterns.

3. Correlation analysis of PAmat scores and concordance for a skewed Xi pattern between BC and hematopoietic lineages found that PAmat scores were strongly correlated between the two tissues, with correlation coefficients approximating 0,5 ($0,46<r<0,56$), implying that the X of same parental origin was preferentially inactivated in both tissues, evidence in favor of a body-wide skewing trait. If XIRs were tissue-specific, correlation coefficients would have been lower or nil. Further, concordance for skewed Xi patterns found that BC skewing was strongly concordant with skewed Xi patterns in leukocytes, i.e., 47-64% concordance, a value higher then expected if the etiology of skewing was tissue-specific. This was further supported by analysis of direction of skewing (i.e., Xp versus Xm) among individuals whom had a skewed Xi pattern in both tissues. As found, there was ≥94% concordance, consistent with both tissues preferentially inactivating the X of same parental origin, suggesting a common mechanism of Xi skewing. Lack of further correlation / concordance

between the two tissues is possibly due to i) tissue-specific etiologies of skewing (AS, clonal hematopoiesis, stochastic processes) and ii) technical variability.

Indirect evidence that the incidence of skewed Xi in BC represents the primary Xi trait stems from the reported incidence of females manifesting X-linked recessive traits. For example, the number of females carrying a Duchenne muscular dystrophy (DMD) mutation whom manifest symptoms has approximated 10% (Norman and Harper, 1989), similar to our incidence of BC Xi skewing. Since severity of DMD has been correlated with the proportion of DMD-positive to DMD-negative cells (Pegoraro et al., 1994), we speculate that as a result of the primary Xi skewing trait, manifestation likely results from the unfortunate activation of cells expressing the mutant X-linked allele.

Intriguingly, that the DS score of BC actually slightly decreased (but insignificantly) with advancing age has raised some concern and several explanations have been put forward. First, and excluding technical variability, it may imply that with advancing age, there is a selection process favoring both Xs to be equally represented, thus suggesting a secondary mechanism of skewing albeit weak occurring in NHTs. Alternatively, assuming that BC XIRs are stable, the finding may imply that females with BC skewing are at a slight viable disadvantage, selectively eliminated with advancing age, thus suggesting BC skewing is a risk-factor for earlier mortality. Nonetheless, and assuming a genetic component to derivation of a primary skewed Xi pattern, that the trait was prevalent in 12,4% of our collection of females suggests that natural selection has acted to maintain trait variance. One functional relevance may be that a primary skewed Xi pattern increases reproductive fitness. However, as seen in Table XXXV (page 173) , the BC XIR was not associated with family data, arguing against the latter hypothesis. Nonetheless, to rule out a primary skewed Xi pattern as a risk factor for early mortality merits a prospective study analysis of longevity differences between primary Xi skewed and non-skewed females.

Further characterization of the AS trait

1. Versus the incidence of skewed Xi (DS score $\geq$ 0,25) in BC (12,4%), a skewed Xi pattern in hematopoietic lineages was more frequent: 36,0% in PMN cells, 36.6% in monocytes, 20,1% in T cells and 26.5% in B cells. The increased incidence of Xi skewing was attributed to the AS trait. Moreover, when the incidence of AS was measured by quantitating deviation from the primary / BC XIR, a significantly higher incidence of AS was observed in PMN (22,7%) and monocytes (27,2%) versus T (11,4%) and B (16,3%) lymphocytes.
2. Although the DS scores of PMN and monocyte lineages increased significantly with advancing age (Figure 13, page 95), a finding consistent with the increased incidence of Xi skewing for

these lineages (Figure 8, page 90), the DS score of T and B lymphocytes did not, thus suggesting that the AS trait is occurring at a much slower pace in these lineages, consistent with preliminary data suggesting that T lymphocytes do not contribute to the AS trait as they are long-lived (Abkowitz et al., 1998). B-estimates for the effect of age on DS scores was small ($\beta \leq 0{,}0016$), suggesting age is a small predictor of DS scores in an elderly group of females. Nonetheless, for better clarification of the age-effect on lymphocyte XIRs, a cohort analysis of females cross-sectioned into various age groups is warranted.

3. As a function of aging and unlike the DS score results, the AS_{DS} scores increased significantly with age for all four hematopoietic lineages, suggesting significant deviation from the primary XIR (BC XIR) for all hematopoietic lineages, particularly for PMN and monocytes, consistent with the highest incidence of skewed Xi patterns for these lineages. Similar to DS score-age results, B-estimates of age effect on AS (AS_{DS} scores) were small ($b \leq 0{,}0024$), consistent with a gradual process of AS (Christensen et al., 2000).

Hematopoietic lineages implicated in the AS trait

To determine the hematopoietic lineage most implicated in the AS trait, phenotypic profiles of individual hematopoietic lineages were compared. First, when analyzing the various hematopoietic lineages, a significantly higher incidence of Xi skewing (DS score $\geq 0{,}25$) was observed in PMN (36,0%) and monocytes (36,6%) versus T (20,1%) and B lymphocytes (26,5%), suggesting myeloid lineages are most involved in the AS trait. Similar results were obtained when AS_{DS} scores were analyzed and when qualitative methods such as AS_{QL} and AS_{QT} were employed to quantitate the incidence of AS. The lower incidence of Xi skewing in T cells is consistent with previous findings (Champion et al., 1997), (Gale et al., 1997), (Tonon et al., 1998) and speculated to reflect longevity of T cells (Abkowitz et al., 1998). Alternatively, since PMN, monocytes, and B-lymphocytes are bone marrow-derived, while T lymphocytes are thymus- and/or extrathymically derived, the microenvironment may also have an influence on derivation of Xi patterns. Interestingly, although the incidence of Xi skewing was higher in B versus T cells, it was lower versus myeloid lineages. If consistent with cell longevity as a variable for implication in AS, the finding suggests that B cells are relatively less long-lived versus T cells. To our knowledge, this is the first study providing Xi ratio analysis of B-lymphocytes as a function of aging.

Interestingly, since results of the quantitative method to measure the incidence of AS, i.e., AS_{QT}, reported an incidence of AS approximately 10 percentage points lower versus the AS_{QL} method, and that this difference of 10% is similar to the incidence of Xi skewing observed in BC (12,4%), implies that the AS_{QT} method may be a more realistic and accurate method to report the true

incidence of AS. Further, this finding supports our method to quantitate AS by subtracting the BC PAmat score from blood PAmat score.

AS represents departure from the primary XIR

Indirect evidence that AS represents deviation from the primary Xi pattern was derived from the age-AS_{DS} correlation analyses. As observed in Figure 15 (page 97), when the line of best-fit was extrapolated to age 30, an AS_{DS} score of approximately 0,1 was observed for all lineages, consistent with strong correlation of PAmat scores between the two tissues in young females. however, as a function of aging, correlation of PAmat scores between the two tissues increased significantly for all four lineages, particularly for PMN and monocytes, consistent with a higher incidence of AS in myeloid versus lymphoid lineages. That the BC DS score was stable with aging suggests that AS represents deviation from the primary XIR.

HSC origin of AS

As presented in Table XII b) and c) (page 105) and Figures 17 and 18 (pages 99, 100), strong correlation of PAmat and AS_{PAmat} scores among the various hematopoietic lineages ($0{,}649 < r > 0{,}881$ for PAmat scores and $0{,}550 < r > 0{,}846$ for AS_{PAmat} scores), is consistent with a common mechanism of XIR derivation among the various hematopoietic lineages, i.e., a HSC origin of AS. In n=3 females demonstrating a skewed phenotype in PB, found was a skewed Xi pattern not only among differentiated hematopoietic lineages (i.e., CD3+, CD56+ and CD14+) but also among committed (CD34+ CD38+) and less committed (CD34+ CD38-) progenitor cells (see Annex 3). Lack of complete (100%) correlation may reflect lineage-specific etiologies of Xi skewing (clonal hematopoiesis, genetic heterogeneity of X-linked genes contributing to AS and/or stochastic processes), technical variability and/or longevity of lymphocytes. To determine to what extent these may contribute to variability in correlation coefficients necessitates further investigation. For example, clonal derivation of cells can be evaluated by identification of genetic alterations (which can provide the lineage with a proliferative advantage) in skewed versus non-skewed females. Genetic heterogeneity can be evaluated by performing quantitative trait loci mapping of XIRs from the various hematopoietic lineages. Identification of linkage to different regions of the genome in a lineage-specific manner is consistent with genetic heterogeneity. Longevity of lymphocytes can be investigated by analyzing the Xi pattern of individual hematopoietic lineages in pre- versus post-bone marrow transplantations, as performed in felines (Abkowitz et al., 1998). Lack of a positive finding to any of the above supports stochastic processes, technical variability and/or unidentified factors in derivation of XIRs among the various hematopietic lineages.

Suggestive evidence that primary skewing and AS are distinct traits

Several observations supporting distinct etiologies of primary skewing and AS traits were made. First, unlike the primary XIR (BC DS score), which was relatively stable with advancing age, AS_{DS} scores of hematopoietic lineages increased significantly with advancing age. Secondly, when the direction (Xi skewing in favor of Xp versus Xm) of AS (AS_{PAmat} score) was compared relative to the primary XIR (BC PAmat score), it was found uninfluenced by the direction of Xi skewing in BC, suggesting that AS does not reflect completion of the PS trait. This was particularly true for PMN and monocytes (Figure 19, page 101;Table XIV, page 107). Interestingly, for T and B lymphocytes, the direction of AS occurred preferentially in the opposite direction of primary skewing (this finding suggests that with advancing age, AS in T and B cells reflects selection of lymphocytes harboring the alternate parental X chromosome active). For example, if Xm was preferentially active in BC tissue, it was relatively less active in lymphocytes. Nonetheless, the latter is still consistent with an X-linked genetic component to derivation of AS_{DS} scores, albeit by a different mechanism versus myeloid lineages.

In light of these findings, it appears that BC Xi skewing and blood skewing (AS) are disparate traits, thus validating our approach to dissect the skewing trait into subphenotypes. This finding possibly clarifies why two loci were mapped to the X chromosome in prior attempts to map the Xi skewing trait (Naumova et al., 1998). We speculate that one region corresponds to the primary skewing trait, and the other to the AS trait. Nonetheless, to derive formal evidence for a genetic component to Xi skewing, familial resemblance, segregation analyses and heritability analyses of both traits was conducted (Results, Part II). As such, deriving evidence of heritability will provide rational support to map the trait(s).

Clinical implications

Interpretation of Xi patterns for myeloproliferative disorders

The demonstration of a clonal (skewed) Xi pattern in myeloid cells but a random Xi pattern in a control tissue; usually skin, cultured fibroblasts or T cells, has been readily used as a marker for a myeloproliferative disorder (AML for example). A problem however arises as both the primary skewing and AS traits can mimic clonal derivation of cells, thus rendering clonality assay results uninterpretable in the absence of a suitable control tissue. Clearly, to rule our primary skewing (as the primary XIR is body-wide), Xi analysis of BC is required. Moreover, to rule out AS, B cells, as they appear more implicated in the AS trait versus T cells, should be used as a control tissue.

Control tissue for X-linked disorders:

Since the degree of clinical manifestation in female carriers of X-linked disorders is correlated with the degree of Xi skewing, Xi pattern analysis of affected tissues (i.e., tissues that express the gene in question) is necessitated. In most instances however, PB is analyzed. However, as there are at least two Xi traits, analysis of the appropriate tissue is required. For example, for immune/hematopoietic disorders, the Xi pattern of PB and/or the specific hematopoietic lineage should be investigated. Alternatively, for disorders affecting NHTs, the affected tissue should be analyzed. However, if this is not possible (due to impossibility of obtaining a tissue-biopsy), BC tissue, as it is representative of the primary (body-wide) XIR, should be used as a surrogate.

Figure 6. Distribution of family size, i.e., number of participant female siblings per family. N=193 families, median sibling size of five, mean of 5,6.

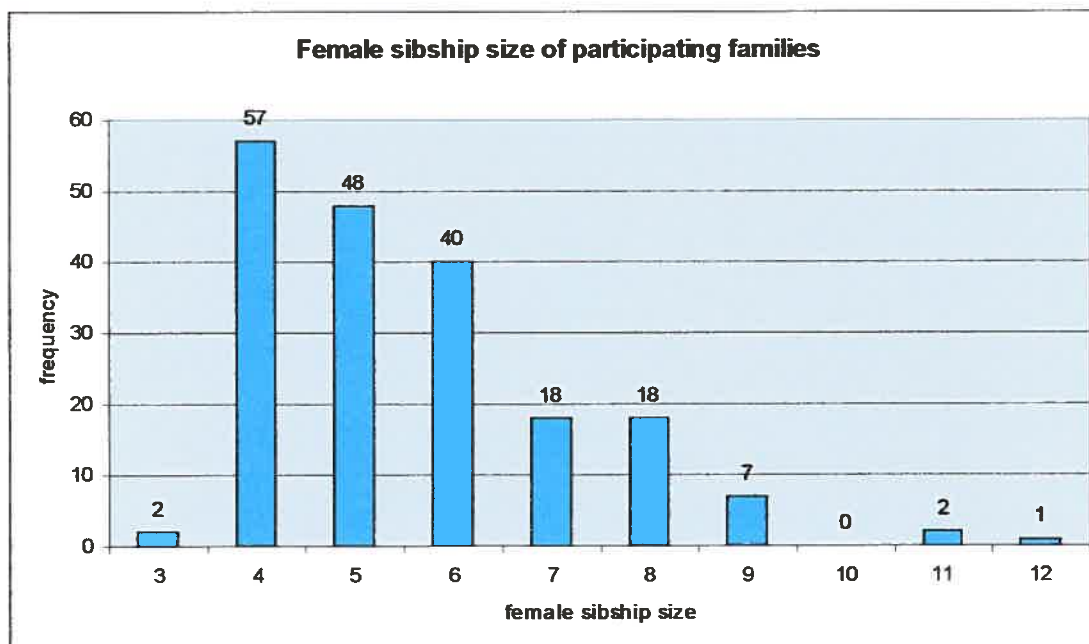


Figure 7. Age distribution of female participants. Ages ranged from 38 to 96 years, with a mean of 63,3 years ± 10,0.

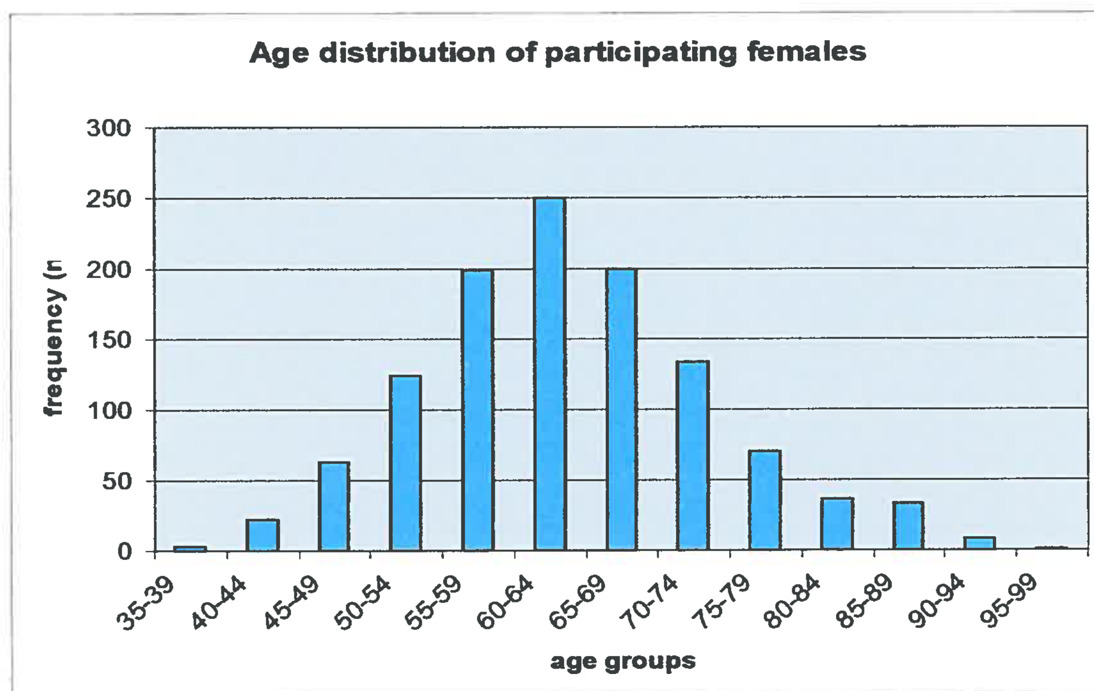


Figure 8. Incidence of skewed X inactivation (DS score ≥ 0,25). The percentage of females skewed for each cell type is indicated above the bar. For buccal cells, the incidence of skewing speculatively corresponds to the frequency of primary skewing (PS). The incidence of skewing in blood hypothetically reflects the contribution from both the primary skewing and AS traits.

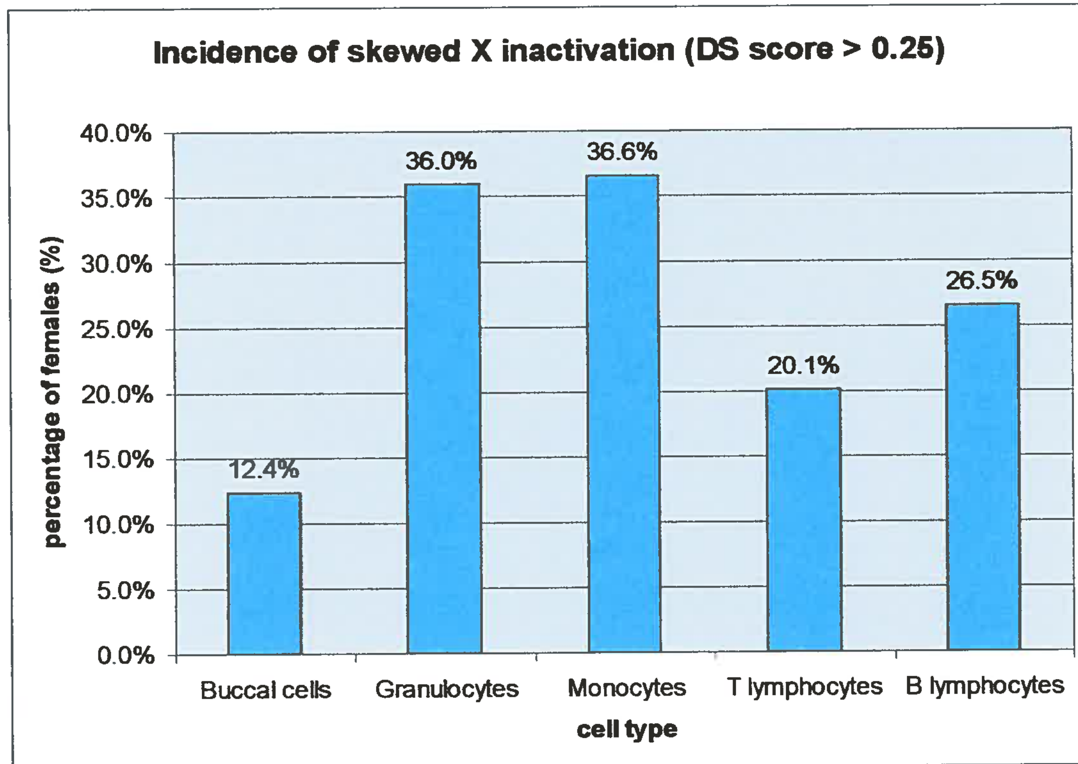
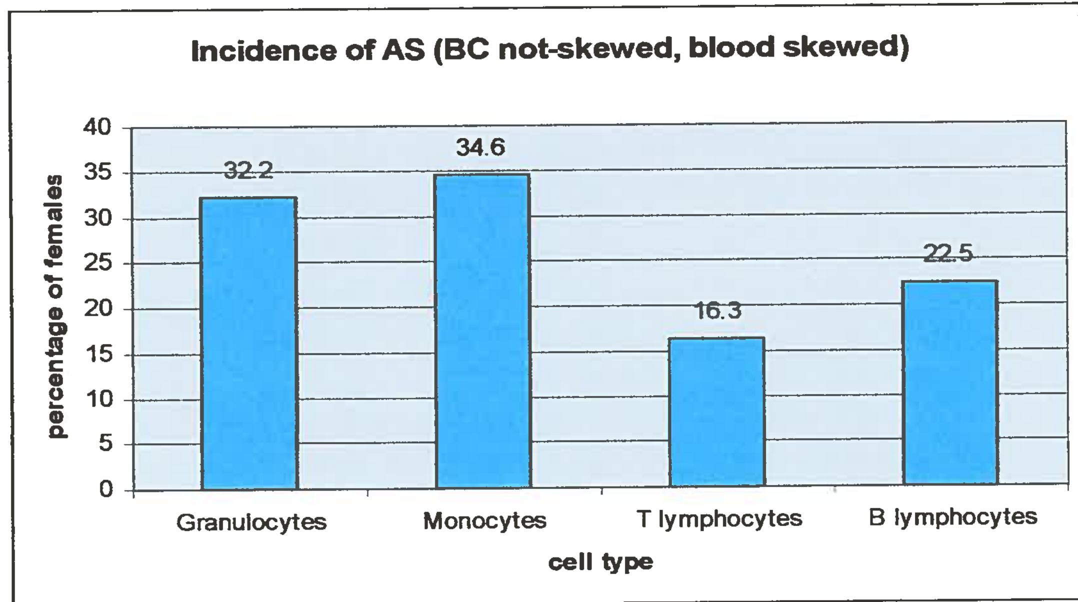


Figure 9. Incidence of AS in hematopoietic lineages. A) Qualitative analyses (skewed in blood, not skewed in BC). B) Quantitative analyses (AS_{DS} score ≥ 0,25).

a) Incidence of AS - qualitative analyses: AS_{QL}



b) Incidence of AS - quantitative analyses: AS_{QT}

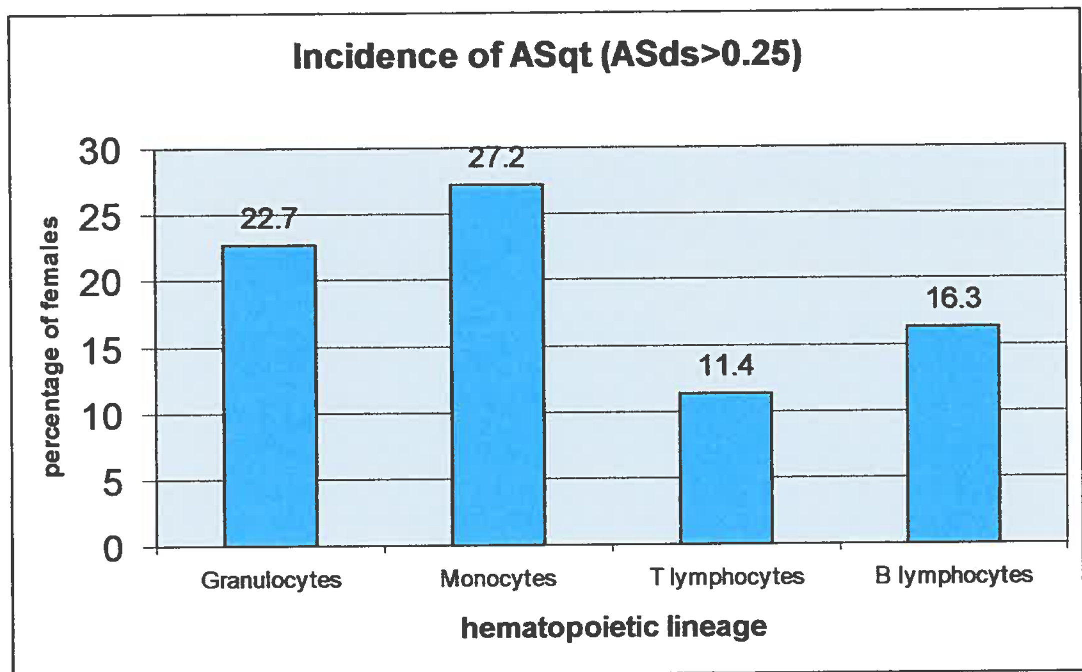


Figure 10. Frequency distribution of DS scores (range: 0 to 0,5).

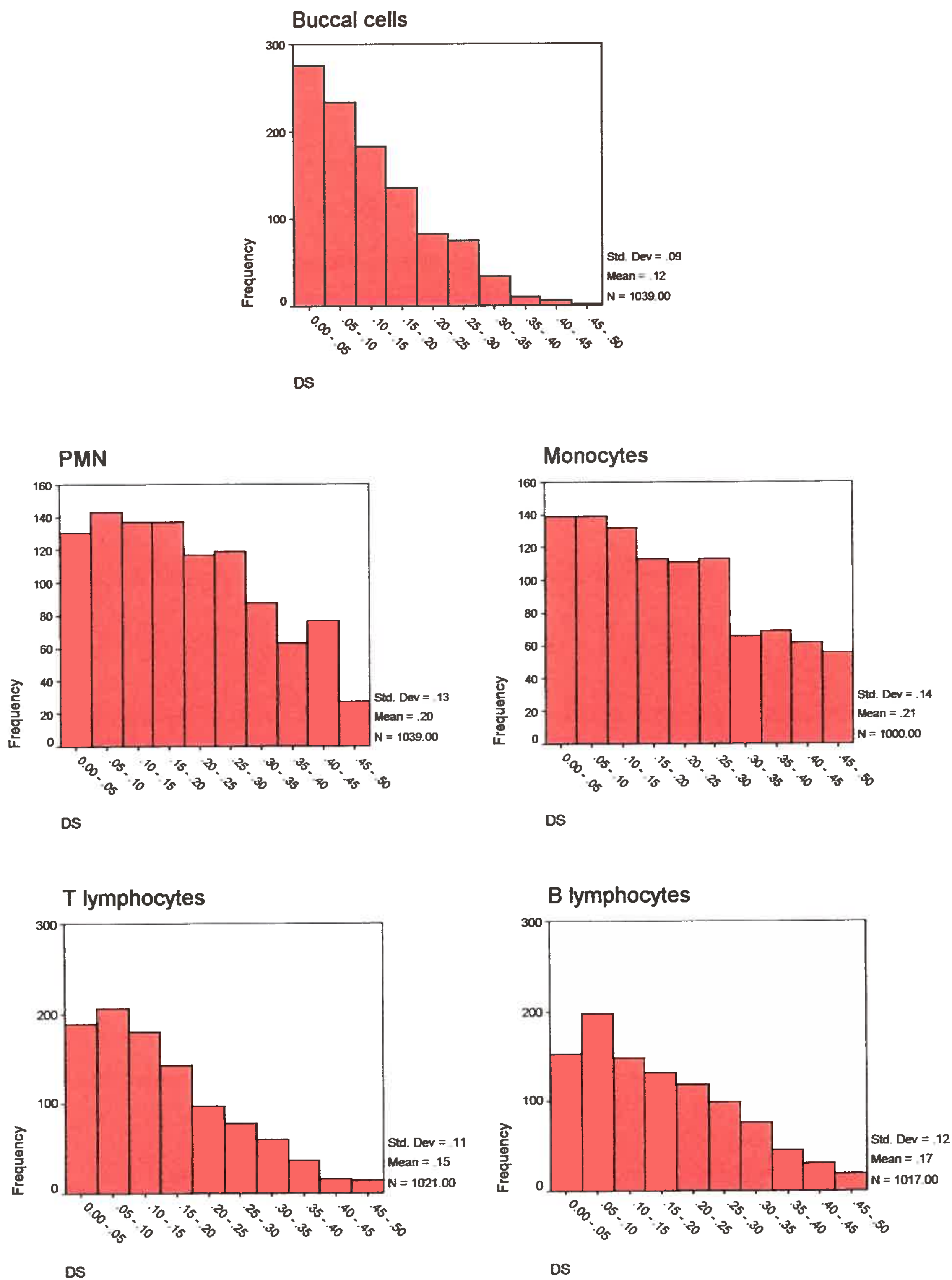
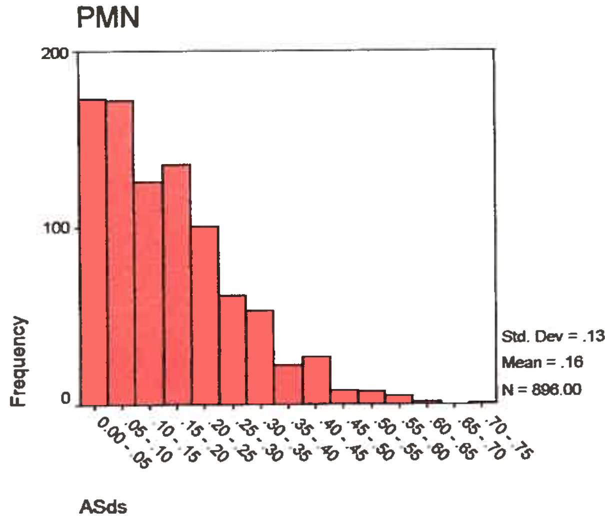
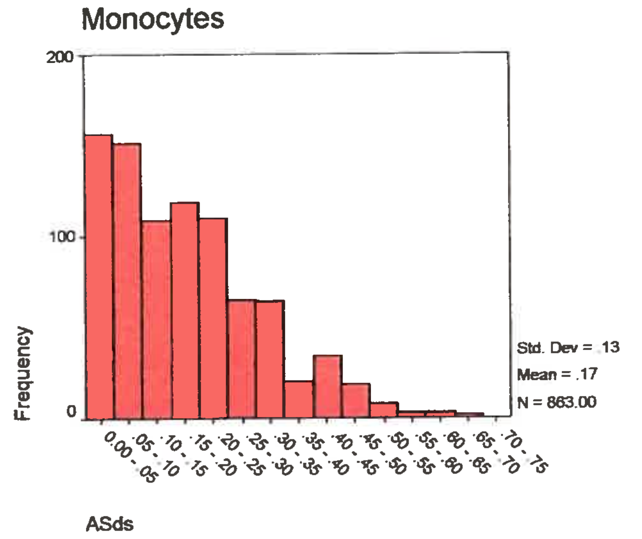
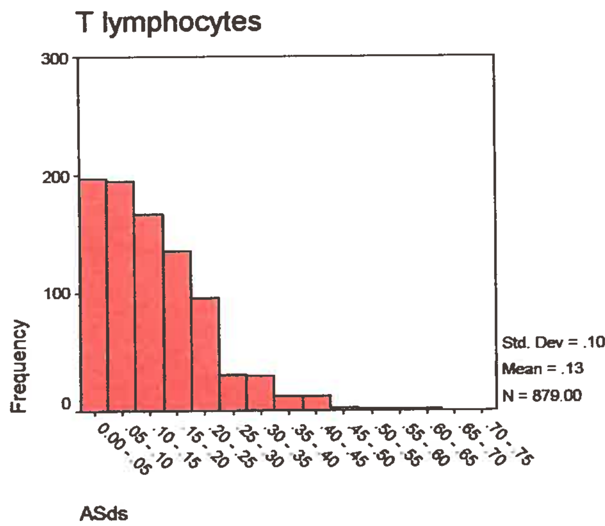


Figure 11. Frequency distribution of AS_{DS} scores (range: 0 to 0,75).







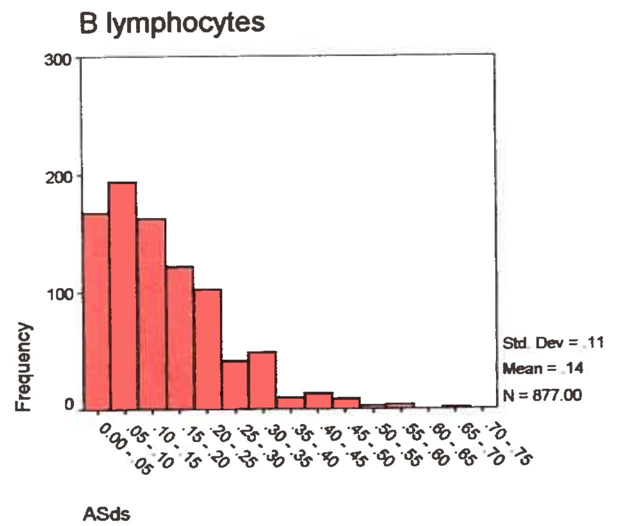
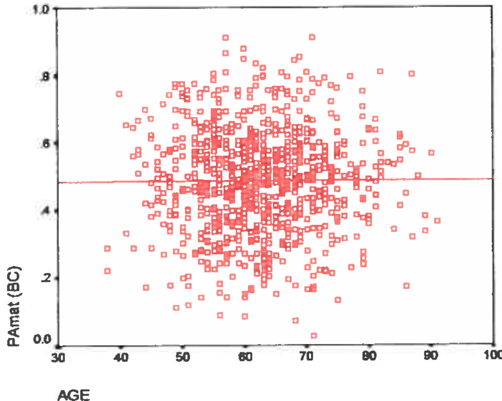
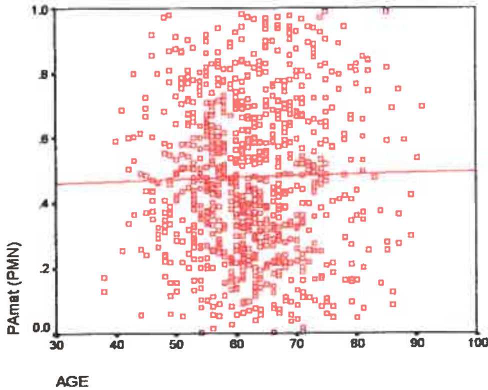


Figure 12. Age-PAmat score scatter-plots. Results of LR method provided. a) BC, b) PMN, c) monocytes, d) T cells, e) B cells.

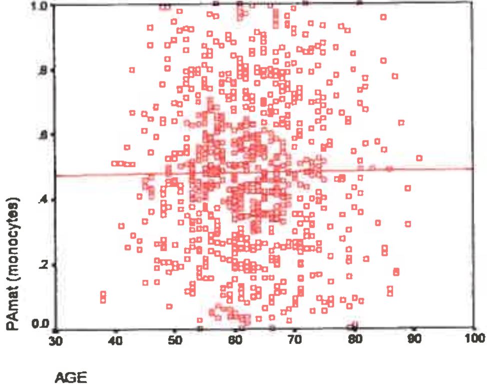
a) BC (p= 0,851, b= 0,0001)



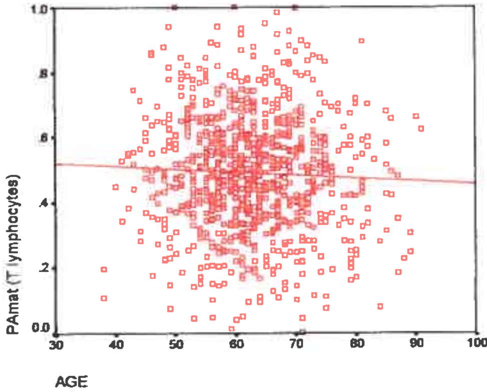
b) PMN (p= 0,556, b=0,0005)



c) monocytes (p= 0,841, b=0,0002)



d) T cells (p= 0,171 b= -0,0009)



e) B cells (p= 0,738, b= -0,0003)

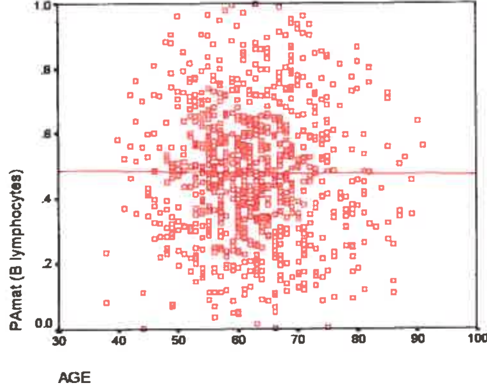
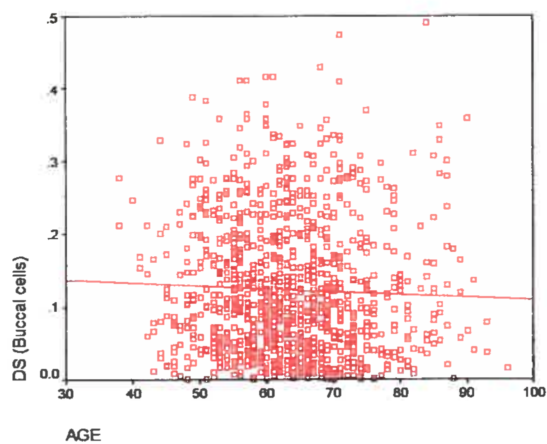
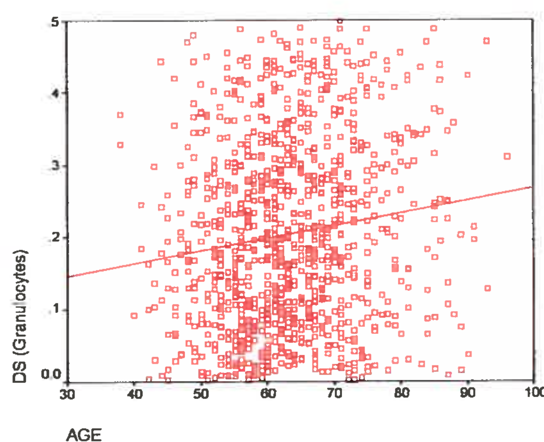


Figure 13. Age-DS score (skewing) scatter-plots. Results of GEE method provided. a) BC, b) PMN, c) monocytes, d) T lymphocytes, e) B lymphocytes.

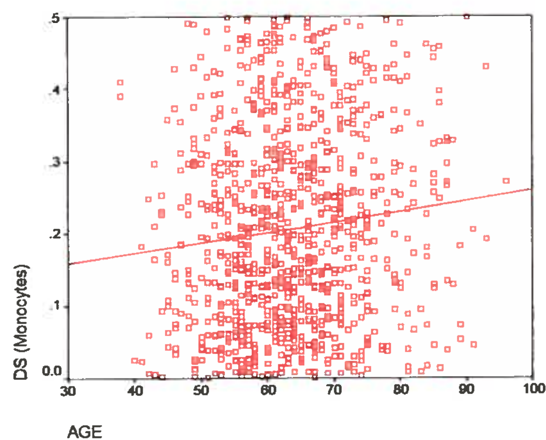
a) BC (p= 0,2069, b= -0,0004)



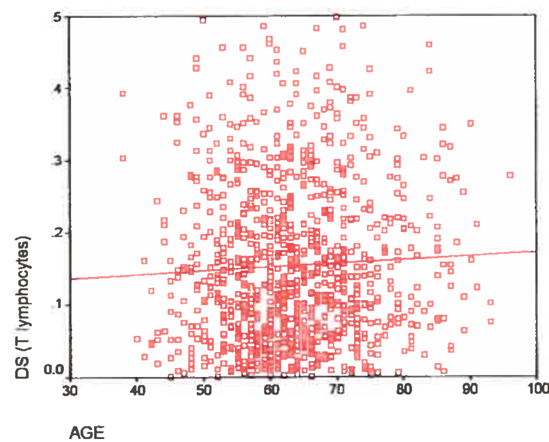
b) PMN (p= 0,0004, b= 0,0016)



c) Monocytes (p= 0,0056, b= 0,0013)



d) T cells (p= 0,1341, b= 0,0006)



e) B cells (p= 0,3481, b= 0,0004)

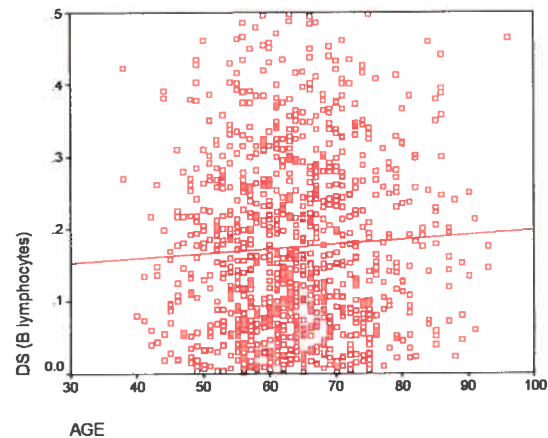
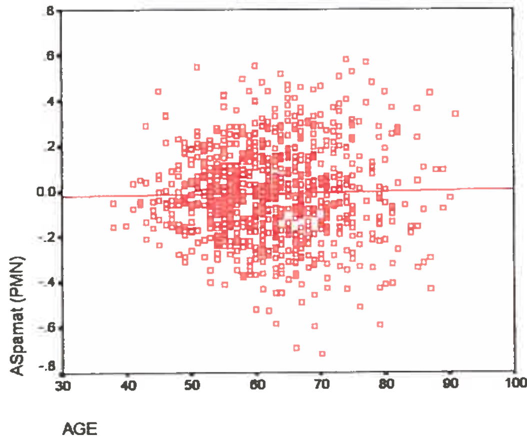
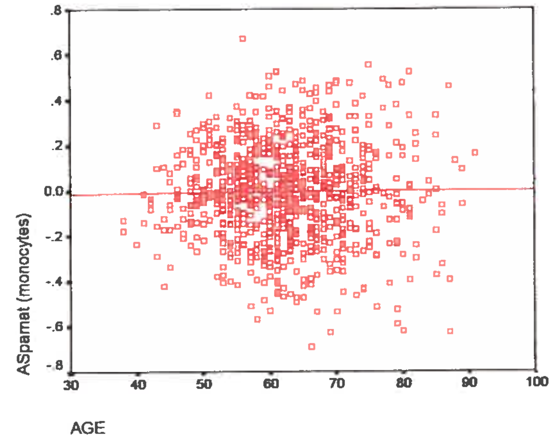


Figure 14. Age-AS_{PAmat} score scatter-plots. Results of LR method provided. a) PMN, b) monocytes, c) T cells, d) B cells.

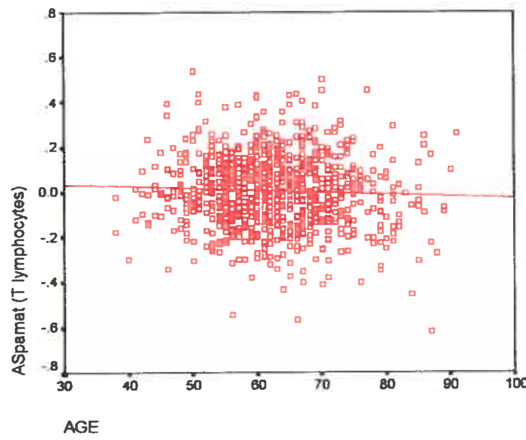
a) PMN (p= 0,582, b= 0,0004)



b) Monocytes (p= 0,782, b= 0,0002)



c) T cells (p= 0,115, b= -0,0009)



d) B cells (p= 0,596, b= -0,0004)

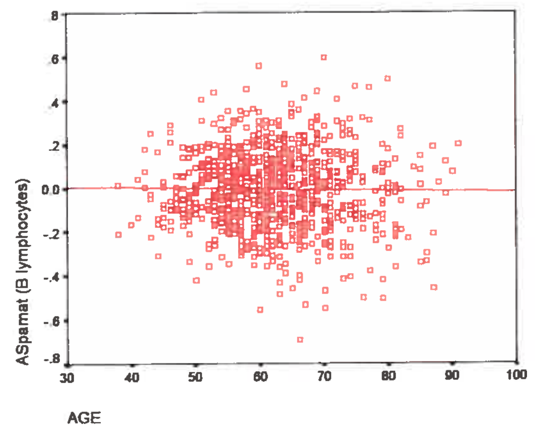
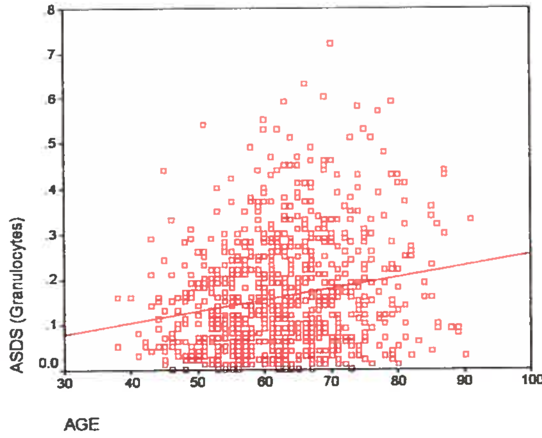
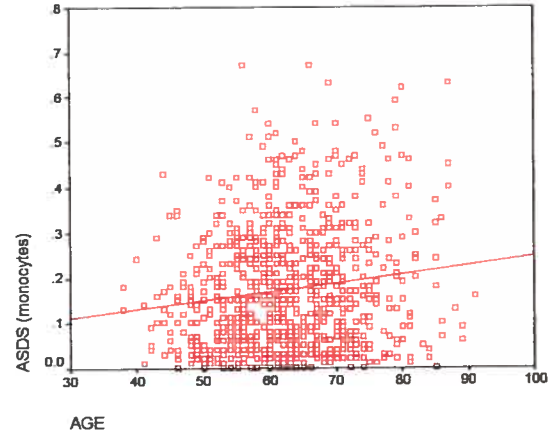


Figure 15. Age-AS_{DS} score (AS) scatter-plots. Results of GEE statistical method provided. a) PMN b) monocytes c) T lymphocytes d) and B lymphocytes. Lower AS_{DS} scores (stronger correlation) in younger versus older females is consistent with AS representing deviation from the primary Xi ratio.

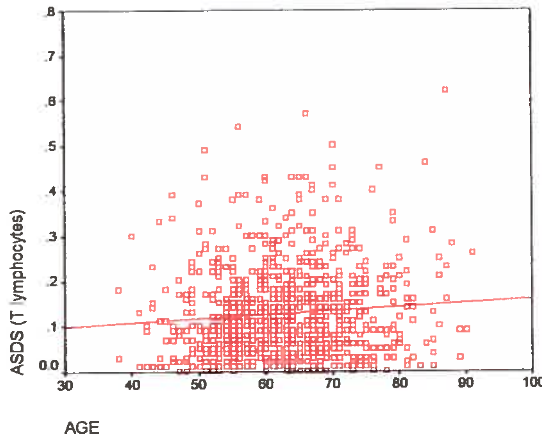
a) PMN ($p<0,0001$, $b=0,0024$)



b) Monocytes ($p=0,0004$, $b=0,0019$)



c) T cells ($p=0,0378$, $b=0,0009$)



d) B cells ($p=0,0021$, $b=0,0013$)

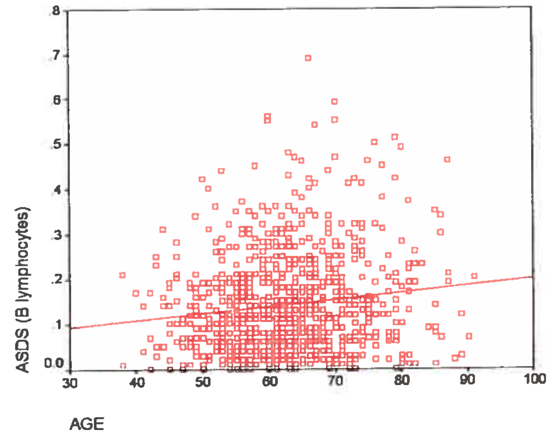
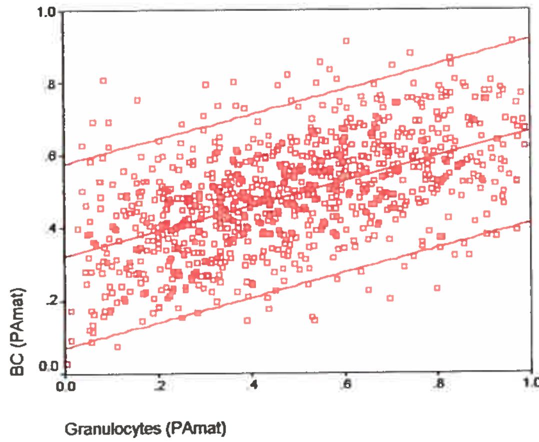
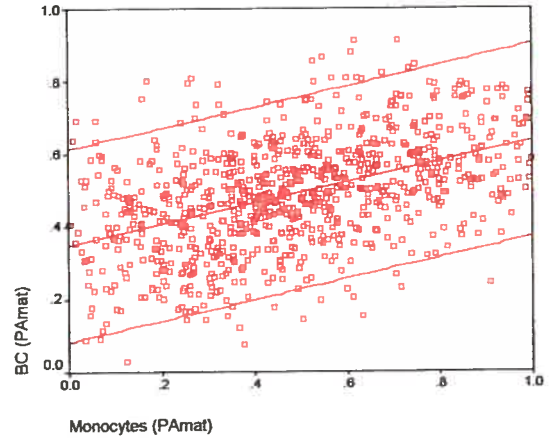


Figure 16. Primary XIR: intra-individual correlation of PAmat scores between BC and the various hematopoietic lineages. Regression prediction line and 95% confidence intervals are shown. See Table XII (page 105) for table format of results.

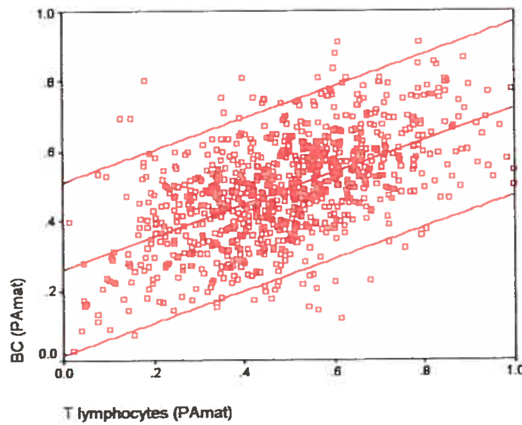
a) BC-PMN correlation
(r = 0,535 , $p<10^{-17}$)



b) BC-monocytes correlation
(r = 0,462 $p<10^{-17}$)



c) BC-T cell correlation
(r=0,563 , $p<10^{-17}$)



d) BC-B cell correlation
(r=0,546, $p<10^{-17}$)

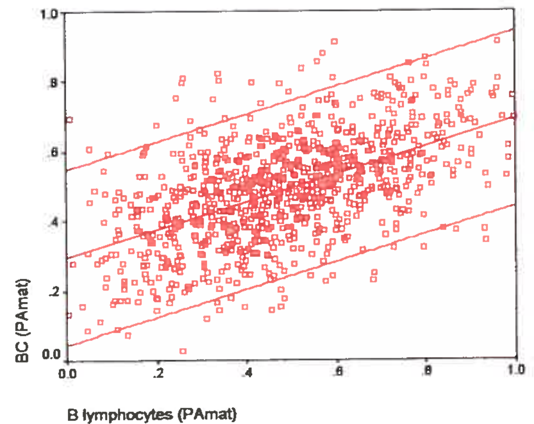
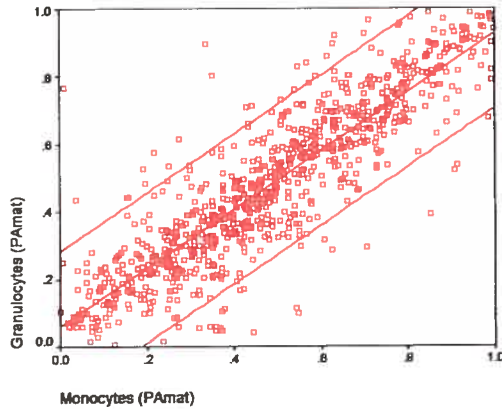
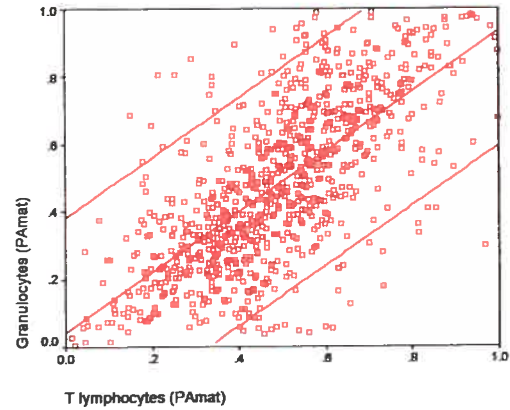


Figure 17. Blood skewing: intra-individual correlation of PAmat scores among hematopoietic lineages. Regression prediction line and 95% confidence intervals are shown. See Table b) XII (page 105) for table format of results.

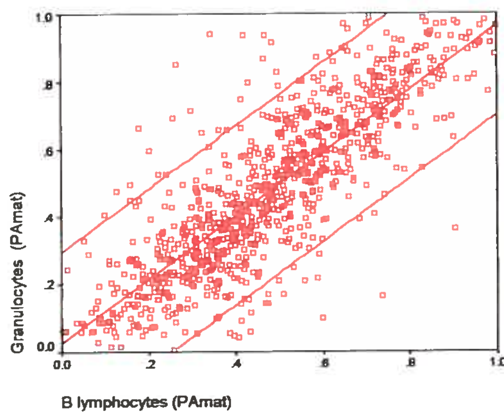
PMN - monocytes ($r=0,881$, $p<10^{-17}$)



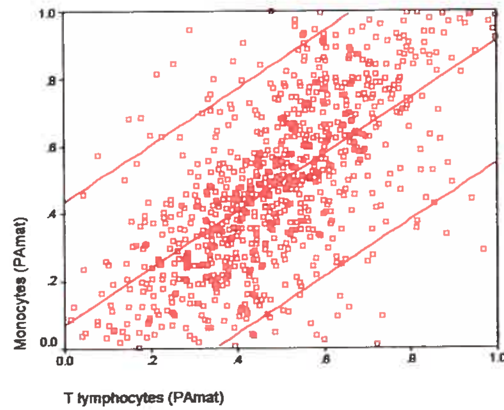
PMN - T cells ($r=0,695$, $p<10^{-17}$)



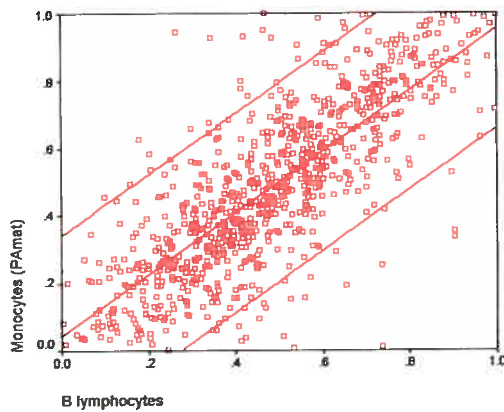
PMN - B cells ($r=0,824$, $p<10^{-17}$)



Monocytes – T cells ($r=0,649$, $p<10^{-17}$)



Monocytes – B cells ($r=0,787$, $p<10^{-17}$)



T cells – B cells ($r=0,712$, $p<10^{-17}$)

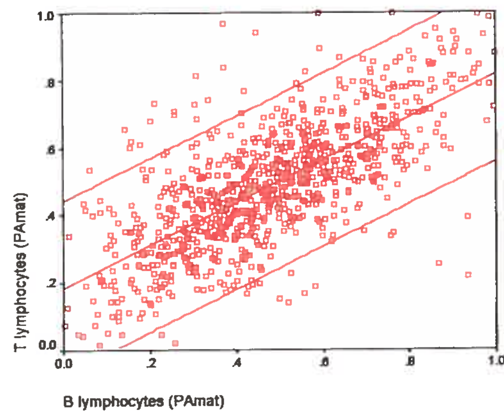
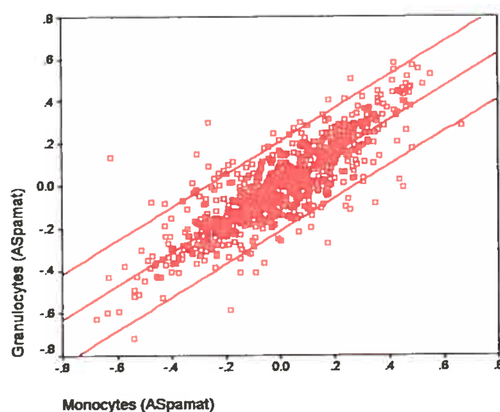
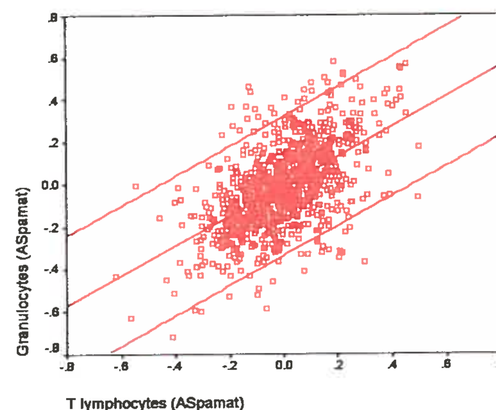


Figure 18. AS: intra-individual correlation of AS_{PAmat} scores among hematopoietic lineages. Regression prediction line and 95% confidence intervals are shown. See Table XII c) (page 105) for table format of results.

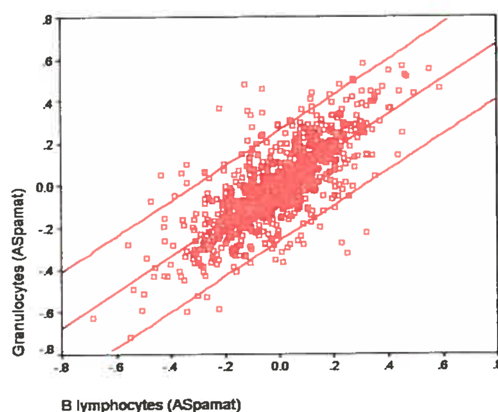
PMN – monocytes (r=0,846, $p<10^{-17}$)



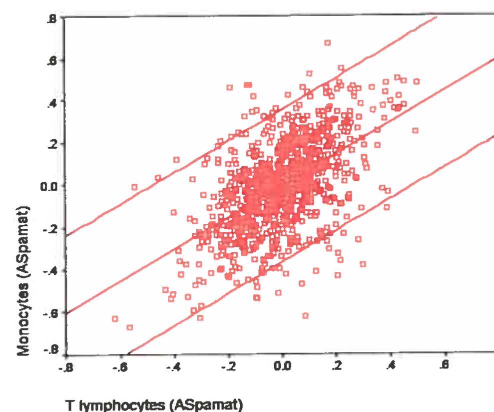
PMN – T cells (r=0,562, $p<10^{-17}$)



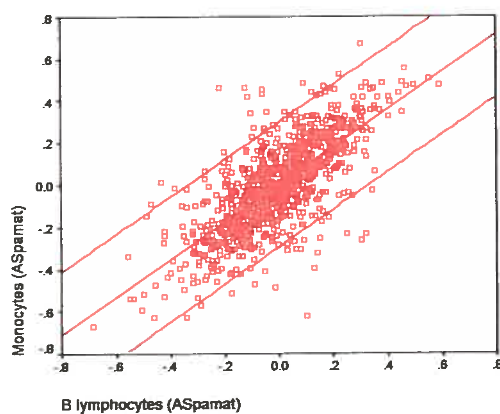
PMN – B cells (r=0,749, $p<10^{-17}$)



monocytes – T cells (r=0,550, $p<10^{-17}$)



Monocytes – B cells (r=0,728, $p<10^{-17}$)



T cells – B cells (r=0,660, $p<10^{-17}$)

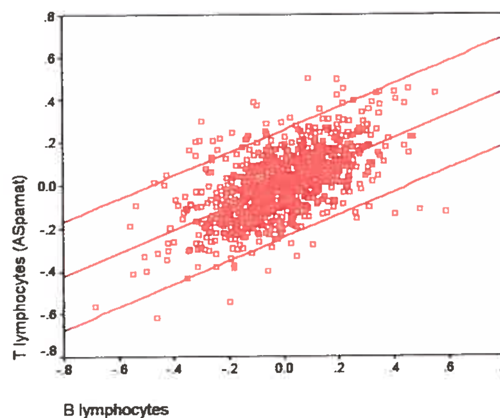
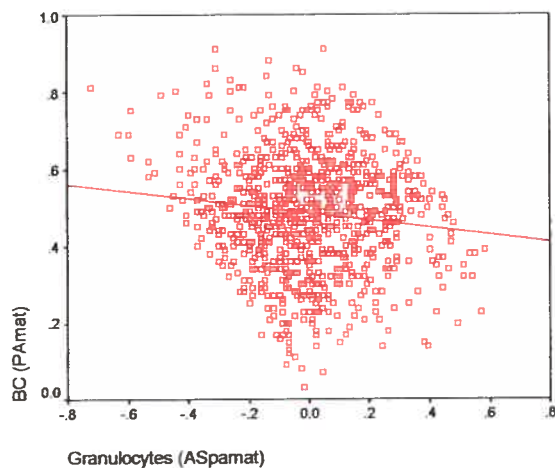
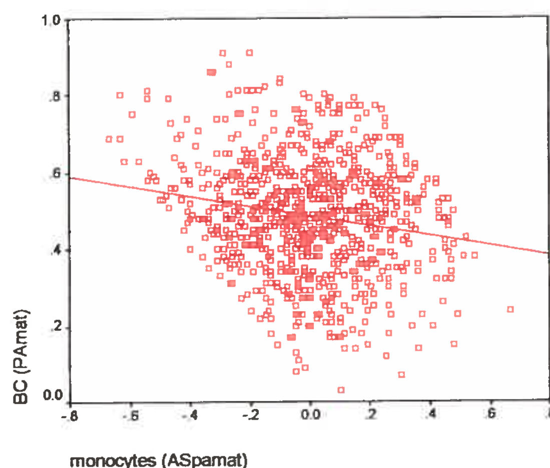


Figure 19. Direction of AS relative to the primary XIR: suggestive evidence for distinct Xi skewing traits. *Method*: The AS_{PAmat} score was analyzed as a function of the primary XIR (BC PAmat score). Scatter plots for each hematopoietic lineage are shown. The BC PAmat score ranged from 0 to 1, a score of 1 defining complete inactivation of the paternal X. In leukocytes, AS_{PAmat} scores ranged from –0,8 to 0,8, a score of +0,8 defining preferential inactivation of the paternal X. As shown, the direction of AS in PMN and monocytes appears to occur random to the BC PAmat score, suggesting the AS trait does not reflect a continuation of the primary Xi trait. In T and B lymphocytes however (particularly for T lymphocytes), direction of AS does not appear normally distributed. In fact, as seen in Table XIV (page 107), T cell AS relative to the BC XIR appears to reflect selection of cells bearing the alternate inactive X. Regression prediction lines are shown.

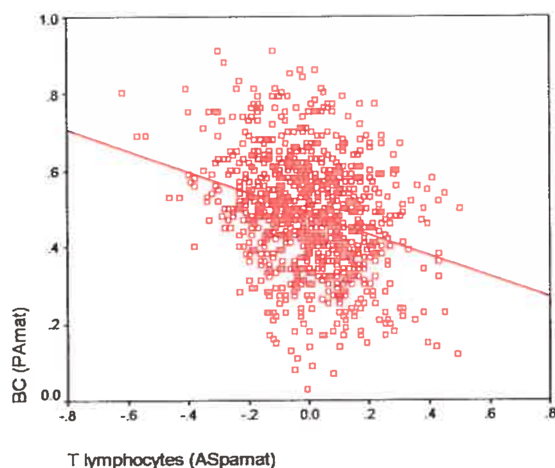
a) BC PAmat – PMN AS_{PAmat}



b) BC PAmat – monocytes AS_{PAmat}



c) BC PAmat – T cells AS_{PAmat}



d) BC PAmat – B cells AS_{PAmat}

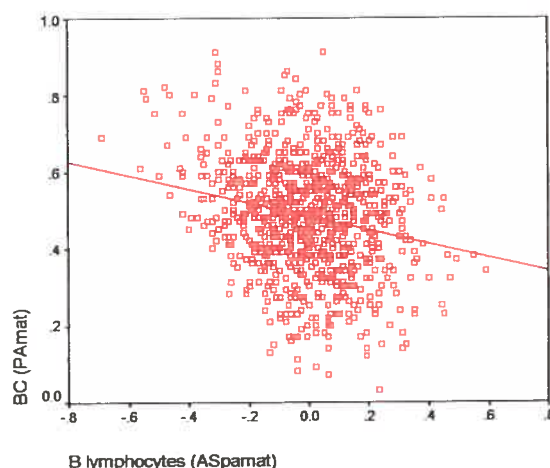


Table VIII. Number of informative females (DNA amplified and heterozygous for the HUMARA clonality assay). For buccal cells and granulocytes, all DNA samples amplified, deriving a heterozygosity rate of 90,8%.

Cell type	Number of samples	Number of informative samples (DNA amplified and heterozygous)
Buccal cells	1144	1039 (90,8%)
PMN (granulocytes)	1144	1039 (90,8%)
CD14 (monocytes)	1144	1000 (87,4%)
CD3 (T lymphocytes)	1144	1026 (89,7%)
CD19 (B lymphocytes)	1144	1022 (89,3%)

Table IX. Incidence of skewing (DS score ≥ 0,25) and AS (three methods given: estimated, AS_{QL}, AS_{QT}). Of the three methods used to measure the incidence of AS (see table below and text for more details), the AS_{QT} score may be a more accurate method as it utilizes a more stringent criterion of skewing.

Cell type	Skewed X inactivation (DS score ≥ 0,25)	Estimated incidence of AS (incidence of blood skewing - incidence of BC skewing)	AS_{QL} -skewed in blood -not-skewed in BC	AS_{QT} (AS_{DS} score ≥ 0,25)
Buccal cell	12,4 %	-	-	-
PMN	36,0 %	23,6 %	32,2 %	22,7 %
Monocyte	36,6 %	24,2 %	34,6 %	27,2 %
T cells	20,1 %	7,7 %	16,3 %	11,4 %
B cells	26,5 %	14,1 %	22,5 %	16,3 %

Table X. Mean X inactivation ratios. a) Mean DS scores and standard deviations. b) Mean AS_{DS} scores and standard deviations.

a) Mean DS scores

Cell type	n	DS score (mean)	min	max	Std. deviation (+/-)
BC	1034	0,123	0,000086	0,5	0,094
PMN	1033	0,204	0,001001	0,5	0,129
Monocytes	995	0,207	0,000066	0,5	0,136
T lymphocytes	1016	0,153	0,000359	0,5	0,111
B lymphocytes	1012	0,174	0,001910	0,5	0,112

b) Mean AS_{DS} scores

Cell type	n	AS_{DS} score (mean)	min	max	Std. deviation (+/-)
PMN	874	0,159	0,000012	0,722287	0,125
Monocytes	840	0,173	0,000557	0,692075	0,133
T lymphocytes	858	0,127	0,000007	0,618584	0,099
B lymphocytes	854	0,142	0,000146	0,692075	0,110

Table XI. Skewing as a function of age. Results of LR, GEE and SOLAR statistical methods are provided (when available).

a) PAmat score

	LR method		**GEE method**		**SOLAR method**
Cell type	*p-value*	*beta*	*p-value*	*beta*	*p-value*
BC	0,851	0,0001			
PMN	0,556	0,0005			
Monocytes	0,841	0,0002			
T lymphocytes	0,171	- 0,0009			
B lymphocytes	0,738	- 0,0003			

b) DS score (PS)

	LR method		**GEE method**		**SOLAR method**
Cell type	*p-value*	*beta*	*p-value*	*beta*	*p-value*
BC	0,1752	- 0,0004	0,2069	- 0,0004	0,29
PMN	<0,0001	0,0018	0,0004	0,0016	0,00009
Monocytes	0,0007	0,0015	0,0056	0,0013	0,005
T lymphocytes	0,1224	0,0006	0,1341	0,0006	0,28
B lymphocytes	0,0827	0,0007	0,3481	0,0004	0,23

c) AS_{PAmat} score

	LR method		**GEE method**		**SOLAR method**
Cell type	*p-value*	*beta*	*p-value*	*beta*	*p-value*
PMN	0,582	0,0004			
Monocytes	0,782	0,0002			
T lymphocytes	0,115	- 0,0009			
B lymphocytes	0,596	- 0,0004			

d) AS_{DS} score (AS)

	LR method		**GEE method**		**SOLAR method**
Cell type	*p-value*	*beta*	*p-value*	*beta*	*p-value*
PMN	<0,0001	0,0025	<0,0001	0,0024	9×10^{-6}
Monocytes	0,0001	0,0019	0,0004	0,0019	0,0035
T lymphocytes	0,0095	0,0009	0,0378	0,0009	0,13
B lymphocytes	0,0002	0,0015	0,0021	0,0013	0,14

Table XII. Intra-individual correlation of X inactivation ratios, a-c.

a) Correlation between BC and blood cells (PAmat score).

Correlation (PAmat)	BC	PMN	Monocytes	T cells	B cells
BC					
Pearson correlation	1,000	0,535	0,462	0,563	0,546
Sig. (2-tailed)		$<10^{-17}$	$<10^{-17}$	$<10^{-17}$	$<10^{-17}$
N	896	896	863	878	876

b) Correlation among blood cells (PAmat score).

Correlation (PAmat)	PMN	Monocytes	T cells	B cells
PMN				
Pearson correlation	1,000	0,881	0,695	0,824
Sig. (2-tailed)		$<10^{-17}$	$<10^{-17}$	$<10^{-17}$
N	896	863	878	876
Monocytes				
Pearson correlation		1,000	0,649	0,787
Sig. (2-tailed)			$<10^{-17}$	$<10^{-17}$
N		863	851	849
T cells				
Pearson correlation			1,000	0,712
Sig. (2-tailed)				$<10^{-17}$
N			878	866
B cells				
Pearson correlation				1,000
Sig. (2-tailed)				
N				876

c) Correlation among blood cells (AS_{PAmat} score).

Correlation (PAmat)	PMN	Monocytes	T cells	B cells
PMN				
Pearson correlation	1,000	0,846	0,562	0,749
Sig. (2-tailed)		$<10^{-17}$	$<10^{-17}$	$<10^{-17}$
N	887	860	872	871
Monocytes				
Pearson correlation		1,000	0,550	0,728
Sig. (2-tailed)			$<10^{-17}$	$<10^{-17}$
N		860	848	847
T cells				
Pearson correlation			1,000	0,600
Sig. (2-tailed)				$<10^{-17}$
N			872	861
B cells				
Pearson correlation				1,000
Sig. (2-tailed)				
N				871

Table XIII. Evidence supporting a primary Xi skewing trait (PS). *Methods*: Females with a skewed Xi pattern in BC were examined for Xi pattern analysis in hematopoietic lineages. We hypothesized that in the event of a primary Xi skewing trait, BC skewing should be highly concordant with leukocyte skewing (body-wide skewing). On the other hand, if Xi-skewing were tissue-specific, the incidence of leukocyte skewing should be unrelated to BC skewing, thus expected to reflect the incidence of skewing observed in our cohort population of females (column 3). As observed in column 4, concordance for a skewed Xi pattern between the two tissues varied from 53,6 to 64,3%, supporting both a body-wide and lineage-specific etiology of skewed Xi patterns. Other etiologies include the AS trait, clonal derivation of cells, stochastic processes, methodologies used and/or technical variability. Moreover, as direction of skewing (preferential inactivation of a parental-specific X chromosome) was highly concordant between the two tissues (94 to 98,6%, column 5) suggests a common mechanism in derivation of skewed Xi patterns between the two tissues, congruent with an X-linked etiology or derivation from a common stem cell pool.

Cell type	Expected incidence of skewing in blood if skewing were body-wide (%)	Expected incidence of skewing in blood if skewing were tissue-specific (%)	Observed incidence of skewed Xi in blood when skewed in BC (%)	Concordance in direction of skewing between the two tissues (%)
PMN	100	36,1	64,3 (83/129)	94,0 (78/83)
Monocytes	100	37,3	53,6 (67/125)	94,0 (63/67)
T lymphocytes	100	20,2	47,2 (60/127)	98,3 (59/60)
B lymphocytes	100	26,8	56,7 (72/127)	98,6 (71/72)

Table XIV. *Direction of AS relative to the BC XIR. Method*: The direction of AS relative to the BC PAmat score was analyzed for each hematopoietic lineage. Briefly, in informative subjects the BC PAmat score was dichotomized as preferential inactivation of a paternal X (PAmat > 0,50) or maternal X (PAmat < 0,50). Next, the PAmat score of blood as a function of BC PAmat was analyzed. Direction of AS in blood was dichotomized as concordant with direction of skewing in BC or discordant with direction of skewing in BC. Results are provided in the table below. As observed, direction of AS in both PMN and monocytes was random relative to the primary (BC) Xi ratio, suggestive evidence that AS does not represent completion of the primary Xi trait in BC tissue. Alternatively, in T and B lymphocytes, direction of AS was preferentially discordant with direction of skewing in BC, i.e., skewing towards a random Xi pattern was observed in B and particularly T cells.

Cell type	**n**	**Concordant in direction of skewing**	**Discordant in direction of skewing**	**Significance (chi-test)**
PMN	906	451	455	0,89
Monocytes	864	419	445	0,38
T lymphocytes	879	365	514	<0,0001
B lymphocytes	877	404	473	0,02

Chapter IV

RESULTS (PART II)

Since the AS and primary (body-wide) skewing traits were prevalent in a large proportion of our French-Canadian female population (reviewed in Part I), a comprehensive investigation was undertaken to identify clinical associations and etiologies. As explored in Section 13 of the introduction, Xi skewing has been associated with several pathophysiological conditions (expression of X-linked disease alleles, RSA, ovarian cancer, autoimmune disorders and plausibly longevity). Other than the unfortunate preferential inactivation of the wild type X in female carriers of X-linked mutant alleles as the cause of disease expression, whether Xi skewing is the cause or consequence of phenotype expression and variability of the above mentioned medical conditions and how it contributes to disease progression has not been clarified and well validated. Insight may be provided by identifying the molecular elements implicated in the etiology of Xi skewing and/or identifying the biomolecules / genes possibly altered as a consequence of Xi skewing. Alternatively, insight may also be gained through a prospective study where two groups of females: one demonstrating skewed Xi patterns (cases) and the other random Xi patterns (controls), are examined for clinical trait expression. We conjecture that if the clinical trait is a consequence of Xi skewing, the case group should demonstrate a different incidence versus controls. Alternatively, to determine whether skewing is the consequence of a medical condition a case-control prospective study where females demonstrating early clinical disease symptoms and in the absence of a skewed Xi pattern are recruited. Likewise, age-matched controls are also recruited. Analysis of Xi patterns as a function of aging between the two groups would provide insight. Regardless, as no prospective studies have been performed in this study, the majority of analyses in this project have been limited to association studies.

Particular medical disorders/biological variables selected for analysis include: blood counts, cancer (all-type), disorders with an autoimmune component, medicinal use and number and occurrence of spontaneous abortions. An interesting facet to these disorders is whether i) the finding can be replicated in our population/family based study and ii) whether the variable is associated with PS and/or AS.

A first step to investigate the etiologies of Xi skewing (be they genetic, environmental and/or stochastic) was to analyze familial resemblance of XIRs. This included both qualitative (familial aggregation of skewing) and quantitative analyses (correlation of XIRs within versus among families by ANOVA and maximum likelihood heritability estimates). The variance component approach of the SOLAR software allowed attribution of trait variance to both genetic and non-genetic factors. Non-genetic covariates taken into consideration were age and smoking habits.

4.1 Characteristics of the sample population

Although demographic and skewing characteristics of our cohort population of females were primarily described in Part I, further characteristics are provided below; in particular, hematopoietic, clinical, smoking and family/parity data.

4.1.1 Hematopoietic indices

Results of blood work and count differentials, as obtained for the majority of participants (97,5%), are provided in Table XV (page 150). Whether these means are particular to French Canadian females and/or to all females of this age group was not thoroughly investigated. See (Tsang et al., 1998) for review of hematopoietic indices as a function of age groups. The frequency distributions of blood counts are depicted in Figure 20 and as shown, the majority of lineages demonstrated a normal but slightly skewed distribution. The normal distribution of blood indices is likely due to variability induced by age, genetic and environmental influences. Whether the distribution departed from Gaussian distribution (possibly tested by using the Shapiro-Wilk statistic and/or Kolmorogov-Smirnov test for normality) was not investigated. The most extreme case was observed for basophils. Artifactual basophil count by the cell-counter has been speculated as the primary determinant.

Since outliers can have a significant and drastic effect on statistical analyses, they were identified and selectively removed for several of the statistical analyses. The criterion for removal was arbitrarily based on extreme values. In some instances however, exclusion was also validated by hematologic standards. The criterion for exclusion were: counts $\geq$ 14,50 for WBC (n=4), counts $\geq$ 11,00 for neutrophils (n=4), counts $\geq$ 4,32 for lymphocytes (n=1), counts $\geq$ 1,10 for monocytes (n=5), counts $\geq$ 0,60 for eosinophils (n=8), a concentration $\geq$ 160 for hemoglobin (n=1), and counts $\leq$ 45 and $\geq$ 500 for platelets (n=2). No significant outliers were identified for MCV. Table XVI (page 151) presents the number of cases removed from each particular hematopoietic lineage and Figure 21(page 148-149) presents distribution of blood counts without extreme outliers, thus depicting the distribution of a "healthier" population. As observed, distributions were less skewed. Although females demonstrating extreme hematopoietic values are likely to represent chronic and/or acute diseases, for instance increased leukocyte, WBC and eosinophil count has been associated with increased mortality (de Labry et al., 1990) (Weiss et al., 1995) (Hospers et al., 2000), on ethical grounds, no clinical follow-up was conducted.

Possible factors contributing to variability of hematopoietic values includes inter-individual variation (age, hormonal balance, genetic variation), intra-individual variation (diet, smoking habits, medicinal use, altitude, exercise and clinical conditions) and possibly laboratory / technical error.

4.1.2 Clinical characteristics

Clinical data was self-reported as no clinical files were investigated for confirmation. Medical conditions included asthma, RA, SLE, anemia and cancer history. Medicinal use included cancer-related therapy (radiotherapy and chemotherapy) and other medicinal types (n=9). A limitation to the cancer analyses was the clustering of all cancer types into one variable. Descriptive statistics for clinical data is provided in Table XVII (page 152).

4.1.2.1 Prevalence of clinical traits versus reference populations:

Rheumatoid arthritis: Since i) RA has a population prevalence of ~ 1% (occurs in all ethnic groups) (Lawrence et al., 1989), ii) is a common clinical problem in the elderly (although anyone can get RA, the disease usually begins in middle adult years), and iii) demonstrates a high female:male ratio distortion (approximately 3-to-1), the rate observed in this study (i.e., 6,1%) may be an appropriate estimate. Moreover, a proportion of false-positives (55%) is speculated, likely to reflect OA (Lynn et al., 1995).

Anemia: Although the diagnostic criterion for anemia in females is a hemoglobin concentration < 120 g/L (World Health Organization), anemia was self-reported in this study (plausibly determined with the aid of a medical doctor). Factors plausibly involved in the etiology of age-related anemia are inflammation, chronic disease, blood loss and nutritional deficiencies (protein-energy malnutrition: folic acid and vitamin B12 deficiency – megaloblastic anemia). However, because of increased iron stores with advancing age, iron-deficiency is an unlikely cause of anemia in the elderly. Epidemiologic analyses have reported i) an annual incidence of 1-2% in the general population, ii) a higher incidence in males versus females and iii) a significantly higher incidence in older individuals: 4-6 fold higher in Caucasians >65 yo (Lipschitz, 1999). Thus, taking these factors into consideration, the reported incidence in this study, i.e., 3,2%, is in agreement with epidemiologic findings.

Asthma: In Canada, 10-15% of the population suffers from asthma, with the incidence on the rise particularly in children. The lower incidence reported in our population therefore (7,8%), may reflect selection of an elderly female study population whom may be less susceptible to asthma.

The incidence of cancer (8,7%) and SLE (0,3%) was not compared with reference populations.

4.1.3 Smoking characteristics

Discrete and quantitative variables of smoking habits are provided in Table XVIII (page 153). Discrete variables included whether the subject was a non-smoker, current smoker or ex-smoker. Quantitative measures of smoke included pack-years (pack/day x number of years smoked), years-smoke and years-stop smoking. Pack-years and years-smoke are the cumulative of both current and ex-smokers. Whether smoking statistics for this population was consistent with reference populations was not investigated.

4.1.4 Family / parity data

Descriptive statistics for family / parity data are provided in Table XIX (page 154). Included were variables such as number of participating families (n=193), number and sex of siblings, number of participating mothers (n=37), number and sex of offspring, sex ratio of offspring and the occurrence / number of miscarriages.

4.2 Confounding variables

Prior to analyzing the associations between skewing and the various biological variables (blood counts, medical conditions, use of medicinal products, parity and smoking characteristics), it was vital to identify potential confounding variables. The latter were investigated by several methods that included linear or logistic regression analysis (LR) and multivariate linear regression (Generalized Linear Model – GLM). Moreover, as the data was correlated, i.e., data was derived from families, a more robust approach for regression analyses was to utilize the General Estimating Equation (GEE method). A further approach, which also took familial correlation into consideration, was the SOLAR statistical method, which tested various variables to determine whether they accounted for a significant proportion of trait variance.

Once confounding variables were identified, rather than selectively removing these individuals from statistical analyses, the confounding variable(s) was (were) included as a covariate(s).

4.2.1 Age as a confounding variable

Age-hematopoietic indices

As blood counts have been previously cited to vary with age (Kelly and Munan, 1977), (Jernigan et al., 1980), (Zauber and Zauber, 1987), we elaborated whether blood counts varied significantly in our population of elderly females (mean age of 63.3). Hematopoietic outliers as displayed in Table XVI (page 151) were excluded from analyses and results provided in Table XX (page 155). As shown, age clearly predicted the blood counts of various hematopoietic lineages: monocytes, eosinophils, hemoglobin concentration and MCV, as it was demonstrated by all three statistical methods. In particular, while monocyte and eosinophil counts increased with age, hemoglobin and MCV decreased. Whether the fluctuations are a normal consequence of the aging process or reflect an underlying abnormality has not been determined. A possible association may also exist between age and WBC, neutrophil and basophil count, as a significant p-value was demonstrated by the SOLAR method and a statistical trend by LR and GEE analyses. Lymphocyte and platelet count demonstrated a statistical trend by the LR and GEE methods but p-values were clearly insignificant by the SOLAR method. The small influence of age on lymphocyte counts is consistent with previous findings (Hall et al., 2000). Thus, since age predicted blood indices for various lineages, it was included as a confounding variable in the blood-skewing analyses when appropriate.

Age-clinical data

Since medicinal use and the prevalence of various medical conditions are age-related, for instance the prevalence of cancer increases significantly with each successive decade, the association between age and clinical data was investigated. As shown in Table XXI (page 156), age predicted (positive association) the occurrence of medical conditions such as all-type cancer ($p=0,037$ LR) and RA ($p=0,006$ LR, $p=0,0034$ SOLAR), consistent with published data. Lack of an association between age and anemia ($p=0,158$) was consistent with published data: prevalence of anemia in elderly females (>59 yo) was similar to that of women of childbearing age (Lipschitz, 1999). Lack of an association between age and asthma is consistent with a non-age-related form of asthma.

Among medicinal use, age demonstrated an inverse relationship (-) with HRT ($p<0,0001$) but a positive relationship (+) with use of med-heart ($p<0,0001$), med anti-inflam ($p<0,0001$) and med-other ($p<0,0001$). Thus age was included as a confounding variable in the skewing-clinical data analyses when appropriate.

Age-smoking habits

The relationship between age and smoking characteristics are provided in Table XXII (page 157). For categorical variables, while the number of current smokers significantly decreased with

advancing age ($p<0{,}0001$), the number of ex-smokers increased ($p=0{,}001$), suggesting that the decrease of current smokers is due to the increased number of ex-smokers. Observation of an inverse relationship (-) between age and the quantitative variables pack-years ($p<0{,}0001$ LR, $p=0{,}00011$ SOLAR) and years-smoke ($p=0{,}0003$ LR) indicates that older females have been exposed to less voluntary smoke (quantity and duration) versus younger females. Whether any biases exist in the data, such as older females being less honest (or having a failed memory) on the quantity and duration of cigarette smoking could not be determined. A positive association (+) between age and years-stop smoking ($p<0{,}0001$, LR) is consistent with the increased number of ex-smokers with advancing age. Thus, age was included as a confounding variable in the smoke-skewing analyses when appropriate.

Age-parity

The relationship between age and parity data (occurrence and number of miscarriages, offspring sex-ratio and number of offspring) was investigated. As shown in Table XXIII a) (page 158), age significantly (+) predicted two miscarriage variables: occurrence ($p=0{,}0001$ LR) and number of miscarriages ($p<0{,}0001$ LR, $p=5\times10^{-7}$ SOLAR). However, adjusting for number of offspring (multivariate analyses – Table b) XXIII, page 158), significance was lost. This finding can be most reasonably explained by the following: since elder females bore more children ($p<0{,}0001$) and assuming that a greater number of pregnancies (offspring) translate into a greater number of miscarriages, that older females had an increased prevalence of miscarriages was not surprising. Offspring sex-ratio was also significantly associated with age ($p=0{,}022$), however, excluding parity data of mothers of subjects selected for study enrollment (n=37), as they represent a selection bias (i.e., families with 4 or more female siblings were selected), the relationship was no longer significant ($p=0{,}197$). Number of offspring significantly increased with advancing age ($p<0{,}0001$), as expected, reflective of the lifestyle of more elderly females. Thus, age was included as a confounding variable in the skewing-parity analyses when appropriate.

4.2.2 Clinical data as a confounding variable

4.2.2.1 Clinical data-blood counts

In order to derive general and/or supportive information on clinical data, the association between clinical data and blood counts was investigated (Table XXIV, page 159-160), thus providing insight into the physiological basis of the clinical trait and possibly supportive evidence in favor of a genuine medical trait. For instance, an association between asthma and eosinophilia is well documented (Bousquet et al., 1990). However, since clinical data was self-reported, a proportion of false positives may confound the findings. In addition, as each medicinal type (n=9) comprises a

heterogeneous list of medications, associations cannot be directed to specific drugs. A similar corollary can be made for medical conditions. That is, each medicinal condition is likely composed of a heterogeneous group of diseases. For example, cancer is comprised of a heterogeneous group of cancers, RA confounded by OA, and anemia possibly confounded by a non-specific fatigue syndrome. Conditions that can mimic asthma include chronic obstructive pulmonary disease (COPD), congestive heart failure, coronary obstruction due to various causes, vocal cord dysfunction and esophageal spasm. Nonetheless, controlling for age when necessary, various associations were identified. Results provided in Table XXIV (page 159-160) are discussed below.

Asthma / med-asthma: Since both asthma and med-asthma demonstrated a positive (+) association with WBC, neutrophil and eosinophil count, evidence for a genuine asthma claim was speculated. However, upon multivariate analyses (controlling for asthma, med-asthma and age - Table XXIV b) (page 159-160), WBC and neutrophil count were no longer associated with asthma and med-asthma, suggesting a spurious finding. The association between asthma and eosinophil count however remained significant (p=0,023), in keeping with published data, consistent with a genuine asthma claim. However, the association between med-asthma and eosinophil count was lost, implying most of the variation was attributable to asthma.

Anemia: Although anemia was self-reported, that an inverse relationship (-) was observed between anemia and both hemoglobin concentration (p=0,0001) and MCV (p=0,002), is consistent with a hypochromic, microcytic form of anemia. Moreover, analysis of CBCs by the principal investigator (L. Busque, MD, FRCPC, hematologist) validated an authentic form of anemia for the majority of cases.

RA: Controlling for age, a positive association (+) was observed between RA and WBC (p=0,042) and neutrophil count (p=0,041). Whether this finding is consistent with the physiological basis of RA, a result of the medicinal products used to treat RA or due to confounding variable(s) was not investigated.

Medicinal products: Several associations were observed between medicinal products and blood counts. However, whether medicinal products had a direct effect on blood counts or was merely a reflection of a clinical condition, which of itself influenced blood counts, could not be determined. Pharmacologic indications and/or prospective studies would be required. Findings are as follows: i) an inverse relationship (-) was observed between chemotherapy intervention and lymphocyte count (p=0,046) and ii) between med-other and hemoglobin concentration (p=0,026). However, since the

findings were limited to a single hematopoietic lineage and the p-values were barely significant, a spurious finding is speculated. A highly significant/positive (+) association was observed between iii) med-anti-inflam use and eosinophil count (p=0,004) and iv) med-anti-convulsants and MCV (p=0,006). However, whether these findings are consistent with published data and/or due to confounding variables was not investigated. v) The use of HRT, med-heart and med-hemato were associated with the blood counts of several hematopoietic lineages, with highly significant p-values in certain instances, thus suggestive of an authentic finding. Whether the finding is consistent with published data and whether these medicinal products have a direct/primary effect on blood counts (therefore blood count the responsive variable) was not investigated. Nonetheless, medical conditions and medicinal use were included as a confounding variable in the blood counts – skewing analyses when appropriate.

4.2.2.2 *Clinical data – parity data analysis*

The relationship between clinical and family data was investigated. A limitation to the analyses however, as mean female age was 63,3, current medicinal use / medical condition may not be reflective of the clinical situation at the time of offspring delivery/miscarriage. Nonetheless, it may reveal consequences associated with family choices such as bearing a low versus high number of children.

As shown in Table XXV (page 161), controlling for age, a positive relationship (+) was found between use of med-heart and the number of offspring (p=0,004), the occurrence of a miscarriage (p=0,010) and a statistical trend with the number of miscarriages (p=0,055). However, when controlling for number of children, the association with miscarriage data was lost, implying only increased parturition as a plausible cause of med-heart use.

Both asthma [a statistical trend with occurrence of a miscarriage (p=0,060) and number of miscarriages (p=0,102)] and med-asthma (p=0,048 with number of miscarriages), demonstrated a positive (+) association with miscarriage data unconfounded by number of children, implying a plausible association between miscarriage and asthma, possibly consistent with an immune/autoimmune component common to both conditions.

RA was associated (+) with number of children (p=0,042), number of miscarriages (p=0,031) and a statistical trend with occurrence of a miscarriage (p=0,156). However, controlling for number of children, the association with miscarriage data was lost, implying that increased parturition may be the only family variable associated with increased occurrence of RA. Assuming that med-anti-

inflam is used by females with RA, the latter finding is compounded by a positive association between med-anti-inflam and number of children (p=0,017).

The relationship between use of HRT and offspring sex-ratio distortion (p=0,043) is likely to be spurious and/or biologically incomprehensible as females on HRT are typically menopausal, therefore beyond the child-bearing age.

4.2.3 Smoke as a confounding variable

To determine whether cigarette smoke was a significant confounding variable, the relationship of smoking habits with i) clinical data, ii) blood counts and iii) parity was investigated.

4.2.3.1 Smoke-clinical data

Since smoking has been i) associated with various medical conditions, for example, a potential risk-factor for RA (Hutchinson et al., 2001), MDS (Bjork et al., 2000) and atherosclerosis (Blann et al., 1998) and ii) demonstrates a suppressive effect on human immunity (Moszczynski et al., 2001), the association between smoking habits and clinical data was investigated. Results are given in Table XXVI (page 162). As discussed below, significance of the results is dependent on several assumptions and the medical condition in question. Age was included as a covariate when necessary.

HRT: A significant inverse relationship (-) was observed between current smoker status and the use of HRT (p=0,034). However, as the association was not detected with other variables of smoke, a spurious finding is speculated.

Med-asthma / asthma: Pack-years of cigarette smoke and years-smoked were both associated (+) with use of med-asthma (p<0,0001 and p=0,002 respectively). In addition, pack-years was also associated (+) with asthma (p=0,002 GLM, p=0,0087 SOLAR). In all instances, the relationship was positively correlated, consistent with published data citing exposure to cigarette smoke with increased severity of symptoms (Landau, 2001). Moreover, assuming smoke is a risk-factor for asthma, that ex-smoker and years-stop smoking were not associated with a decrease in med-asthma/asthma implies that onset of asthma is not reversed by smoke-cessation. However, this finding may be population-dependent and limited to our elderly population of females. Furthermore, that discrete variables of smoke (current and ex-smoker) were not associated with asthma suggests that smoking in of itself may not be sufficient to induce asthma. That is, the quantity (intensity and duration) of smoke may be a more important risk-factor. Alternatively, discrete variables of smoke

may be limited by smaller sample size (n=189 for current smoker versus n=1128 for pack-years), thus statistically less powerful versus quantitative variables.

Med-other: Use of med-other was significantly related (+) with pack-years (p=0,003) and number of years-smoked (p=0,0002). However, whether use of med-other is due to smoking or vice-versa is not known.

Med-heart: Controlling for age, use of med-heart was associated (+) with pack-years (p=0,018) and years-smoked (p=0,003), suggesting that large quantities of and/or many years of cigarette smoking may result in heart complications necessitating medical attention. Smoking has been previously associated with heart disease (Blann et al., 1998), consistent with a causal relationship. Curiously however, discrete smoking variables, current smoker in particular, was not associated with use of med-heart (p=0,891). This may be most reasonably explained by: i) smaller sample size (n=189 for current smokers versus n=1128 for pack-years), suggesting discrete variables of smoke are statistically less powerful versus quantitative variables; ii) current smoking, for reasons that some may be light smokers, is insufficient by itself to induce use of med-heart. Moreover, the results suggest that once use of med-heart has been initiated, the number of years-stop smoking is not statistically sufficient to reverse the use of med-heart (p=0,163).

Anemia: The findings suggest that current smoking may be a risk-factor (+) for anemia (p=0,065) while quitting smoke (ex-smoker) a protective mechanism (-) (p=0,020). That the findings are complementary adds an element of authenticity to the finding. That the association was not observed with the quantitative variables pack-years and years-smoke may most reasonably be explained by confounding variables. That is, since these variables are the cumulative of both current and ex-smokers, their individual effects may be plausibly neutralized. Therefore, reanalysis with the variables pack-years and years-smoke but dichotomized into current and ex-smoker is warranted. Alternatively, the finding could be spurious or confounded by covariates. Paradoxically, smoking is reported to increase hematocrit (Rampling, 1999) and hemoglobin concentration [(Kondo et al., 1993) and present study included – see Table XXVII, page 163], thus plausibly confounding the relationship between smoking and anemia.

RA: Results of the SOLAR statistical method, although those of the GLM were not significant, suggest a significant relationship between RA and current-smoker status (p=0,034) and pack-years of smoke (p=0,039). The former is consistent with data in the literature citing men who are current smokers had an increased risk of developing seropositive RA (Uhlig et al., 1999) while the latter is

consistent with the cumulative exposure of cigarette smoke as a risk-factor for RA (Hutchinson et al., 2001).

4.2.3.2 Smoke-blood counts

Since cigarette smoke influences blood counts ((Corre et al., 1971), (Grimm, Jr. et al., 1985), (Schwartz and Weiss, 1994), (Tsang et al., 1998)) and associated with B-cell lymphocytosis - primarily in women (Delage et al., 2001), the relationship between smoking habits and hematopoietic indices in our population of females was investigated. As presented in Table XXVII (page 163) results of LR, GLM, GEE and most of the SOLAR statistical analyses clearly indicate that current smoker status, pack-years and years-smoke predicted (+) the counts of WBC, neutrophil, lymphocyte, monocyte, basophil, MCV and hemoglobin. Evidence of a relationship for eosinophil count and platelet count with smoking habits was inconsistent, dependent on the statistical method used and smoking variables analyzed. Since platelet count was significantly associated with quantitative variables of smoking habits (pack-years and years-smoked) however implies an authentic finding. Although we failed to demonstrate a consistent association between smoking habits and eosinophil counts, others have not (Mensinga et al., 1994), (Schwartz and Weiss, 1994). Discrepancies may lie in sample size. That a negative association (-) was detected between blood counts and ex-smoker status and years-stop smoking suggests that after removal of the stimuli (i.e., smoke), hematopoietic indices return to original levels, suggesting that smoke is the causal factor. These findings suggest that smoking habits should be included as a confounding variable in the blood-skewing analyses.

4.2.3.3 Smoke-parity

Although the association between smoking habits and parity data is not clear, for instance i) are current smoking habits reflective of past smoking habits when parturition occurred and ii) are smoking habits the cause or consequence of parturition data. Analyses were nonetheless conducted and parity data set as the dependent variable. As shown in Table XXVIII (page 164), other than a negative association between number of children with pack-years and years-smoked, no other significant associations were found, suggesting cigarette smoking is not a significant predictor of parity data or vice-versa.

4.2.4 Parity data as a confounding variable

Parity and blood counts

To determine whether parity data was predictive of blood counts, the association was investigated. As presented in Table XXIX (page 165), other than some marginal associations, no strong or consistent association was observed.

4.2.5 Covariates associated with blood counts: multivariate analyses

As several covariates were associated with hematopoietic indices, multivariate analyses were conducted – results provided in Table XXX (page 166). For WBC, significant covariates were med-heart (p=0,017), ex-smoker (p=0,008) and pack-years of cigarette smoke (p=0,004). Neutrophil count was associated with asthma (p=0,009) and pack-years (p=0,009). The only variable to significantly influence lymphocyte count was ex-smoker (p=0,001). Age was the only variable to predict monocyte count (p<0,0001). Eosinophil count was associated with HRT (p=0,041), med-anti-inflam (p=0,018) and asthma (p=0,043). Age (p=0,024) and current smoker (p=0,001) predicted basophil count. Platelet count was associated with HRT (p=0,035). Hemoglobin was associated with med-asthma (p<0,0001), med-other (p=0,005), anemia (p<0,0001) and pack-years (p=0,043). MCV was predicted by age (p=0,025) and med-anti-conv (p=0,015). Thus for future statistical analyses pertaining to blood counts, the aforementioned variables should be considered.

4.2.6 Maximum likelihood heritability estimates of confounding variables

The phenotypic variance of confounding variables (blood counts, number of miscarriages, smoking habits) attributable to genetic factors was investigated by maximum likelihood estimates (adjusted for age, pack-years of cigarette smoke, current smoker and ever smoker). Although not investigated in this study, familial aggregation and heritability of RA (Deighton and Walker, 1991), asthma (Palmer et al., 2000) and COPD (Khoury et al., 1985) has long been recognized. In fact, utilizing a classic twin study design, 73% of the variation in susceptibility to asthma and 60% of the risk of developing RA can be explained by genetic factors (Skadhauge et al., 1999) (MacGregor et al., 2000).

Blood counts

As presented in Table XXXI (page 167), the variance of all hematopoietic lineages demonstrated significant heritability estimates ($0,14 \leq h^2 \leq 0,66$; $p \leq 0,045$), implying a genetic component to derivation of blood counts. In particular, MCV demonstrated the strongest familial resemblance ($h^2=0,66$; $p<1x10^{-7}$), followed by platelet count ($h^2=0,54$; $p<1x10^{-7}$). Monocyte and lymphocyte counts also demonstrated modest h^2 values (0,44 and 0,43, respectively, $p<1x10^{-7}$). Lack of strong

heritability of basophil count may be attributed to a low mean count (0,033) and a high standard deviation (± 0,055) – see Table XV (page 150).

Parity data

As shown in Table XXXI (page 167), a heritable / genetic component (h^2=0,12; p=0,002) was demonstrated for multiple miscarriage occurrence. Genetic components possibly include genetic aberrations and/or miscarriage susceptibility allele(s).

Smoking habits

As shown in Table XXXI (page 167), phenotypic variance attributable to genetic effects for the quantity of smoke inhaled (pack-years) was significant (h^2=0,18; p=0,00002), suggesting that quantitative variables of smoking habits have a genetic component, consistent with a genetic component to smoking addiction (Munafo et al., 2001).

4.3 Familial aggregation of skewed Xi patterns

To determine whether genetic and/or environmental factor(s) influence derivation of XIRs, familial aggregation of skewed Xi patterns was examined. As such, the recurrence risk (RR) and recurrence risk ratio (RRR) were determined for each skewed Xi phenotype. Confidence intervals were not determined.

4.3.1 Primary skewing trait (DS score ≥ 0,25)

As shown in Table XXXII (page 168), familial aggregation of skewed Xi pattern in BC, PMN and T lymphocytes was not significant (RRR = 1,01; 1,12 and 1,19; respectively), suggesting that genetic and/or shared environmental factors do not contribute to derivation of XIRs for these tissues / lineages. These findings imply a stochastic and/or non-genetic mechanism(s) of skewed Xi pattern derivation.

Nonetheless, and assuming a genetic component to skewed Xi patterns in BC, lack of familial aggregation may be attributed to: i) Small sample size of cases (n=20) versus controls (n=178), thus necessitating a larger case sample size. ii) A too stringent criterion of skewing. For example, in mice, as in the case of extreme heterozygotes (Xce^a /Xce^c heterozygotes), Xce X chromosome skewing achieves a maximum skewing of 20-30%. A lower degree is expected in weaker heterozygotes (Xce^a /Xce^b for example). Therefore, and assuming a similar mechanism of skewing in humans, reanalysis of skewed Xi patterns BC but with a less stringent criterion, a DS score ≥0,18

for example, is warranted. Pertaining to T lymphocytes, and assuming a genetic component to skewed Xi patterns in T lymphocyte, lack of familial clustering may be explained by i) small sample size of cases (n=39) versus controls (n=158), thus requiring a larger sample size. ii) As in the case of BC skewing, a DS score ≥ 0,25 criterion may be too stringent, thus necessitating reanalysis but with a less stringent criterion (e.g. DS score ≥ 0,18).

In monocytes and B lymphocytes on the other hand, evidence for significant familial clustering of skewed Xi patterns was observed (RRR = 1,28; p<0,01 and 1,62; p<0,001 respectively), evidence in favor of a genetic component and/or shared environmental influence(s). To decipher whether the genetic component was plausibly X-linked, an estimate of heritability and correlation in direction of skewed Xi patterns (i.e., Xp versus Xm) among siblings was analyzed (Chapter 4.5.2). The effect of smoke as a possible environmental stimulus was also investigated (Chapter 4.5.1).

Surprisingly, as both monocytes and PMN are of myeloid origin, and strongly correlated for XIRs (Figure 17, page 99), thus speculated to exhibit a similar Xi phenotype, lack of familial aggregation for skewed Xi patterns in PMN is unclear. One possibility may be that analyses lacked sufficient power, thus requiring a larger sample size. Alternatively, the finding may be genuine, implying a strong stochastic etiology to derivation of XIRs in PMN.

4.3.2 AS trait (AS_{DS} score ≥ 0,25)

As presented in Table XXXII (page 168), the AS trait (AS_{DS} score ≥ 0,25) was significantly aggregated within families as demonstrated for all four hematopoietic lineages ($1,57 \leq RRR \leq 2,63$; p<0,01), suggesting that genetic and/or shared environmental factors play a role in derivation of the AS trait. Among the various lineages, aggregation was strongest for T lymphocytes (RRR=2,63), suggesting a strong heritable component. Heritability estimates may provide further insight for the role of genetic factors in derivation of the AS trait (Chapter 4.5.2). To determine whether X-linked genetic factor(s) are plausibly implicated in derivation of the AS trait, correlation in direction of AS (AS_{PAmat} score) among siblings was also analyzed (Chapter 4.5.2.2). The role of cigarette smoke as a possible environmental stimulus in derivation of the AS trait will be reviewed in Chapter 4.5.1.2.

4.4 Physiological relevance of skewed Xi patterns: association with biological variables

The significance of the findings herein stated are dependent on several factors: i) statistical strength (p-value), ii) whether the finding can be demonstrated in more than a single hematopoietic lineage,

iii) the biological significance of the finding and iv) temporality. Moreover, as multiple testing was performed, thus increasing the probability of incurring a type I error, one can raise the argument that results should be interpreted with caution. Therefore, a straight Bonferroni correction for multiple testing should be included. Alternatively a more stringent alpha level of 0,01 may be used.

Both DS and AS_{DS} scores were utilized for the association studies. As introduced in Part I, the BC DS score putatively quantitates the primary XIR. Further, in blood, the DS score putatively quantitates the primary XIR confounded by the AS trait. Thus the AS_{DS} was employed to quantitate the AS trait more effectively. PAmat and AS_{PAmat} scores on the other hand, as they provide a measure of the direction of skewing (Xp versus Xm), were excluded from the analyses. That is, we were interested in the relationship between 'magnitude' of Xi skewing and biological variables and not between 'direction' of Xi skewing and biological variables. The latter however may prove fruitful for analyses wishing to investigate the relationship between parental source of Xi skewing and biological end-points, potentially identifying imprinted X-linked traits.

4.4.1 Hematopoietic indices and Xi skewing

To determine the physiological significance of AS, as modulation of HSC kinetics is speculated, blood counts as a function of XIRs was investigated. Association analyses included LR, GLM and GEE statistical methods. Confounding variables are stated in parentheses of Table XXXIII (page 169-170).

4.4.1.1 DS score

The relationship between DS score and blood counts are provided in Table XXXIII a) (page 169-170).

<u>BC DS score</u>: As shown, other than a marginal association between the BC DS score and MCV ($p=0,049$), the BC DS score did not predict hematopoietic counts for any of the remaining hematopoietic lineages, suggesting that the primary Xi pattern is not associated with variations in the hematopoietic department.

<u>Blood DS score and eosinophil count</u>: With or without adjusting for covariates, a negative relationship (-) was observed between DS scores of all four hematopoietic lineages and eosinophil count (results of LR, GLM and GEE all significant). This result suggests that with increasing DS scores, there are less eosinophils in PB. Lack of an association with the BC DS score limits the

relationship to the hematopoietic department. An association with the AS trait is thus speculated (see below).

Blood DS score and platelet count: Platelet count was negatively associated (-) with blood DS scores, as observed with the DS scores of monocytes, B lymphocytes and a statistical trend with T lymphocytes. Lack of an association with PMN, a cell type well correlated with the PAmat score of monocytes, is presently unclear. Nonetheless, and although the mechanism is unknown, increasing blood DS scores may be associated with vascular alterations, such as reduced efficiency to control bleeding and/or thrombocytopenia.

Blood DS score and hemoglobin: With or without controlling for confounding variables, hemoglobin concentration was significantly associated (+) with PMN and monocyte DS scores (LR, GLM and GEE methods). As these cell types demonstrated the highest incidence of AS, an association with the AS trait was speculated. No relationship with AS_{DS} scores was found [Table b) XXXIII; page 169-170], limiting the association to DS values. Further, since hemoglobin concentration is negatively related to anemia, we queried an association between XIRs and anemia. As seen in Table XXXIV (page 171-172), although myeloid DS scores were not associated with anemia, the AS_{DS} scores of monocytes and B lymphocytes were. However, this association was lost to confounding variables (smoke, ex-smoker and hemoglobin – GLM method). Thus increasing hemoglobin concentration as a function of increasing DS scores appears unrelated to self-reported anemia. Alternatively, the positive association between DS scores and hemoglobin may be associated with intravascular hemolysis. Analysis of DS scores as a function of RBC numbers should derive further insight.

Blood DS score and MCV: A marginal negative (-) association was observed between MCV and T lymphocyte DS scores. However, a spurious finding is speculated as i) direction of relationship (-) was not consistent with the finding in BC which was positive (+), ii) p-values were marginal (p=0,049 for BC and p=0,012 for T cells) and iii) DS scores for the other three hematopoietic lineages were not significant.

4.4.1.2 AS_{DS} score

The relationship between AS and blood counts are provided in Table b) XXXIII (page 169-170).

AS_{DS} score and eosinophil count: Consistent with blood DS score results, eosinophil count was negatively (-) associated with AS_{DS} scores, as detected with various lineages (results of GEE

method: p=0,035 with PMN; p=0,013 with T lymphocytes, and p=0,011 with B lymphocytes). Controlling for additional covariates slightly reduced the significance of p-values (GLM method). Monocyte AS_{DS} scores however, for reasons which are presently unclear, were not associated with eosinophil counts (p=0,357 GEE method). Nonetheless, that both DS and AS_{DS} scores were associated with decreasing eosinophil counts implies an authentic finding. As such, an X-linked genetic factor to eosinophil counts is speculated.

AS_{DS} score and platelet count: A significant/positive association (+) was observed between platelet count and the AS_{DS} score of PMN (p=0,043 GEE method), suggesting that AS may be associated with mild thrombocytosis. However, since the finding was barely significant (p=0,043) and was restricted to a single hematopoietic lineage (PMN), a spurious finding is speculated. Moreover, this finding is in contrast to DS score results in which the relationship between DS scores and platelet count was inversely related and was detected in more than a single hematopoietic lineage.

4.4.2 Clinical data and Xi patterns

To determine whether varying Xi patterns were associated with clinical data, self-reported medical conditions and a partial list of medicinal use were analyzed as a function of DS and AS_{DS} scores.

4.4.2.1 DS score

DS score and asthma: As shown in Table a) XXXIV (page 171-172), a positive relationship (+) was observed between increasing DS scores and asthma. As the association was observed with BC tissue (p=0,032; LR) and hematopoietic lineages: PMN (p=0,032; LR) and possibly monocytes (p=0,054; GEE), an association with the primary skewing trait is speculated. Lack of T and B cell contribution may be attributed to the lower incidence of skewing for these lineages. Controlling for confounding variables (GLM and GEE methods) did not drastically alter p-values. A positive association between med-asthma and the DS scores of PMN (p=0,013), monocytes (p=0,081) and B lymphocytes (p=0,041) further supports the relationship. Lack of a relationship between asthma / med-asthma and AS_{DS} scores [Table b) XXXIV (page 171-172)] delimits the association to the primary skewing (DS score) trait. Since DS scores were not associated with med-allergy, grounds for an association of the primary Xi pattern with allergic diseases in general could not be supported.

DS score and RA: With or without controlling for confounding variables, an inverse relationship (-) was observed between myeloid DS scores and RA. As the observation was limited to the hematopoietic department (PMN and monocytes), we conjecture an association with the AS trait. In

fact, as seen in Table b) XXXIV (page 171-172), increasing AS_{DS} scores for PMN and monocytes were associated with a decreased occurrence of RA (p=0,045 for PMN, GLM method and a statistical trend for monocytes AS_{DS} scores for all three statistical methods). Since med-anti-inflam was not associated with DS and AS_{DS} scores, a general association between DS scores and inflammatory conditions could be concluded.

DS scores and HRT: A significant association (-) was observed between the use of HRT and the DS score of PMN (LR result: p=0,023). However, loss of significance upon multivariate analyses and lack of association with monocytes suggests a spurious finding.

No significant association was observed between skewing (DS score) and cancer, cancer related therapy (chemotherapy, radiotherapy), anemia, SLE, med-allergy, med-anti-inflam and med-heart.

4.4.2.2 AS_{DS} score

AS_{DS} score and anemia: As presented in Table b) XXXIV (page 171-172), although a negative association (-) was detected between AS_{DS} scores of PMN, B lymphocytes and T lymphocytes (statistical trend) and anemia as observed with the LR and GEE statistical methods, the association was lost upon controlling for confounding variables (GLM method), thus attributing trait variance to the latter variables. Interestingly, this finding is consistent with a lack of an association between AS_{DS} scores and hemoglobin concentration and MCV [Table b) XXXIII (page 169-170)].

AS_{DS} score and RA: As presented in Table XXXIV b) (page 171-172), AS_{DS} scores of PMN (p=0,045 for GLM and p=0,064 for GEE) and possibly monocytes (p=0,093 for GLM and p=0,057 for GEE), were associated with a decreased (-) reporting of RA. Similar findings with DS scores suggest that increasing DS and AS_{DS} scores in myeloid lineages plausibly attributable to the AS trait, may be associated with protection from RA. Consistent with an intrinsic X-linked growth advantage model of AS, one possibility may be that genetic variants of the AS gene, in addition to providing the cell with a relative growth advantage, confer reduced susceptibility to RA. Evidence of an X-linked genetic component to RA (Jawaheer et al., 2001) supports the latter.

No significant association was detected between AS_{DS} scores and cancer, chemotherapy, radiotherapy, asthma, SLE and use of HRT, med-allergy, med anti-inflam and med-asthma.

4.4.3 Parity data and Xi patterns

To determine whether the primary skewing trait (DS score) and/or the AS trait (AS_{DS} score) were associated with X-linked disease allele(s), the association between parity data and XIRs was investigated. Our rationale is that X-linked disease allele(s) will induce both a skewed X inactivation pattern in female carriers of the disease allele and the miscarriage of a male conception (as a female can transmit either of her X chromosomes, she is expected to transmit the mutant X half the time. Moreover, as a male is hemizygous for the X, the disease allele is likely to be fully penetrant, possibly inducing the spontaneous abortion of the developing fetus). If X-linked mutant alleles are indeed common in the Québec female population, we speculate that skewed Xi patterns are associated with increased miscarriages and an increased female:male offspring ratio. Results provided in Table XXXV (page 173).

4.4.3.1 DS score

No significant associations were identified between parity data and BC DS scores, implying that primary skewed Xi patterns are unlikely due to the presence of X-linked mutant allele(s). Among hematopoietic lineages, a negative association was observed between PMN DS score and miscarriage occurrence ($p=0,048$) and number of offspring ($p=0,047$), implying that a skewed Xi pattern in blood may be associated with reduced fertility. However, i) controlling for confounding variables (GLM), the association with miscarriage occurrence was lost ($p=0,132$), ii) lack of an association with monocytes, and iii) marginally significant p-values, all imply a spurious finding. Moreover, lack of an association with offspring sex-ratio ($p\geq0,859$) and number of miscarriage ($p\geq0,330$) strongly argue against a general role for X-linked mutant allele(s) in the etiology of skewed Xi patterns. This finding however is not consistent with data in the literature citing a 'high' frequency of skewed X inactivation among females experiencing RSA. Discrepancies are likely due to i) clinical criterion, as we did not collect information on the recurrence frequency of miscarriages. Unlike previous studies that defined RSA as 2-to-3 (or more) consecutive losses under 20 weeks gestation, the criterion utilized in this study was unrelated to gestation period and irrelevant to consequtiveness of the miscarriages. ii) Since DS scores were obtained in an elderly population of females (mean age 63,3) and not derived strictly from those whom have recently underwent parturition, temporal differences in DS scores are speculated. iii) Statistical power may be limited by smaller sample size of females experiencing a miscarriage ($n=352$) versus controls ($n=792$).

4.4.3.2 AS_{DS} score

Similar to the DS score results, a significant inverse relationship (-) was observed between miscarriage occurrence and increasing AS_{DS} scores, as observed for PMN, monocytes and B lymphocytes, particularly when controlling for confounding variables (GLM). Since the current AS_{DS} score may not be reflective of that when parturition occurred (as it occurs at a later age), the finding implies that women who have experienced a miscarriage are less likely to develop AS. Nonetheless, unlike DS score results, several features implicate an authentic finding: i) Like the DS score analyses, PMN were significantly implicated. ii) PMN, monocytes and B lymphocytes (the three hematopoietic lineages most implicated in AS) were consistent in direction of relationship, i.e., inverse (-) relationship. iii) Like the DS score analyses, AS_{DS} analyses also produced a negative relationship. Nonetheless, these findings argue against X-linked mutant alleles as a general cause of AS since X-linked disease alleles would have been associated with i) an 'increased' miscarriage occurrence and ii) other variables of parity (offspring sex-ratio, number of miscarriages and number of children).

4.5 Etiologies of skewed Xi patterns

As skewed Xi patterns for various hematopoietic lineages demonstrated significant familial aggregation (Chapter 4.3), the role of environmental and genetic etiologies were investigated. The role of age in trait evolution has been previously scrutinized (Part I of Results).

4.5.1 Environmental etiology: cigarette smoke

For reasons declared in Chapter 2.2.2, cigarette smoke was investigated as a possible environmental stimulus in the etiology of skewed Xi patterns. Smoking characteristics of participant females are provided in Table XVIII (page 153). XIRs were set as the dependent variable and significant predictors of XIRs were included as covariates.

4.5.1.1 DS score

As shown in Table a) XXXVI (page 174-175), the BC DS score was not associated with smoking variables, an expected finding since the primary XIR is putatively acquired during embryogenesis and relatively stable thereafter (Results, Part I). Among hematopoietic lineages and controlling for confounding variables, other then a marginal association between B cell DS scores and current smoker (p=0,04) and a statistical trend (p=0,07) with pack-years (SOLAR method), no other significant associations were found, suggesting that cigarette smoke is not a significant factor in derivation of blood DS scores. The association identified with B lymphocytes poses an interesting dilemma since i) p-values derived with the other statistical methods (LR, GLM and GEE) were

clearly insignificant and ii) since B lymphocyte DS scores did not significantly increase with advancing age, it is not clear how environmental factors are significantly implicated in derivation of B cell XIRs, thus implying a spurious finding.

4.5.1.2 AS_{DS} scores

As shown in Table b) XXXVI (page 174-175), controlling for confounding variables (GLM method), current smoker status [statistical trend for both PMN (p=0,067) and monocytes (p=0,074)] and pack-years of cigarette smoke (p=0,012 for PMN and p=0,069 for monocytes) predicted AS_{DS} scores, suggesting that cigarette smoke may be involved in the etiology of the AS trait. Lack of an association with T and B lymphocytes may be attributed to the low incidence of AS for these lineages. We hypothesize that a chemical particulate(s) of cigarette smoke, the smoking habit itself, or a covariate associated with cigarette smoke be associated with derivation of the AS trait.

4.5.2 Familial/genetic etiology: sibling correlation of XIRs and heritability coefficients

To derive evidence of a plausible X-linked genetic component to derivation of XIRs, sibling correlation of XIRs (ANOVA) and maximum likelihood heritability estimates (adjusted for age, pack-years of cigarette smoke, current smoker and ever-smoker) were analyzed. The latter derived estimates of the genetic effect (h^2). We speculate that if Xi patterns are influenced by a familial component, the magnitude of skewing (DS and AS_{DS} scores) should be more similar within versus among families. Moreover, if skewing is influenced by an X-linked genetic locus, direction of skewing (that is preferential inactivation of a parental-specific X chromosome) should be more similar within versus among families. The latter analysis was performed by evaluating PAmat and AS_{PAmat} scores.

4.5.2.1 Xi Skewing trait: deviation from random Xi

Primary / BC XIR: As seen in Table a) XXXVII (page 176), although the magnitude of BC skewing (DS score) was not significantly related among siblings (p=0,24 for ANOVA, h^2=0,04), thus consistent with the lack of familial aggregation of skewed Xi patterns in BC (Table XXXII, page 168), PAmat scores were significantly related among siblings ($p=3.3x10^{-9}$ for ANOVA, h^2=0,30). The latter is consistent with an X-linked genetic component to derivation of BC PAmat scores, plausibly consistent with an XCE-like etiology. Autosomal or stochastic etiologies are improbable as they would favor a normal or random distribution of PAmat scores among siblings.

Of special note, results pertaining to DS and AS_{DS} scores should be interpreted with caution as they violate an assumption of ANOVA and SOLAR statistical methods: these variables are not normally distributed as required (see Figures 10, 11; pages 92, 93).

Blood XIR: Among the various hematopoietic lineages [Table a) XXXVII, page 176], both DS scores ($p \leq 2,9x10^{-4}$ for ANOVA, $0,11 \leq h^2 \leq 0,24$) and PAmat scores ($p \leq 2,7x10^{-5}$ for ANOVA, $0,20 \leq h^2 \leq 0,39$) were significantly related among siblings, evidence supporting a genetic, possibly X-linked, component to derivation of blood XIRs. Similar to results in Table XXXII (page 168), familial resemblance of DS scores was strongest for monocytes and B cells ($h^2=0,21$ and $h^2=0,24$, respectively) and weakest for PMN ($h^2=0,11$). Interestingly, T cell PAmat scores demonstrated the strongest familial resemblance ($p=1,6x10^{-12}$ for ANOVA, $h^2=0,39$), implying a strong genetic (X-linked) component.

4.5.2.2 AS trait: Deviation from the BC XIR

Among the various hematopoietic lineages [Table b) XXXVII, page 176], with the exception of the PMN AS_{DS} score, both the AS_{DS} score ($p \leq 5,2x10^{-3}$ for ANOVA, $0,17 \leq h^2 \leq 0,30$), and AS_{PAmat} score ($p \leq 3,1x10^{-5}$ for ANOVA, $0,20 \leq h^2 \leq 0,36$), were significantly related among siblings, compelling evidence of a genetic basis, possibly X-linked, component to derivation of AS scores. Lack of familial correlation for PMN AS_{DS} scores utilizing the SOLAR statistical method is presently unclear as both ANOVA analyses ($p=1,5x10^{-4}$) and familial aggregation results [RRR=1,71; $p<0,005$; Table XXXII (page 168)] demonstrated significant familial resemblance of PMN AS_{DS} scores. Similar to PAmat results, an extremely significant p-value was found for T lymphocyte AS_{PAmat} scores ($p=3,1x10^{-11}$ for ANOVA; $h^2=0,36$), implying a strong X-linked genetic component to T lymphocyte AS.

4.5.3 Segregation analysis of skewed Xi patterns (preliminary findings)

To derive further evidence for an X-linked genetic component to derivation of skewed Xi patterns, segregation analysis of skewed Xi patterns within families was analyzed. We hypothesize that if the trait is X-linked, the X of same parental origin (i.e., direction of skewed Xi) should be preferentially inactivated among siblings more often then expected by chance alone.

4.5.3.1 Segregation analysis of skewed Xi patterns (DS score $\geq 0,25$)

BC tissue: Among families demonstrating a skewed Xi pattern in BC (DS score$\geq$0,25) and informative for parental derivation of X (i.e., Xp versus Xm) (n=91 families), n=27 carried at least two siblings demonstrative of a skewed Xi pattern in BC, thus informative for segregation analysis. Results are presented in Table a) XXXVIII, (page 177-183). As found, when direction of skewed Xi patterns (i.e., preferential inactivation of a parental-specific X; i.e., Xp versus Xm) was tabulated among these families, the X of same parental derivation (i.e., same) tended to be preferentially

inactivated among siblings: 54 (same parental X) versus 8 (alternate parental Xs). These values significantly differ from those expected if skewed Xi patterns were determined exclusively by a stochastic process: 34 (same parental X) versus 28 (alternate parental Xs). This finding is consistent with an X-linked genetic component ($p<0,0001$) to derivation of skewed Xi patterns in BC tissue, plausibly consistent with an XCE-like etiology. Lack of further concordance is congruent with a genetically complex trait.

Interestingly, when parental derivation of the skewed (preferentially inactive) X chromosome was tabulated among all families carrying at least one sibling with a skewed Xi pattern in BC, Xm was found to be preferentially inactivated versus Xp ($n=74$ versus $n=52$, $p=0,050$). Various explanations are provided in the discussion.

PMN: Among families demonstrating a skewed Xi pattern in PMN and informative for parental derivation of X (Xp versus Xm) ($n=144$ families), $n=103$ carried at least two siblings whom demonstrated a skewed Xi pattern in PMN, thus informative for segregation analysis. As presented in Table b), XXXVIII (page 177-183), when direction of Xi skewing (preferential inactivation of a parental-specific X) was tabulated among these families, it appeared that the X of same parental origin was preferentially selected among siblings: 235 (same parental X) versus 61 (alternate parental Xs). These values were significantly higher than expected if skewed Xi patterns were dictated solely by stochastic processes: 163 (same parental X) versus 133 (alternate parental Xs). This finding is consistent with an X-linked genetic component ($p<0,0001$) to derivation of skewed Xi patterns in PMN. Lack of further concordance in direction of skewed Xi patterns among siblings is congruent with a genetically complex trait.

Similar to results observed in BC tissue, when all families containing at least one sibling with a skewed Xi pattern in PMN were tabulated, Xm ($n=197$) tended to be preferentially inactivated versus Xp ($n=141$) ($p=0,002$), consistent with a parent-of-origin effect for skewed Xi patterns in PMN cells.

4.5.3.2 Segregation analysis of the AS trait (AS_{DS} score $\geq 0,25$)

Among families demonstrating an AS phenotype in PMN (AS_{DS} score$\geq0,25$) and informative for parental derivation of X (Xp versus Xm) ($n=110$ families), $n=58$ contained two or more siblings demonstrating an AS phenotype in PMN, thus informative for segregation analysis. As presented in Table c) XXXVIII (page 177-183), when direction of the AS trait (i.e., preferential inactivation of a parental-specific X) was tabulated among these families, it was found that a parental-specific X was

preferentially inactivated among siblings demonstrating an AS phenotype in PMN: 119 (same parental X) versus 29 (alternate parental Xs). This value was significantly different from those expected if the AS trait was exclusively dictated by a stochastic process: 83 (same parental X) versus 65 (alternate parental Xs). This finding is consistent with an X-linked genetic component ($p<0,0001$) to derivation of the AS trait in PMN. Lack of further concordance is congruent with a genetically complex trait.

Unlike DS score results, when parental derivation of the preferentially inactivated X chromosome was tabulated among all families carrying at least one sibling with an AS phenotype in PMN, Xm and Xp were equally inactivated (99 versus 101, respectively), limiting the parent-of-origin effect to DS score results.

DISCUSSION FOR PART II

Descriptive statistics and confounding variables

As presented in Tables (XV-XIX, pages 150-154) and Figures (20-21, pages 146-149), frequency distributions, descriptive statistics and reference values for various biological variables were determined for our elderly population of French-Canadian females. This included hematopoietic indices, clinical characteristics, smoking habits and family/parity data. The distribution of blood counts was analyzed and extreme outliers were identified and removed for some statistical analyses.

Prior to analyzing the relationship between Xi patterns and these biological variables, the influence of potential confounding variables was investigated. As found, age was a significant predictor of particular blood counts (Table XX, page 155), medical conditions/medicinal use (Table XXI, page 156), smoking habits (Table XXII, page 157) and parity data (Table XXIII, page 158). [The influence of age on XIRs was previously analyzed - Part I]. Various associations were identified between clinical data and hematopoietic indices (Table XXIV, page 159-160), thus providing i) physiological background to the clinical trait and ii) data from which to support or refute a self-reported clinical condition (that is, to determine whether variations in blood counts were consistent with published data of the medical trait). In addition, the relationship between medicinal use and blood counts may be used as an indirect method to infer a medical condition. For example, the association between use of med-asthma and blood counts was consistent with the association between asthma and blood counts, thus use of med-asthma was an indirect method to infer a clinical trait (i.e., asthma). However, in some instances, it was not known whether medicinal use had a

direct effect on blood counts, for example the effect of HRT on blood counts. Smoking habits were linked to various clinical conditions (Table XXVI, page 162) (asthma, anemia, RA, use of med-heart and med-other) and associated with variance of blood counts (Table XXVII, page 163), a finding consistent with published data, implying a causal role. Multivariate analyses of blood counts identified covariates significantly associated with variance of blood counts (Table XXX, page 166). Familial resemblance (heritability estimates) identified a significant genetic component to blood counts ($0{,}14 \leq h^2 \leq 0{,}66$), number of miscarriages ($h^2=0{,}12$) and smoking habits ($h^2=0{,}18$), implying a genetic contribution to trait variance (Table XXXI, page 167).

Genetic component to variability of hematopoietic cell numbers

A heritable component to variance of blood counts is consistent with classical twin model approaches, which attributed 61 to 96% of variance in blood cell measures to genetic factors (Evans et al., 1999). In a further study comprised entirely of females, 50% of absolute T lymphocyte count was attributable to genetic/heritable factors (Hall et al., 2000). Since the age of females participating in the former study (12 yo) was younger versus our study (38-96 yo) and the latter (18-80 yo) suggests that genetic effects, particularly so for T lymphocytes counts, decreases with advancing age. That lymphocyte count was predicted by current smoking habits and inversely related to ex-smoker status suggests that non-genetic factors (environment and possibly clonal expansion) play an increasing role in the decrease of lymphocyte levels with advancing age [(statistical trend – Table XX, page 155) and as previously reported (Hall et al., 2000)]. Moreover, that the magnitude of genetic effects may differ between the sexes (Evans et al., 1999), sex-limited attributes, such as genetic linkage to the sex chromosomes, may be a further source of variation. Nonetheless, taking into consideration the various non-genetic factors as covariates, genetic linkage studies and/or genetic association studies are merited to identify genetic determinants implicated in homeostatic regulation of blood counts.

Primary Xi pattern: BC XIR

Evidence of an X-linked genetic component: familial aggregation / correlation analyses

Although skewed Xi patterns in BC tissue were not aggregated within families (RRR=1,01), (possibly owing to small sample size and/or DS score ≥ 0,25 too stringent of a criterion to define a skewed Xi pattern), nor were BC DS scores significantly correlated among siblings (p=0,24; $h^2=0{,}04$) (possibly owing to analysis of a non-normally distributed trait), PAmat scores were significantly correlated among siblings ($p=3{,}3 \times 10^{-9}$) and demonstrated a modest heritability coefficient ($h^2=0{,}30$; $p=1 \times 10^{-7}$). This finding implies preferential inactivation of a parental-specific X chromosome among siblings, consistent with a genetic (possibly X-linked cis-acting) component

to derivation of BC PAmat scores. Autosomal or stochastic processes are unlikely as they would not foster familial correlation of PAmat scores. In addition, segregation analysis of skewed Xi patterns in BC tissue found significant concordance in preferential inactivation of a parent-specific X (Xp versus Xm) among siblings, compelling evidence for an X-linked genetic component.

Candidate gene:

Because the XIC region plays an essential role in Xi, candidate genes likely reside within this region. As discussed in the introduction (Section 7.3.2.1), several observations support an XCE-like etiology in derivation of the primary XIR. As such we speculate that part of the variability of BC PAmat scores reflect genetic variants of XCE-like alleles. Although the latter is consistent with a monogenic trait (see below), several findings in this project support a genetically complex trait. X-linked mutant alleles have been excluded as a major cause of skewed Xi pattern in BC due to lack of an association between BC DS scores and parity data.

Grounds for a genetically complex trait

Assuming i) an X-linked genetic component in derivation of skewed Xi patterns, ii) maternal or paternal inheritance of the skewing gene, iii) genetic linkage of the skewing gene outside of the PAR (thus no genetic recombination expected between the skewing gene and our clonality HUMARA marker on the paternal X) and iv) all female siblings to share the same paternal X, we expected to find some families demonstrating a high incidence of skewed Xi patterns among siblings when the skewing gene was paternally derived (i.e., in 50% of families). However, although some families did demonstrate a high incidence of skewed Xi patterns in BC, in particular family 121 where 5 of 8 siblings were skewed in favor of the same parental X (thus consistent with an X-linked genetic etiology), the finding was exceptional. These findings are consistent with a genetically complex trait. Variability of BC XIRs may plausibly be explained by variable expressivity, incomplete penetrance, polygenic inheritance, genetic heterogeneity, environmental (in utero) influences and/or stochastic processes. In utero influences may include maternal factors such as age, birth order and gestational diseases (e.g., gestational diabetes).

Number of progenitor cells present when Xi is initiated

In light of a genetically complex trait, it appears therefore that XIRs are not strictly reflective of a random/stochastic process dependant on the number of stem cells present at the initiation of X inactivation. Rather, a combination of both a stochastic process and X-linked cis-acting controlling element(s) plausibly play a role in derivation of the primary Xi pattern. We speculate that this X-linked genetic element(s) gives the allusion that a small number of progenitor cells are present when

Xi is initiated. Therefore, we deduce that the number of progenitor present is actually higher than previously suggested.

Linkage analysis of a genetically complex trait

Parametric analysis

One of several methods to map the putative XCE-like gene is by linkage analysis (LA). Genetic linkage refers to the tendency of alleles from two loci on a chromosome to segregate together in a family. The extent of linkage is a function of the distance between two loci, and can be measured by the number of crossover events between the two loci among the observed meioses (i.e., the recombination fraction). The loci are said to be linked if the recombination fraction between them is less than 0,5. Calculation of the lod score estimates the recombination fraction and the significance of the evidence for linkage. Lod score analysis requires specifying a genetic model, which includes mode of inheritance, disease allele frequency and penetrance. In addition, multigeneration families may be required. The trait can be qualitative or quantitative. This model has been successful for one-locus disease models and occasionally for genetically complex traits (early-onset breast cancer for example (Hall JM et al., 1990)).

For the primary Xi trait, several parameters of the genetic model can be derived. Inheritance can be set as X-linked. Setting the trait allele frequency is more difficult as both stochastic (environment) and genetic factors (Table XXXVII, page 176) likely participate in derivation of the primary Xi pattern. Nonetheless, a heritability coefficient of 30% was observed for BC PAmat scores, thus 30% of variability in BC XIRs attributable to X-linked genetic effect(s). Assuming a fully penetrant trait, disease allele frequency can be set to approximately 12,4% (i.e., frequency of skewed Xi pattern in BC) x 30% = 0,037. Alternatively, penetrance may be set lower, thus trait allele frequency set higher. Disease phenotype may be set as affected, unaffected or unknown (the latter is ideal when penetrance is unknown as in complex genetic traits). An affecteds-only approach may also be used for parametric LA of a genetically complex trait since assumptions concerning unaffected individuals are eliminated. Lod scores are summed across families, and a score $\geq 3{,}0$ is indicative of linkage. A lower score requires additional methodologies to narrow the region, this may include multipoint lod score analysis, recruitment of additional families, saturation genotyping of the region and/or analysis of key recombinants. Candidate genes may then be sequenced to identify trait allele(s). As a first approach, saturation genotyping of the XIC region could be performed since the XCE-like element is speculated to reside here.

Nonparametric analysis

Although a parametric approach may be used to map complex genetic traits, it is susceptible to producing misleading results since genetic parameters are not fully known. Alternatively, since nonparametric methodologies of LA do not require knowledge of the genetic model, they have been advocated as the method of choice for mapping susceptibility genes in genetically complex traits. Both qualitative and quantitative traits can be studied by this approach. For common trait alleles, the most powerful affected relative pair sampling unit is affected sibs. For qualitative traits, affected sib-pair linkage is measured by allele sharing (1 or 2 alleles for an X-linked marker) between a known marker and a trait. This is based on the assumption that affected sibs who share marker alleles will be phenotypically similar for traits influenced by a nearby gene. For quantitative traits, testing for linkage between a marker locus and a quantitative trait consists of regressing the squared difference in trait values between two sibs on the number of shared marker alleles. Further, three types of sib-pairs may be analyzed: i) random sampling of sib-pairs, ii) sib-pairs that are both concordant for trait value and iii) sib pairs who are discordant for trait values. The latter appears to be the most powerful (Risch N. and Zhang H, 1995). A more powerful approach are multipoint methods such as variance components approach, which uses sibship data rather than sib pairs (Goldgar DE, 1990). Because the primary Xi trait demonstrates a complex genetic architecture: i.e., genetic and stochastic (environment) factors contribute to trait variance, penetrance unknown, and the fact that most recruited families consist of sib-ships, the sib-pair allele sharing method and/or variance components approach is advocated. Utilizing a panel of X-linked markers, both qualitative and quantitative analyses of the Xi trait can be analyzed for LA as previously performed (Naumova et al., 1998). As a first step, for qualitative analyses, families containing two or more affected sibs may be selected for linkage analysis. This could include both concordant and discordant sib pairs. For quantitative analyses, to obtain maximum information, analyses should utilize all three types of sib-pair models. Once a region of interest is found, and if the region is small enough (1-2 Mbp), one option is to select candidate genes and sequence for functional polymorphisms. Alternatively if this is not possible, the goal is to narrow the regions as much as possible which may include saturation genotyping of the region, recruitment of additional families and reanalysis of the data by LA and/or by allelic association studies (such as TDT).

Allelic association

A further tool to identify susceptibility genes in genetically complex traits is through allelic association. Allelic association refers to significant deviation in occurrence of a marker allele with a trait versus a control group. The association can be biologically true (the *APOE-4* allele in Alzheimer disease for example) or due to tight linkage (i.e., linkage disequilibrium) with a nearby

susceptibility locus, the closer they are the stronger the disequilibrium. Linkage disequilibrium (LD) mapping is especially powerful in isolated populations since trait susceptibility haplotypes in such populations are speculated to have arisen from a common ancestor. LD mapping utilizes all recombination events in that population, thus statistically more powerful versus family-based LA. Allelic association studies can be family and/or population (case-control) based.

Population (case-control) studies compare allele frequency in a group of unrelated affecteds to a matched control group. However spurious associations can arise due to population stratification. Unlike population-based methods, family-based association studies use nuclear family data to estimate control marker allele frequencies, thus avoids possible confounding factors introduced by population sub-structure. The transmission disequilibrium test (TDT) tests for evidence of both linkage and LD when the two co-exist. The TDT is advocated for analysis of genetically complex traits, particularly in relatively isolated populations and for detecting genes of minor or modest effect size. Utilizing mother-father-affected child trios, the TDT evaluates whether the frequency of a particular disease-marker haplotype among affected children deviates from those expected by Hardy-Weinberg equilibrium. Utilizing genotypic data from affected and unaffected siblings the sib-TDT is applicable for late onset disorders (Spielman RS and Ewens WJ, 1998) thus applicable for analysis of the AS trait.

AS trait: hematopoietic XIRs

Evidence of an X-linked genetic component: familial aggregation / correlation analyses

Among hematopoietic lineages, skewed Xi patterns (DS score $\geq$ 0,25) were significantly aggregated within families for monocytes and B lymphocytes. Lack of familial aggregation for PMN and T lymphocytes may be due to small sample size as the criterion to define a skewed Xi pattern may have been too stringent. This may be particularly true for T lymphocytes that demonstrated a low incidence of skewed Xi (20,1%). Pertaining to the AS trait (AS_{DS} score $\geq$ 0,25), although all hematopoietic lineages demonstrated significant familial aggregation, thus evidence in favor of a heritable component to derivation of the AS trait, T lymphocytes demonstrated strongest evidence of familial aggregation. Corroborating data was observed with quantitative analyses: DS, AS_{DS}, PAmat and AS_{PAmat} scores were significantly correlated among siblings, compelling evidence for a genetic (X-linked) component to trait derivation. Strongest familial correlation was observed for B lymphocytes for DS and AS_{DS} scores and strongest for T lymphocytes when PAmat and AS_{PAmat} scores were considered. Surprisingly, although modest evidence for familial aggregation for skewed Xi patterns (DS score $\geq$ 0,25, AS_{DS} score $\geq$ 0,25) in PMN was observed, corroborated by modest familial correlation of DS and AS_{DS} score scores ($h^2 \leq 0,11$), segregation analysis of the AS

phenotype (Table XXXVIII, page 177-183) revealed significant familial aggregation in direction of skewing, compelling evidence for a genetic / X-linked component.

Although deviation from both random Xi (DS score) and from the BC (primary) Xi pattern (i.e., AS_{DS} score) was weaker in lymphocytes versus myeloid lineages (Figures 13 and 15, pages 95 and 97), thus weaker genetic effects speculated for lymphocytic lineages, familial correlation and heritability coefficients of XIRs were approximately similar for all four hematopoietic lineages (Table XXXVII, page 176). This finding implies that utilizing qualitative methods of Xi trait analyses, such as a DS score ≥ 0,25 to define skewed Xi patterns may undermine underlying genetic components to trait variance, thus warranting quantitative trait analyses as the preferred method.

T lymphocyte XIRs: strong evidence of a familial / X-linked genetic component

A stronger heritability coefficient was observed for T lymphocyte PAmat and AS_{PAmat} scores versus PMN, monocytes and B lymphocytes, implying a stronger genetic component to trait derivation, suggesting strong selection of X-linked allele(s) is occurring in T cells. Clonal expansion of T cells is unlikely as this would not foster correlation of PAmat and AS_{DS} scores among siblings. Because T cell DS and AS_{DS} scores were relatively stable with aging suggests that X-linked allele selection occurred primarily prior to age range 38-96, thus an etiology different from AS is speculated.

AS trait: grounds for a genetically complex trait

The above findings provide unequivocal evidence and rationale to map the genetic determinants of the AS trait. Similar to skewed Xi patterns in BC, the AS trait may be analyzed by parametric methods. Inheritance can be set as X-linked; allele frequency can be calculated by factoring the heritability coefficient (Table XXXVII, page 176) by trait frequency (AS_{DS} score ≥ 0,25) (Figure 9, page 91). However, because trait penetrance is unknown, environmental factor(s) (cigarette smoke) possibly implicated, stochastic processes and genetic heterogeneity speculated, a multifactorial (complex) trait is speculated. Consequently, a nonparametric method of linkage analysis is merited.

Candidate gene(s) for AS: modulation of HSC kinetics

Since we postulate modulation of HSC kinetics in derivation of the AS trait and that the X chromosome is peppered with various interleukin receptors, candidate loci include: IL-1R, IL-1-RAPL1, IL-1-RAPL2, IL-2-Rg, IL-3-Ra, IL-9-R, and IL-13-RA2. Other hematopoietic growth factor receptors, signaling peptides or transcription factors include: CSF-2R, BMX, PIG-A, GATA-1 and DKC-1. The role of these genes may be tested by analyzing frequency distribution of polymorphisms in these genes between affecteds and matched control groups. In the event of

positive linkage, a TDT may be pursued. In addition, saturation genotyping may also provide insight. X-linked disease alleles have been excluded as candidate genes since AS was not associated with increased miscarriage occurrence and offspring sex-ratio distortion. However, subtle mutations / polymorphisms in these genes cannot be excluded.

Role of environmental factor(s) in variance of the AS trait: cigarette smoke

Analysis of smoking habits as a potential environmental stimulus in the etiology of AS found current smoker and pack-years of cigarette smoke to be positively associated with the AS_{DS} score of both PMN and monocytes, suggesting cigarette smoke is a potential environmental stimulus in the etiology of AS. That removal of the stimulus (ex-smoker) was negatively associated (statistical trend) with monocyte AS_{DS} scores is a complementary finding, implying a causal role. Slightly smaller p-values with discrete smoking variables may be explained by smaller sample size (thus less power) versus quantitative variables. Absence of an association with T and B lymphocytes may be explained by a lower contribution of these cell types in the AS trait. An undetermined component(s) of cigarette smoke, a biological metabolite activated in response to smoking, and/or a covariate associated with cigarette smoking (diet for example) may be implicated. In light of this finding, the increase of AS_{DS} scores with aging (Part I) may be explained in part by the cumulative exposure of cigarette smoke with advancing age, thus meriting reanalysis of age-XIR analyses controlling for confounding variables.

Since the AS trait is conjectured to reflect modulation of HSC kinetics and that cigarette smoke elicits the production of pro-inflammatory mediators (TNF-α, IL-1, -6, -8) and hematopoietic growth factors (GM-CSF and G-CSF) (reviewed in (Van Eeden and Hogg, 2000)), evidence for a plausible interaction between cigarette smoke and HSC modulation, resulting in AS, is supported. Hypothetically, the X chromosome may carry receptors for any of these mediator(s), and in the heterozygous state, HSCs bearing the 'stronger' receptor will demonstrate a gradual growth advantage, resulting in AS. Further investigation is thus warranted.

The drawback of having environmental factors (cigarette smoke) contributing to trait variance is a reduction in power to detect linkage (Risch NJ and Zhang H, 1996).

Hematologic associations / consequence of AS

Although a relative growth advantage conferred by an X-linked allele(s) in HSCs has been postulated as a mechanism of AS, that the AS trait was not associated with a general increase or decrease of blood counts implies that hematopoiesis is still under the control of 'normal'

hematopoiesis. Rather, the data is consistent with a territorial / 'Darwinian' competition between the two X-linked alleles, with HSCs expressing the stronger allele outcompeting HSC expressing the weaker allele. Nonetheless, blood XIRs were associated with variations in particular lineages, thus possibly associated with lineage-specific hemato-pathologies.

Blood Xi-skewing (DS score) and hemoglobin concentration

Increasing DS scores of PMN and monocytes were associated with an increase of hemoglobin concentration unrelated to anemia. An association with hemolysis (perhaps insufficient to cause clinical hemolysis) is thus speculated. Indeed, evidence for an X-linked genetic component(s) to variability in hemoglobin has been previously demonstrated (Dover et al., 1992), (Huebner et al., 1986). Of interest, mutations in the X-linked *PIG-A* gene [associated with the clinical condition paroxysmal nocturnal hemoglobinuria (PNH)] have been associated with intravascular hemolysis. And although PIG-A mutations are found in normal individuals, albeit a low frequency (0,002%), (Araten et al., 1999), an aplastic BM environment has been speculated as the causative factor for clonal expansion of PIG-A mutant HSCs. The latter is consistent with extrinsic factor(s) in the etiology of clonal expansion (Araten et al., 2002). Moreover, that cytopenias associated with PNH respond to immunosuppresion, thus implying an association with immuno-modulation (possibly autoimmunity favoring selection of the PIG-A mutant clone - i.e., selection against HSCs expressing the wild-type PIG-A gene product), is consistent with the association of the AS trait with autoimmune diseases. We hypothesize that a subtle mutation(s) / polymorphism(s) in the *PIG-A* gene be a candidate locus for skewed Xi patterns in hematopoietic lineages. Supporting evidence stems from the following: like PNH, the AS trait has i) a HSC origin, ii) a putative immuno-modulating role, iii) associated with hemolysis and iv) a role for extrinsic factors (cigarette smoke in the case of AS) in trait evolution.

AS trait (DS and AS_{DS} scores) and eosinophil counts

Increasing DS and AS_{DS} scores of hematopoietic lineages were associated with a decrease in eosinophil counts, suggesting an X-linked genetic basis to eosinophil counts. Physiologically, since eosinophils are implicated in various biological functions, the relation with clinical data merits discussion. Although various allergic diseases are associated with eosinophilia, AS_{DS} scores were not associated with allergic conditions such as asthma and use of med-allergy [Table b) XXXIV (page 171-172)], thus arguing against such a relationship. Alternatively, since the natural role of eosinophils is to defend against parasites, we conjecture that AS (DS and AS_{DS} scores) be associated with an increased occurrence of parasitic infections. However, since parasitic infections are not common in Québec, this type of data was not available from our female participants, thus

meriting investigation in a more relevant population. Nonetheless, employing the extrinsic model of AS, the finding suggests that in lack of particular environmental factors (parasites), X-linked genetic determinant(s) which offer resistance to parasitic infections (i.e., X-linked allelic variants coding for high eosinophil numbers versus low number allele) have a relative growth disadvantage, selectively eliminated with advancing age. An other hypothesis may be that females preferentially expressing X-linked genetic factors coding for high eosinophil numbers demonstrate a survival disadvantage, selectively eliminated with advancing age. In fact, since eosinophilia is associated with increased all-cause mortality (Hospers et al., 2000), we speculate that increasing blood DS/AS_{DS} scores, thus decreased eosinophil count, be associated with increased longevity. The latter may be a significant factor as to why women live longer then men. Although a role for skewed Xi patterns in the longevity differences between males and females has been previously addressed (Christensen et al., 2000), this is the first piece of evidence linking XIRs and eosinophil counts with longevity. A prospective study examining the relationship between XIRs, eosinophil count and longevity is thus merited.

Blood Xi-skewing (DS score) and platelet count

A possible association may exist between increasing blood DS scores and a decrease in platelet count, suggesting skewed Xi patterns in blood may be associated with increased bleeding, mild thrombocytopenia and/or reduced thrombotic complications (thrombosis and hemorrhage). Lack of an association with the AS trait limits the relationship to DS scores of hematopoietic lineages.

Whether all three hematologic findings are related to a single X-linked gene is presently unclear. If limited to a single gene, it is likely to have pleiotropic activity. Alternatively, if multiple genes are implicated, one specific for each hematopoietic lineage, the etiology of skewed Xi patterns is likely to be family-specific. The latter poses a limitation to genetic linkage studies as multiple loci may be identified. To circumvent this problem, Xi traits may be sub-phenotyped by additional clinical / biological features such as hemoglobin concentration, eosinophil and platelet counts for example. As such, genetic linkage studies can be performed on subtypes of skewed Xi traits, thus increasing the power of linkage analysis by plausibly decreasing genetic heterogeneity.

Clinical consequence of the primary Xi pattern

Primary Xi pattern and asthma

Increasing DS scores were associated with an increased reporting of asthma and use of med-asthma. And since the primary XIR is relatively stable with advancing age (as presented in Part I), we

speculate that a skewed Xi pattern in BC tissue poses a risk-factor (rather then a consequence) of asthma susceptibility.: increasing DS scores equates with increased asthma susceptibility. Moreover, since asthma is associated with eosinophilia (De Monchy et al., 1985), we anticipated a positive relationship between BC DS score and eosinophil count. However, as seen in Table a) XXXIII (page 169-170), BC DS scores were not associated with eosinophil counts, arguing against such a relationship. One possibility may be that the clinical condition 'asthma', as it was self-reported, may in fact be a related phenotype (chronic obstructive pulmonary disease (COPD) for example), thus possibly unrelated to eosinophil counts. Interestingly, the negative relationship between eosinophil counts and AS (DS and AS scores) is speculatively unrelated to the relationship between skewing (DS scores) and asthma since the former is linked to the AS trait while the latter to the primary Xi trait.

A plausible association between asthma and skewed Xi patterns is consistent with published data where an X-linked genetic component to asthma susceptibility was reported (Lynch et al., 1999). In this seminal paper, the researchers mapped the gene for the human cysteinyl leukotriene receptor-1 ($CysLT_1$) – an important mediator of human bronchial asthma, to the long arm of the X chromosome (Xq13-Xq21). In addition to the $CysLT_1$ receptor, a second is $CysLT_2$, both of which are activated by cysteinyl leukotrienes (LTC_4, LTD_4, LTE_4) – important mediators of bronchial asthma (Lewis et al., 1990). Biochemical analyses found that activation of the $CysLT_1$ receptor by leukotriene D_4 (LTD_4) resulted in hallmark features of asthma: constriction and proliferation of smooth muscle tissue, eosinophil migration, edema, and damage to the mucosal lining of the lung. Consequently, treatment of asthma has included pharmacologic development of $CysLT_1$ – selective antagonists. Another important X-linked mediator of asthma is interleukin-13.

XCE-like locus may influence expression of X-linked genes

Although the mechanism by which a primary skewed Xi pattern increases the susceptibility to asthma/COPD is presently unclear, one can theorize that upregulation of the X-linked $CysLT_1$ receptor gene may pose a risk-factor for asthma susceptibility. Since the molecular consequences of a primary skewed Xi pattern are unknown, one can theorize that the *Xce* locus, in addition to influencing overall probability of undergoing Xi, increases expression of particular (or chromosome-wide) gene expression. Supporting evidence was drawn from twin studies which demonstrated stronger correlation of X-linked gene expression in MZ versus DZ twins (Brewer et al., 1967), thus consistent with a role of genetic factors influencing quantitative X-linked gene expression. Consequently, increased expression of $CysLT_1$ for example, will translate into an increased susceptibility to asthma/COPD. Quantitative expression analysis of the $CysLT_1$ receptor in skewed versus non-skewed females is thus merited. That the $CysLT_1$ receptor is expressed in

both PB leukocytes and smooth muscle tissue is consistent with our results, namely an association of asthma with increasing DS scores in both hematopoietic lineages and a NHT (i.e., BC). Alternatively, since T lymphocyte release of IL-14 and IL-4 can also up-regulate $CysLT_1$ receptor expression, a further possibility may be that the PS trait affects T lymphocyte immune function.

In spite of a positive association between increasing BC DS scores and asthma susceptibility, eosinophil count was not associated with BC DS scores. Interestingly however, eosinophil count was 'inversely' related to increasing DS and AS_{DS} scores in hematopoietic lineages. One reason may be the use of anti-allergic therapies such as corticosteroids, anti-leukotrienes and anti-interleukin-5 among asthmatic patients, resulting in downregulation of eosinophil survival/counts ((Wallen et al., 1991), (Gauvreau et al., 2000)). However, controlling for med-asthma (Table XXXIII, page 169-170) did not alter the relationship between eosinophil count and DS / AS_{DS} scores, evidence against the latter. As such, the relationship between the primary Xi pattern and asthma susceptibility is likely unrelated to decreasing eosinophil counts as a function of increasing blood DS / AS_{DS} scores.

Preferential inactivation of Xm versus Xp

When parental derivation of the skewed (preferentially inactive) X chromosome was tabulated among all families carrying at least one sibling with a skewed Xi pattern in BC, Xm was found to be preferentially inactivated versus Xp (n=74 versus n=52, p=0,050). Several explanations are possible. 1) Females with the paternal X preferentially inactivated in BC tissue have a survival disadvantage, and are relatively eliminated with advancing age. Parental derivation of the skewed X in young females versus older females may provide insight. 2) Alternatively, the finding may suggest a parent-of-origin effect. As such, perhaps the maternally derived X, relative to the paternal one, is more susceptible to the inactivation signal during embryogenesis. If consistent with an XCE-like etiology of skewed Xi pattern, a possible mechanism may be that the action of *XCE*-like alleles are dependent on parental derivation. That is, consistent with the model in mice where Xce allele strength is correlated with methylation status, it is possible that XCE-like alleles are preferentially methylated (thus relatively stronger) when paternally versus maternally derived, resulting if preferential inactivation of the maternal X. This hypothesis is consistent with stronger social (verbal) skills in Turner syndrome females whom have acquired Xp relative to Xm, consistent with the notion that some X-linked genes are preferentially expressed when paternally versus maternally derived. 3) A final possibility may be that the findings are artifactual, possibly reflecting a bias in PCR amplification of the inactive Xm versus the inactive Xp. That is, perhaps Xp is relatively more heterochromatic versus Xm, therefore less accessible to PCR amplification. However, that the DNA

is denatured to 94°C prior to amplification (thus relaxing secondary DNA structures) argues against the latter.

Clinical associations of the AS trait

AS and susceptibility to RA

Both increasing DS and AS_{DS} scores of hematopoietic lineages were associated with a significant decrease in self-reported RA. Since RA has an autoimmune component, the finding supports published data linking skewed Xi patterns with protection from autoimmune-related phenotypes (Martin-Villa et al., 1999). This finding implies that a RA-related gene(s) resides on the X chromosome, in keeping with published data (Jawaheer et al., 2001). A candidate gene in vicinity to marker DXS6897 has been postulated. Three models have been invoked to elucidate the findings. i) Extrinsic factor(s) influencing selection of X-linked alleles. Hypothetically, similar to negative selection of autoreactive T lymphocytes in the thymus or negative selection of HSC bearing wild-type PIG-A expression (versus *PIG-A* mutations) in an aplastic BM environment, perhaps the BM environment (stromal layer, macrophages, fibroblasts and/or endothelium), for whatever reason, may be dictating negative selection of HSC bearing X-linked RA susceptibility allele(s). ii) Intrinsic factor(s) (X-linked genetic variants) governing HSC growth kinetics. Assuming that a relative growth advantage is gained by HSC expressing an X-linked RA protective allele versus a RA susceptibility allele, the AS trait may reflect the growth competition conferred by HSCs expressing RA protective allele(s). Nonetheless, in light that RA demonstrates a female preponderance possibly exacerbated by sex hormones, if our hypothesis holds true, we speculate that with advancing age, since the frequency of the AS trait increases with aging, female preponderance of RA should decrease. Epidemiologic analysis of RA as a function of advancing age is warranted for confirmation. iii) Another possibility may be that females bearing X-linked RA susceptibility allele(s) (versus RA protective allele) on the preferentially active X are selectively eliminated with advancing age. Thus AS should reflect survival of females bearing RA-protective allele(s).

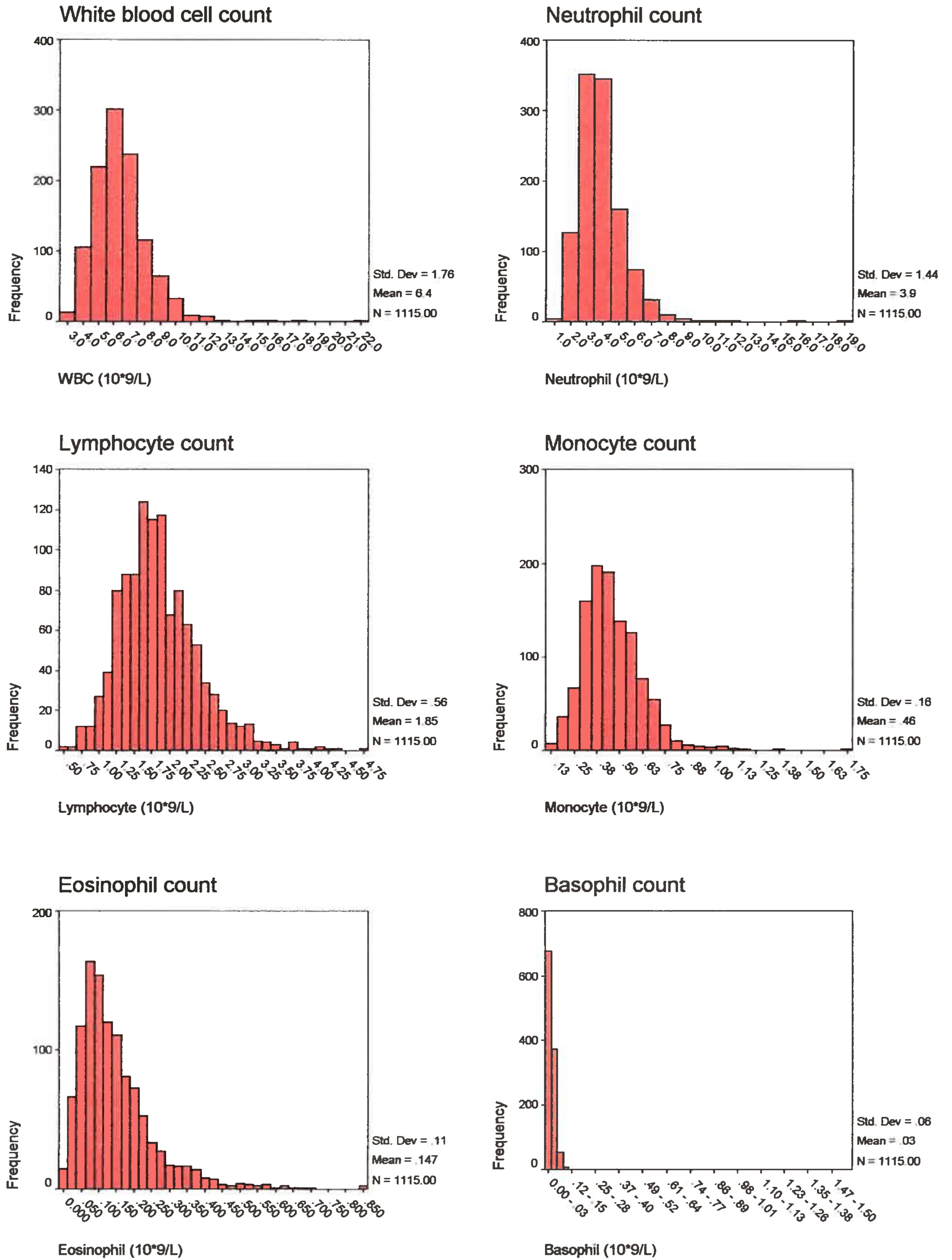
AS and immune-modulation

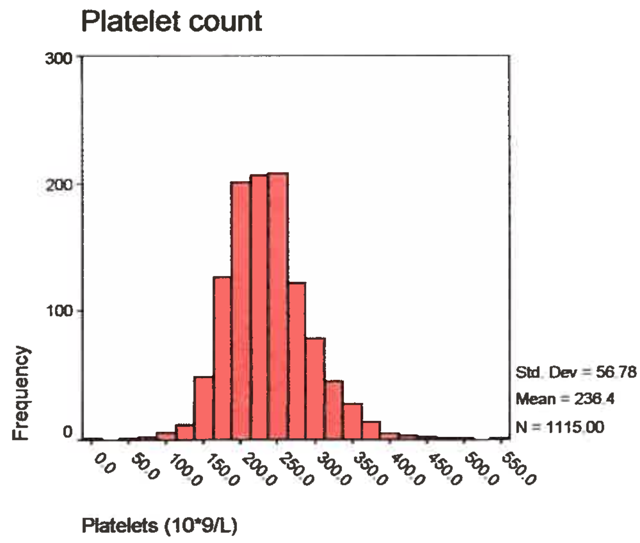
Interestingly, increasing AS_{DS} scores were associated with a 'decreased' occurrence of miscarriages. Because multiple miscarriage occurrences have been associated with autoimmune disorders (Shelton et al., 1994), we hypothesize that the AS trait may be associated with immune-modulation, specifically with a 'less' autoreactive immune system. Supportive findings are reduced eosinophil counts, and reduced reporting of RA with increasing DS and AS_{DS} scores. A candidate gene may reside in the region of Xq28 as a small deletion mapped to this region has been associated with a high rate of SA (Pegoraro et al., 1997). However, that about 70% of human conceptions are lost

before implantation, thus not expected to be reported as miscarriage, may open the possibility that perhaps miscarriage occurrence is equal between skewed and non-skewed females but occurring much earlier in skewed versus non-skewed females. Although this argues against AS playing a role in 'decreased' autoreactivity (as it supports 'increased' autoreactivity to allo-antigens), it nonetheless still supports a role of AS in immune-modulation.

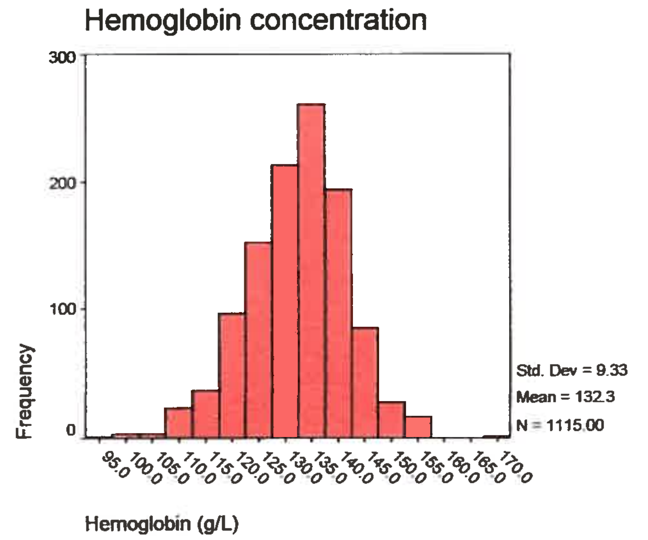
Unlike previous studies citing an association between skewed Xi patterns and invasive ovarian cancer (Buller et al., 1999), we found no significant association between XIRs and all-type cancer. However, reanalysis of XIRs as a function of individual cancer subtype (breast, ovarian/uterus, skin, lymphatic, digestive) requires further elucidation.

Figure 20. Frequency distribution of blood counts (all informative females).

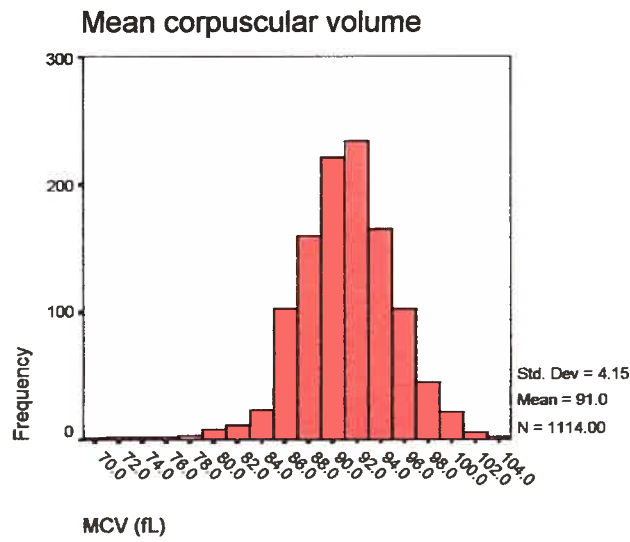




Platelet count
Frequency
300
200
100
0
0.0
50.0
100.0
150.0
200.0
250.0
300.0
350.0
400.0
450.0
500.0
550.0
Std. Dev = 56.78
Mean = 236.4
N = 1115.00
Platelets (10*9/L)

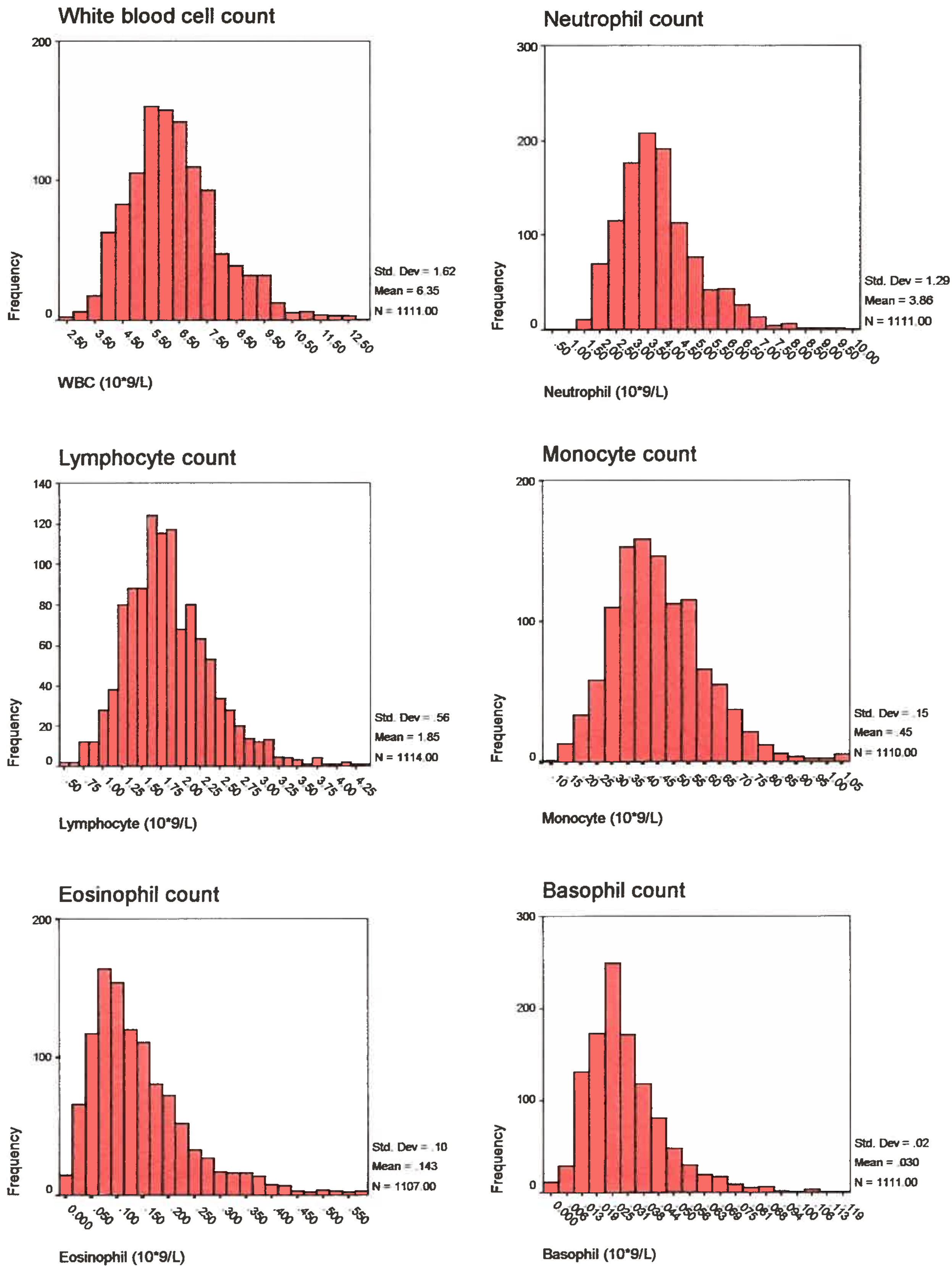


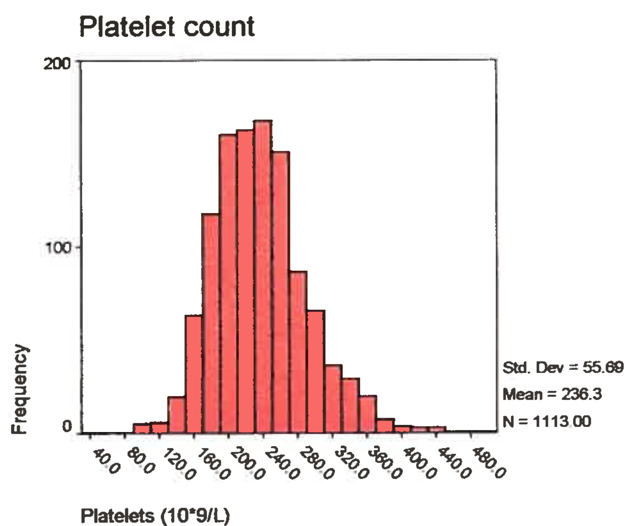
Hemoglobin concentration
Frequency
300
200
100
0
95.0
100.0
105.0
110.0
115.0
120.0
125.0
130.0
135.0
140.0
145.0
150.0
155.0
160.0
165.0
170.0
Std. Dev = 9.33
Mean = 132.3
N = 1115.00
Hemoglobin (g/L)

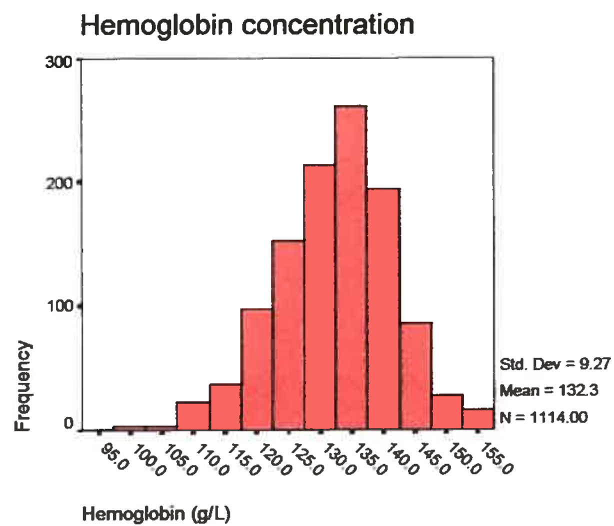


Mean corpuscular volume
Frequency
300
200
100
0
70.0
72.0
74.0
76.0
78.0
80.0
82.0
84.0
86.0
88.0
90.0
92.0
94.0
96.0
98.0
100.0
102.0
104.0
Std. Dev = 4.15
Mean = 91.0
N = 1114.00
MCV (fL)

Figure 21. Frequency distribution of blood counts without outliers.







No outliers found for mean corpuscular volume
(see Figure 20 for frequency distribution of MCV)

Table XV. Descriptive statistics for hematologic indices. All informative females (97,5%) were included. Mean values, range, mean proportion and standard deviations were reported.

parameter	n	Mean Count	Range	Mean proportion
White blood cell	1115	6,387 ± 1,758 (10^9 / L)	2,500 - 22,100	
Neutrophil	1115	3,899 ± 1,441 (10^9 / L)	0,576 - 18,697	0,602 ± 0,082
Lymphocyte	1115	1,848 ± 0,561 (10^9 / L)	0,504 - 4,715	0,296 ± 0,075
Monocyte	1115	0,457 ± 0,162 (10^9 / L)	0,121 - 1,760	0,073 ± 0,022
Eosinophil	1115	0,147 ± 0,108 (10^9 / L)	0 - 0,859	0,023 ± 0,017
Basophil	1115	0,033 ± 0,055 (10^9 / L)	0 - 1,550	0,005 ± 0,010
Platelet	1115	236,4 ± 56,8 (10^9 / L)	9 - 539	
Hemoglobin	1115	132,33 ± 9,33 g/l	93 - 169	
MCV	1114	91,04 ± 4,15 fL	70,3 - 103,6	

Abbreviations: MCV, mean corpuscular volume

Table XVI. Outliers removed from hematopoietic variables for association analyses. The criterion for removal from study was based on extreme values as depicted in Figure 20 (page 146-147). Several of these outliers were confirmed by hematological standards. Individuals outlined in yellow signify more than a single cell type with an extreme value. No outliers were found for MCV.

ID	fam	letter	WBC	neutrophil	lymphocyte	monocyte	eosinophil	basophil	platelet	hemoglobin
527	10	g	22,1	18,70					9	
312	19	c					0,85			
461	19	f	17,9	15,86						
329	24	c			4,72	1,16				
367	31	xa				1,11				
665	37	d					0,65			
3	41	c	15,5	11,93						
14	46	d				1,36				
91	62	b					0,86			
330 ?	63	g	14,6							
53	74	a				1,76				
417	80	e						0,20		
654	95	f					0,63			
198	113	xc					0,61			
428	116	c		11,02						
474	120	a					0,70			
996	127	e				1,11				
997	142	h						0,15		
697	142	b							539	
881	158	b						1,55		
435	161	b					0,64			
747	168	xa						0,90		
614	172	b								169
542	176	a					0,67			
Total excluded			4	4	1	5	8	4	2	1

Table XVII. Descriptive statistics of clinical data. Medical conditions and medicinal use were self-reported.

Clinical characteristics / medicinal use	**n**	**mean**	**range**
Mean date of birth	1144	1936,4	1908 - 1961
Mean age	1144	63,3 ±10,0	38-96
Ever had cancer	99	8,7%	
- chemotherapy	13	1,1%	
- radiotherapy	38	3,3 %	
Asthma	89	7,8%	
Anemia	36	3,2%	
SLE	3	0,3%	
Rheumatoid arthritis	70	6,1%	
Hormone replacement therapy (HRT)	539	47,1%	
Med-allergy	9	0,8 %	
Med-anti-conv	20	1,7 %	
Med-anti-inflam	192	16,8 %	
Med-asthma	38	3,3 %	
Med-other	438	38,3 %	
Med-heart	441	38,5 %	
Med-hemato	44	3,8 %	
Med-vit-other	749	65,5 %	

Table XVIII. Descriptive statistics of smoking habits. Discrete variables included current smoker and ex-smoker. The continuous variables years-smoke and pack-years (pack/day x years-smoked) are the cumulative of both current and ex-smoker. If the subject was a non-smoker, a false value of 0 was assigned for years-smoke and pack-years. Non-smokers were given a null value for years-stop smoking.

Smoking characteristic	**n**	**mean**	**range**
Discrete variables:			
Current smoker	189	16,5 %	
Ex-smoker	378	33,0 %	
Continuous variables:			
Pack-years (current and ex)	1128	7,4 ± 13,7	0-99,2
Years-smoke (current and ex)	1139	10,2 ± 14,5	0-60
Years-stop smoking	546	13,5 ± 13,2	0-65

Table XIX. Descriptive statistics of family / parity data.

	n	mean	range
No. of participation families	193	-	-
No. of female siblings / family	-	7,1 ± 2,1	4-14
No. of participating female siblings / family	-	5,6 ± 1,6	3-12
Number of male siblings	905	4,7 ± 2,2	0-11
Participating mothers	37	19,2 %	-
Living children	3758	2,9 ± 2,0	0-18
Daughters	1816	1,6 ± 1,8	0-11
Sons	1762	1,5 ± 1,6	0-12
Offspring sex-ratio	932	0,49 ± 0,32	-
Ever had spontaneous abortion/miscarriage	352	30,8 %	-
Number of miscarriages		0,49 ± 0,95	0-9
0	794		
1	218		
2	90		
3	26		
4	8		
5	3		
6	0		
7	3		
8	1		
9	1		

Note: two females (id 439, 571) had missing values for number of miscarriages.

Table XX. Association between age and hematopoietic indices. Blood outliers were removed from analyses. The relationship was analyzed by linear regression (LR), Generalized Estimating Equation (GEE) and the SOLAR methods. When significant and known ($p \leq 0,05$), the direction of relationship (+ or -) is given in parentheses. Values in gray indicate a statistical trend but p-values are insignificant.

Cell type (dependant variable)	**n**	**Age**		
		LR *p-value*	**GEE** *p-value*	**SOLAR** *p-value*
WBC count	1111	0,090 (+)	0,075	0,017
Neutrophil count	1110	0,052 (+)	0,071	0,046
Lymphocyte count	1114	0,076 (-)	0,092	0,82
Monocyte count	1110	<0,0001 (+)	<0,0002 (+)	$2,8 \times 10^{-7}$
Eosinophil count	1107	0,0002 (+)	0,0026 (+)	0,0015
Basophil count	1111	0,059 (+)	0,113	0,016
Platelet count	1113	0,074 (-)	0,126	0,28
Hemoglobin conc.	1114	<0,0001 (-)	<0,0001 (-)	0,0043
MCV (fL)	1114	0,0005 (-)	0,0033 (-)	0,038

Abbreviations: conc, concentration; MCV, mean corpuscular volume

Table XXI. Association between age and clinical data: medical conditions and medicinal use. The relationship was analyzed by LR, GEE and SOLAR methods. Clinical data was collected on 1144 females. When significant (p≤0,05) and known, the direction of relationship is given in parentheses.

Medical condition / medicinal use	**Age**		
	LR *p-value*	**GEE** *p-value*	**SOLAR** *p-value*
Cancer	0,037 (+)	n.d.	n.d.
- chemotherapy	0,308	n.d.	n.d.
- radiotherapy	0,157	n.d.	n.d.
asthma	0,558	n.d.	0,8
anemia	0,158	n.d.	n.d.
SLE	0,947	n.d.	n.d.
RA	0,006 (+)	n.d.	0,0034
HRT	<0,0001 (-)	n.d	n.d.
Med-allergy	0,801	n.d.	n.d.
Med-anti-conv	0,887	n.d.	n.d.
Med-anti-inflam	<0,0001 (+)	n.d.	n.d.
Med-asthma	0,193	n.d	n.d.
Med-other	<0,0001 (+)	n.d.	n.d.
Med-heart	<0,0001 (+)	n.d.	n.d.
Med-hemato	0,696	n.d.	n.d.
Med-vit-other	0,078	n.d.	n.d.

Abbreviations: n.d., not determined

Table XXII. Association between age and smoking characteristics. The relationship was analyzed by LR, GEE and the SOLAR method. When significant (p≤0,05), the direction of relationship is given in parentheses. Years-smoke and pack-years of smoke are the cumulative of current and ex-smoker. Smoking habit is the dependent variable.

	Age		
	LR	**GEE**	**SOLAR**
Smoking habit	*p-value*	*p-value*	*p-valus*
Current smoker	<0,0001 (-)	n.d.	n.d.
Ex-smoker	0,001 (+)	n.d.	n.d.
Years-smoke (current and ex)	0,0003 (-)	n.d.	n.d.
Pack-years smoke (current and ex)	<0,0001 (-)	n.d.	0,00011
Years-stop smoking	<0,0001 (+)	n.d.	n.d.

Abbreviations: n.d., not determined

Table XXIII. Association between age and parity data. The relationship was analyzed by LR, GEE and the SOLAR statistical methods. When significant (p≤00,05), the direction of relationship is given in parentheses. a) bi-variate analyses, b) multi-variate analyses

a) bi-variate analyses

Parity data (dependant variable)	**Age** **LR** *p-value*	**GEE** *p-value*	**SOLAR** *p-value*
Had a miscarriage	0,0001(+)	n.d.	n.d.
Number of miscarriages	<0,0001(+)	n.d.	$5x10^{-7}$
Offspring sex-ratio (including mothers)	0,022 (+)	n.d.	n.d.
Offspring sex-ratio (excluding mothers)	0,197	n.d.	n.d.
Number of children	<0,0001(+)	n.d.	n.d.

b) Multi-variate analysis (GLM)

Parity data	**Age** Dependant variable *p-value*
Had a miscarriage	0,871
Number of miscarriages	0,144
Offspring sex-ratio (excluding mothers)	0,416
Number of children	<0,0001 (+)

Abbreviation: n.d., not determined

Table XXIV. Association between clinical data and hematopoietic indices (with outliers). Table a) LR and/or GLM (adjusted for age) analyses. Table b) association between blood counts and asthma / med-asthma: multivariate analyses. When p-values were significant (p≤0,05), the direction of relationship was provided. Blood count was the dependent variable.

a) LR and/or GEE analyses (adjusting for age when necessary)

Blood count (Dependent variable)	Chemo	Radio	HRT	Med-allergy	Med-anti-conv	Med-anti-inflam	Med-asthma	Med-other	Med-heart	Med-hema	Med-vit-other	Cancer	Asthma	Anemia	SLE	RA
WBC -GLM (age)	0,082	0,267	0,294	0,563	0,665	0,209	**0,024 +**	0,121	**0,007 +**	0,198	0,824	0,282	**0,008 +**	0,551	0,191	**0,042 +**
Neutrophil -GLM (age)	0,250	0,371	0,089	0,926	0,810	0,355	**0,006 +**	0,110	0,096	0,521	0,801	0,516	**0,004 +**	0,972	0,285	**0,041 +**
Lymphocyte -LR	**0,046 -**	0,484	0,041	0,204	0,577	0,609	0,409	0,957	**0,038 +**	**0,027 -**	0,757	0,167	0,881	0,119	0,244	0,337
Monocytes -GLM (age)	0,209	0,199	0,060	0,108	0,555	0,886	0,189	0,386	**0,007 +**	0,203	0,910	0,266	0,087	0,240	0,685	0,765
Eosinophil -GLM (age)	0,975	0,408	**0,045 -**	0,197	0,881	**0,004 +**	**0,036 +**	0,204	**0,006 +**	0,328	0,318	0,891	**0,003 +**	0,824	0,880	0,886
Basophil -GLM (age)	0,452	0,998	0,303	0,430	0,934	0,289	0,975	0,571	0,286	0,810	0,121	0,663	0,865	0,862	0,704	0,344
Platelet -LR	0,063	0,300	**0,014 +**	0,500	0,336	0,182	0,968	0,571	0,104	0,895	0,147	0,259	0,576	0,740	0,280	0,599
Hemoglobin -GLM (age)	0,529	0,514	0,196	0,388	0,841	0,515	**0,004 +**	**0,026 -**	**0,034 ?**	**0,001 -**	0,529	0,979	0,324	**0,0001 -**	0,888	0,770
MCV -GLM (age)	0,933	0,634	**0,0005 +**	0,170	**0,006 +**	0,103	0,457	0,311	0,252	0,258	0,324	0,053	0,084	**0,002 -**	0,591	0,459

Abbreviations: WBC, white blood cells; MCV, mean corpuscular volume; chemo, chemotherapy; radio, radiotherapy; HRT, hormonotherapy; nd, not determined.

b) Association between asthma / med-asthma: multivariate analysis

	Blood count (dependant variable)		
	WBC	Neutrophil	Eosinophil
Age	0,124	**0,001** (+)	**0,0002** (+)
Asthma	0,082	0,082	**0,023** (+)
Med-asthma	0,273	0,124	0,497

Table XXV. Association between clinical data and parity. Age was included as a confounding variable when significant (see Tables XXI and XXIII, pages 156, 158). When significant (p≤0,05), the direction of relationship is given in parentheses. Parity data was the dependent variable.

Medical condition or medicinal therapy (dependant variable)	**Parity data** had a miscarriage	number of miscarriages	offspring sex-ratio	number of children
Chemotherapy				
-GLM (age)	0,622	0,965	0,598	0,831
Radiotherapy				
-GLM (age)	0,503	0,841	0,358	0,577
HRT				
-GLM (age)	0,733	0,846	0,043 (+)	0,300
Med-allergy				
-GLM (age)	0,355	0,189	0,276	0,623
Med-anti-conv				
-GLM (age)	0,927	0,980	0,377	0,202
Med-anti-inflam				
-GLM (age)	0,111	0,158	0,844	0,017 (+)
Med-asthma				
-GLM (age)	0,297	0,048 (+)	0,892	0,223
Med-other				
-GLM (age)	0,988	0,200	0,284	0,400
Med-heart				
-GLM (age)	0,010 (+)	0,055 (+)	0,516	0,004 (+)
-GLM (age, num of children)	0,057	0,200	-	-
Med-hemato				
-GLM (age)	0,576	0,419	0,749	0,188
Med-vit-others				
-GLM (age)	0,103	0,934	0,643	0,190
Cancer				
-GLM (age)	0,150	0,141	0,498	0,084
Asthma				
-GLM (age)	0,060 (+)	0,102 (+)	0,211	0,473
Anemia				
-GLM (age)	0,851	0,608	0,744	0,272
SLE				
-GLM (age)	0,929	0,764	0,747	0,222
RA				
-GLM (age)	0,156 (+)	0,031 (+)	0,887	0,042 (+)
-GLM (age, num of children)	0,346	0,083	-	-

Table XXVI. Association between smoking habits and clinical data. The relationship was analyzed by LR, GLM and SOLAR method in some instances. Age was included as a covariate when necessary. When significant (p≤0,05), the direction of relationship is given in parentheses. Direction of association is not given for the SOLAR method.

	Smoking habit				
Medical condition or medicinal therapy (dependant variable)	Current-smoker	Ex-smoker	Pack-years	Years-smoke	Years-stop smoke
Chemotherapy					
-LR	0,389	0,802	0,246	0,196	0,618
Radiotherapy					
-LR	0,920	0,682	0,187	0,054	0,063
HRT					
-GLM (age)	0,034 (-)	0,124	0,525	0,849	0,119
Med-allergy					
-LR	0,181	0,181	0,541	0,878	0,758
Anti-convulsants					
-LR	0,673	0,371	0,246	0,517	0,380
Anti-inflammatory					
-GLM (age)	0,763	0,780	0,160	0,105	0,275
Med-asthma					
-LR	0,902	0,620	<0,0001 (+)	0,002 (+)	0,408
Med-other					
-GLM (age)	0,162	0,507	0,003 (+)	0,0002 (+)	0,442
Med-heart					
-GLM (age)	0,891	0,949	0,018 (+)	0,003 (+)	0,163
Med-hematopoietic					
-LR	0,348	0,512	0,139	0,287	0,351
HRT					
-GLM (age)	0,055	0,112	0,521	0,847	0,078
Med-vit-others					
-LR	0,195	0,315	0,740	0,705	0,504
Cancer					
-GLM	0,532	0,308	0,975	0,246	0,084
Asthma					
-LR	0,930	0,698	0,002 (+)	0,175	0,466
-SOLAR	0,16	nd	0,0087	nd	nd
Anemia					
-LR	0,065 (+)	0,020 (-)	0,592	0,393	0,152
SLE					
-LR	0,441	0,346	0,528	0,798	0,388
RA					
-GLM (age)	0,184	0,086	0,577	0,700	0,682
-SOLAR	0,034	nd	0,039	nd	nd

Table XXVII. Association between smoking habits and blood counts (without outliers). Statistical methods included: LR, GLM, GEE and SOLAR. When significant ($p \leq 0,05$), the direction of relationship (+ or -) is given in parentheses. Hematopoietic indices were the dependent variable.

Blood count (Dependent variable)	**Smoking habit** Current-smoker	Ex-smoker	Pack-years	Years-smoke	Years-stop smoke
WBC					
-GLM (age)	<0,0001 +	<0,0001 -	<0,0001 +	<0,0001 +	0,002 -
-GEE (age)	<0,0001 +	0,819	<0,0001 +	nd	nd
-SOLAR	0,007	nd	0,0003	nd	nd
Neutrophil					
-GLM (age)	<0,0001 +	<0,0001	<0,0001 +	<0,0001 +	0,017 -
-GEE (age)	<0,0001 +	0,894	<0,0001 +	nd	nd
-SOLAR	0,088	nd	0,0009	nd	nd
Lymphocyte					
-LR	<0,0001 +	<0,0001 -	<0,0001 +	<0,0001 +	0,001 -
-GEE (age)	0,0007 +	0,500	<0,0001 +	nd	nd
-SOLAR	0,002	nd	0,031	nd	nd
Monocytes					
-GLM (age)	0,002 +	0,020 -	<0,0004 +	<0,0001 +	0,072
-GEE (age)	0,0042 +	0,716	0,0043 +	nd	nd
-SOLAR	0,12	nd	0,10	nd	nd
Eosinophil					
-GLM (age)	0,671	0,566	0,137	0,046 +	0,707
-GEE (age)	0,646	0,088	0,246	nd	nd
-SOLAR	0,96	nd	0,39	nd	nd
Basophil					
-GLM (age)	0,022 +	0,410	<0,0001 +	<0,0001 +	0,659
-GEE (age)	0,036 +	0,171	0,0012 +	nd	nd
-SOLAR	0,0002	nd	0,40	nd	nd
Platelet					
-LR	0,048 +	0,212	0,028 +	0,001 +	0,392
-GEE (age)	0,213	0,957	0,109	nd	nd
-SOLAR	0,47	nd	0,12	nd	nd
Hemoglobin					
-GLM (age)	<0,0001 +	<0,0001 -	<0,0001 +	<0,0001 +	0,0001 -
-GEE (age)	<0,0001 +	0,939	<0,0001 +	nd	nd
-SOLAR	0,0038	nd	0,0014	nd	nd
MCV					
-GLM (age)	<0,0001 +	<0,0001 -	0,001 +	<0,0001 +	0,002 -
-GEE (age)	<0,0001 +	0,0122 -	0,0002 +	nd	nd
-SOLAR	0,00003	nd	0,084	nd	nd

Table XXVIII. Association between smoking habits and parity. When age was a significant confounding variable (see Tables XXII and XXIII, pages 157, 158), it was included as a confounding variable. When significant ($p \leq 0,05$), the direction of relationship (+ or -) is given in parentheses. Parity data was the dependent variable.

	Smoking habit				
Parity data (Dependent variable)	Current-smoker	Ex-smoker	Pack-years	Years-smoke	Years-stop smoke
Had a miscarriage					
-GLM (age)	0,555	0,511	0,328	0,562	0,261
Number of miscarriages					
-GLM (age)	0,535	0,444	0,544	0,466	0,746
-SOLAR	0,40	-	0,40	-	-
Offspring sex-ratio					
-GLM (age)	0,916	0,722	0,809	0,796	0,710
Number of children					
-GLM (age)	0,419	0,111	0,064 (-)	0,020 (-)	0,155

Table XXIX. Association between parity and blood counts (without outliers). When significant (see Table XX, page 155), age was included as a confounding variable, stated in parentheses next to variable name. When significant ($p \leq 0,05$), the direction of relationship (+ or -) is given in parentheses. Blood count was the dependent variable.

Blood count (dependant variable)	Parity data			
	Had a miscarriage	Number of miscarriages	Offspring sex-ratio	Number of children
WBC - GLM (age)	0,470	0,235	0,507	0,976
Neutrophil - GLM (age)	0,919	0,445	0,486	0,641
Lymphocyte - LR	0,373	0,563	0,442	0,844
Monocytes - GLM (age)	0,711	0,990	0,156	0,183
Eosinophil - GLM (age)	0,165	0,377	0,126	0,038 (+)
Basophil - GLM (age)	0,281	0,192	0,048 (+)	0,075
Platelet - LR	0,527	0,859	0,257	0,533
Hemoglobin - GLM (age)	0,706	0,082	0,411	0,363
MCV - GLM (age)	0,202	0,519	0,911	0,080

Table XXX. Covariates associated with hematopoietic indices: multivariate analyses (GLM).

covariates	WBC count	covariates	Basophil count
Age	0,377	Age	0,024 (+)
Med-asthma	0,380	Smoker	0,001 (+)
Med-heart	0,017 (+)	Pack-years	0,918
Asthma	0,069		
RA	0,928		
Smoker	0,296		
Ex-smoker	0,008 (-)		
Pack-years	0,004 (+)		
	Neutrophil count		**Platelets count**
Age	0,098	HRT	0,035 (+)
Med-asthma	0,807	Smoker	0,263
Asthma	0,009 (+)	Pack-years	0,104
RA	0,980		
Smoker	0,618		
Ex-smoker	0,056		
Pack-years	0,009 (+)		
	Lymphocyte count		**Hemoglobin conc**
Chemotherapy	0,333	Age	0,063
Med-heart	0,168	Med-asthma	<0,0001 (+)
Med-hematopoietic	0,833	Med-other	0,005 (-)
Smoker	0,052	Med-heart	0,180
Ex-smoker	0,001 (-)	Med-hemato	0,896
Pack-years	0,058	Anemia	<0,0001 (-)
		Smoker	0,751
		Ex-smoker	0,116
		Pack-years	0,043 (+)
	Monocyte count		**MCV**
Age	<0,0001 (+)	Age	0,025 (-)
Med-heart	0,145	HRT	0,073
Smoker	0,663	Med-anti-conv	0,015 (+)
Ex-smoker	0,560	Anemia	0,084
Pack-years	0,177	Smoker	0,776
		Ex-smoker	0,055
		Pack-years	0,609
	Eosinophil count		
Age	0,179		
HRT	0,041 (-)		
Med-anti-inflam	0,018 (+)		
Med-asthma	0,634		
Med-heart	0,062		
Asthma	0,043 (+)		
Number of children	0,337		

Table XXXI. Heritability estimates of confounding variables. Included in the analyses were blood counts, number of miscarriages and pack-years of smoke. Provided are heritability coefficients (h^2), standard error (SE) and corresponding p-values.

Variable	Heritability coefficient h^2 +/- SE	p-value
Blood counts		
White blood cell	0,27 +/- 0,056	$1x10^{-7}$
Neutrophil	0,22 +/- 0,05	$1x10^{-7}$
Lymphocyte	0,43 +/- 0,06	$1x10^{-7}$
Monocyte	0,40 +/- 0,06	$1x10^{-7}$
Eosinophil	0,25 +/- 0,056	$1x10^{-7}$
Basophil	0,14 +/- 0,08	0,045
Platelet	0,54 +/- 0,067	$1x10^{-7}$
Hemoglobin	0,38 +/- 0,061	$1x10^{-7}$
MCV	0,66 +/- 0,067	$1x10^{-7}$
Parity data		
Number of miscarriages	0,12 +/- 0,05	0,002
Smoking characteristics		
Pack-years of smoke	0,18 +/- 0,05	0,00002

Table XXXII. Familial aggregation of the skewing (DS score $\geq 0,25$) and AS (AS_{DS} score $\geq 0,25$) traits. **Methods:** A proband was randomly selected from each family and categorized as case (skewed) or control (random X inactivation). The relative risk (RR) of skewing in remaining siblings was then determined for cases (RR_{case}) and controls ($RR_{control}$). The relative risk ratio (RRR), derived by the following ratio: RR_{case} / $RR_{control}$, provides an indication as to whether the trait is aggregated within families. A RRR-score >1,00 is normally indicative of familial aggregation of the trait. Environmental versus genetic influences however, are not discernible. RRR-scores highlighted in bold are significant (p-values given).

Cell type and skewing trait	**RR_{case}**	**$RR_{control}$**	**RRR**	**p-value (chi^2-value)**
BC skewing	11/81 = 0,1358	96/716 = 0,1341	1,01	p>0,95 (0,002)
PMN skewing	118/302 = 0,3907	179/510 = 0,3510	1,12	p>0,10 (1,66)
Monocyte skewing	115/265 = 0,4340	171/505 = 0,3386	**1,28**	p<0,01 (6,76)
T cell skewing	37/154 = 0,2403	129/638 = 0,2022	1,19	p>0,25 (1,08)
B cell skewing	75/191 = 0,3927	144/595 = 0,2420	**1,62**	p<0,001 (16,32)
PMN AS	39/117 = 0,3333	113/580 = 0,1948	**1,71**	p<0,005 (10,59)
Monocyte AS	66/191 = 0,3455	104/472 = 0,2203	**1,57**	p<0,001 (11,19)
T cell AS	15/58 = 0,2586	61/621 = 0,0982	**2,63**	p<0,001 (17,74)
B cell AS	23/100 = 0,2300	83/584 = 0,1421	**1,62**	p<0,01 (5,03)

Table XXXIII. Association between skewing and blood counts (without outliers). Analyses included LR, GLM and GEE. Confounding variables for GLM and GEE are stated in parentheses. P-values and the type of relationship (+ or -) for DS score (a) and AS_{DS} score (b) are provided.

A) **DS score**

Blood count (dependent variable)	**BC** *p-value*	**PMN** *p-value*	**Mono** *p-value*	**T cells** *p-value*	**B cells** *p-value*
WBC					
-LR	0,242	0,468	0,675	0,584	0,326
-GLM (age, m-ast, m-hrt, ast, RA, smk, pk-yr, x-smk)	0,628	0,745	0,458	0,757	0,355
-GEE (age, smk, pk-yr, x-smk)	0,194	0,802	0,387	0,688	0,161
Neutrophil					
-LR	0,352	0,352	0,953	0,722	0,399
-GLM (age, m-ast, ast, RA, smk, pk-yr, x-smk)	0,943	0,965	0,764	0,829	0,227
-GEE (age, smk, pk-yr, x-smk)	0,272	0,911	0,475	0,843	0,180
Lymphocyte					
-LR	0,409	0,593	0,710	0,096	0,751
-GLM (chemotherapy, m-hrt, m-hem, smk, pk-yr, x-smk)	0,230	0,451	0,634	0,332	0,902
-GEE (age, smk, pk-yr, x-smk)	0,407	0,507	0,984	0,107	0,857
Monocyte					
-LR	0,572	0,074	0,473	0,765	0,472
-GLM (age, m-hrt, smk, pk-yr, x-smk)	0,300	0,820	0,608	0,853	0,997
-GEE (age, smk, pk-yr, x-smk)	0,270	0,433	0,299	0,792	0,404
Eosinophil					
-LR	0,248	0,024 (-)	0,014 (-)	0,019 (-)	0,0019 (-)
-GLM (age, HRT, m-anti-inflam, m-ast, m-hrt, ast, num of offspr)	0,466	0,004 (-)	0,004 (-)	0,010 (-)	0,0008 (-)
-GEE (age, smk, pk-yr, x-smk)	0,091	0,007 (-)	0,004 (-)	0,041 (-)	0,0002 (-)
Basophil					
-LR	0,504	0,135	0,851	0,265	0,983
-GLM (age, smk, pk-yrs)	0,653	0,327	0,596	0,241	0,989
-GEE (age, smk, pk-yr, x-smk)	0,587	0,396	0,446	0,271	0,698
Platelet					
-LR	0,260	0,492	0,006 (-)	0,060 (-)	0,054 (-)
-GLM (HRT, pk-yr)	0,378	0,650	0,010 (-)	0,073 (-)	0,084 (-)
-GEE (age, smk, pk-yr, x-smk)	0,155	0,902	0,049 (-)	0,170	0,028 (-)
Hemoglobin					
-LR	0,614	0,048 (+)	0,022 (+)	0,785	0,312
-GLM (age, m-ast, m-oth, m-hrt, m-hem, an, smk, pk-yr, x-smk)	0,218	0,015 (+)	0,003 (+)	0,859	0,326
-GEE (age, smk, pk-yr, x-smk)	0,666	0,025 (+)	0,048 (+)	0,833	0,427
MCV					
-LR	0,330	0,118	0,533	0,091	0,400
-GLM (age, HRT, m-anti-conv, an, smk, pk-yr, x-smk)	0,049 (+)	0,689	0,974	0,268	0,479
-GEE (age, smk, pk-yr, x-smk)	0,663	0,116	0,920	0,012 (-)	0,209

B) AS_{DS} score

Blood count (dependent variable)	PMN *p-value*	Mono *p-value*	T cells *p-value*	B cells *p-value*
WBC				
-LR	0,637	0,722	0,103	0,495
-GLM (age, m-ast, m-hrt, ast, RA, smk, pk-yr, x-smk)	0,327	0,649	0,123	0,478
-GEE (age, smk, pk-yr, x-smk)	0,142	0,463	0,053	0,246
Neutrophil				
-LR	0,668	0,705	0,190	0,627
-GLM (age, m-ast, ast, RA, smk, pk-yr, x-smk)	0,158	0,476	0,073	0,277
-GEE (age, smk, pk-yr, x-smk)	0,107	0,207	0,073	0,234
Lymphocyte				
-LR	0,927	0,306	0,780	0,883
-GLM (chemotherapy, m-hrt, m-hem, smk, pk-yr, x-smk)	0,841	0,400	0,852	0,850
-GEE (age, smk, pk-yr, x-smk)	0,858	0,108	0,941	0,818
Monocyte				
-LR	0,758	0,491	0,919	0,590
-GLM (age, m-hrt, smk, pk-yr, x-smk)	0,406	0,657	0,772	0,899
-GEE (age, smk, pk-yr, x-smk)	0,234	0,111	0,764	0,212
Eosinophil				
-LR	0,112	0,384	0,027 (-)	0,064 (-)
-GLM (age, HRT, m-anti-inflam, m-ast, m-hrt, ast, num of offspr)	0,141	0,482	0,065 (-)	0,050 (-)
-GEE (age, smk, pk-yr, x-smk)	0,035 (-)	0,357	0,013 (-)	0,011 (-)
Basophil				
-LR	0,158	0,283	0,461	0,324
-GLM (age, smk, pk-yrs)	0,394	0,457	0,450	0,434
-GEE (age, smk, pk-yr, x-smk)	0,315	0,542	0,475	0,591
Platelet				
-LR	0,066 (+)	0,716	0,132	0,615
-GLM (HRT, pk-yr)	0,065 (+)	0,730	0,118	0,618
-GEE (age, smk, pk-yr, x-smk)	0,043 (+)	0,788	0,061	0,714
Hemoglobin				
-LR	0,800	0,935	0,457	0,921
-GLM (age, m-ast, m-oth, m-hrt, m-hem, an, smk, pk-yr, x-smk)	0,418	0,463	0,379	0,974
-GEE (age, smk, pk-yr, x-smk)	0,665	0,984	0,203	0,962
MCV				
-LR	0,667	0,486	0,068	0,225
-GLM (age, HRT, m-anti-conv, an, smk, pk-yr, x-smk)	0,873	0,136	0,840	0,534
-GEE (age, smk, pk-yr, x-smk)	0,510	0,208	0,190	0,059

Abbreviations: m-ast, med-asthma; m-hrt, med-heart; ast, asthma; smk, current smoker; pk-yr, pack-years of cigarette smoke; x-smk, ex-smoker; m-hem, med-hemato; m-oth, med-other; an, anemia; m-anti-conv, med-anti-conv; num of offspr, number of offspring

Table XXXIV. Association between skewing and clinical data. Analyses included LR, GLM and GEE. Confounding variables are printed in parentheses. P-values and direction of relationship (+ or -) are given when known for DS score (a) and AS_{DS} score (b).

A) **DS score**

Clinical data (dependant variable)	**BC** *p-value*	**PMN** *p-value*	**Mono** *p-value*	**T cells** *p-value*	**B cells** *p-value*
Cancer					
-LR	0,594	0,650	0,438	0,156	0,611
-GLM (age)	0,546	0,814	0,567	0,180	0,671
-GEE (age)	0,479	0,703	0,440	0,133	0,680
Chemotherapy					
-LR	0,876	0,131	0,524	0,683	0,685
-GLM (lymphocyte)	0,898	0,131	0,523	0,601	0,714
Radiotherapy					
-LR	0,889	0,537	0,323	0,145	0,743
-GLM (years-smoke, years stop-smoke)	0,080	0,700	0,544	0,960	0,816
Asthma					
-LR	0,032 (+)	0,032 (+)	0,123	0,143	0,160
-GLM (WBC, neut, eos, pk-yr)	0,051 (+)	0,028 (+)	0,113	0,129	0,104
-GEE (age, eos, pk-yr, pk-yr sec)	0,063	0,016 (+)	0,054 (+)	0,101	0,105
Anemia					
-LR	0,691	0,902	0,291	0,634	0,742
-GLM (hemo, MCV, smk, x-smk)	0,409	0,268	0,913	0,773	0,545
-GEE (age, MCV)	0,605	0,940	0,242	0,887	0,605
SLE					
-LR	0,647	0,578	0,176	0,522	0,237
RA					
-LR	0,546	0,048 (-)	0,040 (-)	0,317	0,829
-GLM (age, WBC, neut, smk, x-smk)	0,534	0,062 (-)	0,059 (-)	0,784	0,399
-GEE (age)	0,604	0,007 (-)	0,010 (-)	0,326	0,708
HRT					
-LR	0,634	0,023 (-)	0,488	0,745	0,214
-GLM (age, eos, plat, MCV, smk)	0,365	0,226	0,980	0,081	0,408
Med-allergy					
-LR	0,721	0,397	0,261	0,314	0,531
Med-anti-inflam					
-LR	0,218	0,610	0,627	0,909	0,269
-GLM (age, eos)	0,205	0,748	0,297	0,805	0,121
Med-asthma					
-LR	0,331	0,028 (+)	0,111	0,771	0,112
-GLM (WBC, neut, eos, hemo, pk-yr)	0,304	0,013 (+)	0,081 (+)	0,453	0,041 (+)
Med-heart					
-LR	0,765	0,084	0,135	0,885	0,749
-GLM (age, WBC, lymp, mono, eos, hemo, pk-yr)	0,679	0,996	0,783	0,968	0,543

B) AS_{DS} score

Clinical condition (dependant variable)	PMN *p-value*	Mono *p-value*	T cells *p-value*	B cells *p-value*
Cancer				
-LR	0,797	0,489	0,147	0,632
-GLM (age)	0,525	0,697	0,200	0,813
-GEE (age)	0,559	0,654	0,245	0,810
Chemotherapy				
-LR	0,053	0,513	0,884	0,506
-GLM (lymphocyte)	0,056	0,567	0,900	0,498
Radiotherapy				
-LR	0,854	0,863	0,468	0,967
-GLM (years smoke, years stop smoke)	0,911	0,685	0,287	0,895
Asthma				
-LR	0,196	0,487	0,692	0,756
-GLM (WBC, neut, eos, pk-yr)	0,182	0,396	0,813	0,707
-GEE (age, eos, pk-yr, pk-yr sec)	0,130	0,378	0,978	0,514
Anemia				
-LR	0,277	0,026 (-)	0,066 (-)	0,033 (-)
-GLM (hemo, MCV, smk, x-smk)	0,775	0,318	0,143	0,701
-GEE (age, MCV)	0,292	0,002 (-)	0,219	0,050 (-)
SLE				
-LR	0,103	0,590	0,138	0,266
RA				
-LR	0,095	0,083	0,782	0,540
-GLM (age, WBC, neut, smk, x-smk)	0,045 (-)	0,093 (-)	0,587	0,632
-GEE (age)	0,064 (-)	0,057 (-)	0,818	0,471
HRT				
-LR	0,711	0,365	0,341	0,556
-GLM (age, eos, plat, MCV, smk)	0,888	0,211	0,303	0,761
Med-allergy				
-LR	0,740	0,176	0,987	0,424
Med-anti-inflam				
-LR	0,697	0,111	0,279	0,276
-GLM (age, eos)	0,336	0,055	0,147	0,209
Med-asthma				
-LR	0,808	0,643	0,070	0,505
-GLM (WBC, neut, eos, hemo, pk-yr)	0,666	0,588	0,139	0,572
Med-heart				
-LR	0,893	0,709	0,168	0,771
-GLM (age, WBC, lymp, mono, eos, hemo, pk-yr)	0,051	0,291	0,041 (-)	0,533

Abbreviations: eos, eosinophil count; pk-yr, pack-years of cigarette smoke; pack-yrs cig sec, pack-years of cigarette smoke second-hand smoke exposure; RA, rheumatoid arthritis, SLE, systemic lupus erythematosus; HRT, hormone replacement therapy

Table XXXV. Parity data and skewing (search for X-linked disease alleles). Analyses included LR, GLM and GEE. Confounding variables are given in parentheses. P-values and direction of relationship (+ or -) are provided for DS scores (a) and AS_{DS} scores (b).

A) **DS score**

Parity data (dependant variable)	**BC** *p-value*	**PMN** *p-value*	**Mono** *p-value*	**T cells** *p-value*	**B cells** *p-value*
Offspring sex-ratio					
-LR	0,610	0,948	0,305	0,316	0,979
-GLM (HRT)	0,599	0,859	0,289	0,311	0,964
-GEE (age)	0,590	0,913	0,379	0,355	0,934
Miscarriage occurrence					
-LR	0,508	0,140	0,454	0,793	0,444
-GLM (age, med-hrt, num of child)	0,692	0,132	0,336	0,731	0,455
-GEE (age)	0,393	0,048 (-)	0,267	0,687	0,302
Number of miscarriages					
-LR	0,907	0,429	0,984	0,383	0,419
-GLM (age, med-asthma, RA, num of child)	0,857	0,384	0,876	0,498	0,479
-GEE (age)	0,687	0,330	0,738	0,480	0,540
Number of offspring					
-LR	0,349	0,954	0,544	0,722	0,941
-GLM (age, med-anti-inflam, med-hrt, RA)	0,119	0,047 (-)	0,456	0,727	0,553

B) **AS_{DS} score**

Parity data (dependant variable)	**PMN** *p-value*	**Mono** *p-value*	**T cells** *p-value*	**B cells** *p-value*
Offspring sex-ratio				
-LR	0,957	0,656	0,571	0,882
-GLM (HRT)	0,978	0,696	0,522	0,819
-GEE (age)	0,841	0,701	0,359	0,972
Miscarriage occurrence				
-LR	0,046 (-)	0,118	0,247	0,141
-GLM (age, med-hrt, num of child)	0,030 (-)	0,036 (-)	0,260	0,041 (-)
-GEE (age)	0,019 (-)	0,057 (-)	0,235	0,096
Number of miscarriages				
-LR	0,290	0,443	0,422	0,474
-GLM (age, med-asthma, RA, num of child)	0,197	0,253	0,355	0,212
-GEE (age)	0,167	0,288	0,356	0,419
Number of offspring				
-LR	0,592	0,123	0,865	0,070
-GLM (age, med-anti-inflam, med-hrt, RA)	0,259	0,578	0,620	0,437

Table XXXVI. Association between smoking habits and skewing: role of smoke as a plausible environmental stimulus. Analyses included LR, GLM and GEE. Confounding variables are provided in parentheses. P-values and direction of relationship (+ or -) are provided for DS scores (a) and AS_{DS} scores (b).

A) **DS score**

Cell type – skewing (dependant variable)	**Current smoker** *p-value*	**Ex-smoker** *p-value*	**Pack-yrs** *p-value*
BC			
-LR	0,277	0,740	0,903
-GLM (asthma)	0,284	0,730	0,918
-GEE (age)	0,375	0,278	0,920
-SOLAR	0,07	-	0,19
PMN			
-LR	0,384	0,551	0,978
-GLM (age, eos, hemo, asthma, RA, HRT, m-ast, SA)	0,167	0,600	0,941
-GEE (age)	0,077	0,972	0,581
-SOLAR	0,11	-	0,71
Monocytes			
-LR	0,840	0,620	0,497
-GLM (age, eos, plat, hemo, RA)	0,534	0,702	0,714
-GEE (age)	0,569	0,921	0,949
-SOLAR	0,60	-	0,70
T cells			
-LR	0,353	0,251	0,082
-GLM (eos)	0,346	0,233	0,151
-GEE (age)	0,297	0,022 (-)	0,089
-SOLAR	0,20	-	0,09
B cells			
-LR	0,314	0,082	0,264
-GLM (eos)	0,230	0,579	0,571
-GEE (age)	0,214	0,089	0,385
-SOLAR	0,04	-	0,07

B) AS_{DS} score

Cell type – skewing (dependant variable)	**Current smoker** *p-value*	**Ex-smoker** *p-value*	**Pack-yrs** *p-value*
PMN			
-LR	0,375	0,457	0,079 (+)
-GLM (age, SA)	0,067 (+)	0,204	0,012 (+)
-GEE (age)	0,061	0,787	0,009 (+)
-SOLAR	0,78	-	0,59
Monocytes			
-LR	0,325	0,202	0,168
-GLM (age, anemia)	0,074 (+)	0,081 (-)	0,069 (+)
-GEE (age)	0,083	0,569	0,032 (+)
-SOLAR	0,74	-	0,61
T cells			
-LR	0,664	0,470	0,658
-GLM (age, eosinophil)	0,268	0,342	0,999
-GEE (age)	0,591	0,581	0,720
-SOLAR	0,81	-	0,72
B cells			
-LR	0,553	0,565	0,766
-GLM (age, anemia)	0,189	0,318	0,491
-GEE (age)	0,319	0,343	0,386
-SOLAR	0,52	-	0,72

Table XXXVII. Familial resemblance of XIRs: familial correlation and heritability estimates by ANOVA and SOLAR statistical methods trait analysis included i) DS and AS_{DS} scores (magnitude of Xi skewing) and ii) PAmat and AS_{PAmat} scores (magnitude and direction of skewing). a) Deviation from random Xi (DS and PAmat scores). For the primary / BC XIR, although the magnitude of skewing (DS score) was not significantly related among siblings, direction of skewing (PAmat score) was, implying an X-linked genetic component. A significant h^2 coefficient supports the latter. Among hematopoietic lineages, both DS and PAmat scores were significantly related within families (ANOVA results) and demonstrated significant heritability coefficients (SOLAR results), consistent with an X-linked genetic component to derivation of blood XIRs. T cells interestingly, demonstrated the strongest familial resemblance for PAmat scores, implying a strong X-linked genetic component. b) AS trait (AS_{DS} and AS_{PAmat} scores). With the exception of the PMN AS_{DS} score, all hematopoietic lineages demonstrated significant resemblance of AS_{DS} and AS_{PAmat} scores within families. This was complemented by modest heritability coefficients, consistent with an X-linked genetic component in derivation of AS scores. An extremely significant p-value for T cell AS_{PAmat} scores and a strong heritability coefficient imply a strong X-linked genetic component to derivation of T cell AS.

a) Xi skewing trait: deviation from random Xi

	BC *p-value*	**PMN** *p-value*	**Monocytes** *p-value*	**T cells** *p-value*	**B cells** *p-value*
DS					
ANOVA: p-value	0,242	$2,9 \times 10^{-4}$	$2,8 \times 10^{-6}$	$5,2 \times 10^{-5}$	$1,3 \times 10^{-7}$
SOLAR: p-value (h^2)	0,18 (0,035)	0,005 (0,11)	4×10^{-6} (0,21)	4×10^{-6} (0,20)	3×10^{-7} (0,24)
PAmat					
ANOVA	$3,3 \times 10^{-9}$	$1,1 \times 10^{-5}$	$2,7 \times 10^{-5}$	$1,6 \times 10^{-12}$	$7,3 \times 10^{-6}$
SOLAR: p-value (h^2)	1×10^{-7} (0,30)	$1,1 \times 10^{-5}$ (0,22)	5×10^{-5} (0,20)	1×10^{-7} (0,385)	$9,2 \times 10^{-7}$ (0,20)

b) AS trait: deviation from the primary Xi pattern

	PMN *p-value*	**Monocytes** *p-value*	**T cells** *p-value*	**B cells** *p-value*
AS_{DS}				
ANOVA: p-value	$1,5 \times 10^{-4}$	$5,2 \times 10^{-3}$	$2,1 \times 10^{-4}$	$1,5 \times 10^{-5}$
SOLAR: p-value (h^2)	0,058 (0,069)	$1,2 \times 10^{-5}$(0,20)	$5,6 \times 10^{-5}$ (0,17)	1×10^{-7} (0,30)
AS_{PAmat}				
ANOVA: p-value	$8,2 \times 10^{-6}$	$7,8 \times 10^{-6}$	$3,1 \times 10^{-11}$	$3,1 \times 10^{-5}$
SOLAR: p-value (h^2)	$6,3 \times 10^{-6}$ (0,23)	6×10^{-5} (0,23)	1×10^{-7} (0,36)	7×10^{-5} (0,20)

Table XXXVIII. Segregation analysis of skewed Xi patterns. a) skewed Xi pattern in BC tissue (DS score ≥ 0,25), b) skewed Xi pattern in PMN (DS score ≥ 0,25), c) AS trait in PMN (AS_{DS} score ≥ 0,25). Families where at least one sibling demonstrating a skewed Xi pattern (X preferentially inactivated) and where parental derivation (Xp versus Xm) was known were included in the tables below. 'Family ID' refers to family number; 'Number skewed' refers to the number of Xi skewed siblings per family, 'X skewed (observed)' refers to the number of siblings demonstrating a skewed Xi pattern for a parental specific X (Xp versus Xm); 'Direction of skewing (observed)' refers to the same values as in column 'X skewed (observed)', however only when n≥2 siblings demonstrated a skewed Xi pattern were included. Further, preferential inactivation of a parental-specific X was redressed as 'same' (i.e., versus the index case, the number of siblings Xi skewed in favor of the same parental X) and 'different' (i.e., versus the index case, the number of siblings Xi skewed in favor of the alternate parental X); 'Direction of skewing (expected)' refers to the expected number of siblings demonstrating a skewed Xi pattern in favor of the same parental X (same) versus number of individuals expected to inactivate the alternate parental X (different) if skewed Xi patterns were determined exclusively by stochastic processes.

a) skewed Xi pattern in BC tissue (DS score ≥ 0,25)

Family ID	Number skewed	X skewed (observed)		Direction of skewing (observed)		Direction of skewing (expected)	
		Xm	Xp	Same	Different	Same	Different
2	¼	1	0				
4	1/6	1	0				
9	¼	1	0				
10	1/7	0	1				
14	2/5	1	1	1	1	1	1
15	2/8	2	0	2	0	1	1
17	1/6	1	0				
26	1/10	1	0				
27	1/5	0	1				
29	¼	1	0				
33	2/7	2	0	2	0	1	1
41	1/5	0	1				
44	1/6	0	1				
47	1/6	1	0				
50	1/3	1	0				
53	2/4	2	2	2	0	1	1
59	1/8	1	0				
64	¼	0	1				
68	2/9	2	0	2	0	1	1
71	2/9	1	1	1	1	1	1
73	¼	1	0				
78	2/6	0	2	2	0	1	1
79	2/5	2	0	2	0	1	1
80	1/5	1	0				
82	1/5	1	0				
83	1/5	0	1				
87	¼	1	0				
91	¼	1	0				
94	1/3	0	1				
96	2/5	1	1	1	1	1	1

98	1/3	1	1				
99	2/5	2	0	2	0	1	1
100	2/3	0	2	2	0	1	1
102	¼	1	0				
103	2/5	2	0	1	1	1	1
104	3/5	3	0	3	0	2	1
107	1/8	1	0				
108	¼	1	0				
113	1/6	1	0				
114	¼	1	0				
115	1/1	1	0				
120	2/4	2	0	2	0	1	1
121	5/8	5	0	5	0	3	2
122	1/6	0	1				
123	3/6	0	3	3	0	2	1
124	¼	1	0				
125	1/6	0	1				
128	¼	1	0				
131	¼	0	1				
134	2/4	0	2	2	0	1	1
136	1/6	1	0				
137	¼	0	1				
141	1/5	0	1				
142	¼	1	0				
143	1/5	1	0				
144	1/5	1	0				
145	3/6	1	2	2	1	2	1
147	¼	0	1				
148	¼	0	1				
149	1/7	1	0				
152	2/5	1	1	1	1	1	1
153	2/6	2	0	2	0	1	1
154	¼	0	1				
155	2/5	2	0	2	0	1	1
160	1/5	0	1				
162	¼	0	1				
168	2/3	0	2	2	0	1	1
169	1/5	0	1				
170	1/6	0	1				
172	2/7	0	2	2	0	1	1
174	1/8	0	1				
175	1/5	1	0				
176	1/7	0	1				
178	1/6	1	0				
185	¼	1	0				
187	1/6	0	1				
188	1/5	1	0				
189	1/5	1	0				
191	3/7	0	3	3	0	2	1
194	1/7	0	1				
201	½	0	1				
204	1/11	1	0				
205	1/6	1	0				
207	2/6	1	1	1	1	1	1
213	1/8	1	0				
215	1/6	1	0				
216	¼	0	1				
217	2/6	0	2	2	0	1	1

218	1/5	1	0				
220	1/9	1	0				
222	3/9	2	1	2	1	2	1
91	126/492	74	52	54	8	34	28

b) Skewed Xi pattern in PMN (DS score ≥ 0,25)

Family ID	Number skewed	X skewed (observed)		Direction of skewing (observed)		Direction of skewing (expected)	
		Xm	Xp	Same	Different	Same	Different
2	2/4	2	0	2	0	1	1
3	¾	2	1	2	1	2	1
4	4/6	4	0	4	0	2	2
5	2/4	1	1	1	1	1	1
6	2/4	1	1	1	1	1	1
7	2/3	0	2	2	0	1	1
9	¼	1	0				
10	2/7	1	1	1	1	1	1
12	¼	0	1				
14	4/5	4	0	4	0	2	2
15	3/8	2	1	2	1	2	1
16	1/3	0	1				
17	1/6	1	0				
20	¼	0	1				
21	¼	1	0				
24	¾	0	3	3	0	2	1
25	1/3	1	0				
26	6/10	2	4	4	2	3	3
27	2/5	1	1	1	1	1	1
29	2/4	2	0	2	0	1	1
31	1/6	1	0				
33	1/7	1	0				
35	2/5	2	0	2	0	1	1
36	¾	1	2	2	1	2	1
38	1/1	1	0				
41	2/5	0	2	2	0	1	1
42	3/5	0	3	3	0	2	1
43	3/6	2	1	2	1	2	1
44	5/6	4	1	4	1	3	2
46	1/5	1	0				
50	2/3	1	1	1	1	1	1
53	¾	2	1	2	1	2	1
54	1/3	1	0				
56	2/6	2	0	2	0	1	1
57	¼	0	1				
58	4/5	2	2	2	2	2	2
59	4/8	3	1	3	1	2	2
61	1/8	0	1				
62	1/5	1	0				
63	4/7	3	1	3	1	2	2
64	4/4	2		2	2	2	2
65	1/3	0	2				
66	1/6	0	1				
67	2/5	1	1	1	1	1	1
68	3/9	2	1	2	1	2	1
69	1/6	0	1				
70	4/6	3	1	3	1	2	2

71	4/9	4	1	4	0	2	2
72	1/5	1	0				
73	¼	1	0				
74	¼	0	1				
76	2/4	2	0	2	0	1	1
77	¼	0	1				
78	2/6	0	2	2	0	1	1
79	3/5	2	1	2	1	2	1
80	2/5	2	0	2	0	1	1
81	¾	2	1	2	1	2	1
82	4/5	3	1	3	1	2	2
83	4/5	1	3	3	1	2	2
91	4/4	4	0	4	0	2	2
92	2/5	0	2	2	0	1	1
93	4/6	3	1	3	1	2	2
94	2/3	1	1	1	1	1	1
95	1/6	0	1				
96	3/5	1	2	2	1	2	1
97	¼	1	0				
98	2/3	1	1	1	1	1	1
99	1/5	1	0				
100	2/3	0	2	2	0	1	1
101	2/4	1	1	1	1	1	1
102	¼	1	0				
103	2/5	2	0	2	0	1	1
104	5/5	5	0	5	0	3	2
105	3/7	0	3	3	0	2	1
107	2/8	2	0	2	0	1	1
108	¼	1	0				
111	4/4	2	2	2	2	2	2
113	2/6	0	2	2	0	1	1
114	2/4	2	0	2	0	1	1
115	1/1	1	0				
116	2/8	1	1	1	1	1	1
117	2/5	2	0	2	0	1	1
118	1/6	0	1				
120	2/4	2	0	2	0	1	1
121	3/8	3	0	3	0	2	1
122	2/6	0	2	2	0	1	1
123	3/6	0	3	3	0	2	1
124	4/4	3	1	3	1	2	2
125	2/6	1	1	1	1	1	1
127	4/6	3	1	3	1	2	2
128	1/5	1	0				
129	2/8	1	1	1	1	1	1
131	2/4	1	0	2	0	1	1
133	2/4	1	1	1	1	1	1
136	1/6	1	0				
137	2/4	2	1	1	1	1	1
138	2/6	0	2	2	0	1	1
140	2/5	2	0	2	0	1	1
141	1/5	1	0				
142	2/2	2	0	2	0	1	1
143	2/5	2	0	2	0	1	1
144	1/5	1	0				
145	5/6	1	4	4	1	3	2
147	2/4	1	1	1	1	1	1
148	2/4	1	1	1	1	1	1

149	1/7	0	1				
150	¾	0	3	3	0	2	1
152	3/5	2	1	2	1	2	1
153	2/6	2	0	2	0	1	1
154	¼	0	1				
155	3/5	3	0	3	0	2	1
158	2/3	2	0	2	0	1	1
160	3/5	1	2	2	1	2	1
163	2/5	2	0	2	0	1	1
166	3/5	3	0	2	0	1	1
169	5/5	0	5	5	0	3	2
170	4/6	2	2	2	2	2	2
172	3/7	0	3	3	0	2	1
173	4/9	3	1	3	1	2	2
174	2/8	1	1	2	0	1	1
175	2/5	0	2	2	0	1	1
176	3/7	1	2	2	1	2	1
178	4/6	4	0	4	0	2	2
181	2/7	1	1	1	1	1	1
184	2/4	1	1	1	1	1	1
186	5/9	2	3	3	2	3	2
187	6/6	3	3	3	3	3	3
188	3/5	0	3	3	0	2	1
189	2/5	1	1	1	1	1	1
191	4/7	1	3	3	1	2	2
194	2/7	1	1	1	1	1	1
195	1/5	1	0				
200	3/8	0	3	3	0	2	1
202	1/5	0	1				
204	3/11	3	0	3	0	2	1
205	3/6	2	1	2	1	2	1
209	1/1	1	0				
212	1/5	1	0				
215	2/6	2	0	2	0	1	1
216	2/4	0	2	2	0	1	1
217	1/6	0	1				
218	1/5	1	0				
220	4/9	4	0	4	0	2	2
222	5/9	4	1	4	1	3	2
144	338/758	197	141	235	61	163	133

c) AS trait in PMN (AS_{DS} score ≥ 0,25)

Family ID	Number skewed	X skewed (observed)		Direction of skewing (observed)		Direction of skewing (expected)	
		Xm	Xp	Same	Different	Same	Different
2	1/1	0	1				
3	¾	2	1	2	1	2	1
4	1/6	1	0				
5	¼	0	1				
6	2/4	1	1	1	1	1	1
10	1/7	0	1				
12	¼	0	1				
14	4/5	4	0	4	0	2	2
15	3/8	0	3	3	0	2	1
16	1/3	0	1				
17	1/6	1	0				

21	¼	1	0				
25	1/3	0	1				
26	4/10	1	3	3	1	2	2
29	¼	1	0				
31	2/6	1	1	1	1	1	1
35	2/5	2	0	2	0	1	1
42	2/5	0	2	2	0	1	1
43	1/6	1	0				
44	2/6	2	0	2	0	1	1
46	½	0	1				
47	1/6	1	0				
50	1/3	0	1				
53	¼	0	1				
58	4/5	2	2	2	2	2	2
59	3/8	2	1	2	1	2	1
63	2/7	1	1	1	1	1	1
64	2/4	1	1	1	1	1	1
66	1/6	0	1				
67	2/5	1	1	1	1	1	1
68	2/9	1	1	1	1	1	1
71	2/9	2	0	2	0	1	1
72	2/5	2	0	2	0	1	1
74	¼	0	1				
75	1/5	0	1				
76	¼	1	0				
80	1/3	1	0				
81	¾	2	1	2	1	2	1
82	1/5	0	1				
83	3/5	1	2	2	1	2	1
91	¼	1	0				
93	2/6	1	1	1	1	1	1
94	1/3	0	1				
95	1/6	0	1				
96	1/5	0	1				
97	¼	1	0				
98	2/3	1	1	1	1	1	1
101	¼	0	1				
104	1/5	1	0				
105	3/7	0	3	3	0	2	1
107	1/8	1	0				
113	2/6	0	2	2	0	1	1
114	¼	1	0				
115	3/3	0	3	3	0	2	1
116	2/8	0	2	2	0	1	1
117	2/5	2	0	2	0	1	1
118	1/6	0	1				
120	¼	0	1				
121	2/8	0	2	2	0	1	1
123	2/6	2	0	2	0	1	1
124	2/4	1	1	1	1	1	1
127	2/6	1	1	1	1	1	1
129	2/8	1	1	1	1	1	1
131	2/4	2	0	2	0	1	1
133	2/4	1	1	1	1	1	1
137	¼	1	0				
138	1/6	0	1				
140	2/5	2	0	2	0	1	1
141	1/5	0	1				

142	2/4	1	1	1	1	1	1
143	2/5	2	0	2	0	1	1
144	1/5	1	0				
145	3/6	0	3	3	0	2	1
147	¼	1	0				
149	1/7	0	1				
150	¾	0	3	3	0	2	1
152	2/5	1	1	1	1	1	1
155	1/5	0	1				
158	1/3	1	0				
160	3/5	2	1	2	1	2	1
162	¼	1	0				
163	1/5	1	0				
166	3/5	3	0	3	0	2	1
168	2/3	2	0	2	0	1	1
169	4/5	0	4	4	0	2	2
173	4/9	3	1	3	1	2	2
174	5/8	5	0	5	0	3	2
175	3/5	0	3	3	0	2	1
176	4/7	2	2	2	2	2	2
178	3/6	3	0	3	0	2	1
181	2/7	0	2	2	0	1	1
185	¼	0	1				
186	2/9	1	1	1	1	1	1
187	3/6	1	2	2	1	2	1
188	3/5	0	3	3	0	2	1
195	1/1	0	1				
200	2/8	0	2	2	0	1	1
201	½	1	0				
202	1/5	0	1				
204	2/11	2	0	2	0	1	1
205	1/6	0	1				
207	1/6	1	0				
209	1/1	1	0				
213	2/2	2	0	2	0	1	1
214	1/6	0	1				
215	3/6	1	2	2	1	2	1
216	¼	0	1				
217	2/6	2	0	2	0	1	1
220	1/9	1	0				
222	3/9	1	2	2	1	2	1
110	200/571	99	101	119	29	83	65

Chapter V

CONCLUSION, PERSPECTIVES AND FUTURE DIRECTIONS

5.1 Conclusion

To further elucidate the etiologies and characterization of Xi phenotypes in human females, a cohort population of French Canadian females were recruited for study analysis. This was a monumental work necessitating recruitment of 193 families (1144 females) from various regions of the province of Québec Canada, the largest study of its kind. Because analysis of Xi phenotypes in human females is confounded by two unrelated phenotypes: the primary Xi trait versus the secondary (or acquired) Xi pattern of blood cells, methodologies to effectively distinguish and quantitate the two were developed. Careful analysis of the literature suggested that Xi patterns of NHTs were generally body-wide and relatively resistant to secondary factors of skewing (X-linked disease alleles). Consequently, the Xi pattern of a NHT (i.e., BC tissue) was used to further characterize the primary Xi trait and used as a control tissue to quantitate AS. Because a higher incidence of skewing occurs in PB versus NHTs, we speculated that the acquired from of skewing (i.e., AS) represents deviation from the primary Xi trait. Thus methodologies were derived to further characterize and quantitate the AS trait. In fact it was quantitated by measuring deviation from the primary (BC) XIR. This is the first study to our knowledge utilizing such a measure of the AS trait.

Primary Xi trait

12,4% of females demonstrated a skewed Xi pattern in BC tissue, a low incidence of skewing similar to that found in cord blood of neonates. Age- DS score analysis of BC tissue suggested that the incidence of skewing was relatively stable between the ages 38-96 yo, implying a stable Xi trait. Convincing support for a primary (body-wide) Xi pattern was derived by significant intraindividual correlation of PAmat scores between BC tissue and hematopoietic lineages. Lack of complete correlation however may be explained by the AS trait occurring in hematopoietic.

Genetic (X-linked) component to variability of BC XIRs

Because the X-linked *Xce* locus in mice plays a role in derivation of the primary Xi pattern, we speculated a similar mechanism may be occurring in humans. That PAmat scores of BC tissue was significantly similar between siblings (h^2=0,30) and that the X of same parental derivation (i.e., Xp versus Xm) was preferentially inactivated among siblings is consistent with an X-linked genetic component (possibly XCE-like) influencing the primary Xi pattern. If so we speculate genetic linkage of skewed Xi patterns in BC tissue to the XIC region. These findings provide clear and rational support to map the trait(s) influencing variability of the primary Xi pattern, plausibly identifying the long sought XCE locus and/or factors implicated in the choice step of Xi.

Clinical findings associated with the BC XIR

The primary Xi pattern (BC DS score) was associated with an increased reporting of asthma / COPD, suggesting a primary skewed Xi pattern may pose a risk factor for asthma / COPD susceptibility. A possible mechanism may be upregulation of X-linked asthma /COPD susceptibility alleles as a consequence of a primary skewed pattern.

AS trait

Great undertakings were initiated to characterize the AS trait as a function of independent hematopoietic lineages, the first such extensive study of its kind. Two sets of variables were used to quantitate the AS trait of hematopoietic lineages: i) deviation from random Xi (DS and PAmat scores) and ii) deviation from the primary Xi pattern (AS_{DS} and AS_{PAmat} scores). Versus BC tissue, a higher incidence of skewing was observed in hematopoietic lineages. And because DS score-age analyses found the incidence of skewing of hematopoietic lineages to increase with age, the higher incidence of skewing was attributed to the AS trait. Utilizing both sets of variables to measure AS, a higher incidence of skewing was observed in PMN, monocytes and B cells versus T lymphocytes. The higher incidence of skewing in PMN versus T cells was previously reported and may reflect longevity differences between short-lived myeloid versus long-lived T cells and/or other etiologies.

Hematopoietic stem cell origin of AS

Supporting evidence for a HSC origin of AS was derived by strong intraindividual correlation of PAmat and AS_{PAmat} scores among PMN, monocytes and B cells.

Genetic (X-linked) component to variability of AS values

Significant correlation of PAmat and AS_{PAmat} scores among siblings and significant concordance in direction of AS (i.e., preferential inactivation of a parental-specific X) among siblings, with h^2 coefficients ranging from $0,39 \geq h^2 \geq 0,20$ is consistent with both genetic (plausibly X-linked) and stochastic (possibly environmental factors such as cigarette smoke) influencing the AS trait, consistent with a multifactorial trait. Evidence of a genetic factor provides clear and rational support to map the AS gene(s). X-linked candidate genes possibly include any gene that can influence HSC kinetics: growth factors, transcription factors, intracellular transduction molecules, chemokines and related receptors. Alternatively, the entire X chromosome may be scanned for genetic linkage.

Hematologic associations of the AS trait

Although AS was not associated with variability of total WBC count, AS was associated with variability of particular lineages. i) Increasing DS scores in hematopoietic lineages were associated

with increasing hemoglobin concentration, suggesting AS may reflect selection of X-linked allele(s) coding for low hemoglobin expression, a candidate gene could be *PIG-A*. ii) Increasing DṠ and AS_{DS} scores were associated with decreasing eosinophil counts, suggesting a gene coding for variability in eosinophil numbers is X-linked and that the AS trait may reflect selection of X-linked alleles coding for low eosinophil numbers. Since eosinophils are implicated in all-cause mortality, the finding suggests AS may be associated with increased longevity.

Clinical associations of AS

AS (DS and AS_{DS} scores) was associated with a decreased reporting of RA, suggesting that the X chromosome may code for a RA susceptibility allele (consistent with published data). The finding may suggest that the AS trait reflects selection of X-linked RA protective allele(s).

Because the AS trait was associated with a decreased reporting of miscarriage occurrence, in addition to decreased eosinophil count and decreased reporting of RA, a role for AS in decreased autoimmune activity (or reduced disease susceptibility) is insinuated. Genetic identification of the blood skewing/AS gene(s) may provide elucidation of genetic determinant(s) for these disorders, possibly identifying immune-modulating gene(s). Whether a single gene of pleiotropic activity and/or whether genetic heterogeneity is responsible for the various clinical features of AS merits investigation. If several genes are implicated, sub-phenotypes of skewed Xi patterns based on clinical features will assist in the genetic linkage studies, since individuals with a more common phenotype are more likely to share a common genetic predisposition. If limited to a single gene however, the finding is consistent with one of immune-modulation, thus affecting several hematopoietic and immunological functions. An X-linked transcription factor controlling the expression of various immune/hematopoietic related genes (which need not be X-linked) is plausible. The mechanism however by which this X-linked allele(s) is favored as a function of aging remains elusive. Consistent with the X-linked allele competitive growth model of AS, one possibility may be that the X-linked protective disease allele(s) confers an intrinsic growth advantage to HSC. Alternatively, extrinsic factors such as cigarette smoke and/or the BM microenvironment may dictate survival kinetics of HSCs expressing particular X-linked allele(s). Other extrinsic factors which may be considered for further investigation include: bioactive compounds (hormones, growth factors and cytokines), infectious agents, vaccines, diet, medicinal products, pollutants and lifestyle factors (stress, psychological temperament and physical activity). A final possibility may be that, in part, AS reflects survival of females bearing X-linked disease-protective allele(s) .

5.2 Perspectives and Future Directions

Aside from the various 'additional works merited / warranted' mentioned in the Results section, additional analyses should include the following.

Since the DS score of BCs slightly decreased with advancing age, analysis of younger females may provide further insight as to whether a significant decrease occurs with advancing age. For a more powerful analysis, XIR-age analysis may be analyzed as a cross-sectional study of various age groups.

Because cigarette smoke may be an environmental factor in the etiology of AS, in vitro studies to identify the biologically active constituent(s) and/or the biological metabolite(s) altered in response to cigarette smoke is warranted. However, as cigarette smoke contains over 3800 chemicals, many biologically active (Vineis and Caporaso, 1995), identifying the active compound(s) appears overwhelmingly tedious.

Segregation analysis of the HUMARA locus among families where the mother demonstrated a skewed Xi pattern (DS score $\geq 0{,}25$) in the PMN fraction (n=16) found that the preferentially inactive X was preferentially transmitted to daughters more often than expected by chance. That is, of the 94 female offspring of these mothers, 59 received the HUMARA allele of the preferentially inactive X versus 35 the preferentially active X (59 of 94 meioses, $p=0{,}013$). This finding implies a transmission-ratio distortion (TRD) of the maternal Xs. A similar finding was observed among male offspring of skewed mothers (Naumova et al., 1995). Possible mechanisms include meiotic drive (biased segregation during meiosis), gametic selection (differential success of gamates achieving fertilization) and postzygotic survival selection based on particular genotypes, with the latter seemingly playing an important role (Zollner S et al., 2004). Regardless of the mechanism, our finding suggests that the two maternal Xs in oocytes of skewed mothers are not equal. One X (the one that is preferentially inactive in maternal tissue) demonstrates better survival kinetics when paired with the paternally derived X versus the X that is preferentially active in maternal tissues. This finding insinuates that a gene(s) related to fertility resides on the X chromosome, consistent with published data (Lanasa and Hogge, 2000), (Di Pasquale E et al., 2004). Paradoxically, it appears that while the X-linked fertility allele provides a relative survival advantage to developing embryos, it provides a relative growth disadvantage to adult HSCs. Nonetheless, it implicates the AS gene in both embryo survival kinetics and HSC growth kinetics.

Since blood skewing / AS may be associated with altered expression of immune related gene(s) (cytokines, growth factors, associated receptors and transcription factors), expression analysis of these factors in skewed versus non-skewed, by enzyme-linked immunoabsorbant assay, may provide pivotal evidence.

In the event that power to detect linkage is insufficient, and providing additional families are available, an efficient method to recruit additional families is to analyze a single sibling per family. If skewed (XIR score ≥ 0,25), the rest of the family could be recruited. On the other hand, if a random XIR is found, the family could be dropped from enrollment. This is based on our findings that familial aggregation is consistent with a genetic contribution. Although a biased approach, it is efficient for genetic linkage analyses.

To help determine whether RA is a cause, a consequence or mutually exclusive with the AS trait, a prospective study is merited. Ideally, a case (skewed) – control (non-skewed) group of females (matched for age) without RA could be monitored for disease onset as a function of aging. However, that the prevalence of both traits increase concurrently with advancing age, analysis may prove inconclusive.

Because the HUMARA locus is reportedly associated with various diseases / conditions, and that allele sizes have been determined for 90,8% of our study participants, the relationship between allele size and self-reported diseases could be investigated. Moreover, and assuming i) that the HUMARA locus is a candidate gene for the skewing/AS trait and ii) that different alleles of the CAG_n polymorphism in exon 1 represent a gradient of alleles - each conferring a relative but different growth advantage to HSC (akin to the relation between HUMARA allele length and androgen receptor activity in prostate cancer) - we hypothesize that the magnitude of skewing (DS and AS_{DS} scores) may be associated with HUMARA allele lengths. For example, XIRs could be analyzed as a function of the difference (allele 1 – allele 2) in HUMARA allele sizes.

That the incidence of the AS trait reported in our Québec population was similar to that observed in European populations (35-56%) (Gale et al., 1997), (Tonon et al., 1998), (Christensen et al., 2000) suggests a commonly found trait, ruling out founder effects to explain the high frequency of skewed Xi patterns. Alternatively, the wide geographic range of AS argues in favor of a heterozygote advantage. Nonetheless, as various regions of the province of Québec demonstrate characteristics of a founder population, particularly the region of Saguenay, Lac-St-Jean, the incidence of skewed Xi could be analyzed as a function of region, allowing potential detection of a founder effect.

The heterozygote advantage model may be invoked to account for the high frequency of skewed Xi by several models. i) Assuming a genetic basis to variability of the AS trait, one method is to analyze reproductive fitness measures among carriers of the trait allele versus non-carriers. In fact, for the primary Xi pattern, no increased reproductive fitness was associated with the BC Xi pattern, thus no link can be made with the primary Xi trait. Alternatively, increasing AS_{DS} scores were associated with reduced reporting of miscarriage occurrence. However, since the AS trait is late-onset, thus unlikely to reflect the Xi pattern at time of parturition, and assuming that time of miscarriage occurrence is the same between females with a low versus a high AS_{DS} score, the finding implies that women whom have incurred a miscarriage are less likely to undergo AS. Since the AS trait may reduce female susceptibility to RA and is associated with reduced eosinophil count, invokes a link between AS and reduced autoimmune complications. As such, although elderly females themselves may not be reproductively active, that they are relatively healthier may imply better caring of their children / grandchildren. Consequently, and speculatively, better conditions promoting active reproductive choices for their children. Consequently, AS susceptibility allele(s) are selectively passed on to future generations. Further data corroborating this model can be invoked by analyzing other measures of improved reproductive fitness and measures of improved health in females demonstrating the AS trait versus non-AS females. ii) A further method to support the heterozygote advantage model is to identify factors associated with the AS susceptibility allele(s) which increases the carrier's probability of attaining age of reproductive capability. Consistent with a role of the immune system, such factors could include increased resistance to childhood diseases. iii) Genetic data to corroborate the heterozygote advantage model includes 1) estimation of allele age, 2) number of trait-associated alleles and 3) inferred geographic origin of alleles (Risch N et al., 2003).

Because an association between XIRs and diseases with a putative X-linked genetic component was observed (i.e., RA, asthma, miscarriages), investigation of other disease/syndromes with an X-linked genetic component is warranted. Further, assuming diseases demonstrating a female preponderance have an X-linked genetic component, the investigation of female preponderant diseases (osteoporosis and Alzheimer's disease (Gao et al., 1998) for example) as a function of skewing is justified. These studies may provide further elucidation in the role of skewed Xi in female health.

As a genetic (X-linked) component to variance of XIRs was presented, segregation analysis may be used to evaluate whether a major gene contributes to variability of the phenotype.

Reference List

Abkowitz,J.L., Catlin,S.N., and Guttorp,P. (1996). Evidence that hematopoiesis may be a stochastic process in vivo. Nat. Med. *2* , 190-197.

Abkowitz,J.L., Ott,R.M., Holly,R.D., and Adamson,J.W. (1988). Clonal evolution following chemotherapy-induced stem cell depletion in cats heterozygous for glucose-6-phosphate dehydrogenase. Blood *71*, 1687-1692.

Abkowitz,J.L., Persik,M.T., Shelton,G.H., Ott,R.L., Kiklevich,J.V., Catlin,S.N., and Guttorp,P. (1995). Behavior of hematopoietic stem cells in a large animal. Proc. Natl. Acad. Sci. U. S. A *92*, 2031-2035.

Abkowitz,J.L., Taboada,M., Shelton,G.H., Catlin,S.N., Guttorp,P., and Kiklevich,J.V. (1998). An X chromosome gene regulates hematopoietic stem cell kinetics. Proc. Natl. Acad. Sci. U. S. A *95*, 3862-3866.

Abrahamson,G., Boultwood,J., Madden,J., Kelly,S., Oscier,D.G., Rack,K., Buckle,V.J., and Wainscoat,J.S. (1991). Clonality of cell populations in refractory anaemia using combined approach of gene loss and X-linked restriction fragment length polymorphism-methylation analyses. Br. J. Haematol. *79*, 550-555.

Achille,A., Scarpa,A., Montresor,M., Scardoni,M., Zamboni,G., Chilosi,M., Capelli,P., Franzin,G., and Menestrina,F. (1995). Routine application of polymerase chain reaction in the diagnosis of monoclonality of B-cell lymphoid proliferations. Diagn. Mol. Pathol. *4*, 14-24.

Adamson,J.W., Fialkow,P.J., Murphy,S., Prchal,J.F., and Steinmann,L. (1976). Polycythemia vera: stem-cell and probable clonal origin of the disease. N. Engl. J. Med. *295*, 913-916.

Adler,D.A., Rugarli,E.I., Lingenfelter,P.A., Tsuchiya,K., Poslinski,D., Liggitt,H.D., Chapman,V.M., Elliott,R.W., Ballabio,A., and Disteche,C.M. (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. U. S. A *94*, 9244-9248.

Albertini,R.J. and DeMars,R. (1974). Mosaicism of peripheral blood lymphocyte populations in females heterozygous for the Lesch-Nyhan mutation. Biochem. Genet. *11*, 397-411.

Allen,R.C., Nachtman,R.G., Rosenblatt,H.M., and Belmont,J.W. (1994). Application of carrier testing to genetic counseling for X-linked agammaglobulinemia. Am. J. Hum. Genet. *54*, 25-35.

Allen,R.C., Zoghbi,H.Y., Moseley,A.B., Rosenblatt,H.M., and Belmont,J.W. (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*, 1229-1239.

Almasy,L. and Blangero,J. (1998). Multipoint quantitative-trait linkage analysis in general pedigrees. Am. J. Hum. Genet. *62*, 1198-1211.

Anderson,C.L. and Brown,C.J. (1999). Polymorphic X-chromosome inactivation of the human TIMP1 gene. Am. J. Hum. Genet. *65*, 699-708.

Anger,B., Janssen,J.W., Schrezenmeier,H., Hehlmann,R., Heimpel,H., and Bartram,C.R. (1990). Clonal analysis of chronic myeloproliferative disorders using X-linked DNA polymorphisms. Leukemia *4*, 258-261.

Ansell,J.D., Samuel,K., Whittingham,D.G., Patek,C.E., Hardy,K., Handyside,A.H., Jones,K.W., Muggleton-Harris,A.L., Taylor,A.H., and Hooper,M.L. (1991). Hypoxanthine phosphoribosyl transferase deficiency, haematopoiesis and fertility in the mouse. Development *112*, 489-498.

Aral,B., de Saint,B.G., Al Garawi,S., Kamoun,P., and Ceballos-Picot,I. (1996). Novel nonsense mutation in the hypoxanthine guanine phosphoribosyltransferase gene and nonrandom X-inactivation causing Lesch-Nyhan syndrome in a female patient. Hum. Mutat. *7*, 52-58.

Araten,D.J., Bessler,M., McKenzie,S., Castro-Malaspina,H., Childs,B.H., Boulad,F., Karadimitris,A., Notaro,R., and Luzzatto,L. (2002). Dynamics of hematopoiesis in paroxysmal nocturnal hemoglobinuria (PNH): no evidence for intrinsic growth advantage of PNH clones. Leukemia *16*, 2243-2248.

Araten,D.J., Nafa,K., Pakdeesuwan,K., and Luzzatto,L. (1999). Clonal populations of hematopoietic cells with paroxysmal nocturnal hemoglobinuria genotype and phenotype are present in normal individuals. Proc. Natl. Acad. Sci. U. S. A *96*, 5209-5214.

Aruffo,A., Hollenbaugh,D., Wu,L.H., and Ochs,H.D. (1994). The molecular basis of X-linked agammaglobulinemia, hyper-IgM syndrome, and severe combined immunodeficiency in humans. Curr. Opin. Hematol. *1*, 12-18.

Asimakopoulos,F.A., Gilbert,J.G., Aldred,M.A., Pearson,T.C., and Green,A.R. (1996). Interstitial deletion constitutes the major mechanism for loss of heterozygosity on chromosome 20q in polycythemia vera. Blood *88*, 2690-2698.

Asplund,A., Guo,Z., Hu,X., Wassberg,C., and Ponten,F. (2001). Mosaic pattern of maternal and paternal keratinocyte clones in normal human epidermis revealed by analysis of X-chromosome inactivation. J. Invest Dermatol. *117*, 128-131.

Avner,P. and Heard,E. (2001). X-chromosome inactivation: counting, choice and initiation. Nat. Rev. Genet. *2*, 59-67.

Avner,P., Prissette,M., Arnaud,D., Courtier,B., Cecchi,C., and Heard,E. (1998). Molecular correlates of the murine Xce locus. Genet. Res *72*, 217-224.

Azofeifa,J., Voit,T., Hubner,C., and Cremer,M. (1995). X-chromosome methylation in manifesting and healthy carriers of dystrophinopathies: concordance of activation ratios among first degree female relatives and skewed inactivation as cause of the affected phenotypes. Hum. Genet. *96*, 167-176.

Azofeifa,J., Waldherr,R., and 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.

Bagnara,G.P., Bonsi,L., Strippoli,P., Bonifazi,F., Tonelli,R., D'Addato,S., Paganelli,R., Scala,E., Fagiolo,U., Monti,D., Cossarizza,A., Bonafe,M., and Franceschi,C. (2000). Hemopoiesis in healthy old people and

centenarians: well-maintained responsiveness of CD34+ cells to hemopoietic growth factors and remodeling of cytokine network. J. Gerontol. A Biol Sci. Med. Sci. *55*, B61-B66.

Bailey,J.A., Carrel,L., Chakravarti,A., and Eichler,E.E. (2000). From the cover: molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis [see comments]. Proc. Natl. Acad. Sci. U. S. A *97*, 6634-6639.

Bamba,M., Sugihara,H., Okada,K., Bamba,T., and Hattori,T. (1998). Clonal analysis of superficial depressed-type gastric carcinoma in humans. Cancer *83*, 867-875.

Bamforth,F., Machin,G., and 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.

Bandmann,H.J., Breit,R., and Perwein,E. (1984). Klinefelter's Syndrome., B.Bandmann, ed. (Heidelberg: Springer-Verlag Berlin).

Barr,M.L. and Bertram,E.G. (1949). A morphological distinction between neurones of the male and female, and the behavioor of the nucleolar satellite during accelerated necleoprotein synthesis. Nature *163*, 676.

Beard,C., Li,E., and Jaenisch,R. (1995). Loss of methylation activates Xist in somatic but not in embryonic cells. Genes Dev. *9* , 2325-2334.

Belickova,M., Schroeder,H.W., Jr., Guan,Y.L., Brierre,J., Berney,S., Cooper,M.D., and Prchal,J.T. (1994). Clonal hematopoiesis and acquired thalassemia in common variable immunodeficiency. Mol. Med. *1*, 56-61.

Belmont,J.W. (1995). Insights into lymphocyte development from X-linked immune deficiencies. Trends Genet. *11*, 112-116.

Belmont,J.W. (1996). Genetic control of X inactivation and processes leading to X- inactivation skewing. Am. J. Hum. Genet. *58*, 1101-1108.

Beutler,E., Yeh,M., and Fairbanks,V.F. (1962). The normal human female as a mosaic of X-chromosome activity: Studies using the gene gor G-6-PD deficieny as a marker. Proc Natl Acad Sci USA *48*, 9.

Beutler,E. and Kuhl,W. (1990). The NT 1311 polymorphism of G6PD: G6PD Mediterranean mutation may have originated independently in Europe and Asia. Am. J. Hum. Genet. *47*, 1008-1012.

Beutler,E., Vulliamy,T., and Luzzatto,L. (1996). Hematologically important mutations: glucose-6-phosphate dehydrogenase. Blood Cells Mol. Dis. *22*, 49-56.

Bird,A.P. (1986). CpG-rich islands and the function of DNA methylation. Nature *321*, 209-213.

Bjork,J., Albin,M., Mauritzson,N., Stromberg,U., Johansson,B., and Hagmar,L. (2000). Smoking and myelodysplastic syndromes. Epidemiology *11*, 285-291.

Blann,A.D., Kirkpatrick,U., Devine,C., Naser,S., and McCollum,C.N. (1998). The influence of acute smoking on leucocytes, platelets and the endothelium. Atherosclerosis *141*, 133-139.

Boggs,B.A. and Chinault,A.C. (1994). Analysis of replication timing properties of human X-chromosomal loci by fluorescence in situ hybridization. Proc. Natl. Acad. Sci. U. S. A *91*, 6083-6087.

Boggs,B.A., Connors,B., Sobel,R.E., Chinault,A.C., and Allis,C.D. (1996). Reduced levels of histone H3 acetylation on the inactive X chromosome in human females. Chromosoma *105*, 303-309.

Borsani,G., Tonlorenzi,R., Simmler,M.C., Dandolo,L., Arnaud,D., Capra,V., Grompe,M., Pizzuti,A., Muzny,D., and Lawrence,C. (1991). Characterization of a murine gene expressed from the inactive X chromosome. Nature *351*, 325-329.

Boumil,R.M. and Lee,J.T. (2001). Forty years of decoding the silence in X-chromosome inactivation. Hum. Mol. Genet. *10*, 2225-2232.

Bousquet,J., Chanez,P., Lacoste,J.Y., Barneon,G., Ghavanian,N., Enander,I., Venge,P., Ahlstedt,S., Simony-Lafontaine,J., Godard,P., and . (1990). Eosinophilic inflammation in asthma. N. Engl. J. Med. *323*, 1033-1039.

Boyd,Y. and Fraser,N.J. (1990). Methylation patterns at the hypervariable X-chromosome locus DXS255 (M27 beta): correlation with X-inactivation status. Genomics *7*, 182-187.

Brewer,G.J., Gall,J.C., Honeyman,M., Gershowitz,H., Shreffler,D.C., Dern,R.J., and Hames,C. (1967). Inheritance of quantitative expression of erythrocyte glucose-6- phosphate dehydrogenase activity in the negro--a twin study. Biochem. Genet. *1*, 41-53.

Brockdorff,N., Ashworth,A., Kay,G.F., Cooper,P., Smith,S., McCabe,V.M., Norris,D.P., Penny,G.D., Patel,D., and 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,G.F., McCabe,V.M., Norris,D.P., Cooper,P.J., Swift,S., and Rastan,S. (1992). The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. Cell *71*, 515-526.

Brown,C.J., Ballabio,A., Rupert,J.L., Lafreniere,R.G., Grompe,M., Tonlorenzi,R., and Willard,H.F. (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,C.J., Hendrich,B.D., Rupert,J.L., Lafreniere,R.G., Xing,Y., Lawrence,J., and Willard,H.F. (1992). The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. Cell *71*, 527-542.

Brown,C.J., Lafreniere,R.G., Powers,V.E., Sebastio,G., Ballabio,A., Pettigrew,A.L., Ledbetter,D.H., Levy,E., Craig,I.W., and Willard,H.F. (1991a). Localization of the X inactivation centre on the human X chromosome in Xq13. Nature *349*, 82-84.

Brown,C.J. and Robinson,W.P. (2000). The causes and consequences of random and non-random X chromosome inactivation in humans. Clin. Genet. *58*, 353-363.

Brown,C.J. and Willard,H.F. (1994). The human X-inactivation centre is not required for maintenance of X-chromosome inactivation. Nature *368*, 154-156.

Brown,R.M. and Brown,G.K. (1993). X chromosome inactivation and the diagnosis of X linked disease in females. J. Med. Genet. *30*, 177-184.

Brown,S. and Rastan,S. (1988). Age-related reactivation of an X-linked gene close to the inactivation centre in the mouse. Genet. Res *52*, 151-154.

Buller,R.E., Sood,A.K., Lallas,T., Buekers,T., and Skilling,J.S. (1999). Association between nonrandom X-chromosome inactivation and BRCA1 mutation in germline DNA of patients with ovarian cancer [see comments]. J. Natl. Cancer Inst. *91*, 339-346.

Busque,L. and Gilliland,D.G. (1998). X-inactivation analysis in the 1990s: promise and potential problems. Leukemia *12*, 128-135.

Busque,L., Mio,R., Mattioli,J., Brais,E., Blais,N., Lalonde,Y., Maragh,M., and Gilliland,D.G. (1996). Nonrandom X-inactivation patterns in normal females: lyonization ratios vary with age. Blood *88*, 59-65.

Busque,L., Zhu,J., DeHart,D., Griffith,B., Willman,C., Carroll,R., Black,P.M., and Gilliland,D.G. (1994). An expression based clonality assay at the human androgen receptor locus (HUMARA) on chromosome X. Nucleic Acids Res. *22*, 697-698.

Buzin,C.H., Mann,J.R., and 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.

Carr,D.H. (1971). Chromosome studies in selected spontaneous abortions. Polyploidy in man. J. Med. Genet. *8*, 164-174.

Carrel,L., Cottle,A.A., Goglin,K.C., and Willard,H.F. (1999). A first-generation X-inactivation profile of the human X chromosome [see comments]. Proc. Natl. Acad. Sci. U. S. A *96*, 14440-14444.

Carrel,L. and Willard,H.F. (1996). An assay for X inactivation based on differential methylation at the fragile X locus, FMR1. Am. J. Med. Genet. *64*, 27-30.

Carrel,L. and Willard,H.F. (1999). Heterogeneous gene expression from the inactive X chromosome: an X-linked gene that escapes X inactivation in some human cell lines but is inactivated in others. Proc. Natl. Acad. Sci. U. S. A *96*, 7364-7369.

Catlin,S.N., Abkowitz,J.L., and Guttorp,P. (1997). Hematopoiesis as competitive exclusion process. Blood *abstract 3352*.

Cattanach,B.M. and Isaacson,J.H. (1967). Controlling elements in the mouse X chromosome. Genetics *57*, 331-346.

Cattanach,B.M. and Papworth,D. (1981). Controlling elements in the mouse. V. Linkage tests with X-linked genes. Genet. Res *38*, 57-70.

Cattanach,B.M. and Perez,J.N. (1970). Parental influence on X-autosome translocation-induced variegation in the mouse. Genet. Res. *15*, 43-53.

Cattanach,B.M., Pollard,C.E., and Perez,J.N. (1969). Controlling elements in the mouse X-chromosome. I. Interaction with the X-linked genes. Genet. Res. *14*, 223-235.

Cattanach,B.M., Rasberry,C., Evans,E.P., Dandolo,L., Simmler,M.C., and Avner,P. (1991). Genetic and molecular evidence of an X-chromosome deletion spanning the tabby (Ta) and testicular feminization (Tfm) loci in the mouse. Cytogenet. Cell Genet. *56*, 137-143.

Cattanach,B.M. and Williams,C.E. (1972). Evidence of non-random X chromosome activity in the mouse. Genet. Res. *19*, 229-240.

Cazzola,M., May,A., Bergamaschi,G., Cerani,P., Rosti,V., and Bishop,D.F. (2000). Familial-skewed X-chromosome inactivation as a predisposing factor for late-onset X-linked sideroblastic anemia in carrier females. Blood *96*, 4363-4365.

Cazzola,M., Sacchi,F., Pagani,A., Marconi,M., Ciriello,M.M., Fietta,A., Clivio,A., and Ascari,E. (1985). X-linked chronic granulomatous disease in an adult woman. Evidence for a cell selection favoring neutrophils expressing the mutant allele. Haematologica *70*, 291-295.

Champion,K.M., Gilbert,J.G., Asimakopoulos,F.A., Hinshelwood,S., and Green,A.R. (1997). Clonal haemopoiesis in normal elderly women: implications for the myeloproliferative disorders and myelodysplastic syndromes. Br. J. Haematol. *97*, 920-926.

Chang,C.S., Kokontis,J., and Liao,S.T. (1988). Molecular cloning of human and rat complementary DNA encoding androgen receptors. Science *240*, 324-326.

Chang,H.W., Chong,S.M., Peh,S.C., and Lee,S.H. (1999). X-chromosome inactivation analysis of isolated Reed-Sternberg cells in nodular sclerosing Hodgkin's disease. Br. J. Haematol. *107*, 641-647.

Chao,W., Huynh,K.D., Spencer,R.J., Davidow,L.S., and Lee,J.T. (2002). CTCF, a candidate trans-acting factor for X-inactivation choice. Science *295*, 345-347.

Chen,T.C., Kuo,T., and Chan,H.L. (2000). Dermatofibroma is a clonal proliferative disease. J. Cutan. Pathol. *27*, 36-39.

Cheng,P.C., Gosewehr,J.A., Kim,T.M., Velicescu,M., Wan,M., Zheng,J., Felix,J.C., Cofer,K.F., Luo,P., Biela,B.H., Godorov,G., and Dubeau,L. (1996). Potential role of the inactivated X chromosome in ovarian epithelial tumor development. J. Natl. Cancer Inst. *88*, 510-518.

Cheshier,S.H., Morrison,S.J., Liao,X., and Weissman,I.L. (1999). In vivo proliferation and cell cycle kinetics of long-term self- renewing hematopoietic stem cells. Proc. Natl. Acad. Sci. U. S. A *96*, 3120-3125.

Chetritt,J., Paradis,V., Dargere,D., Adle-Biassette,H., Maurage,C.A., Mussini,J.M., Vital,A., Wechsler,J., and Bedossa,P. (1999). Chester-Erdheim disease: a neoplastic disorder. Hum. Pathol. *30*, 1093-1096.

Chitnis,S., Monteiro,J., Glass,D., Apatoff,B., Salmon,J., Concannon,P., and Gregersen,P.K. (2000). The role of X-chromosome inactivation in female predisposition to autoimmunity. Arthritis Res *2*, 399-406.

Christensen,K., Kristiansen,M., Hagen-Larsen,H., Skytthe,A., Bathum,L., Jeune,B., Andersen-Ranberg,K., Vaupel,J.W., and Orstavik,K.H. (2000). X-linked genetic factors regulate hematopoietic stem-cell kinetics in females. Blood *95*, 2449-2451.

Clemson,C.M., Chow,J.C., Brown,C.J., and Lawrence,J.B. (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.

Clemson,C.M., McNeil,J.A., Willard,H.F., and Lawrence,J.B. (1996). XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. J. Cell Biol. *132*, 259-275.

Conley,M.E., Brown,P., Pickard,A.R., Buckley,R.H., Miller,D.S., Raskind,W.H., Singer,J.W., and Fialkow,P.J. (1986). Expression of the gene defect in X-linked agammaglobulinemia. N. Engl. J. Med. *315*, 564-567.

Cooper,G.S., Dooley,M.A., Treadwell,E.L., St Clair,E.W., Parks,C.G., and Gilkeson,G.S. (1998). Hormonal, environmental, and infectious risk factors for developing systemic lupus erythematosus. Arthritis Rheum. *41*, 1714-1724.

Corre,F., Lellouch,J., and Schwartz,D. (1971). Smoking and leucocyte-counts. Results of an epidemiological survey. Lancet *2*, 632-634.

Costanzi,C. and Pehrson,J.R. (1998). Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. Nature *393*, 599-601.

Costanzi,C., Stein,P., Worrad,D.M., Schultz,R.M., and Pehrson,J.R. (2000). Histone macroH2A1 is concentrated in the inactive X chromosome of female preimplantation mouse embryos. Development *127*, 2283-2289.

Courtier,B., Heard,E., and Avner,P. (1995). Xce haplotypes show modified methylation in a region of the active X chromosome lying 3' to Xist. Proc. Natl. Acad. Sci. U. S. A *92*, 3531-3535.

Csankovszki,G., Panning,B., Bates,B., Pehrson,J.R., and Jaenisch,R. (1999). Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X inactivation [letter]. Nat. Genet. *22*, 323-324.

Curnutte,J.T., Hopkins,P.J., Kuhl,W., and Beutler,E. (1992). Studying X inactivation. Lancet *339*, 749.

Cutolo,M. and Masi,A.T. (1998). Do androgens influence the pathophysiology of rheumatoid arthritis? Facts and hypotheses. J. Rheumatol. *25*, 1041-1047.

D'Adda,T., Candidus,S., Denk,H., Bordi,C., and Hofler,H. (1999). Gastric neuroendocrine neoplasms: tumour clonality and malignancy- associated large X-chromosomal deletions. J. Pathol. *189*, 394-401.

Daniels,R., Zuccotti,M., Kinis,T., Serhal,P., and Monk,M. (1997b). XIST expression in human oocytes and preimplantation embryos. Am. J. Hum. Genet. *61*, 33-39.

Daniels,R., Zuccotti,M., Kinis,T., Serhal,P., and Monk,M. (1997a). XIST expression in human oocytes and preimplantation embryos [see comments]. Am. J. Hum. Genet. *61*, 33-39.

Davidson,R.G., Nithowsky,H.M., and Childs,B. (1963). Demonstration of two populations of cells in the human female heterozygous for glucose-6-phosphate dehdrogenase variant. Proc. Natl. Acad. Sci. U. S. A *50* , 481.

de Labry,L.O., Campion,E.W., Glynn,R.J., and Vokonas,P.S. (1990). White blood cell count as a predictor of mortality: results over 18 years from the Normative Aging Study. J. Clin. Epidemiol. *43*, 153-157.

De Monchy,J.G., Kauffman,H.F., Venge,P., Koeter,G.H., Jansen,H.M., Sluiter,H.J., and De Vries,K. (1985). Bronchoalveolar eosinophilia during allergen-induced late asthmatic reactions. Am. Rev. Respir. Dis. *131*, 373-376.

Debrand,E., Heard,E., and Avner,P. (1998). Cloning and localization of the murine Xpct gene: evidence for complex rearrangements during the evolution of the region around the Xist gene. Genomics *48*, 296-303.

Deighton,C.M. and Walker,D.J. (1991). The familial nature of rheumatoid arthritis. Ann. Rheum. Dis. *50*, 62-65.

Delabesse,E., Oksenhendler,E., Lebbe,C., Verola,O., Varet,B., and Turhan,A.G. (1997). Molecular analysis of clonality in Kaposi's sarcoma. J. Clin. Pathol. *50*, 664-668.

Delage,R., Jacques,L., Massinga-Loembe,M., Poulin,J., Bilodeau,D., Mignault,C., Leblond,P.F., and Darveau,A. (2001). Persistent polyclonal B-cell lymphocytosis: further evidence for a genetic disorder associated with B-cell abnormalities. Br. J. Haematol. *114*, 666-670.

Desmarais,D., Zhong,Y., Chakraborty,R., Perreault,C., and Busque,L. (1998). Development of a highly polymorphic STR marker for identity testing purposes at the human androgen receptor gene (HUMARA). J. Forensic Sci. *43*, 1046-1049.

Devriendt,K., Matthijs,G., Legius,E., Schollen,E., Blockmans,D., van Geet,C., Degreef,H., Cassiman,J.J., and Fryns,J.P. (1997). Skewed X-chromosome inactivation in female carriers of dyskeratosis congenita. Am. J. Hum. Genet. *60*, 581-587.

Di Pasquale E, Beck-Peccoz P, and Persani L. (2004). Hypergonadotropic ovarian failure associated with an inherited mutation of human bone morphogenetic protein-15 (BMP15) gene. Am J Hum Genet *75*, 106-111.

Disteche,C.M. (1999). Escapees on the X chromosome. Proc. Natl. Acad. Sci. U. S. A *96*, 14180-14182.

Disteche,C.M., Eicher,E.M., and Latt,S.A. (1981). Late replication patterns in adult and embryonic mice carrying Searle's X-autosome translocation. Exp Cell Res *133*, 357-362.

Donner,A. and Eliasziw,M. (1991). Methodology for inferences concerning familial correlations: a review. J. Clin. Epidemiol. *44* , 449-455.

Dover,G.J., Smith,K.D., Chang,Y.C., Purvis,S., Mays,A., Meyers,D.A., Sheils,C., and Serjeant,G. (1992). Fetal hemoglobin levels in sickle cell disease and normal individuals are partially controlled by an X-linked gene located at Xp22.2. Blood *80*, 816-824.

Dow,L.W., Martin,P., Moohr,J., Greenberg,M., Macdougall,L.G., Najfeld,V., and Fialkow,P.J. (1985). Evidence for clonal development of childhood acute lymphoblastic leukemia. Blood *66*, 902-907.

Driscoll,D.J. and Migeon,B.R. (1990). Sex difference in methylation of single-copy genes in human meiotic germ cells: implications for X chromosome inactivation, parental imprinting, and origin of CpG mutations. Somat. Cell Mol. Genet. *16*, 267-282.

Dyer,K.A., Canfield,T.K., and Gartler,S.M. (1989). Molecular cytological differentiation of active from inactive X domains in interphase: implications for X chromosome inactivation. Cytogenet. Cell Genet. *50* , 116-120.

Edwards,A., Hammond,H.A., Jin,L., Caskey,C.T., and Chakraborty,R. (1992). Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics *12*, 241-253.

Eggan,K., Akutsu,H., Hochedlinger,K., Rideout,W., III, Yanagimachi,R., and Jaenisch,R. (2000). X-Chromosome inactivation in cloned mouse embryos. Science *290*, 1578-1581.

Egusa,Y., Fujiwara,Y., Syahruddin,E., Isobe,T., and Yamakido,M. (1998). Effect of age on human peripheral blood stem cells. Oncol. Rep. *5*, 397-400.

el Kassar,N., Hetet,G., Briere,J., and Grandchamp,B. (1997). Clonality analysis of hematopoiesis in essential thrombocythemia: advantages of studying T lymphocytes and platelets. Blood *89*, 128-134.

el Kassar,N., Hetet,G., Li,Y., Briere,J., and Grandchamp,B. (1995). Clonal analysis of haemopoietic cells in essential thrombocythaemia. Br. J. Haematol. *90*, 131-137.

Endo,Y., Sugimura,H., and Kino,I. (1995). Monoclonality of normal human colonic crypts. Pathol. Int. *45*, 602-604.

Enomoto,T., Haba,T., Fujita,M., Hamada,T., Yoshino,K., Nakashima,R., Wada,H., Kurachi,H., Wakasa,K., Sakurai,M., Murata,Y., and Shroyer,K.R. (1997). Clonal analysis of high-grade squamous intra-epithelial lesions of the uterine cervix. Int. J. Cancer *73*, 339-344.

Epstein,C.J. (1972). Expression of the mammalian X chromosome before and after fertilization. Science *175*, 1467-1468.

Epstein,C.J., Smith,S., Travis,B., and Tucker,G. (1978). Both X chromosomes function before visible X-chromosome inactivation in female mouse embryos. Nature *274*, 500-503.

Evans,D.M., Frazer,I.H., and Martin,N.G. (1999). Genetic and environmental causes of variation in basal levels of blood cells. Twin. Res *2*, 250-257.

Favier,R., Lavergne,J.M., Costa,J.M., Caron,C., Mazurier,C., Viemont,M., Delpech,M., and Valleix,S. (2000). Unbalanced X-chromosome inactivation with a novel FVIII gene mutation resulting in severe hemophilia A in a female. Blood *96*, 4373-4375.

Fearon,E.R., Kohn,D.B., Winkelstein,J.A., Vogelstein,B., and Blaese,R.M. (1988). Carrier detection in the Wiskott Aldrich syndrome. Blood *72*, 1735-1739.

Fey,M.F., Liechti-Gallati,S., von Rohr,A., Borisch,B., Theilkas,L., Schneider,V., Oestreicher,M., Nagel,S., Ziemiecki,A., and Tobler,A. (1994). Clonality and X-inactivation patterns in hematopoietic cell populations detected by the highly informative M27 beta DNA probe [see comments]. Blood *83*, 931-938.

Fey,M.F., Peter,H.J., Hinds,H.L., Zimmermann,A., Liechti-Gallati,S., Gerber,H., Studer,H., and Tobler,A. (1992). Clonal analysis of human tumors with M27 beta, a highly informative polymorphic X chromosomal probe. J. Clin. Invest *89*, 1438-1444.

Fialkow,P.J. (1973). Primordial cell pool size and lineage relationships of five human cell types. Ann. Hum. Genet. *37*, 39-48.

Fialkow,P.J., Faguet,G.B., Jacobson,R.J., Vaidya,K., and Murphy,S. (1981). Evidence that essential thrombocythemia is a clonal disorder with origin in a multipotent stem cell. Blood *58*, 916-919.

Fialkow,P.J., Gartler,S.M., and Yoshida,A. (1967). Clonal origin of chronic myelocytic leukemia in man. Proc. Natl. Acad. Sci. U. S. A *58*, 1468-1471.

Filosa,S., Giacometti,N., Wangwei,C., De Mattia,D., Pagnini,D., Alfinito,F., Schettini,F., Luzzatto,L., and Martini,G. (1996). Somatic-cell selection is a major determinant of the blood-cell phenotype in heterozygotes for glucose-6-phosphate dehydrogenase mutations causing severe enzyme deficiency. Am. J. Hum. Genet. *59*, 887-895.

Fujita,M.Q., Hashida,N., Shin,M., Nakanishi,H., Yoshihara,W., and Aozasa,K. (1998). Clonal composition of malignant fibrous histiocytoma: analysis by PCR- based assay of the human androgen receptor gene (HUMARA). Oncology *55*, 600-606.

Gaetani,G.F., Ferraris,A.M., Galiano,S., Giuntini,P., Canepa,L., and D'Urso,M. (1982). Primary thrombocythemia: clonal origin of platelets, erythrocytes, and granulocytes in a GdB/GdMediterranean subject. Blood *59*, 76-79.

Gale,R.E., Fielding,A.K., Harrison,C.N., and Linch,D.C. (1997). Acquired skewing of X-chromosome inactivation patterns in myeloid cells of the elderly suggests stochastic clonal loss with age. Br. J. Haematol. *98*, 512-519.

Gale,R.E. and Linch,D.C. (1998). Clonality studies in acute myeloid leukemia. Leukemia *12*, 117-120.

Gale,R.E., Wheadon,H., Boulos,P., and Linch,D.C. (1994). Tissue specificity of X-chromosome inactivation patterns. Blood *83*, 2899-2905.

Gale,R.E., Wheadon,H., and Linch,D.C. (1991). X-chromosome inactivation patterns using HPRT and PGK polymorphisms in haematologically normal and post-chemotherapy females. Br. J. Haematol. *79*, 193-197.

Gale,R.E., Wheadon,H., and Linch,D.C. (1992). Assessment of X-chromosome inactivation patterns using the hypervariable probe M27 beta in normal hemopoietic cells and acute myeloid leukemic blasts. Leukemia *6* , 649-655.

Gandini,E. and Gartler,S.M. (1969). Glucose-6-phosphate dehydrogenase mosaicism for studying the development of blood cell precursors. Nature *224*, 599-600.

Gandini,E., Gartler,S.M., Angioni,G., Argiolas,N., and Dell'Acqua,G. (1968). Developmental implications of multiple tissue studies in glucose-6- phosphate dehydrogenase-deficient heterozygotes. Proc. Natl. Acad. Sci. U. S. A *61*, 945-948.

Ganesan,S., Silver,D.P., Greenberg,R.A., Avni,D., Drapkin,R., Miron,A., Mok,S.C., Randrianarison,V., Brodie,S., Salstrom,J., Rasmussen,T.P., Klimke,A., Marrese,C., Marahrens,Y., Deng,C.X., Feunteun,J., and Livingston,D.M. (2002). BRCA1 supports XIST RNA concentration on the inactive X chromosome. Cell *111*, 393-405.

Gao,S., Hendrie,H.C., Hall,K.S., and Hui,S. (1998). The relationships between age, sex, and the incidence of dementia and Alzheimer disease: a meta-analysis. Arch. Gen. Psychiatry *55*, 809-815.

Gardner,R.L. and Lyon,M.F. (1971). X chromosome inactivation studied by injection of a single cell into the mouse blastocyst. Nature *231*, 385-386.

Gartler,S.M., Dyer,K.A., and Goldman,M.A. (1992). Mammalian X chromosome inactivation. Mol. Genet. Med. *2*, 121-160.

Gartler,S.M., Gandini,E., Angioni,G., and Argiolas,N. (1969). Glucose-6 phosphate dehydrogenase mosaicism: utilization as a tracer in the study of the development of hair root cells. Ann. Hum. Genet. *33*, 171-176.

Gartler,S.M., Liskay,R.M., Campbell,B.K., Sparkes,R., and Gant,N. (1972). Evidence for two functional X chromosomes in human oocytes. Cell Differ. *1*, 215-218.

Gartler,S.M. and Riggs,A.D. (1983). Mammalian X-chromosome inactivation. Annu. Rev. Genet. *17*, 155-190.

Gartler,S.M., Rivest,M., and Cole,R.E. (1980). Cytological evidence for an inactive X chromosome in murine oogonia. Cytogenet. Cell Genet. *28*, 203-207.

Gauvreau,G.M., Wood,L.J., Sehmi,R., Watson,R.M., Dorman,S.C., Schleimer,R.P., Denburg,J.A., and O'Byrne,P.M. (2000). The effects of inhaled budesonide on circulating eosinophil progenitors and their expression of cytokines after allergen challenge in subjects with atopic asthma. Am. J. Respir. Crit Care Med. *162*, 2139-2144.

Giguere,Y., Dewailly,E., Brisson,J., Ayotte,P., Laflamme,N., Demers,A., Forest,V.I., Dodin,S., Robert,J., and Rousseau,F. (2001). Short polyglutamine tracts in the androgen receptor are protective against breast cancer in the general population. Cancer Res *61*, 5869-5874.

Gilbert,C.W., Muldol,S., Lajtha,L.G., and Rowley,J. (1962). Time-sequence of the human chromosome duplication. Nature (Lond.) *195*, 869.

Gilbert,S.L. and Sharp,P.A. (1999). Promoter-specific hypoacetylation of X-inactivated genes. Proc. Natl. Acad. Sci. U. S. A *96* , 13825-13830.

Gilleece,M.H. and Dexter,T.M. (1993). Effect of Campath-1H antibody on human hematopoietic progenitors in vitro. Blood *82*, 807-812.

Gilliland,D.G., Blanchard,K.L., and Bunn,H.F. (1991a). Clonality in acquired hematologic disorders. Annu. Rev. Med. *42*, 491-506.

Gilliland,D.G., Blanchard,K.L., Levy,J., Perrin,S., and Bunn,H.F. (1991b). Clonality in myeloproliferative disorders: analysis by means of the polymerase chain reaction. Proc. Natl. Acad. Sci. U. S. A *88*, 6848-6852.

Goldgar DE (1990). Multipoint analysis of human quantitative genetic variation. Am J Hum Genet *47*, 957-967.

Goodship,J., Carter,J., and Burn,J. (1996). X-inactivation patterns in monozygotic and dizygotic female twins. Am. J. Med. Genet. *61*, 205-208.

Goto,T., Wright,E., and Monk,M. (1997). Paternal X-chromosome inactivation in human trophoblastic cells. Mol. Hum. Reprod. *3*, 77-80.

Graves,J.A. (1982). 5-azacytidine-induced re-expression of alleles on the inactive X chromosome in a hybrid mouse cell line. Exp Cell Res *141*, 99-105.

Green,A.J., Sepp,T., and Yates,J.R. (1996). Clonality of tuberous sclerosis harmatomas shown by non-random X- chromosome inactivation. Hum. Genet. *97*, 240-243.

Gregg,X.T., Kralovics,R., and Prchal,J.T. (2000). A polymorphism of the X-linked gene IDS increases the number of females informative for transcriptional clonality assays. Am. J. Hematol. *63*, 184-191.

Grimm,R.H., Jr., Neaton,J.D., and Ludwig,W. (1985). Prognostic importance of the white blood cell count for coronary, cancer, and all-cause mortality. JAMA *254*, 1932-1937.

Haig,D. (1993). Genetic conflicts in human pregnancy. Q. Rev. Biol. *68*, 495-532.

Hall JM, Lee MK, Newman B, Morrow JE, Anderson LA, Huey B, and King MC (1990). Linkage of early-onset familial breast cancer to chromosome 17q21. Science *250*, 1684-1689.

Hall,M.A., Ahmadi,K.R., Norman,P., Snieder,H., MacGregor,A.J., Vaughan,R.W., Spector,T.D., and Lanchbury,J.S. (2000). Genetic influence on peripheral blood T lymphocyte levels. Genes Immun. *1*, 423-427.

Happle,R. (1985). Lyonization and the lines of Blaschko. Hum. Genet. *70*, 200-206.

Harada,M., Suzuki,M., Ikeda,T., Kaneko,T., Harada,S., and Fukayama,M. (1997). Clonality in nevocellular nevus and melanoma: an expression-based clonality analysis at the X-linked genes by polymerase chain reaction. J. Invest Dermatol. *109*, 656-660.

Harper,M.I., Fosten,M., and Monk,M. (1982). Preferential paternal X inactivation in extraembryonic tissues of early mouse embryos. J. Embryol. Exp. Morphol. *67*, 127-135.

Harris,A., Collins,J., Vetrie,D., Cole,C., and 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,C.N., Gale,R.E., and Linch,D.C. (1998). Quantification of X-chromosome inactivation patterns using RT-PCR of the polymorphic iduronate-2-sulphatase gene and correlation of the results obtained with DNA-based techniques. Leukemia *12*, 1834-1839.

Harrison,D.E. and Astle,C.M. (1982). Loss of stem cell repopulating ability upon transplantation. Effects of donor age, cell number, and transplantation procedure. J. Exp Med. *156*, 1767-1779.

Harrison,K.B. (1989). X-chromosome inactivation in the human cytotrophoblast. Cytogenet. Cell Genet. *52*, 37-41.

Harrison,K.B. and Warburton,D. (1986). Preferential X-chromosome activity in human female placental tissues. Cytogenet. Cell Genet. *41*, 163-168.

Heard,E., Clerc,P., and Avner,P. (1997). X-chromosome inactivation in mammals. Annu. Rev. Genet. *31*, 571-610.

Heard,E., Rougeulle,C., Arnaud,D., Avner,P., Allis,C.D., and Spector,D.L. (2001). Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X inactivation. Cell *107*, 727-738.

Hendriks,R.W., Chen,Z.Y., Hinds,H., Schuurman,R.K., and Craig,I.W. (1992). An X chromosome inactivation assay based on differential methylation of a CpG island coupled to a VNTR polymorphism at the 5' end of the monoamine oxidase A gene. Hum. Mol. Genet. *1*, 187-194.

Hirota,Y., Okamura,S., Kimura,N., Shibuya,T., and Niho,Y. (1988). Haematopoiesis in the aged as studied by in vitro colony assay. Eur. J. Haematol. *40*, 83-90.

Hitzeroth,H.W. and Bender,K. (1981). Age-dependency of somatic selection in South African Negro G-6-PD heterozygotes. Hum. Genet. *58*, 338-343.

Honda,K., Kashima,K., Daa,T., Yokoyama,S., and Nakayama,I. (2000). Clonal analysis of the epithelial component of Warthin's tumor. Hum. Pathol. *31*, 1377-1380.

Hong,Y.K., Ontiveros,S.D., Chen,C., and Strauss,W.M. (1999). A new structure for the murine Xist gene and its relationship to chromosome choice/counting during X-chromosome inactivation. Proc. Natl. Acad. Sci. U. S. A *96*, 6829-6834.

Hong,Y.K., Ontiveros,S.D., and Strauss,W.M. (2000). A revision of the human XIST gene organization and structural comparison with mouse Xist. Mamm. Genome *11*, 220-224.

Horny,H.P., Ruck,P., Krober,S., and Kaiserling,E. (1997). Systemic mast cell disease (mastocytosis). General aspects and histopathological diagnosis. Histol. Histopathol. *12*, 1081-1089.

Hospers,J.J., Schouten,J.P., Weiss,S.T., Postma,D.S., and Rijcken,B. (2000). Eosinophilia is associated with increased all-cause mortality after a follow-up of 30 years in a general population sample. Epidemiology *11*, 261-268.

Huebner,K., Isobe,M., Gasson,J.C., Golde,D.W., and Croce,C.M. (1986). Localization of the gene encoding human erythroid-potentiating activity to chromosome region Xp11.1----Xp11.4. Am. J. Hum. Genet. *38*, 819-826.

Hutchinson,D., Shepstone,L., Moots,R., Lear,J.T., and Lynch,M.P. (2001). Heavy cigarette smoking is strongly associated with rheumatoid arthritis (RA), particularly in patients without a family history of RA. Ann. Rheum. Dis. *60*, 223-227.

Ingerslev,J., Schwartz,M., Lamm,L.U., Kruse,T.A., Bukh,A., and Stenbjerg,S. (1989). Female haemophilia A in a family with seeming extreme bidirectional lyonization tendency: abnormal premature X-chromosome inactivation? Clin. Genet. *35*, 41-48.

Jacobs,P.A. and Migeon,B.R. (1989). Studies of X-chromosome inactivation in trisomies. Cytogenet. Cell Genet. *50*, 75-77.

Jacobson,R.J., Salo,A., and Fialkow,P.J. (1978). Agnogenic myeloid metaplasia: a clonal proliferation of hematopoietic stem cells with secondary myelofibrosis. Blood *51*, 189-194.

Janssen,J.W., Anger,B.R., Drexler,H.G., Bartram,C.R., and Heimpel,H. (1990). Essential thrombocythemia in two sisters originating from different stem cell levels. Blood *75*, 1633-1636.

Janssen,J.W., Buschle,M., Layton,M., Drexler,H.G., Lyons,J., van den,B.H., Heimpel,H., Kubanek,B., Kleihauer,E., and Mufti,G.J. (1989). Clonal analysis of myelodysplastic syndromes: evidence of multipotent stem cell origin. Blood *73*, 248-254.

Jawaheer,D., Seldin,M.F., Amos,C.I., Chen,W.V., Shigeta,R., Monteiro,J., Kern,M., Criswell,L.A., Albani,S., Nelson,J.L., Clegg,D.O., Pope,R., Schroeder,H.W., Jr., Bridges,S.L., Jr., Pisetsky,D.S., Ward,R., Kastner,D.L., Wilder,R.L., Pincus,T., Callahan,L.F., Flemming,D., Wener,M.H., and Gregersen,P.K. (2001). A genomewide screen in multiplex rheumatoid arthritis families suggests genetic overlap with other autoimmune diseases. Am. J. Hum. Genet. *68*, 927-936.

Jazaeri,A.A., Yee,C.J., Sotiriou,C., Brantley,K.R., Boyd,J., and Liu,E.T. (2002). Gene expression profiles of BRCA1-linked, BRCA2-linked, and sporadic ovarian cancers. J. Natl. Cancer Inst. *94*, 990-1000.

Jeppesen,P. and Turner,B.M. (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.

Jernigan,J.A., Gudat,J.C., Blake,J.L., Bowen,L., and Lezotte,D.C. (1980). Reference values for blood findings in relatively fit elderly persons. J. Am. Geriatr. Soc. *28*, 308-314.

Jimbo,H., Hitomi,Y., Yoshikawa,H., Yano,T., Momoeda,M., Sakamoto,A., Tsutsumi,O., Taketani,Y., and Esumi,H. (1997). Evidence for monoclonal expansion of epithelial cells in ovarian endometrial cysts. Am. J. Pathol. *150*, 1173-1178.

Jin,H., May,M., Tranebjaerg,L., Kendall,E., Fontan,G., Jackson,J., Subramony,S.H., Arena,F., Lubs,H., Smith,S., Stevenson,R., Schwartz,C., and Vetrie,D. (1996). A novel X-linked gene, DDP, shows mutations in families with deafness (DFN-1), dystonia, mental deficiency and blindness. Nat. Genet. *14*, 177-180.

Johnston,C.M., Nesterova,T.B., Formstone,E.J., Newall,A.E., Duthie,S.M., Sheardown,S.A., and Brockdorff,N. (1998). Developmentally regulated Xist promoter switch mediates initiation of X inactivation. Cell *94*, 809-817.

Johnston,P.G. and Cattanach,B.M. (1981). Controlling elements in the mouse. IV. Evidence of non-random X- inactivation. Genet. Res. *37*, 151-160.

Jorde,L.B. (1995). Linkage disequilibrium as a gene-mapping tool. Am. J. Hum. Genet. *56*, 11-14.

Kattar,M.M., Kupsky,W.J., Shimoyama,R.K., Vo,T.D., Olson,M.W., Bargar,G.R., and Sarkar,F.H. (1997). Clonal analysis of gliomas. Hum. Pathol. *28*, 1166-1179.

Kay H.E.M (1965). How many cell generations? Lancet *2*, 418.

Kay,G.F., Penny,G.D., Patel,D., Ashworth,A., Brockdorff,N., and Rastan,S. (1993). Expression of Xist during mouse development suggests a role in the initiation of X chromosome inactivation. Cell *72*, 171-182.

Keith,D.H., Singer-Sam,J., and Riggs,A.D. (1986). Active X chromosome DNA is unmethylated at eight CCGG sites clustered in a guanine-plus-cytosine-rich island at the 5' end of the gene for phosphoglycerate kinase. Mol. Cell Biol. *6*, 4122-4125.

Kelly,A. and Munan,L. (1977). Haematologic profile of natural populations: red cell parameters. Br. J. Haematol. *35*, 153-160.

Keohane,A.M., O'Neill,L.P., Belyaev,N.D., Lavender,J.S., and Turner,B.M. (1996). X-Inactivation and histone H4 acetylation in embryonic stem cells. Dev. Biol. *180*, 618-630.

Khalifa,M.M., Struthers,J.L., Maurice,S., Harrison,K., and Duncan,A.M. (2001). Methylation of HpaII site at the human DXS16 locus on Xp22 as an assay for abnormal patterns of X inactivation. Am. J. Med. Genet. *98*, 64-69.

Khoury,M.J., Beaty,T.H., Tockman,M.S., Self,S.G., and Cohen,B.H. (1985). Familial aggregation in chronic obstructive pulmonary disease: use of the loglinear model to analyze intermediate environmental and genetic risk factors. Genet. Epidemiol. *2*, 155-166.

Kondo,H., Kusaka,Y., and Morimoto,K. (1993). [Effects of lifestyle on hematologic parameters; I. Analysis of hematologic data in association with smoking habit and age]. Sangyo Igaku *35*, 98-104.

Kratzer,P.G. and Chapman,V.M. (1981). X chromosome reactivation in oocytes of Mus caroli. Proc. Natl. Acad. Sci. U. S. A *78*, 3093-3097.

Kratzer,P.G. and Gartler,S.M. (1978). HGPRT activity changes in preimplantation mouse embryos. Nature *274*, 503-504.

Kreipe,H., Jaquet,K., Felgner,J., Radzun,H.J., and Parwaresch,M.R. (1991). Clonal granulocytes and bone marrow cells in the cellular phase of agnogenic myeloid metaplasia. Blood *78*, 1814-1817.

Krietsch,W.K., Fehlau,M., Renner,P., Bucher,T., and Fundele,R. (1986). Expression of X-linked phosphoglycerate kinase in early mouse embryos homozygous at the Xce locus. Differentiation *31*, 50-54.

Kristiansen,M., Langerod,A., Knudsen,G.P., Weber,B.L., Borresen-Dale,A.L., and Orstavik,K.H. (2002). High frequency of skewed X inactivation in young breast cancer patients. J. Med. Genet. *39*, 30-33.

Kuiper,G.G., Faber,P.W., van Rooij,H.C., van der Korput,J.A., Ris-Stalpers,C., Klaassen,P., Trapman,J., and Brinkmann,A.O. (1989). Structural organization of the human androgen receptor gene. J. Mol. Endocrinol. *2*, R1-R4.

Kutsche,R. and Brown,C.J. (2000). Determination of X-chromosome inactivation status using X-linked expressed polymorphisms identified by database searching. Genomics *65*, 9-15.

La Spada,A.R., Wilson,E.M., Lubahn,D.B., Harding,A.E., and Fischbeck,K.H. (1991). Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. Nature *352*, 77-79.

Lafreniere,R.G., Brown,C.J., Rider,S., Chelly,J., Taillon-Miller,P., Chinault,A.C., Monaco,A.P., and Willard,H.F. (1993). 2.6 Mb YAC contig of the human X inactivation center region in Xq13: physical linkage of the RPS4X, PHKA1, XIST and DXS128E genes. Hum. Mol. Genet. *2*, 1105-1115.

Lanasa,M.C. and Hogge,W.A. (2000). X chromosome defects as an etiology of recurrent spontaneous abortion. Semin. Reprod. Med. *18*, 97-103.

Lanasa,M.C., Hogge,W.A., Kubik,C., Blancato,J., and Hoffman,E.P. (1999). Highly skewed X-chromosome inactivation is associated with idiopathic recurrent spontaneous abortion. Am. J. Hum. Genet. *65*, 252-254.

Landau,L.I. (2001). Parental smoking: asthma and wheezing illnesses in infants and children. Paediatr. Respir. Rev. *2*, 202-206.

Lange,K., Westlake,J., and Spence,M.A. (1976). Extensions to pedigree analysis. III. Variance components by the scoring method. Ann. Hum. Genet. *39*, 485-491.

Latham,K.E. (1999). Epigenetic modification and imprinting of the mammalian genome during development. Curr. Top. Dev. Biol. *43* , 1-49.

Lau,A.W., Brown,C.J., Penaherrera,M., Langlois,S., Kalousek,D.K., and Robinson,W.P. (1997). Skewed X-chromosome inactivation is common in fetuses or newborns associated with confined placental mosaicism. Am. J. Hum. Genet. *61*, 1353-1361.

Lawrence,H.J., Broudy,V.C., Magenis,R.E., Olson,S., Tomar,D., Barton,S., Fitchen,J.H., and Bagby,G.C., Jr. (1987). Cytogenetic evidence for involvement of B lymphocytes in acquired idiopathic sideroblastic anemias. Blood *70*, 1003-1005.

Lawrence,R.C., Hochberg,M.C., Kelsey,J.L., McDuffie,F.C., Medsger,T.A., Jr., Felts,W.R., and Shulman,L.E. (1989). Estimates of the prevalence of selected arthritic and musculoskeletal diseases in the United States. J. Rheumatol. *16*, 427-441.

Lee,J.T., Davidow,L.S., and Warshawsky,D. (1999b). Tsix, a gene antisense to Xist at the X-inactivation centre [see comments]. Nat. Genet. *21*, 400-404.

Lee,J.T. and Lu,N. (1999). Targeted mutagenesis of Tsix leads to nonrandom X inactivation. Cell *99*, 47-57.

Lee,J.T., Lu,N., and Han,Y. (1999a). Genetic analysis of the mouse X inactivation center defines an 80-kb multifunction domain. Proc. Natl. Acad. Sci. U. S. A *96*, 3836-3841.

Lee,J.T., Strauss,W.M., Dausman,J.A., and Jaenisch,R. (1996). A 450 kb transgene displays properties of the mammalian X-inactivation center. Cell *86*, 83-94.

Lee,P.S., Sabbath-Solitare,M., Redondo,T.C., and Ongcapin,E.H. (2000). Molecular evidence that the stromal and epithelial cells in pleomorphic adenomas of salivary gland arise from the same origin: clonal analysis using human androgen receptor gene (HUMARA) assay. Hum. Pathol. *31*, 498-503.

Lee,S.D., Shroyer,K.R., Markham,N.E., Cool,C.D., Voelkel,N.F., and Tuder,R.M. (1998). Monoclonal endothelial cell proliferation is present in primary but not secondary pulmonary hypertension. J. Clin. Invest *101*, 927-934.

Lee,S.T., McGlennen,R.C., and Litz,C.E. (1994). Clonal determination by the fragile X (FMR1) and phosphoglycerate kinase (PGK) genes in hematological malignancies. Cancer Res. *54*, 5212-5216.

Leppig,K.A., Brown,C.J., Bressler,S.L., Gustashaw,K., Pagon,R.A., Willard,H.F., and Disteche,C.M. (1993). Mapping of the distal boundary of the X-inactivation center in a rearranged X chromosome from a female expressing XIST. Hum. Mol. Genet. *2*, 883-887.

Lewis,R.A., Austen,K.F., and Soberman,R.J. (1990). Leukotrienes and other products of the 5-lipoxygenase pathway. Biochemistry and relation to pathobiology in human diseases. N. Engl. J. Med. *323*, 645-655.

Li,M. and Cannizzaro,L.A. (1999). Identical clonal origin of synchronous and metachronous low-grade, noninvasive papillary transitional cell carcinomas of the urinary tract. Hum. Pathol. *30*, 1197-1200.

Li,M., Cordon-Cardo,C., Gerald,W.L., and Rosai,J. (1996). Desmoid fibromatosis is a clonal process. Hum. Pathol. *27*, 939-943.

Liang,K.Y., Beaty,T.H., and Cohen,B.H. (1986). Application of odds ratio regression models for assessing familial aggregation from case-control studies. Am. J. Epidemiol. *124*, 678-683.

Lifschytz,E. and Lindsley,D.L. (1972). The role of X-chromosome inactivation during spermatogenesis (Drosophila-allocycly-chromosome evolution-male sterility-dosage compensation). Proc. Natl. Acad. Sci. U. S. A *69*, 182-186.

Linder,D. and Gartler,S.M. (1965). Glucose-6-phosphate dehydrogenase mosaicism: utilization as a cell marker in the study of leiomyomas. Science *150*, 67-69.

Lipschitz,D.A. (1999). The Effect of Age on Hematopoiesis and the Work-Up of Anemia in the Elderly. Hematology *1999*, 504-509.

Liu,Y., Phelan,J., Go,R.C., Prchal,J.F., and Prchal,J.T. (1997). Rapid determination of clonality by detection of two closely-linked X chromosome exonic polymorphisms using allele-specific PCR. J. Clin. Invest *99*, 1984-1990.

Lobaccaro,J.M., Lumbroso,S., Belon,C., Galtier-Dereure,F., Bringer,J., Lesimple,T., Namer,M., Cutuli,B.F., Pujol,H., and Sultan,C. (1993). Androgen receptor gene mutation in male breast cancer. Hum. Mol. Genet. *2*, 1799-1802.

Looijenga,L.H., Gillis,A.J., Verkerk,A.J., van Putten,W.L., and Oosterhuis,J.W. (1999). Heterogeneous X inactivation in trophoblastic cells of human full-term female placentas. Am. J. Hum. Genet. *64*, 1445-1452.

Lubahn,D.B., Joseph,D.R., Sullivan,P.M., Willard,H.F., French,F.S., and Wilson,E.M. (1988). Cloning of human androgen receptor complementary DNA and localization to the X chromosome. Science *240*, 327-330.

Lucas,D.R., Shroyer,K.R., McCarthy,P.J., Markham,N.E., Fujita,M., and Enomoto,T.E. (1997). Desmoid tumor is a clonal cellular proliferation: PCR amplification of HUMARA for analysis of patterns of X-chromosome inactivation. Am. J. Surg. Pathol. *21*, 306-311.

Luhovy,M., Liu,Y., Belickova,M., Prchal,J.F., and Prchal,J.T. (1995). A novel clonality assay based on transcriptional polymorphism of X chromosome gene p55. Biol. Blood Marrow Transplant. *1*, 81-87.

Luzzatto,L., Usanga,E.A., Bienzle,U., Esan,G.F., and Fusuan,F.A. (1979). Imbalance in X-chromosome expression: evidence for a human X-linked gene affecting growth of hemopoietic cells. Science *205*, 1418-1420.

Luzzatto,L., Usanga,F.A., and Reddy,S. (1969). Glucose-6-phosphate dehydrogenase deficient red cells: resistance to infection by malarial parasites. Science *164*, 839-842.

Lynch,K.R., O'Neill,G.P., Liu,Q., Im,D.S., Sawyer,N., Metters,K.M., Coulombe,N., Abramovitz,M., Figueroa,D.J., Zeng,Z., Connolly,B.M., Bai,C., Austin,C.P., Chateauneuf,A., Stocco,R., Greig,G.M., Kargman,S., Hooks,S.B., Hosfield,E., Williams,D.L., Jr., Ford-Hutchinson,A.W., Caskey,C.T., and Evans,J.F. (1999). Characterization of the human cysteinyl leukotriene CysLT1 receptor. Nature *399*, 789-793.

Lynn,A.H., Kwoh,C.K., Venglish,C.M., Aston,C.E., and Chakravarti,A. (1995). Genetic epidemiology of rheumatoid arthritis. Am. J. Hum. Genet. *57*, 150-159.

Lyon,M.F. (1960). J. Hered *51*, 116.

Lyon,M.F. (1961). Gene action in the X-chromosome of the mouse [*Mus musculus* L.]. Nature 190:372. Nature *190*, 372-373.

Lyon,M.F. (1962). Sex chromatin and gene action in the mammalian X-chromosome. Am. J. Hum. Genet. *14*, 135-148.

Lyon,M.F. (1998). X-chromosome inactivation: a repeat hypothesis. Cytogenet. Cell Genet. *80*, 133-137.

Lyon,M.F. and Rastan,S. (1984). Parental source of chromosome imprinting and its relevance for X chromosome inactivation. Differentiation *26*, 63-67.

MacGregor,A.J., Snieder,H., Rigby,A.S., Koskenvuo,M., Kaprio,J., Aho,K., and Silman,A.J. (2000). Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. Arthritis Rheum. *43*, 30-37.

Marahrens,Y., Panning,B., Dausman,J., Strauss,W., and Jaenisch,R. (1997). Xist-deficient mice are defective in dosage compensation but not spermatogenesis [see comments]. Genes Dev. *11*, 156-166.

Marcus,S., Steen,A.M., Andersson,B., Lambert,B., Kristoffersson,U., and Francke,U. (1992). Mutation analysis and prenatal diagnosis in a Lesch-Nyhan family showing non-random X-inactivation interfering with carrier detection tests. Hum. Genet. *89*, 395-400.

Martin-Villa,J.M., Corell,A., Ramos-Amador,J.T., Ruiz-Contreras,J., and Arnaiz-Villena,A. (1999). Higher incidence of autoantibodies in X-linked chronic granulomatous disease carriers: random X-chromosome inactivation may be related to autoimmunity. Autoimmunity *31*, 261-264.

Mashal,R.D., Fejzo,M.L., Friedman,A.J., Mitchner,N., Nowak,R.A., Rein,M.S., Morton,C.C., and Sklar,J. (1994). Analysis of androgen receptor DNA reveals the independent clonal origins of uterine leiomyomata and the secondary nature of cytogenetic aberrations in the development of leiomyomata. Genes Chromosomes. Cancer *11*, 1-6.

Mathioudakis,G., Storb,R., McSweeney,P.A., Torok-Storb,B., Lansdorp,P.M., Brummendorf,T.H., Gass,M.J., Bryant,E.M., Storek,J., Flowers,M.E., Gooley,T., and Nash,R.A. (2000). Polyclonal hematopoiesis with variable telomere shortening in human long-term allogeneic marrow graft recipients. Blood *96*, 3991-3994.

McDonald,J.A. and Kelley,W.N. (1972). Lesch-Nyhan syndrome: absence of the mutant enzyme in erythrocytes of a heterozygote for both normal and mutant hypoxanthine-guanine phosphoribosyl transferase. Biochem. Genet. *6*, 21-26.

McLaren,A. (1972). Numerology of development. Nature *239*, 274-276.

Mensinga,T.T., Schouten,J.P., Rijcken,B., Weiss,S.T., and van der,L.R. (1994). Host factors and environmental determinants associated with skin test reactivity and eosinophilia in a community-based population study. Ann. Epidemiol. *4*, 382-392.

Mermoud,J.E., Popova,B., Peters,A.H., Jenuwein,T., and Brockdorff,N. (2002). Histone H3 lysine 9 methylation occurs rapidly at the onset of random X chromosome inactivation. Curr. Biol *12*, 247-251.

Migeon,B.R. (1972). Stability of X chromosomal inactivation in human somatic cells. Nature *239*, 87-89.

Migeon,B.R. (1993). The postulated X-inactivation center at Xq27 is most reasonably explained by ascertainment bias: heterozygous expression of recessive mutations is a powerful means of detecting unbalanced X inactivation [letter; comment]. Am. J. Hum. Genet. *52*, 431-434.

Migeon,B.R. (1994). X-chromosome inactivation: molecular mechanisms and genetic consequences. Trends Genet. *10*, 230-235.

Migeon,B.R. (1998). Non-random X chromosome inactivation in mammalian cells. Cytogenet. Cell Genet. *80*, 142-148.

Migeon,B.R., Chowdhury,A.K., Dunston,J.A., and 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,B.R. and Do,T.T. (1979). In search of non-random X inactivation: studies of fetal membranes heterozygous for glucose-6-phosphate dehydrogenase. Am. J. Hum. Genet. *31*, 581-585.

Migeon,B.R., Luo,S., Stasiowski,B.A., Jani,M., Axelman,J., Van Dyke,D.L., Weiss,L., Jacobs,P.A., Yang-Feng,T.L., and Wiley,J.E. (1993). Deficient transcription of XIST from tiny ring X chromosomes in females with severe phenotypes. Proc. Natl. Acad. Sci. U. S. A *90*, 12025-12029.

Migeon,B.R., Wolf,S.F., Axelman,J., Kaslow,D.C., and Schmidt,M. (1985). Incomplete X chromosome dosage compensation in chorionic villi of human placenta. Proc. Natl. Acad. Sci. U. S. A *82*, 3390-3394.

Miller,A.P. and Willard,H.F. (1998). Chromosomal basis of X chromosome inactivation: identification of a multigene domain in Xp11.21-p11.22 that escapes X inactivation. Proc. Natl. Acad. Sci. U. S. A *95*, 8709-8714.

Mohandas,T., Sparkes,R.S., and Shapiro,L.J. (1981). Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation. Science *211*, 393-396.

Mohandas,T.K., Passage,M.B., Williams,J.W., III, Sparkes,R.S., Yen,P.H., and Shapiro,L.J. (1989). X-chromosome inactivation in cultured cells from human chorionic villi. Somat. Cell Mol. Genet. *15*, 131-136.

Monk,M., Boubelik,M., and Lehnert,S. (1987). Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. Development *99*, 371-382.

Monk,M. and Harper,M.I. (1979). Sequential X chromosome inactivation coupled with cellular differentiation in early mouse embryos. Nature *281*, 311-313.

Monk,M. and Kathuria,H. (1977). Dosage compensation for an X-linked gene in pre-implantation mouse embryos. Nature *270*, 599-601.

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

Monteiro,J., Derom,C., Vlietinck,R., Kohn,N., Lesser,M., and Gregersen,P.K. (1998). Commitment to X inactivation precedes the twinning event in monochorionic MZ twins. Am. J. Hum. Genet. *63*, 339-346.

Morrison,S.J., Wandycz,A.M., Akashi,K., Globerson,A., and Weissman,I.L. (1996). The aging of hematopoietic stem cells. Nat. Med. *2*, 1011-1016.

Moszczynski,P., Zabinski,Z., Moszczynski,P., Jr., Rutowski,J., Slowinski,S., and Tabarowski,Z. (2001). Immunological findings in cigarette smokers. Toxicol. Lett. *118*, 121-127.

Munafo,M., Johnstone,E., Murphy,M., and Walton,R. (2001). New directions in the genetic mechanisms underlying nicotine addiction. Addict. Biol *6*, 109-117.

Murry,C.E., Gipaya,C.T., Bartosek,T., Benditt,E.P., and Schwartz,S.M. (1997). Monoclonality of smooth muscle cells in human atherosclerosis. Am. J. Pathol. *151*, 697-705.

Mutter,G.L., Chaponot,M.L., and Fletcher,J.A. (1995). A polymerase chain reaction assay for non-random X chromosome inactivation identifies monoclonal endometrial cancers and precancers. Am. J. Pathol. *146*, 501-508.

Nash,R., Storb,R., and Neiman,P. (1988). Polyclonal reconstitution of human marrow after allogeneic bone marrow transplantation. Blood *72*, 2031-2037.

Naumova,A.K., Olien,L., Bird,L.M., Slamka,C., Fonseca,M., Verner,A.E., Wang,M., Leppert,M., Morgan,K., and Sapienza,C. (1995). Transmission-ratio distortion of X chromosomes among male offspring of females with skewed X-inactivation. Dev. Genet. *17*, 198-205.

Naumova,A.K., Olien,L., Bird,L.M., Smith,M., Verner,A.E., Leppert,M., Morgan,K., and 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.

Naumova,A.K., Plenge,R.M., Bird,L.M., Leppert,M., Morgan,K., Willard,H.F., and Sapienza,C. (1996b). Heritability of X chromosome--inactivation phenotype in a large family. Am. J. Hum. Genet. *58*, 1111-1119.

Naumova,A.K., Plenge,R.M., Bird,L.M., Leppert,M., Morgan,K., Willard,H.F., and Sapienza,C. (1996a). Heritability of X chromosome--inactivation phenotype in a large family [see comments]. Am. J. Hum. Genet. *58*, 1111-1119.

Nesbitt,M.N. (1971). X-chromosome inactivation mosaicism in the mouse. Dev Biol *26*, 252-263.

Nesbitt,M.N. and Gartler,S. (1971). The application of genetic mosaicism to developmental problems. Annu Rev Genet *5*, 143-162.

Nesterova,T.B., Slobodyanyuk,S.Y., Elisaphenko,E.A., Shevchenko,A.I., Johnston,C., Pavlova,M.E., Rogozin,I.B., Kolesnikov,N.N., Brockdorff,N., and Zakian,S.M. (2001). Characterization of the genomic Xist locus in rodents reveals conservation of overall gene structure and tandem repeats but rapid evolution of unique sequence. Genome Res. *11*, 833-849.

Newall,A.E., Duthie,S., Formstone,E., Nesterova,T., Alexiou,M., Johnston,C., Caparros,M.L., and Brockdorff,N. (2001). Primary non-random X inactivation associated with disruption of Xist promoter regulation. Hum. Mol. Genet. *10*, 581-589.

Niho,S., Suzuki,K., Yokose,T., Kodama,T., Nishiwaki,Y., and Esumi,H. (1998). Monoclonality of both pale cells and cuboidal cells of sclerosing hemangioma of the lung. Am. J. Pathol. *152*, 1065-1069.

Niho,S., Yokose,T., Suzuki,K., Kodama,T., Nishiwaki,Y., and Mukai,K. (1999). Monoclonality of atypical adenomatous hyperplasia of the lung. Am. J. Pathol. *154*, 249-254.

Nilsson-Ehle,H., Swolin,B., and Westin,J. (1995). Bone marrow progenitor cell growth and karyotype changes in healthy 88- year-old subjects. Eur. J. Haematol. *55*, 14-18.

Norman,A. and Harper,P. (1989). A survey of manifesting carriers of Duchenne and Becker muscular dystrophy in Wales. Clin. Genet. *36*, 31-37.

Nowell,P.C. and Hungerford,D.A. (1960). A minute chromosome in human chronic granulocytic leukemia. Science *132*, 1497.

Nyhan,W.L., Bakay,B., Connor,J.D., Marks,J.F., and Keele,D.K. (1970). Hemizygous expression of glucose-6-phosphate dehydrogenase in erythrocytes of heterozygotes for the Lesch-Nyhan syndrome. Proc. Natl. Acad. Sci. U. S. A *65*, 214-218.

Ogawa,T., Kitagawa,M., and Hirokawa,K. (2000). Age-related changes of human bone marrow: a histometric estimation of proliferative cells, apoptotic cells, T cells, B cells and macrophages. Mech. Ageing Dev. *117*, 57-68.

Ogawa,Y. and Lee,J.T. (2003). Xite, X-inactivation intergenic transcription elements that regulate the probability of choice. Mol. Cell *11*, 731-743.

Ohno,S., Kaplan,W.D., and Kinosita,R. (1959). Formation of the sex chromatin by a single X chromosome in liver cells of *Rattus norvegicus*. Exp Cell Res *18*, 415.

Ohno,S. (1967). Sex chromosomes and sex-linked genes. (Berlin: Springer-Verlag).

Okamoto,T., Okada,M., Wada,H., Kanamaru,A., Kakishita,E., Hashimoto,T., and Furuyama,J. (1998). Clonal analysis of hematopoietic cells using a novel polymorphic site of the X chromosome. Am. J. Hematol. *58*, 263-266.

Orstavik,K.H., Orstavik,R.E., Eiklid,K., and 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,K.H., Orstavik,R.E., and Schwartz,M. (1999). Skewed X chromosome inactivation in a female with haemophilia B and in her non-carrier daughter: a genetic influence on X chromosome inactivation? J. Med. Genet. *36*, 865-866.

Osborne,R.J. and Leech,V. (1994). Polymerase chain reaction allelotyping of human ovarian cancer. Br. J. Cancer *69*, 429-438.

Oscier,D. (1997). Atypical chronic myeloid leukemias. Pathol. Biol (Paris) *45*, 587-593.

Palmer,L.J., Burton,P.R., James,A.L., Musk,A.W., and Cookson,W.O. (2000). Familial aggregation and heritability of asthma-associated quantitative traits in a population-based sample of nuclear families. Eur. J. Hum. Genet. *8*, 853-860.

Panning,B., Dausman,J., and Jaenisch,R. (1997). X chromosome inactivation is mediated by Xist RNA stabilization. Cell *90*, 907-916.

Panning,B. and Jaenisch,R. (1996). DNA hypomethylation can activate Xist expression and silence X-linked genes. Genes Dev. *10*, 1991-2002.

Paradis,V., Dargere,D., Bonvoust,F., Rubbia-Brandt,L., Ba,N., Bioulac-Sage,P., and Bedossa,P. (2000). Clonal analysis of micronodules in virus C-induced liver cirrhosis using laser capture microdissection (LCM) and HUMARA assay. Lab Invest *80*, 1553-1559.

Paradis,V., Laurendeau,I., Vidaud,M., and Bedossa,P. (1998b). Clonal analysis of macronodules in cirrhosis. Hepatology *28*, 953-958.

Paradis,V., Laurendeau,I., Vieillefond,A., Blanchet,P., Eschwege,P., Benoit,G., Vidaud,M., Jardin,A., and Bedossa,P. (1998a). Clonal analysis of renal sporadic angiomyolipomas. Hum. Pathol. *29*, 1063-1067.

Paradis,V., Laurent,A., Flejou,J.F., Vidaud,M., and Bedossa,P. (1997). Evidence for the polyclonal nature of focal nodular hyperplasia of the liver by the study of X-chromosome inactivation. Hepatology *26*, 891-895.

Park,W.W. (1957). The occurrence of sex chromatin in early human and macaque embryos. J. Anat *91*, 369.

Parolini,O., Ressmann,G., Haas,O.A., Pawlowsky,J., Gadner,H., Knapp,W., and Holter,W. (1998). X-linked Wiskott-Aldrich syndrome in a girl. N. Engl. J. Med. *338*, 291-295.

Paulli,M., Bergamaschi,G., Tonon,L., Viglio,A., Rosso,R., Facchetti,F., Geerts,M.L., Magrini,U., and Cazzola,M. (1995). Evidence for a polyclonal nature of the cell infiltrate in sinus histiocytosis with massive lymphadenopathy (Rosai-Dorfman disease). Br. J. Haematol. *91*, 415-418.

Pegoraro,E., Schimke,R.N., Arahata,K., Hayashi,Y., Stern,H., Marks,H., Glasberg,M.R., Carroll,J.E., Taber,J.W., and Wessel,H.B. (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., and Hoffman,E.P. (1997). Familial skewed X inactivation: a molecular trait associated with high spontaneous-abortion rate maps to Xq28 [see comments]. Am. J. Hum. Genet. *61*, 160-170.

Peltonen,L., Palotie,A., and Lange,K. (2000). Use of population isolates for mapping complex traits. Nat. Rev. Genet. *1*, 182-190.

Penny,G.D., Kay,G.F., Sheardown,S.A., Rastan,S., and Brockdorff,N. (1996). Requirement for Xist in X chromosome inactivation [see comments]. Nature *379*, 131-137.

Percec,I. and Bartolomei,M.S. (2002). Genetics. Do X chromosomes set boundaries? Science *295*, 287-288.

Percec,I., Plenge,R.M., Nadeau,J.H., Bartolomei,M.S., and Willard,H.F. (2002). Autosomal dominant mutations affecting X inactivation choice in the mouse. Science *296*, 1136-1139.

Pereira,L.V. and Zatz,M. (1999). Screening of the C43G mutation in the promoter region of the XIST gene in females with highly skewed X-chromosome inactivation. Am. J. Med. Genet. *87*, 86-87.

Piao,Z., Park,Y.N., Kim,H., and Park,C. (1997). Clonality of large regenerative nodules in liver cirrhosis. Liver *17*, 251-256.

Plenge,R.M., Hendrich,B.D., Schwartz,C., Arena,J.F., Naumova,A., Sapienza,C., Winter,R.M., and Willard,H.F. (1997). A promoter mutation in the XIST gene in two unrelated families with skewed X-chromosome inactivation. Nat. Genet. *17*, 353-356.

Plenge,R.M., Percec,I., Nadeau,J.H., and Willard,H.F. (2000). Expression-based assay of an X-linked gene to examine effects of the X- controlling element (Xce) locus. Mamm. Genome *11*, 405-408.

Plenge,R.M., Tranebjaerg,L., Jensen,P.K., Schwartz,C., and Willard,H.F. (1999). Evidence that mutations in the X-linked DDP gene cause incompletely penetrant and variable skewed X inactivation. Am. J. Hum. Genet. *64*, 759-767.

Posnett,D.N., Sinha,R., Kabak,S., and Russo,C. (1994). Clonal populations of T cells in normal elderly humans: the T cell equivalent to "benign monoclonal gammapathy" [published erratum appears in J Exp Med 1994 Mar 1;179(3):1077]. J. Exp. Med. *179*, 609-618.

Prchal,J.T., Carroll,A.J., Prchal,J.F., Crist,W.M., Skalka,H.W., Gealy,W.J., Harley,J., and Malluh,A. (1980). Wiskott-Aldrich syndrome: cellular impairments and their implication for carrier detection. Blood *56*, 1048-1054.

Prchal,J.T. and Guan,Y.L. (1993). A novel clonality assay based on transcriptional analysis of the active X chromosome. Stem Cells (Dayt) *11 Suppl 1*, 62-65.

Prchal,J.T., Prchal,J.F., Belickova,M., Chen,S., Guan,Y., Gartland,G.L., and Cooper,M.D. (1996). Clonal stability of blood cell lineages indicated by X-chromosomal transcriptional polymorphism. J. Exp. Med. *183*, 561-567.

Prissette,M., El Maarri,O., Arnaud,D., Walter,J., and Avner,P. (2001). Methylation profiles of DXPas34 during the onset of X-inactivation. Hum. Mol. Genet. *10*, 31-38.

Puck,J.M., Nussbaum,R.L., and Conley,M.E. (1987). Carrier detection in X-linked severe combined immunodeficiency based on patterns of X chromosome inactivation. J. Clin. Invest *79*, 1395-1400.

Puck,J.M., Stewart,C.C., and Nussbaum,R.L. (1992). Maximum-likelihood analysis of human T-cell X chromosome inactivation patterns: normal women versus carriers of X-linked severe combined immunodeficiency. Am. J. Hum. Genet. *50*, 742-748.

Puck,J.M. and Willard,H.F. (1998). X inactivation in females with X-linked disease. N. Engl. J. Med. *338*, 325-328.

Quade,B.J., McLachlin,C.M., Soto-Wright,V., Zuckerman,J., Mutter,G.L., and Morton,C.C. (1997). Disseminated peritoneal leiomyomatosis. Clonality analysis by X chromosome inactivation and cytogenetics of a clinically benign smooth muscle proliferation. Am. J. Pathol. *150*, 2153-2166.

Rabkin,C.S., Janz,S., Lash,A., Coleman,A.E., Musaba,E., Liotta,L., Biggar,R.J., and Zhuang,Z. (1997). Monoclonal origin of multicentric Kaposi's sarcoma lesions [see comments]. N. Engl. J. Med. *336*, 988-993.

Rampling,M.W. (1999). Haemorheological disturbances in hypertension: the influence of diabetes and smoking. Clin. Hemorheol. Microcirc. *21*, 183-187.

Raskind,W.H., Jacobson,R., Murphy,S., Adamson,J.W., and Fialkow,P.J. (1985). Evidence for the involvement of B lymphoid cells in polycythemia vera and essential thrombocythemia. J. Clin. Invest *75*, 1388-1390.

Raskind,W.H., Tirumali,N., Jacobson,R., Singer,J., and Fialkow,P.J. (1984). Evidence for a multistep pathogenesis of a myelodysplastic syndrome. Blood *63*, 1318-1323.

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

Rastan,S. (1982a). Timing of X-chromosome inactivation in postimplantation mouse embryos. J. Embryol. Exp. Morphol. *71*, 11-24.

Rastan,S. (1983). Non-random X-chromosome inactivation in mouse X-autosome translocation embryos--location of the inactivation centre. J. Embryol. Exp. Morphol. *78*, 1-22.

Rastan,S. and Robertson,E.J. (1985). X-chromosome deletions in embryo-derived (EK) cell lines associated with lack of X-chromosome inactivation. J. Embryol. Exp. Morphol. *90*, 379-388.

Ray,P.F., Winston,R.M., and Handyside,A.H. (1997). XIST expression from the maternal X chromosome in human male preimplantation embryos at the blastocyst stage. Hum. Mol. Genet. *6*, 1323-1327.

Rebbeck,T.R., Kantoff,P.W., Krithivas,K., Neuhausen,S., Blackwood,M.A., Godwin,A.K., Daly,M.B., Narod,S.A., Garber,J.E., Lynch,H.T., Weber,B.L., and Brown,M. (1999). Modification of BRCA1-associated breast cancer risk by the polymorphic androgen-receptor CAG repeat. Am. J. Hum. Genet. *64*, 1371-1377.

Richler,C., Soreq,H., and Wahrman,J. (1992). X inactivation in mammalian testis is correlated with inactive X- specific transcription. Nat. Genet. *2*, 192-195.

Riggs,A.D. (1990). DNA methylation and late replication probably aid cell memory, and type I DNA reeling could aid chromosome folding and enhancer function. Philos. Trans. R. Soc. Lond B Biol. Sci. *326*, 285-297.

Rinaldi,A., Filippi,G., and Siniscalco,M. (1976). Variability of red cell phenotypes between and within individuals in an unbiased sample of 77 heterozygotes for G6PD deficiency in Sardinia. Am. J. Hum. Genet. *28*, 496-505.

Risch N, Tang H, Katzenstein H, and Ekstein J. (2003). Geographic distribution of disease mutations in the Ashkenazi Jewish population supports genetic drift over selection. Am J Hum Genet *72*, 812-822.

Risch N. and Zhang H (1995). Extreme discordant sib pairs for mapping quantitative trait loci in humans. Science *268*, 1584-1589.

Risch NJ and Zhang H (1996). Mapping quantitative trait loci with extreme discordant sib pairs: sampling considerations. Am J Hum Genet *58*, 836-843.

Risch,N. (1990). Linkage strategies for genetically complex traits. II. The power of affected relative pairs. Am. J. Hum. Genet. *46*, 229-241.

Robinson,W.A., Lemon,M., Elefanty,A., Harrison-Smith,M., Markham,N., and Norris,D. (1998). Human acquired naevi are clonal. Melanoma Res *8*, 499-503.

Ropers,H.H., Wienker,T.F., Grimm,T., Schroetter,K., and Bender,K. (1977). Evidence for preferential X-chromosome inactivation in a family with Fabry disease. Am. J. Hum. Genet. *29*, 361-370.

Ropers,H.H., Wolff,G., and Hitzeroth,H.W. (1978). Preferential X inactivation in human placenta membranes: is the paternal X inactive in early embryonic development of female mammals? Hum. Genet. *43*, 265-273.

Rosen-Wolff,A., Soldan,W., Heyne,K., Bickhardt,J., Gahr,M., and Roesler,J. (2001). Increased susceptibility of a carrier of X-linked chronic granulomatous disease (CGD) to Aspergillus fumigatus infection associated with age- related skewing of lyonization. Ann. Hematol. *80*, 113-115.

Rupert,J.L., Brown,C.J., and Willard,H.F. (1995). Direct detection of non-random X chromosome inactivation by use of a transcribed polymorphism in the XIST gene. Eur. J. Hum. Genet. *3*, 333-343.

Russell,L.B. (1963). Mammalian X-chromosome action: Inactivation limited in spread and in region of origin. Science *140*, 976-978.

Russell,L.B. and Cacheiro,N.L. (1978). The use of mouse X-autosome translocations in the study of X-inactivation pathways and nonrandomness. Basic Life Sci. *12*, 393-416.

Sado,T., Wang,Z., Sasaki,H., and Li,E. (2001). Regulation of imprinted X-chromosome inactivation in mice by Tsix. Development *128*, 1275-1286.

Saint-Basile,G., Schlegel,N., Caniglia,M., Le Deist,F., Kaplan,C., Lecompte,T., Piller,F., Fischer,A., and Griscelli,C. (1991). X-linked thrombocytopenia and Wiskott-Aldrich syndrome: similar regional assignment but distinct X-inactivation pattern in carriers. Ann. Hematol. *63*, 107-110.

Samuel,K., Clarke,A.R., Ansell,J.D., and Hooper,M.L. (1993). Age-dependent selection against hypoxanthine phosphoribosyl transferase- deficient cells in mouse haematopoiesis. Development *118*, 859-863.

Sandler,D.P. (1987). Epidemiology of acute myelogenous leukemia. Semin. Oncol. *14*, 359-364.

Sangha,K.K., Stephenson,M.D., Brown,C.J., and Robinson,W.P. (1999). Extremely skewed X-chromosome inactivation is increased in women with recurrent spontaneous abortion. Am. J. Hum. Genet. *65*, 913-917.

Saunders,M.J., Jowitt,S.N., and Yin,J.A. (1995). Clonality studies in patients undergoing allogeneic and autologous bone marrow transplantation for haematological malignancies. Bone Marrow Transplant. *15*, 81-85.

Saxena,A., Alport,E.C., Custead,S., and Skinnider,L.F. (1999). Molecular analysis of clonality of sporadic angiomyolipoma. J. Pathol. *189*, 79-84.

Schmidt,M. and Migeon,B.R. (1990). Asynchronous replication of homologous loci on human active and inactive X chromosomes. Proc. Natl. Acad. Sci. U. S. A *87*, 3685-3689.

Schwartz,J. and Weiss,S.T. (1994). Cigarette smoking and peripheral blood leukocyte differentials. Ann. Epidemiol. *4*, 236-242.

Sheardown,S.A., Duthie,S.M., Johnston,C.M., Newall,A.E., Formstone,E.J., Arkell,R.M., Nesterova,T.B., Alghisi,G.C., Rastan,S., and Brockdorff,N. (1997). Stabilization of Xist RNA mediates initiation of X chromosome inactivation [see comments]. Cell *91*, 99-107.

Shelton,A.J., Harger,J.H., Dorman,J.S., Kuller,L.H., LaPorte,R.E., and Gill,T.J., III (1994). Association between familial autoimmune diseases and recurrent spontaneous abortions. Am. J. Reprod. Immunol. *32*, 82-87.

Shih,L.Y., Lin,T.L., Dunn,P., Wu,J.H., Tseng,C.P., Lai,C.L., Wang,P.N., and Kuo,M.C. (2001). Clonality analysis using X-chromosome inactivation patterns by HUMARA- PCR assay in female controls and patients with idiopathic thrombocytosis in Taiwan. Exp Hematol. *29*, 202-208.

Short,M.A., Evans,P.A., Shiach,C.R., Jack,A., Richards,S., and Morgan,G.J. (1996). 32P-incorporation PCR for the detection of rearrangements at the TCR- gamma locus. Diagn. Mol. Pathol. *5*, 26-32.

Simmler,M.C., Cattanach,B.M., Rasberry,C., Rougeulle,C., and Avner,P. (1993). Mapping the murine Xce locus with (CA)n repeats. Mamm. Genome *4*, 523-530.

Singer-Sam,J., Chapman,V., LeBon,J.M., and Riggs,A.D. (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. U. S. A *89*, 10469-10473.

Skadhauge,L.R., Christensen,K., Kyvik,K.O., and Sigsgaard,T. (1999). Genetic and environmental influence on asthma: a population-based study of 11,688 Danish twin pairs. Eur. Respir. J. *13*, 8-14.

Sleddens,H.F., Oostra,B.A., Brinkmann,A.O., and Trapman,J. (1992). Trinucleotide repeat polymorphism in the androgen receptor gene (AR). Nucleic Acids Res. *20*, 1427.

Smith,C.I. and Notarangelo,L.D. (1997). Molecular basis for X-linked immunodeficiencies. Adv. Genet. *35*, 57-115.

Spielman RS and Ewens WJ (1998). A sibship test for linkage in the presence of association: the sib transmission/disequilibrium test. Am J Hum Genet *62*, 450-458.

Spurdle,A.B., Dite,G.S., Chen,X., Mayne,C.J., Southey,M.C., Batten,L.E., Chy,H., Trute,L., McCredie,M.R., Giles,G.G., Armes,J., Venter,D.J., Hopper,J.L., and Chenevix-Trench,G. (1999). Androgen receptor exon 1 CAG repeat length and breast cancer in women before age forty years. J. Natl. Cancer Inst. *91*, 961-966.

Stewart,J.J. (1998). The female X-inactivation mosaic in systemic lupus erythematosus. Immunol. Today *19*, 352-357.

Tada,T., Obata,Y., Tada,M., Goto,Y., Nakatsuji,N., Tan,S., Kono,T., and Takagi,N. (2000). Imprint switching for non-random X-chromosome inactivation during mouse oocyte growth. Development *127*, 3101-3105.

Tai,H.H., Gordon,J., and McBurney,M.W. (1994). Xist is expressed in female embryonal carcinoma cells with two active X chromosomes. Somat. Cell Mol. Genet. *20*, 171-182.

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

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

Takagi,N., Wake,N., and Sasaki,M. (1978). Cytologic evidence for preferential inactivation of the paternally derived X chromosome in XX mouse blastocysts. Cytogenet. Cell Genet. *20*, 240-248.

Tan,S.S., Godinho,L., and Tam,P.P. (2000). Cell lineage analysis. X-inactivation mosaics. Methods Mol. Biol. *135*, 289-295.

Tan,S.S., Williams,E.A., and Tam,P.P. (1993). X-chromosome inactivation occurs at different times in different tissues of the post-implantation mouse embryo [published erratum appears in Nat Genet 1993 Jul;4(3):320]. Nat. Genet. *3*, 170-174.

Tate,J.E., Mutter,G.L., Boynton,K.A., and Crum,C.P. (1997). Monoclonal origin of vulvar intraepithelial neoplasia and some vulvar hyperplasias. Am. J. Pathol. *150*, 315-322.

Tefferi,A., Thibodeau,S.N., and Solberg,L.A., Jr. (1990). Clonal studies in the myelodysplastic syndrome using X-linked restriction fragment length polymorphisms. Blood *75*, 1770-1773.

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

Tihy,F., Vogt,N., Recan,D., Malfoy,B., Leturcq,F., Coquet,M., Serville,F., Fontan,D., Guillard,J.M., and Kaplan,J.C. (1994). Skewed inactivation of an X chromosome deleted at the dystrophin gene in an asymptomatic mother and her affected daughter. Hum. Genet. *93*, 563-567.

Till,J.E. and McCulloch,E.A. (1980). Hemopoietic stem cell differentiation. Biochim. Biophys. Acta *605*, 431-459.

Tonon,L., Bergamaschi,G., Dellavecchia,C., Rosti,V., Lucotti,C., Malabarba,L., Novella,A., Vercesi,E., Frassoni,F., and Cazzola,M. (1998). Unbalanced X-chromosome inactivation in haemopoietic cells from normal women. Br. J. Haematol. *102*, 996-1003.

Tregouet,D.A. and Tiret,L. (2000). Applications of the estimating equations theory to genetic epidemiology: a review. Ann. Hum. Genet. *64*, 1-14.

Tsang,C.W., Lazarus,R., Smith,W., Mitchell,P., Koutts,J., and Burnett,L. (1998). Hematological indices in an older population sample: derivation of healthy reference values. Clin. Chem. *44*, 96-101.

Tsukamoto,N., Morita,K., Maehara,T., Okamoto,K., Karasawa,M., Omine,M., and Naruse,T. (1993). Clonality in myelodysplastic syndromes: demonstration of pluripotent stem cell origin using X-linked restriction fragment length polymorphisms. Br. J. Haematol. *83*, 589-594.

Tsukamoto,N., Morita,K., Maehara,T., Okamoto,K., Sakai,H., Karasawa,M., Naruse,T., and Omine,M. (1994). Clonality in chronic myeloproliferative disorders defined by X- chromosome linked probes: demonstration of heterogeneity in lineage involvement. Br. J. Haematol. *86*, 253-258.

Turhan,A.G., Humphries,R.K., Phillips,G.L., Eaves,A.C., and Eaves,C.J. (1989). Clonal hematopoiesis demonstrated by X-linked DNA polymorphisms after allogeneic bone marrow transplantation [see comments]. N. Engl. J. Med. *320*, 1655-1661.

Uehara,S., Tamura,M., Nata,M., Ji,G., Yaegashi,N., Okamura,K., and Yajima,A. (2000). X-chromosome inactivation in the human trophoblast of early pregnancy. J. Hum. Genet. *45*, 119-126.

Uhlig,T., Hagen,K.B., and Kvien,T.K. (1999). Current tobacco smoking, formal education, and the risk of rheumatoid arthritis. J. Rheumatol. *26*, 47-54.

van Dijk,J.P., Heuver,L., Stevens-Linders,E., Jansen,J.H., Mensink,E.J., Raymakers,R.A., and de Witte,T. (2002). Acquired skewing of Lyonization remains stable for a prolonged period in healthy blood donors. Leukemia *16*, 362-367.

Van Eeden,S.F. and Hogg,J.C. (2000). The response of human bone marrow to chronic cigarette smoking. Eur. Respir. J. *15*, 915-921.

Vickers,M.A., McLeod,E., Spector,T.D., and Wilson,I.J. (2001). Assessment of mechanism of acquired skewed X inactivation by analysis of twins. Blood *97*, 1274-1281.

Vineis,P. and Caporaso,N. (1995). Tobacco and cancer: epidemiology and the laboratory. Environ. Health Perspect. *103*, 156-160.

Vogelstein,B., Fearon,E.R., Hamilton,S.R., and Feinberg,A.P. (1985). Use of restriction fragment length polymorphisms to determine the clonal origin of human tumors. Science *227*, 642-645.

Vogelstein,B., Fearon,E.R., Hamilton,S.R., Preisinger,A.C., Willard,H.F., Michelson,A.M., Riggs,A.D., and Orkin,S.H. (1987). Clonal analysis using recombinant DNA probes from the X-chromosome. Cancer Res. *47* , 4806-4813.

Vogrincic,G.S., O'Connell,J.X., and Gilks,C.B. (1997). Giant cell tumor of tendon sheath is a polyclonal cellular proliferation. Hum. Pathol. *28*, 815-819.

Vulliamy,T.J., Knight,S.W., Dokal,I., and Mason,P.J. (1997). Skewed X-inactivation in carriers of X-linked dyskeratosis congenita. Blood *90*, 2213-2216.

Wada,H., Enomoto,T., Tsujimoto,M., Nomura,T., Murata,Y., and Shroyer,K.R. (1998). Carcinosarcoma of the breast: molecular-biological study for analysis of histogenesis. Hum. Pathol. *29*, 1324-1328.

Wallen,N., Kita,H., Weiler,D., and Gleich,G.J. (1991). Glucocorticoids inhibit cytokine-mediated eosinophil survival. J. Immunol. *147*, 3490-3495.

Walsh,D.S., Peacocke,M., Harrington,A., James,W.D., and Tsou,H.C. (1998). Patterns of X chromosome inactivation in sporadic basal cell carcinomas: evidence for clonality. J. Am. Acad. Dermatol. *38*, 49-55.

Wang,J.C., Doedens,M., and Dick,J.E. (1997). Primitive human hematopoietic cells are enriched in cord blood compared with adult bone marrow or mobilized peripheral blood as measured by the quantitative in vivo SCID-repopulating cell assay. Blood *89*, 3919-3924.

Wareham,K.A., Lyon,M.F., Glenister,P.H., and Williams,E.D. (1987). Age related reactivation of an X-linked gene. Nature *327*, 725-727.

Warshawsky,D., Stavropoulos,N., and Lee,J.T. (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. U. S. A *96*, 14424-14429.

Weiss,S.T., Segal,M.R., Sparrow,D., and Wager,C. (1995). Relation of FEV1 and peripheral blood leukocyte count to total mortality. The Normative Aging Study. Am. J. Epidemiol. *142*, 493-498.

Welshons,W.J. and Russell,L.B. (1959). Proc. Natl. Acad. Sci. U. S. A *45*, 560.

Wengler,G., Gorlin,J.B., Williamson,J.M., Rosen,F.S., and Bing,D.H. (1995). Nonrandom inactivation of the X chromosome in early lineage hematopoietic cells in carriers of Wiskott-Aldrich syndrome. Blood *85*, 2471-2477.

West,J.D. and Chapman,V.M. (1978). Variation for X chromosome expression in mice detected by electrophoresis of phosphoglycerate kinase. Genet. Res *32*, 91-102.

West,J.D., Frels,W.I., Chapman,V.M., and Papaioannou,V.E. (1977). Preferential expression of the maternally derived X chromosome in the mouse yolk sac. Cell *12*, 873-882.

Wiggans,R.G., Jacobson,R.J., Fialkow,P.J., Woolley,P.V., III, Macdonald,J.S., and Schein,P.S. (1978). Probable clonal origin of acute myeloblastic leukemia following radiation and chemotherapy of colon cancer. Blood *52*, 659-663.

Willard,H.F. (1996b). X chromosome inactivation and X-linked mental retardation. Am. J. Med. Genet. *64*, 21-26.

Willard,H.F. (1996a). X chromosome inactivation, XIST, and pursuit of the X-inactivation center. Cell *86*, 5-7.

Willman,C.L., Busque,L., Griffith,B.B., Favara,B.E., McClain,K.L., Duncan,M.H., and Gilliland,D.G. (1994). Langerhans'-cell histiocytosis (histiocytosis X)--a clonal proliferative disease [see comments]. N. Engl. J. Med. *331*, 154-160.

Wolf,S.F., Dintzis,S., Toniolo,D., Persico,G., Lunnen,K.D., Axelman,J., and Migeon,B.R. (1984). Complete concordance between glucose-6-phosphate dehydrogenase activity and hypomethylation of 3' CpG clusters: implications for X chromosome dosage compensation. Nucleic Acids Res. *12*, 9333-9348.

Wolf,S.F. and Migeon,B.R. (1982). Studies of X chromosome DNA methylation in normal human cells. Nature *295*, 667-671.

Wutz,A. and Jaenisch,R. (2000). A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. Mol. Cell *5*, 695-705.

Yano,T., Jimbo,H., Yoshikawa,H., Tsutsumi,O., and Taketani,Y. (1999). Molecular analysis of clonality in ovarian endometrial cysts. Gynecol. Obstet. Invest *47 Suppl 1*, 41-45.

Yong,E.L., Lim,J., Qi,W., Ong,V., and Mifsud,A. (2000). Molecular basis of androgen receptor diseases. Ann. Med. *32*, 15-22.

Yu,H., Bharaj,B., Vassilikos,E.J., Giai,M., and Diamandis,E.P. (2000). Shorter CAG repeat length in the androgen receptor gene is associated with more aggressive forms of breast cancer. Breast Cancer Res Treat. *59*, 153-161.

Zauber,N.P. and Zauber,A.G. (1987). Hematologic data of healthy very old people. JAMA *257*, 2181-2184.

Zhu,J., Frosch,M.P., Busque,L., Beggs,A.H., Dashner,K., Gilliland,D.G., and Black,P.M. (1995). Analysis of meningiomas by methyla. Cancer Res. *55*, 3865-3872.

Zhu,J.J., Leon,S.P., Folkerth,R.D., Guo,S.Z., Wu,J.K., and Black,P.M. (1997). Evidence for clonal origin of neoplastic neuronal and glial cells in gangliogliomas. Am. J. Pathol. *151*, 565-571.

Zollner S, Wen X, Hanchard NA, Herbert MA, Ober C, and Pritchard JK. (2004). Evidence for extensive transmission distortion in the human genome. Am J Hum Genet *74(1):62-72*, 62-72.

APPENDIXES

Annex 1

HERITABILITY OF NONMRANDOM X-INACTIVATION IN HEALTHY FEMALES

Heritability of Primary Nonrandom X-Inactivation in Healthy Females: Implication in Primordial Hematopoietic Stem Cell Pool Size

Robert Mio, Alain Bonnardeaux and Lambert Busque

Research Center, Maisonneuve-Rosemont Hospital, University of Montreal, Montreal, Canada.

This work was supported by the Bayer/Canadian Red Cross Society Research and Development Fund. L.B. and A.B. are scholars of the fonds de la recherche en santé du Quebec (FRSQ). A.B. is scholar of the Kidney Foundation of Canada.

Correspondence Address

Lambert Busque MD
Research Center
Hôpital Maisonneuve-Rosemont
5415 de l'Assomption blvd.
Montreal, Quebec, Canada H1T 2M4
Tel: 514-252-3495;
Fax: 514-254-5094

Abstract

Nonrandom X-inactivation occurs in a significant proportion of normal females. This phenomenon is thought to be caused by the small number of cells present at time of X-inactivation, most of which stochastically inactivate the same X-chromosome. However, it is also possible that nonrandom X-inactivation is genetically determined as in the mouse *Xce* model. In this model, alleles of different strengths have an influence, *in cis*, on which X chromosome is to be inactivated. To test for heritability of nonrandom X-inactivation in humans, we examined X-inactivation ratios at the human androgen receptor (HUMARA) locus in 142 heterozygous female neonate-mother pairs for concordance of nonrandom X-inactivation. Neonates were categorized into two groups: those exhibiting random X-inactivation ($1.0 \leq$ ratio < 3.0) or those with nonrandom X-inactivation (ratio ≥ 3.0). The incidence of skewing in the mothers of non-skewed neonates was 20.8% (26/125), while the incidence was 58.8% (10/17) in mothers of skewed neonates ($p = 0.0018$). Concordance for nonrandom X-inactivation between neonates and their respective mothers indicates that primary nonrandom X-inactivation is a heritable trait in humans, possibly determined genetically. This implies that the estimation of the primordial hematopoietic stem cell pool size cannot be made accurately if based exclusively on the stochastic model of nonrandom X-inactivation. No evidence for a rare *XIST* mutation predisposing to skewing was detected. Further studies are required to determine whether genetic factors are implicated in primary nonrandom X-inactivation.

Key words: X chromosome inactivation, nonrandom X inactivation, HUMARA, Xce, *XIST, PGK*

Introduction

X-chromosome inactivation allows gene dosage compensation between XX females and XY males.[1] Analysis of X chromosome deletions and X-autosome translocations have suggested X-inactivation to be dependent on a *cis* residing segment, termed the X-inactivation center (*Xic*),[2,3] regionalized to Xq13.[4] Brown et al identified a gene, localized to the same interval as *Xic*, that was expressed exclusively from the inactive X (Xi).[4,5] The gene was termed *XIST* for X-inactivation specific transcript. The human *XIST* gene does not appear to encode a protein, but may function as a structural RNA.[6] Penny et al have recently provided clear evidence for the involvement of *Xist* in X-inactivation: X chromosomes bearing a knockout *Xist* gene fail to undergo *cis* inactivation while the normal X chromosome would undergo X-inactivation.[7]
In the developing normal embryo, both X-chromosomes of the newly fertilized egg are in the active state.[8] The first occurrence of *XIST* expression was found at the 4-cell stage and is limited to the trophoectoderm of the developing blastocyst and is imprinted with exclusive expression of the unmethylated paternal allele.[9,10] This is followed by preferential paternal X-inactivation in extraembryonic tissues. A genome wide demethylation mechanism thought to occur between the 8 cell and the blastocyst stage is believed responsible for erasure of the parental imprints,[10,11] resulting in random *XIST* expression of the inner cell

mass, hence random X-inactivation as seen at the gastrula stage. Once X-inactivation has occurred, it is stably inherited through successive cell divisions.[12] Random X-inactivation results in females being mosaics for X-linked alleles, with a balanced expression between maternal (Xm) and paternal (Xp) alleles.[13] Any significant deviation from the theoretical 1:1 ratio between the Xp/Xm is termed nonrandom X-inactivation (synonymous for excessive Lyonization or skewing).

Clonal disorders such as leukemias (reviewed in 14), carrier states for a deleterious mutation on the X chromosome such as in the Wiskott-Aldrich syndrome (reviewed in 15) and possibly stem cell depletion,[16-18] cause skewed patterns of X-inactivation. Furthermore, close to 25% of normal females exhibit nonrandom X-inactivation patterns in blood-derived cells.[19-21] These females are presumed to be excessively Lyonized and exhibit no evidence for one of the causes of skewing above stated. The exact etiology of primary nonrandom X-inactivation in normal females is not known. The most accepted hypothesis is based on a pure stochastic model of X-inactivation, where inactivation occurs when the primordial cell pool is composed of as few as 10-20 cells.[19,22] Accordingly, the variance of the distribution of X-inactivation patterns in females will follow a binomial distribution with a significant proportion carrying unbalanced X-chromosome inactivation patterns. In contrast to the situation in humans, a genetic model for nonrandom X-inactivation has been well documented in mice.[23] The X chromosome controlling element (*Xce*) can bias the choice of which X is to be inactivated during embryogenesis, thereby interfering with the randomness of X-inactivation. *Xce* has three well characterized alleles classified in a gradient of strength: $Xce^c > Xce^b > Xce^a$,[24] whereby heterozygotes manifest nonrandom X-inactivation. *Xce* maps to or within the vicinity of the *Xic* region,[25] and has been shown both biochemically[24] and cytologically[26] to cause primary nonrandom X-inactivation rather than secondary selection of cells. Furthermore, the effects of different *Xce* alleles can modify the imprinted preferential inactivation of Xp in extraembryonic tissues.[27] Recently, *Xce* and *Xist* were found to be separable genetic loci.[28]

As in the mouse model, demonstration of heritability of nonrandom X-inactivation in humans may provide evidence for a genetic basis of skewing. Although the familial clustering of skewed X inactivation has been previously reported (Orstavik, Pegoraro, Plenge, Naumova), thus implying a genetic (X-linked) genetic basis to skewing, selection against X-linked mutant alleles is still a plausible mechanism, thus warranting further investigation. Nonetheless, the lack of highly informative X-inactivation assays has precluded such investigations in the past. The identification of a highly polymorphic CAG short tandem repeat in the first exon of the X-linked human androgen receptor gene (HUMARA) coupled with differential methylation sites allowing the distinction between the active (Xa) and inactive X (Xi) chromosome has provided a powerful means of analyzing X-inactivation patterns in more than 90% of females.[29] The HUMARA assay has been validated in numerous studies,[30,31,32] and has allowed us to search for evidence of heritability of primary nonrandom X-inactivation by analyzing the relationship between the degree of nonrandom X-inactivation in mother-daughter (neonate) pairs. Neonate-mother pairs were chosen because we have recently documented

that the incidence of nonrandom X-inactivation increases with age in blood cells of normal females.[33] The incidence in female neonates was 8.6%, increasing to 16.4% in females close to 30 years old, and to 37.9% in females aged 60 years and over. As neonates have not been submitted to environmental exposure to the same extent as older females, they represent a group where a genetic influence on primary nonrandom X-inactivation might be easier to detect.

Materials and Methods

Normal female population.

Peripheral blood from 177 female neonate-mother pairs (total of 354 female individuals) was screened for heterozygosity at the HUMARA locus under the auspices of institutional approval for analysis of anonymous discarded blood samples. Only samples where both the mother and daughter were heterozygous for the HUMARA locus were included in this study. The mother group was comprised of healthy females coming to the obstetric department for delivery. The neonatal group was comprised of at term healthy singleton babies. For this group, cord blood was selected instead of peripheral blood. Each of these females were hematologically healthy, had a normal CBC and WBC differential, including normal red cell indices.

DNA extraction.

Peripheral blood was processed for DNA extraction as follows: 3-5 ml of peripheral blood was lysed with Triton X-100 (Sigma, St.-Louis) lysis solution (0.32 M Sucrose; 10mM Tris pH 7.5; 5 mM $MgCl_2$; 1% Triton X-100), digested with proteinase K in SDS buffer (Proteinase K 2 mg/ml, 5% SDS) at 37°C for 24 hours. Then DNA was extracted with phenol, phenol/chloroform and chloroform, precipitated with 2 volumes of absolute ethanol and 1/10 volume of 3 M sodium acetate, and resuspended in Tris-EDTA (TE) buffer.

HUMARA clonality assay.

The HUMARA assay was carried out as previously described[29] with minor modifications. Briefly, genomic DNA was pre-cut by mixing sample DNA (100ng-1 µg in 2 µl) with *Hpa* II (1 µl, high concentration, 40 U/µl), *Rsa* I (0.5 µl, high concentration, 40 U/µl), L buffer (2 µl, Boehringer Mannheim) and H_2O (14.5 µl). An auto-control was prepared in the same way except that *Hpa* II was omitted from the mix. *PCR amplification of the HUMARA locus:* 2µl of digested DNA were added to 23 µl of a PCR mix containing buffer, dNTPs (200 µM each); 12.5 pmol each primer: HUMARA I (5'-GCTGTGAAGGTTGCTGTTCCTCAT-3'), and HUMARA II: (5'-TCCAGAATCTGTTCCAGAGCGTGC-3') ; DMSO (1.0 µl, Sigma); □-^{32}P end labeled HUMARA I primer (1.25 pmol), Taq polymerase (0.5 units, Perkin Elmer-Cetus); H_2O to final volume of 23 µl. Samples were amplified on a programmable thermal cycler (MJ Research, Inc.); initial DNA denaturation at 94°C for 3 min., then 28 cycles starting with 94°C

for 45 sec, 60°C for 30 sec and 72°C for 30 sec. Amplified PCR products (8-10 µl) were electrophoresed on a 4% acrylamide-urea-formamide denaturing gel at 60 watts for approximately 3 hours.

Quantitation of alleles.

The allele ratio was defined as the ratio between the two X-linked alleles in a given sample. The corrected ratio (CR) was defined as the allele ratio of the pre-cut sample divided by the allele ratio of the non-precut sample of the same specimen. The CR compensates for potential preferential amplification of one of the two alleles. Gels were exposed to high performance autoradiography film overnight at -80°C. Developed film was scanned on a LKB Bromma Ultroscan XL laser densitometer (Pharmacia AB, Bromma, Sweden) and the ratios between the two X-linked alleles were measured using Gel Scan XL software (Pharmacia LKB Biotechnology AB, Bromma, Sweden).

Phosphoglycerate kinase (PGK)-PCR clonality assay (PPCA)

This technique was performed as described previously.[34]

Hha I restriction digest.

The restriction assay was carried out as previously described (Plenge) with minor modifications. Briefly, PB DNA (30 ng) were added to 23 µl of a PCR mix containing buffer; dNTPs (200 µM each); 12.5 pmol of each primer: G7R (5'-GAAGTTGTGACTCCTGGTCT-3') and G10R (5'-GAGAGATCTTCAGTCAGGAAG-3'); DMSO (4%), Taq DNA polymerase (0.5 units, BM); H_2O to final volume of 23 µl. Samples were amplified on a programmable thermal cycler (MJ Research, Inc.); initial DNA denaturation at 94°C for 3 min., then 35 cycles starting with 94°C for 45 sec, 55°C for 30 sec and 72°C for 30 sec. Amplified PCR products (10 µl) were electrophoresed on a 3% etidium bromide-stained agarose gel.

Criteria for nonrandom X-inactivation

Criteria for nonrandom X-inactivation is arbitrary. However, to allow comparison with previous literature we chose CR $\geq$ 3:1 which corresponds to greater than 75% expression of one allele. This ratio was widely accepted in the literature.[19,20,33,35,36]

Statistical analysis.

Overall comparison of the two neonate categories (CR <3.0, CR $\geq$3.0), with respect to allele ratios in the mother population, was performed *via* chi-square analysis or Fisher exact test when appropriate.

Results

Incidence of nonrandom X-inactivation.

DNA samples of 142 of the 177 (80.2%) mother-neonates pairs were both heterozygous at the HUMARA locus. The allelic ratio between the two X-linked alleles at the HUMARA locus were determined for each individual. The overall incidence of skewing (CR $\geq$ 3:1) in the maternal population was 25.4% (36/142). These females were aged from 16 to 40 years old with a mean at 27.2 y.o. The incidence of skewing was 12.0% (17/142) in the neonate population. This increased incidence of skewing with age was statistically significant (p=0.0038, □^{2}= 8.37).

Concordance in nonrandom X-inactivation between neonates and their mothers.

Allelic ratios for the neonate population were scored into two categories: skewed (n=17) or non-skewed (n=125). For the 17 skewed neonates, 10 (58.8%) had mothers that were also skewed. In contrast, of the 125 non-skewed neonates, only 26 (20.8%) had mothers that were skewed (Fig 1). Statistically, the incidence of skewing in the mothers of skewed neonates was significantly different from that of non-skewed neonates (p=0.0018, Fisher's exact test, 2-tail). Figure 2 shows some examples of skewed mother-daughter pairs. The mean age of mothers of skewed neonates was 26.53 years old, which was not significantly different from mothers of non-skewed neonates: 26.99 years old.

Parental origin of the inactive X chromosome in skewed neonates.

Of the 17 neonates found skewed (CR □ 3:1) by HUMARA analysis, 8 had the paternal X skewed (preferentially inactive), and 9 had the maternal X skewed (data not shown).

Concordance of activation states in skewed mother-neonate pairs

Analysis of the activation status of the transmitted X chromosome from parent to daughter is limited to the analysis of maternal X's. Ten of 17 skewed neonates had skewed mothers, thereby allowing this analysis. In 7 pairs the activation state was retained: the transmitted X was inactive in the mother and inactive in the daughter (Fig 2A), or the transmitted X was active and remained active in the daughter (Fig 2B). The activity state was discordant in 3 pairs: the inactive X was transmitted but became preferentially active in the daughter (Fig 2C), or the active X was transmitted and became preferentially inactive in the daughter (Fig 2D).

PGK-PCR clonality analysis.

Of the 17 neonates skewed by HUMARA analysis, 7 were informative (heterozygous) for PGK analysis. All 7 were also skewed by PGK analysis (data not shown). Similarly, of the 36 mothers skewed by HUMARA analysis, 15 were informative and skewed at the PGK locus. The 7 non-skewed mothers of skewed neonates were also tested, 3 were PGK informative and exhibited a random X-inactivation pattern. Lastly, 26 non-skewed daughters of skewed mothers were tested. Seven were PGK informative: 6 exhibited random X-inactivation patterns while one had a skewed pattern of X inactivation. Therefore, 31 of 32 (96.9%) had concordant results at both loci.

Xist minimal promoter mutation.

A C-to-G mutation at position -43 in the minimal promoter of the Xist gene creates a novel *Hha* I restriction enzyme site. An internal *Hha* I site serves as an internal control. Of the 17 mother-neonate pairs analyzed, none exhibited the Xist promoter mutation (data not shown).

Discussion

In this study we have documented, for the first time, a highly significant correlation between the degree of nonrandom X-inactivation in female neonates and their mothers. We found that a neonate with nonrandom X-inactivation has a 59% chance of having a mother exhibiting the same phenomenon. Only 20% of neonates with random X-inactivation had a mother with a skewed pattern of inactivation. It has been possible to document convincingly the heritability of nonrandom X-inactivation on account of the HUMARA assay's highly informative polymorphism. More than 80% of mother-neonate pairs were heterozygous, thus evaluable for concordance. In comparison, the analysis of other loci, such as PGK, would have allowed the analysis of only 10 % of mother-daughter pairs, which may have lead to an acertainement bias. Furthermore, the skewing documented in the neonate population may correspond to primary nonrandom X-inactivation since their blood cells have not been subjected to selective pressure or environmental factors to the same extent as in older females. The presence of phenocopies, as documented by the increase of skewing with age, would have obscured the heritability of the trait if older subjects were analyzed.

The aggregation of skewing in mother-daughter pairs suggests that nonrandom X-inactivation is a heritable trait in humans. If X-inactivation was purely stochastic and nonrandom X-inactivation was a rare statistical event where most cells have inactivated the same X only by chance, no concordance between the mother and daughter would have been expected. However, heritability of skewing, as for any case of familiality, could be caused by shared environment or by genetic factors, either autosomal, mitochondrial or X-linked. The analysis of this observation is difficult since nonrandom X-inactivation may be caused by several different mechanisms occurring at different times in the development of the embryo (Table 1). Several of these mechanisms may be heritable.

The possibility that heritability of skewing in humans is caused by a *cis*-acting element that biases the choice of the X to be inactivated has to be thoughtfully considered. The evolutionary conservation of several mechanisms and sequences responsible for X-inactivation in eutherians reinforces the possibility that the mouse *Xce* may have an human equivalent. In fact, the data obtained in the mother-daughter pairs does not exclude an *XCE* model in humans. Strong *XCE* alleles can be transmitted by either parent and concordance of activation state between a skewed mother and daughter is expected if the *XCE* locus is genetically linked to the HUMARA locus. The discordance of activation state in 3 of the 10 pairs could be due to the effects of recombination (genetic distance) between the HUMARA locus and a putative *XCE* locus, or alternatively, to the effects of the paternal *XCE* allele, whose "strength" cannot be assessed relative to the maternal allele. Another possibility is that these three discordant mothers are phenocopies, i.e. manifest skewing for other

reasons, and that the father is transmitting the strong *XCE* allele. Is there further evidence for an *XCE* model in humans? The only supporting evidence would come from a unique family reported by Naumova,[37] where an inherited pattern of nonrandom X-inactivation was observed in a healthy family. Seven daughters were skewed with the paternal X being the preferentially active one. If the *XCE* model is operational in this family, the phenotype may be caused by an exceptionally strong allele that can completely bias the choice of the X to be inactivated. However, the occurrence of such strong alleles is probably very rare: we have analyzed 69 families composed of 387 individuals and found no such pattern of transmission (Mio R and Busque L, unpublished). One possible limitation in the analysis of skewing in multigenerational families is the increased incidence of skewing with age. This phenomenon is particularly significant after the age of 60. Since the cause(s) of acquired skewing are not known, several of these skewed females may represent phenocopies and obscure the evidence for specific genetic transmission of the skewing trait. This acquired skewing phenomenon was evident in the mother-daughters pairs since the incidence of nonrandom X-inactivation was 12% in the neonates and 25.4% in the mothers (p=0.0038). However, the mother's age had no effect on the incidence of skewing in the neonates since the mean age of mothers of skewed neonates was the same as mothers of non-skewed neonates.

Mutations in the Xist promoter have also been implicated as a heritable cause of skewing. However, the mutation was not detectable in our 17 mother-skewed neonate pairs. Furthermore, In the study by Plenge et al., of 1166 independent X chromosomes analyzed, only one was found to harbor the mutation. This suggests it is a rare mutation and does not explain the majority of skewed individuals in the population.

Carrier state for mutant X-linked alleles is also associated with heritable skewing of the X chromosome, since cells with the active X carrying the deleterious gene are selected against. However, if the mother is transmitting the mutant allele, it is expected that the maternal allele will be inactive in the daughter. Thus all daughters of skewed females will have the paternal X in the active state. This was not the case for 4 of the 10 daughters of skewed mothers. Furthermore, the mutant allele model would imply that at least 12% of the normal female population carry deleterious genes on the X chromosome. There is no evidence for such a high incidence of X-linked diseases in the normal population.

Another theoretical mechanism possibly associated with heritability of skewing is the presence of a trans-acting factor that may influence the choice of the X to be inactivated. This factor could be X-linked, autosomal, or even mitochondrial. The analysis of the mode of transmission has shown that a potential skewing gene may be transmitted by either parent. This would be less compatible with mitochondrial inheritance but would favor autosomal or X-linked inheritance. However, there is no definite documentation of a trans-acting factor that regulates X-inactivation. Although we have mainly considered genetic factors so far, several other mechanisms may be heritable. Even a pure stochastic model of X-inactivation may show heritability if the number of cells present at the time of X-inactivation varies in families. For example, if X-inactivation occurs when the number of embryoblasts is consistently lower in certain families, this will lead

to an increased incidence of nonrandom X-inactivation in these families, even in the absence of influence on the choice of the X to be inactivated.

The documentation of heritability, irrespective of its cause, may force reevaluation of certain assertions made in the past. For example, the incidence of nonrandom X-inactivation in normal females has been used to evaluate the number of embryological cells contributing to the hematopoietic stem cell pool, assuming a pure stochastical model of X-inactivation. The number of contributing cells estimated by this method is relatively low: 5-18.[19,22,41-44] However, if other causes of skewing such as a genetic predisposition, as suggested by heritability of the trait, accounts for a significant proportion of skewed females, the estimation of the primordial pool size has been underestimated.

The documentation of heritability of nonrandom X-inactivation justifies the efforts of mapping and identifying the skewing gene(s). This will gain insight in the mechanism of X-inactivation and the interactions between genetic factors and nonrandom X-inactivation. However, due to a multiplicity of causes giving rise to skewed patterns of X-inactivation, mapping of the skewing gene(s) may be a difficult task. Mapping strategies should focus on identifying a population with low phenocopies. The better understanding of primary nonrandom X-inactivation will help characterize the etiology of acquired skewing documented in aging normal females. This may lead to a better understanding of the biology of hematopoiesis with age.

Acknowledgments

We would like to thank Dr. E.A. Drobetsky and Dr. C. Perreault for critically reviewing the manuscript.

Footnotes

1. *Abbreviations used in this paper:* CR, corrected ratio; HUMARA, human androgen-receptor gene; PGK, phosphoglycerate kinase; Xa, active X chromosome; Xi, inactive X chromosome; Xm, maternal X chromosome; Xp, paternal X chromosome; XCE, X chromosome controlling element; XIC, X-inactivation center; XIST, X-inactivation specific transcript.

Figure 1. Concordance of skewing in mother-daughter pairs. Neonates were categorized according to the allelic ratios obtained at the HUMARA locus: skewed (ratio $\geq$ 3.0) or non-skewed (ratio < 3.0). The incidence of skewing in respective mothers was 58.8 % (10/17) for skewed neonates and 20.8% (10/17) for non-skewed neonates (p=0.0018).

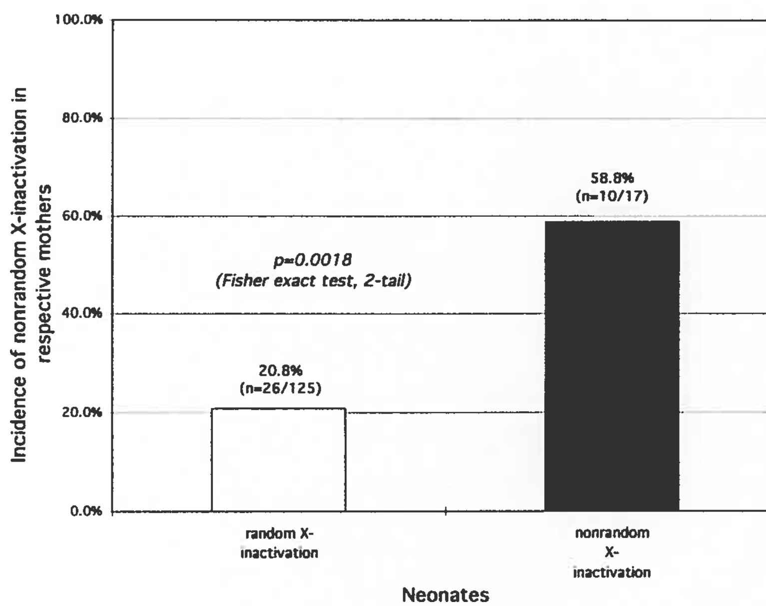


Figure 2. Representative results of the HUMARA clonality assay in mother-daughter pair analysis. Of seventeen skewed neonates, ten had mothers that were also skewed - four pairs are shown above (samples A-D). For the remaining seven neonates, mothers exhibited random X inactivation patterns (data not shown). Lane 1 (-) contains the auto-control, lane 2 (+) the *Hpa*II digested sample. Samples A and B, the status of the maternally derived X is retained in the daughter. Samples C and D, the activation state of the maternally derived X is not retained in the daughter. Results are expressed as corrected ratios (C.R.). Ratios greater than or equal to 3 represent nonrandom X-inactivation pattern. M, mother; D, daughter (neonate).

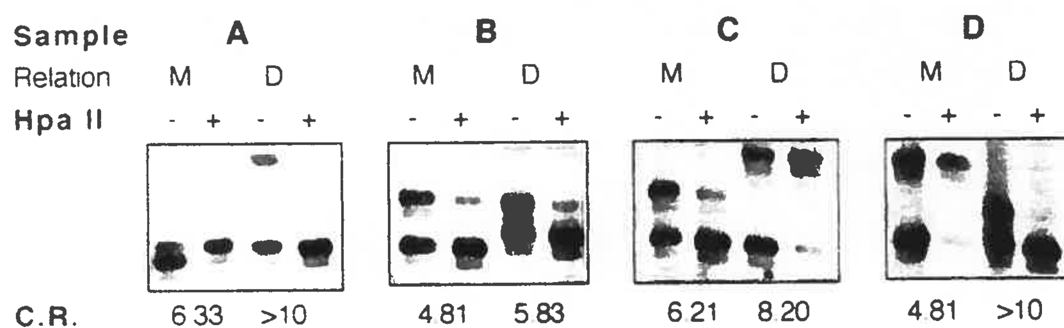


Table I. Possible Causes of Nonrandom X-Inactivation

Developmental Period	Possible Causes of Skewing	Supporting Evidence
Embryogenesis:		
Time of X-inactivation	A. stochastic: random X-inactivation of a small number of cells (embryoblasts) at time of X-inactivation.	- High incidence of skewing in the population (19-21).
	B. genetically determined: i) *cis*-acting element (*Xce* model) (23).	- Heritability of skewing in a large family (37). - Heritability of skewing in mother-daughter pairs (this study).
	ii) selection against cells harboring mutant X-linked allele(s) in the active state.*	- Skewed X-inactivation patterns in female carriers of X-linked mutant alleles (reviewed in 15).
Tissue differentiation	C. small number of cells present at time of X-inactivation and/or unequal partition of progenitor cells contributing to a particular organ.	- Tissue specificity of X-inactivation patterns (21,35,38). - Lineage-specific timing in the initiation of the inactivation process (39).
Post-natal:		
long-term hematopoiesis	D. shrinkage of the pluripotent stem cell pool (environmental ?).	- Stem cell depletion induced clonal dominance (16-18). - Acquired skewing: increased incidence of skewing with age (33).
	E. acquired clonal hematopoiesis (environmental ?).	- Increased incidence of acquired hematologic malignancies with age (40).

* selection against cells harboring active X-linked mutant alleles may occur at any developmental period.

References

1. Lyon MF: Gene action in the X-chromosome of the mouse [*Mus musculus* L.]. Nature 190:372, 1961

2. Russel LB: Mammalian X-chromosome action: inactivation limited in spread and in region of origin. Science 140:976, 1963

3. Cattanach BM: Control of X chromosome inactivation. Annu rev genet 9:1, 1975

4. Brown CJ, Lafreniere RG, Powers VE, Sebastio G, Ballabio A, Pettigrew AL, Ledbetter DH, Levy E, Craig IW, Willard HF: Localization of the inactivation centre on the human X chromosome in Xq13. Nature 349:82, 1991

5. Brown CJ, Ballabio A, Rupert JL, Lafreniere RG, Grompe M, Tonlorenzi R, Willard HF: A gene from the region of the human X inactivation center is expressed exclusively from the inactive X chromosome. Nature 349:38, 1991

6. Brown CJ, Hendrich BD, Rupert JL, Lafrenière RG, Xing Y, Lawrence J, Willard HF: The human *XIST* gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. Cell 71:527, 1992

7. Penny GD, Kay GF, Sheardown SA, Rastan S, Brockdorff N: Requirement for *Xist* in X chromosome inactivation. Nature 379:131 1996

8. Epstein CJ, Smith S, Travis B, Tucker G: Both X chromosomes function before visible X-chromosome inactivation in female mouse embryos. Nature 274:500, 1978

9. Kay GF, Penny GD, Patel D, Ashworth A, Brockdorff N, Rastan S: Expression of *Xist* during Mouse development suggests a role in the initiation of X chromosome inactivation. Cell 72:171, 1993

10. Kay GF, Barton SC, Surani MA, Rastan S: Imprinting and X chromosome counting mechanisms determine *Xist* expression in early mouse development. Cell 77:639, 1994

11. Monk M, Boubelik M, Lehnert S: Temporal and regional changes in DNA methylation in the embryonic, extraembryonic, and germ cell lineages during mouse development. Development 99:371, 1987

12. Migeon BR: Stability of X-chromosomal inactivation in human somatic cells. Nature 239:87, 1972

13. Lyon MF: Sex chromatin and gene action in the mammalian X-chromosome. Am J Hum Genet 14:135, 1962

14. Gilliland DG, Blanchard KL, Bunn HF: Clonality in acquired hematologic disorders. Annu Rev Med 42:491, 1991

15. Belmont JW: Insights into lymphocyte development from X-linked immune deficiencies. Trends in Genet 11:112, 1995

16. Abkowitz JL, Ott RM, Holly RD, Adamson JW: Clonal evolution following chemotherapy-induced stem cell depletion in cats heterozygous for glucose-6-phosphate dehydrogenase. Blood 71:1687, 1988

17. Abkowitz JL, Persik MT, Shelton GH, Ott RL, Kiklevich JV, Catlin SN, Guttorp P: Behavior of hematopoietic stem cells in a large animal. Proc Natl Acad Sci USA 92: 2031, 1995

18. Abkowitz JL, Catlin SN, Guttorp P: Evidence that hematopoiesis may be a stochastic process *in vivo*. Nature med 2:190, 1996

19. Gale RE, Wheadon H, Linch DC: X-Chromosome inactivation patterns using HPRT and PGK polymorphisms in haematologically normal and post-chemotherapy females. Br J Heamatol 79:193, 1991

20. Gale RE, Wheadon H, Linch DC: Assessment of X-chromosome inactivation patterns using the hypervariable probe M27ß in normal hemopoietic cells and in acute myeloid leukemic blasts. Leukemia 6:649, 1992

21. Fey MF, Peter H-J, Hinds HL, Zimmermann A, Liechti-Gallati S, Gerber H, Studer H, Tobler A: Clonal analysis of human tumors with M27ß, a highly informative polymorphic X chromosomal probe. J Clin Invest 89: 1438, 1992

22. Fialkow PJ: Primordial cell pool size and lineage relationships of five human cell types. Ann Hum Genet 37:39, 1973

23. Cattanach BM, Isaacson JH: Controlling elements in the mouse X-chromosome. Genetics 57:331, 1967

24. Johnston PG, Cattanach BM: Controlling elements in the mouse IV. Evidence of non-random X-inactivation. Genet Res 37:151, 1981

25. Cattanach BM, Rasberry C, Evans EP, Dandolo L, Simmler MC, Avner P: Genetic and molecular evidence of an X-chromosome deletion spanning the tabby (Ta) and testicular feminization (*Tfm*) loci in the mouse. Cytogenet Cell Genet 56:137, 1991

26. Rastan S: Primary nonrandom X-inactivation caused by controlling elements in the mouse demonstrated at the cellular level. Genet Res 40:139, 1982

27. Rastan S, Cattanach BM: Interaction between the Xce locus and imprinting of the paternal X chromosome in mouse yolk-sac endoderm. Nature 303:635, 1983

28. Simmler M-C, Cattanach BM, Rasberry C, Rougeulle C, Avner P: Mapping the murine *Xce* with $(CA)_n$ repeats. Mamm genome 4:523, 1993

29. Allen RC, Zoghbi HY, Moseley AB, Rosenblatt HM, Belmont JW: Methylation of *Hpa*II and *Hha*I sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. Am J Hum Genet 51:1229, 1992

30. Busque L, Zhu J, DeHart D, Griffith B, Willman C, Carroll R, McL.Black P, Gilliland DG: An expression based clonality assay at the human androgen receptor locus (HUMARA) on chromosome X. Nucleic Acids Res 22:697, 1994

31. Willman CL, Busque L, Griffith BB, Favara BE, McClain KL, Duncan MH, Gilliland DG: Langerhans'-cell histiocytosis (Histiocytosis X) - a clonal proliferative disease. N Engl J Med 331:154, 1994

32. 11. Busque L, Gilliland DG: Clonal evolution in acute myeloid leukemia. Blood 82:337, 1993

33. Busque L, Mio R, Mattioli J, Brais E, Blais N, Lalonde Y, Maragh M, Gilliland DG: Nonrandom X-inactivation patterns in normal females: Lyonization ratios vary with age. Blood 88:59, 1996

34. Busque L, and Gilliland DG: The PGK-PCR based clonality assay (PPCA), in Adrian J. Harwood (ed): Methods in molecular biology: Protocols for gene analysis, vol 31.Totowa NJ, Humana Press 1994, p 237

35. Gale RE, Wheadon H, Boulos P, Linch DC: Tissue specificity of X-chromosome inactivation patterns. Blood 83:2899,1994

36. Vogelstein B, Fearon ER, Hamilton SR, Preisinger AC, Willard HF, Michelson AM, Riggs AD, Orkin SH: Clonal analysis using recombinant DNA probes from the X-chromosome. Cancer Res 47:4806, 1987

37. Naumova AK, Plenge RM, Bird LM, Leppert M, Morgan K, Willard HF, Sapienza C: Heritability of X chromosome-Inactivation phenotype in a large family. Am J Hum Genet 58:1111, 1996

38. Gartler SM, Gandini E, Angioni G, Argiolas N: Glucose-6-phosphate dehydrogenase mosaicism: utilization as a tracer in the study of development of hair root cells. Ann Hum Genet 33:171, 1969

39. Tan S-S, Williams EA, Tam PPL: X-chromosome inactivation occurs at different times in different tissues of the post-implantation mouse embryo. Nature Genet 3:170, 1993

40. Sandler D: Epidemiology of acute myelogenous leukemia. Semin Oncol 14:359, 1987

41. Puck JM, Steward CC, Nussbaum RL: Maximum-likelihood analysis of human T-cells X chromosome inactivation patterns: normal women vs carriers of X-linked severe combined immunodeficency. Am J Hum Genet 50:742, 1992

42. Nance WE: Genetic tests with a sex-linked marker: glucose-6-phosphate dehydrogenase. Cold Spring Harbor Symp Quant Biol 29:415, 1964

43. Nesbitt MN, Gartler SM: The application of genetic mosaicism to developmental problems. A rev genet 5:143, 1971

44. Gandini E, Gartler SM, Angioni G, Argiolas N, Dell'Acqua G: Developmental implications of multiple tissue studies in glucose-6-phosphate dehydrogenase-deficient heterozygotes. Proc Natl Acad Sci USA 61:945, 1968

Annex 2

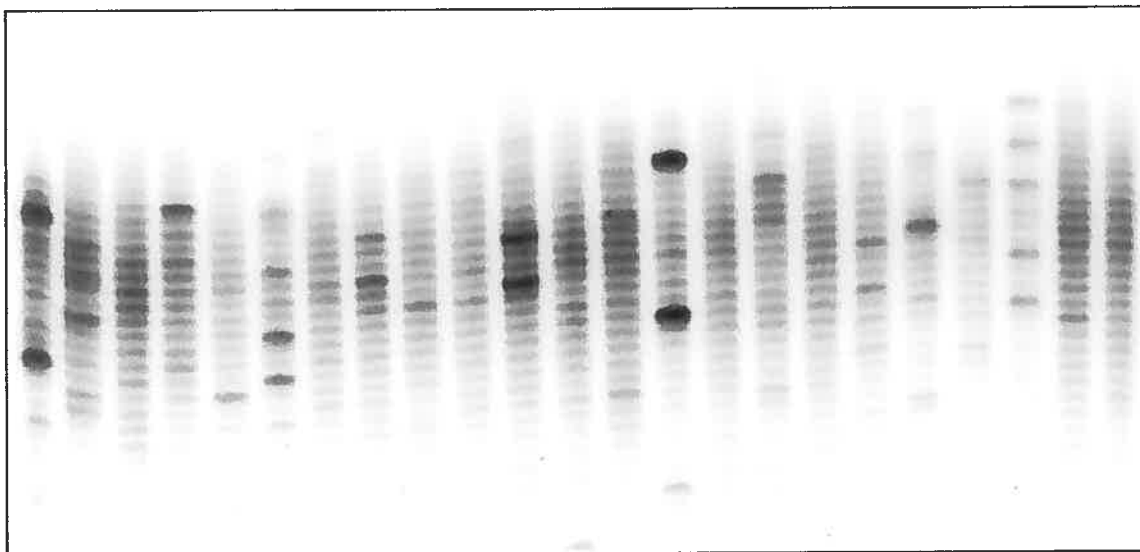
SKEWED Xi AND LYMPHOCYTE CLONALITY

SKEWED Xi AND LYMPHOCYTE CLONALITY

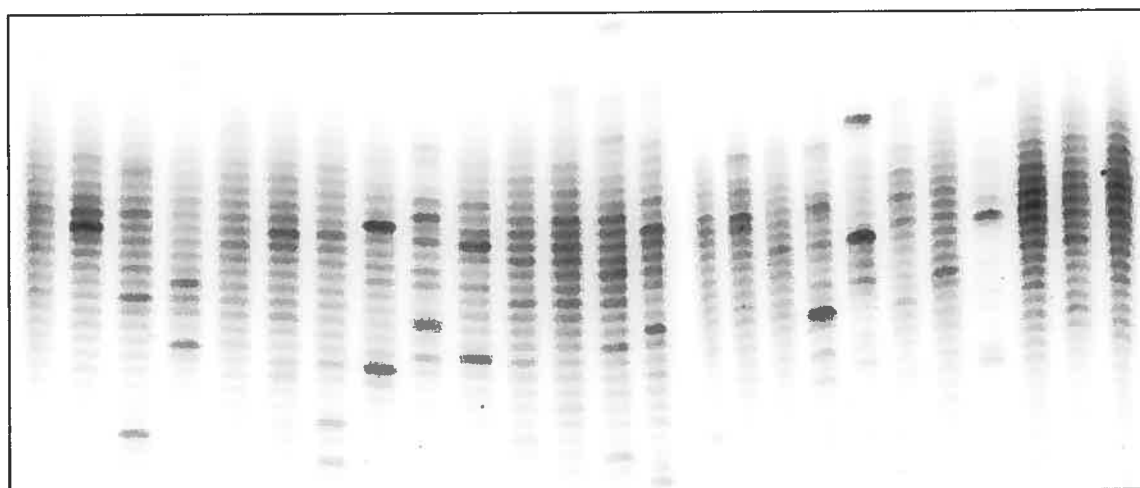
Since the frequency of clonal populations of lymphocytes increases with age (Posnett et al., 1994), we set out to determine whether lymphocyte clonality was contributing to the AS trait. In particular, the clonal status of T and B lymphocytes (as determined by TCR-γ and IgH gene rearrangements) was analyzed as a function of the Xi phenotype (skewed versus non-skewed). For T lymphocyte clonality, T-cell receptor gene rearrangements at the (TCR)-γ chain locus were analyzed (Short et al., 1996). For B lymphocyte clonality, rearrangements of immunoglobulin heavy chain (IgH) genes were analyzed (Achille et al., 1995). ***Methods***: For TCR-γ and IgH gene rearrangements, DNA from 23 Xi-skewed (≥60yo) and 25 non-skewed (≥60yo) females were investigated (Xi skewing status was determined by HUMARA analysis of PMN cells). The criterion to delineate a clonal population was the visual assessment of a single or two predominant amplification products (bands). ***Results***: **TCR-γ gene rearrangement analysis**. As shown in Figure Annex 2-1 a) and b) (page xx), versus the positive controls (Figure Annex 2-1 c) (page xx), clonal populations of T lymphocytes were equally observed in both Xi-skewed and non-skewed females, suggesting T lymphocyte clonality does not contribute to skewed Xi patterns. **IgH gene rearrangement analysis.** As shown in Figure Annex 2-2 a) and b) (page xxi), the majority of females do not support clonal populations of B lymphocytes. Further, the frequency of B lymphocyte clonality was equally distributed among Xi-skewed and non-skewed females, suggesting B lymphocyte clonality does not contribute to skewed Xi patterns. ***Conclusions***: lymphocyte clonality (determined gene rearrangements) does not appear to contribute to a skewed Xi pattern when PMN cells are analyzed. Nonetheless, as the criterion for skewed Xi was based on PMN analysis, a more insightful study may be to analyze lymphocyte clonality in females whom demonstrate a skewed Xi pattern in lymphocytes.

Figure Annex 2-1. T lymphocyte clonality assessed by TCR-γ gene rearrangements.

a) skewed females (≥60 yo) (n=23)



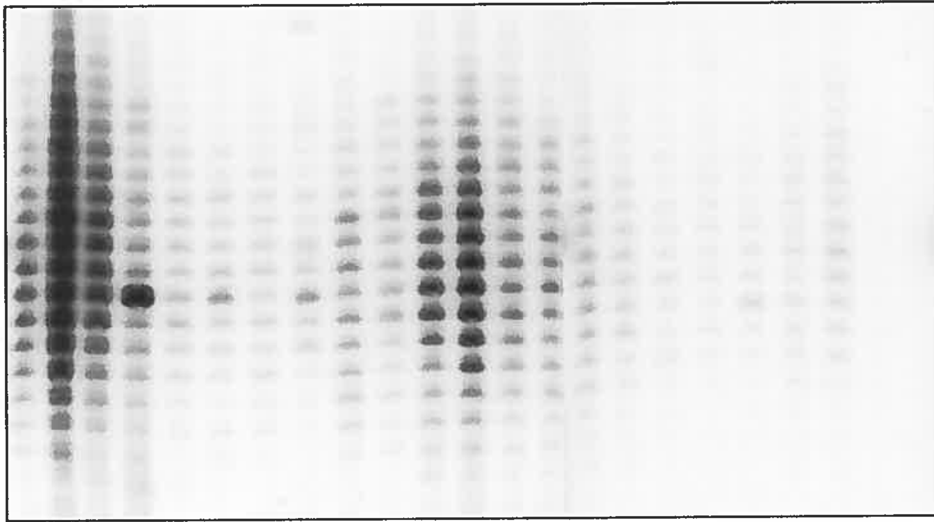
b) non-skewed females (≥60 yo) (n=25)



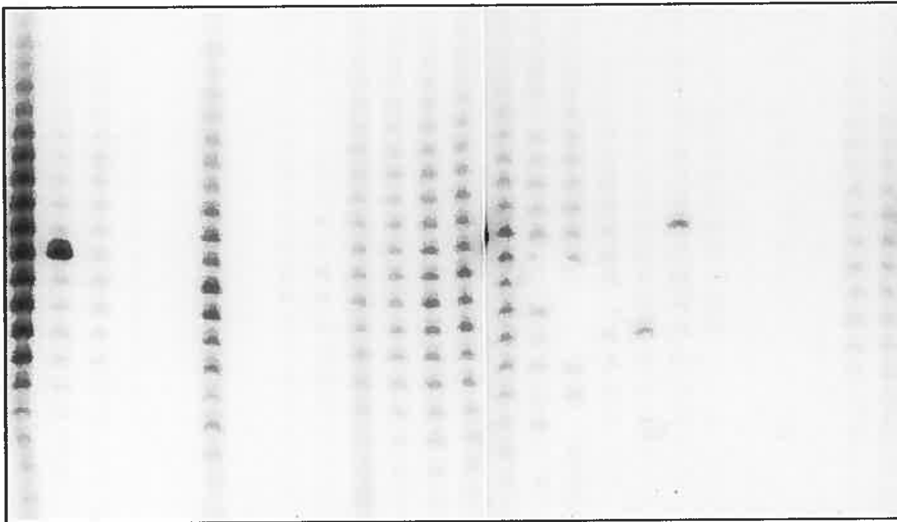
c) positive controls (lymphoma) (n=2)

Figure Annex 2-2. B lymphocyte clonality assessed by immunoglobulin heavy chain gene rearrangements.

a) skewed females (≥60 yo) (n=23)



b) non-skewed females (≥60 yo) (n=25)



c) positive controls (B cell lymphoma) (n=2)

Annex 3

EVIDENCE OF A HSC ORIGIN OF SKEWING IN HUMANS

Hematopoietic stem cell origin of skewing in humans

To determine the cell origin of Xi skewing in hematopoietic lineages, the XIR of progenitor cells was compared to that of more differentiated cell types. ***Methods:*** With informed consent, PB samples of eligible female BM donors were evaluated for Xi status. Based on a skewed Xi pattern, three healthy female volunteers were selected for analyses. A small aliquot of BM aspirate was dedicated for study analysis. Pure populations of committed progenitor cells (CD34+ 38+), less committed progenitor cells (CD34+ 38-), and differentiated cell populations: T cells (CD3+), NK cells (CD56+), B cells (CD20+) and monocytes (CD14+) were obtained by florescent activated cell sorting (FACS). The Xi pattern was determined by the HUMARA clonality assay. ***Results and Conclusion:*** based on visual examination, XIRs of individual lineages within each individual were highly concordant, consistent with derivation from a common progenitor, evidence in favor of a HSC origin of AS. However, since only blood cells were analyzed, a limitation to the study was whether Xi skewing was primary or secondary/acquired in origin.

Annex 4

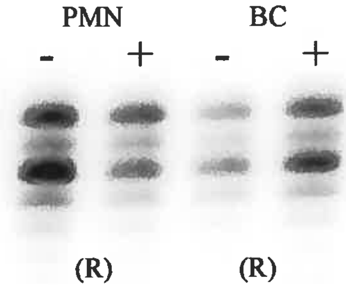
PILOT STUDY: EVIDENCE FOR TWO Xi TRAITS IN ELDERLY HUMAN FEMALES

EVIDENCE FOR TWO Xi TRAITS IN ELDERLY HUMAN FEMALES

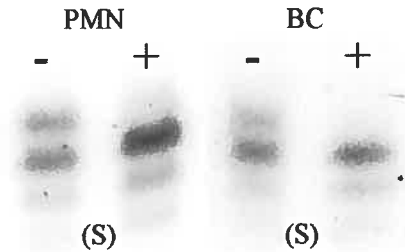
To derive preliminary evidence in favor of two Xi traits in human females, the incidence of skewed Xi and concordance for a skewed Xi pattern among various tissues was analyzed. ***Methods:*** A small cohort of elderly females entering the ambulatory center for complete blood counts were accrued for study enrollment. To obtain females demonstrating the AS trait, elderly females (based on visual appearance) were primarily selected. With written informed consent (page xxviii), two biological samples were obtained from volunteering females (n=52): PB and buccal cells (BC). PB was separated into polymorphonuclear (PMN) and mononuclear (MLL) layers by cellular fractionation (see Chapter 2.5 and 2.6 for biological and cellular fractionation procedures). DNA was isolated and the Xi pattern determined by the HUMARA clonality assay (Chapter 2.9 for methods). Seven subjects were removed from analyses due to non-informativeness (homozygosity for the HUMARA locus), 2 additional samples were excluded due to poor PCR results, rendering a total of 43 informative females. Mean age of participants was 50,7 yo. ***Results and Conclusions:*** 34.9% of females had a skewed Xi pattern in PMN cells and 20,9% in BC samples (see Table Annex 4, page xxvii) for individual subject results). This difference was significantly different (p=0,05). The higher incidence of Xi-skewing in PMN is consistent with the AS trait. A skewed Xi pattern in BC was highly concordant with a skewed Xi pattern in leukocytes (7/9 or 78%), consistent with the properties of the primary Xi trait. In contrast, Xi-skewing in blood (n=15) was less frequently concordant with skewing in BC (7/15 or 47%) (see Figure Annex 4 (page xxvi) for sample results), suggesting that the AS trait is unrelated to the primary Xi trait.

Figure Annex 4. Intraindividual concordance of Xi patterns. a) Random Xi in both PMN and BC, illustrative of a primary Xi pattern. b) Skewed Xi pattern in both PMN and BC, consistent with the properties of a skewed primary Xi pattern. c) Skewed Xi pattern in PMN but random Xi in BC, consistent with the properties of the AS trait.

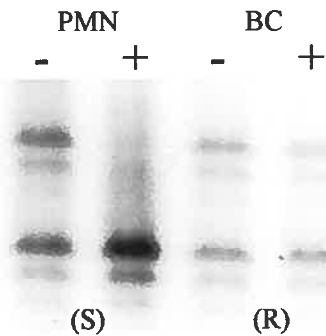
a) Concordant for random X inactivation



b) Concordant for skewed X inactivation: body-wide skewing



c) Discordant for X inactivation patterns: AS



Abbreviations: -, mock digest; + , *Hpa II* digest; R, random X inactivation; S, skewed X inactivation.

Table Annex 4. Xi patterns in polymorphonuclear cells (PMN) versus buccal cells.

Identifier code	X inactivation phenotype		Age
	PMN	Buccal cells	
591	R	R	62
592	R	R	28
594	R	R	49
597	R	R	52
598	R	R	49
600	R	R	48
602	R	R	67
607	R	R	39
610	R	R	46
612	R	R	50
614	R	R	31
616	R	R	77
618	R	R	22
619	R	R	36
620	R	R	24
628	R	R	48
629	R	R	41
647	R	R	34
651	R	R	56
653	R	R	55
655	R	R	78
659	R	R	58
663	R	R	79
678	R	R	72
680	R	R	85
2650 A	R	R	31
587	S	R	79
595	S	R	46
599	S	R	51
609	S	R	59
627	S	R	44
656	S	R	63
719	S	R	20
808 A	S	R	39
601	R	S	43
624	R	S	45
588	S	S	79
605	S	S	65
608	S	S	46
611	S	S	49
617	S	S	47
621	S	S	59
625	S	S	27

Abbreviations: PMN, polymorphonuclear cells, R, random X inactivation, S skewed X inactivation

Centre de recherche
GUY-BERNIER
HMR Hôpital Maisonneuve-Rosemont
Centre hospitalier affilié à l'Université de Montréal

Etude sur le skewing des cellules hématopoiétiques.

Un consentement éclairé a été obtenu pour chacune des patientes enrollées dans l'étude selon les directives du comité d'étique de l'hôpital Maisonneuve-Rosemont. Ces consentements sont conservés dans les filières du Dr. Denis Claude Roy.

Les patients ayant eu un frottis buccal ont une copie de leur leur consentement dans cette filière.

Lambert Busque MD

Pavillon Maisonneuve 5415, boul. de l'Assomption, Montréal, Qc H1T 2M4 Tél. : (514) 252-3400 poste 3329, Télécopieur : (514) 252-3430
Pavillon Rosemont 5689, boul. Rosemont, Montréal, Qc H1T 2H1 Tél. : (514) 252-3400 poste 4959, Télécopieur : (514) 252-3821

FORMULAIRE DE CONSENTEMENT

Dons de sang de sujets normaux

- Nous sommes à la recherche d'échantillons de sang provenant de sujets normaux pour étudier les cellules sanguines. De cette façon, nous croyons pouvoir améliorer notre compréhension sur les leucémies et lymphomes. Seulement des prises de sang sont nécessaires.

- Les leucémies et les lymphomes sont des cancers impliquant des cellules du sang. Nous avons développé des méthodes d'analyse retraçant les gènes que nous croyons impliqués dans le développement de ces cancers. Cependant, nous ignorons si la présence de ces gènes est directement reliée à la maladie ou s'il s'agit, parfois, d'une observation de laboratoire sans conséquence. Pour éclaircir ce point, nous voulons déterminer si ces gènes sont présents chez les individus normaux, et ce en fonction de leur âge. En somme, nos recherches visent à mieux comprendre ces maladies, leur apparition et leur développement.

- En participant à ces études, vous contribuerez à l'avancement des connaissances sur les maladies sanguines telles que les leucémies et les lymphomes.

POUR ÊTRE ADMISSIBLE, IL FAUT:

N'avoir aucune maladie du sang

Ne pas avoir ou avoir eu un cancer.

EFFET SECONDAIRE

Les effets secondaires sont ceux reliés à la prise de sang. Il y aura de la douleur au moment de la prise de sang à cause de la piqûre. Rarement, il peut y avoir une ecchymose (bleu) au site de la ponction, plus rarement encore, il peut y avoir un hématome ou une infection au site de ponction.

TYPES D'ANALYSES

Nous utiliserons les cellules du sang, leur ADN et dérivés afin:

- De caractériser des molécules associées à certains cancers (bcl-2, bcr-abl, ras, c-fms, p53, etc.).

- D'étudier la structure et le comportement des chromosomes (polymorphisme, inactivation du chromosome X, mutation, etc.).

- D'isoler et étudier des cellules actives du sang face à certains cancers (lymphocytes, cellules NK, etc.) .

Bien qu'il n'y aura pas de commercialisation directe ou indirecte des spécimens ou des données génétiques, il est possible que les connaissances acquises lors de ces études puissent éventuellement donner lieu à des applications diagnostiques ou thérapeutiques lesquelles pourraient être commercialisées.

CONFIDENTIALITÉ

Toutes les mesures seront prises pour que les résultats soient gardés confidentiels. Si ces données étaient publiées, l'anonymat sera conservé. Cependant, vous pourrez être informés des résultats d'analyses. De plus, si vous acceptez, il nous sera possible de communiquer avec vous afin d'effectuer des analyses sanguines supplémentaires.

Voici votre choix de type de participation (cocher une case seulement):

☐ J'accepte de donner un échantillon de mon sang pour cette étude. De plus, j'accepte d'être éventuellement contacté et pourrai être informé des résultats par lettre confidentielle.

☐ J'accepte de donner de façon confidentielle un échantillon de mon sang, mais je ne veux pas être contacté et ne serai pas informé des résultats de mes analyses.

☐ Je refuse de participer à cette recherche

____

Signature du sujet NOM (Caractères d'imprimerie)(No. téléphone) DATE

____

Signature du témoin NOM (Caractères d'imprimerie) DATE

Pour de plus amples informations, vous pouvez contacter les docteurs Denis-Claude Roy ou Lambert Busque responsables de ces études, Département d'Hématologie-Immunologie, Hôpital Maisonneuve-Rosemont, au 252-3495.

Annex 5

FORMULAIRE DE CONSENTEMENT

HMR Hôpital Maisonneuve-Rosemont
Centre affilié à l'Université de Montréal

Adresse postale : 5415, boul. de l'Assomption
Montréal (Québec) H1T 2M4

FORMULAIRE DE CONSENTEMENT

Projet :

IDENTIFICATION D'UN GÈNE CONTRÔLANT LA PRODUCTION DES CELLULES SANGUINES, SITUÉ SUR LE CHROMOSOME X.

INTRODUCTION :

Le laboratoire d'hématologie moléculaire du centre de recherche de l'Hôpital Maisonneuve-Rosemont se spécialise dans l'étude des maladies du sang (leucémies, lymphomes) et les greffes de moelles osseuses à l'aide d'analyses basées sur l'inactivation du chromosome X chez les femmes.

Nous avons récemment découvert que près de 40% des femmes normales âgées de plus de 60 ans ont une variation de la normale qui fait que un des deux chromosome X est préférentiellement inactivé au niveau des cellules de la moelle osseuse (hématopoïèse clonale, skewing). Le projet de recherche vise à mieux comprendre ce phénomène et potentiellement à identifier un gène important pour la production des cellules du sang situé sur le chromosome X.

Ce projet nous permettra de mieux comprendre le vieillissement de la moelle osseuse et possiblement les causes des cancers de la moelle osseuse. De plus, il est possible que nous identifions un ou des gènes importants pour la production des cellules sanguines, ce qui pourrait un jour favoriser le développement de nouveaux médicaments.

RECRUTEMENT DE FAMILLES (4 sœurs ou plus, dont l'une âgée de 60 ans ou plus).

Afin de poursuivre nos recherches, nous devons recruter des familles n'ayant pas de maladies du sang. Nous avons besoin de familles où il y a un minimum de 4 sœurs, et dont au moins une des sœurs est âgée de 60 ans ou plus.

POURQUOI SOMMES-NOUS INTÉRESSÉS PAR LES FEMMES ?

Nous utilisons un test basé sur la présence de 2 chromosomes X, c'est pourquoi seulement les femmes peuvent participer à notre étude.

POURQUOI RECRUTER DES FAMILLES ?

Car il est plus facile de suivre la transmission des gènes dans une famille afin d'identifier un ou des gène (s) responsable(s) de l'anomalie des cellules du sang.

COMMENT PARTICIPER À L'ÉTUDE ?

- Être en bonne santé et ne pas avoir de maladie du sang ;
- Ne pas avoir de cancer actif ;
- Avoir 3 sœurs ou plus, dont l'une est âgée de 60 ans ou plus ;
- Demander vous-même à vos sœurs si elles veulent participer à l'étude ;
- Consentir à une prise de sang ;
- Consentir à un frottis buccal (se faire gratter doucement l'intérieur de la joue avec une petite brosse) ;

Une infirmière clinicienne prendra rendez-vous avec vous. Elle fera le prélèvement à votre domicile au moment où vous serez disponible. Ce n'est pas un problème si vos sœurs n'habitent pas la même région que vous, notre infirmière se déplacera pour faire les prélèvements.

RISQUES

Les seuls risques sont ceux d'une prise de sang. Il y aura de la douleur au moment de la prise de sang à cause de la piqûre. Rarement, il peut y avoir une ecchymose (bleu) au site de la ponction, plus rarement encore, il peut y avoir un hématome ou une infection au site de ponction.

TYPES D'ANALYSES:

Nous utiliserons les cellules du sang, leur ADN (acide déoxyribonucléique) et dérivés afin:

- De trouver la cause de l'anomalie de clonalité (skewing) des cellules sanguines chez les personnes de plus de 60 ans en étudiant la structure et le comportement des chromosomes (inactivation du chromosome X, polymorphisme, mutation, etc.).
- De faire des analyses connexes permettant de mieux comprendre les maladies du sang.

CONSERVATION DES ÉCHANTILLONS ET DES DONNÉES SCIENTIFIQUES

Les spécimens et les données seront détruits après 50 ans ou plus tôt selon les directives du sujet. La conservation de l'ADN est sujette aux cas fortuits et à la force majeure.

CONFIDENTIALITÉ :

Toutes les mesures seront prises pour que les résultats demeurent confidentiels et soient gardés au laboratoire de recherche.

Les résultats de cette étude seront éventuellement publiés dans une revue spécialisée de médecine. Toutefois, il n'y aura aucune mention de l'identité des sujets analysés (confidentialité). De plus vous avez le droit de vous retirer de l'étude en tout temps en communicant avec le responsable de l'étude, le Docteur Lambert Busque MD FRCPC au service d'hématologie de l'Hôpital Maisonneuve-Rosemont (252-3400 poste 3741).

Je comprends que je ne recevrai pas d'information médicale provenant de ce projet de recherche.

☐ J'accepte d'être recontactée si le projet requiert d'autres spécimens, pour répondre aux questions d'ordre médical ou encore pour de nouvelles études.
Je comprends qu'il est de ma responsabilité d'informer le responsable de l'étude de tout changement d'adresse.

☐ Je refuse d'être recontactée si le projet requiert d'autres spécimens, pour répondre aux questions d'ordre médical ou encore pour de nouvelles études.

Commercialisation

Il n'y aura pas de commercialisation directe ou indirecte des spécimens ou des données génétiques. Ceux-ci demeureront la propriété de l'Hôpital Maisonneuve-Rosemont. Toutefois, il est possible que les connaissances acquises lors de ces études donnent lieu éventuellement à des applications diagnostiques ou thérapeutiques lesquelles pourraient être commercialisées. Il n'y aura aucune retombée économique pour les participants à l'étude.

Une compagnie, dont le Docteur Lambert Busque est actionnaire minoritaire, détiendra toute propriété intellectuelle éventuelle qui pourrait résulter du présent projet.

ÉVALUATION ÉTHIQUE

Le Comité d'éthique de la recherche de l'Hôpital Maisonneuve-Rosemont a approuvé le présent projet.

☐ Ayant été informée et ayant eu l'occasion de poser mes questions, j'accepte de donner un échantillon de sang pour l'étude de l'inactivation du chromosome X et pour des analyses connexes permettant de mieux comprendre l'inactivation du chromosome X.

☐ Je refuse de participer à cette recherche

____

Signature du sujet Date

____

Nom (caractères d'imprimerie)

____

Adresse

____

____

(no téléphone)

____

Signature du témoin Nom (caractères d'imprimerie) Date

Signature du chercheur : ______________________________________

Pour de plus amples informations, vous pouvez contacter le Dr Lambert Busque MD FRCPC responsable de l'étude au service d'hématologie, Hôpital Maisonneuve-Rosemont ((514) 252-3400 poste 3741) ou Linda Lizotte au (514) 252-3400 poste 4685 (ou au 1-800-726-3403). Advenant un problème avec le déroulement de cette étude vous pouvez joindre une personne indépendante de l'équipe de recherche. La représentante du comité d'éthique est le Dr. Yvette Bonny (514-252-3400).

Annex 6

QUESTIONNAIRE MÉDICAL

QUESTIONNAIRE

1. DATE DE NAISSANCE: ____________________
2. VOTRE MÈRE EST-ELLE VIVANTE ? ☐ OUI ☐ NON
3. VOTRE PÈRE EST-IL VIVANT ? ☐ OUI ☐ NON
4. COMBIEN DE SOEURS AVEZ-VOUS ? __________
5. COMBIEN DE FRÈRES AVEZ-VOUS ? __________
6. COMBIEN DE FILLES AVEZ-VOUS ? __________
7. COMBIEN DE FILS AVEZ-VOUS ? __________
8. AVEZ-VOUS FAIT DES FAUSSES COUCHES ? ☐ OUI ☐ NON

   COMBIEN ? __________
9. FAITES-VOUS DE L'ANÉMIE ? ☐ OUI ☐ NON
10. AVEZ-VOUS DÉJÀ EU UN CANCER ? ☐ OUI ☐ NON
11. PRENEZ-VOUS DES HORMONES? ☐ OUI ☐ NON
12. ÊTES-VOUS FUMEUSE ? ☐ OUI ☐ NON

Famille : ______________________

Individu : ______________________

Paroisse mariage des grands-parents paternels :____________________

Paroisse mariage des grands-parents maternels :____________________

Cause de décès :

Père __________ Frère __________ Fils __________

Mère __________ Sœur __________ Fille __________

Cancer

Quand : ________

Type : ______________________________________

Traitement : radiothérapie ☐ ________________________

Chimiothérapie ☐ ________________________

Si fumeuse (au moment de la prise de sang)

Nb d'année ________

Nb de paquet par jour (au moment de la prise de sang) ________

Si non-fumeuse :

A déjà fumée? Oui☐ Non☐

Si oui Combien de temps? ________

Combien de paquet\jour ________

Arrêt depuis quand? ________

Fumée secondaire : Oui ☐ Non ☐

Si oui

Combien de temps? ________

Arrêt depuis quand? ________

Personne fume combien de paquet par jour en sa présence?________

Souffrez-vous de maladies tel que :

Asthme : ☐

Lupus : ☐

Arthrite rhumatoïde : ☐ (Attention gonflement des articulations et non arthrose)

Liste de tous les <u>médicaments</u> et <u>produits naturels</u> pris au moment de la prise de sang (vitamines, aspirine, pression etc.)

Hormonothérapie : ____

____
____
____
____
____
____
____
____
____
____

Commentaire :

____
____
____
____
____
____
____
____
____
____
____
____