

Organisation of the mammalian genome in the nucleus

Nicola L Mahy

PhD
University of Edinburgh
2001



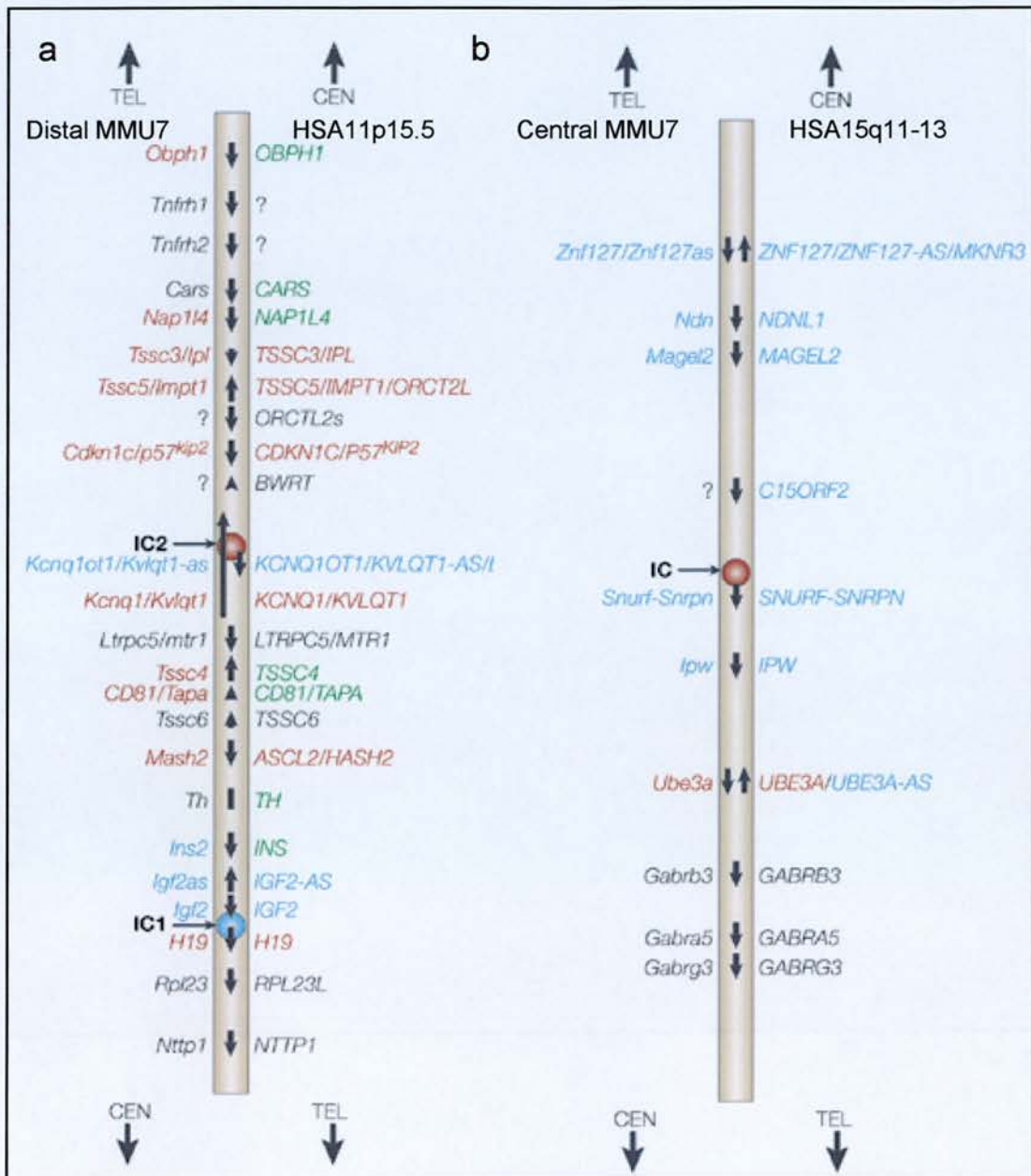


Figure 1.1 Imprinting clusters in human and mouse genomes.

Human chromosomes a) 11p15.5 and b) 15q11-13 and orthologous clusters on a) mouse chromosome distal 7 and b) central 7. The relative and transcriptional orientation of genes are indicated by arrows. The imprinting status is shown in red (maternally expressed), blue (paternally expressed), black (biallelic expression) and green (imprinted expression status not yet known). ? indicates that the orthologues of the mouse or human genes are not known. a) The BWS cluster comprises about 1 Mb, and b) the PWA/AS cluster roughly 2 Mb. Imprinting centres (IC) are marked by circles coloured according to the parental origin of the imprint. (Taken from Reik and Walter, 2001).

“Maps embody a perspective of that which is known and a perception of that which may be worth knowing”.

John Noble Wilford in ‘The Mapmakers’.

ACKNOWLEDGEMENTS

Firstly, thank you to Wendy for getting me started in the lab and for introducing me to the wondrous world of chromosomes. Thank you also to Paul for making this project possible, without images and scripts there would be no data to discuss! Shelagh has been my FISH mentor, and has also taught me appreciation of chocolate, shopping, and wee trips, so thank you to Shelagh for being a great friend, and willing chauffeur! This thesis was also made possible by vast quantities of tea and cake (especially of the chocolate variety), so thank you to everyone who ever provided an excuse for another for cake! A big thank you also to the extended Bickmore lab, to Heidi, Sue, Jon, DJ, Lisa, Nick, Graham, Muriel and Marie, for making the lab a fun place in which to work, and socialise.

Dancing and parties have been a recurrent theme in Edinburgh. Firstly, I must acknowledge the founder-members of rent-a-crowd for dancing and spangly tops. Andie deserves recognition for putting up with me as a flat-mate for our 4 years in Edinburgh. Together we've enjoyed many nights of mischief, and our quest for entertainment led us to become party-organisers extraordinaire. Not that we were recognisable as sensible scientists at any of these events: any excuse to don a wig and various 'props(?)' have meant that our themed parties provided much gossip, for all sorts of reasons...but parties are no fun unless you know fun people to invite, so I'd like to thank all my friends in Edinburgh, especially fellow PhD students, for entertaining evenings full of drinking, dancing (especially Ceilidhing) and dressing-up!

Away from Edinburgh, I want to thank Heidi in London for being a great long-distance friend and travelling companion, and for organising many nights out on the town in exotic locations around the world, about which I wish I could remember more!

Finally, I must thank my family. My sister, for being one of my best friends, and also being wise and studying a subject which takes you to more exotic locations (Paris and South America) than mine (Glasgow and Norwich), and therefore giving me an excuse to leave the world of science (sometimes brightened only by my pink lab coat!) and visit. Last, but unquestionably not least, I acknowledge the unfailing love and support of my parents, who have given me the confidence to achieve my goals, safe in the knowledge that they will always be there for me, for which I am eternally grateful.

ABSTRACT

Transcription factor binding to specific DNA sequences is insufficient to explain the control of gene expression. Attention has turned to large-scale spatial organisation and functional compartmentalisation of chromosomes, and of the nucleus itself, in the quest to understand how the expression of complex genomes is regulated.

Chromosomes occupy discrete domains in interphase and there has been speculation that transcription, RNA processing and RNA transport occur in a space between chromosomes which links to the nuclear pores, the inter-chromosome domain (ICD). In support of this model, studies of a small number of isolated sequences found genes to be positioned at the surface of chromosomes, independent of their transcriptional status. However, nascent transcripts and both early- (gene-rich) and late-replicating (gene-poor) DNA appear to be distributed throughout chromosome territories. Consequently, the ICD compartment model has been modified to include space between the surfaces of subchromosomal domains. Conversely, it has been suggested that differential localisation of gene sequences adjacent to centromeric heterochromatin can facilitate transcriptional repression.

To determine whether genes occupy specific sites within interphase chromosome territories, I have used 2D and 3D fluorescence *in situ* hybridisation (FISH) to investigate the spatial organisation of stretches of DNA from the short arm of human chromosome 11 (HSA11p), which differ in gene content. I found that gene sequences contained in 1Mb of DNA from the moderately gene-rich WAGR (Wilms' tumour, aniridia, genitourinary, mental retardation syndrome) locus of 11p13 are not localised at the periphery of the interphase HSA11p territory as compared to intergenic, non-expressed sequences from the same region. The entire locus is positioned approximately one third of the way along the radius of the HSA11p domain. The organisation of this 1Mb of DNA remains unchanged in cell lines expressing different combinations of genes, indicating that transcriptional status does not play a significant role in the interphase positioning of this locus in human cell lines.

In contrast, I found that DNA from gene-rich 11p15.5 adopts an unusual configuration in the interphase nucleus. Sequences from this region appear to localise outside of the main body of the HSA11p territory. I have examined possible explanations for this extrusion of chromatin from the chromosome domain. It is not characteristic of all chromatin containing telomere

sequences, neither is it characteristic of all loci subject to imprinting. However, the extrusion of chromatin from a chromosome territory appears to be characteristic of gene-rich regions of the human genome. A recent study of the gene-rich MHC locus on chromosome 6 observed looping of chromatin from this region and found its incidence to be transcription dependent.

The publication of the complete sequence of chromosomes HSA21 and HSA22 provided the ideal tool to further test the hypothesis that gene-density plays a role in the organisation of chromatin. Analysis of the interphase position of gene-rich and gene-poor regions across both chromosomes corroborated my hypothesis of chromosomal organisation and also led me to suggest that these gene-rich chromatin loops are specifically orientated towards the centre of the nucleus.

My observations prompt speculation that there is a connection between gene density and the position adopted by chromatin within the nucleus and I suggest that the study of single genes from scattered chromosomal locations is insufficient to determine patterns of organisation; context is an important consideration. This study reinforces the idea that the sequence environment of artificially positioned genes will influence the success of transgenic technologies and may also help to explain the 'position effect' basis of genetic diseases resulting from chromosomal rearrangement.

CONTENTS

PAGE

DECLARATION	ii
ACKNOWLEDGEMENTS	iv
ABSTRACT	v
CONTENTS	vii
LIST OF FIGURES	xiii
LIST OF TABLES AND TEXT BOXES	xvi
ABBREVIATIONS	xviii

CHAPTER 1: INTRODUCTION	1
1.1 The molecular basis of chromatin structure	1
1.2 DNA and histone modifications	2
1.2.1 DNA methylation	2
1.2.2 Core histone modification	3
1.2.2.1 Histone acetylation	5
1.2.2.2 Histone phosphorylation	6
1.2.2.3 Histone methylation	7
1.2.2.4 The histone code hypothesis	9
1.3 Building and remodelling the 30 nm chromatin fibre	9
1.3.1 Methyl CpG binding proteins	10
1.3.2 Acetylated histone binding proteins and transcriptional activation	10
1.3.3 Histone deacetylase-containing complexes and transcriptional repression	11
1.3.4 Su(var) group proteins and establishment of silent chromatin	11
1.3.5 Polycomb proteins and propagation of silent chromatin	13
1.3.6 ATP-dependent chromatin remodelling	14
1.3.7 Boundary elements	16
1.3.7.1 Insulators	16
1.3.7.2 LCRs	18
1.3.8 A role for unconventional RNAs in chromatin structure	19
1.4 Higher-order chromatin architecture	19
1.4.1 Territorial organisation of interphase chromosomes	21
1.4.2 Models of large-scale chromatin folding	22
1.4.2.1 The radial array	22
1.4.2.2 The 'folded chromonema' model	24
1.4.2.3 The random walk/ giant loop model	24
1.4.2.4 The 'multi-loop subcompartment' model	24
1.5 The role of large-scale chromatin structure in transcription regulation	25
1.5.1 Lessons from model, non-mitotic chromosomes	25
1.5.2 Chromatin compaction in X-inactivation	26
1.5.3 Visualising the effects of transcriptional activation on chromatin in live cells	27
1.6 Transcription and the chromosome territory	29
1.6.1 Interphase chromosome subdomains	29
1.6.2 The inter-chromosomal model (ICD)	30

1.6.3	Are active genes displayed on the surface of chromosome territories?	32
1.6.4	The 'modified' ICD hypothesis	33
1.7	The spatial distribution of chromatin in the nucleus and implications for regulation of transcription	34
1.7.1	The role of the nuclear matrix	34
1.7.2	The relative positioning of chromosome territories	37
1.7.2.1	Homologous chromosome alignment in Dipterans	37
1.7.2.2	The nucleolus and acrocentric human chromosomes	37
1.7.2.3	Non-random positioning of human chromosomes within the nucleus	38
1.7.3	Sites of transcriptional repression in the nucleus	39
1.7.3.1	The nuclear periphery	39
1.7.3.2	Centromeric heterochromatin	40
1.7.4	The nuclear interior and transcription	43
1.7.5	Positioning of genes relative to other nuclear bodies	44
1.8	Dynamics of large-scale chromatin organisation	45
1.8.1	Chromatin motion in the nuclei of living cells	45
1.8.2	Cell cycle changes	46
1.8.3	Quiescence and senescence	47
1.8.4	Differentiation	48
1.9	Proposed research	49
1.9.1	Background to proposal	49
1.9.2	The initial proposal	50
1.9.3	Project development enabled by emerging resources	51

CHAPTER 2: MATERIALS AND METHODS **53**

2.1	Mammalian cell culture	53
2.1.1	Cell counting	53
2.1.2	Thawing cells from storage in liquid nitrogen	53
2.1.3	Culture of transformed human cell lines	53
2.1.3.1	Suspension cells	53
2.1.3.2	Monolayer cells	54
2.1.4	Primary human fibroblast culture	54
2.1.5	Mouse embryonic stem cell culture	54
2.1.6	Preparation of mouse embryonic fibroblasts (MEFs)	54
2.2	Preparation of nuclei for fluorescence <i>in situ</i> hybridisation	55
2.2.1	Harvesting and fixing cells in 3:1 methanol:acetic acid	55
2.2.2	Preparation of three-dimensionally preserved nuclei	55
2.3	Preparation of nucleic acids	56
2.3.1	Isolation of RNA	56
2.3.2	Preparation of 1 st strand cDNA	56
2.3.3	Isolation of DNA from cosmid clones	57
2.3.4	Isolation of Bacterial Artificial Chromosome (BAC) and P1 phage-based Artificial Chromosome (PAC) clones	57
2.3.5	Ethanol precipitation of DNA	58
2.3.6	Measuring quality and quantity of DNA and RNA	58
2.3.7	Reverse transcription polymerase chain reaction (RT PCR)	58
2.3.8	Electrophoresis of DNA	60
2.4	Preparation of fluorescence <i>in situ</i> hybridisation (FISH) probes	61
2.4.1	Preparation of labelled mouse and human chromosome paints by PCR	61
2.4.2	Nick translation	62

2.4.3	Removal of unincorporated label	62
2.4.4	Quantifying label incorporation	62
2.5	Fluorescence <i>in situ</i> hybridisation (FISH) on MAA-fixed nuclei	63
2.5.1	Slide preparation	63
2.5.2	Pepsin treatment	64
2.5.3	Hybridisation	64
2.5.4	Washing and detection	65
2.6	FISH on three-dimensionally preserved nuclei	66
2.7	Immunofluorescence	67
2.8	Capture and analysis of 2D images	68
2.9	Capture and analysis of 3D images	68
2.9.1	Image capture and processing	68
2.9.2	Script analysis of 3D images	69
2.9.3	Visualisation of domains within image stacks by 3D surface rendering and movie making to view stacks by rotation through 360°	70
2.10	Statistical analysis	71
 CHAPTER 3: THE ORGANISATION OF 1 MB FROM 11p13 WITHIN INTERPHASE HSA11p		 73
3.1	Introduction	73
3.1.1	The positioning of genes within a chromosome territory	73
3.1.2	Is chromosome territory organisation dependent on transcription?	74
3.1.3	Investigating territory organisation using contiguous sequence from HSA11p and MMU2	74
3.2	The WAGR locus	77
3.2.1	The human WAGR locus	77
3.2.2	The murine WAGR locus	77
3.3	Analysis tools	79
3.3.1	Cell line selection and expression analysis	79
3.3.2	DNA probe selection	79
3.3.3	Image capture	81
3.4	Organisation within an interphase chromosome territory	83
3.4.1	Devising 2D analysis scripts	83
3.4.2	Using the raw data output from script analysis	90
3.5	Sequences across 1 Mb of 11p13 are found in a similar position within HSA11p	91
3.5.1	The organisation of WAGR within a transformed cell line: lymphoblastoid cells	91
3.5.2	The organisation of the WAGR locus is identical in a primary cell line	94
3.5.3	The position of WAGR within HSA11p is not altered by increased transcription from the locus	97
3.6	The spatial organisation of the murine WAGR locus within interphase MMU2 is comparable to that of the syntenic human region in HSA11p	100
3.7	Resolution in 2D MAA-fixed nuclei is sufficient to detect any patterns of localisation across the WAGR locus	104
3.8	Discussion	106
3.8.1	Genes are not located in a different position within a chromosome territory as compared to non-transcribed sequences	106
3.8.2	Gross organisation of interphase chromosomes is conserved from mouse to human	106

3.8.3	Small-scale changes in transcription do not affect gross organisation of interphase chromosomes	107
-------	---	-----

CHAPTER 4: THE CONTRASTING ORGANISATION OF OTHER LOCI FROM HSA11p **109**

4.1	Introduction	109
4.1.1	11p15.5 is very gene dense	109
4.1.2	Human clusters of imprinted genes	112
4.1.3	The murine region of synteny to HSA11p15.5	112
4.2	Localisation of the INSULIN/IGF2/TH genes relative to the HSA11p domain	112
4.3	INS/IGF2 loci adopt a similar position relative to the HSA11p domain in other cell lines	115
4.4	Mapping the extent of the 'looping' structure	117
4.5	Correlation of raw data from this analysis with a previous study of 'looping' chromatin	121
4.6	The HSA11p complex probe comprehensively paints the entire length of the chromosome arm, as visualised at metaphase	125
4.7	The position of other telomeric sequences within interphase chromosome domains	127
4.8	Using a cell line with an inversion within HSA11p to further define patterns of organisation	128
4.9	The localisation of sequences outside of chromosome domains is not characteristic of imprinted regions	132
4.10	Is there a connection between gene-density and the position adopted by chromatin relative to the HSA11p territory?	135
4.11	The organisation of the murine region of synteny to 11p15.5 within MMU7	137
4.12	Discussion	
4.12.1	The unusual configuration of chromatin from 11p15.5 is not just a position effect	141
4.12.2	The unusual configuration of chromatin from 11p15.5 is not common to other imprinted loci	142
4.12.3	Organisation of chromatin within an interphase chromosome may be influenced by gene density	142
4.12.4	The organisation of the syntenic region in the mouse	144
4.12.5	Is interplay between gene density and transcription activity important in organising chromosome territories?	144
4.12.6	Extension of analysis to three dimensions is required	145

CHAPTER 5: THREE DIMENSIONAL ANALYSIS OF THE ORGANISATION OF HSA11p **147**

5.1	Introduction	147
5.2	Development of a 3D analysis technique	148
5.3	The organisation of WAGR in 3D preserved nuclei confirms 2D data	150

5.4 The unusual configuration of 11p15.5 is apparent in 3D but is less pronounced	155
5.5 Gene-poor 11p14 is more internal within HSA11p in 3D analysis	158
5.6 Discussion	160
5.6.1 Genes can be transcribed from within the interior of a chromosome territory	160
5.6.2 Extrusion of gene-rich chromatin from the surface of the HSA11p chromosome territory is apparent in 3D analysis	162
5.6.3 The internal position of 11p14 within HSA11p in 3D analysis provides further evidence for a role for gene density in chromatin organisation within interphase chromosomes	163

CHAPTER 6: DO OTHER REGIONS OF THE HUMAN GENOME EXEMPLIFY PATTERNS OF INTERPHASE ORGANISATION SUGGESTED BY HSA11p? 165

6.1 Introduction	165
6.1.1 Identifying gene-rich human loci	165
6.1.2 Using finished sequence from the human genome: HSA21q and HSA22q	168
6.2 Loci from gene-rich 11q13 and 16p13 are frequently positioned outside of their chromosome territories	169
6.2.1 11q13	169
6.2.2 16p13	171
6.3 Utilising the finished sequence of HSA22q to investigate patterns of interphase chromosome organisation	173
6.4 Using the finished sequence of HSA21q to investigate interphase chromosome organisation	177
6.5 Gene densities of regions of the genome may correlate with the size of chromatin loops extending from the surface of chromosome domains	182
6.6 Discussion	184
6.6.1 Gene density contributes to the organisation of chromatin within a chromosome territory	184
6.6.2 What mechanisms might cause decondensation of chromatin from the surface of a chromosome territory?	186
6.6.3 The decondensed nature of 16p13 correlates with previous descriptions of an open chromatin structure for α -globin genes	188

CHAPTER 7: DISSECTING PATTERNS OF ORGANISATION WITHIN THE NUCLEUS 189

7.1 Introduction	189
7.1.1 The nuclear interior has been associated with gene-rich chromatin	189
7.1.2 Nuclear compartments associated with transcriptional repression	190
7.2 Orientation of the WAGR locus within the nucleus	191
7.3 Determining the nuclear orientation of gene-rich chromatin extruded from HSA11	194
7.4 Chromatin from the HSA11p15.5 conserved region of synteny on MMU7 is also orientated towards the centre of the nucleus	201

7.5 Gene-poor regions from HSA21 and HSA22 are orientated towards the nuclear periphery	203
7.6 Correlating the area occupied by interphase chromosomes with their gene richness	205
7.7 Discussion	211
7.7.1 The nuclear distribution of chromatin with respect to gene density	211
7.7.2 Is there intermingling between chromatin from different chromosomes?	213
7.7.3 The nuclear area occupied by a chromosome territory is related to gene density	213
CHAPTER 8: THESIS OVERVIEW AND DISCUSSION OF FUTURE DIRECTIONS	216
8.1 The human genome—from primary sequence to 3D organisation	216
8.2 Discussion of my observations in relation to current models of nuclear organisation	218
8.2.1 All genes are not preferentially positioned at the surface of a chromosome territory	218
8.2.2 Is the chromosomal or nuclear position of a gene dependent on transcription status?	220
8.2.3 Gene-rich chromatin is preferentially located at the surface of a chromosome 'territory', and sometimes appears outside of the chromosome domain	223
8.2.4 Gene-rich chromatin is preferentially positioned within the nuclear interior	225
8.2.5 How accurate is the concept of a chromosome 'territory'?	226
8.3 Implications of this research	228

FIGURES

PAGE

CHAPTER 1

Figure 1.1	Imprinting clusters in human and mouse genomes	4
Figure 1.2	The histone code hypothesis	8
Figure 1.3	Modulation of insulator activity	17
Figure 1.4	Models of LCR function	20
Figure 1.5	Whole chromosome paints, chromosome arm paints and probes to individual genes	23
Figure 1.6	Dosage compensation in <i>C. elegans</i>	28
Figure 1.7	The inter- <u>chromosome</u> domain model	31
Figure 1.8	The inter- <u>chromatin</u> domain model	35
Figure 1.9	Localisation of silenced genes with heterochromatin: cause or consequence?	42

CHAPTER 3

Figure 3.1	G-banded ideograms of human chromosome 11 (HSA11) and mouse chromosome 2 (MMU2)	76
Figure 3.2	Comparison of the human, murine and <i>Fugu rubripes</i> WAGR loci	78
Figure 3.3	Expression profiles of WAGR genes in human cell lines	80
Figure 3.4	Assessing chromosome paints by FISH to metaphase chromosomes	82
Figure 3.5	Detecting sequences from the human WAGR locus within the territory of chromosome 11p in interphase nuclei	84
Figure 3.6	Analysing an image using the 'cosmid territory' script	87
Figure 3.7	Analysing an image using the 'escaping cosmid' script	89
Figure 3.8	The organisation of the WAGR locus within HSA11p in nuclei of lymphoblastoid cells	93
Figure 3.9	Visualising the organisation of WAGR within HSA11p of primary fibroblasts	95
Figure 3.10	The organisation of the WAGR locus within HSA11p in primary fibroblast nuclei	96
Figure 3.11	Analysis of the WAGR locus in human cell lines expressing <i>WT1</i> or <i>PAX6</i>	98
Figure 3.12	Comparison of the spatial organisation of WAGR in cell lines expressing different combinations of genes from the locus	99
Figure 3.13	Embryonic stem cell nuclei hybridised with an MMU2 chromosome paint and BACs specific to genes of the WAGR locus	102
Figure 3.14	Position of sequences spanning the WAGR locus within interphase MMU2 chromosome of mouse embryonic stem cells	103
Figure 3.15	Combining PAC probes from the WAGR locus to visualise large stretches of DNA within the interphase nucleus	105

CHAPTER 4

Figure 4.1	G-banded ideograms of human chromosome 11 (HSA11) and mouse chromosome 2 (MMU2)	110
Figure 4.2	A map of the genes and markers of HS11p15	111
Figure 4.3	The organisation of INS and IGF2 genes in relation to HSA11p in lymphoblastoid, primary fibroblast, and lens epithelium cell nuclei	114
Figure 4.4	Sequences from distal 11p15.5 are extruded from the HSA11p territory in lymphoblastoid and primary fibroblast cells	119
Figure 4.5	Correlation of qualitative and quantitative 2D analysis of the organisation of 11p15.5 in lymphoblastoid cell nuclei	123
Figure 4.6	The HSA11p-specific paint hybridises to sequences spanning the entire length of the 11p chromosome arm	126
Figure 4.7	Analysis of a lymphoblastoid cell line, CV581, with an inversion within HSA11 (inv(p11.2p15.5))	130
Figure 4.8	FISH analysis of the organisation of the imprinted gene cluster on HSA15q11-13 in lymphoblastoid nuclei	133
Figure 4.9	Analysis of the organisation of the murine BWS locus within MMU7 in embryonic stem cell nuclei	138
Figure 4.10	Comparing the distribution of loci relative to MMU2 and MMU7 chromosomes in ES cell nuclei	140

CHAPTER 5

Figure 5.1	Reconstructing 3D fibroblast nuclei from image plane stacks	149
Figure 5.2	Comparing 2D and 3D analysis of the organisation of the WAGR locus within interphase HSA11p in primary fibroblast nuclei	152
Figure 5.3	Comparing the distributions of signals from WAGR within HSA11p in 2D and 3D analysis of primary fibroblasts	154
Figure 5.4	Reconstruction of a 3D fibroblast nucleus showing INS/IGF2 at the surface of HSA11p	157
Figure 5.5	The distributions of signals from 11p15 within HSA11p in 2D and 3D analysis of primary fibroblasts	159

CHAPTER 6

Figure 6.1	The distribution of CpG islands in the human genome	166
Figure 6.2	G-banded ideograms of human chromosomes 6,11,16,21 and 22	167
Figure 6.3	The spatial organisation of 11q13 in two and three dimensions	170
Figure 6.4	The spatial organisation of 16p13.3	172
Figure 6.5	Sequence location and spatial organisation at interphase of gene-rich and gene-poor regions from HSA22q	175
Figure 6.6	The spatial distribution of gene-rich and gene-poor regions of HSA22q within lymphoblastoid cell nuclei	176
Figure 6.7	Sequence location and spatial organisation at interphase of gene-	

	rich and gene-poor regions from HSA21q	180
Figure 6.8	The spatial distribution of gene-rich and gene-poor regions of HSA21q within lymphoblastoid cell nuclei	181
Figure 6.9	Correlation between gene density and intra-chromosome locus position	183

CHAPTER 7

Figure 7.1	Determining the orientation of WAGR within the nucleus	193
Figure 7.2	'Relative erosion' script for assessing the relative positions of the chromosome territory and a locus within the interphase nucleus	196
Figure 7.3	Erosion analysis shows that the tail of chromatin extruded from the Surface of the HSA11p territory orientates towards the centre of the nucleus	198
Figure 7.4	Erosion analysis comparing loci from three regions of HSA11p	200
Figure 7.5	Erosion analysis of mouse chromosomes MMU2 (Wt1) and MMU7 (BWS region)	202
Figure 7.6	Erosion analysis shows that gene-rich loops of chromatin from HSA21q and HSA22q orientate towards the centre of the nucleus	204
Figure 7.7	Correlation of gene density with the area occupied by a chromosome territory within the nucleus	209

TABLES

PAGE

CHAPTER 1

Table 2.1	Primers for amplification of human WAGR gene sequences	59
Table 2.2	Human WAGR RT-PCR programmes	60
Table 2.3	Antibodies and fluorochrome conjugates used for FISH	67

CHAPTER 3

Table 3.1	Comparing the position of sequences from WAGR relative to the chromosome periphery in different cell lines	92
Table 3.2	Mean intra-chromosomal positions of loci spanning murine WAGR within nuclei of mouse embryonic stem cells	101

CHAPTER 4

Table 4.1	Position of INS/IGF2 relative to the interphase HSA11p domain	113
Table 4.2	Patterns of INS/IGF2 signals relative to the interphase HSA11p domain within lymphoblastoid cells	116
Table 4.3	Mean positions of loci spanning more than 9 megabases of distal HSA11p relative to the HSA11p territory	118
Table 4.4	The frequency with which a sequence from HSA11p15.5 locates inside of the HSA11p domain correlates with the mean distance of the signal from the nearest edge of the chromosome territory	122
Table 4.5	Location of telomere sequences in relation to their respective interphase chromosome domains	128
Table 4.6	Comparison of the position of loci from 11p15 in a cell line containing an inversion within HSA11p, and a control cell line	129
Table 4.7	Mean positions of sequences from imprinted regions of the genome within their respective chromosome domains	134
Table 4.8	Mean position of sequences from 11p14 relative to WAGR (11p13) in lymphoblastoid cells	136

CHAPTER 5

Table 5.1	The position of sequences from WAGR relative to the HSA11p chromosome periphery of primary fibroblasts in 2D and 3D	151
Table 5.2	The position of sequences from 11p15 relative to the HSA11p chromosome periphery of primary fibroblasts in 2D and 3D	156

CHAPTER 6

Table 6.1 Positions of BACs corresponding to gene-rich and gene-poor regions of chromosome 22, relative to the interphase HSA22q territory	174
Table 6.2 Positions of BACs corresponding to gene-rich and gene-poor regions of chromosome 22, relative to the interphase HSA22q territory	179

CHAPTER 7

Table 7.1 Using PACs spanning the locus in 11p13 to determine whether WAGR is specifically orientated within the nucleus	192
Table 7.2 Calculating a mean chromosome territory radius	206
Table 7.3 Correlation between the gene density and area of a chromosome territory	208

TEXT BOXES

PAGE

CHAPTER 3

Box 3.1 The 'cosmid territory' script used to determine the position of a locus signal within a chromosome domain in 2D	86
Box 3.2 The 'escaping cosmid' script used to determine the distance of a locus signal outside of a chromosome domain in 2D	88

CHAPTER 7

Box 7.1 'Relative erosion' script for determining the relative 2D positions of a locus and its chromosome domain within the nucleus	195
--	-----

ABBREVIATIONS

2D	Two dimensional
3D	Three dimensional
A	adenine (purine)
ATP	adenosine triphosphate
BAC	bacterial artificial chromosome
BrdU	5-bromo-2-deoxyuridine
BSA	bovine serum albumin
BWS	Beckwith-Weiderman syndrome
C	cytosine (pyrimidine)
C-terminal	carboxy-terminal
CBP	CREB-binding protein
CTCF	CCCTC-binding factor
CCD	charged couple device camera
<i>C. elegans</i>	<i>Caenorhabditis elegans</i>
CHRAC	<i>chromatin accessibility complex</i>
chromodomain	<i>chromatin organisation modifier</i>
CpG	cytosine and guanine
CSD	chromo shadow domain
DAPI	4, 6-diamidino-2-phenylindole
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
dGTP	deoxyguanosine triphosphate
Dig	Digoxigenin
<i>Drosophila</i>	<i>Drosophila melanogaster</i>
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethyl sulphoxide
DNA	deoxyribose nucleic acid
DNase	dexyribonuclease
DNMT	DNA methyl transferase
dNTP	deoxynucleoside triphosphate
DMR	differentially methylated region
DTT	Dithiothreitol

dUTP	deoxyuridine triphosphate
EM	Electron microscopy
ES cells	Embryonic stem cells
FCS	Foetal calf serum
FISH	fluorescence in situ hybridisation
FITC	fluorescein isothiocyanate fluorochrome
FRAP	fluorescence recovery after photobleaching
G	guanine (purine)
G-band	chromosome pattern produced by Giemsa staining
G1	growth phase 1 of the cell cycle (pre-replication)
G2	growth phase 2 of the cell cycle (post-replication)
GC	guanine and cytosine
GFP	green fluorescent
HAT	histone acetyltransferase
HCl	hydrochloric acid
HDAC	histone deacetylase
HDM	histone demethylase
HMT	histone methyltransferase
<i>INS/IGF2</i>	cosmid containing Insulin, IGF2 and TH genes
kb	kilobase pairs of DNA
KCl	potassium chloride
LBR	lamin B receptor
LCR	locus control region
M	molar
MAA	methanol:acetic acid
MAR	matrix attachment region
Mb	megabase pairs of DNA
MBD	methyl binding domain
MeCP	Methyl binding protein
MEF	Mouse embryonic fibroblast
mRNA	messenger RNA
N-terminal	amino-terminal
NaCl	sodium chloride
NOR	nucleolar organising region
NURF	<i>nucleosome remodelling factor</i>

OD	optical density
p	chromosome short arm
PAC	P1 phage-based artificial chromosome
PBS	phosphate buffered saline
PCAF	p300/CBP-associated factor
PcG	Polycomb group
PCR	polymerase chain reaction
PEV	position effect variation
pFA	para-Formaldehyde
PML	promyelocytic leukaemia
PWS/AS	Prader-Willi/Angeman syndrome
q	chromosome long arm
R-band	chromosome pattern produced by reverse Giemsa staining
rDNA	rRNA-encoding DNA
RNA	ribose nucleic acid
RNase	ribonuclease
rRNA	ribosomal RNA
RNP	ribonucleoprotein
RSTS	Rubenstein-Taybi syndrome
RT	reverse transcriptase
RT PCR	reverse transcription polymerase chain reaction
S-phase	synthesis phase of the cell cycle (DNA replication)
SAGA	Spt-Ada-Gcn5 acetyltransferase
SEM	Standard error of the mean
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SDS	sodium dodecyl sulphate
SMC	structural maintenance of chromosomes
<i>S. Pombe</i>	<i>Schizosaccharomyces pombe</i>
SSC	standard saline citrate (1x: 150 mM NaCl, 15 mM tri-Na Citrate, pH 7.4) T thymidine (pyrimidine)
T-band	telomeric band (extreme form of R-band)
TAE	Tris, EDTA acetic acid buffer (1x 90 mM Tris-HCl, 90 mM glacial acetic acid, 2 mM EDTA pH8.0)
TE	Tris, EDTA buffer (10 mM Tris, 1 mM EDTA, pH8.0)
TEM	Transmission electron microscopy

TR	Texas red PBS
trxG	trithorax group
UV	ultra violet
WAGR	<u>W</u> ilm's tumour, <u>a</u> niridia, <u>g</u> enitourinary anomalies, mental <u>r</u> etardation
Xa	active X chromosome
Xi	inactive X chromosome
<i>Xenopus</i>	<i>Xenopus laevis</i>

CHAPTER 1

INTRODUCTION

1.1 THE MOLECULAR BASIS OF CHROMATIN STRUCTURE

In the nuclei of diploid human cells two copies of 3.3×10^9 bp of DNA, each measuring roughly 2m when extended, is folded so as to occupy a sphere barely 15 μm in diameter. To achieve this feat of organisation, DNA is packaged with proteins into a hierarchical structure called chromatin and the highest level of packaging yields a compaction ratio of $>20,000:1$ in terms of the ratio of linear, B-form DNA to the length of the fully compacted metaphase chromosome. The nucleosome and chromatin fibre provide the common structural framework for transcriptional control in eukaryotes and the folding of DNA within these structures can both promote and impede transcription.

The universal repeating element of chromatin is the nucleosome, containing 146 bp DNA (in human cells) and a core histone protein octamer (pairs of H2A, H2B, H3 and H4) (Finch *et al.*, 1977; reviewed (Belmont *et al.*, 1999)). The four core histone molecules are among the most evolutionarily conserved eukaryotic proteins known. However, the N-terminal tails of core histones are not as structured as the C-terminal histone fold domains and make more flexible contacts with DNA and adjacent nucleosomes that allow for dynamic changes in the accessibility of the underlying genome (section 1.2.2).

Nucleosomal arrays *in vitro* adopt an extended 10 nm diameter conformation or condense into a compact 30 nm fibre depending on the ionic strength of the medium. It was originally thought that formation of the chromatin fibre depended critically on the presence of linker histones e.g. H1, chromatin-associated proteins that bind to the exterior of nucleosomes and dramatically stabilise the highly condensed states of chromatin fibres, (Thoma *et al.*, 1979). Two principal concepts of fibre architecture were developed: solenoids in which the linker DNA continues the supercoil established in the nucleosome, and zigzag models in which the linker crosses the fibre (Widom, 1998; Zlatanova *et al.*, 1999; Thomas, 1999). However, chromatin can fold to the highest levels of compaction in the absence of H1 (Carruthers *et*

al., 1998) and neither chromosome condensation, nor nuclear assembly, are affected by the absence of linker histones in *Xenopus* extracts (Dasso *et al.*, 1994; Ohsumi *et al.*, 1993) or *Tetrahymena* (Shen *et al.*, 1995). These data argue against a causal role for histone H1 in these processes and there is evidence to suggest that core histone tails rather than linker histones are critical for both local and global chromatin condensation (Carruthers and Hansen, 2000), (de la Barre *et al.*, 2000).

Despite decades of intensive study, only limited information is available for the structure of the 30nm fibre. Electron microscopy (EM) techniques have not been able to resolve higher-order chromatin structures above the level of the 'beads on a string' conformation of nucleosomes and DNA (Kornberg and Lorch, 1999). One exception to this is the lampbrush and polytene chromosomes of Dipterian insects (section 1.6). No single model has been substantiated but new techniques are enabling ultrastructural visualisation of a complex 3D chromatin architecture (Gobbi *et al.*, 1999) (Woodcock and Horowitz, 0 AD; Woodcock and Dimitrov, 2001) and also 3D chromatin architecture in relation to other nuclear components, such as DNA polymerase alpha (Lattanzi *et al.*, 1998)

1.2 DNA AND HISTONE MODIFICATIONS

It has been demonstrated that chromatin folding, at least in vitro, should be viewed in terms of a dynamic equilibrium between compaction levels (Hayes and Hansen, 2001). Nucleosomes and chromatin fibres, are not structurally inert entities, but undergo conformational changes such as dissociation, histone modification and allosteric transitions which are likely to contribute to transcriptional control (review, Wolffe and Guschin, 2000) and the regulatory flexibility of other processes such as replication, recombination and repair. These changes are related to modifications of chromatin structure at the level of DNA sequence and histones.

1.2.1 DNA methylation

Methylation of the cytosine of CpG dinucleotides is involved in the control of gene expression and is essential for mammalian development (Li *et al.*, 1992; Okano *et al.*, 1999). Three DNA cytosine methyltransferases (DNMT1, -3a and -3b) have been identified (Bestor

et al., 1988; Okano *et al.*, 1998) and these epigenetic 'markers' on DNA can be copied after replication, resulting in heritable changes in chromatin structures. DNA methylation is normally associated with gene silencing, (Bird and Wolffe, 1999) and is also associated with X-inactivation and genomic imprinting (Ng and Bird, 1999).

In vertebrates, 60-90% of all CpGs are methylated, leaving a minor part of the genome methylation free. Many of the remaining non-methylated CpGs (about 15% of all CpGs in human DNA) are found in CpG islands which usually include functional promoters, and are predominantly found in the early replicating (R-bands) of the genome (Craig and Bickmore, 1994).

DNA methylation has a role in mammalian X-inactivation where transcription from one copy of the X chromosome of females is almost completely silenced (Section 1.5.2). Similarly, methylation is also associated with genomic imprinting, a phenomenon where the monoallelic expression pattern of a gene is dependent on whether it has been inherited maternally or paternally (Reik and Walter, 2001). Imprinted genes are frequently found in clusters in both the human and mouse genome (Figure 1.1), which contributes to the control of gene expression (Section 1.3.7.1). All imprinted genes studied so far have differentially methylated regions (DMR) in which the maternal and paternal DNA copies are methylated differently (Neumann *et al.*, 1995), therefore, methylation is suspected to be the signal from the parental germ cells that results in allele-specific expression. However, 7 of 18 imprinted genes studied have methylated DMRs on the active allele (<http://www.mgu.har.mrc.ac.uk/imprinting/imptables.html#impgene>). Promoter methylation may also contribute to immobilisation of mammalian transposons, suppression of transcriptional noise, and the control of tissue-specific gene expression, but decisive evidence for these ideas is lacking (Jones and Takai, 2001).

1.2.2 Core histone modification

The packaging of genomic DNA with histone proteins into chromatin provides a potent obstacle for other protein-DNA interactions, and chromatin has to be remodelled to permit enzymatic events like transcription, replication, DNA repair or recombination ((Workman and Kingston, 1998; Wolffe and Guschin, 2000). The structure of chromatin is dynamic and localised decondensation and remodelling facilitates or inhibits access of other proteins to DNA. Post-translational modifications of histone proteins play a significant role in

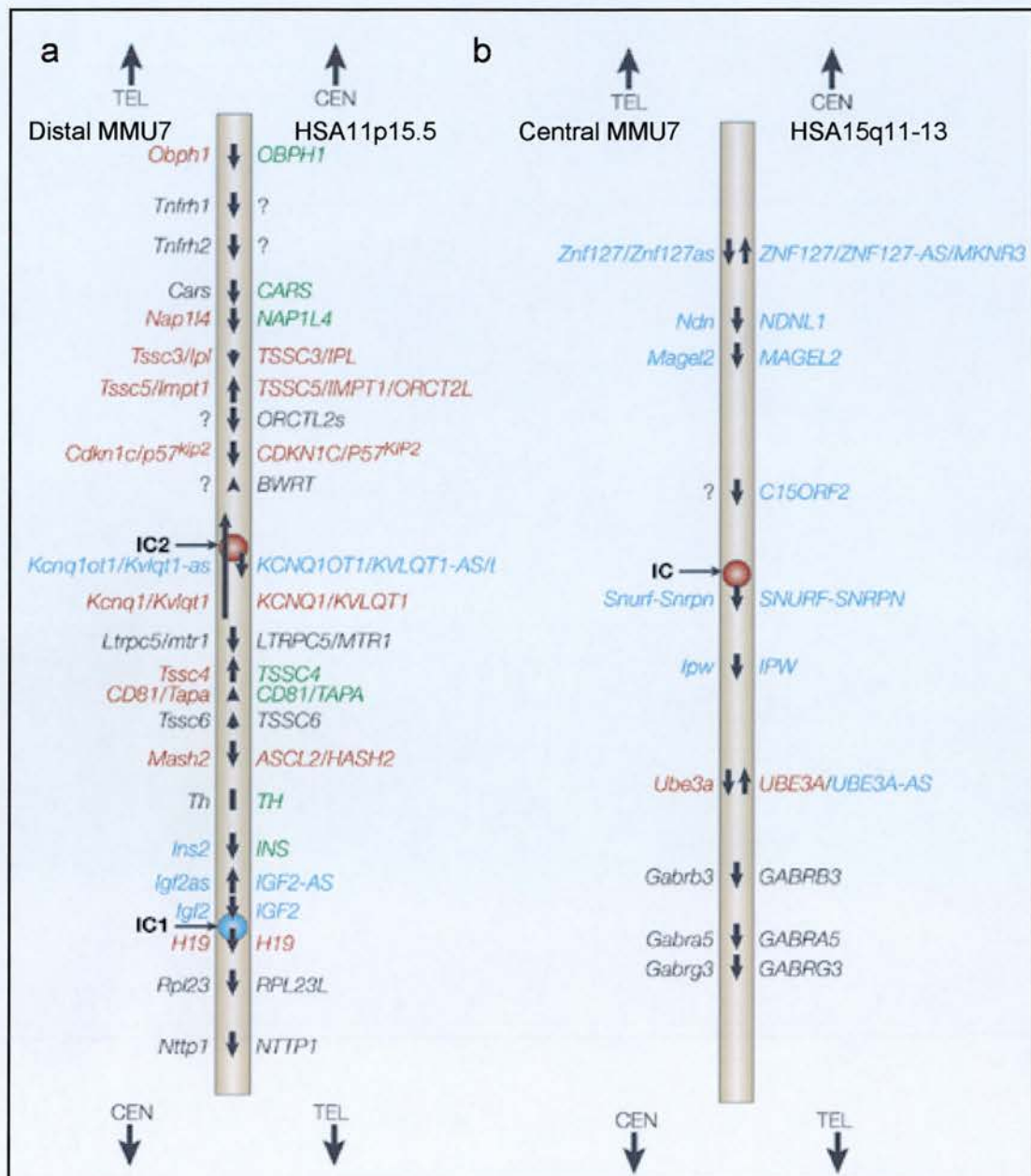


Figure 1.1 Imprinting clusters in human and mouse genomes.

Human chromosomes a) 11p15.5 and b) 15q11-13 and orthologous clusters on a) mouse chromosome distal 7 and b) central 7. The relative and transcriptional orientation of genes are indicated by arrows. The imprinting status is shown in red (maternally expressed), blue (paternally expressed), black (biallelic expression) and green (imprinted expression status not yet known). ? indicates that the orthologues of the mouse or human genes are not known. a) The BWS cluster comprises about 1 Mb, and b) the PWA/AS cluster roughly 2 Mb. Imprinting centres (IC) are marked by circles coloured according to the parental origin of the imprint. (Taken from Reik and Walter, 2001).

regulating nuclear function. These modifications include acetylation, methylation ubiquitination and phosphorylation. The most extensively studied sites of modification are contained in the well-defined very basic N-terminal domains, which extend through the DNA gyres into the space surrounding the nucleosomes.

1.2.2.1 Histone acetylation

Histone acetylation and deacetylation has long been linked to transcriptional activity (Allfrey VG *et al.*, 1964) and has been the most intensely studied histone modification. Acetylated histones are associated with transcriptional activity, and hypoacetylated histones with repression of transcription. In cultured mammalian cells, heterochromatin is deficient in acetylated H4 relative to euchromatin (O'Neill and Turner, 1995). Similarly, the inactive X in female mammals is underacetylated (Jeppesen and Turner, 1993) and the hyper-active X in male *Drosophila melanogaster* larvae is more highly acetylated on Lys16 than female X chromosomes or autosomes (Bone *et al.*, 1994).

A molecular link between acetylation and transcription was first provided when the conserved transcriptional regulator Gcn5 was found to have histone acetyltransferase (HAT) activity (Brownell *et al.*, 1996). Conversely, the yeast transcriptional repressor Rpd3 was found to be a histone deacetylase (HDAC), and the related protein HDAC1 was the first human HDAC to be discovered. Subsequently a wealth of other proteins with HAT and histone deacetylase activity have subsequently been identified (reviewed in Sterner and Berger, 2000). The exact mechanism by which acetylation affects the biophysical properties of chromatin remains unclear, but the acetylation of the core histone N-termini reduces the stability of interaction of the histone tails with nucleosomal DNA, facilitating the binding of transcription factors to their recognition elements (Hansen *et al.*, 1998; Wolffe and Hayes, 1999; Morales and Richard-Foy, 2000). However, recent data suggest that acetylation does not have the anticipated large inhibitory effects on histone-DNA interactions and stimulatory effect on transcription factor access (reviewed in Marmorstein, 2001) and is perhaps required for maintenance of the active state, but not its establishment (Cavalli and Paro, 1999). This, and data discussed in the rest of section 1.2.2, has led to the idea that the biological function of the histone tails might be considerably more complex. The idea that acetylation of histone tails might provide epigenetic marks by which the degree of gene activity is maintained from one cell generation to the next was first put forward by Bryan Turner (Turner, 1998)

(discussed in section 1.2.2.4). Recent data indicate a role for histone acetylation in signalling for other chromatin regulatory and transcriptional events.

A role for acetylation in X-inactivation has been determined: hyperacetylation of chromatin upstream of *Xist* facilitates the promoter switch that leads to stabilisation of the *Xist* transcript and the subsequent deacetylation of this region is essential for the further progression of X inactivation (O'Neill *et al.*, 1999). Underacetylation of the inactive X is also important in the stabilisation of the inactive state (Keohane *et al.*, 1998).

In addition to a role in higher order chromatin packing, histone acetylation is also associated with genome replication and nucleosome assembly, higher-order and interactions of non-histone proteins with nucleosomes (reviewed in Grant, 2001).

1.2.2.2 Histone phosphorylation

The significance of histone phosphorylation is less well understood than acetylation, but has been implicated in processes as diverse as transcription, DNA repair, apoptosis and chromosome condensation (reviewed in Cheung *et al.*, 2000). A role for histone phosphorylation in conflicting nuclear processes such as transcription activation and chromosome condensation may appear paradoxical. However, in all cases, histone phosphorylation may serve to open up the chromatin fibre to allow access of nuclear factors that promote progression of a regulated nuclear mechanism; downstream factors may serve to either condense or decondense chromatin.

At present it is not clear how H3 phosphorylation affects gene expression, but links between chromatin structure and increased levels of phosphorylation have indicated an opening of the chromatin structure. Histone phosphorylation may act in a manner analogous to that of histone acetylation: addition of a negatively-charged group to an N-terminal histone tail may perturb chromatin structure either by disruption of electrostatic interactions between the basic H3 tails and the negatively charged DNA backbone, or by serving as recognition sites for recruitment of transcription factors or regulatory complexes.

Kinases that act on histones have recently begun to be identified. A newly isolated histone kinase Snf1, is a previously known transcription factor that functions at genes regulated by acetylation (Lo *et al.*, 2001). Phosphorylation of histone H3 has been implicated the

immediate-early (IE) response to mitogens. This results in rapid transcription of a subset of genes in the nucleus, which require no new transcription or translation for their induction (elicited by an intracellular MAP kinase signalling pathway), via the kinase MSK1 (Thomson *et al.*, 1999). The proportion of H3 targeted by MSK1 is also especially susceptible to hyperacetylation (Clayton *et al.*, 2000), which indicates that phosphorylation and acetylation of histone H3 may act synergistically.

1.2.2.3 Histone methylation

Histones H3 and H4 have long been known to be substrates for methylation, but the biological significance of this covalent modification was unknown. Proteins responsible for this methylation activity, histone methyltransferases (HMTs), have only recently been identified and the significance of this modification for signalling downstream events to build chromatin structures is becoming increasingly apparent. Importantly, unlike the dynamic 'on-off' nature of histone acetylation, the majority of data available suggests that the methyl modification is relatively irreversible (Byvoet *et al.*, 1972), so the existence of a global histone demethylase (HDM) seems unlikely.

Recently it was shown that the human and *S. Pombe* homologues of the *Drosophila* heterochromatin protein Su(var)3-9 (SUV39H1 and Clr4, respectively) are H3-specific lysine methyltransferases, providing evidence for the involvement of histone methylation in transcriptional regulation (Rea *et al.*, 2000), reviewed in (Zhang and Reinberg, 2001; Rice and Allis, 2001) (see Section 1.3.4). A nuclear receptor coactivator-associated protein, CARM1, is an H3-specific arginine methyltransferase (Chen *et al.*, 1999). Similar to the apparent conflicting role of histone phosphorylation, so histone methylation also induces divergent transcriptional outputs. The SU(VAR)3-9 protein family seems to mediate transcriptional silencing (section 1.3.4) as methylated H3Lys9 is a binding site for HP1, thus directly linking histone modification to formation of higher-order chromatin structure. Paradoxically, methylation of Lys4 of H3 has also been correlated with transcriptionally active nuclei in *Tetrahymena* (Strahl *et al.*, 1999), but as suggested by the histone code hypothesis (Strahl and Allis, 2000) (Section 1.2.2.4), it probably depends on which residues are modified (Figure 1.2).

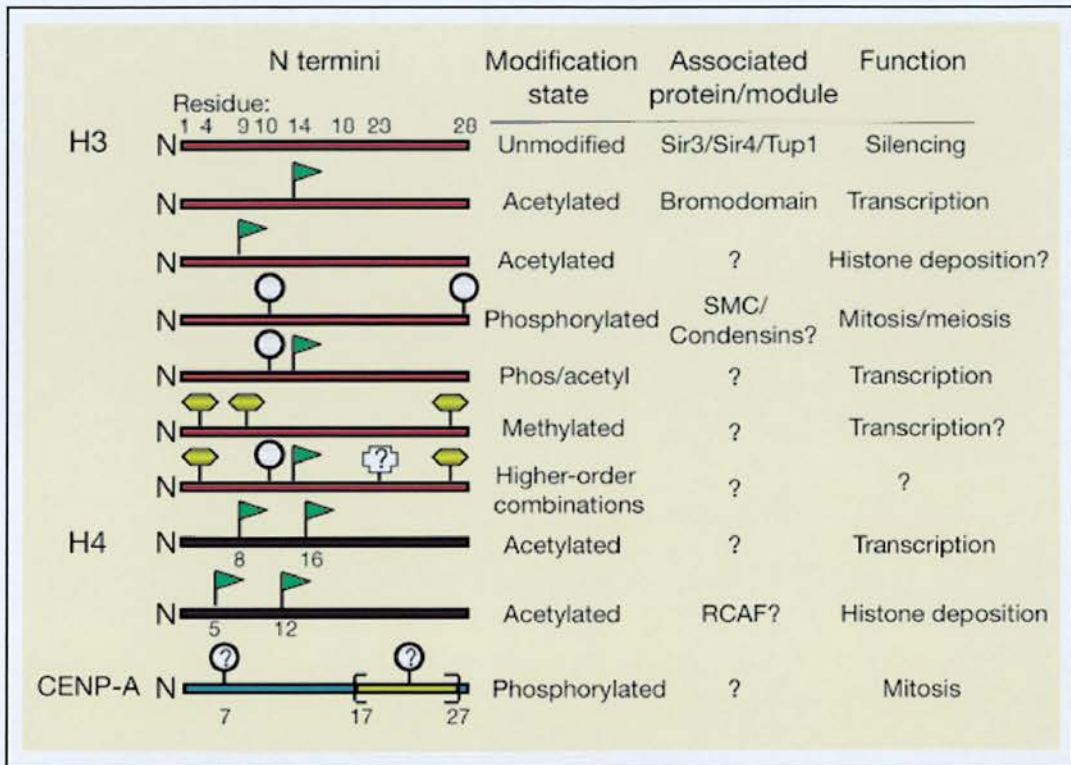


Figure 1.2 The histone code hypothesis

Histone modifications occur at selected residues, and some of the patterns shown have been closely linked to a biological event (for example, acetylation and transcription). Emerging evidence suggests that distinct H3 (red) and H4 (black) tail modifications act sequentially or in combination to regulate unique biological outcomes. How this hierarchy of multiple modifications extends (depicted as 'higher-order combinations'), or how distinct combinatorial sets of modifications are established or maintained in localised regions of the chromatin fibre is not known. Relevant proteins or protein domains known to interact or associate with distinct modifications are indicated. The CENP-A tail domain (blue) might also be subjected to mitosis-related marks such as phosphorylation; the yellow bracket depicts a motif in which serines and threonines alternate with proline residues. (Taken from Strahl and Allis 2000).

1.2.2.4 The histone code hypothesis

Both the molecular mechanisms of action and biological functions of the core histone N-termini in chromatin appear to be complex (discussed in Hansen *et al.*, 1998). At the level of the chromatin fibre, multiple distinct functions of the N-termini are required to achieve alterations in the level of higher order chromatin condensation, some of which apparently involve protein-protein rather than protein-DNA interactions. Therefore, the global structure and function of any given region of the genome may be determined predominantly by the core histone N-termini and their specific interaction partners.

Appreciation of the importance of other core histone modifications led to the extension of the idea that histone acetylation is important in marking histones to signal downstream transcription events (Turner, 1998; Turner, 2000), and suggestion of the ‘histone code hypothesis’ (Strahl and Allis, 2000). This model predicts that multiple histone modifications, acting in a combinatorial or sequential fashion on one or multiple histone tails, specify unique downstream functions. Such modifications serve as marks for the recruitment of different proteins or protein complexes to regulate diverse chromatin functions, such as gene expression, DNA replication and chromosome segregation i.e. other proteins ‘read the histone code’ and bind accordingly.

1.3 BUILDING AND REMODELLING THE 30 NM CHROMATIN FIBRE

As discussed above, modification of DNA and histones serves to ‘mark’ chromatin and signal for downstream mechanistic events. Proteins bind to nucleosomes as a consequence of these modifications, and subsequent layers of protein–DNA and protein–protein interactions ensue. Recruited proteins may have enzymatic activities that serve to modify the histone code further, or may elicit chromatin remodelling, such as ATP-dependent nucleosome remodelling complexes. Other proteins with the ability to read the histone code are structural proteins that bind to build higher-order chromatin structures.

1.3.1 Methyl CpG binding proteins

A family of methyl-CpG binding proteins, which are thought to mediate the biological consequences of DNA methylation, has been identified and characterised (Hendrich and Bird, 1998). This family is characterised by a methyl-binding domain (MBD), of which the founder member is the transcriptional repressor MeCP2. Repression by MeCP2 can act at a distance from a promoter and involves the targeting of protein complexes that contain histone deacetylases (section 1.2.2.2) (Jones *et al.*, 1998; Nan *et al.*, 1998), resulting in a block to transcription initiation. MeCP2 is mutated in Rett syndrome (Hendrich and Bickmore, 2001), demonstrating the importance of interpreting DNA methylation signals. Targeting of histone deacetylases is probably also important in transcriptional repression by other family members MBD1 and MBD2 (Ng *et al.*, 1999; Ng *et al.*, 2000).

1.3.2 Acetylated histone binding proteins and transcriptional activation

After the identification of yeast Gcn5 as a HAT, human HATs including GCN5, the orthologous human protein PCAF, p300/CBP and TAF250 were subsequently found in multisubunit complexes that function at promoters. These complexes are thought to act locally, gene by gene, to modulate transcription. Biochemical and genetic evidence support a role for histone acetylation in transcription regulation via increased access of transcription factors to DNA through structural changes in nucleosomes or nucleosomal arrays. This is probably achieved by acetylases acting synergistically with ATP-dependent chromatin remodelling proteins (section 1.3.6), rather than as a direct consequence of histone acetylation (reviews Wu and Grunstein, 2000; Sterner and Berger, 2000; Roth *et al.*, 2001). The bromodomain found in transcription factors and HATs such as TAFII250, PCAF and GCN5 is a protein domain that allows for the direct recognition of histone tails when they are acetylated at specific lysine residues (Winston and Allis, 1999; Dhalluin *et al.*, 1999; Owen *et al.*, 2000; Jacobson *et al.*, 2000). A further HAT, CREB-binding protein (CBP) that can acetylate all the core histones in addition to proteins such as p53 and components of the basal transcription machinery, is mutated in Rubenstein-Taybi syndrome (RSTS) (Hendrich and Bickmore, 2001). This provides further evidence for the role of chromatin components that serve to silence or activate gene expression, in human genetic disorders.

1.3.3 Histone deacetylase-containing complexes and transcriptional repression

The link between histone deacetylation and transcription repression first became apparent when a mammalian histone deacetylase, HDAC1, turned out to be related to a known yeast transcriptional repressor RPD3 (Taunton *et al.*, 1996). Inhibition of HDAC activity was subsequently shown to disrupt the formation of highly condensed heterochromatic regions in yeast (Ekwall *et al.*, 1997). The yeast silencing protein Sir2 has been found to have NAD-dependent HDAC activity, and the deacetylation of lysine 16 of H4 seems important for the interaction of such silencing proteins and subsequent spreading of heterochromatin (Imai *et al.*, 2000; Landry *et al.*, 2000; Smith *et al.*, 2000).

Like HATs, HDACs do not work alone and exist in multiprotein complexes, e.g. the mammalian Sin3 complex, which comprises at least seven subunits including HDAC1, HDAC2 (Zhang *et al.*, 1998). Many mammalian proteins have been identified that repress via the Sin3 complex (reviewed in Ng and Bird, 2000), and among the best-understood are the nuclear hormone receptors (eg SMRT and NcoR) which, in the absence of hormone, bind to specific promoters and repress transcription (reviewed in Xu *et al.*, 1999).

1.3.4 Su(var) group proteins and establishment of silent chromatin

Constitutive heterochromatin is genetically inert, traditionally defined as the portion of a chromosome that remains condensed throughout the cell cycle. It consists of highly repetitive DNA and, in the human genome, is found at the centromeres of every chromosome, on the long arm of the Y and at the pericentric regions of chromosomes 1, 9 and 16 (Miklos and John, 1979).

Facultative heterochromatin describes euchromatin that assumes the characteristics of heterochromatin in a local, developmentally controlled manner, therefore eliciting temporary silencing of parts of the genome. Thus, heritable but reversible changes in gene expression can occur without alterations in DNA sequence, the best example of which is X-inactivation. Epigenetic inheritance of an inactivated state can be propagated during mitosis and through the germ-line during meiosis (Grewal and Klar, 1996; Cavalli and Paro, 1998). This has been termed position-effect variegation (PEV) and is implicated in human disease (Kleinjan and van Heyningen, 1998). Genetic screens in *Drosophila* (Reuter and Spierer, 1992) and *S.*

pombe (Thon and Klar, 1992; Allshire *et al.*, 1994) have identified 30-40 loci involved in modifying PEV. Among the modifier genes identified in these model systems, one subclass suppresses variation (the Su(var) group) and includes chromatin-associated components, best characterised by the heterochromatin protein HP1 (Su(var)2-5 in *Drosophila*) (Eissenberg *et al.*, 1990) as well as gene products such as HDACs and protein phosphatases (PPTases).

The discovery that a Su(var) protein (*Drosophila* Su(var)3-9) is a histone lysine (H3 Lys9) methyltransferase (as are its homologues SUV39H1 in humans, Suv39h1 in mouse, and Clr4 in *S. Pombe*), provided a critical link between histone methylation and heterochromatin assembly (Rea *et al.*, 2000). At the carboxy-terminal, Su(var)3-9 protein family members have a SET (Su(var)3-9, Enhancer of zeste, Trithorax) domain, also found, as the name suggests, in other chromatin-associated proteins, including members of both the PcG and trxG (sections 1.3.5 and 1.3.6). The heterochromatin protein HP1 is a potential methyl-histone binding protein. Both HP1 and its *S. Pombe* homologue Swi6 are required for heterochromatin formation and colocalise with SUV39H1 and Clr4 respectively (Eissenberg and Elgin, 2000; Ekwall *et al.*, 1996; Aagaard *et al.*, 1999). These, and many other heterochromatin-associated proteins, share an evolutionarily conserved domain known as the chromodomain (*chromatin organisation modifier*, also found in members of both the PcG and trxG (Sections 1.3.5 and 1.3.6)) at the amino-terminal (Eissenberg and Elgin, 2000). Some also have a chromo shadow domain (CSD) (Jones *et al.*, 2000). It has recently been shown that the CD but not the CSD of HP1 preferentially binds to methylated H3 Lys9 *in vitro* (Lachner *et al.*, 2001; Bannister *et al.*, 2001), but the CSD is essential for HP1 function. Thus, the chromodomain of the heterochromatin-associated protein HP1 targets the Lys9 of histone H3 for heterochromatin assembly and gene silencing. Combining this and other data, a pathway for heterochromatin assembly leading to epigenetic silencing is beginning to emerge (reviewed in Rice and Allis, 2001; Marmorstein, 2001).

An important question is how extensively this packaging mechanism is used on a single gene scale in facultative as well as constitutive heterochromatin. A recent study suggests a role for the SUV39H1-HP1 complex in Rb-mediated gene repression (Nielsen *et al.*, 2001). It has been shown that Rb is necessary to direct methylation of histone H3 by SUV39H1 at the cyclin E promoter, and is necessary for binding of HP1 to the cyclin E promoter to facilitate silencing.

Just as the histone code hypothesis is supported by the fact that the acetyl modification can serve to recruit bromodomain-containing HATs and other coactivators to chromatin for transcription, it appears that a similar conserved pathway may exist for histone methylation and heterochromatin assembly via chromodomain-containing proteins.

1.3.5 Polycomb proteins and propagation of silent chromatin

Although the transcriptional ground state of chromatin is generally repressive to transcription, in certain situations it is necessary to maintain and extend this repression in development. *Polycomb* (*PcG*) and *Trithorax* (*trxG*) group proteins have been shown to regulate the expression of homeotic genes in *Drosophila* and the *PcG/trxG* maintenance system is evolutionarily conserved from *Drosophila* through to mammals and homeotic transformations have been observed in mutant mice (Akasaka *et al.*, 2001). This system is thought to regulate many target genes in addition to homeotic genes, indicating that epigenetic maintenance of activated or repressed states might be a fundamental developmental mechanism (reviewed in Pirrotta, 1998). *PcG* proteins do not bind DNA directly, but have preferred sites of binding at the PRE (*Polycomb response element*). The simplest model for how *PcG* proteins recognise transcriptionally silent states is that molecular features of the chromatin in repressed genes are compatible with, or stabilise, assembly of *PcG* complexes (reviewed in Francis and Kingston, 2001). Two general models have been suggested for how discrete sites of *PcG* protein binding act over large distances to communicate repression to the promoter. These include direct contact by looping, or by spreading of the structure to the promoter. *PcG*-mediated repression may also involve recruitment to a nuclear *PcG* domains, either near the nuclear periphery (Gerasimova *et al.*, 2000) or with centromeric heterochromatin (Saurin *et al.*, 1998; reviewed in Francastel *et al.*, 2000; Gasser, 2001) (see section 1.7.3). In addition to repression through direct interactions with chromatin, there is evidence that links histone modifications to *PcG* function: hyperacetylation might be involved in excluding *PcG*-mediated repression (Cavalli and Paro, 1999). The mouse *PcG* protein *Eed* interacts with histone deacetylases, and has recently been implicated in the stable maintenance of imprinted X-inactivation in extra-embryonic tissues (Wang *et al.*, 2001).

1.3.6 ATP-dependent chromatin remodelling

The *trxG* genes were defined as suppressers of *PcG* phenotypes and participate in regulatory complexes that act at the level of the chromatin template to maintain active gene expression and to counteract *PcG* repression. Several *trxG* proteins have been directly linked to transcriptional activation. *MLL*, a mammalian homologue of *trx* has been implicated in transcriptional maintenance in mammalian development and is a common target of chromosomal translocations in human acute leukaemias (Yu *et al.*, 1998). There are also *trxG* proteins in SWI/SNF-related chromatin-remodelling complexes, e.g. the *trx* gene *Brm* encodes the DNA-dependent ATPase in the *Drosophila* SWI/SNF complex and the equivalent ATP-utilising subunit of human SWI/SNF is BRG1 or BRM1 (Kal *et al.*, 2000).

A similar class of proteins antagonising PEV modifiers to enhance variegation were also identified as the (E(var) group) (Reuter and Spierer, 1992), which counteract the Su(var)-induced silent chromatin state. Several E(var) gene products were also identified as components of ATP-dependent nucleosome-remodelling machines (Tsukiyama and Wu, 1997; Kal *et al.*, 2000), which target histone tails and are thought to act synergistically with histone acetyltransferases to establish a chromatin structure permissive for subsequent events (Fry and Peterson, 2001).

ATP-dependent chromatin remodelling complexes have been separated into two families that are defined by the different features of their ATPase subunits. The 'SWI/SNF' family all contain a protein related to the yeast protein Swi2/Snf2, the founding member of this conserved family (Kingston and Narlikar, 1999), whereas the 'ISWI' family contain an ATPase related to the *Drosophila* ISWI protein.

To date, very little work has been carried out at a molecular level to examine how individual genes or chromosome loci are packaged within the larger chromatin masses of differentiated cells, and whether changes in chromatin condensation directly influence gene regulation during development (Peterson and Workman, 2000). One of the best characterised examples is the yeast HO promoter (Cosma *et al.*, 1999; Krebs *et al.*, 1999), where SWI/SNF is targeted to the gene through an interaction with the transcription factor Swi5p. SWI/SNF action then results in the recruitment of the transcriptional complex SAGA, which has HAT activity that can acetylate chromatin over 1kb. SWI/SNF-dependent SAGA (Spt-Ada-Gcn5 acetyltransferase) recruitment is observed most frequently for genes expressed during

mitosis, suggesting a role for SWI/SNF in unfolding condensed chromatin (Krebs *et al.*, 2000). If genes targeted by SWI/SNF late in mitosis include *trx-G*-regulated genes, this activity could contribute to re-establishing active states after mitosis. Studies suggest that SWI/SNF acts prior to HATs, and that SWI/SNF action is required for establishing the active state of the gene, and may be continually required to maintain transcription of their target genes (Biggar and Crabtree, 1999). Maintenance of transcription may therefore require continual SWI/SNF-directed remodelling, or SWI/SNF might be necessary only after relatively rare events that result in gene inactivation (e.g. condensation of chromatin in mitosis), depending if the 'ground state' is permissive for transcription.

Biochemical assays utilising a *Drosophila in vitro* chromatin assembly system were established independently to recapitulate different aspects of chromatin dynamics, and allowed the identification and purification of distinct chromatin remodelling and assembly factors. Each of these contained the ISWI ATPase protein, originally identified by sequence similarity to Brahma, the *Drosophila* counterpart of yeast SWI2/SNF2 and include NURF (*nucleosome remodelling factor*), ACF (*ATP-utilising chromatin assembly and remodelling factor*), and CHRAC (*chromatin accessibility complex*) (Varga-Weisz and Becker, 1998). Yeast (ISWI1 and ISWI2) and human (hSNFL and SNF2H) counterparts have since been identified and complexes are currently being characterised.

Biochemical studies have shown that an activated nucleosome formed by a chromatin-remodelling complex can meet three fates: histone octamer transfer to another DNA molecule, sliding of the nucleosomes along DNA, or generation of other remodelled intermediates (Flaus and Owen-Hughes, 2001; Lorch *et al.*, 2001). There is also growing evidence for a role for ATP-dependent remodelling activities in the repression of transcription, possibly at the level of chromatin fibre formation (Tyler and Kadonaga, 1999). The NuRD (nucleosome remodelling and histone deacetylase) complex has both ATP-dependent remodelling and HDAC activity, as well as methyl-CpG-binding activity via MBD3, suggesting that as with HAT proteins, HDAC activity is mechanistically linked to chromatin remodelling.

The human protein ATRX is a member of the SNF2 family of helicase/ATPases (Gibbons *et al.*, 1995; Gibbons *et al.*, 1997), which suggests that ATRX may act as a transcriptional regulator by directly affecting chromatin structure. The protein has been shown to interact with a murine homologue of HP1 (mHP1 α) in a yeast two-hybrid screen and is localised to pericentromeric heterochromatin during interphase and mitosis (McDowell *et al.*, 1999). At

metaphase, some ATRX protein is also localised at the ribosomal DNA (rDNA) arrays on the short arms of human acrocentric chromosomes and interestingly, mutations in ATRX give rise to changes in the pattern of methylation of several highly repeated sequences including the rDNA arrays, a Y-specific satellite and subtelomeric repeats (Gibbons *et al.*, 2000). This protein therefore provides a link between the processes of chromatin remodelling, DNA methylation and gene expression in mammalian development. Mutations in the *ATRX* gene are linked to the genetic disorder ATR-X syndrome, which may be result from a defect in the organisation of repressive chromatin (Hendrich and Bickmore, 2001).

In addition to ATRX, HP1 has also been shown to interact with transcriptional regulators, such as the lymphoid differentiation-specific transcription factors Ikaros (see section 1.7.3.2) (Brown *et al.*, 1997; Lamond and Earnshaw, 1998). In turn, Ikaros been linked to complexes containing components of the Swi/Snf family of chromatin remodelling complexes and HDACs -1 and -2 (Kim *et al.*, 1999), indicating a role for differential histone acetylation and chromatin remodelling in promotion of differentiation.

1.3.7 Boundary elements

1.3.7.1 Insulators

Insulators mark the boundaries of chromatin domains by limiting the range of action of enhancers and silencers. These elements are defined in an enhancer blocking assay that measures the effect of placing an element between an enhancer and promoter in transgenic fruit flies (Kellum and Schedl, 1992). The connection between enhancer blocking activity and chromatin boundary function was made after two elements *scs* and *scs'*, that appeared to mark the ends of a chromatin domain at the *Drosophila hsp70* locus, were found to function as positional enhancer blockers (Kellum and Schedl, 1992). Insulator activity can be modulated by nearby regulatory elements, bound cofactors or covalent modification of the DNA. The *Drosophila* retrotransposon *gypsy* was also found to be an insulator and its activity resides in binding sites for the protein suppressor of Hairy wing (*su(Hw)*) (Gerasimova and Corces, 1998). The *gypsy* insulator was also found to contain an additional protein, *Mod(mdg4)*, which can modulate the activity of the insulator: mutations in *mod(mdg4)* have properties characteristic of a *trxG* gene (Figure 1.3a). In addition, mutations in *trxG* genes augment insulator effects on adjacent enhancers, whereas mutations in *PcG* genes have the opposite result. A model in which *PcG* and *trxG* proteins regulate

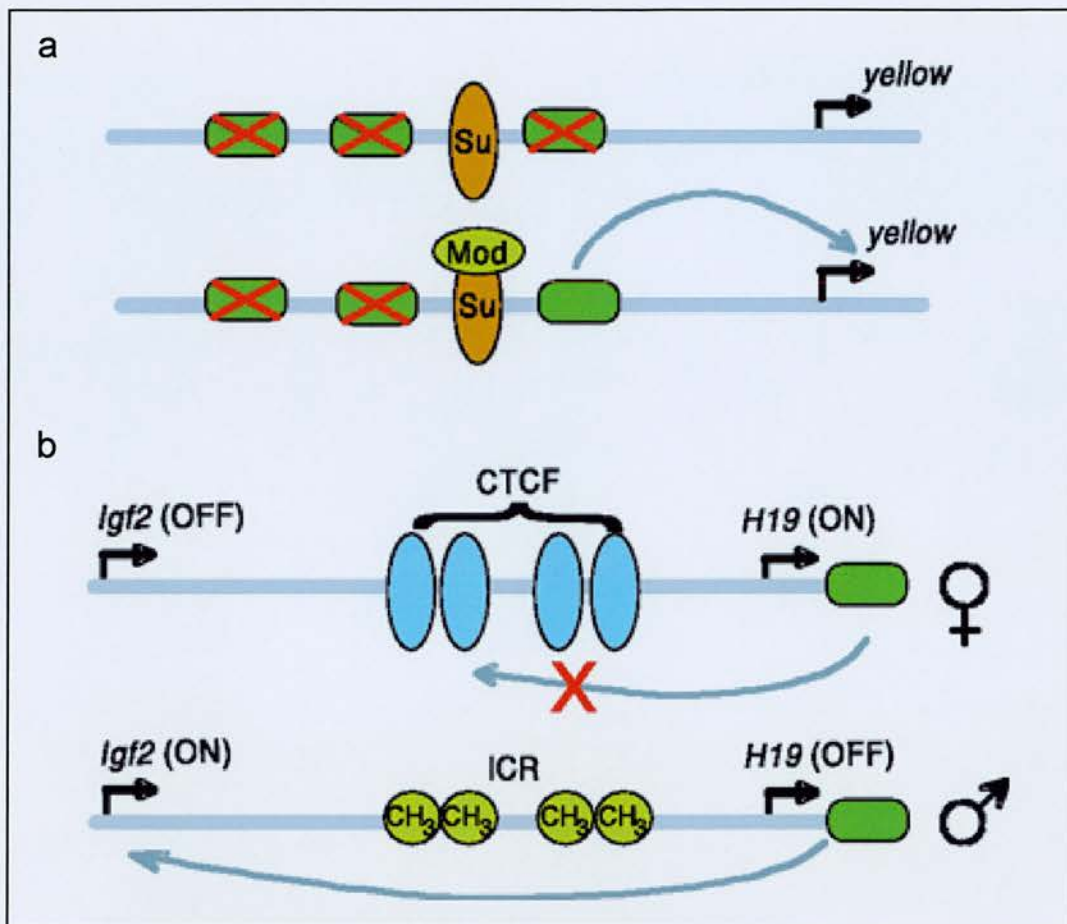


Figure 1.3 Modulation of insulator activity

The insulators are shown as ellipses.

a) When Su(Hw) binding sites (orange) in the *gypsy* retrotransposon are inserted between enhancers (green) at the *Drosophila yellow* locus, the effect in the absence of the protein product of *mod(mdg4)* is to silence expression from all enhancers (top). Positional enhancer blocking at this locus requires the modifying interaction of *mod(mdg4)* protein with Su(Hw) at this site. With *mod(mdg4)* present, only enhancers upstream of the insulator are blocked, and the *yellow* gene is transcribed.

b) The mouse *Igf2/H19* locus contains four binding sites (blue) for the protein CTCF (seven in humans). CTCF binding is associated with known insulator elements. On the maternally inherited allele, the CTCF sites serve to block the action of a downstream enhancer (green) on the *Igf2* promoter. Consequently, *Igf2* is not expressed from the maternal allele. On the paternal allele, the CTCF sites are methylated, and CTCF does not bind, inactivating the insulator. The enhancer is now free to activate *Igf2* expression from the paternal allele. (Inactivation of *H19* expression on the paternal allele is controlled by a separate mechanism). (Taken from Bell *et. al.*, 2001)

insulator function by establishing domains of higher order chromatin organisation has been suggested and it is thought that chromatin insulators might regulate gene expression by controlling the subnuclear organisation of DNA (see section 1.7.3.1) (Ghosh *et al.*, 2001).

A number of insulators have now been identified in other species (Schmidt *et al.*, 2000). A boundary element identified in yeast implicates the SMC (structural maintenance of chromosome) proteins in insulator function (Donze *et al.*, 1999). Vertebrate insulators all contain CCCTC sites and the protein that binds to these sites, CCCTC-binding factor (CTCF), is thought to be responsible for enhancer blocking activity (Bell *et al.*, 1999). Recently an insulator has been found to play a role of in the mechanism of imprinting at the *Igf2/H19* locus (reviewed in Bell *et al.*, 2001) (Figure 1.3b). The paternal allele is differentially methylated in a region between the two genes (the imprinted control region, ICR) and since the activity of this insulator depends on the presence of functional CTCF binding sites, there is no insulation on the methylated paternal allele, therefore no CTCF binding and the *Igf2* gene is expressed.

1.3.7.2 LCRs

The β -globin locus control region (LCR) is the founding member of a novel class of *cis*-acting regulatory elements that confer high level, tissue specific, site of integration-independent, copy number-dependent expression on linked transgenes located in ectopic chromatin sites (Grosveld *et al.*, 1987, reviewed in Li *et al.*, 1999; Kioussis and Festenstein, 1997). Copy number-dependent expression is widely considered to be indicative of open chromatin structure i.e. DNA that is accessible to transcription factors, and it is thought that LCRs initiate a pathway of decondensation, possibly mediated by histone acetylation and recruitment of chromatin remodelling complexes (Madisen *et al.*, 1998). LCRs control the level of transcription by ensuring that the locus is active and also protect against position effects and negative chromatin modifications and higher-order structures.

LCRs appear to be comprised of multiple elements, including enhancers, insulators, chromatin domain-opening elements and tissue-specificity elements; these can be clustered together or scattered throughout a locus. Individually these elements differentially affect gene expression; but it is their collective interaction that functionally defines an LCR. LCRs are frequently associated with DNaseI hypersensitive sites and are responsible for the temporal regulation of clustered tissue-specific genes by competition for transcription

factors. Evidence suggests that the LCR interacts with only one β -globin gene promoter at a time, and that it 'flip-flops' between two or more promoters, depending on the stage of development (Li *et al.*, 1999). These interactions might also influence the formation of the higher-order fibre, especially if the LCRs are looping to interact with other enhancer sequences (Figure 1.4).

1.3.8 A role for unconventional RNAs in chromatin structure

Unconventional RNAs are being encountered in novel epigenetic regulatory mechanisms e.g. the heterochromatinisation of Xi in X inactivation is initiated by the Xist RNA (Meller, 2000) (from which protein is not translated). Similarly, the product of the maternally expressed *H19* gene in the imprinted *Ins2/Igf2* gene cluster is a non-coding RNA (Tilghman, 1999). No function has been ascribed to the RNA, but it is conserved and abundantly expressed. The *Dlk/Gtl2* locus was also recently found to have a nearly identical arrangement of a paternally imprinted protein-coding gene adjacent to a maternally imprinted non-coding RNA, suggesting selection for these unusual RNAs (Schmidt *et al.*, 2000). In addition to genic transcription, the phenomenon of intergenic transcription is increasingly receiving attention, but the extent to which this occurs in the human genome unknown. The significance of intergenic transcription is not clear, but one might speculate a role in maintenance of local chromatin structure with a high density of transcription units in a region of chromatin inducing decondensation of chromatin.

1.4 HIGHER-ORDER CHROMATIN ARCHITECTURE

Data described above (described in sections 1.1 to 1.3) indicates that chromatin structure determined by molecular mechanisms plays an essential role in the control of gene expression. It is further becoming apparent that higher-order chromosomal architecture and three-dimensional genome organisation is implicated in the control of gene expression by facilitating action of these mechanisms. The nucleus is compartmentalised and designed for rapid and plastic response to signalling events in the cell.

Ideally, any model of large-scale chromatin folding would unify mitotic and interphase chromosome structure and predict the structural transitions accompanying cell cycle-driven

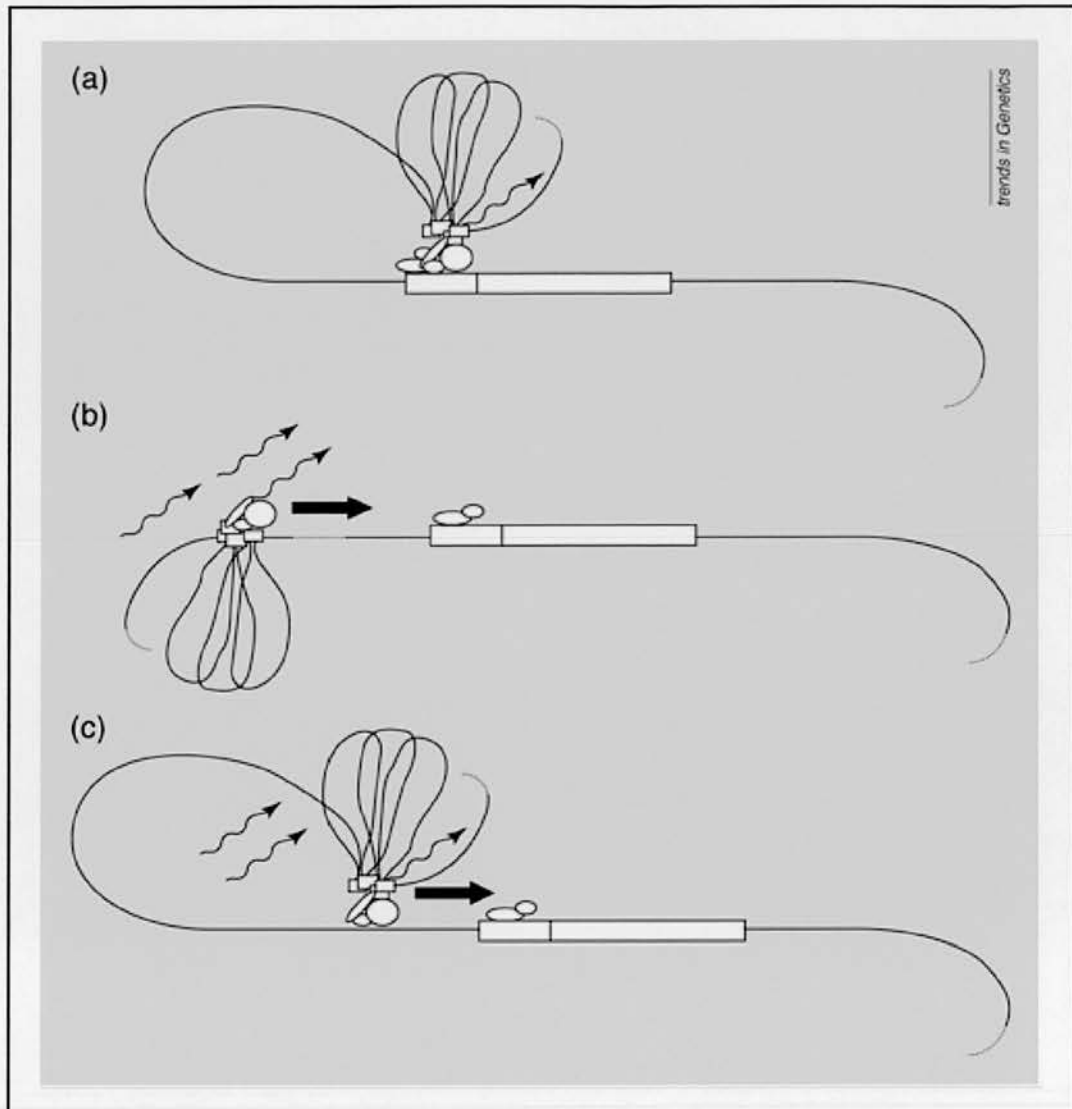


Figure 1.4 Models of LCR function

a) Looping model. Sequence-specific activation factors bound at the LCR, coactivators and a globin gene promoter, are brought together to form a transcription complex by looping of the DNA between the LCR and the distal gene.

b) Tracking model. Sequence specific activators bound at the LCR track down the helix until they encounter the appropriate stage-specific promoter bound by other activators, then transcription is initiated.

c) Facilitated tracking model. Sequence-specific activators bind the LCR and a loop is formed bringing the factors into contact with a DNA region distal to the LCR and proximal of the gene promoter. Tracking occurs until the correct promoter and associated transcription factors are encountered, then a transcription complex is formed and transcription begins. A globin gene is represented as a rectangular box, promoter sequences are to the left of the vertical line. The four erythroid HS sites are shown as small boxes, and the core flanking regions are displayed as looped lines that constrain formation of the LCR holocomplex. Transcription factors and co-activators are represented as different sized circles and ovals; wavy arrows indicate transcripts.

(Taken from Li *et. al.*, 1999).

chromosome condensation/decondensation. During interphase most chromatin is less condensed than at mitosis, and individual chromosomes escape simple light-microscopic detection. As eukaryotic DNA is expressed and replicated in the context of an interphase nucleus it seems biologically most meaningful to understand chromosome organisation in this state. Recent research has begun to define patterns of organisation within interphase chromosomes.

1.4.1 Territorial organisation of interphase chromosomes

At the turn of the century, a few cytologists proposed the idea of a territorial organisation of chromosomes in nuclei of animal and plant cells (Rabl, 1885; Boveri, 1909, for review see Cremer *et al.*, 1993; Chevret *et al.*, 2000). As recently as 25 years ago this concept was mainly ignored, and the interphase nucleus was viewed by many as a sack containing chromatin fibres from the different chromosomes intermingling in the nucleoplasm (Comings, 1968; Vogel and Schroeder, 1974). Only very general features of euchromatic and heterochromatic DNA could be distinguished in interphase nuclei and these patterns bore no obvious relation to individual mitotic chromosomes, the recognisable genomic structures studied for many years.

In 1885, Rabl proposed his theory on the internal structure of the nucleus based on observations in *Salamandra maculata* (Rabl, 1885). Today, the 'Rabl configuration' refers to the idea that chromosomes maintain their anaphase-telophase orientation, with centromeres localised at one side of the nucleus and telomeres at the other, as well as their individuality. This organisation is apparent in *Drosophila* polytene chromosomes (Hochstrasser *et al.*, 1986) and early embryonic nuclei (Marshall *et al.*, 1996). In the G2 phase of the cell cycle, telomeres of *Drosophila* polytene chromosomes and of *S. pombe* chromosomes also show comparable localisation at the nuclear periphery (Hochstrasser *et al.*, 1986; Funabiki *et al.*, 1993).

The interphase chromosome territory hypothesis, in which each chromosome occupies a discrete and distinct region within an interphase nucleus, is now well established. Initial evidence came from UV-laser micro-irradiation experiments on Chinese hamster cells cultured *in vitro*. After pulse-labelling with [3H]thymidine, DNA synthesis could be detected in the micro-irradiated part of the nucleus. DNA damage was shown, by autoradiography and by antibodies specific for UV-irradiated DNA, to be restricted to a few chromosomes when

cells were followed to the subsequent mitosis (Cremer *et al.*, 1982; Rappold *et al.*, 1984, reviewed in Cremer *et al.*, 1993).

Development of the fluorescence *in situ* hybridisation (FISH) technique made it possible to visualise whole chromosomes directly in rodent-human hybrid cells using human genomic DNA as a probe (Manuelidis, 1985; Schardin *et al.*, 1985). Even in these hybrid cells, human chromosomes occupied defined domains. Later, FISH 'painting' using chromosome-specific DNA sequences as probes confirmed the existence of chromosome territories in both tumour and normal human cells (Pinkel *et al.*, 1986; Lichter *et al.*, 1988, reviewed in Manuelidis, 1990; Cremer *et al.*, 1993) and also in an increasing number of animal and plant species (Heslop-Harrison and Bennett, 1990). Further advances in fluorescence microscopy and generation of new DNA probes now enable us to visualise individual genes, selected chromosome segments as well as entire single chromosomes to study the three-dimensional organisation of chromosome territories in detail (Lichter *et al.*, 1991) (Figure 1.5).

1.4.2 Models of large-scale chromatin folding

Despite intensive efforts to decode the linear structure of complex genomes, still relatively little is known about three-dimensional genome organisation within nuclei of higher organisms. Various models have been suggested, based on a common theme: chromatin loops and hierarchical folding as a means of maintaining the integrity of the chromosome and facilitating dramatic changes in condensation during the cell cycle and during transcriptional activation.

1.4.2.1 The radial array model

The radial-loop, helical-coil model of mitotic chromosome structure (Paulson and Laemmli, 1977) has been extended to interphase chromosomes (Manuelidis and Chen, 1990). In this model, the 30nm fibre is arranged as pairs of 30 kb loops which extend outwards from a central 'matrix' to form an array. Ten such loops form a 'radial array' with a diameter of 240 nm, the observed measurement of an interphase chromatin fibre and these are coiled, more tightly in G-band regions than in R-bands. This model predicted that metaphase chromosome bands and sub-bands would be preserved structural units in the interphase chromosome. It is not clear, however, whether the same SAR/MAR sequences are attached to an underlying scaffold in both mitotic and interphase chromosomes.

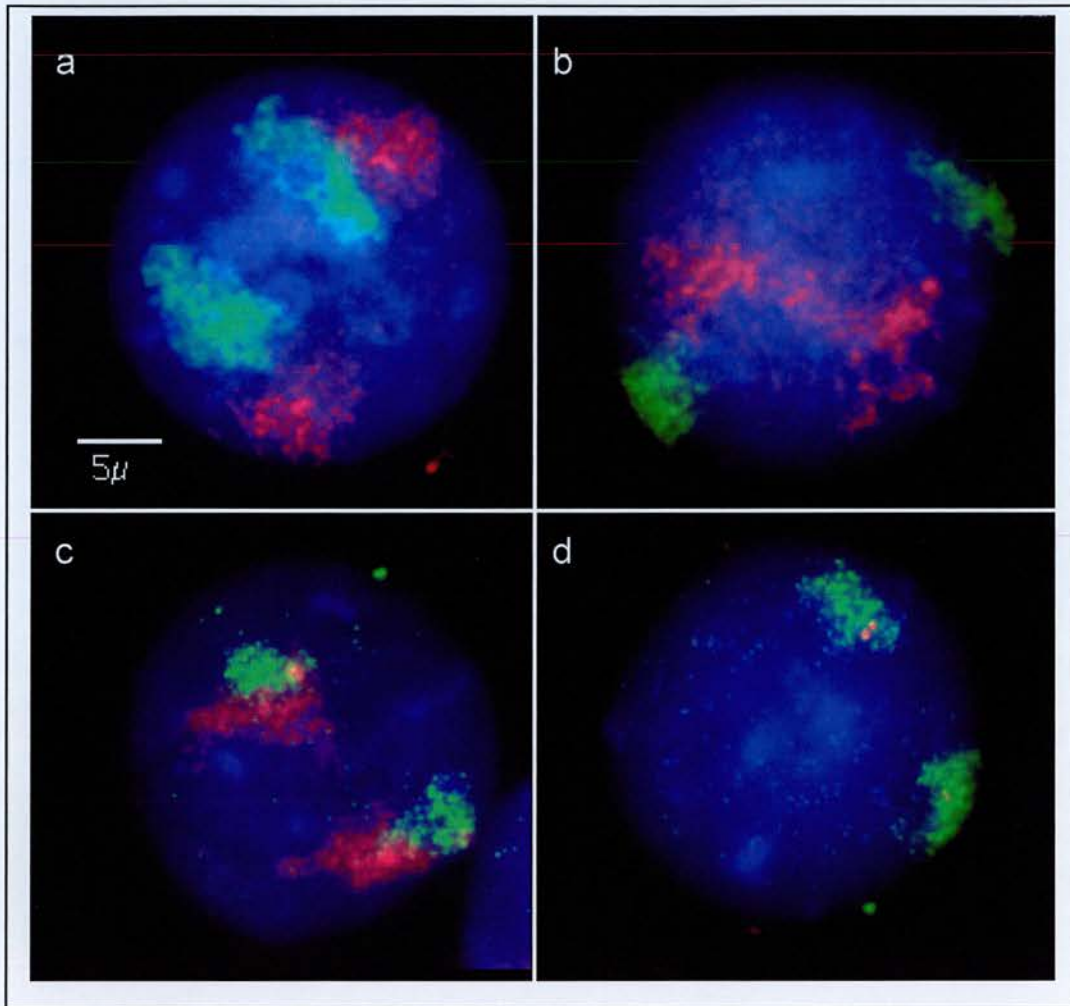


Figure 1.5 FISH using chromosome paints, chromosome arm paints and probes to individual genes.

FISH on MAA-fixed lymphoblastoid nuclei counterstained with DAPI, using:
 a) whole chromosome paints specific to HSA6 (green) and HSA8 (red). Note the more internal position of HSA6 (S. Boyle).

b) whole chromosome paints specific to HSA18 (green) and HSA19 (red) (S. Boyle). Note the peripheral, position of HSA18 and the more internal location of HSA19. In addition, although the two chromosomes are of similar physical size, the area of the two territories is vastly different. Gene-rich HSA19 occupies a much larger area of the nucleus than the gene-poor HSA18 territory.

c) HSA11p paint (green) and HSA11q paint (red). Note that separate domains are formed by the two chromosome arms, with little intermingling between p arm and q arm chromatin (Section 1.4.1).

d) HSA11p arm paint (green) and a cosmid probe specific to the *WT1* gene from 11p13 (red).

1.4.2.2 The 'folded chromonema' model

An early alternative model proposed a successive, helical coiling of 10 nm chromatin fibres into 30-50 nm tubes, and of these into 200 nm diameter tubes, which coil into ~600 nm metaphase chromatids (Sedat and Manuelidis, 1978). This model has recently been modified to the 'folded chromonema' model, based on *in vivo* light microscopy combined with transmission electron microscopy (TEM) ultrastructural analysis of folding intermediates during the transition into and out of mitosis (Belmont and Bruce, 1994; Li *et al.*, 1998a). In this model, 10 and 30 nm chromatin fibres fold to form a ~100 nm diameter chromonema fibre, which then folds into a 200-300 nm diameter prophase chromatid, which itself coils to form the metaphase chromosome. This model involves no anchoring of DNA loops to a chromosome axis, but chromatin is postulated to be organised by nuclear scaffolding proteins, or by non-specific inter-fibre interactions, possibly modulated by histone modification (section 1.2).

1.4.2.3 The random walk/ giant loop model

An important feature of G1 phase chromosomes is that they behave approximately as ideal Gaussian chains, which obey random walk statistics (van den *et al.*, 1992; Sachs *et al.*, 1995; Yokota *et al.*, 1995). This model is based on statistical analysis of the mean separation between two chromosome sites as a function of genomic distance. This relationship is biphasic with respect to increasing genomic distance between sites, which was taken to be indicative of two organisational levels of chromatin packaging (Yokota *et al.*, 1995). In the first level, flexible chromatin loops of several Mb are attached to a supple backbone and at the second level, both the backbone and loops are folded according to a random-walk path with loop sizes adapted to fit measured interphase distances. Again, differences in compaction of G- and R-band chromatin was observed and it was suggested that this arrangement would result in a degree of overlap between chromosomal domains with no maintenance of interchromosomal channels (Yokota *et al.*, 1997).

1.4.2.4 The 'multi-loop subcompartment' model

This model is based on computer simulations of chromatin folding under different conditions (Munkel *et al.*, 1999). Interphase chromosomes were simulated as a flexible fibre folded into 120 kb loops and subsequent arrangement of ten of these loops, held in place by 'stiff spring'

anchoring proteins, into rosette-like subcompartments. These subcompartments are connected by chromatin segments of approximately 120 kb. This model assumes a random folding of the chromatin chain, but does not exclude organisation of individual subcompartments related to the function of the chromatin. It is interesting to observe that the predicted subcompartment size is similar to observed subchromosomal foci corresponding to early- and late-replicating DNA (section 1.6.1).

1.5 THE ROLE OF LARGE-SCALE CHROMATIN STRUCTURE IN TRANSCRIPTION REGULATION

1.5.1 Lessons from model, non-mitotic chromosomes

As discussed (sections 1.1 and 1.4.1), spatially distinct 30 nm chromatin fibres cannot be visualised within interphase nuclei by TEM (Woodcock and Horowitz, 1995). Instead, chromatin is packaged into higher-order chromatin domains within which the linear fibre organisation is not evident. The exceptions to this technical limitation are special interphase chromosome types; these include lampbrush chromosomes in meiotic nuclei, and polytene chromosomes in many Dipterans. The distinct banding pattern of polytene chromosomes, combined with the transient appearance of decondensed "puffs" associated with high transcriptional activity, immediately suggest the existence of defined genomic regions with different higher order chromatin compaction levels. EM ultrastructural analysis of a Balbiani Ring puff in *Chironimus* polytene chromosomes has reinforced these earlier light microscopy observations (Ericsson *et al.*, 1989). Abrupt transitions in banding patterns and chromatin folding are seen within several kilobases of the 5' and 3' ends of the transcription units.

Inferences based on studies of polytene chromosomes have been supported by the visualisation of ribonucleoprotein- (RNP-) covered loops of highly transcriptionally active loci in lampbrush chromosomes (Callan, 1982). Again, sharp demarcations between the decondensed, transcriptionally active loops and the highly condensed chromatin along the chromosome axis are seen. These observations indicate that very distinct and sudden transitions in higher order chromatin structure can occur in a sequence-dependent manner. The appearance of lampbrush chromosome loops has reinforced models of decondensed

chromatin with looped domains associated with gene activity. Polymerase densities on these lampbrush loops and polytene puffs are often so high that nucleosomes may be substantially lost from the transcribed regions. In contrast, the majority of human RNA Pol II-transcribed genes in somatic cells are expressed with less than one polymerase per transcription unit (Jackson *et al.*, 1998). The organisation of these transcribed regions in the context of chromatin is still unknown.

1.5.2 Chromatin compaction in X-inactivation

The most striking example of long-range transcriptional repression coincident with changes in large-scale chromatin structure is the formation of the Barr body during mammalian X chromosome inactivation. Great advances have been made in deciphering the mechanisms that initiate the events of X-inactivation involving transcription of the Xist RNA (Section 1.3.8) and subsequent changes in DNA methylation and histone acetylation (Sections 1.2 and 1.3). However, we still do not understand the role, if any, that large-scale chromatin structure plays in achieving repression of transcription from the inactive X (Xi). Early cytological studies indicated that facultative and constitutive heterochromatin is more compact than euchromatin in interphase nuclei. The inactive X chromosome (Xi) appears as a dense mass by DNA staining methods both in light microscopy and by examination in the TEM. Using FISH, Xa and Xi interphase chromosomes differ in their shape and surface structure but still have similar volumes (Eils *et al.*, 1996) suggesting that silencing, in this case, is not simply a matter of chromosome compaction. Histone H2A variants MacroH2A1 and MacroH2A2 have been identified which are enriched in Xi chromatin (Costanzi and Pehrson, 1998; Chadwick and Willard, 2001). A specific role for these H2A variants is unclear, however the timing of MacroH2A1.2 accumulation on the Xi suggests it is not necessary for the initiation or propagation of random X-inactivation (Mermoud *et al.*, 1999).

Distinct dosage compensation mechanisms have evolved independently in other species but have a common mechanistic basis: specific molecules decorate the X chromosome and play a role in remodelling the chromatin structure of the dosage-compensated chromosome in one sex, but not in the other. In *Drosophila*, the male X chromosome is hyper-activated by an RNA-protein complex which effects site-specific acetylation of histone H4 and results in the up-regulation of gene transcription 2-fold relative to each X-chromosome in female cells (Willard and Salz, 1997). A global change in the chromosome structure of the hyper-

activated X is suggested by its more diffuse cytological appearance (Gorman *et al.*, 1993; Kelley *et al.*, 1995).

Conversely, in nematodes, a distinct protein complex is recruited to both X chromosomes in the interphase nuclei of hermaphrodite worms, which reduces gene transcription levels to those of the single X chromosome in males. This protein complex is related to the 13S condensin complex; key components are DPY-27 (dumpy protein-27), MIX-1 (mitosis and X-associated protein-1). Both of these are SMC (structural maintenance of chromosomes) proteins (Lieb *et al.*, 1996; Chuang *et al.*, 1996), putative ATPases that resemble molecular motors and are implicated in a wide variety of higher-order chromosome dynamics, including chromosome condensation (Hirano, 1998) (Figure 1.6). Although the role of DPY-27 is specific to dosage compensation, MIX-1 has an essential role in chromosome segregation and localises to all chromosomes at mitosis. Other components of the worm dosage compensation complex (DPY-26 and -28) also have some homology to non-SMC subunits of the 13S condensin complex. Thus, dosage compensation in nematodes appears to provide a direct link between chromatin compaction and regulation of transcription (Figure 1.6). Normal mitotic chromosome condensation is accompanied by a cessation of transcription, though it is unclear what role compaction of chromatin *per se* has in this. A mitotic-like condensation of interphase X chromosomes may directly decrease the transcription of X-linked genes in hermaphrodite worms, or might act as a trigger to modify the X-linked chromatin in other ways.

It will be interesting to see whether gene repression through modulation of higher-order chromatin structure is found at other loci and in other species. For example, there is genetic evidence that SMC proteins play a role in maintaining boundary-insulator element function (Section 1.3.7) at the silent mating type loci in yeast (Donze *et al.*, 1999).

1.5.3 Visualising the effects of transcriptional activation on chromatin in live cells

Recently a system has been developed that enables visualisation of amplified chromosome regions or even entire chromosome arms in living cells. By inserting lac operator repeats adjacent to a selectable marker and combining this with gene amplification, GFP-lac repressor staining representing distinct large-scale chromatin fibres can be selectively seen extending for up to 5 μm in living cells (Robinett *et al.*, 1996). Direct analysis of a late

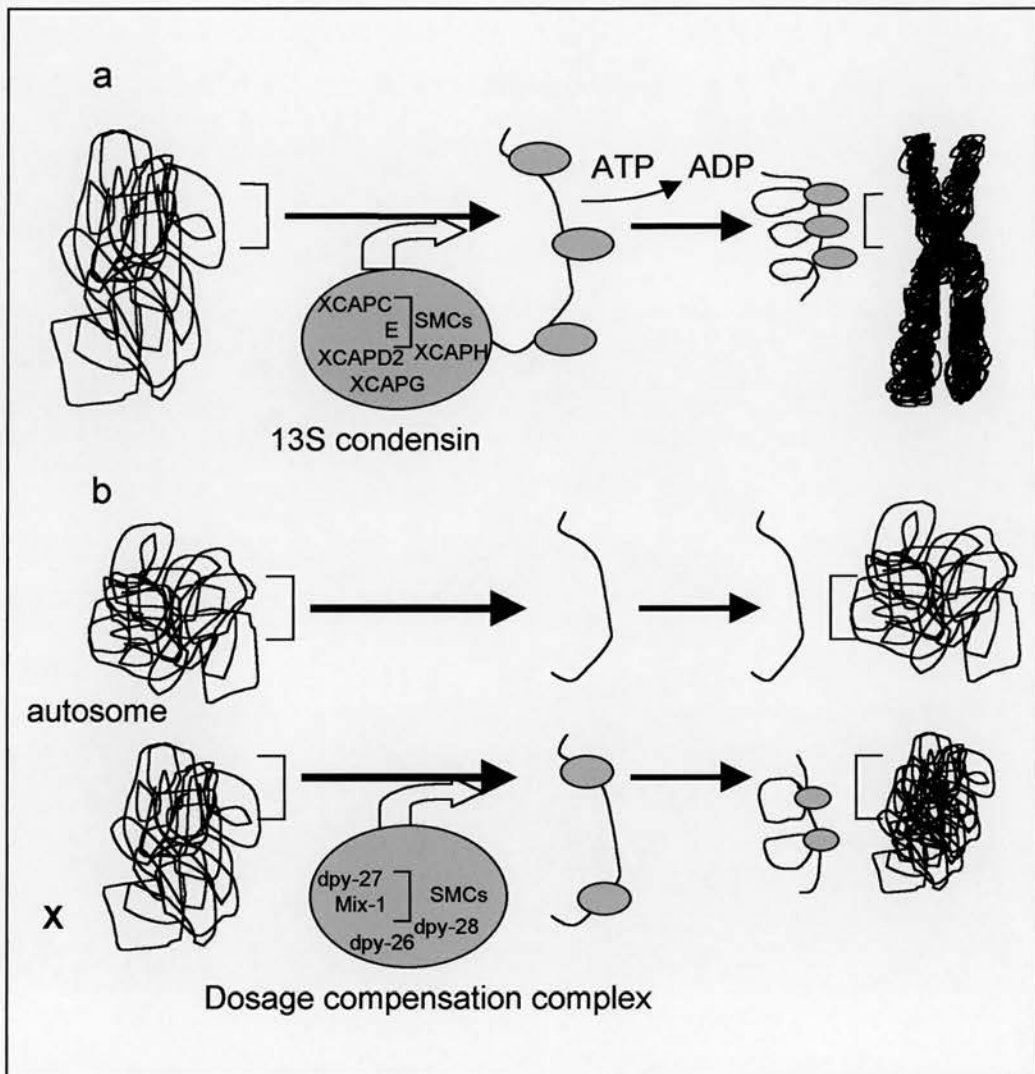


Figure 1.6 Dosage compensation in *C. elegans*.

a) The 13S condensin complex is required for mitotic chromosome condensation. Key components of this complex are two SMC proteins (called XCAPC and E in *Xenopus laevis*) that may use the energy generated by ATP hydrolysis to drive coiling of the interphase chromatid fibre, and so condense it for mitosis.

b) In dosage compensation in the nematode worm *C. elegans*, a dosage complex is recruited to the X chromosomes, but not to the autosomes, of hermaphrodite animals during interphase. Two components of this complex (DPY-27 and Mix-1) are SMC proteins. In the model, the SMC proteins in this complex bring about a partial condensation of the X chromosomes that either directly or indirectly dampens the transcription of X-linked genes.

replicating, heterochromatic amplified chromosome arm has revealed compaction ratios only several fold lower than metaphase values during late stages of DNA replication (Li *et al.*, 1998). After targeting a strong transcriptional activator to this region, large-scale chromatin fibres with compaction ratios near 1000:1, which show extremely high levels of transcriptional activity, can be seen within interphase nuclei (Tumbar *et al.*, 1999). What impact this level of organisation has on transcriptional activity remains unclear, but large-scale decondensation appears to be a direct cause or consequence of transcription. Similarly, combining this system with two colour variants of GFP enabled examination of the spatial organisation and timing of gene expression (Tsukamoto *et al.*, 2000). On gene activation, dynamic morphological changes, from a condensed to an open chromatin structure, were observed.

1.6 TRANSCRIPTION AND THE CHROMOSOME TERRITORY

1.6.1 Interphase chromosome subdomains

Probes for individual genes, selected chromosome segments and entire single chromosomes are now being used to study the three-dimensional organisation of interphase chromosomes in detail (Cremer *et al.*, 1993; Bridger and Bickmore, 1998) (Figure 1.7) (Section 1.4.1). These studies have revealed different levels of organisation within chromosome territories: separate domains are formed by individual chromosome arms (Dietzel *et al.*, 1998; Scherthan *et al.*, 1998; Visser and Aten, 1999). A recent technique involving microinjection of fluorescent nucleotides into cells, followed by chases over several cell generations has enabled visualisation of spatially distinct, labelled, and compact chromosome territories in living cells (Zink *et al.*, 1998).

Mammalian chromosomal domains replicate at defined, developmentally regulated times during S-phase. It has been suggested that mammalian genomes are compartmentalised according to the programme of DNA replication timing (Ferreira *et al.*, 1997; Berezney and Wei, 1998) and that these compartments are established at late telophase/early G1 (Dimitrova and Gilbert, 1999). Defined domains are formed within interphase chromosomes by early and late-replicating chromatin (Visser *et al.*, 1998) and the G- and R-bands of mitotic chromosomes persist in interphase as distinct domains termed subchromosomal foci

(Zink *et al.*, 1999). Replication foci have been proposed to persist throughout the cell cycle (Sparvoli *et al.*, 1994; Jackson and Pombo, 1998). Together these suggest that structural, transcriptional and replicational domains share topographical boundaries as basic units of chromosome organisation.

1.6.2 The inter-chromosomal domain model (ICD)

Specific species of RNA have been found in accumulations, either spherical or track-like in morphology (Lawrence *et al.*, 1989; Lawrence *et al.*, 1993). The transcripts contained within some of these RNA structures appear to have been released from a discrete genomic site and their morphology suggests that they are channelled within the nucleoplasm (Xing *et al.*, 1995; Lampel *et al.*, 1997). It has also been shown that single species of RNA are located immediately outside of the chromosome from which the RNA originated and concentrations of splicing components visualised by Sm staining were also found to be associated with the periphery of chromosome territories and were excluded from their interior (Zirbel *et al.*, 1993). These observations, in combination with the fact that chromosomes have a distinct territorial organisation, gave rise to the interchromosome domain (ICD) model (Cremer *et al.*, 1993; Zirbel *et al.*, 1993) (Figure 1.7). This model describes chromosome territories as having strict boundaries, around which a series of channels connect the chromosome with the remainder of the nucleus, and also to the cytoplasm via the nuclear pores. It was hypothesised that the macromolecular machinery required for transcription, RNA splicing, and transport is contained in this compartment and that packaging of chromatin into the interior of chromosome territories away from the transcriptional machinery provides a mechanism for gene repression (Cremer *et al.*, 1993).

Channels have been visualised indirectly using human cells stably transfected with *Xenopus* vimentin engineered to contain a nuclear localisation signal (Bridger *et al.*, 1998; Reichenzeller *et al.*, 2000). NLS-vimentin was seen to progressively extend, with time, out from speckles into strictly orientated intranuclear, filamentous arrays, localised almost exclusively outside of chromosome territories with specific nuclear RNAs, coiled bodies and PML bodies. An interchromosomal space was also suggested from electron microscopy examination of chromosomes during the cell cycle (Chai and Sandberg, 1988).

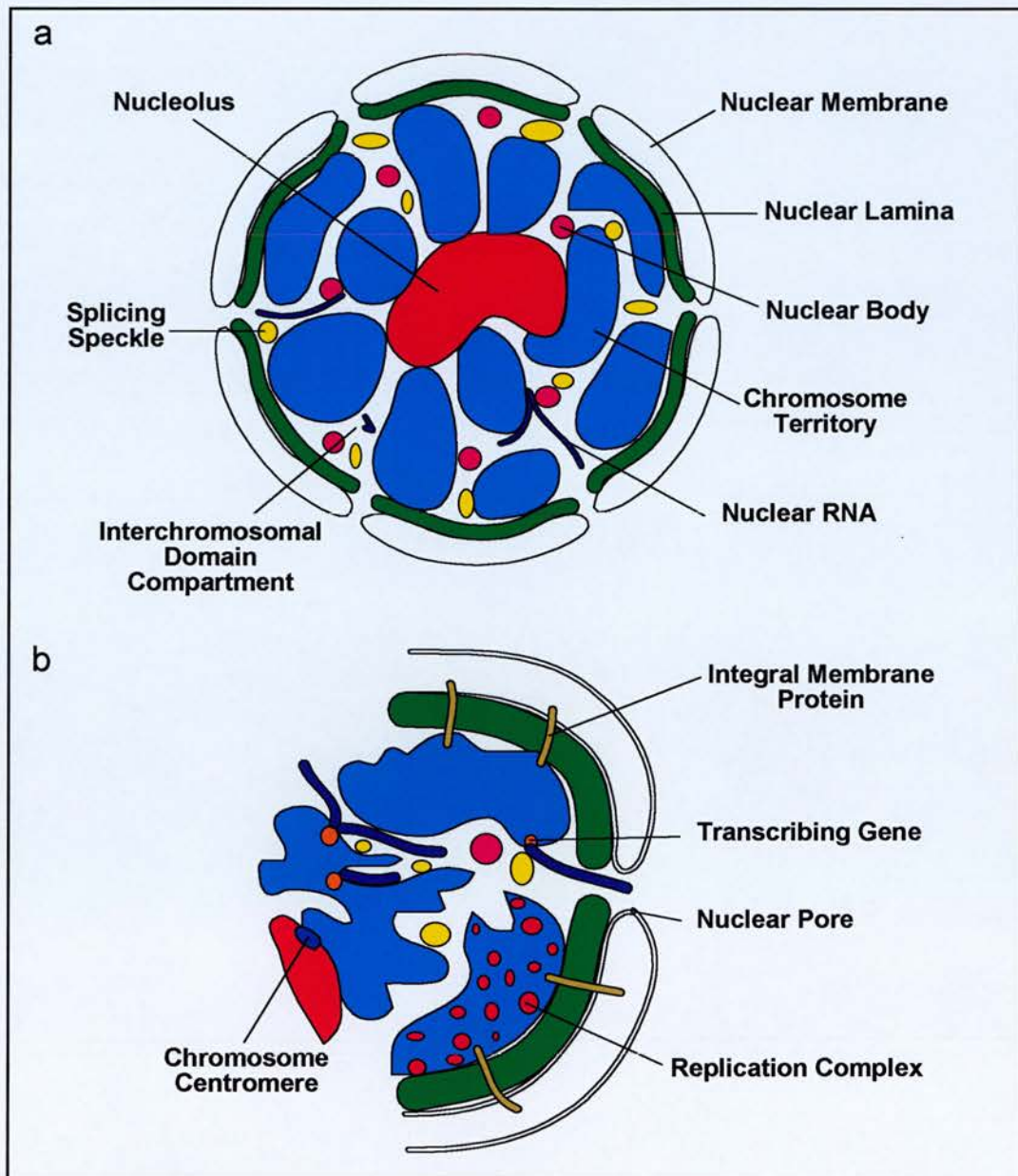


Figure 1.7 The Inter-chromosome domain model

a) The nucleus, bound by the nuclear membrane and the inner lamina, is a compartmentalised structure. Chromosomes occupy discrete domains known as chromosome territories (blue), and the nucleolus (red) is formed by the specific interaction of rDNA on the short arms of acrocentric chromosomes. The space between chromosomes has been termed the interchromosome domain (ICD) (Section 1.6.2) and contains concentrations of components involved in metabolic processes e.g. splicing speckles, transcription factories etc., and also nuclear bodies such as coiled bodies and PML bodies.

b) The nuclear membrane contains integral membrane proteins such as LBR and LAP2 that might have a role in positioning chromosomes within the nucleus (Section 1.7). Transcription occurs at the surface of chromosome territories in the ICD, and specific nascent RNAs are extruded from the surface of chromosome territories.

(Taken from Bridger and Bickmore, 1998).

1.6.3 Are active genes displayed on the surface of chromosome territories?

A prediction of the original ICD model was that genes should be displayed on the surface of interphase chromosome territories, so as to be accessible for interaction with the transcriptional machinery residing in the ICD. Several groups have tested the validity of this model by analysing the positions of a small number of individual expressed and non-expressed sequences within their respective chromosomes.

In one such study, three human genes (*DMD*, *MYH7* and *HBB*) were all found to be located at the surface of their respective chromosome territories and, in contrast, a non-coding DNA sequence from HSA11 was found more internally positioned (Kurz *et al.*, 1996). A prediction of the ICD hypothesis was that packaging of chromatin into the interior of chromosome territories away from the transcriptional machinery might provide a mechanism for gene repression. However, genes were found at the surface of chromosome territories irrespective of their transcription status and a second non-coding sequence from HSA1 was also found positioned at the periphery. Kurz *et al.*, interpreted the positioning of a non-coding sequence at the territory surface to be due to the surrounding DNA being a gene-rich R-band, suggesting that the relationship between gene positioning and the ICD might not be clear cut.

In a second study, correlation between the positions of two genes and their transcription activity was found (Dietzel *et al.*, 1999). The two *SLC25A6* (previously *ANT3*) homologues, which are not subject to inactivation, and the active *SLC25A5* (previously *ANT2*) homologue were close to the surfaces of painted Xa and Xi territories. In contrast, the methylated and inactive *SLC25A5* gene was found in a more interior location within the Xi territory. These data suggests a role for transcription status in gene positioning, however, it is possible that the Xi territory and its methylated sequences might be subject to alternative organisational constraints, and therefore may not be representative of patterns of organisation found in other chromosome territories.

A recent study has investigated the organisation of large genomic regions from human chromosome 6, including the gene-rich Major Histocompatibility Complex (MHC) in 6p21, and also gene-poor 6p24. (Volpi *et al.*, 2000). All regions analysed were found predominantly at the periphery of HSA6, irrespective of their expression status. In addition,

several Mb of gene-rich chromatin were frequently observed looping out from the surface of the chromosome. This configuration was found to be cell type dependent, and related to the number of genes in that region. This is similar to the previous observation of thin chromatin protrusions carrying amplified ERBB-2 sequences extending from the surface of HSA17 (Park and De Boni, 1998). Thus, studying individual genes is insufficient to discern patterns of organisation, and sequence context is very important.

Chromosome territories or chromosomal subdomains appeared to be well defined at the resolution of chromosome painting and light microscopy. However, looping of chromatin, as observed from the surface of both HSA6 and HSA17, is not compatible with the 'solid' chromosome territory structure inferred by the ICD model (Cremer *et al.*, 1993). These loops of chromatin will probably project into other chromosome domains, although further research is required to determine whether there are specific sites of intermingling between chromosomes. Chromosome territories exhibit various shapes with occasional finger-like protrusions which have been seen embedded in other territories at a small number of sites (Visser and Aten, 1999). However, intertwining of chromatin from two different chromosomes in the context of the interphase nucleus is thought to be unlikely (Cremer and Cremer, 2001). These ideas of territory structure are closer to the picture generated using higher resolution methods such as TEM, in which nuclei contain highly decondensed chromatin masses, (Belmont *et al.*, 1999).

1.6.4 The 'modified' ICD hypothesis

With the exception of Xi, both early- (gene-rich) and late-replicating (gene-poor) DNA are distributed throughout their chromosome territories, implying that gene-rich DNA is also located within the chromosome interior (Visser *et al.*, 1998). Similarly, GC-rich (gene-rich) DNA sequences show no preferential localisation within chromosome territories although AT-rich (gene-poor) sequences were positioned more internally within chromosome territories (Tajbakhsh *et al.*, 2000).

Sites of transcription appear to be present throughout chromosome territories, with the obvious exception of the human Xi chromosome (Abranches *et al.*, 1998; Verschure *et al.*, 1999). Furthermore, TEM localisation of uridine or BrUTP incorporation has shown heavy labelling at the edge of condensed, large-scale chromatin domains rather than at the surface of chromosomes *per se* (Fakan and Nobis, 1978; Wansink *et al.*, 1996).

Together, these data seem to suggest little requirement for genes to locate to the surface of a chromosome to facilitate transcription. Rather, locally compacted and unfolded regions of an interphase chromosome form distinct subdomains, and chromatin folding is organised in such a way that transcriptionally active DNA is at the surface of large-scale chromatin fibres throughout chromosome territories. In addition, decondensed chromatin fibres can extend to a significant distance from the surface of a chromosome domain. Experimental evaluation of the ICD model has led to it being reformulated as the inter-chromatin domain model (Cremer *et al.*, 2000) (Figure 1.8).

1.7 THE SPATIAL DISTRIBUTION OF CHROMATIN IN THE NUCLEUS AND IMPLICATIONS FOR REGULATION OF TRANSCRIPTION

Gene activity can be directly influenced by the proteins that package DNA into chromatin and by enzymes that modify both those proteins and the DNA itself. The positioning of genes within chromosome territories has also been implicated in transcriptional control and evidence is growing that the organisation of chromatin within the nucleus in relation to other nuclear compartments can impinge on transcription.

A territorial organisation of interphase chromosomes raises the question of whether there are distinct spatial relationships between the territory of one chromosome and that of another, and also between chromosome territories and other nuclear bodies. The diffusional constraints on chromatin movement *in vivo* (Marshall *et al.*, 1997) mean that the physical proximity of different chromosomes in interphase, and their interactions with nuclear substructure, may be important in determining the likelihood of any two chromosomes meeting and interacting.

1.7.1 The role of the nuclear matrix

Just as the tethering of chromatin loops within chromosomes is suggested to maintain the structural integrity of chromosome territories, so the formation of a continuous matrix network could help stabilise 3D chromatin architecture in 3D cells. Major chromatin movements have been observed during differentiation of neuronal cells (De Boni, 1994;

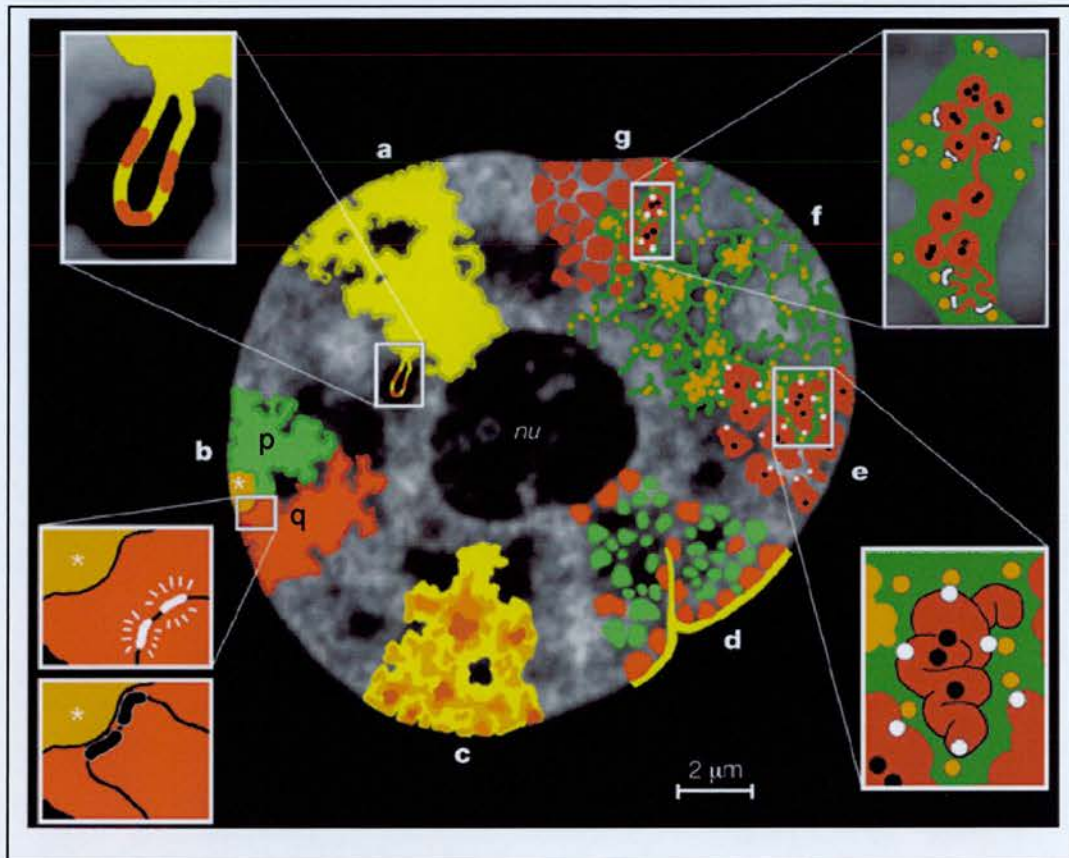


Figure 1.8 The Inter-chromatin domain model

The original inter-chromosome domain model has been modified to incorporate more recent experimental data that indicates significant sub-chromosomal organisation. Features of -chromosomal structure are indicated (a-g), on an optical section of a living HeLa cell. (Taken from Cremer and Cremer, 2001).

a) Chromosome territories have complex surfaces and giant gene-rich chromatin loops emanating from the territory surface have been observed (Section 1.6.3).
 b) Separate domains are formed chromosome arms (p and q) and by the centromere (asterisk). Association with centromeric heterochromatin has been implicated in gene silencing (inset) (Section 1.7.3.2).

c) Chromosome territories have variable chromatin intensity (brown high intensity, yellow low intensity).

d) Early- (green) and late- (red) replicating chromatin form separate domains within a chromosome territory, each comprising about 1 Mb (Section 1.6.1). Gene-poor chromatin (red) is often clustered at the nuclear lamina (yellow) and around the nucleolus (nu) (Section 1.7).

e) Higher-order chromatin structures (Section 1.4.2). Inset: a topological view of gene regulation indicates that active genes (white) are at the surface of chromatin fibres whereas silenced genes (black) may be located towards the chromatin interior (Section 1.6).

f) This model predicts that the ICD (green) contains protein complexes for transcription, splicing, DNA replication and DNA repair (orange).

g) Chromosome territory with 1 Mb domains (red) within the ICD network (green). Inset: a topological relationship between the ICD and active and inactive genes. Top: active genes (white) are located at the surface of these domains, whereas silenced genes (black) are located in the interior. Bottom: alternatively, closed 100 kb chromatin domains with silenced genes are transformed into an open configuration on transcriptional activation.

Manuelidis, 1990), so to allow for such movements, a nuclear matrix must have a dynamic structure. The form of the nuclear matrix is a controversial issue and its existence has even been questioned (van Driel *et al.*, 1995; Pederson, 2000).

It has been argued that the nuclear matrix is an insoluble network dispersed throughout the nucleus which is operationally defined as being resistant to high salt or detergents (Berezney *et al.*, 1995). It is believed that protein machinery for transcription, RNA splicing and DNA replication is associated with a structural matrix in the nucleus and that this plays a fundamental role in the organisation of these processes (Jackson and Cook, 1995; Berezney and Wei, 1998). Chromatin loops are suggested to have both stable and transient attachments to the nuclear matrix at 'matrix attachment regions (MARs) (Jackson *et al.*, 1992; Craig *et al.*, 1997), and the same or different MARs may have a role in organisation of interphase and mitotic chromosome architecture (section 1.4.2). Stable attachments are thought to provide structural integrity and shape and are associated with non-transcribed DNA, while the transient attachments have a role in juxtaposing DNA with macromolecular machinery for transcription, and perhaps, in turn to coupled nuclear translation (Iborra *et al.*, 2001).

Chromosome territories were observed to remain intact after an extraction procedure which removed most of the histone and other soluble nuclear proteins; however, disruption of the nuclear matrix with RNase I, followed by 2M NaCl, resulted in total disruption (Ma *et al.*, 1999). The authors argue that RNA and/or RNP interactions, together with a small subset of acidic nuclear proteins, called the 'chromosome territory anchor proteins' (CTAPs), which are released with the 2M NaCl treatment, are essential. Cremer co-workers argue that the topology and content of the ICD provide the *in vivo* equivalent of nuclear matrix preparations (Cremer *et al.*, 2000). They propose a second interpretation of the data from Ma and co-workers: that the ICD space in the living cell nucleus contains protein and RNP aggregates only in a discontinuous soluble form and that high salt treatment results in the formation of an insoluble *in situ* matrix. This dynamic view allows the possibility of establishing a continuous 3D network that may be dissolved at another time to allow territory movements or the reordering of early and late replicating chromatin domains, and then become re-established later, perhaps to preserve a distinct 3D architecture of a terminally differentiated cell type.

Of particular interest is the dramatically different association of chromosomes 18 and 19 with the nuclear matrix. Gene-rich HSA19 is more tightly associated with the nuclear matrix

than gene-poor HSA18 (Croft *et al.*, 1999), which implicates some component of the transcription machinery in the nuclear matrix architecture.

1.7.2 The relative positioning of chromosome territories

1.7.2.1 Homologous chromosome alignment in Dipterans

In *Drosophila* and other Dipterian insects, homologous chromosomes are usually paired at interphase (Metz, 1916) and have also been shown to interact genetically (Tartof and Henikoff, 1991; Gemkow *et al.*, 1998). Suppression or enhancement of a phenotype (trans-sensing) can be observed when this pairing is disrupted by chromosomal rearrangements (Henikoff, 1997; Wu and Morris, 1999). Homologous chromosome alignment is prominent in meiosis (Roeder, 1995) but for most organisms, it is not a normal characteristic of somatic nuclear organisation (Manuelidis and Borden, 1988; Leitch *et al.*, 1994). However, homologous loci of human and mouse genes subject to imprinting are transiently associated during late S-phase (LaSalle and Lalande, 1996). It has been suggested that these trans-interactions between oppositely imprinted chromosome regions are important for the maintenance of imprinting (Riesselmann and Haaf, 1999). Subtelomeric chromosome regions, which contain a high density of genes, also show an elevated incidence of somatic pairing in human nuclei (Stout *et al.*, 1999).

1.7.2.2 The nucleolus and acrocentric human chromosomes

The paradigm for a distinct spatial relationship of one chromosome to another occurs at the nucleolus, a specialised nuclear domain where ribosomal DNA (rDNA) is transcribed and pre-ribosomes are assembled. It is organised by “nucleolar organiser regions” (NORs) which, in humans, are localised at the short arms (“p” arms) of acrocentric chromosomes (HSA13, 14, 15, 21 and 22), which contain the ribosomal RNA genes (Henderson *et al.*, 1972). In the interphase nucleus rDNA-containing chromosomes are intrinsic to the nucleolus and also remain associated with one another through successive cell cycles (Bobrow and Heritage, 1980).

1.7.2.3 Non-random positioning of human chromosomes within the nucleus

Several lines of evidence suggested that the interphase organisation of chromosomes within the nucleus is non-random, and may be cell type- and cell cycle-specific. Barr bodies, directly distinguishable in cell nuclei by intense DAPI staining were noted to cluster near the nuclear periphery (Belmont *et al.*, 1986) and initial 3D studies on interphase nuclei suggested a non-random distribution (Manuelidis and Borden, 1988; Manuelidis, 1990). Support for this idea came from statistical analysis of metaphase spreads (Haaf and Schmid, 1991), based on the assumption that the chromosome distribution at metaphase reflects the topology of chromosomes in the preceding interphase. Invariant positions of acrocentric chromosomes on metaphase plates supports this hypothesis (Klein *et al.*, 1998). Measurements between homologous chromosomes in primary fibroblasts exhibited a broad distribution as a population, but were strongly correlated within each pair of cells derived from the same mother cell (daughters) (Sun and Yokota, 1999). This suggests that chromosomes are immobile at a global scale in the G1 interphase nuclei. It has also been suggested that the positioning of chromosomes within the interphase nucleus is size-dependent with larger chromosomes localising towards the nuclear periphery (Sun *et al.*, 2000).

The first specific 'nuclear addresses' assigned to chromosomes were to gene-rich human chromosome 19, which localised towards the interior of the nucleus, and HSA 18 which was positioned towards the periphery (Croft *et al.*, 1999). A subsequent study of the complete human karyotype found that this trend for gene-rich chromosomes to be positioned within the nuclear interior of differentiated lymphocyte nuclei applied to all chromosomes (Boyle *et al.*, 2001) (see section 1.7.4).

Support of the idea that mammalian chromosomes regularly adopt defined addresses within the nucleus and that chromosomes align to build up higher order compartments within the interphase nucleus is growing. However, the significance of such an organisation for gene and genome function is, so far, only implied. How chromosome territories are arranged in a highly defined nuclear architecture remain unclear (Spector, 1996). Transcription of rDNA does not appear to direct association of human acrocentric chromosomes to form the nucleolus, as human acrocentric chromosomes from which there is no transcription in a mouse background, are also targeted to the mouse nucleolus (Sullivan *et al.*, 2001). Rather, it appears that the nucleolus is formed by association of chromosomes containing NORs (Section 1.7.2.2). However, chromosome 19, also localises adjacent to the nucleolus,

(Bridger *et al.*, 2000), but is devoid of ribosomal genes, so must be positioned by an alternative mechanism. Components of the nuclear lamina have chromatin binding properties, which could serve to position certain chromosome domains to the nuclear periphery (section 1.7.3.1). However, the intranuclear organisation of chromosomes was not altered in cells lacking the integral nuclear membrane protein emerin, suggesting that emerin is not necessary for localising chromosome at the nuclear periphery (Boyle *et al.*, 2001), although other components of the lamin may be involved.

It is interesting to note that 15-20 Mb of DNA is sufficient to confer the characteristic localisation of human chromosomes 18 and 19 (Croft *et al.*, 1999). This suggests that the even a small region of chromatin from a particular chromosome has a 'nuclear address' sufficient to determine its location within the nucleus; there is not just one site on a chromosome which dictates the overall position of the whole chromosome.

1.7.3 Sites of transcriptional repression in the nucleus

1.7.3.1 The nuclear periphery

Rabl chromosome organisation in the nucleus, where centromeres are localised at one side of the nucleus and telomeres at the other (Section 1.4.1) is apparent in *Drosophila* polytene chromosomes (Hochstrasser *et al.*, 1986) and early embryonic nuclei (Marshall *et al.*, 1996). However, in larval imaginal disks, chromosomes maintain this Rabl orientation for less than 2 hours after mitosis and then reorient, but still maintain heterochromatic areas at the nuclear periphery (Csink and Henikoff, 1998). A peripheral localisation of telomeres has also been observed in *Trypanosoma* (Chung *et al.*, 1990), fission yeast (Funabiki *et al.*, 1993b), and some, but not all, plant species (Heslop-Harrison *et al.*, 1993; Dong and Jiang, 1998). The functional significance of this configuration is unclear, but the nuclear periphery has been associated with transcriptional repression in lower eukaryotes.

In *Saccharomyces cerevisiae*, heritable inactivation of genes occurs at the silent mating type loci and at telomeres, which cluster near the nuclear periphery. Telomeres are bound by at least seven proteins, including Rap1p, three Sir proteins and the yKu complex (Bourns *et al.*, 1998; Gravel *et al.*, 1998) and these proteins have been shown to interact with nuclear pore complexes (Galy *et al.*, 2000). Disruption of this complex leads to loss of both telomeric gene silencing and peripheral localisation (Gravel *et al.*, 1998; Maillet *et al.*, 1996). Sir

proteins which confer position-effect variegation (PEV) (section 1.3.4), are maintained at a high local concentration in this nuclear compartment (Gotta *et al.*, 1996; Cryderman *et al.*, 1999, reviewed Cockell and Gasser, 1999). In one study, tethering or targeting of a silencer-flanked reporter gene to the nuclear envelope facilitated its repression in the absence of an effective silencing protein complex (Andrulis *et al.*, 1998), suggesting a causal role for localisation at the periphery in gene silencing.

In mammalian somatic cells, centromere and telomere positioning is non-Rabl and is probably cell-type specific (Manuelidis and Borden, 1988; Billia and De Boni, 1991; He and Brinkley, 1996). Heterochromatin is often found juxtaposed to the nuclear envelope in higher eukaryotes, but it is not known whether this association actively facilitates repression of associated genes. Indirect evidence from *Drosophila* suggests that it may. Human HP1-type chromo domain proteins (section 1.3.4) make specific cell-cycle regulated interactions with the lamin β receptor (LBR) *in vitro*, providing a molecular link between the nuclear periphery and heterochromatin (Ye *et al.*, 1997). In addition, *Drosophila* Su(Hw) and mod(mdg4) proteins, factors that bind and modulate the activity of widely distributed insulator elements (section 1.3.7.2), are found in clusters near the nuclear lamina and can influence the nuclear position of genes not usually clustered together and this latered nuclear position is accompanied by global changes in transcription (Gerasimova and Corces, 1998; Gerasimova *et al.*, 2000). However, human telomeres have a dispersed nuclear distribution, with a lower density in the outer 25% of nuclear volume than expected for a random distribution (Luderus *et al.*, 1996).

1.7.3.2 Centromeric heterochromatin

In *Drosophila*, as in yeast, association of a variegated locus with a heterochromatic environment (Csink and Henikoff, 1996; Dernburg *et al.*, 1996) accompanies PEV. A mutant allele of the brown gene, bwD, which was generated by the insertion of ~1Mb of centromeric heterochromatic DNA (Platero *et al.*, 1998) causes trans-inactivation of the normal brown allele, with which it is paired during interphase (Talbert *et al.*, 1994). Silencing correlates with relocation of both copies to centromeric heterochromatin in the nucleus (Csink and Henikoff, 1996; Dernburg *et al.*, 1996). In the embryonic nuclei, bwD is spatially separated from other heterochromatin and association occurs during larval development. These data suggest that transcriptional repression can be accomplished by relocating genes to heterochromatic late-replicating compartments of the nucleus.

Studies in human and mouse cells have also shown that association with centromeric heterochromatin can accompany the silencing of certain genes (Figure 1.9). In mouse lymphoid cells, the protein Ikaros may mediate specific gene silencing. Ikaros localises to centromeric heterochromatin foci in the nucleus and is dynamically redistributed during the cell cycle. Genes differentially expressed during different stages of B cell maturation show a correlation of expression status with Ikaros association: Ikaros association is evident only in cells in which those genes are not transcribed (Brown *et al.*, 1997; Brown *et al.*, 1999). However, a recent study has shown that the Ikaros proteins initiate gene silencing through a direct effect on the promoter (Sabbattini *et al.*, 2001). Localisation to pericentromeric heterochromatin is likely to affect the action of Ikaros on regulatory sequences rather than cause the heterochromatinisation of the gene. In addition, at least part of the cellular pool of Ikaros is complexed to NURD, a chromatin remodelling complex, suggesting that Ikaros function may also involve the remodelling of chromatin (O'Neill *et al.*, 2000; Kim *et al.*, 1999).

Transcriptional enhancers may maintain gene expression by preventing (trans)genes from localising close to centromeric heterochromatin in the nucleus, and/or by recruiting (trans)genes to places in the nucleus that are permissive for stable, inherited transcription (Francastel *et al.*, 1999). Therefore, it is possible that genetic domains may traffic between repressive (heterochromatic) and transcriptionally liberal (euchromatic) compartments. Further chromatin studies show that it is not active transcription itself that correlates with the 'release' of the transgene from a heterochromatin environment, but the hyperacetylated state of its associated histones (Schubeler *et al.*, 2000). In addition, it has been suggested that juxtaposition to repetitive DNA is not incompatible with expression, but that a strong activator is required to overcome the repressed state (Lundgren *et al.*, 2000). These observations raise the question of which genes are sequestered in these transcriptionally repressed domains, as topographical constraints on chromatin structure would prevent all silenced genes from localising to pericentric heterochromatin.

The precise spatial positioning of genes may merely be the consequence of regulatory events that lead to the repressed state, and it remains to be proven that the subnuclear localisation of a gene contributes actively to the establishment and or inheritance of a particular chromatin

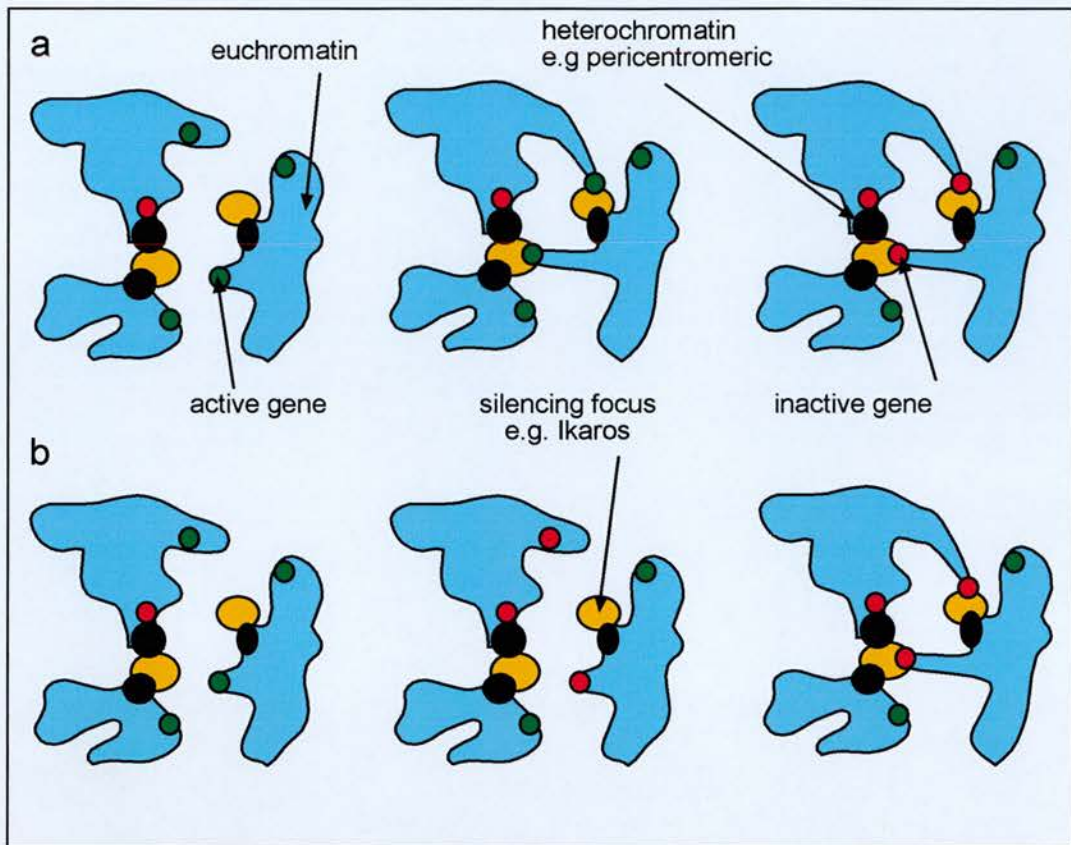


Figure 1.9 Localisation of silenced genes with heterochromatin: cause or consequence?

Active (green) and inactive (red) genes are distributed throughout euchromatic chromatin (blue). However, some genes change their nuclear position within the nucleus/chromosome environment dependent on their activity, and become localised with heterochromatin (black) and may subsequently become silenced. Pericentric heterochromatin is further associated with proteins such as Ikaros (yellow) which facilitate this gene silencing (Section 1.7.3). a) Recruitment of an active gene to a site of heterochromatin results in its silencing. Note however, that genes can still be transcribed when localised with centromeric heterochromatin (middle), as long as they are at the surface (Section 1.7.3.2). Inactivation may be a result of exposure to a concentrated pool of silencing factors, such as histone modification enzymes, or nucleosome remodelling complexes, or by removal from a pool of available transcription factors.

b) A transient gene inactivation allows a gene to be drawn into heterochromatin. Subsequently it is subject to heritable gene silencing.

Adapted from Bridger and Bickmore 1998 and Mahy *et. al.*, 2000.

state (Figure 1.9). It is also not clear what proportion or type of genes are amenable to regulation at this level.

1.7.4 The nuclear interior and transcription

A trend for gene-rich chromosomes to be positioned within the nuclear interior of differentiated nuclei, first noted for HSA19 (Croft *et al.*, 1999), was found to apply to all chromosomes (Boyle *et al.*, 2001). This preferential localisation of gene-rich chromatin was supported by the previous observation that the most highly acetylated isoforms of histone H4, generally associated with transcriptionally active chromatin (sections 1.2.2 and 1.3.2), were enriched in the nucleus interior (Sadoni *et al.*, 1999). In addition, DNA replicated during the first half of S-phase (representing generally gene-rich chromatin) is distributed throughout the nuclear interior, but is excluded from the perinuclear and perinucleolar regions (Ferreira *et al.*, 1997).

An artificial system utilising the GFP-based lac repressor-lac operator recognition system (section 1.5.3), enabled a heterochromatic chromosome arm to be engineered by gene amplification (Tumbar *et al.*, 1999). This chromosome domain was peripherally located through most of the cell cycle. However, targeting a transcription factor to this arm resulted in a shift to a more interior nuclear position and dramatic uncoiling. This data perhaps suggests that transcription factors target chromatin to the nucleus interior, which might be a preferential site for active transcription.

Arguing against preferential positioning of transcription in the nucleus interior, sites of transcription have been visualised throughout the entire nucleus (reviewed (Szentirmay and Sawadogo, 2000)). Such studies have used both antibodies against hyperphosphorylated (active) forms of RNA polymerase II (Bregman *et al.*, 1995; Zeng *et al.*, 1997) and detection of BrUTP incorporation (developed by Wansink *et al.*, 1993 and Jackson *et al.*, 1993). Transcription foci are dispersed, seemingly randomly throughout the nucleus with no preference in either the nuclear interior or the periphery. Similarly, in a further study, transcription factors were found distributed in punctate patterns throughout the nucleoplasm (Grande *et al.*, 1997). In wheat nuclei genes are concentrated towards the telomeres, which are positioned at one side of the nucleus due to the adoption of a Rab1 configuration by chromosomes within the interphase nucleus. However, localisation of sites of transcription were not correlated with the positioning of coding sequence and the distribution of foci of

BrUTP incorporation was found to be uniform across the nucleus (Abranches *et al.*, 1998). This suggests that there may be high levels of transcription occurring throughout the genome (perhaps intergenic transcription) and could be interpreted as providing indirect evidence that although transcription is generally non-uniform within the nucleus, gene-specific transcription may indeed be concentrated in certain gene-rich areas of the nucleus (the nuclear interior).

There is evidence that replication and transcription sites are clustered into separate zones in the cell nucleus (Wei *et al.*, 1998, reviewed Berezney and Wei, 1998). Since actively transcribed genes are preferentially replicated in early S-phase, these findings have implications for understanding the co-ordination of replication and transcription of genes. It was suggested that replication and transcription are arranged in successively higher levels of organisation in the cell nucleus: from individual sites that contain multiple genes to zones that contain numerous sites, to 3D networks that are the structural basis for the co-ordination of replication and transcription in the mammalian cell.

1.7.5 Positioning of genes relative to other nuclear bodies

Several genes appear to be specifically co-localised with different types of nuclear domains, including interchromatin granules, splicing speckles, coiled bodies, cleavage bodies and PML (promyelocytic leukaemia) nuclear bodies (reviewed in Lamond and Earnshaw, 1998; Dundr and Misteli, 2001). These observations obviously have implications for chromosome territory architecture, as associations would need to occur within the spatial constraints of the chromosomes.

The lac operator/repressor system has been used in combination with two colour variants of GFP, to visualise a gene and its protein product directly in living cells. Interestingly, the integrated gene locus was surrounded by a PML body and the association was transcription independent, but dependent upon the localisation of specific proteins to the locus (Tsukamoto *et al.*, 2000). The PML gene encodes a protein associated with a distinct subnuclear domain called PML bodies, which appear as dense fibrillar rings in normal cells, but are fragmented in promyelocytic leukaemia. Various functions have been attributed to the PML nuclear body and recent findings indicate that it can act as a transcription co-factor (Zhong *et al.*, 2000). In addition, it has recently been shown that gene-rich loops of

chromatin containing the MHC locus, extruded from the surface of interphase HSA6 (section 1.6.3), may be directed towards PML bodies (C. Shiels, personal communication).

Active genes, such as fibronectin, β -actin and collagen 1 α 1, may be associated preferentially with clusters of speckles which are enriched in factors required for the formation and maturation of mRNA (Xing *et al.*, 1993; Xing *et al.*, 1995; Huang and Spector, 1991). Many coiled bodies are also associated with specific gene loci, including the histone gene clusters, U1 and U2 small nuclear RNA genes, and the U3 small nucleolar RNA genes (Smith *et al.*, 1995; Frey *et al.*, 1999), but the functional significance of this association is unclear.

Cleavage bodies containing the RNA 3'-processing factors CstF and CPSF can be associated with a coiled body. This association changes during the cell cycle and can be correlated with both the spatial juxtaposition and transcription activity of the cell cycle-regulated histone gene cluster (Schul *et al.*, 1999). The PML nuclear body is also frequently closely associated with the coiled body/cleavage body doublet (Schul *et al.*, 1996) and with DNA that is replicated in a particular stage late in S-phase (Grande *et al.*, 1996).

1.8 DYNAMICS OF LARGE-SCALE CHROMATIN ORGANISATION

1.8.1 Chromatin motion in the nuclei of living cells

Four-dimensional analyses suggested that both Brownian-like movements as well as occasional directed movements of chromatin foci occur in living cell nuclei (Shelby *et al.*, 1996; Zink and Cremer, 1998; Bornfleth *et al.*, 1999). Use of a lacO lacR GFP system (section 1.5.3) has enabled the position of a marked locus to be studied in living cells and the distance a locus travels as a function of time provides a measure of the type of motion, whether it is active or passive, and its characteristics. In yeast, chromatin was observed to diffuse passively with a diffusion constant $D=5 \times 10^{-12}$ cm²/sec within a confined nuclear subvolume with a radius of 0.3 μ m and a similar value was found in *Drosophila* (Marshall *et al.*, 1997). This "constrained diffusional motion" of the chromatin fibre motion was independent of ATP hydrolysis.

Recent studies have used fluorescence recovery after photobleaching (FRAP) analysis where fluorescence is initially bleached in a spot using a focused laser beam, and the subsequent recovery of fluorescence is measured as a reporter of the diffusion of unbleached molecules into the bleached zone. The nuclear diffusion constants of GFP fused to the pre-mRNA splicing factor SF2/ASF, the chromatin protein HMG-17, and the nucleolar protein fibrillarin were measured in living cells using FRAP (Phair and Misteli, 2000). The diffusion constant for the chromatin-associated protein was significantly lower than for the other three proteins, which suggests that non-chromatin proteins move more rapidly and freely within the nucleus. Again no evidence was found for energy-dependent mobility. It should be borne in mind, however, that methods such as FRAP assay bulk populations of molecules and so might not identify small subpopulations of molecules with significantly different properties.

Of particular interest is identifying the molecular forces behind directed chromatin movement, which may involve the action of nuclear motor proteins on chromatin. Candidates for motor proteins inside the nucleus include actin and myosin (Milankov and De Boni, 1993), and also the SMC family of ATPases that have been proposed to act as motors in chromatin condensation (section 1.5.2) (Figure 1.6). The ability of RNA polymerases to act as motors to move DNA has also been demonstrated; active positioning of chromatin by locally applied forces may thus turn out to be a widespread phenomenon.

1.8.2 Cell cycle changes

Large-scale chromatin structure is subject to dynamic change during the cell cycle with compaction ratios falling from 20,000:1 in the metaphase chromosome to as low as 40:1 within individual 30 nm chromatin fibres. Changes of large-scale chromatin structure during the cell cycle have been observed by light microscopy for more than a century, where a progressive decondensation, accompanied by increases in nuclear volume, through G1 into early S phase is followed by a progressive condensation during late S and G2 (Kendall *et al.*, 1977). A similar trend is observed by TEM in mouse cells (Leblond and El Alfy, 1998) and a more detailed TEM study of chromosome decondensation reveals the uncoiling and straightening of a 100-130 nm chromonema fibre during early G1, followed by a transition to a 60-80nm fibre in late G1 early S phase (Belmont and Bruce, 1994).

The chromatin structure at a specific locus may follow a precise choreography of cell cycle associated changes, suggesting an ordered and reproducible organisation of DNA within

large-scale chromatin domains. The use of the GFP-based lacR-lacO recognition system has identified cell cycle-specific conformations of a large heterochromatic HSR (homogeneously staining region) (Li *et al.*, 1998); an abrupt decondensation in mid to late S phase appears to slightly precede the onset of DNA replication of this amplified chromosome arm. Interestingly, this decondensation was accompanied by a repositioning from the nuclear periphery to the nuclear interior. However, this could be an artefact of such a large amplified array, and may not be physiologically representative.

Changes in intranuclear localisation of native chromosome regions also occur at distinct stages of the cell cycle. Early replicating, gene-rich R bands localise preferentially in the nuclear interior while late replicating, gene-poor G/Q bands are found preferentially at the nuclear periphery (Croft *et al.*, 1999; Sadoni *et al.*, 1999) (section 1.7.4). This pattern is established during the first few hours of G1, which seems to be a particularly dynamic time (Bridger *et al.*, 2000; Dimitrova and Gilbert, 1999) (section 4.2). The intriguing observation that targeting a transcriptional activator to a specific chromosome site changes its normal cell-cycle program of intranuclear positioning suggests a possible link among nuclear positioning of large-scale chromatin domains, transcriptional activation, and the cell cycle (Tumbar and Belmont, 2001).

1.8.3 Quiescence and senescence

While most studies of large-scale chromatin and nuclear architecture have focused on proliferating, transformed cells in culture, it is likely that the nuclear organisation in these cells differs significantly from terminally differentiated, quiescent cells which comprise the majority of tissue mass in multicellular organisms. Quiescence, also known as G0, is a non-dividing growth state that occurs when cells cease DNA synthesis and withdraw from the cell cycle either temporarily or permanently in terminal differentiation. Major changes in chromatin organisation occur on transition to G0.

In quiescent lymphocytes, a sheet of compact chromatin covers the nuclear envelope, and large masses of chromatin associate with the nucleolus (Lopez-Velazquez *et al.*, 1996). Human chromosome 18, relatively gene-poor and located at the nuclear periphery in proliferating diploid fibroblasts is positioned in the nuclear interior during G0. Re-entry of cells into the cell cycle is also associated with changes in large-scale chromatin structure and changes in specific chromosome positioning. Interestingly, a return to a peripheral

localisation for HSA18 does not occur during the G0-G1 transition, but instead occurs in early G1 of the next cell division (Bridger *et al.*, 2000). A change in location relative to the nuclear periphery has also been seen for the centromere of human chromosome 11 during the G0 to G1 transition in peripheral lymphocytes (Tagawa *et al.*, 1997). Large-scale changes in positioning of individual genes have also been reported. Several inactive genes, which are found away from centromeres in nuclei of quiescent lymphocytes, are repositioned to heterochromatic centromeres on transition from G0 to G1 (Brown *et al.*, 1999) (see section 4.6). The mechanism behind these movements is not yet known.

Changes in nuclear organisation have also been reported in senescent cells. An age-related reactivation of genes from the inactive X-chromosome is observed in mice (Wareham *et al.*, 1987), suggesting a general decondensation of chromatin in senescence. A connection between metabolism, loss of gene silencing and increased rates of ageing has also been implied for yeast and perhaps higher eukaryotes (Imai *et al.*, 2000). Other studies, conversely, suggest an increase in DNA condensation with age (Preumont *et al.*, 1983). In *Drosophila*, instability of heterochromatin-mediated position-effect variegation (PEV) has also been observed in post-mitotic cells (Lu *et al.*, 1996), further evidence that changes in chromatin compaction occur during differentiation. Gene-poor human chromosome 18 is preferentially localised towards the nuclear periphery in young, proliferating fibroblasts but is found to relocate to the nuclear interior in senescence (Bridger *et al.*, 2000). This change in position, however, is not accompanied by any detectable change in chromosome volume, so chromatin compaction status in senescence remains open to debate. Using genetic tools, it should be possible to investigate the direct impact of heterochromatin maintenance on life span, providing further insight into this exciting research area.

1.8.4 Differentiation

Global changes in chromatin compaction and distribution within the nucleus are hallmarks of differentiation. In some cell types, these features are so reproducible that identification is possible from nuclear morphology alone: during mouse oogenesis, four stages of chromatin organisation are distinguishable, which correlate with specific stages of the ovarian follicle development (Mattson and Albertini, 1990). A link between increased chromatin compaction and terminal stages of differentiation in vertebrate haematopoietic cells has been identified in the form of MENT (myeloid and erythroid nuclear termination stage-specific protein). MENT is a heterochromatin-associated, serpin family protein, which associates with, and

perhaps facilitates compaction of, peripheral heterochromatin during hematopoiesis (Grigoryev *et al.*, 1999).

In vitro differentiation of myoblasts into myotubes results in the redistribution of centromeres from a random location within the nucleus to the nuclear periphery. This correlates with the increase in heterochromatin at the nuclear periphery observed by TEM (Chaly and Munro, 1996). It has long been observed that nuclear architecture of chromatin is drastically altered during differentiation of neuronal cells (Manuelidis, 1984; Manuelidis, 1990) and centromeres occupy a perinucleolar position during the first day of *in vitro* neuronal growth, but then redistribute at the nuclear periphery (Choh and De Boni, 1996). During differentiation of neuronal PC12 cells, pre-mRNA, Cajal and bodies are translocated towards the nuclear periphery (Janevski *et al.*, 1997). Chromocentres, formed by centromeres from different combinations of human chromosomes, are also characteristic of terminally differentiated myeloid or lymphoid cells (Alcobia *et al.*, 2000). As discussed earlier, differentiation of lymphoid cells is accompanied by characteristic patterns of localisation of developmental specific genes with pericentric heterochromatin, as cells progress along different differentiation pathways (section 1.7.3.2).

1.9 PROPOSED RESEARCH

1.9.1 Background to proposal

As discussed in sections 1.1 to 1.4, chromatin organisation up to the level of the 30nm fibre is a central player in the regulation of gene expression. As I began my PhD, realisation that the impact of chromatin structure on transcriptional regulation does not stop at the 30 nm fibre meant that research on the higher-order chromatin architecture within the context of the nucleus had become an area of intense research. Up to 1998, when I began my project, work relating chromosome territory structure and genome function had been largely correlative. Studies of specific loci had considered only single genes from scattered chromosomal locations. In addition, they were been treated as isolated sequences, rather than part of the linear DNA sequence that comprises the genome. I believed that experiments that directly assay large-scale chromatin structure with respect to gene density and expression at specific, extended loci were needed.

It was well documented that chromosomes occupy discrete domains in interphase and there has been speculation that transcription, RNA processing and RNA transport occur in a space between chromosomes which links to the nuclear pores, the inter-chromosome domain (ICD) (Zirbel *et al.*, 1993). In support of this model, studies of a small number of isolated sequences found genes to be positioned at the surface of chromosomes, independent of their transcriptional status (Kurz *et al.*, 1996).

1.9.2 The initial proposal

This project aimed to address the question of gene positioning within interphase chromosomes by addressing the organisation of sequences in interphase chromosomes in the context of long stretches of contiguous sequence. Genes are not isolated sequences and conclusions drawn from the study of individual gene sequences from scattered chromosomal locations seemed an over-simplification. I proposed to develop an analysis technique to compare the position of loci with respect to the nearest edge of the chromosome territory in both two- and three-dimensional analysis using HSA11p as a model system to determine patterns of interphase organisation of contiguous stretches of the genome. HSA11p has a relatively even distribution of R- (gene-rich) and G- (gene-poor) bands and also an unusually gene-rich sub-telomeric T band, exemplified by 11p13, 11p14 and 11p15.5 respectively.

1 Mb of sequence corresponding to the WAGR (Wilm's tumour, aniridia, genitourinary anomalies, rental retardation) locus from 11p13 was chosen, to determine to what extent the position of a gene versus an adjacent non-transcribed sequence differed within an interphase chromosome. This same contiguous region could also be examined in cell lines expressing different combinations of genes from this locus to determine the influence of transcription status on the positioning of sequence within chromosome territories.

The idea that all genes might be positioned at the surface of an interphase chromosome would incur problems of spatial constraint imposed on chromatin in regions of the genome where genes are closely clustered over many megabases. Therefore, the organisation of two other regions from HSA11p which differ greatly in gene content was considered in an attempt to dissect patterns of organisation of gene-rich and gene-poor chromatin. The organisation within the HSA11p territory of the very gene-dense region within 11p15, containing the imprinted Insulin and Igf2 genes and the gene-poor G-band 11p14 were addressed in an analogous manner.

It was known that murine chromosomes have a territorial organisation in interphase, similar to that of human chromosomes, but the organisation of these domains had not been investigated. I proposed to address the organisation of syntenic regions in the mouse to the human WAGR and BWS loci (located on MMU2 and MMU7, respectively), aiming to draw parallels to the human analysis and to make comparisons between the architecture of chromatin in murine and human nuclei.

I proposed to begin this project using 2D analysis as this had previously been shown to successfully predict 3D patterns of organisation and allows processing and subsequent statistical analysis of a large number of nuclei (Croft *et al.*, 1999; Bridger *et al.*, 2000; Boyle *et al.*, 2001). However, I aimed to later develop an analogous 3D analysis technique to determine the validity of 2D data and to draw parallels between experimentation on material subjected to different fixation procedures.

1.9.3 Project development enabled by emerging resources

The genome sequencing projects will have a potential impact on this area of research and I have been using the emerging sequence to interpret my data. In the future, the availability of finely mapped probes to all regions of genomes and access to large, contiguous stretches of sequence will facilitate the correlation of linear sequence to higher-order chromatin structure and 3D spatial distribution within the nucleus. I proposed to use the probe and sequence facilities currently available to begin such correlation of linear sequence with chromatin architecture. As the sequences of HSA21 and HSA22 became available during this study, I utilised this resource to begin to draw parallels between patterns of organisation of interphase chromosome architecture, observed for my model system HSA11p, and other chromosomes of the human karyotype. As the human genome sequence emerges, it will be possible to use this data to identify gene rich and gene poor regions directly, however, during this study, the accuracy of the published sequence database was limiting.

Contemporary with the development of my project, it became clear that nascent transcripts and both early- (gene-rich) and late-replicating (gene-poor) DNA appeared to be distributed throughout chromosome territories (Visser *et al.*, 1998; Verschure *et al.*, 1999; Zink *et al.*, 1999). However, these studies relied on inference of gene density from replication timing or isochore identity of DNA. Sites of transcription were visualised throughout the nucleus and

protein complexes were observed throughout the nucleus and were not obviously excluded from large regions, which might correspond to chromosome territories (Abranches *et al.*, 1998; Phair and Misteli, 2000). In addition, the MHC genes from HSA6 were observed, not only at the surface of a chromosome territory, but on an external loop of chromatin seen emanating from the chromosome surface (Volpi *et al.*, 2000). Up-regulation of transcription of genes from this locus correlated with an increased incidence of chromatin looping, suggesting a role for transcription in organisation of chromatin within the nucleus.

Consequently, the ICD compartment model has recently been modified to include space between the surfaces of subchromosomal domains (Cremer *et al.*, 2000). However, no further study has been published of the organisation of genes within interphase chromosomes in light of these developments. My data, interpreted in light of studies published by other groups also supports this modified ICD hypothesis.

CHAPTER 2

MATERIALS AND METHODS

2.1 MAMMALIAN CELL CULTURE

2.1.1 Cell Counting

Cells were counted using a haemocytometer (Weber Scientific International Ltd.) which has a chamber 0.1mm in height with an etched grid of 1mm^2 , subdivided into 400 squares. Therefore the total volume defined by the grid was 1×10^{-4} ml and cell concentrations per ml were obtained by multiplying the total number of cells over the grid by 10^4 .

2.1.2 Thawing cells from storage in liquid nitrogen

Cells were stored in liquid nitrogen, suspended in freezing medium (6% DMSO in foetal calf serum (FCS) (Sigma)) at concentrations of approximately 10^6 cells/ml. To thaw, tubes were immediately incubated at 37°C in a beaker of water after excess nitrogen gas had been expelled from the tube by loosening the cap. Cells were diluted in 5ml culture medium and the medium changed after 24 hours incubation at 37°C .

2.1.3 Culture of transformed human cell lines

All cells were incubated at 37°C with 5% CO_2 and tissue culture medium contained penicillin (10000 units/ml) and streptomycin ($650\mu\text{g/ml}$).

2.1.3.1 Suspension cells:

FATO (46 XY) and CV581 (46 XX) are human lymphoblast cell lines and were grown as suspension cultures in RPMI (Life Technologies) supplemented with 10% FCS. Cells were split 1:3 in fresh medium every 2-3 days.

2.1.3.2 Monolayer cells:

The human lens epithelium cell line CD5a (gift of A. Prescott, University of Dundee) was grown in DMEM (Life Technologies) supplemented with 10% FCS and 2mM glutamine. The human ovarian carcinoma cell line COV434 (Berg-Bakker *et al.*, 1993) was grown in 1:1 DMEM: F10 (Life Technologies) supplemented with 10% FCS. Cells from each of these lines were allowed to grow to near confluence before splitting. Medium was poured off, the cells washed twice in PBS (phosphate buffered saline) then covered in a thin layer of 10% trypsin-EDTA (Sigma) and placed at 37°C for 5 minutes. Gentle agitation dislodged the cells, medium was added and the cells were pelleted at 250g for 5 minutes, then re-plated or fixed.

2.1.4 Primary human fibroblast culture

A normal male primary fibroblast cell line, 1HD (Bridger *et al.*, 1998), was obtained from Dr. Joanna Bridger, Brunel (made from foreskin tissue). The cells were grown in DMEM supplemented with 10% FCS. Cultures were grown to almost confluence and split using trypsin-EDTA, as described above. Cells were seeded at concentrations of 1.5×10^5 cells per slide and 5×10^5 cells per 75cm² tissue culture flask. Primary cell lines can be maintained in culture for a limited number of passages before they senesce.

2.1.5 Mouse Embryonic Stem cell culture

The mouse ES cell line, E14 (Hooper *et al.*, 1987) was maintained in MEM/BHK21 medium supplemented with 0.23% sodium bicarbonate, 1x MEM non-essential amino acids, 2 mM glutamine, 1 mM sodium pyruvate, 50 µM β-mercaptoethanol (all Life Technologies), 10% FCS (SIGMA), 100U/ml soluble DIA/LIF (Sigma). Cells were seeded on a gelatinised (0.1% gelatine in PBS) tissue culture surface, fed fresh medium 24 hours later and passaged at a 1:10 dilution the following day.

2.1.6 Preparation of mouse embryonic fibroblasts (MEFs)

Mouse embryonic fibroblasts (MEFs) were prepared by dissecting out the liver and spleen of embryonic day 12.5 embryos and placing the organs in a culture dish with 5 ml media (High glucose DMEM containing 15% FCS, 1% L-glutamate, 1% non-essential amino acids (NEAA) (Life Technologies), 4% β-mercaptoethanol (Sigma). The tissues were

homogenised using a 20-gauge needle then an additional 5ml media added. After 2-3 days in culture, fibroblasts were seen growing out from the tissue masses at which point a brief treatment with trypsin-EDTA was performed. The volume of the resultant cell suspension was made up to 10ml and then left to settle for 1-2 minutes to allow lumps of tissue to settle out. The remaining cells in suspension were passage 1 (P1) fibroblasts and were plated out, then split 1:3 (P2) when confluent. This protocol is based on that of Jenuwein and Muller, 1987.

2.2 PREPARATION OF NUCLEI FOR FLUORESCENCE *IN SITU* HYBRIDISATION

2.2.1 Harvesting and fixing cells in 3:1 methanol:acetic acid (MAA)

After harvesting, cells were suspended in a mildly hypotonic solution, 0.033M KCl, 0.017M tri-NaCitrate. The hypotonic solution was added drop-wise while continually agitating the tube to a volume of 10ml (the concentration of cells in hypotonic should be $<2 \times 10^7$ /ml). The cell suspension was left at room temperature for 10 minutes before centrifuging at 400g for 5 minutes. Cells were then fixed with fresh 3:1 methanol: glacial acetic acid. 2ml of fix were added drop-wise to the cells while the tube was agitated. A further 8ml of fix were added and the tubes were placed on ice for a minimum of 15 minutes or stored at -20°C overnight. Cells were fixed twice more and stored indefinitely at -20°C .

2.2.2 Preparation of three-dimensionally preserved nuclei

1HD human fibroblasts were grown on permafrost microscope slides (BDH laboratory supplies) that had been previously been boiled in detergent for 10 minutes, washed 10 times each in tap water and distilled water, then stored in methanol. Each slide was flamed and placed in a quadriperm slide chamber (Satorius). 1.5×10^5 cells were seeded on each slide in 5ml medium. Seeding at this density results in approximately 70% confluency (the optimum density of cells for 3D FISH) after a further 16 hours in culture.

Slides were washed 3 times for 3 minutes in PBS, then fixed in 4% paraformaldehyde (pFA) dissolved in PBS for 10 minutes (by heating the suspension to 60°C to dissolve the pFA and then cooling to room temperature again before use). Slides were then washed 3 times for 3 minutes (3 x 3 minutes) in PBS and cells permeabilised in 0.5% saponin/ 0.5% Triton X-100 in PBS for 14 minutes. All incubations were carried out at room temperature. The slides were washed 3 x 3 minutes, in PBS before incubation in 20% glycerol in PBS, either for 30 minutes at room temperature or overnight at 4°C.

Slides were dipped into liquid nitrogen until completely frozen, allowed to thaw slowly to room temperature, then returned to 20% glycerol for about 3 minutes before being frozen again. This freeze-thaw procedure was repeated 5 times and, after the fifth freezing, slides were either taken directly through the 3D FISH protocol (Section 2.6) or stored indefinitely at -70°C.

2.3 PREPARATION OF NUCLEIC ACIDS

2.3.1 Isolation of RNA

RNA was prepared from cultured cells using a Bio/RNA-Xcell total RNA isolation kit (Bio/Gene Ltd.). 5×10^6 cells were harvested prior to RNA isolation (as described in Section 2.1.3) and the kit used as per the manufacturer's recommended protocol. The RNA concentration was measured as described in Section 2.3.6.

2.3.2 Preparation of 1st strand cDNA

10µg RNA in dH₂O were added to 400ng hexanucleotide primers (Roche), 100µM dNTPs (final concentration), 100µM Dithiothreitol (DTT) (final concentration), 80U RNase inhibitor (Roche) in a final volume of 38µl. This was split into 2 and 20U M-MuLV reverse transcriptase (RT) enzyme added one tube, the other left as a control. The two reactions were incubated at 42 °C for 2 hours.

2.3.3 Isolation of DNA from cosmid clones

Bacterial stocks stored at -70°C in 25% glycerol were streaked onto L-agar plates containing the appropriate antibiotic (ampicillin at 25 $\mu\text{g}/\text{ml}$, chloramphenicol at 25 $\mu\text{g}/\text{ml}$, or kanamycin at 30 $\mu\text{g}/\text{ml}$) and grown overnight at 37°C . 3ml cultures were inoculated with single colonies and grown overnight in 'snap top' 15ml tubes, shaking at 37°C . 1.5ml of the culture suspension was spun down in an eppendorf tube at 12000g for 2 minutes, then the supernatant removed and the pellet resuspended in 200 μl cold GTE buffer (50mM glucose, 25mM Tris-HCl pH 8.0, 10mM EDTA pH 8.0) with approximately 50 $\mu\text{g}/\text{ml}$ lysozyme and left at room temperature for 5 minutes. 400 μl cold lysis buffer were added (0.2M NaOH, 1% SDS), the solution mixed by inversion and left on ice for 5 minutes. 300 μl acetate buffer were added (3M potassium, 5M acetate made by adding 60ml 5M KAc to 11.5ml glacial acetic acid and 28.5ml distilled water), the suspension mixed by vortexing gently and then left on ice for 5 minutes before centrifuging for 5 minutes at 4°C , 13000g. The supernatant was transferred to a fresh eppendorf tube and an equal volume of phenol:chloroform added (1:1) and mixed by inversion. After centrifugation for 2 minutes at 4°C , 12000g in a microcentrifuge, the top aqueous layer was transferred to a fresh eppendorf tube and an equal volume of isopropanol added. After storage at -20°C for 1 hour and centrifugation for 15 minutes at 4°C , the pellet was washed twice in 70% ethanol and resuspended in 50 μl 50mM Tris-EDTA (TE) pH 7.4, (10mM Tris-Cl pH7.4, 1 mM EDTA pH8.0)

2.3.4 Isolation of Bacterial Artificial Chromosome (BAC) and Plasmid artificial chromosome (PAC) clones

BAC and PAC clone stocks, stored at -70°C in 25% glycerol, were streaked onto L-agar plates containing the appropriate antibiotic, at concentrations listed in 2.3.3. A rapid alkaline lysis mini-prep method based on a Quiagen-Tip method as modified by P. de Jong was used. 5ml cultures were grown overnight, as described for cosmid isolation above, then spun at 2000g at room temperature. Supernatants were discarded and the cells resuspended in 300 μl 15mM Tris HCl pH 8.0, 10mM EDTA, 100 $\mu\text{g}/\text{ml}$ RNase A and the suspension transferred to an eppendorf tube. 300 μl 0.2M NaOH containing 1% SDS were added and the solution left at room temperature for 5 minutes. 300 μl of 3M KOAc, pH 5.5 were then added and the suspension left on ice for 5 minutes then spun at 10000rpm for 10 minutes at 4°C . The

resultant supernatant was transferred to an eppendorf containing 800µl ice cold isopropanol, mixed by inversion then spun at 13000rpm for 15 minutes. After washing twice in 70% ethanol and air drying, the pellet was resuspended in 50µl TE.

2.3.5 Ethanol precipitation

DNA was precipitated by adding 1/3 volume of 3M sodium acetate and 3 volumes of 100% ethanol. Samples were incubated at -20°C for a minimum of 1 hour or overnight before centrifuging at 12000g for 15 minutes at 4°C. DNA pellets were washed twice in 70% ethanol, air-dried and suspended in the appropriate volume of TE buffer. Samples were stored at 4°C.

2.3.6 Measuring quality and quantity of DNA and RNA

DNA and RNA concentration was measured spectrophotometrically and/or by gel electrophoresis alongside DNA of known concentration (Section 2.3.8). DNA was diluted 1:100 in TE and transferred to a quartz cuvette. The absorbance or optical density (OD) at 260nm (A_{260}) was measured for DNA and for RNA. An OD of 1 corresponds to ~50µg/ml of DNA and to 38µg/ml RNA. To determine the purity of nucleic acid, the A_{280} was also measured. Pure DNA has an A_{260} / A_{280} of 1.8 and values lower than these indicate contamination by proteins, RNA or phenol. Pure RNA has an A_{260} / A_{280} of approximately 1.8-2.0.

2.3.7 Reverse transcription polymerase chain reaction (RT PCR)

Primers were obtained for PCR amplification of each of the human WAGR locus genes (*WT1*, *RCN*, *PAX6* and *PAXNEB*) from cDNA of FATO, 1HD, CD5a and COV434 cell lines (Sections 2.3.1 and 2.3.2). Sequences of these primers and the sizes of the fragments generated are recorded in Table 2.1. In each reaction, 10 ng DNA template 200ng of primers, 200 µM dNTPs and 3U of Taq polymerase (Roche) were used in 1x Cetus buffer II and 250 µM MgCl₂ and (Roche). The PCR program conditions are described in Table 2.2.

Table 2.1 Primers for amplification of human WAGR gene sequences

Gene	Primer	Sequence (5' to 3')	Size of fragment (bp)	Source/ reference
WT1	T39 (D609), exon 1	CAA ACA GCA GCC GAG CTG G	686	Williamson, 1996
	T41 (D610), exon 6	GCA CAT CCT GAA TGC CTC TG		
RCN	T37 (G829)	GGA CAA CCA GAG CTT CCA GTA CGA	795	Kent <i>et al.</i> , 1997
	T38 (G830)	TTG GCT TCC CAC AAA CAT RTT CCA		
PAX6	B251 (M463), exon 7	TCA ATA AAC AGA GTT CTT	898	Hanson <i>et al.</i> , 1993
	C402 (M465), exon 13	AAT CTT GGC CAG TAT TGA		
PAXNEB	O23, exon 7	CTC TCT TTC AGC CTG CAT C	396	D.A. Kleinjan
	O41, exon 10	TCC CTA GTC CAG GT GC		

Table 2.2 Human WAGR RT-PCR Programmes

Gene	Step 1		Step 2		Step 3		Step 4	
	Temp (°C)	Time (sec)	Temp (°C)	Time (sec)	Temp (°C)	Time (sec)	Temp (°C)	Time (sec)
<i>WT1</i>	94	300	94	60	60	45	72	60
<i>RCN</i>	94	300	94	60	58	45	72	90
<i>PAX6</i>	94	300	94	60	58	45	72	90
<i>PAX-NEB</i>	94	300	94	45	56	45	72	45

All PCR products were resolved on 1.5% agarose gels (Section 2.3.8).

2.3.8 Electrophoresis of DNA

The electrophoresis buffer was 1x TAE (90 mM Tris-HCl, 90 mM glacial acetic acid, 2 mM EDTA pH8.0). Horizontal agarose gels (0.5-1% in 1xTAE) were used to resolve DNA molecules of 0.2 to 50 kb. To resolve small fragments (0.05 to 5Kb), horizontal 1.5-2% gels were used.

To ensure effective loading of DNA samples onto gels, 1 volume of 6x loading buffer (15% Ficoll 400, 0.25% Bromophenol Blue, 0.25% xylene cyanol) was added to 5 volumes of DNA, to increase the density of the samples. Dyes included in this buffer also provided convenient size markers whose electrophoretic mobility depended on gel concentration. Commercially available DNA size markers (ϕ X174 DNA *HaeIII* digest and λ DNA *HindIII* digest, Promega) were diluted to 50ng/ μ l and 500ng loaded per well.

The DNA-intercalating fluorescent stain, ethidium bromide (2,7-diamino-10-ethyl-9-phenyl-phenanthridium bromide), was used to stain DNA gels. Gels were soaked in TAE containing ethidium bromide, with mild agitation for 30 minutes, then de-stained in dH₂O for a further

30 minutes. Alternatively, ethidium bromide was added to molten agarose before pouring of a gel. Stained DNA was visualised with UV light from a transilluminator. Gels were photographed using an Appligene (Oncor) television camera and images printed using a Mitsubishi video copy processor.

2.4 PREPARATION OF FLUORESCENCE *IN SITU* HYBRIDISATION (FISH) PROBES

DNA was labelled using biotin-16-dUTP or digoxigenin-11-dUTP. These analogues were incorporated into DNA by PCR or by nick translation. Following either method, unincorporated nucleotides were removed and efficiency of labelling was assessed.

2.4.1 Preparation of labelled mouse and human chromosome paints by PCR

Human (gifts of Michael Bittner, (Guan *et al.*, 1996) and mouse (obtained from I Solovei) micro-dissected chromosome arms were amplified on a PTC-225 DNA Engine (MJ Research).

Biotin-16-dUTP or digoxigenin-11-dUTP was incorporated directly by PCR into chromosome paints by adding these analogues to the PCR reaction mix (used for HSA11p, 15q, and 16p and MMU2 and MMU7). For direct labelling of chromosome paints, 1 µl of PCR template (from a second round of amplification of original template stock) was added to 5µl each of 2 mM dATP, dCTP, dGTP, 3 µl 0.5 mM dTTP, 3 µl 1 mM biotin- or dig-dUTP, 400 ng primer (Human 5'CCGACTCGAGNNNNNNATGTGG^{3'}; Mouse 5'CCGACTCGAGNNNNNNNTACACC^{3'}), 3 U Taq polymerase (Roche), 5 µl 25mM MgCl₂, 5µl Cetus PCR buffer II (10x) and the volume of the reaction made up to 50µl with distilled water. Amplification conditions were 3 minutes at 94 °C then 30 cycles of 94 °C for 30 seconds, 56 °C for 30 seconds and 72 °C initially for 2 minutes extending by 3 seconds per cycle.

Alternatively, when the PCR products were too large (bigger than about 500bp), cold PCR products were precipitated using ethanol and NaAcetate (Section 2.3.5), pH5.5, then biotin- or dig-labelled by nick translation (Section 2.4.2) (used for HSA21q and 22q).

All chromosome PCR fragments were resolved on 2% agarose gels in TAE (Section 2.3.8).

2.4.2 Nick translation

1-1.5 μ g of DNA were added to 4 μ l 10x nick translation salts (0.5M Tris-HCl pH7.5, 0.1M MgSO₄, 1mM DTT, 500 μ g/ml Bovine serum albumin (BSA)), 4 μ l each of 2mM dATP, dGTP and dCTP, 2 μ l of 0.5mM dTTP and 4 μ l 1mM biotin-16-dUTP or digoxigenin-11-dUTP (Roche). DNase I (Gibco BRL) was freshly diluted to a concentration of 20units/ml in dH₂O at 4°C and 2 μ l added to the reaction mixture to give a final concentration of 1unit/ml. After the addition of 1 μ l DNA polymerase I (Gibco BRL, 10units/ μ l), dH₂O was added to make the total volume of reaction mixture 40 μ l. The reaction was mixed thoroughly and allowed to proceed at 16°C for 90 minutes. Placing at -20°C stopped the reaction.

2.4.3 Removal of unincorporated label

Quick spin columns (Roche) containing G50 Sephadex beads were used to remove any biotin-16-dUTP or dig-11-dUTP that remained free in solution as per the manufacturer's instructions.

2.4.4 Quantifying label incorporation

Gridded nitrocellulose membranes were prepared by soaking briefly in dH₂O followed by 20x SSC (3M NaCl, 0.3M tri-sodium citrate, pH7.4) for 10 minutes then allowing to air dry. Labelled DNA probes were diluted to 1x10⁻³ and 1x10⁻⁴ in TE and 1 μ l of each was spotted twice onto a gridded membrane. After the spots had dried, a further 1 μ l was added to one of each dilution. On the same membrane 20, 10, 2 and 1pg of appropriately labelled lambda DNA standards (Roche) were spotted. DNA was cross-linked onto the membrane by exposure to 30 mJ of UV irradiation.

The membrane was immersed in buffer 1 (0.1M Tris-HCl pH7.5, 0.15M NaCl) for 5 minutes at room temperature, then in 3% BSA in buffer 1 at 60°C for 45 minutes. 10µl streptavidin-alkaline phosphatase (Boehringer) and/or anti-digoxigenin-alkaline phosphatase (Boehringer) were added to 10ml of buffer 1 and placed in a sealed polythene bag with the membrane for 30 minutes at room temperature with continuous agitation. The membrane was washed twice for 15 minutes in buffer 1 then for 5 minutes in 0.1M Tris-HCl, pH9.5. The colour reaction was developed by incubation of the membrane, in a sealed polythene bag, with 5ml of 0.1M Tris-HCl, pH9.5 and 2 drops from bottles 1-3 from the alkaline phosphatase substrate kit IV (Vector). The substrates in this colour reaction are 5-bromo-4-chloro-3-indolyl phosphate and nitroblue tetrazolium, which produce a blue reaction product. A complete colour reaction was observed within a few hours and an estimate of the concentration of DNA labelled probe was made by comparison with the lambda standards.

2.5 FLUORESCENCE *IN SITU* HYBRIDISATION (FISH) ON MAA-FIXED NUCLEI

The technique of fluorescence *in situ* hybridisation (FISH) is used to determine the chromosomal origin of isolated DNA (Langer-Safer *et al.*, 1982). Methods used here were developed by Dr. J.A. Fantes, Edinburgh University (Fantes *et al.*, 1992).

2.5.1 Slide preparation

Glass slides were prepared by soaking in detergent solution and washing with dH₂O before storage in a dilute solution of HCl in ethanol. Immediately prior to slide making, slides were polished with muslin.

Methanol:acetic acid fixed cells (Section 2.2.1) were removed from storage at -20°C, left to warm to room temperature for 30 minutes and centrifuged at 400g for 5 minutes. Fresh MAA fix was added until the cell suspension reached a 'milky' density. A single drop of cell suspension from a narrow-mouthed pastette was dropped onto a horizontal microscope slide from a height of about 30cm. The spread of cells on the slide was improved by coating the slides with a thin layer of moisture, usually by breathing. An air humidity of ~50% also

aided spreading. Spreading was monitored by phase contrast microscopy. Slides were stored for 2-6 days prior to hybridisation. When fixed material of more than 2 years old was to be used, or there was a high level of cytoplasm around the nuclei, slides were treated with pepsin prior to hybridisation (Section 2.5.2).

2.5.2 Pepsin treatment

MAA nucleus preparations may not hybridise with a high efficiency after the 2D FISH protocol if they still have cytoplasm attached. To increase the efficiency of hybridisation, these slides were treated with pepsin. Slides were dehydrated in acetone for 5 minutes then dried under vacuum. RNase treatment was then carried out as in Section 2.5.3. 43µl of 11M HCl was added to 50ml of dH₂O and pre-heated to 37°C then 125µl of 2% pepsin was added and the slides incubated for 5 minutes before washing in PBS with 50mM MgCl₂. Slides were subsequently dehydrated through an ethanol series and hybridisation was carried out as described in Section 2.5.3.

2.5.3 Hybridisation

Slides were mounted vertically in a metal slide rack and subsequent incubations carried out in 200ml glass troughs. Slides were first treated with 100µg/ml RNase in 2x SSC for 1 hour at 37°C, washed briefly in 2x SSC and dehydrated through an ethanol series (2 minutes each in 70%, 90% and 100% ethanol). The slides were left to dry under vacuum for 10 minutes before being heated in a 70°C oven for 5 minutes and immediately denatured in 70% formamide, 2x SSC at 70°C for 2-3 min. After passing through 70% ethanol at 4°C and an ethanol series as above, slides were again vacuum dried.

Concurrent with slide preparation the probes were prepared. Labelled DNA (Section 2.4) (200ng human or mouse chromosome paint, 50ng cosmid, 100ng BAC, 100ng PAC), usually suspended in TE, were precipitated with 5µg salmon sperm DNA and mouse or human C₀t 1 DNA (the amount of which varied from 5 to 50µg, depending on the potential repeat content of the probe) (Gibco BRL). After the addition of two volumes of ethanol, probes were spun down under vacuum until they had precipitated, and re-suspended in 10µl hybridisation mix (50% deionised formamide, 10% dextran sulphate, and 1% Tween 20, in 2x SSC). Commercial probes were usually provided in or with hybridisation buffer and did not require

addition of salmon sperm DNA or human C₀t 1 DNA. All probes were denatured at 70°C for 5 minutes and reannealed at 37°C for 15 minutes before spotting onto pre-cleaned coverslips. The denatured slides were carefully laid onto the appropriate coverslip and sealed with rubber solution (TipTop) before placing in a metal tray in a 37°C water bath overnight.

2.5.4 Washing and detection

Slides were washed in glass racks in 200ml troughs. After removal of the rubber cement, they were immersed in 2x SSC at 45°C for 3 minutes, with gentle agitation to facilitate detachment of the coverslips. Slides were washed a further 3x in the same buffer then for 4x 3 minutes in 50% formamide/2x SSC at 45°C and 4x 3 minutes 0.1xSSC at 60°C before transferring to 4x SSC/ 0.1% Tween 20. Detection was carried out in a moist chamber pre-heated to 37°C. Biotin was detected with sequential layers of fluorochrome-conjugated avidin (Fluorescein isothiocyanate- (FITC) or Texas Red- (TR) avidin), biotinylated anti-avidin, and a further layer of fluorochrome-conjugated avidin. Digoxigenin was detected with sequential layers of FITC-or Rhodamine-conjugated anti-digoxigenin and FITC- or TR-conjugated anti-sheep antibodies. Detection reagents were diluted in SSCM (4x SSC, 5% Marvel dried skimmed milk) to the appropriate concentration (Table 2.3). After incubation with 40µl SSCM for 5 minutes at room temperature, 40µl of the appropriate detection layer were applied to each slide and covered with a square of parafilm. Slides were incubated in the moist chamber at 37°C for 45 minutes, followed by 3 washes of 2 minutes in 4xSSC/0.1% Tween20 at 37°C.

All slides containing human cells were mounted with 1µg/ml 4,6-diamidino-2-phenylindole (DAPI), in Vectashield (Vector), 0.5µg/ml for mouse. Coverslips were sealed with rubber solution (Pang) and slides were stored in the dark at 4°C.

2.6 FISH ON THREE-DIMENSIONALLY PRESERVED NUCLEI

This method was used to maintain the three-dimensional architecture of nuclei (Kurz *et al.*, 1996; Croft *et al.*, 1999) through the process of FISH.

Cells were grown and fixed on slides, as described in Section 2.2.2. Slides were placed in PBS for at least 2x 30 minutes then incubated in 2x SSC (300mM NaCl, 30mM tri-sodium citrate, pH7.4) containing 100µg/ml RNase for 1 hour at 37°C. Following this, slides were returned briefly to PBS and then placed in 0.1M HCl in dH₂O for 7 minutes before being returned to PBS again. Slides were then denatured in 70% formamide, 2xSSC (pH7.0) for 3 minutes then 50% formamide, 2xSSC (pH7.0) for 1 minute, both at 75 to 78°C. Probes were applied immediately after the second formamide incubation. Probes were prepared, hybridised and detected as described in 2.5.3 and 2.5.4. Chromosome domains were always detected using FITC-conjugated secondary antibodies and cosmid, BAC or PAC probes using Rho- or TR-conjugated secondary antibodies.

Table 2.3 Antibodies and fluorochrome-conjugates used for FISH

Antibody or fluorochrome-conjugate	Animal in which antibody raised	Source	Stock concentration (mg/ml)	Dilution
FITC-avidin D CS	Goat	Vector	2.0	1:500
TR-avidin D CS		Vector	2.0	1:500
Biotinylated anti-avidin D	Goat	Vector	0.5	1:100
Anti-sheep-FITC	Goat	Vector	1.5	1:100
Anti-digoxigenin-FITC F(ab')₂	Sheep	Roche	0.2	1:40
Rhodamine-anti-digoxigenin	Sheep	Roche	0.2	1:20
Anti-sheep-TR	Rabbit	Vector	0.5	1:100

CS - cell sorting grade

2.7 IMMUNOFLUORESCENCE

For immunofluorescent detection of WT1 epitopes using 3:1 fixed nuclei, slides were prepared as described in Section 2.5.1 and aged overnight. Slides were rinsed in PBS, then blocked for 20 minutes in 5% donkey serum/PBS (serum to dilute antibodies should be from the animal in which secondary antibodies were raised) under parafilm. A 1:250 dilution of a 0.2 mg/ml stock of the C19 antibody (Santa Cruz Biotechnology) was made in 5% serum/PBS and hybridised to the slides under parafilm for 45 minutes at room temperature. Slides were then washed 3 times in PBS. A 1.0 mg/ml stock of a FITC-anti-rabbit secondary antibody (raised in donkey) (Jackson labs), was diluted at 1:150 in 5% serum/PBS and applied in the same manner. The slides were washed again and then mounted in DAPI.

2.8 CAPTURE AND ANALYSIS OF 2D IMAGES

After FISH or immunofluorescence, 2D slides were examined using a Zeiss Axioplan fluorescence microscope with a 100 Watt mercury bulb and equipped with a triple band-pass filter (Chroma #83000). Grey scale images for each fluorochrome were collected with a cooled CCD camera (Princeton Instruments Pentamax).

Microscope slides were scanned in a methodical manner, beginning at the top left hand corner, scanning to the right, then moving downwards and scanning the next row of nuclei from right to left, etc. Fifty bin2 images were collected of consecutive nuclei that fulfilled criteria for images to be suitable for analysis. Nuclei had to be intact and contain two chromosome domains in which locus signals were visible; in addition, analysis scripts (Section 3.4.1) required images of single nuclei not touching any other.

2.9 ANALYSIS OF 3D IMAGES

2.9.1 Image capture and processing

3D stacks of three colour images (cell nucleus DAPI, chromosome territory FITC, locus TR) were captured using an Axioplan microscope fitted with a 100 Watt mercury bulb, Ludl filter wheel Chroma filter set #81000 and motorised stage, and attached to a cooled CCD Kodak KAF 1401e sensor camera (Princeton Instruments). A script was devised (P. Perry) to capture bin 1 resolution level images: The DAPI excitation filter is selected and a “live” focus window displayed. The plane of best focus is interactively selected, as is a region of interest for capture. The number of optical slices to be collected, the z distance between each optical slice and the exposure times for each fluorochrome signal (from an automatic test capture in the plane of best focus) are also determined. The microscope focus motor then moves the stage downward for half the total depth of the focus series to the starting point for capture. To compensate for backlash in the focus mechanism, the stage is moved downwards a further 200 microns, followed by an upward movement of the same distance. Image

capture starts, collecting the specified DAPI FITC or Texas Red images at each focal plane and placing each into a stack file, then moving the stage upwards before repeating the same capture sequence. After the final plane has been captured, the stage is returned to the original "best focus" focal plane. The stack files are then merged to provide a colour stack file that can be animated or projected. For each nucleus 25 to 30 bin1 image planes were captured at 0.5 μm intervals, so as to include the whole of the nucleus in the image stack.

Signals from the chromosome territory (detected using FITC) and the cosmid, BAC or PAC probe (detected using TR) were converted to grey-level images and analysed by segmenting the regions corresponding to the nuclear territory and probe. The segmentation of each 3D region or domain was performed manually using the program MAPaint, developed at the MRC Human Genetics Unit for 3D-image segmentation. This program allows manual delineation using simple line drawing or direct thresholding of the grey-level image. For each pair of images the cosmid and nuclear territory signals were delineated and saved to disk as independent 3D binary images ('.wlz' files). The delineation of the chromosome domain was established using thresholding, which automatically segments connected regions within the image that have the same signal intensities. Therefore the painted chromosome domain could be distinguished from background signal covering the rest of the nucleus. The small probe signal was delineated using the "paint-ball" facility, as the limits of these small signals were obvious within the image planes.

2.9.2 Script analysis of 3D image stacks

Files corresponding to matching chromosome territory and locus signal domain pairs were then passed to a program (written by R. Baldock, MRC Human Genetics Unit) that then calculated the following quantities:

- volume of the probe domain (PD)
- volume of the chromosome territory domain (CTD)
- distance between the centres-of-mass (CM) of each domain
- minimum distance between the CM of the PD and the surface of the CTD
- minimum distance between the CM of the PD and the surface of the CTD
- minimum distance between the surfaces of the two domains

In each case the spatial co-ordinates were normalised using the measured voxel size (x 1, y 1, z 5) to compensate for 0.5 μm spacing between consecutive image planes. Distances

between the centre of the probe domain and the surface of the chromosome territory domain were annotated as negative when the probe was not enclosed within the chromosome domain (as in 2D analysis). Data outputs were imported into Microsoft Excel for statistical analysis (Section 2.10).

The programs used in this analysis are freely available for academic non-commercial use for Unix machines (MRC Human Genetics Unit Mouse Atlas Project, <http://genex.hgu.mrc.ac.uk/>).

2.9.3 Visualisation of domains within image stacks by 3D surface rendering and movie making to view stacks by rotation around 360°

Segmentation of the chromosome territory and small probe signal 3D domains was performed as described in Section 2.9.1. For visualisation of the chromosome and probe domains in context, the nucleus was also segmented, using the same ‘thresholding’ facility, and fuzzy edges were refined using the paintball feature. The delineation of the nucleus domain required a certain amount of interpretation in image planes which were out of focus, but no quantitative analysis was done on the position of chromosomes or probe signals within the nucleus domain, as defined using the programs described in Sections 2.9.

To visualise the probe, chromosome territory and nuclear domains I used the AVS/Express visualisation system (Advanced Visual Systems Inc. USA) to produce a surface-rendered 3D view. Each domain file recorded true (x , y , z) co-ordinates which were read into AVS/Express and converted into a 3D surface. The system then allowed interactive manipulation of the viewing parameters of each surface including orientation, scale, colour, lighting and transparency. The resulting 3D view showed the nucleus as transparent, the chromosome territory as green and the locus probe as red.

This 3D view was used to make an MPEG* movie using the AVS/Express “animator module”. The 3D image was rotated around the x -axis and ten key frames were chosen, each 36 degrees apart. The “animator module” was then used to create twenty interpolated frames between each key frame to create a smooth animation. The AVS/Express “image capture” module saved each of the 200 images were then re-formatted as an MPEG movie file.

*MPEG: This ‘Moving Picture Experts Group’ format used to image sequences is a public domain compression format with viewers for all workstations (<http://www.mpeg.org>).

2.10 STATISTICAL ANALYSIS

The statistical analyses described here were performed by A. Carothers, MRC Human Genetics Unit. Output data from the 2D scripts described in 2.8 (run on images of fifty nuclei per probe) and the 3D script described in 2.9.2 (run on >15 images for each probe) were analysed in terms of the response variable, y , defined as:

$$y = (\text{probe centre to chromosome edge/territory radius})$$

This represents the distance of the centre of the cosmid, BAC or PAC probe signal to the nearest edge of the chromosome territory as defined by a chromosome paint, normalised for a territory radius defined as $(\text{territory area}/\pi)^{1/2}$ for 2D images and $(3 \times \text{territory volume}/4\pi)^{1/3}$ for 3D images.

Data sets for the WAGR locus (11p13 and MMU2 probes, Chapters 3 and 5) were analysed by conventional analysis of variance (ANOVA) using MINITAB (Release 12.21). Sources of variation were: PROBES crossed with G1/G2 status; NUCLEI nested within PROBES; TERRITORIES nested within NUCLEI; SIGNALS nested within TERRITORIES. Preliminary inspection showed that negative values of y were strongly skewed violating a basic assumption required for valid ANOVA. To avoid this problem, negative values of y were transformed by:

$$y_T = \{\exp(3y)-1\}/3$$

The effect was to leave small values of y almost unchanged, whereas large negative values were reduced in magnitude making the overall distribution of y_T closer to Normal. Positive values of y were left unchanged.

The results of all analyses on WAGR showed that (i) the component of variance between TERRITORIES nested within NUCLEI (V_{TER}) was significantly greater than that between

SIGNALS nested within TERRITORIES (V_{SIG}) and (ii) the component of variance between NUCLEI nested within PROBES (V_{NUC}) was not significantly greater than V_{TER} . The territory was therefore taken as the basic unit of observation. It was also found that in all analyses there were no significant differences arising from G1/G2 status and these were also therefore ignored. The effect of these simplifications was that experiments could be regarded simply as a series of replicate measurements on each probe, with each territory as the unit of replication.

CHAPTER 3

THE ORGANISATION OF 1 MB FROM 11p13 WITHIN INTERPHASE HSA11p

3.1 INTRODUCTION

Within the context of the interphase nucleus chromosomes occupy defined territories. While it is becoming increasingly obvious that the nuclear position of a gene can influence its expression (Section 1.7), the specific organisation of DNA sequences within interphase chromosome territories remains unclear.

3.1.1 The positioning of genes within a chromosome territory

The proposal of the ICD model of nuclear organisation (Section 1.6.2) was the impetus for studying the position of genes within chromosome territories. The first study to assess the interphase positioning of genes within chromosome territories considered only a small number of gene loci from scattered chromosomal locations, but suggested that genes are preferentially located at the surface of interphase chromosomes, independent of transcription status (Kurz *et al.*, 1996). In contrast, a non-expressed anonymous fragment was found randomly distributed or localised preferentially in the territory interior. The most extreme interpretation of this data is that most, or all, genes lie on the surface of chromosome territories. Depending on how the 'periphery' of a chromosome domain is defined (half the volume of a sphere is located in an outer shell with a radius 1/5 of the radius of the whole sphere), there is plenty of non-coding DNA in a mammalian genome available to fill the interior chromosome space. But the idea that genes are specifically located at the surface of an interphase chromosome and non-coding DNA is more internally positioned does not fit so neatly when one considers regions of a mammalian genome where genes are packed very closely over many megabases. Similarly, entire genomes of other species have a high gene density and therefore little non-coding DNA to fill the interior of a chromosome domain, e.g. *Fugu rubripes* (Brenner *et al.*, 1993; Elgar *et al.*, 1996).

3.1.2 Is chromosome territory organisation dependent on transcription?

Dramatic changes in overall nuclear morphology and intranuclear chromatin distribution accompany differentiation, mitogen stimulation and transformation, although the functional significance of these changes remains unclear (Brown *et al.*, 1997; Brown *et al.*, 1999; Brown *et al.*, 2001; Bridger *et al.*, 2000). Therefore, is chromatin specifically rearranged within a chromosome territory when the transcription profile of a cell is altered? Two studies, to date, have reported transcription-dependent positioning of genes within an interphase chromosome. On the X-chromosome, an active gene was found in a more peripheral location within the Xa chromosome territory than its counterpart Xi, in which the gene is inactive (Dietzel *et al.*, 1999). Secondly, transcriptional up-regulation of genes from the MHC locus led to an increased incidence of extrusion of a large chromatin loop from the surface of chromosome 6, on which MHC genes were located (Volpi *et al.*, 2000). These studies suggest that active transcription contributes to the positioning of genes within an interphase chromosome domain.

Sites of transcription appear throughout the entire nucleus and are not excluded from within chromosome territories (Grande *et al.*, 1997; Abranches *et al.*, 1998; Verschure *et al.*, 1999; Cmarko *et al.*, 1999, reviewed Szentirmai and Sawadogo, 2000). In addition, both early (gene-rich) and late replicating (gene-poor) DNA is distributed throughout chromosome territories (Visser *et al.*, 1998) with no exclusion of gene-rich DNA from the chromosome interior (Tajbakhsh *et al.*, 2000). Taking these data into consideration, there seems to be little requirement for genes to be strictly exposed on the surface of interphase chromosomes, and therefore little need for large-scale chromatin re-modelling on altered transcription status of individual genes. The question of gene positioning within interphase chromosome territories therefore remains open.

3.1.3 Investigating territory organisation using contiguous sequence from HSA11p and MMU2

To assess in more detail the position of genes versus non-coding DNA within a chromosome territory, and also to determine whether large-scale chromatin organisation is associated with gene expression status, I decided that it was important to analyse a contiguous stretch of DNA. On beginning this work there were no published studies in the literature that attempted to dissect patterns of organisation of a large stretch of DNA within a single interphase

chromosome domain in a methodical manner. One such study has since been published (Volpi *et al.*, 2000).

I decided to proceed initially with analysis in two dimensions (2D), to enable statistically significant examination of large numbers of nuclei. Data produced using 2D FISH analysis has previously been used effectively to predict three dimensional organisation (Croft *et al.*, 1999; Volpi *et al.*, 2000; Boyle *et al.*, 2001). The mathematical modelling principals (stereo projection) behind using 2D analysis to predict 3D organisation have also been considered (Carothers, 2000).

I chose to initially focus on the short arm of human chromosome 11 (HSA11p) and in particular on band 11p13. HSA11 is a metacentric chromosome of middling size, consisting of 142 Mb DNA (<http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/maps.cgi?org=hum&chr=11>) of which approximately 56Mb are contained in HSA11p (excluding centromeric sequences). It contains a relatively even distribution of both G- and R-bands and genomic sequence so far predicts the presence of 988 genes (e.g. <http://www.ncbi.nlm.nih.gov:80/cgi-bin/Entrez/maps.cgi?org=hum&chr=11>). However, as of September 2001, finished human sequence represents only 47.1% of the genome and known genes still remain to be mapped onto the draft sequence (NCBI). Cytogenetic maps suggest there are up to approximately 1150 genes present on HSA11 (<http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/maps.cgi?ORG=hum&MAP0=gene&CHR=11&BEG=&END=&MAP1=ideogr&MAP2=cntg&MAP3=est&MAP4=&MAP5=&MAP6=&VERBOSE=ON&SIZE=20>). 11p13 is a moderately gene-rich R-band midway along the short arm of HSA11 (Figure 3.1a).

No study has been made so far of the organisation of murine chromosome territories. As in the human nucleus, mouse chromosomes occupy discrete domains at interphase. I decided to address the organisation of the mouse nucleus in parallel, to see what comparisons could be made between mammalian species. If any patterns of spatial organisation that I observe in human chromosomes is of fundamental significance, I would expect them to be conserved in evolution. The murine region of conserved synteny to 11p13 is located in MMU2, the second largest mouse chromosome (209 Mb) (<http://www.ncbi.nlm.nih.gov/genome/seq/MmProgress.shtml>), which contains an estimated 750 genes (NCBI <http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/maps.cgi?org=mouse&chr=2>), although finished sequence for this chromosome is only 0.7% complete (as of September 2001, NCBI) (Figure 3.1b).

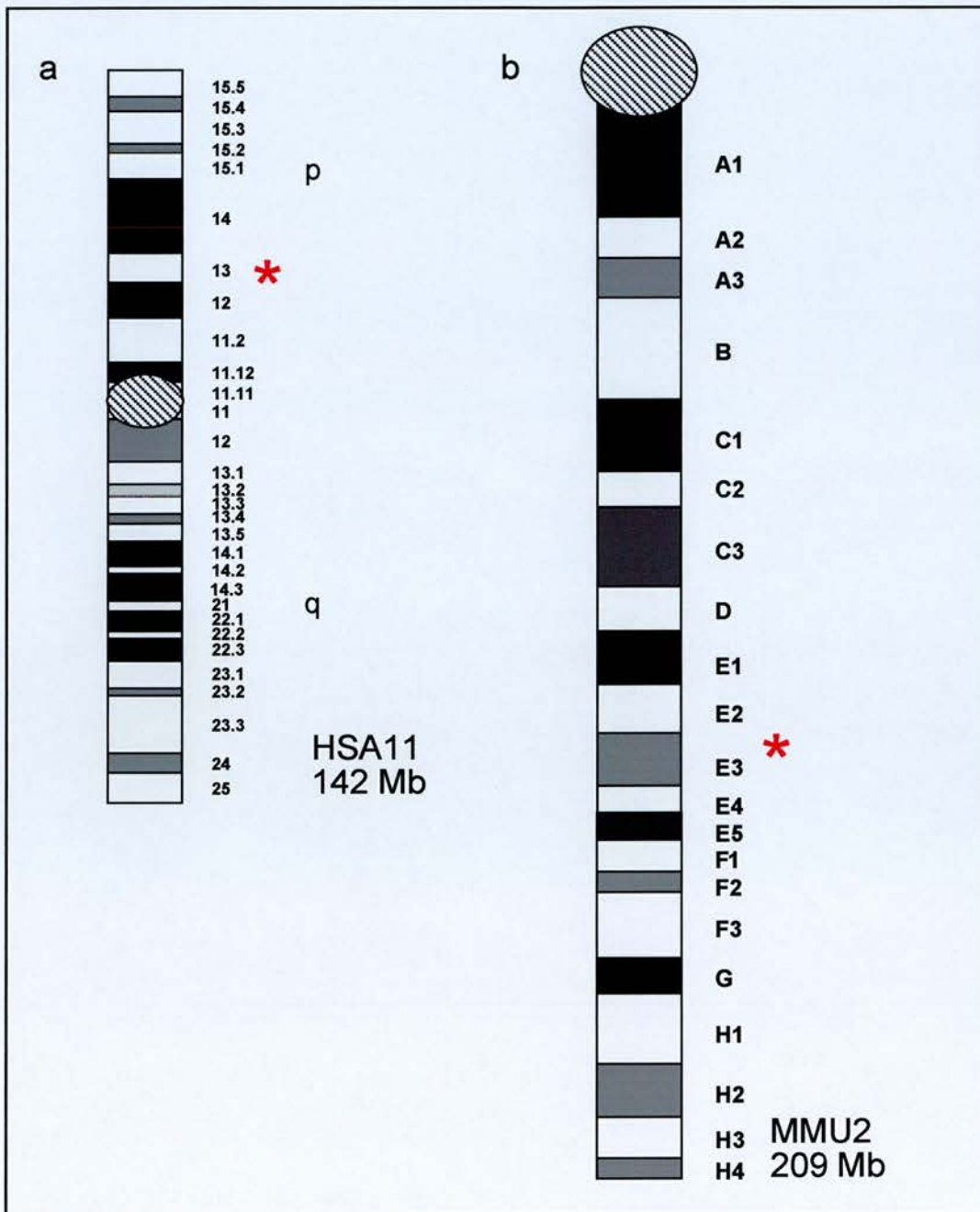


Figure 3.1 G-banded ideograms of human chromosome 11 (HSA11) and mouse chromosome 2 (MMU2).

Standard G-banded ideograms of (a) HSA11 (taken from ISCN 1995) and (b) MMU2 (taken from Sawyer *et al.*, 1987) are illustrated. G-bands are revealed by Giemsa staining and these are generally late replicating, AT-rich, low in gene density and appear to have a 'closed' chromatin structure. R-bands are generally early replicating, GC-rich, high in gene density and appear to have a more 'open' chromatin structure. Diagrams are to-scale and the band position of the WAGR locus on each chromosome is indicated by a red asterisk.

3.2 THE WAGR LOCUS

3.2.1 The human WAGR locus

Distal 11p13 was well mapped and sequenced before the effort to sequence the entire genome because of its involvement in WAGR (Wilm's tumour, aniridia, genitourinary anomalies, mental retardation) syndrome (Compton *et al.*, 1990; Rose *et al.*, 1990; Fantes *et al.*, 1995). The WAGR locus spans approximately 1Mb (Fantes *et al.*, 1995; Gawin *et al.*, 1999) and contains at least 4 genes: Wilms' tumour 1 (*WT1*), reticulocalbin (*RCN*), *PAX6* and pax6-neighbour gene (*PAXNEB*) (Miles *et al.*, 1998; Kleinjan *et al.*, 2001). The expression of two of these genes (*WT1* and *PAX6*) is tissue specific (for *WT1* review, see Little *et al.*, 1999), for *PAX6* review, see Strachan and Read, 1994; Nishina *et al.*, 1999), and two are ubiquitously expressed (*RCN* and *PAXNEB*) (Kent *et al.*, 1997; Kleinjan *et al.*, 2001). There are also large (~300kb) intergenic stretches of DNA from which no transcripts have been identified (Kent *et al.*, 1997; Gawin *et al.*, 1999) (Figure 3.2a). Additional evidence for absence of genes in these regions is suggested by their compaction in the *Fugu rubripes* genome (Figure 3.2c) (Miles *et al.*, 1998).

3.2.2 The murine WAGR locus

Sequencing of MMU2 is in progress, and from data available at the time of writing, gene order, spacing and exon structure of the murine WAGR locus in cytogenetic band E2/E3 of MMU2 appear to be closely conserved with the human locus (P. Gautier, personal communication) (Figure 3.2b). Part of this region is mutated in small eye alleles of the mouse, and small eye mice have a phenotype similar to that of genetic diseases associated with mutations of human *PAX6* (Kent *et al.*, 1997). Synteny is confirmed by the compression of the murine gene spacing in the corresponding region from *F. rubripes* (Miles *et al.*, 1998 and P. Gautier, personal communication), including conservation of the unusually large terminal intron between exons 9 and 10 of the *PAXNEB* gene (Kleinjan *et al.*, 2001a) (Figure 3.2c).

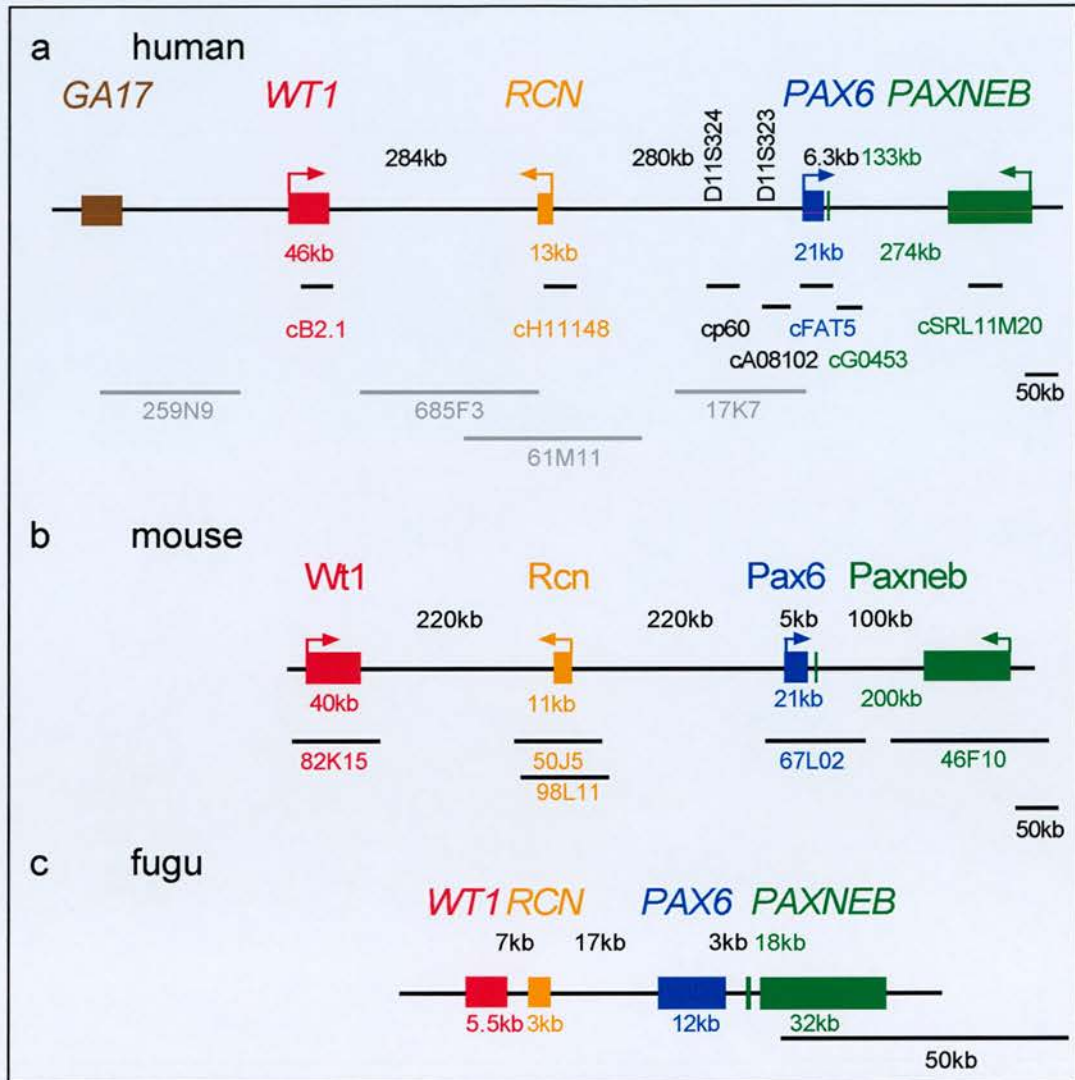


Figure 3.2 Comparison of the human, murine and *Fugu rubripes* WAGR loci.

a) A map of the human WAGR locus at HSA11p13 is annotated with the positions of cosmids and PACs mapped to genes and intergenic sequences used in analysis of the organisation of this region (Fantes *et al.*, 1995; Gawin *et al.*, 1999). Cosmids corresponding to genes are labelled in colour, cosmids representing intergenic, non-transcribed sequences are labelled in black and PACs are indicated in grey.

b) BACs encompassing genes of the murine WAGR locus are coloured according to the corresponding genes on a scale representation of the locus. BACs were selected from a 129 library (Genome Systems Inc.) (D.A. Kleinjan). Gene sizes and intergenic distances are estimated as sequencing is in progress (P. Gaultier, personal communication).

c) The corresponding *Fugu* sequence is illustrated. Note that the genes are in the same order but more compact. The *Fugu* locus is described in detail elsewhere (Miles *et al.*, 1999).

The large intergenic sequences are conserved across all three species, as is the unusually large terminal intron of the *PAXNEB* gene (D.A. Kleinjan, personal communication). Sizes of genes are indicated below in corresponding colours. The length of intergenic sequences are indicated above the line (black), including the large terminal exon of the *PAXNEB* (green) on each map (a-c).

3.3 ANALYSIS TOOLS

3.3.1 Cell line selection and expression analysis

Initial analysis was carried out in a human male lymphoblastoid cell line, FATO. FATO cells have a stable karyotype (46XY) and their nuclei are spherical in shape. High-quality images have previously been produced in FISH experiments using nuclei from this cell line, essential for assessing the validity of the analysis technique I proposed to develop. I confirmed that this cell line expressed only the ubiquitously transcribed genes of the WAGR locus, *RCN* and *PAXNEB*, by RT-PCR (Figure 3.3a). A primary fibroblast cell line, 1HD (Bridger *et al.*, 1998), also expresses the same combination of genes from the locus and was used in subsequent analyses (Figure 3.3a).

Two further human cell lines were chosen which express one of the tissue-restricted genes from WAGR, in addition to *RCN* and *PAXNEB*. A lens epithelium cell line, CD5A (gift of A. Prescott, University of Dundee), known to express the *PAX6* gene (A. Seawright, personal communication), and an ovarian carcinoma (granulosa) cell line (COV434) expressing *WT1* (Berg-Bakker *et al.*, 1993), were chosen. Expression status of all 4 genes of the WAGR locus was confirmed by RT-PCR in both cell lines (FIGURE 3.3a). Expression of *WT1* and *PAX6* in COV434 and CD5a cell lines respectively was assessed to be in 100% of the cell population. For *WT1*, this was demonstrated by immunofluorescence using the C19 antibody, raised against the C-terminal 19 amino acids of the protein (Section 2.7) (Figure 3.3b). For *PAX6*, immunohistochemistry was performed using antibodies raised against the protein (A. Seawright) (Figure 3.3c).

To examine organisation of the murine locus, I made a mouse embryonic fibroblast (MEF) cell line (Section 2.1.7). However, I found that 2D analysis of this cell-line could not easily be completed due to the small size of the nuclei. An ES cell line, E14 was used for murine analysis in 2D (Hooper *et al.*, 1987).

3.3.2 DNA probe selection

In order to study organisation of the WAGR locus in mouse and human nuclei, cosmid, BAC, and PAC probes mapped to genes and intergenic sequences spanning more than 1 Mb, were used in combination with complex chromosome-specific probes in FISH analysis.

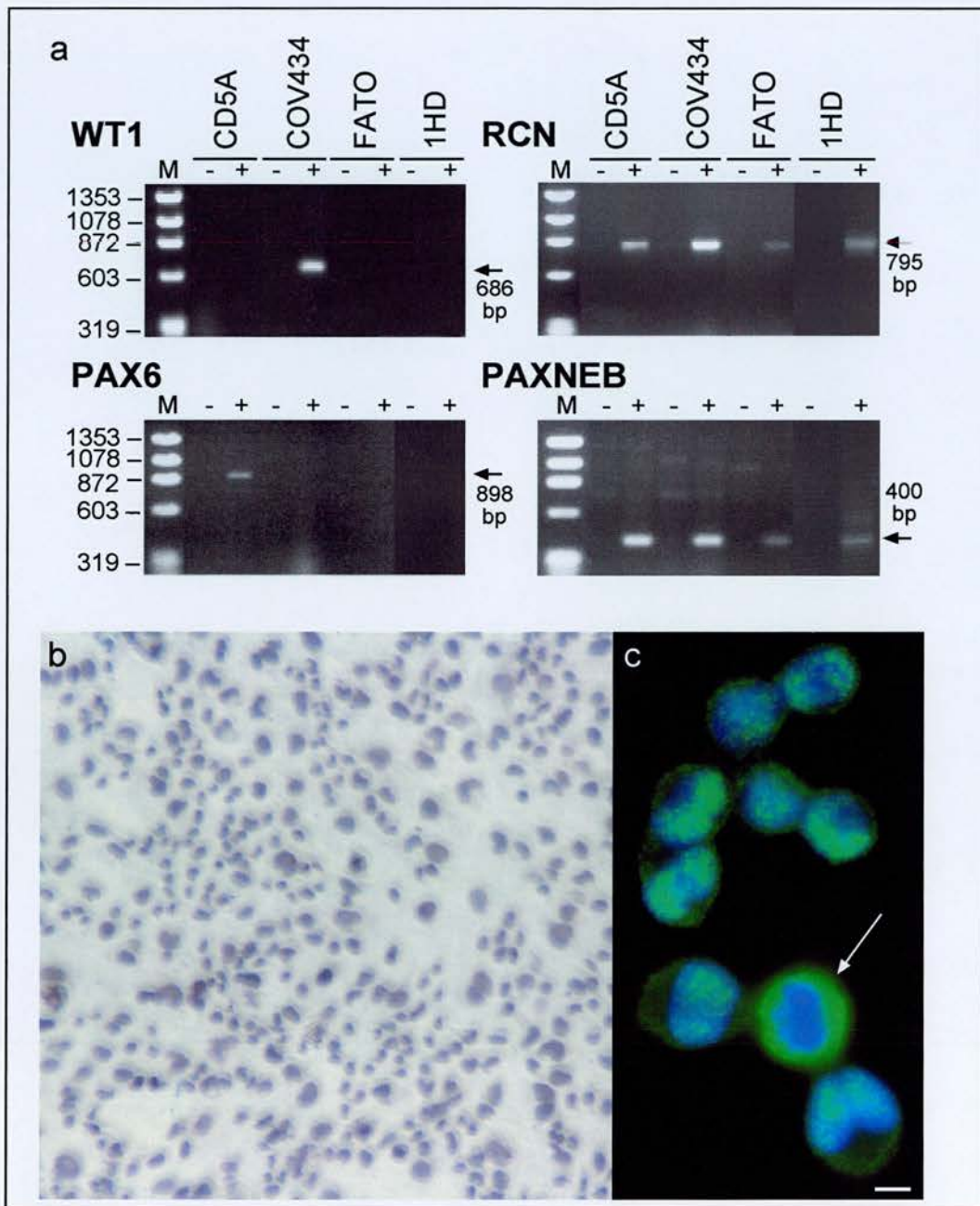


Figure 3.3 Expression profiles of WAGR genes in human cell lines.

a) RT-PCR analysis of the expression status of the four genes from WAGR in four cell lines. RT-PCR reactions were performed on total RNA samples prepared from each cell line (+RT) and a control RNA sample to which the reverse transcriptase enzyme had not been added to the reaction (-RT), was used as a control. The names of the cell lines are indicated above the corresponding gel lanes. *RCN* and *PAXNEB* are ubiquitously expressed genes and were transcribed in all the cell lines (795 and 400bp bands respectively). *WT1* and *PAX6* are tissue restricted genes and, of the cell lines analysed here, were only expressed in COV434 (ovarian carcinoma) and CD5a (lens epithelium) cell lines respectively, as indicated by the production of a 686bp band (*WT1*) and a 898bp band (*PAX6*). Note the absence of bands in -RT control reactions for all genes analysed.

b) The proportion of cells expressing *PAX6* in a population was assessed by immunohistochemistry, using antibodies raised against the protein and 100% of cells were shown to express *PAX6* protein (A. Seawright).

c) Similarly, *WT1* expression was assessed by immunofluorescence using an antibody against the C-terminal 19 amino acids (C19) (green). Nuclei were counterstained using DAPI. Note the exclusion of *WT1* protein from mitotic chromosomes (arrow).

Chromosome-specific or chromosome arm-specific paints were produced by PCR amplification of catch-linkered microdissected chromosome fragments (Section 2.4.1). These paints contained a complex mixture of DNA sequences, 200-500bp in length, complimentary to sequence along the length of a single chromosome or chromosome arm. Genome-wide repeat elements present in the complex probe were suppressed by adding an excess of human or mouse Cot1 DNA to the probe mixture prior to denaturation. Thus, when used in combination, in a single hybridisation reaction, these sequences 'paint' only the chromosome or chromosome arms of interest (Figure 3.4).

For the human WAGR locus, cosmid probes that have been mapped to each of the four genes of the locus and also to one of the large intergenic regions (Fantès *et al.*, 1995), were chosen (Figure 3.2a). PAC probes spanning the same region were also obtained (Niederfuhr *et al.*, 1998; Gawin *et al.*, 1999). Similarly, for the murine locus, BAC clones specific to each of the four genes of the locus were selected from a 129 library (Genome Systems, Inc) (D. A. Kleinjan) (Figure 3.2b).

Chromosome paints and probes specific to sequences within the WAGR locus were used together in FISH on MAA fixed material (Section 2.5). The DNA probes were labelled with biotin or digoxigenin (dig). Hybridisation was detected using fluorescently labelled avidin (a molecule with high affinity to biotin) or fluorescently labelled anti-dig antibody (Section 2.5.4). I experimented with labelling probes with different combinations of biotin and dig-conjugated dUTP, then detecting both of these nucleotide analogues with either Texas Red- (TR) or Fluorescein isothiocyanate- (FITC) coupled antibodies, to see which combination produced the best images for analysis. I found that biotin-conjugated nucleotide analogues were more efficiently incorporated into probe fragments in labelling reactions (Section 2.4). As the complexity of chromosome paint probes is much higher than that of cosmid or BAC probes, I decided to use biotin-labelled chromosome paints and dig-labelled cosmid, BAC or PAC probes for all subsequent FISH experiments. Additionally, FITC is a much brighter fluorochrome than TR, so biotin-labelled chromosome paints were detected with FITC-conjugated antibodies and dig-labelled small probes with TR-conjugated antibody.

3.3.3 Image capture

After FISH, the DNA of all 2D MAA-fixed nuclei was counterstained using DAPI and examined as described above (Section 2.8). Microscope slides were scanned in a methodical manner, beginning at the top left hand corner, scanning to the right, then moving downwards and scanning the next row of nuclei from right to left, etc. Fifty images were collected of consecutive nuclei that fulfilled criteria for images to be suitable for analysis. Nuclei had to

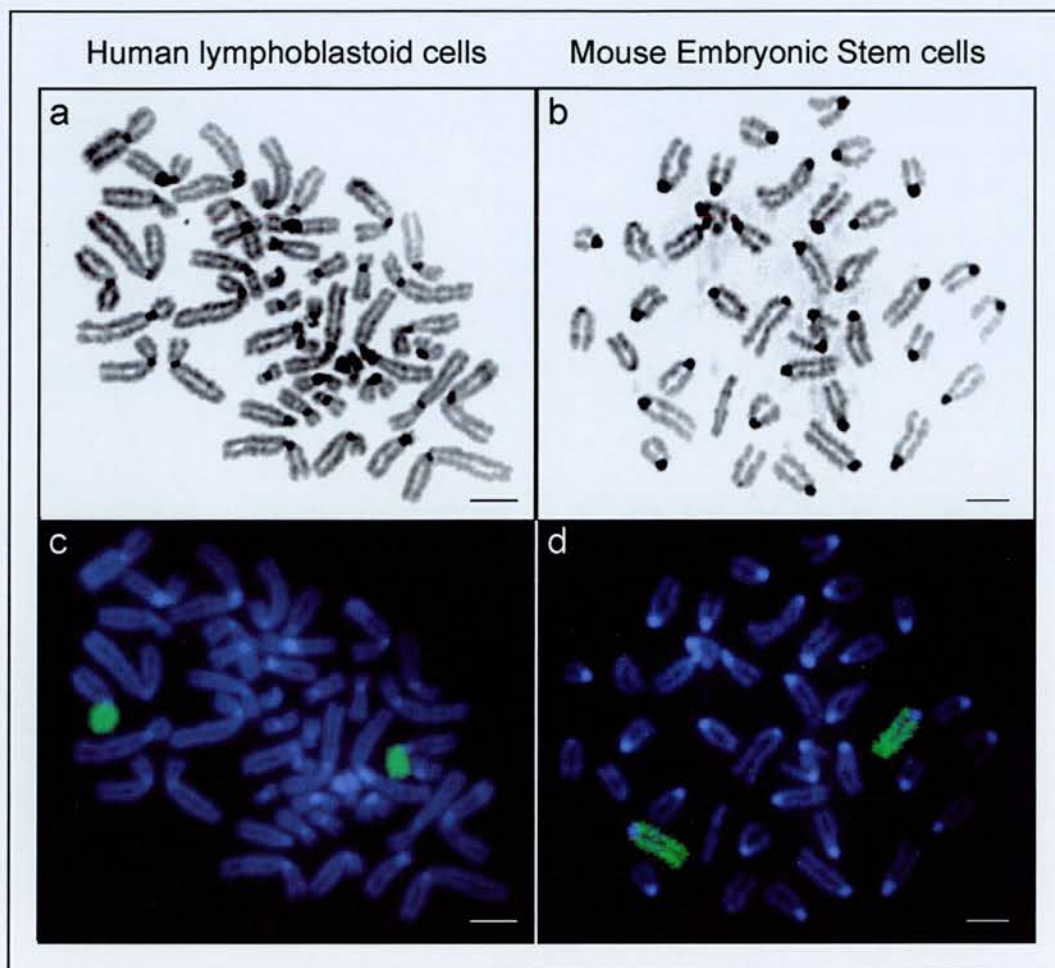


Figure 3.4 Assessing chromosome paints by FISH to metaphase chromosomes.

Grey-scale images of DAPI-stained complete human lymphoblastoid cell (a) and mouse embryonic stem cell (b) metaphase chromosome spreads are shown. Chromosome paints were PCR amplified from microdissected chromosome arms, biotin labelled and detected using avidin-FITC (green). The same metaphase spreads are shown hybridised with a HSA11p chromosome arm paint (c) and a MMU2 chromosome paint (d). Note the absence of hybridisation to centromeres in both (c) and (d). Scale bars represent 5 μ m.

be intact and contain two chromosome domains in which locus signals were visible; in addition, analysis scripts (Section 3.4.1) required images of single nuclei not touching any other. After FISH (Section 3.3.2), each nucleus contained two painted chromosomes (visible in green) as the FATO lymphoblastoid cell line has a stable diploid karyotype. The cosmid, BAC or PAC probe signals were discernible as red or orange spots within the chromosome territory, depending on whether the signal was on the surface of the green chromosome domain (red) or buried within (orange) (Figure 3.5).

3.4 ORGANISATION WITHIN AN INTERPHASE CHROMOSOME TERRITORY

3.4.1 Devising 2D analysis scripts

Three previous studies have considered the position of genes within interphase chromosomes. Kurz and co-workers dissected the position of three isolated genes and one non-transcribed sequence within their respective chromosome territories, in three dimensions (Kurz *et al.*, 1996). Gene localisation was determined in a qualitative manner by assessing whether a gene was located either in the interior or the exterior of a chromosome territory. The chromosome exterior was considered to be a region at the chromosome periphery with a diameter twice that of the FISH signal representing the gene or non-transcribed sequence. Half of this 'exterior' region was within the chromosome boundary, as defined by a chromosome paint. The other half of the 'exterior' was just outside of the chromosome boundary, therefore including locus signals positioned outside of the visible chromosome domain. The remaining chromosome volume constituted the 'interior' area of the territory. A second study assessed simply whether each probe signal was inside or outside the limits, or at the periphery of, chromosome territories defined by a chromosome paint, and compared the percentages of signals in each category in both 2D and 3D (Volpi *et al.*, 2000). The third study (Dietzel *et al.*, 1999) examined the three-dimensional positions of the X-linked adenine nucleotide translocase genes, ANT2 and ANT3 within both the active and inactive X-chromosome territories. This study was quantitative and measured the distance of the centre of a probe signal to the nearest edge of a chromosome domain in 3D.

I decided that the most informative way to analyse the organisation of a large stretch of DNA within an interphase chromosome, making full use of the technologies available to me, would be to measure the distance of the probe signals from the nearest edge of the chromosome territory. Initial analysis was carried out in two-dimensions so that large

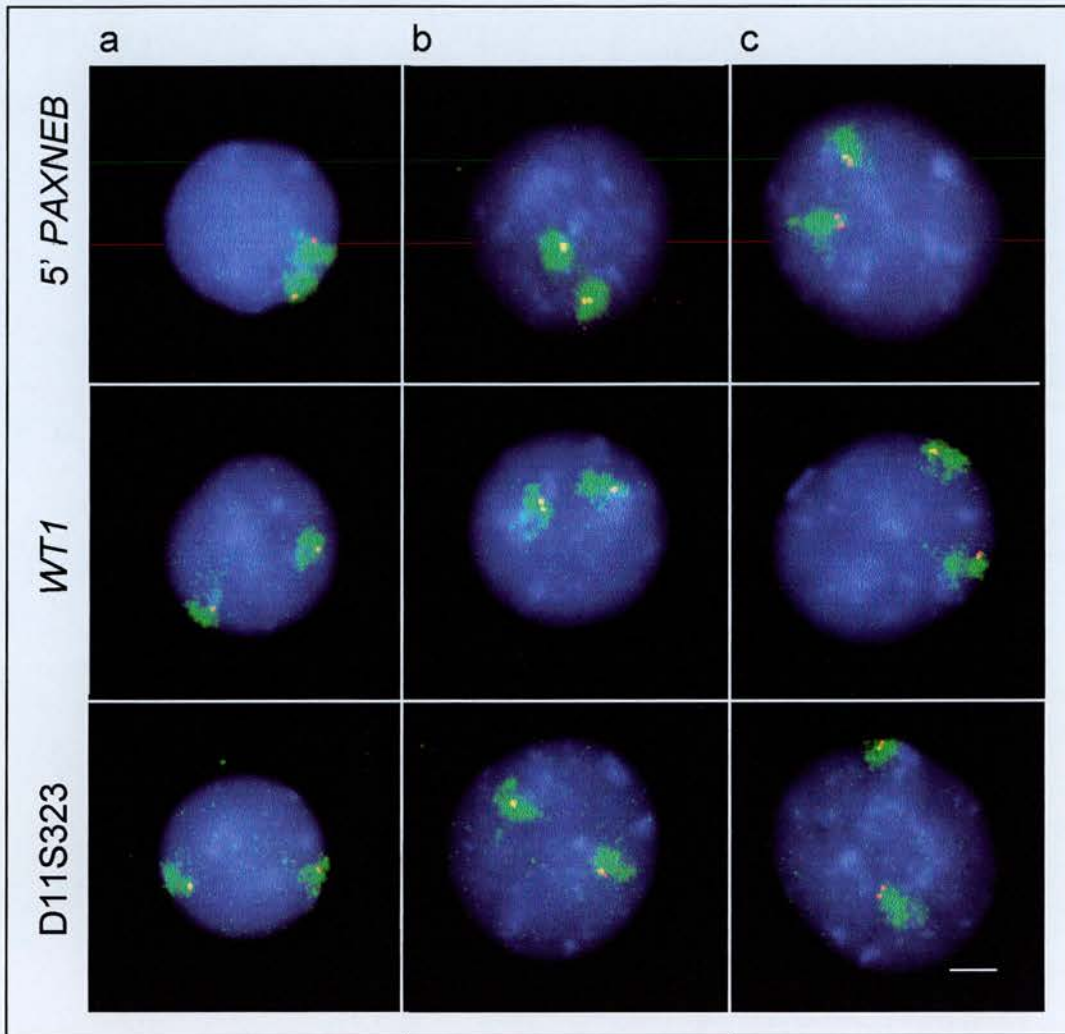


Figure 3.5 Detecting sequences from the human WAGR locus within the territory of chromosome 11p in interphase nuclei.

MAA-fixed lymphoblastoid nuclei were hybridised with a biotin-labelled HSA11p paint (green) and a dig-labelled cosmid (red) specific to the ubiquitously expressed gene *PAXNEB*, the tissue-restricted gene *WT1* or the intergenic sequence D11S324 from the human WAGR locus and counterstained with DAPI (blue). Nuclei in G1- or S-phase of the cell cycle contain 2 chromosome domains with 1 locus signal each (a); nuclei in early S-phase have 1 chromosome domain with 2 signals (locus already replicated) and one containing only 1 signal (yet to be replicated) (b); later S- or G2-phase nuclei had 2 locus signals in each chromosome domain (c). The scale bar represents 5 μ m. Note the increase in nuclear area for each set of nuclei as the cell cycle progresses from G1 to G2 (a-c).

numbers of nuclei could be analysed in a statistical manner with the idea of expanding this analysis to three dimensions later (Chapter 5). Two scripts were developed by Dr Paul Perry, MRC Human Genetics Unit, Edinburgh, using the Digital Scientific software, IPLab Spectrum. The 'cosmid territory' script (Box 3.1) (Figure 3.6), was designed to assess the location of a locus signal positioned within the limits of a chromosome domain, as defined by a chromosome paint. The 'escaping cosmid' script (Box 3.2) (Figure 3.7), was appropriate for analysis of locus signals apparently outside of the area of fluorescence defining the chromosome territory. The steps involved in these scripts are briefly described below (Boxes 3.1 and 3.2) and depicted in Figures 3.6 and 3.7.

Box 3.1 The 'cosmid territory' script: used to determine the position of a locus signal within a chromosome domain in 2D.

Cosmid territory script (Figure 3.6)

- i) The DAPI stained nucleus was automatically segmented from background and the area of this segment calculated (Figure 3.6b). Background was also removed from FITC and TR images by calculating the mean FITC and TR pixel intensities within the area of the DAPI segmented nucleus and subtracting them from the FITC and TR images.
- ii) Locus-specific TR fluorescent signal was segmented and a region of interest manually defined around one of the signals (each locus signal was treated independently, each territory containing either one or two depending on whether the locus had been replicated). The centroid co-ordinates of the locus signal were calculated (Figure 3.6c).
- iii) FITC fluorescence from the chromosome territory was then segmented (Figure 3.6d) and a region of interest manually selected to include all of the visible territory (Figure 3.6e).
- iv) The grey value threshold for the chromosome signal was manually determined to a level which best represented the chromosome domain in the original image and the area of the chromosome territory calculated (Figure 3.6f and g).
- v) A segmentation disc was dilated out from the locus signal centroid, then eroded until a pixel with zero signal intensity from the territory was found: this represented the nearest edge of the chromosome to the locus. The distance from the centre of the locus to the nearest territory edge was thus calculated as the radius of this disc (Figure 3.6h).
- vi) Another segmentation disc was then dilated out from the locus signal centroid and eroded until a pixel with zero intensity from the DAPI signal was found. The radius of this disc defined the distance from the centre of the locus signal to the nearest edge of the nucleus (Figure 3.6i).
- vii) Further measurements for the distances of the nucleus centroid to both the centre of the locus signal and the centroid of the chromosome signal were also calculated (Figure 3.6j). These data were displayed in an output table (Figure 3.6k), which was exported to Microsoft Excel for manipulation.

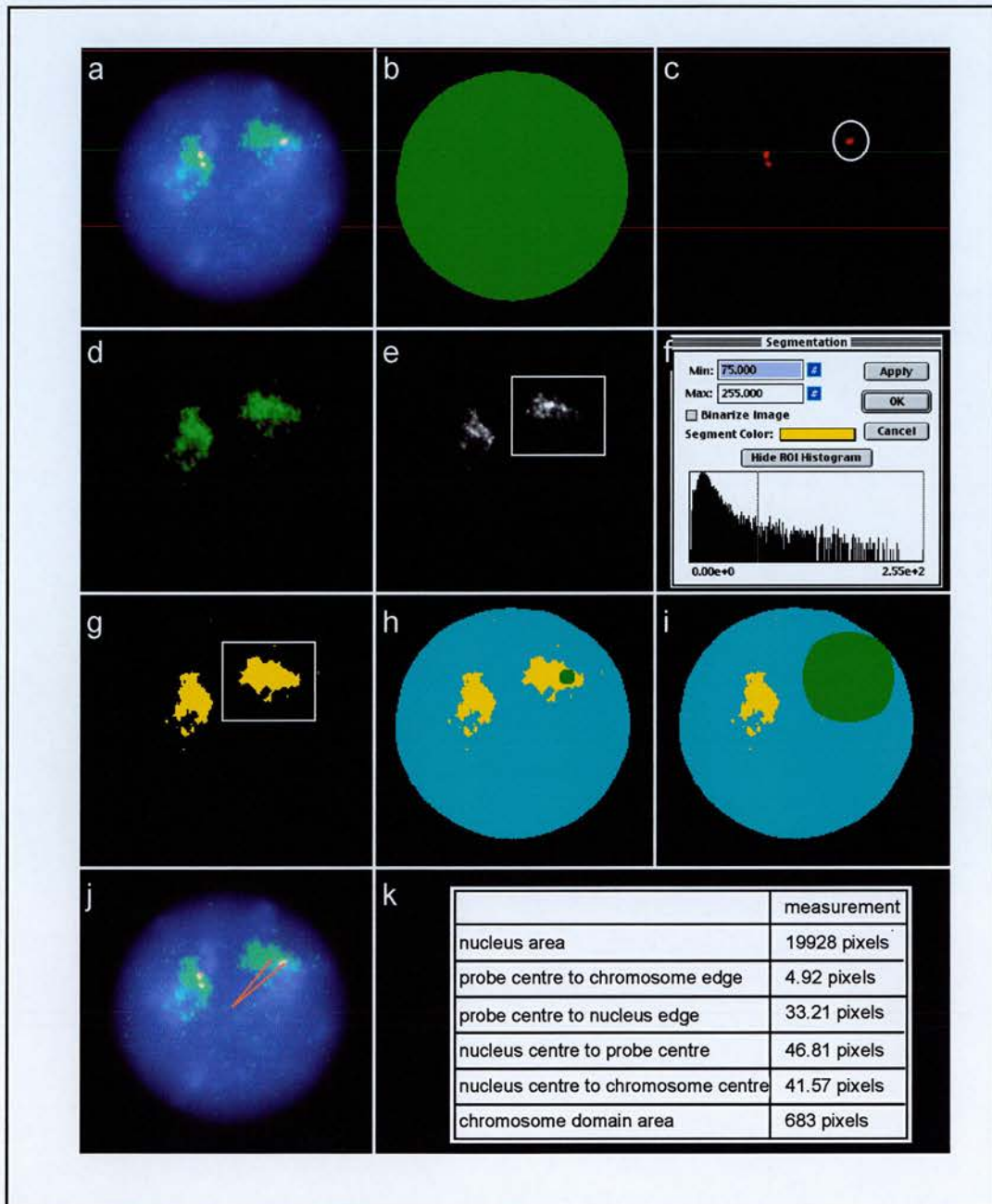


Figure 3.6 Analysing an image using the 'cosmid territory' script.

This script was used for analysis of two dimensional images (produced by FISH on MAA-fixed nuclei) where the locus signal was within the fluorescence domain representing the chromosome territory. The nucleus illustrated (a) depicts a biotin-labelled HSA11p chromosome arm paint (green) and a digoxigenin-labelled cosmid mapped to the *PAX6* gene (red) in a lymphoblastoid cell nucleus. See Box 3.1 for details of each step.

Box 3.2 The 'escaping cosmid' script used to determine the distance of a locus signal outside of a chromosome domain in 2D.

Escaping cosmid script (Figure 3.7)

i-iv) as described above for the cosmid territory script (Box 1) (Figure 3.7b-g).

v) A segmentation disc was dilated out from the locus signal centroid then eroded until a pixel containing territory signal was found. This identified the nearest edge of the territory to the locus and the distance from the centre of the locus to the nearest chromosome edge was calculated as the radius of this disc (Figure 3.7h). The value was assigned a negative sign, to indicate that the locus appeared outside of the visible limits of the chromosome territory.

Vi and vii) As described in Box 3.1 above (Figure 3.7i).

If neither of the above scripts produced a measurement, the locus probe signal was taken to be absolutely at the periphery of the chromosome domain. The distance of the centre of the probe signal to the nearest edge of the chromosome was therefore zero.

Each chromosome domain contained either one or two locus signals (Figure 3.5); before chromatid replication locus FISH signals appear as single dots within a chromosome territory (G1 or early S-phase nuclei) and after replication, in late S-phase or G2 nuclei, double dots are produced. This was used to distinguish G1/early S-phase nuclei (both chromosome domains contain only a single probe signal) from those in late S-phase or G2 (either one or both chromosome domains contain a signal doublet) (Bickmore and Carothers, 1995). On export to Microsoft Excel, each data entry was annotated to indicate whether the nucleus was approximately in G1 or G2 phase of the cell cycle.

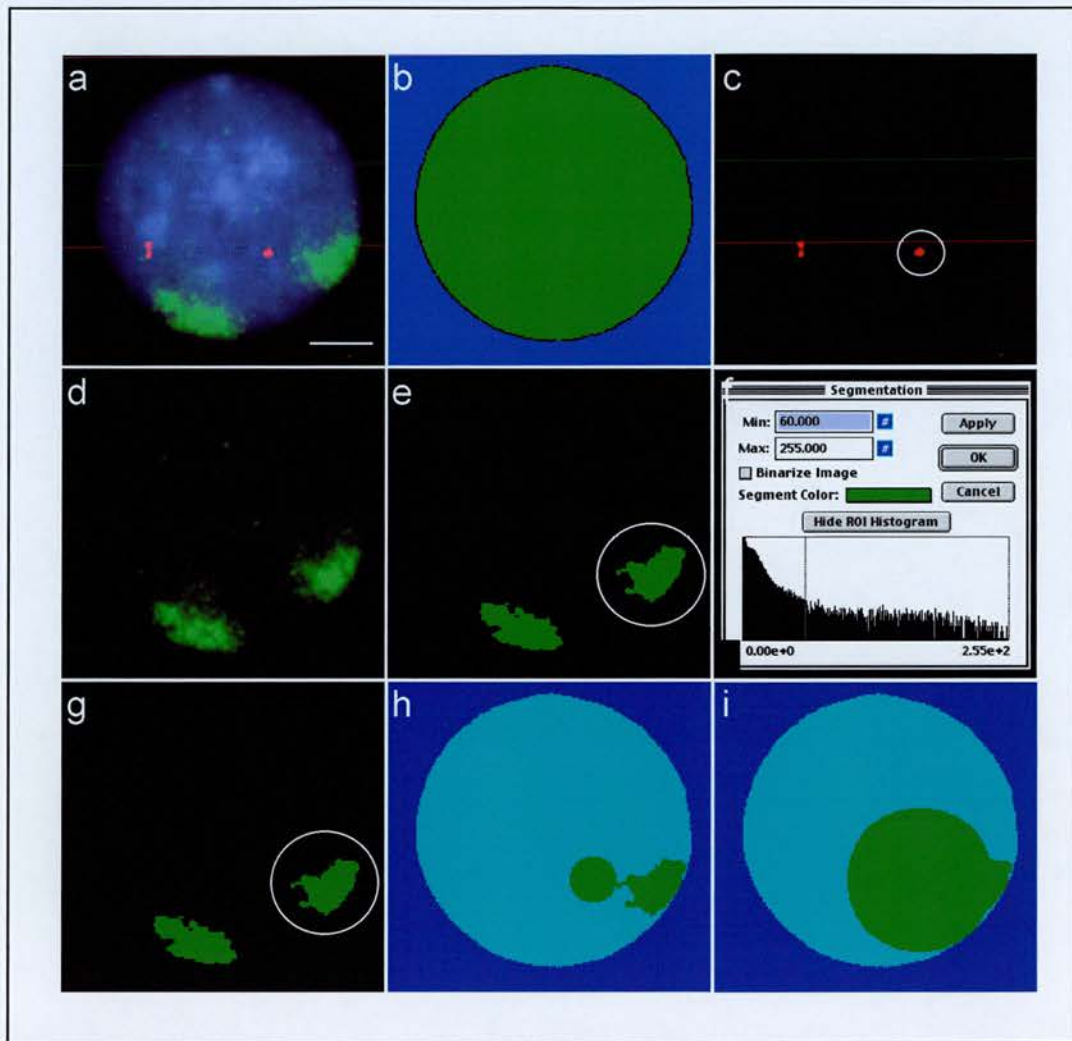


Figure 3.7 Analysing an image using the 'escaping cosmid' script.

This script was used for analysis of two dimensional images (produced by FISH on MAA-fixed nuclei) where the locus signal appeared outside of the chromosome fluorescence domain. The nucleus illustrated (a) depicts a biotin-labelled HSA11p chromosome arm (green) and a digoxigenin-labelled cosmid to the *INS* gene (red) in a lymphoblastoid cell nucleus. At the end of the script data are displayed in an output table (as depicted in Figure 3.6k) and are directly exported to Microsoft Excel for manipulation. See Box 3.2 for details.

3.4.2 Using the raw data output from script analysis

Mean nuclear areas for the data sets representing G1 and G2 nuclei were calculated. In the lymphoblastoid cell line, FATO, the mean nuclear area of late S/G2 nuclei (measured in pixels) was 1.18 times the mean area of G1 nuclei (n=815). Subsequent analysis of primary fibroblast cells produced a similar figure: G2 nuclei were 1.27 times the size of G1 nuclei (n=823). These values for increase in nuclear area are compatible with a doubling in volume (1.18^3 is 1.64, 1.27^3 is 2.06). This increase in nuclear size as nuclei move through the cell cycle from G1, through S-phase when DNA is replicated to G2 is clearly illustrated (Figure 3.5). I decided it was therefore necessary to account for variation in nuclear size within an unsynchronised cell population (and hence chromosome domain area) in analysis (Boxes 3.1 and 3.2). Measuring the absolute distance between two points within a nucleus is not particularly meaningful when nuclear area differs. Normalisation would also account for variation in chromosome territory size that may occur between cell types and enable direct comparison of data between cell lines.

A value for the hypothetical chromosome radius was calculated from the territory area (determined by the scripts described in Boxes 3.1 and 3.2), by assuming the territory to be circular (Section 2.10). A value of 0.0 thus described a sequence at the edge of a chromosome territory (a value assigned when neither script produces a value for the distance to the nearest chromosome edge), and 1.0 denoted a locus positioned at the centre of a chromosome. Data presented in a previous analysis that similarly measured distances of genes from the nearest chromosome edge, did not account for variation in chromosome size between nuclei at different stages of the cell cycle (Dietzel *et al.*, 1999). However, Dietzel and co-workers claim that unpublished data where normalisation was included supported the raw data published.

3.5 SEQUENCES ACROSS 1MB OF 11p13 ARE FOUND IN A SIMILAR POSITION WITHIN HSA11p

3.5.1 The organisation of WAGR within a transformed cell line: lymphoblastoid cells

The positions of seven cosmids spanning the human WAGR locus (Figure 3.2) within the HSA11p chromosome territory were analysed by FISH to 2D preparations of lymphoblastoid cell nuclei. Fifty images were analysed per cosmid (100 chromosome territories) and a value describing the mean intra-chromosomal position, \pm standard error of the mean (SEM), was calculated (Section 2.10) for each cosmid. This described the position of each locus within the HSA11p chromosome domain in lymphoblastoid cells along a calculated chromosome territory radius. These data are recorded in Table 3.1, along with the corresponding absolute distance of each probe to the chromosome edge in microns (calculated from the mean chromosome radius, see Table 7.2) and are displayed graphically in Figure 3.8.

Table 3.1 Comparing the position of sequences from WAGR relative to the chromosome periphery in different cell lines

For each probe, fifty images were analysed using the scripts described in Boxes 3.1 and 3.2. Mean values were calculated for the normalised position of probes relative to the nearest edge of the chromosome domain and the corresponding position distance in microns calculated, both expressed \pm SEM. (Section 2.10). n/d indicates a result not determined. Measurements corresponding to expressed genes are underlined.

Gene/locus probe (chromosomal location)	Lymphoblastoid (FATO)		Primary fibroblast (1HD)		Lens epithelium (CD5a)		Ovarian carcinoma (COV434)	
	Mean probe location as proportion of chromosome radius	Mean distance of probe from edge of chromosome (μ m)	Mean probe location as proportion of chromosome radius	Mean distance of probe from edge of chromosome (μ m)	Mean probe location as proportion of chromosome radius	Mean distance of probe from edge of chromosome (μ m)	Mean probe location as proportion of chromosome radius	Mean distance of probe from edge of chromosome (μ m)
WT1	0.274 ± 0.033	0.511 ± 0.062	0.335 ± 0.019	0.608 ± 0.034	0.258 ± 0.028	0.406 ± 0.044	0.331 ± 0.027	0.586 ± 0.048
RCN	<u>0.306 ± 0.025</u>	<u>0.571 ± 0.047</u>	<u>0.336 ± 0.018</u>	<u>0.610 ± 0.033</u>	<u>0.280 ± 0.024</u>	<u>0.441 ± 0.038</u>	<u>0.301 ± 0.022</u>	<u>0.533 ± 0.039</u>
D11S324	0.313 ± 0.025	0.584 ± 0.047	0.240 ± 0.030	0.436 ± 0.054	n/d	n/d	0.315 ± 0.021	0.558 ± 0.037
D11S323	0.310 ± 0.027	0.579 ± 0.050	0.317 ± 0.019	0.576 ± 0.034	0.311 ± 0.025	0.490 ± 0.039	0.327 ± 0.027	0.579 ± 0.048
PAX6	0.306 ± 0.028	0.571 ± 0.052	0.291 ± 0.023	0.528 ± 0.042	<u>0.308 ± 0.022</u>	<u>0.485 ± 0.035</u>	0.323 ± 0.029	0.572 ± 0.051
PAXNEB 3'	<u>0.354 ± 0.018</u>	<u>0.660 ± 0.033</u>	<u>0.289 ± 0.024</u>	<u>0.525 ± 0.044</u>	<u>0.291 ± 0.023</u>	<u>0.458 ± 0.036</u>	<u>0.271 ± 0.026</u>	<u>0.450 ± 0.046</u>
PAXNEB 5'	<u>0.344 ± 0.024</u>	<u>0.642 ± 0.045</u>	<u>0.343 ± 0.020</u>	<u>0.623 ± 0.036</u>	<u>0.260 ± 0.027</u>	<u>0.409 ± 0.043</u>	<u>0.335 ± 0.018</u>	<u>0.593 ± 0.032</u>

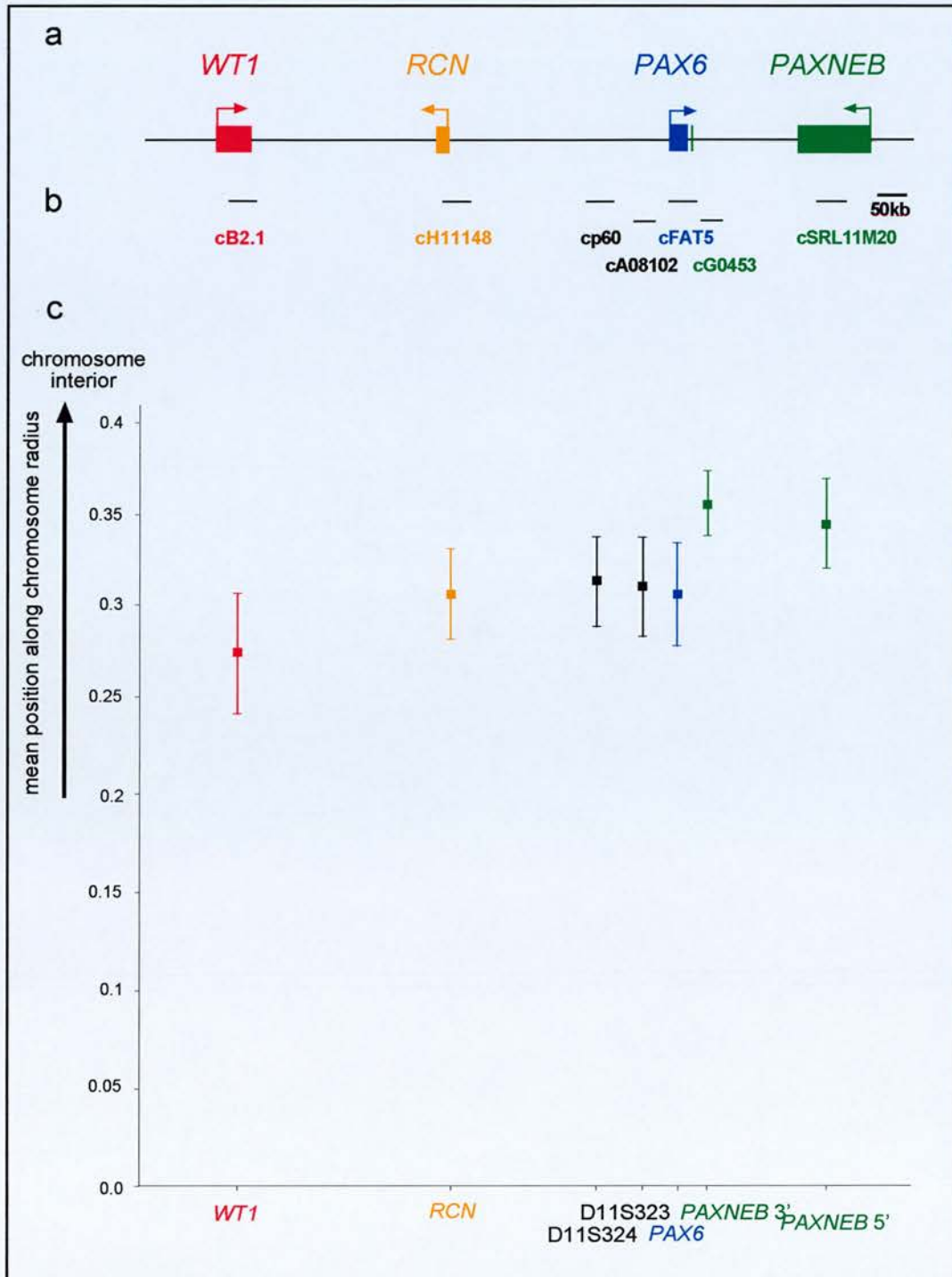


Figure 3.8 The organisation of the WAGR locus within HSA11p in nuclei of lymphoblastoid cells.

The organisation of the human WAGR locus is illustrated (a) and the relative positions of the cosmids used in analysis of the organisation of this region within HSA11p are indicated (b). c) The mean normalised position of each probe along the radius of the interphase 11p domain is depicted graphically. Error bars indicate the standard error of the mean (SEM). A value of 0 would represent the edge of the chromosome and 1.0 the chromosome centre. Labelled across the x axis are the names of the genes or intergenic markers to which the plots correspond.

Using analysis of variance (Section 2.10), no significant difference was found between the position of any of the probes across the megabase locus within the 11p chromosome domain of lymphoblastoid nuclei ($P=0.682$) (A. Carothers). In addition, there was no variation between nuclei determined to be in G1 or G2 of the cell cycle (Section 2.10), indicating that there is no significant cell-cycle related change in the position of the WAGR locus within the HSA11p territory. A previous study suggested that all gene sequences were located at the surface of their respective chromosome domains regardless of transcriptional status (Kurz *et al.*, 1996). Similarly, I found no difference in the position of actively transcribed genes (*RCN* and *PAXNEB*) as compared to non-transcribed genes (*WT1* and *PAX6*). However, I found no preference for gene sequences to be located more towards the periphery of the chromosome as compared to intergenic, non-transcribed sequences (D11S323 and D11S324) within the megabase of DNA studied, which was suggested by Kurz and co-workers. I also found that the ubiquitously expressed genes *RCN* and *PAXNEB* (confirmed to be expressed in the cell line studied here, Figure 3.3) were not at the chromosome surface. As 2D data can be used to predict 3D spatial organisation (Carothers, 2000), this indicates that genes can be transcribed within a chromosome territory.

3.5.2 The organisation of the WAGR locus is identical in a primary cell line

Transformed cell lines can be maintained in culture for hundreds of generations, thus providing continuity for studies that may span many years. However, there is evidence to suggest that transformed cell lines may undergo some remodelling of nuclear architecture (A. Fisher, personal communication). The exact nature of these changes in nuclear structure are uncharacterised, but it is an important consideration when performing fine detail organisational analyses such as described here. Therefore, I decided to assess whether my data, obtained from analysis of a transformed cell line, was reproducible in a primary cell line.

A primary dermal fibroblast cell line, 1HD (Bridger *et al.*, 1998), was obtained. Analysis of the 1HD cell line was performed on nuclei from early-passage cells, up to a maximum of passage 13. I confirmed that these fibroblasts, like the FATO lymphoblastoid cell line, express only the ubiquitously expressed genes from WAGR, *RCN* and *PAXNEB* by RT-PCR analysis (Figure 3.3a). The organisation of WAGR was assessed in 2D preparations of 1HD nuclei using the set of seven cosmid probes spanning the locus (Figure 3.2a) used to analyse organisation within lymphoblastoid cell nuclei were (Section 3.5.1). Fifty images were analysed per cosmid (100 chromosome territories) (Figure 3.9), using analysis scripts (Boxes

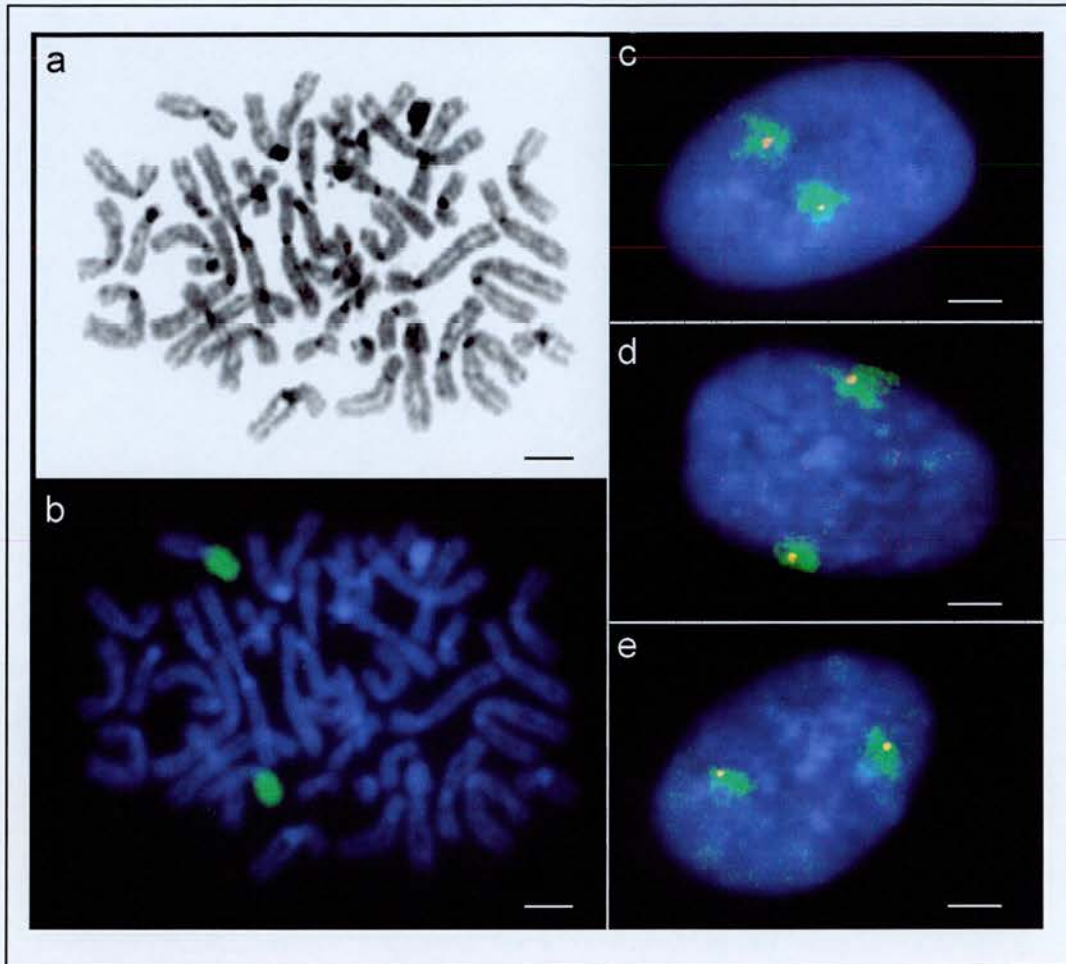


Figure 3.9 Visualising the organisation of WAGR within HSA11p of primary fibroblasts.

a) A grey-scale representation of a DAPI-stained metaphase chromosome spread. b) The specificity of a biotin-labelled HSA11p paint on primary fibroblast chromosomes is shown on the same metaphase spread after FISH and detection using avidin-FITC (green). Chromosomes were counterstained using DAPI (blue). The same paint was used in FISH on fibroblast nuclei (c-e) in combination with dig-labelled probes to the WAGR locus (red), corresponding to ubiquitously expressed genes e.g. *PAXNEB* (c), tissue restricted genes e.g. *WT1* (d) and intergenic regions e.g. cytogenetic marker D11S323 (e). DNA was again counterstained using DAPI (blue). Scale bars represent 5µm.

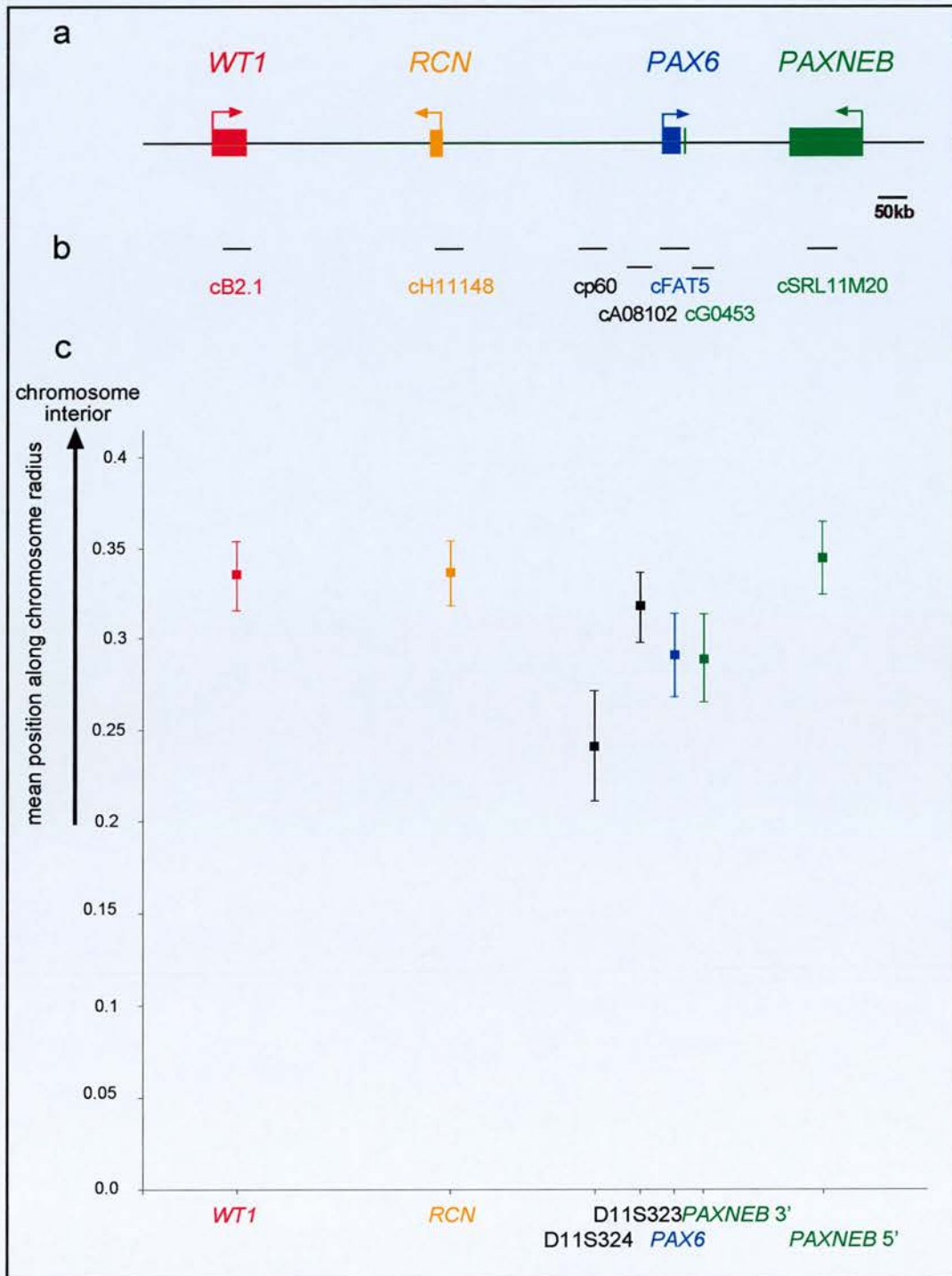


Figure 3.10 The organisation of the WAGR locus within HSA11p in primary fibroblast nuclei.

The organisation of the human WAGR locus is illustrated (a) and the relative positions of the cosmids used in analysis of this region within HSA11p are indicated (b). c) The mean normalised position of each probe along the radius of the interphase 11p domain is graphically depicted. Error bars indicate the standard error of the mean (SEM). A value of 0 represents the edge of the chromosome and 1.0 the chromosome centre. Labeled across the x axis are the names of the genes or intergenic markers to which the plots correspond.

3.1 and 3.2) and a value describing the mean intra-chromosomal position, \pm standard error of the mean (SEM), was calculated (Section 2.10) for each cosmid. Data describing the mean positions of each of the cosmid probes within HSA11p in primary fibroblast nuclei are recorded in Table 3.1 and displayed in Figure 3.10.

From analysis of variance (Section 2.10, A. Carothers), there was no significant difference between the mean positions of probes across the megabase locus within the 11p domain of primary fibroblast nuclei, normalised for a calculated chromosome radius ($P=0.054$). This P value was low, but not significant to a 5% confidence interval; and is mostly due to the fact that the D11S324 probe localised more externally than others from the locus did. Therefore, WAGR adopted a similar intrachromosomal position within a primary cell line as that described within a transformed lymphoblastoid cell line (Section 3.5.1).

3.5.3 The position of WAGR within HSA11p is not altered by increased transcription from the locus

Previous analyses, which assessed the organisation of chromatin with respect to changes in gene expression, have suggested that there may be a connection between transcription status and genome organisation. IFN γ treatment of primary fibroblast cells induces transcription of MHC genes; increased transcription from this locus appears to coincide with an increased incidence of extrusion of chromatin containing these genes from the HSA6 territory (Volpi *et al.*, 2000). In a second study, the position of a transgene relative to a pericentromeric heterochromatin complex was shown to be altered with changing transcriptional activity of the transgene, indicating that transcription factors can initiate changes in higher order chromatin structure during the earliest stages of gene activation (Lundgren *et al.*, 2000).

Therefore, I performed analysis of the organisation of WAGR within HSA11p on 2D MAA-fixed nuclei from cell lines expressing additional genes from WAGR. CD5a, a lens epithelial cell line, expresses Pax6 and COV434, an ovarian carcinoma cell line expresses Wt1 (Section 3.3.1). I showed expression of Pax6 or Wt1 in these cell populations by RT-PCR and, in addition, showed that 100% of cells express the proteins by immunohistochemistry and immunofluorescence respectively (Figure 3.3). Analysis of the organisation of WAGR was performed using the seven cosmids spanning the 1 Mb locus (Figure 3.2a), as described (Section 3.5.1) (Figure 3.11). The mean positions of the probes were calculated and compared to values obtained for the FATO and 1HD cell lines expressing only the ubiquitous genes from the WAGR locus, *RCN* and *PAXNEB* (Sections 3.5.1 and 3.5.2) (Table 3.1 and Figure 3.12).

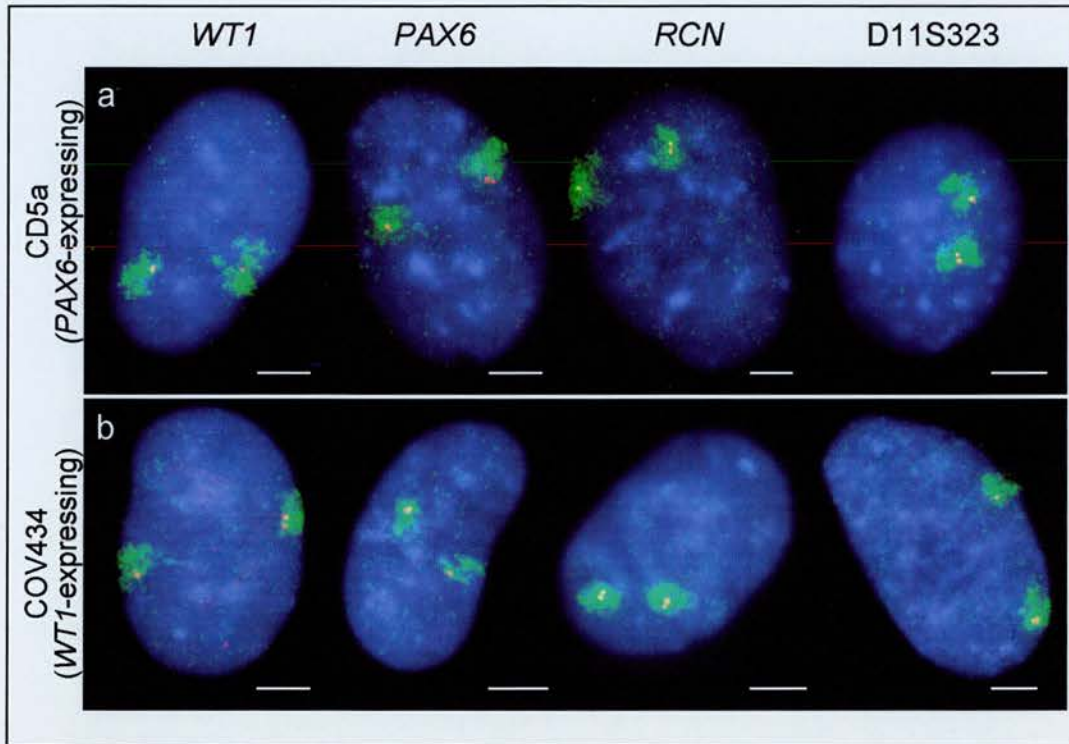


Figure 3.11 Analysis of the WAGR locus in human cell lines expressing WT1 or PAX6.

MAA-fixed nuclei from a lens epithelium cell line (CD5a) (a) which expresses the *PAX6* gene (Figure 3.3) in addition to the ubiquitously expressed WAGR genes *RCN* and *PAXNEB*, and an ovarian carcinoma cell line (COV434) (b) which expresses *WT1*, *RCN* and *PAXNEB* were hybridised with biotin-labelled HSA11p paint (green) and a dig-labelled cosmid (red) specific to the tissue-restricted genes *WT1* or *PAX6*, the ubiquitously expressed gene *RCN*, or the intergenic sequence D11S323 from the WAGR locus. Note that there is no obvious difference in these two dimensional images between the position of the *WT1* or *PAX6* genes in the two cell lines, in which they are both differentially expressed. The scale bars represent 5 μ m.

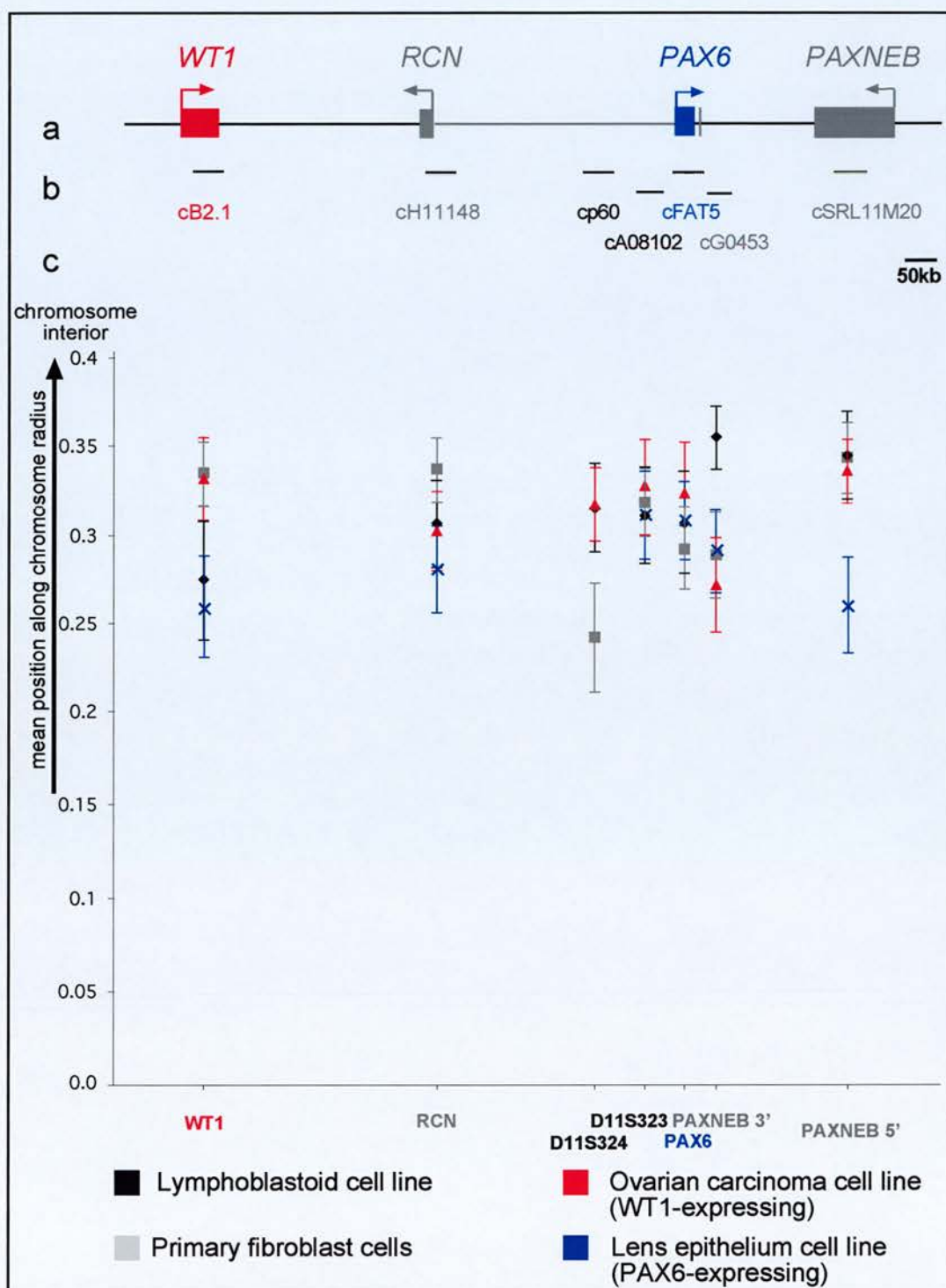


Figure 3.12 Comparison of the spatial organisation of WAGR in cell lines expressing different combinations of genes from the locus.

a) A map of the human WAGR locus with the relative positions of cosmids used in analysis of the organisation of this region within HSA11p indicated (b). Genes differentially expressed in the cell lines considered are coloured.

c) The mean normalised positions of probes spanning WAGR are plotted (\pm SEM). All four cell lines express the *RCN* and *PAXNEB* genes; in addition, the ovarian carcinoma cell line expresses *WT1* (red) and the lens epithelium cell line expresses *PAX6* (blue). A value of 0.0 represents the edge of the chromosome domain and 1.0 the chromosome centre.

By analysis of variance of data for both cell lines (A. Carothers), there was found to be no significant difference between the mean probe positions across the 1 Mb of sequence (COV434 $P=0.458$, CD5a $P=0.302$). The *WT1* gene was not found in a significantly different position compared to the rest of the probes across the megabase locus in the *WT1*-expressing cell line ($P=0.911$). Similarly, in the *PAX6*-expressing cell line, the position of the *PAX6* gene was not significantly different to the other loci ($P=0.957$). In both cell lines, the expressed tissue restricted gene is not noticeably more external than in other cell lines where it is not expressed.

In summary, I found no preference for gene sequences to be located more towards the periphery of the chromosome as compared to intergenic, non-transcribed sequences within the megabase of DNA studied. All loci across the megabase of DNA representing WAGR in 11p13 had a mean position approximately one third of the way along a radius of the interphase 11p chromosome domain in all cell lines examined in two dimensions. The specific position of the gene whose transcription status was different was not significantly altered with respect to the rest of the locus in cell lines where expression status of that gene was altered. These 2D data infer that transcription factors and the transcription machinery must be able to access genes located inside chromosome territories; large-scale 3D chromosome territory structure does not appear to present a barrier to transcription.

3.6 THE SPATIAL ORGANISATION OF THE MURINE WAGR LOCUS WITHIN INTERPHASE MMU2 IS COMPARABLE TO THAT OF THE SYNTENIC HUMAN REGION AT HSA11p

The murine WAGR region of conserved synteny to HSA11p13 is located on mouse chromosome 2 (MMU2). Sequencing is currently in progress, but gene order, spacing and exon structure of the murine WAGR locus in cytogenetic bands MMU2E2-E3 appear to be very closely conserved to the human locus, (P. Gautier, personal communication). An MMU2 chromosome paint was hybridised in combination with a BAC corresponding to each gene from mouse WAGR (D.A. Kleinjan) (illustrated in Figure 3.2), to MAA-fixed mouse embryonic stem (ES) cell nuclei (E14 cell line) (Figure 3.13). Fifty images were captured for each probe and the distance of probe signals from the nearest edge of the chromosome territory were analysed using scripts (Boxes 3.1 and 3.2). Mean distances of each of the four genes from the nearest edge of the chromosome domain were calculated and are recorded in Table 3.2 and displayed in Figure 3.14.

Table 3.2 Mean intra-chromosomal positions of loci spanning murine WAGR within nuclei of mouse embryonic stem cells

Fifty images were captured of mouse ES nuclei after FISH using an MMU2 paint in combination with a probe mapped to genes of the WAGR locus. Scripts (Boxes 3.1 and 3.2) were used to analyse the position of the probe signal in relation to the chromosome domain. Mean distances of the genes from the chromosome edge were also measured in microns (see Table 7.2).

Gene	Mean probe location as proportion of chromosome radius	Mean distance of probe from edge of chromosome (μm)
WT1	0.358 \pm 0.017	0.947 \pm 0.045
3' RCN (50J5)	0.307 \pm 0.020	0.812 \pm 0.053
5' RCN (98L11)	0.315 \pm 0.021	0.833 \pm 0.056
PAX6	0.341 \pm 0.019	0.902 \pm 0.050
PAXNEB	0.326 \pm 0.018	0.863 \pm 0.048

Analysis of variance (Section 2.10), indicated that there is no significant difference between the position of genes across the murine WAGR locus within the MMU2 interphase chromosome domain of undifferentiated ES cells ($P=0.317$) (A. Carothers). Combining all data, the mean position for the WAGR locus in ES cells was 0.329 \pm 0.009 of a chromosome radius from the nearest chromosome edge, compared to 0.317 \pm 0.010 in human cells. Therefore, both the murine and human WAGR loci adopt a similar position within their respective chromosome territories, approximately one third of the distance along the calculated chromosome radius ($P=0.928$).

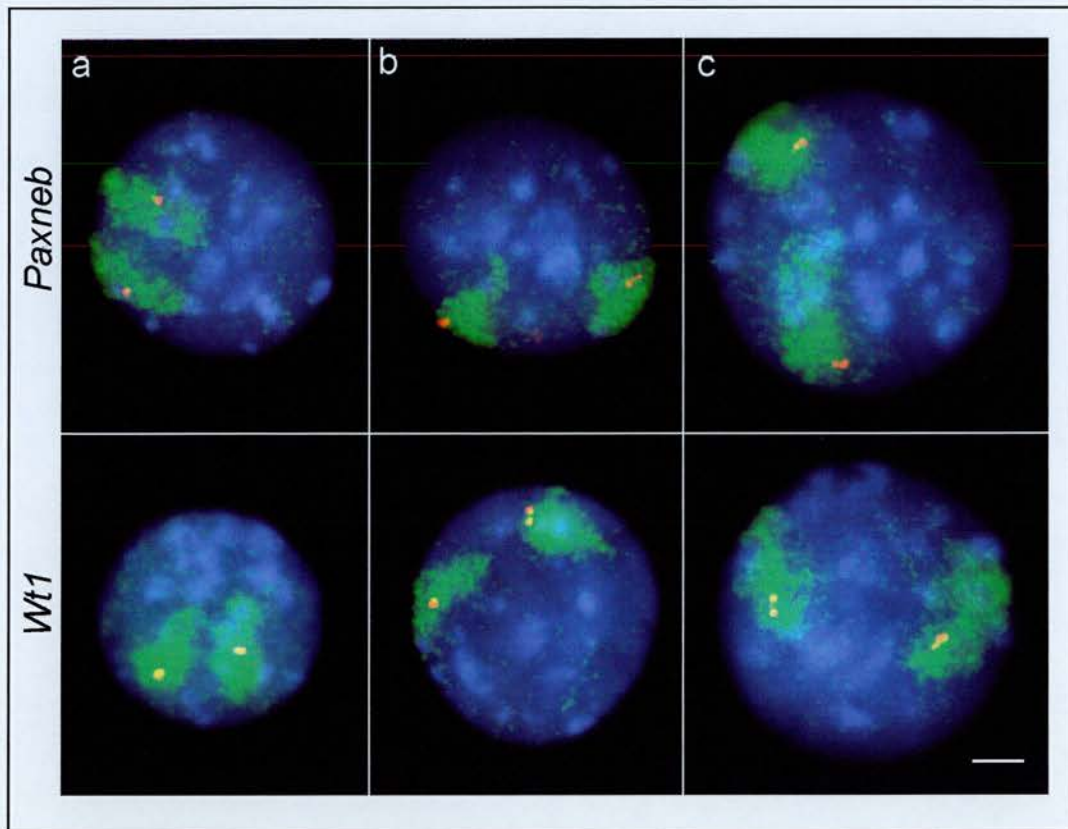


Figure 3.13 Analysing the WAGR locus within MMU2 territories of ES cell nuclei.

MAA-fixed ES cell nuclei were hybridised with biotin-labelled MMU2 paint (green) and a dig-labelled BAC probe (red), specific to the ubiquitously expressed gene Paxneb and the tissue-restricted gene Wt1. Nuclei in G1- or early S-phase of the cell cycle contained two chromosome domains with one locus signal each (a); nuclei undergoing S-phase showed one chromosome domain with two signals (locus already replicated) and one containing only one signal (yet to be replicated) (b); late S- or G2-phase nuclei in which both loci had already undergone replication had two BAC signals in each chromosome domain (c). The scale bar represents 5 μ m.

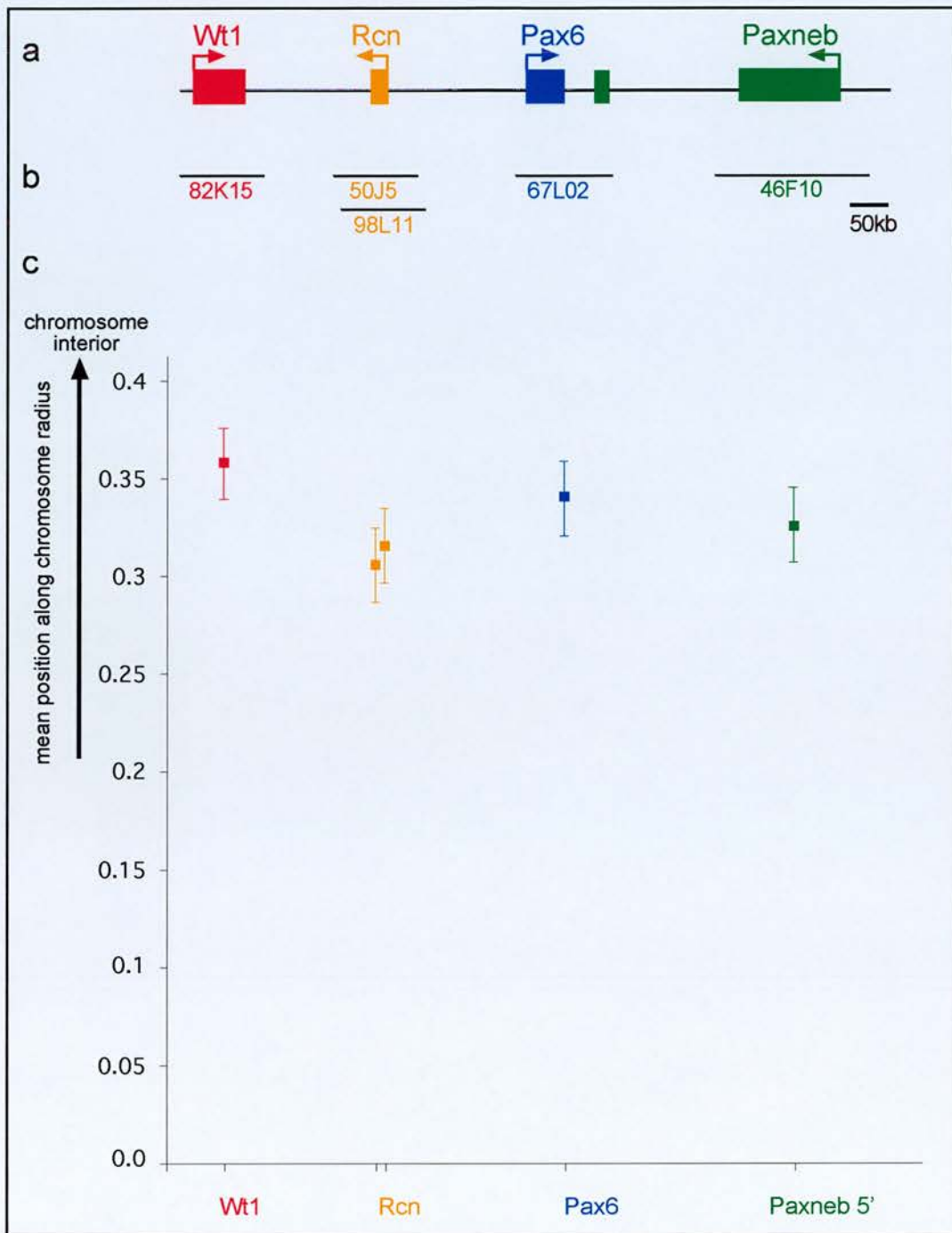


Figure 3.14 Position of sequences spanning the WAGR locus within the interphase MMU2 chromosome domain of mouse embryonic stem cells.

The murine WAGR locus is illustrated (a) and the relative positions of the BAC probes used in analysis of the organisation of this region within the interphase MMU2 domain are indicated (b). c) The mean normalised position of each BAC signal along the radius of the interphase MMU2 domain against the position of the probe across the WAGR locus is graphically depicted. Error bars indicate the standard error of the mean (SEM). A value of 0 represents the chromosome edge and 1.0 the chromosome centre.

3.7 RESOLUTION IN 2D MAA FIXED NUCLEI IS SUFFICIENT TO DETECT ANY PATTERNS OF LOCALISATION ACROSS THE WAGR LOCUS

The data presented in this Chapter show that there is no difference in the position of loci spanning 1 Mb of sequence from 11p13 within the HSA11p interphase chromosome domain, in four different cell lines. In light of this, I was worried that perhaps my analysis did not provide sufficient resolution to be able to detect significant differences in organisation at the sub-chromosomal level. *WT1* and *PAX6* are separated by 700kb of genomic DNA (Fantes *et al.*, 1995) (Figure 3.2). It has previously been shown that sites separated by 50kb or more are resolved in interphase nuclei using conventional fluorescence microscopy (Trask *et al.*, 1989). Previous data using FISH to 2D preparations has shown that, within human lymphoblast nuclei, signals from two cosmids encompassing *WT1* and *PAX6*, respectively, are separated by an average of $0.88^{±0.05}$ μm (n=50) (Fantes, 1997).

In the same lymphoblastoid cells I calculated the mean radius of the hybridisation signal from a paint for HSA11p as $1.866^{±0.006}$ μm (n=815). Therefore, *PAX6* and *WT1* loci could be separated by more than 0.5 of a chromosome radius. To confirm this, PACs spanning this region (Niederfuhr *et al.*, 1998; Gawin *et al.*, 1999) (Figure 3.2), were labelled with biotin and digoxigenin and used in combination with the HSA11p paint in FISH experiments. PAC hybridisation signals could be seen extending within chromosome territories for more than $1\mu\text{m}$ (Figure 3.15). Therefore, while there was no significant difference in the mean position of probes spanning the 1 Mb locus within the HSA11p domain, the DNA was not tightly package or coiled in one position, it was strung out within the chromosome, as expected.

In addition, the mean-square interphase distance ($\langle r^2 \rangle$) between the *WT1* and *PAX6* loci was intermediate to the values reported for gene-poor (G-band) and very gene-rich (T-band) regions of the human genome (Yokota *et al.*, 1997), at $0.89^{±0.11}$ μm . This is consistent with the moderate gene-density of 11p13, an R-band (Fantes *et al.*, 1995).

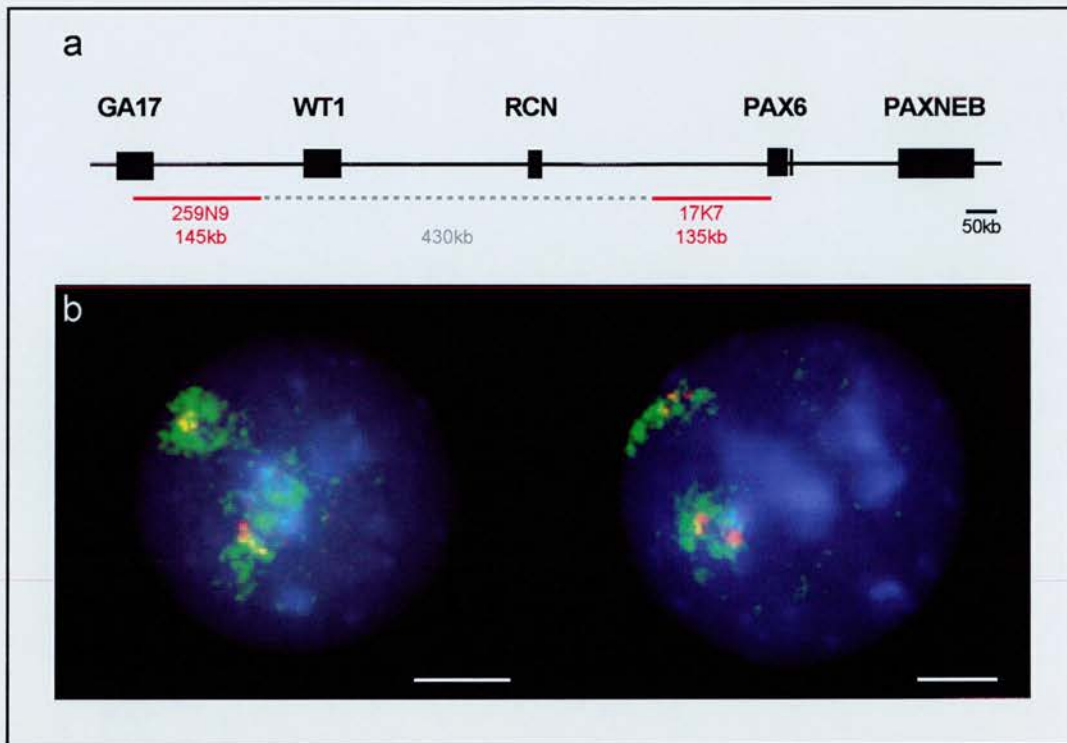


Figure 3.15 Combining PAC probes from the WAGR locus to visualise large stretches of DNA within the interphase nucleus.

PAC probes corresponding to sequences at either end of the WAGR locus (a) were used in FISH on lymphoblastoid nuclei in combination with a HSA11p chromosome paint. PAC signals (red) could be seen extending across the HSA11p domain (green) (b). This indicates that WAGR sequences are not tightly packaged, rather chromatin from 11p13 is strung out within the chromosome territory. Scale bars represent 5µm.

3.8. DISCUSSION

3.8.1 Genes are not located in a different position within a chromosome territory as compared to non-transcribed sequences

In this Chapter I have examined the positioning of a contiguous stretch of DNA within an interphase chromosome domain in four human cell lines. The sequence analysed spans the well-characterised WAGR locus in HSA11p13, which consists of at least 4 genes and includes two large, non-transcribed intergenic regions of DNA. My 2D analysis found no discernible difference in the localisation of actively transcribed genes, silent genes or non-transcribed intergenic sequences across this 1Mb of DNA within the 11p-chromosome territory.

Previous studies have assessed the position of genes and non-coding sequences within interphase chromosome territories. One such study (Kurz *et al.*, 1996), found three coding regions of the human genome to be located at the surface of their respective chromosome domains, independent of their transcription status and a single non-transcribed sequence was randomly distributed, or more internally positioned. The most extreme interpretation of this data is that most, or all, genes lie on the surface of interphase chromosomes at interphase, and that the chromosome interior is filled with non-coding, intergenic sequence, consistent with the ICD model. However, I have shown here that there is no significant difference in the location of protein-coding and non-coding sequence from 11p13, within HSA11p.

3.8.2 Gross organisation of interphase chromosomes is conserved from mouse to human

Extending the study of the WAGR locus to the syntenic region in the mouse is the first time that the organisation of murine chromosome domains has been considered. As this region is highly conserved in mammals, and also across to *Fugu*, it was interesting to note that the distribution of sequences across the murine locus within the MMU2 interphase chromosome was consistent with that of the human. Within MMU2, as within HSA11p, the mean WAGR position was one third of the distance along a calculated chromosome domain radius. These observations provided evidence for a non-random organisation of chromosome territories across mammalian species.

3.8.3 Small-scale changes in transcription do not affect gross organisation of interphase chromosomes

Two studies indicated a role for active transcription in organising genes within a chromosome domain (Dietzel *et al.*, 1999; Volpi *et al.*, 2000). Volpi and co-workers noted a difference in the incidence of extrusion of chromatin containing the MHC locus from the surface of HSA6 in cell lines with different expression profiles of MHC genes. In addition, inducing transcription from this region in primary fibroblasts, using IFN γ , directly increased the incidence of chromatin looping (Volpi *et al.*, 2000). In a second study, a transcription status-related trend was observed for the position of a single gene within the active and inactive X chromosome domains. The SLC25A5 (previously ANT2) gene was more peripheral within the Xa chromosome from which it is actively transcribed (Dietzel *et al.*, 1999) than within the Xi domain from where it is not transcribed.

Using cell lines that express the tissue-restricted *WT1* or *PAX6* genes, I found no significant alteration of the position adopted by the WAGR locus within a chromosome territory that correlated with increased gene expression. The entire locus was positioned approximately one third of a chromosome radius away from the territory surface in all cell lines studied, regardless of the transcriptional profile of WAGR. Genes that are inactive in lymphoblastoid and primary fibroblast cells were not relocated within HSA11p when actively transcribed, but rather remained within the HSA11p chromosome territory in a position similar to those of neighbouring ubiquitously active genes and adjacent non-coding sequences. Therefore, transcription machinery must be able to readily access sequences within the interior of chromosome territories.

On considering why my results seem to differ from studies that detected a transcription-dependent difference in territory organisation, I noted that these differences were ascertained within the context of chromosomes subject to large-scale transcriptional differences (Volpi *et al.*, 2000; Dietzel *et al.*, 1999). The MHC locus consists of families of genes with the same, or a related, function; therefore all genes are subject to the same regulation kinetics (Volpi *et al.*, 2000). This may be an important consideration, perhaps indicating that the observed looping of chromatin from the surface of the HSA6 territory results from rapid transcriptional up-regulation of many genes in a relatively small region. In addition, there is evidence to suggest that the EDC locus at 1q21, another region consisting of families of genes with the same function and therefore the same regulation kinetics, is also subject to a similar 'chromatin looping' from the surface of HSA1 on transcription up-regulation (R. Williams, personal communication). In the second case, there is little transcription from Xi

and it is questionable how widely applicable observations made from study of Xi chromosome domains will be to the rest of the human karyotype. The Xi chromosome adopts a unique structure which differs from the Xa territory (Dietzel *et al.*, 1998; Dietzel *et al.*, 1999) and is also generally devoid of sites of transcription (Verschure *et al.*, 1999).

Together these data indicate that alteration of large-scale chromatin structure may occur only on dramatic alteration of the transcription profile of a locus, perhaps during differentiation and commitment of cells along a particular developmental pathway (Brown *et al.*, 1997; Brown *et al.*, 1999). In general, there may be little gross re-modelling of chromosome domain organisation, detectable by light microscopy, to accommodate small/local changes in gene expression within mammalian cell lines. There is evidence to suggest that changes in chromatin structure at the nucleosome level are correlated with changes in gene activity, but we still lack experimental paradigms that show unequivocally whether changes in higher-order chromatin architecture are the cause or consequence of orchestrated changes in gene expression.

My data suggested that it is insufficient to study isolated loci from scattered genomic locations to dissect patterns of organisation within interphase chromosomes. It is perhaps not surprising that I found no significant difference in the organisation across this 1 Mb of sequence that correlates with the nature of the particular locus under consideration; genes are not isolated pieces of DNA and should not be treated as such. Context is obviously important and it is becoming increasingly apparent over just how great a distance DNA control elements can exert an effect (Kleinjan *et al.*, 2001b).

To address this apparent spatial problem I decided to continue utilising HSA11p as a model system to dissect patterns of organisation in the interphase nucleus and to turn my attention to the very gene-rich sub-telomeric R-band 11p15.5. This analysis forms the subject of Chapter 4.

CHAPTER 4

THE CONTRASTING INTERPHASE ORGANISATION OF OTHER LOCI FROM HSA11p

4.1 INTRODUCTION

In the previous Chapter I concluded from my analysis of 11p13 that the study of individual genes is insufficient to draw conclusions about links between transcription and higher-order chromosome structure. I suggest that it is important to consider more long-range contexts of genes. To address the organisation of the gene-dense regions of the mammalian genome, I decided to investigate the R-band 11p15, specifically the very gene-dense sub-telomeric T-band 11p15.5.

4.1.1 11p15.5 is very gene dense

11p13 (the R-band studied in Chapter 3) is moderately gene-rich, containing 4 genes/Mb DNA (Figure 3.2). In contrast, 11p15.5 (Figure 4.1a) contains at least 47 genes within 4.5Mb DNA (Figure 4.2) equivalent to more than ten genes per megabase. This estimate was achieved after assimilation of data from a variety of papers (Redeker *et al.*, 1994; Reid *et al.*, 1997; Hu *et al.*, 1997; Alders *et al.*, 1997; Paulsen *et al.*, 1998; Lee *et al.*, 1999; Kato *et al.*, 1999; Engemann *et al.*, 2000; Paulsen *et al.*, 2000; Onyango *et al.*, 2000) and also from the Human Genome Sequencing Project (<http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/maps.cgi?org=hum&chr=11>). Although the Human Genome Project is nearing completion, the extent to which the current sequence is accurately assembled and annotated varies considerably from one region to another. It was therefore not possible to use this resource alone to accurately predict the order and number of genes in 11p15.5 and maps from the literature were the most reliable resource. At the time of writing, the majority of sequence available for chromosome HSA11 was draft; finished sequence comprised only 47.1% of the human genome.

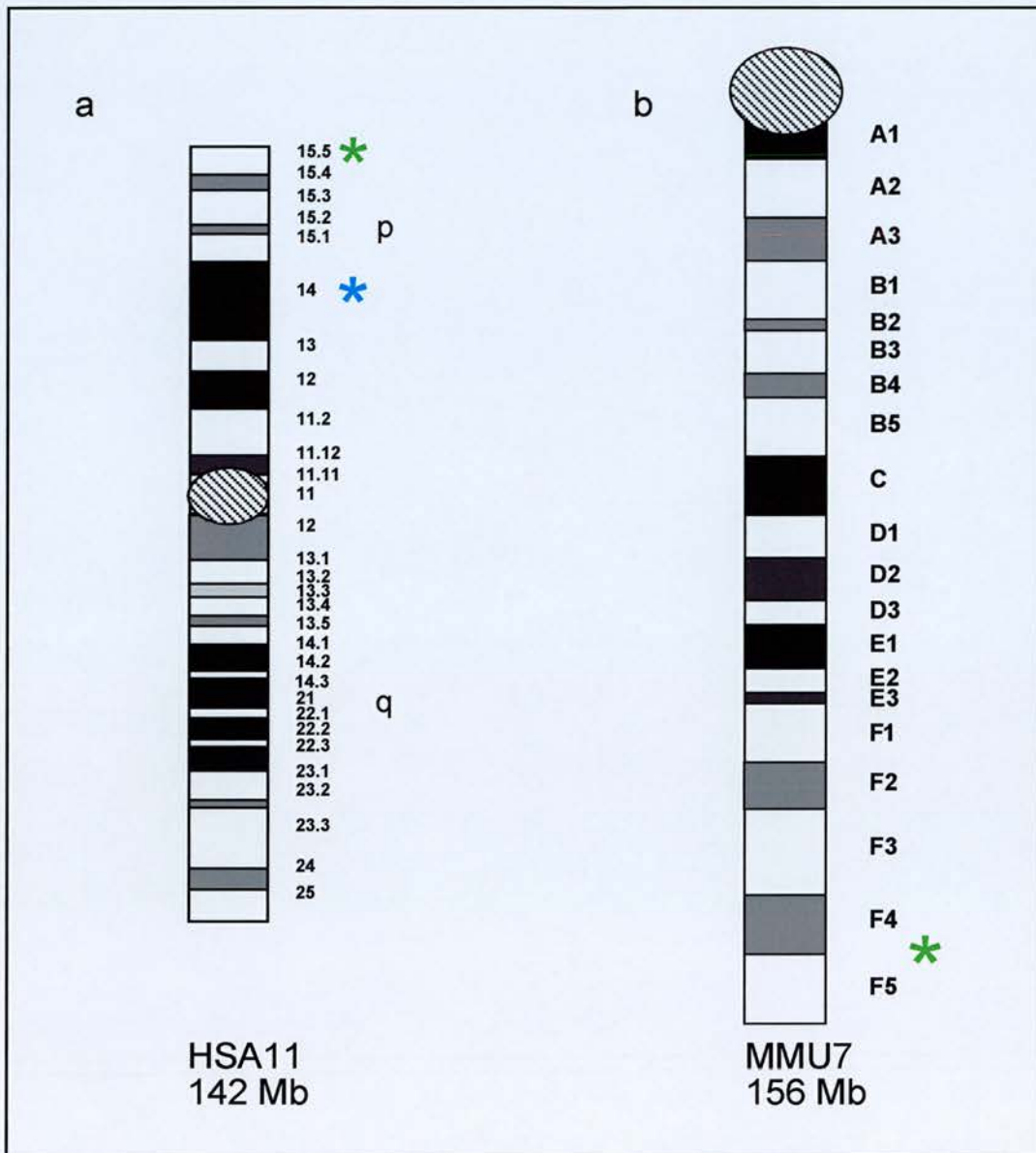


Figure 4.1 G-banded ideograms of human chromosome 11 (HSA11) and mouse chromosome 7 (MMU7).

Standard G-banded ideograms of (a) HSA11 taken from ISCN 1995 and (b) MMU7 (taken from Sawyer *et al.*, 1987). Diagrams are to scale and the bands containing the Insulin and Igf2 gene region on each chromosome are indicated by a green asterisk. The gene-poor band, 11p14, is indicated by a blue asterisk.

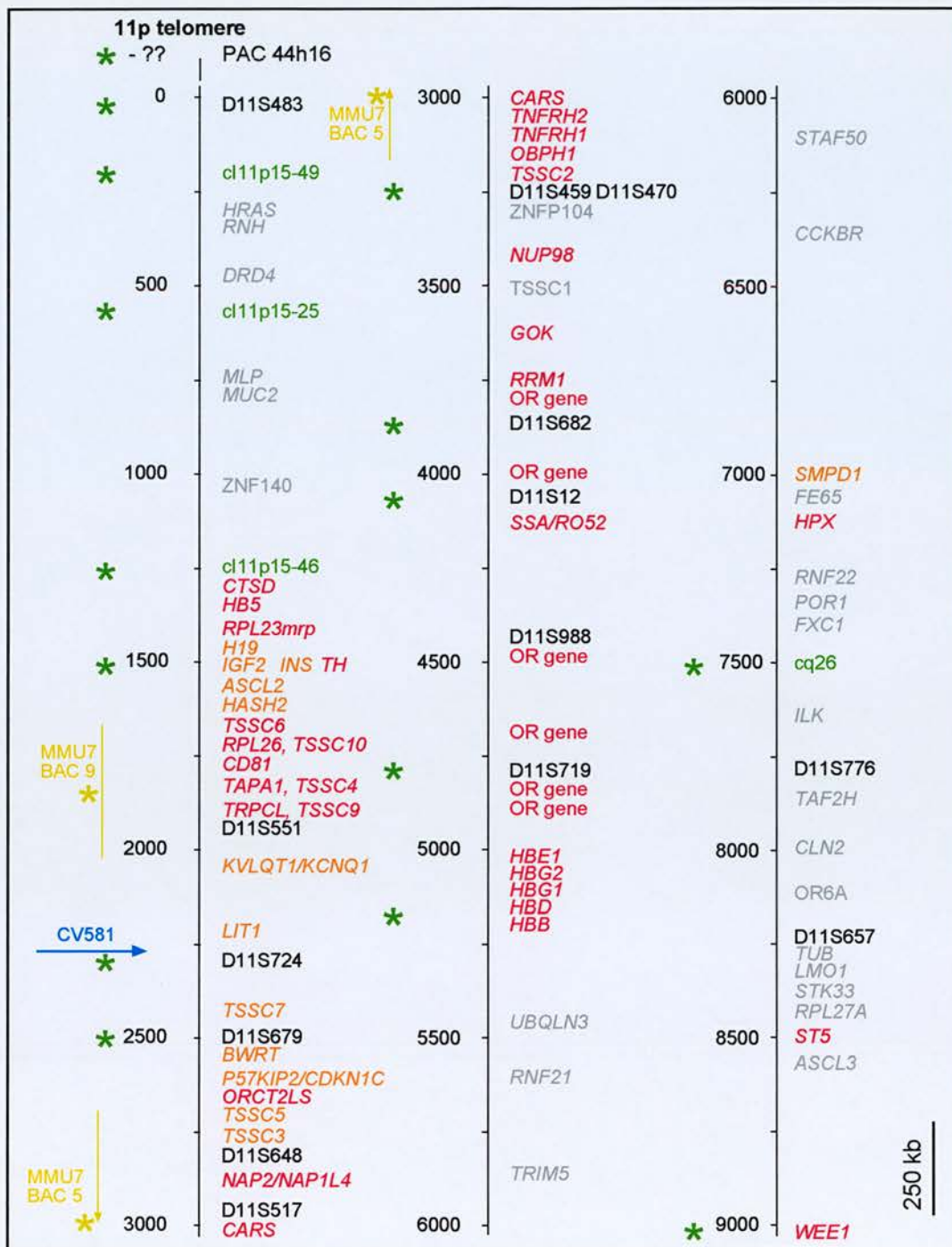


Figure 4.2 A map of the genes and markers of HSA11p15.

The relative positions of known genes and markers across 9Mb of DNA from HSA11p15 are illustrated. Genes mapped to high resolution are indicated in red and those known to be subject to imprinting are highlighted in orange. Genes identified from sequencing of the human genome are marked in grey, however, the positions of these genes are approximate, according to NCBI. Other markers are indicated in black, and cytogenetically mapped cosmid probes used for analysis are coloured in green. Probes that I have used for analysis of the organisation of this region are marked by an asterisk (human green; murine yellow). OR indicates the position of an olfactory receptor family gene. The position of the CV581 breakpoint in 11p15.5 is indicated (blue).

This map is approximately to scale and the relative order of genes has been assimilated from various references (Alders *et al.*, 1997; Hu *et al.*, 1997; Reid *et al.*, 1997; Buettner *et al.*, 1998; Paulsen *et al.*, 1998; Bepler *et al.*, 1999; Kato *et al.*, 1999; Lee *et al.*, 1999; Engemann *et al.*, 2000; Onyango *et al.*, 2000; Paulsen *et al.*, 2000 and the NCBI website: <http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/maps.cgi?org=hum&chr=11>). The map runs from the 11p telomere (top left), and spans 9 Mb to the WEE1 gene (bottom right). The scale represent kilobases of DNA.

4.1.2 Human clusters of imprinted genes

In the human, there are two main clusters of imprinted genes (genes expressed from only the maternal or the paternal allele) (Figure 1.1). One is contained in 11p15.5; the other is located on the long arm of human chromosome 15 (HSA15q) in cytogenetic bands 15q11-13. Disruption of the imprinting of genes in 11p15.5 is implicated in Beckwith-Wiederman syndrome (BWS) (Weksberg *et al.*, 1993); review Maher and Reik, 2000) and disruption of imprinting in 15q11-13 is associated with Prader-Willi and Angelmann syndromes (PWAS) (Cassidy *et al.*, 2000).

4.1.3 The murine region of synteny to HSA11p15.5

The murine region of conserved synteny to 11p15 is located on mouse chromosome 7 (MMU7), bands F3-F4 (Figure 4.1c). This is the seventh largest mouse chromosome, consisting of 156Mb of DNA (<http://www.ncbi.nlm.nih.gov/genome/seq/MmProgress.shtml>) and contains 789 genes (NCBI: http://www.ncbi.nlm.nih.gov/genome/guide/M_musculus.html). The order of genes and the size of this syntenic region are closely conserved to the human and the relative transcriptional orientations of the genes is also conserved (Paulsen *et al.*, 1998; Paulsen *et al.*, 2000; Engemann *et al.*, 2000). Like in the human, two non-imprinted mouse genes (Nap114 and Rpl231) flank the murine imprinted region; however, the orientation of the locus relative to the telomere is reversed (Onyango *et al.*, 2000).

4.2 LOCALISATION OF THE *INSULIN/IGF2/TH* GENES RELATIVE TO THE HSA11p DOMAIN

I decided to examine the organisation of chromatin from 11p15.5 within the HSA11p chromosome territory. I wanted to see whether there is a difference in the position adopted by gene dense sequences within a chromosome territory (11p15.5 is a T-band containing an average of >10 genes/Mb DNA) (Figure 4.2) compared to DNA from a moderately gene-rich region (exemplified by the WAGR locus with 4 genes/Mb, in R-band 11p13).

I selected the cosmid INSUL-IGF2 (Pagter-Holthuizen *et al.*, 1987), mapped to distal 11p15.5 (Figure 4.2), that contains the adjacent *INS*, *IGF2* and *TH* genes. I will refer to this cosmid as INS/IGF2 from now on. I used INS/IGF2, labelled with Dig, in combination with a HSA11p-specific paint labelled with biotin in a FISH experiment on lymphoblastoid cells. It was immediately obvious (Figure 4.3) that this region of HSA11p is subject to a radically different spatial organisation, as compared to chromatin from HSA11p13. Fifty images were analysed using the scripts described in Chapter 3 (Boxes 3.1 and 3.2). The mean position of the INS/IGF2 signal, expressed as distance from the nearest edge of the chromosome territory normalised for a calculated territory radius, was $-0.324^{\pm 0.069}$ (Table 4.1), the negative value indicating that the probe was frequently positioned outside of the defined chromosome territory. This suggested to me that 11p15.5 adopts an unusual spatial configuration in the interphase nucleus of lymphoblastoid cells and I decided to explore this in other cell types.

Table 4.1 Position of *INS/IGF2* relative to the interphase HSA11p chromosome domain.

Fifty images were captured depicting lymphoblastoid nuclei after FISH using a HSA11p paint in combination with a probe to the insulin gene. The scripts described in Chapter 3 (Boxes 3.1 and 3.2) were used to analyse the position of the locus signal in relation to the chromosome domain. Mean positions of the gene relative to the chromosome edge were calculated (Section 2.10).

Cell Line	Position of <i>INS/IGF2</i>	
	Mean position along territory radius	Mean distance from chromosome edge (μm)
Lymphoblastoid (FATO)	$-0.324^{\pm 0.069}$	$-0.605^{\pm 0.129}$
Primary fibroblast (1HD)	$-0.403^{\pm 0.079}$	$-0.732^{\pm 0.143}$
Lens epithelium (CD5a)	$-0.309^{\pm 0.082}$	$-0.486^{\pm 0.129}$

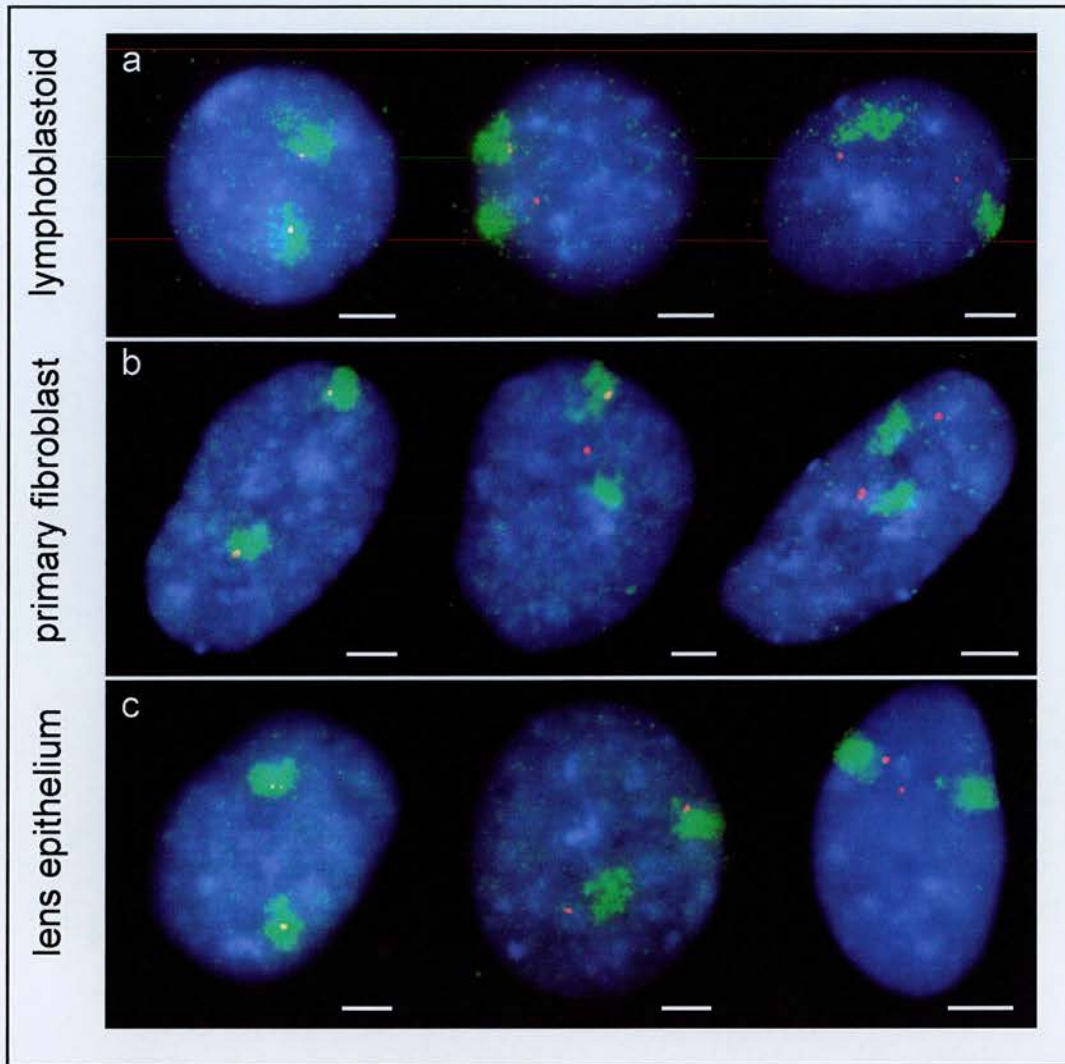


Figure 4.3 The organisation of *INS/IGF2/TH* genes in relation to HSA11p in lymphoblastoid, primary fibroblast and lens epithelium cell nuclei.

MAA-fixed nuclei were hybridised with biotin-labelled HSA11p paint (green) and a dig-labelled cosmid (red) containing the *INS* and *IGF2* genes. It is obvious from these images that chromatin from 11p15.5 is subject to different organisational constraints compared to 11p13 (Chapters 3 and 4). Cosmid probes frequently localised outside of the visible HSA11p domain, sometimes at a great distance from the edge of the chromosome territory. The probe was found both within and outside of the HSA11p territory in nuclei from all three cell lines examined: lymphoblastoid (a) primary fibroblast (b) and lens epithelium (c). Scale bars represent 5 μ m.

4.3 *INS/IGF2* LOCI ADOPT A SIMILAR POSITION RELATIVE TO THE HSA11p DOMAIN IN OTHER CELL LINES












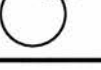
The positioning of *INS/IGF2* gene loci in relation to the interphase HSA11p chromosome territory was analysed in a lens epithelium transformed cell line (CD5a) and also in primary fibroblasts (1HD) (Figure 4.3 and Table 4.1). As in lymphoblasts, *INS/IGF2* signals were frequently positioned outside of the chromosome territory in these cell lines, which included a primary cell line, indicating that this phenomenon is not confined to immortalised cell lines maintained for many generations in culture.

INS and *IGF2* were not observed at a distance from their native chromosome territories in every nucleus examined, however. A negative value for the distance of the locus from the nearest chromosome edge (indicating localisation outside of the chromosome) was produced in 60.8% of cases. This can be broken down further to examine the frequency with which neither, both, or one of the alleles was found outside of its chromosome domain (Table 4.2).

Table 4.2 Patterns of *INS/IGF2* signals relative to the interphase HSA11p domain within lymphoblastoid cells.

The data set produced from analysis of fifty FISH images combining HSA11p chromosome paint with a probe to the *INS/IGF2* genes was analysed in terms of the combinations of allele signals positioned within or outside of each chromosome domain. If both signals of a doublet were inside or outside the territory, they were scored as ‘in’ or ‘out’ and were only differentiated from single signals if one signal was within the chromosome territory, and one was outside of the territory.



INS/IGF2 position vs. territory		% nuclei with this pattern
Territory 1	Territory 2	
In 	In 	18.2
Out 	Out 	22.7
In 	Out 	50.0
Out 	Doublet in/out 	9.1
In 	Doublet in/out 	0
Doublet in/out 	Doublet in/out 	0

In 50% of nuclei, one allele was positioned outside of the chromosome domain and one was within the chromosome. In a further 23% of nuclei, both alleles were positioned outside of the HSA11p territory and 18% of nuclei had both alleles within the HSA11p chromosome domain. This variable localisation of the *INS/IGF2* probe at a distance from the 11p domain prompts speculation that perhaps it is a dynamic phenomenon. One possibility was a cell cycle stage-dependent localisation. However, this was found not to be the case as there was no significant difference in the position of *INS/IGF2*, or any other locus studied, relative to the HSA11p domain when G1 and G2 nuclei were compared (Section 2.10) (see Section 3.4.1 for details on how G1 and G2 nuclei were identified).

4.4 MAPPING THE EXTENT OF THE 'LOOPING' STRUCTURE

On the basis of this intriguing result, I decided to extend my analysis of 11p15, in order to define the limits of this postulated 'loop' of chromatin emanating from the surface of the chromosome domain. A previous study (Volpi *et al.*, 2000) noted a similar configuration of chromatin containing the MHC locus from HSA6p21. In that analysis, the chromatin loops contained several megabases of DNA extending from the surface of the chromosome territory, as defined by a chromosome 6 or 6p paint. Therefore, I decided that in order to define the limits of the suggested chromatin loop from 11p15, I should consider probes covering at least 5 megabases of sequence proximal of the HSA11p telomere.

Cosmids spanning approximately 9 Mb from the distal cytogenetic marker D11S483, towards the centromere of chromosome 11, up to the *WEE1* gene in band 11p15.3 were obtained (gifts of Marielle Alders, Amsterdam) (Redeker *et al.*, 1994; Alders *et al.*, 1997 and references therein). A HSA11p telomere-specific PAC (44h16, Genome Systems Inc.) was also obtained (the sequence distance of the telomere from D11S483 is undetermined) (Figure 4.2). Annotation of this 9 Mb region is incomplete, but several detailed sequence studies have been published. The most distal 4.5 Mb of HSA11p are well characterised due to involvement of this region in several genetic diseases, including BWS (Onyango *et al.*, 2000 and references therein). The more proximal 5 Mb of sequence up to the *WEE1* gene is less well characterised, but some genes have been mapped within this region. A high-resolution physical map beginning just distal of *TSSC1* and spanning 1.4 Mb to beyond the beta globin gene cluster has been produced (Bepler *et al.*, 1999). Twenty-two putative genes were identified including the beta globin genes within a cluster of olfactory receptor genes (Buettner *et al.*, 1998; Bepler *et al.*, 1999).

The mean positions of probes spanning this 9 Mb of sequence relative to the nearest edge of the HSA11p territory in lymphoblastoid and primary fibroblast nuclei were calculated by analysis of FISH images (as described in Section 3.5.1). Results are recorded in Table 4.3 and depicted in Figure 4.4.

Table 4.3 Mean positions of loci spanning more than 9 megabases of distal HSA11p relative to the HSA11p territory.

Probes were labelled with digoxigenin and used in FISH on MAA fixed nuclei from both a lymphoblastoid cell line (FATO) and primary fibroblasts (1HD), in combination with HSA11p-biotin chromosome arm paint. Fifty images were captured for each probe and the positions of locus signals within chromosome domains were analysed using the scripts described in Chapter 3. Mean positions of loci relative to the nearest chromosome edge were calculated (\pm SEM) (Section 2.10). n/d indicates a result not determined.

Gene/ locus from HSA11p	Lymphoblastoid cell line		Primary fibroblast cell line	
	Mean probe position along chromosome radius	Mean distance of probe from edge of chromosome (μ m)	Mean probe position along chromosome radius	Mean distance of probe from edge of chromosome (μ m)
WAGR (p13)	0.315 \pm 0.010	0.588 \pm 0.019	0.282 \pm 0.010	0.512 \pm 0.018
WEE1 (p15.3)	n/d	n/d	0.106 \pm 0.042	0.192 \pm 0.076
cq26	0.099 \pm 0.044	0.185 \pm 0.082	n/d	n/d
pLCR (β -globin, p15.4)	0.160 \pm 0.035	0.299 \pm 0.065	n/d	n/d
D11S719	0.126 \pm 0.043	0.235 \pm 0.080	0.153 \pm 0.032	0.278 \pm 0.058
D11S12	-0.043 \pm 0.052	-0.080 \pm 0.097	0.125 \pm 0.033	0.227 \pm 0.060
D11S682	0.037 \pm 0.048	0.069 \pm 0.090	n/d	n/d
D11S470	-0.307 \pm 0.069	-0.573 \pm 0.129	n/d	n/d
D11S679	-0.251 \pm 0.076	-0.468 \pm 0.142	-0.465 \pm 0.076	-0.844 \pm 0.134
D11S724	-0.474 \pm 0.102	-0.885 \pm 0.190	n/d	n/d
INS	-0.325 \pm 0.069	-0.607 \pm 0.129	-0.403 \pm 0.079	-0.732 \pm 0.143
Cl-11p15-46	-0.573 \pm 0.107	-1.069 \pm 0.120	-0.467 \pm 0.074	-0.845 \pm 0.134
Cl-11p15-25	-0.460 \pm 0.068	-0.858 \pm 0.127	n/d	n/d
Cl-11p15-49	-0.640 \pm 0.090	-1.194 \pm 0.168	n/d	n/d
D11S483	-0.689 \pm 0.089	-1.286 \pm 0.166	-0.809 \pm 0.084	-1.469 \pm 0.153
11p tel (p15.5)	-0.750 \pm 0.090	-1.340 \pm 0.168	-0.744 \pm 0.086	-1.351 \pm 0.157

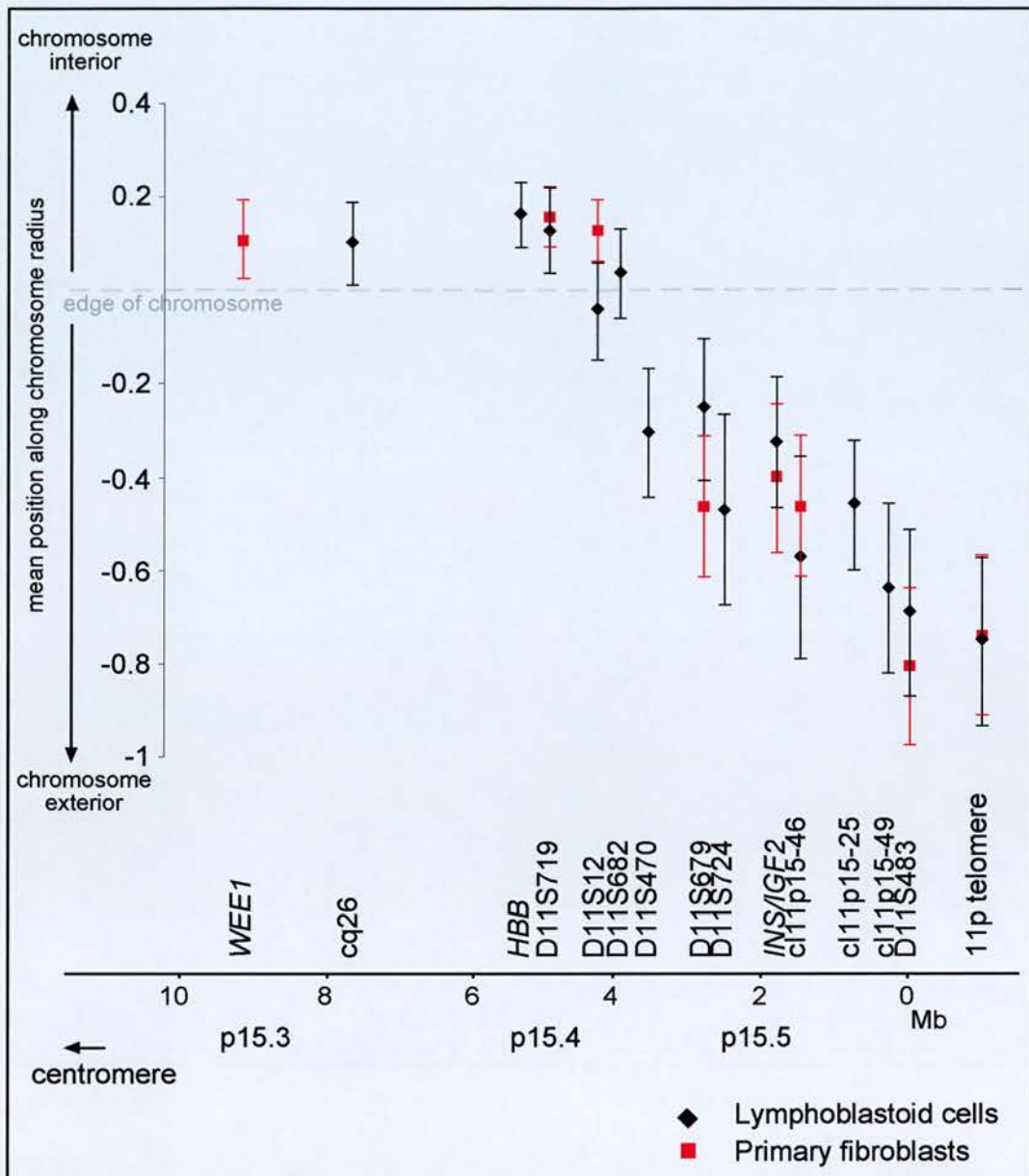


Figure 4.4 Sequences from distal 11p15.5 are extruded from the HSA11p territory in lymphoblastoid and primary fibroblast cells.

The organisation of sequences from HSA11p15 relative to the interphase HSA11p chromosome domain is illustrated graphically for both lymphoblastoid (black) and primary fibroblast cells (red). The mean normalised position of each probe is represented \pm SEM and the relative positions of the probes along the x axis are to scale, as indicated. A value of 0 represents the edge of the chromosome and 1.0 the chromosome centre. The actual distance of the most distal marker, D11S483, to the telomere is undetermined. Note the very similar configuration adopted by chromatin from 11p15.5 outside of the HSA11p domain in both cell lines.

Data presented in Table 4.3 and Figure 4.4 suggests that the configuration adopted by DNA from 11p15.5 is not actually a looping structure. Rather it appears that a tail of DNA extends from the surface of the chromosome domain. This contrasts with the organisation of the MHC locus on chromosome 6 (Volpi *et al.*, 2000), where a defined loop was observed which re-entered the chromosome domain before the gene-poor G-band 6p24. Analysis of the mean positions of probes spanning 11p15.5 suggested that the extrusion of chromatin from the surface of HSA11p begins around the markers D11S12 and D11S682. Both probes localised at the surface of the chromosome domain with a mean position $-0.043^{±0.052}$ and $0.037^{±0.048}$ of a territory radius away from the chromosome edge respectively ($-0.080^{±0.097}$ μm and $0.069^{±0.090}$ μm). The tail of chromatin outside of the main chromosome territory extended to the telomere, which had a mean position $-0.750^{±0.090}$ of a territory radius away from the chromosome edge ($-1.340^{±0.168}$ μm), in lymphoblastoid cells. A similar organisation was also found in primary fibroblast nuclei where extrusion of chromatin started around D11S12 and extended to the telomere. The telomere had a mean position $-0.744^{±0.086}$ of a territory radius away from the chromosome edge ($-1.351^{±0.157}$ μm) outside of the chromosome domain, very similar to that of the telomere in lymphoblastoid cells.

I have presented evidence from three cell lines for the existence of a tail of chromatin emanating from the surface of the interphase HSA11p chromosome domain. It has previously been suggested that delicate structural features of chromosome territories, such as tiny extensions from the surface, might have been overlooked in FISH analysis of interphase chromosome structure as they are difficult to detect with chromosome painting (Eils *et al.*, 1996). Using FISH analysis and light microscopy, I suggest that the only way to visualise such projections would be to identify loci, which lie on these stretches of chromatin, and to analyse their position relative to their native chromosome domains. Few studies have investigated the interphase organisation of specific sequences within interphase chromosomes at such a level and this is probably why this phenomenon has only been reported on two other occasions (Volpi *et al.*, 2000 and R. Williams, personal communication).

4.5 CORRELATION OF RAW DATA FROM THIS ANALYSIS WITH A PREVIOUS STUDY OF 'LOOPING' CHROMATIN

The only published study which described a similar phenomenon identified a looping structure adopted by chromatin from the MHC locus in HSA6p21 (Volpi *et al.*, 2000). The organisation of this region was assessed by calculating the % of locus signals positioned outside of the HSA6 or HSA6p domain for probes spanning 4 Mb rather than actual distances. I applied this method of analysis to my data and compared it to the calculated mean distance of the loci from the chromosome surface (Table 4.4 and Figure 4.5). I wanted to test the relationship between a qualitative (Volpi *et al.*, 2000) and a quantitative (my study) method of analysis.

Table 4.4 The frequency with which a sequence from 11p15.5 locates inside of the HSA11p domain correlated closely with the mean distance of the signal from the nearest edge of the chromosome territory.

The percentages of signals that localised within the HSA11p domain, as defined by a chromosome paint, were calculated and compared to the mean distance of each probe from the nearest edge of the chromosome.

Gene/ locus from HSA11p	Lymphoblastoid cell line		Primary fibroblast cell line	
	% signals localised within chromosome domain	Mean distance of probe from edge of chromosome (μm)	% signals localised within chromosome domain	Mean distance of probe from edge of chromosome (μm)
WAGR (11p13)	89.6	0.588 \pm 0.019	89.3	0.512 \pm 0.018
WEE1 (p15.3)	n/d	n/d	59.9	0.192 \pm 0.076
cq26	65.9	0.185 \pm 0.082	n/d	n/d
PLCR (β -globin, p15.4)	87.18	0.299 \pm 0.065	n/d	n/d
D11S719	59.2	0.235 \pm 0.080	66.4	0.278 \pm 0.058
D11S12 (p15.5)	54.4	-0.080 \pm 0.097	57.4	0.227 \pm 0.060
D11S682	65.7	0.069 \pm 0.090	n/d	n/d
D11S470	34.7	-0.573 \pm 0.129	n/d	n/d
D11S679	48.2	-0.468 \pm 0.142	32.5	-0.844 \pm 0.134
D11S724	41.2	-0.885 \pm 0.190	n/d	n/d
INS	39.2	-0.607 \pm 0.129	27.8	-0.732 \pm 0.143
Cl-11p15-46	29.1	-1.069 \pm 0.120	25.0	-0.845 \pm 0.134
Cl-11p15-25	22.4	-0.858 \pm 0.127	n/d	n/d
Cl-11p15-49	26.3	-1.194 \pm 0.168	n/d	n/d
D11S483	23.9	-1.286 \pm 0.166	12.4	-1.469 \pm 0.153
11p tel	16.3	-1.340 \pm 0.168	13.1	-1.351 \pm 0.157

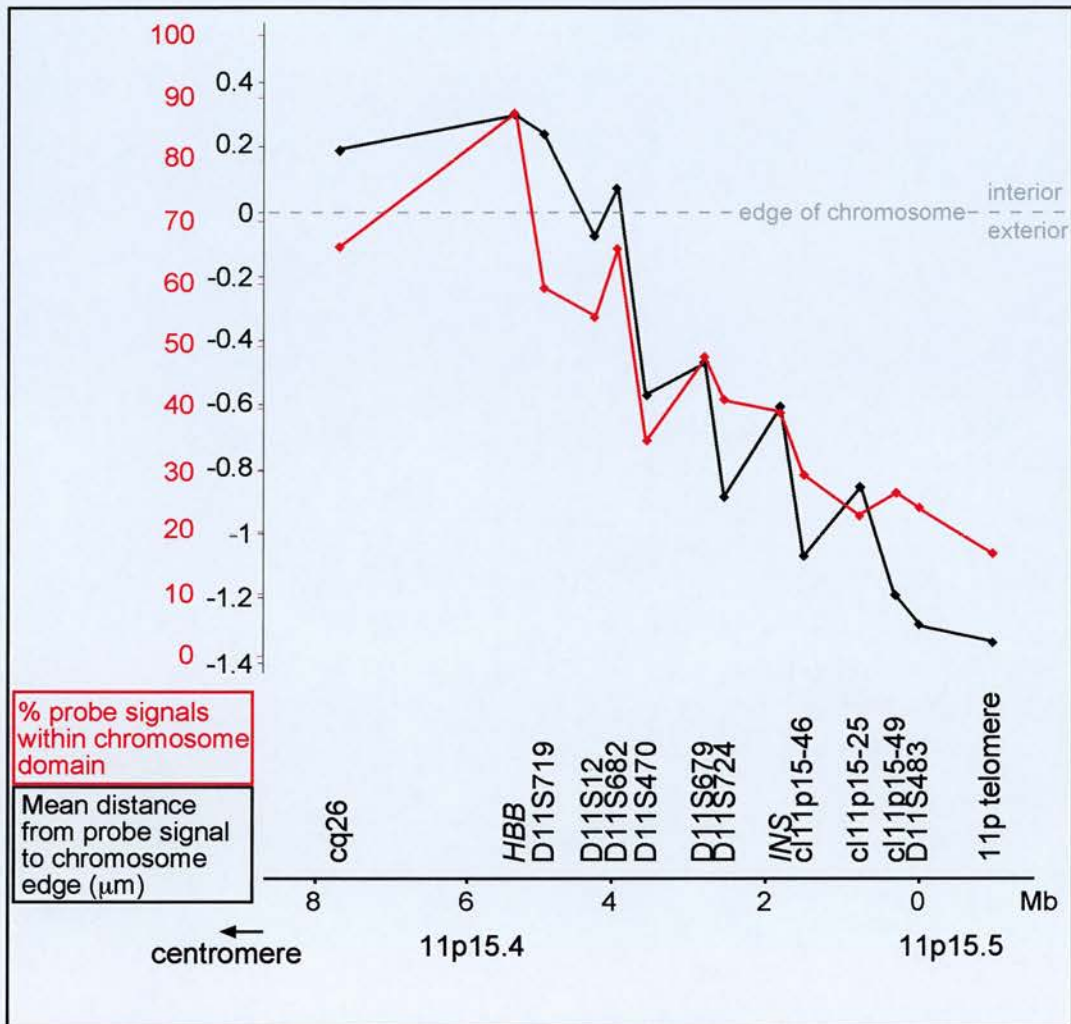


Figure 4.5 Correlation of qualitative and quantitative 2D analysis of the organisation of 11p15.5 in lymphoblastoid cell nuclei.

Two alternative methods for analysing output data for probe localisation were compared. One graph trace (red) represents a qualitative analysis where the % of signals found within the HSA11p interphase chromosome territory were calculated. The black trace represents a quantitative analysis where the mean distances of loci from the nearest edge of the interphase chromosome domain were calculated. Negative values indicate that the average locus position was outside of the HSA11p chromosome domain, as visualised by a chromosome paint. Note the close correlation between the two analysis methods.

Comparing the mean distance of loci away from the chromosome to the percentage of signals localising inside of the territory (Figure 4.5) showed the correlation to be very strong, validating two independently developed methods of analysis. Ten percent of WAGR probes apparently localised outside of the chromosome domain, although the mean position was well within the chromosome; whereas, the % of signals outside of the territory was significantly higher for 11p15.5 probes, especially for the most distal loci.

The quantitative analysis technique presented in this thesis enabled the spatial position of the loop of chromatin to be visualised in real terms. Therefore, I could begin to examine the extent of the decondensation of 11p15 by relating physical DNA length (kb) with nuclear positioning of loci from the edge of the chromosome territory (μm) (Table 4.3). I observed a tail of chromatin emerging from the chromosome 11p domain, beginning around cytogenetic markers D11S12 and D11S682. As the exact distance of the telomere from the next probe, D11S483 is not quantified I will concentrate on known distances: D11S12 and D11S483 are 4 Mb apart. Therefore, the mean positions of loci separated by 4Mb of linear sequence were approximately $1.323^{\pm 0.141}$ μm apart within the nucleus (Figure 4.4).

A previous study estimated the spatial separation of loci with known physical separations (Yokota *et al.*, 1995). In this study, it was calculated that loci 150 kb apart on the linear sequence could be separated by up to 0.2 μm in the interphase nucleus. Applying these calculations to my data estimates that D11S12 and D11S483 could in fact be separated by up to 5.33 μm . However, the relationship between sequence and spatial separation was found to be biphasic with a transition at approximately 2 Mb. This biphasic relationship was interpreted to indicate the existence of two organisational levels, the first an arrangement of chromatin in large loops of several Mb in size (perhaps equivalent to the 30 nm fibre) and the second specific attachments of these loops, or coiling of the 30 nm fibres predicted by various models of interphase chromatin organisation (Sedat and Manuelidis, 1978; Manuelidis, 1990; Manuelidis and Chen, 1990; Belmont and Bruce, 1994). Interestingly, in the most extreme case, my analysis found D11S483 5.93 μm outside of the HSA11p chromosome territory, which correlated closely with the 5.33 μm estimated using evidence provided for the first level of organisation by Yokota *et al.*. Taken together these data suggest that perhaps the entire 4 Mb of chromatin can be extended to a level similar to that of the 30 nm fibre and attachments or higher levels of organisation into which chromatin is normally folded within an interphase territory are diminished in this gene-dense region.

4.6 THE HSA11p COMPLEX PROBE COMPREHENSIVELY PAINTS THE ENTIRE LENGTH OF THE CHROMOSOME ARM, AS VISUALISED AT METAPHASE

11p15.5 is the most distal band of the short arm of human chromosome 11. Therefore, it could be argued that the reason the most telomeric sequences from this band appeared to localise outside of the chromosome domain was simply because the HSA11p chromosome probe did include, or did not comprehensively 'paint' sequences at the end of the chromosome. To show that this was not the case, I used the, IPLab Spectrum extension 'GraphPolygon' (Digital Scientific) to compare relative signal intensities of the chromosome paint and the locus probe along metaphase chromosome arms (Figure 4.6). From this analysis it was clear that the green paint signal extended right to the end of the visible chromosome, ending at a point coincident with a probe to the telomere of HSA11p on metaphase chromatids (Figure 4.6d and e). In interphase nuclei, however, the 11p telomere had a mean position $-1.340^{\pm 0.168}$ μm outside of the visible chromosome domain. When the position of a probe (c11p15-46, red) which localised a mean distance of $-1.069^{\pm 0.120}$ μm outside of the interphase HSA11p chromosome domain was compared to the HSA11p paint signal (green) at metaphase, it was clear that the paint extended well beyond the locus probe (Figure 4.6i and j). Since sequences that localise outside of the interphase chromosome domain are represented in the 11p chromosome paint, the chromatin fibre on which they were positioned at interphase must be too fine to be visible using FISH and light microscopy techniques.

The ideal control for this was to analyse the localisation of the centromere of chromosome 11, the most proximal sequence on HSA11p, to see whether it too was positioned outside of the HSA11p interphase territory. An HSA11p-specific alphoid probe was obtained (Fantes, 1997) and used in FISH in combination with the HSA11p paint. Analysis of fifty images using scripts (Boxes 3.1 and 3.2) produced a mean position at the periphery (but inside of) the HSA11p domain ($0.071^{\pm 0.078}$ of a chromosome radius from the surface, equivalent to $0.133^{\pm 0.146}$ μm). Alphoid sequences are not represented in the HSA11p paint (Figure 3.4), and at metaphase the alphoid signal would appear beyond the signals corresponding to the chromosome paint (data not shown). Therefore this result provides further evidence that the

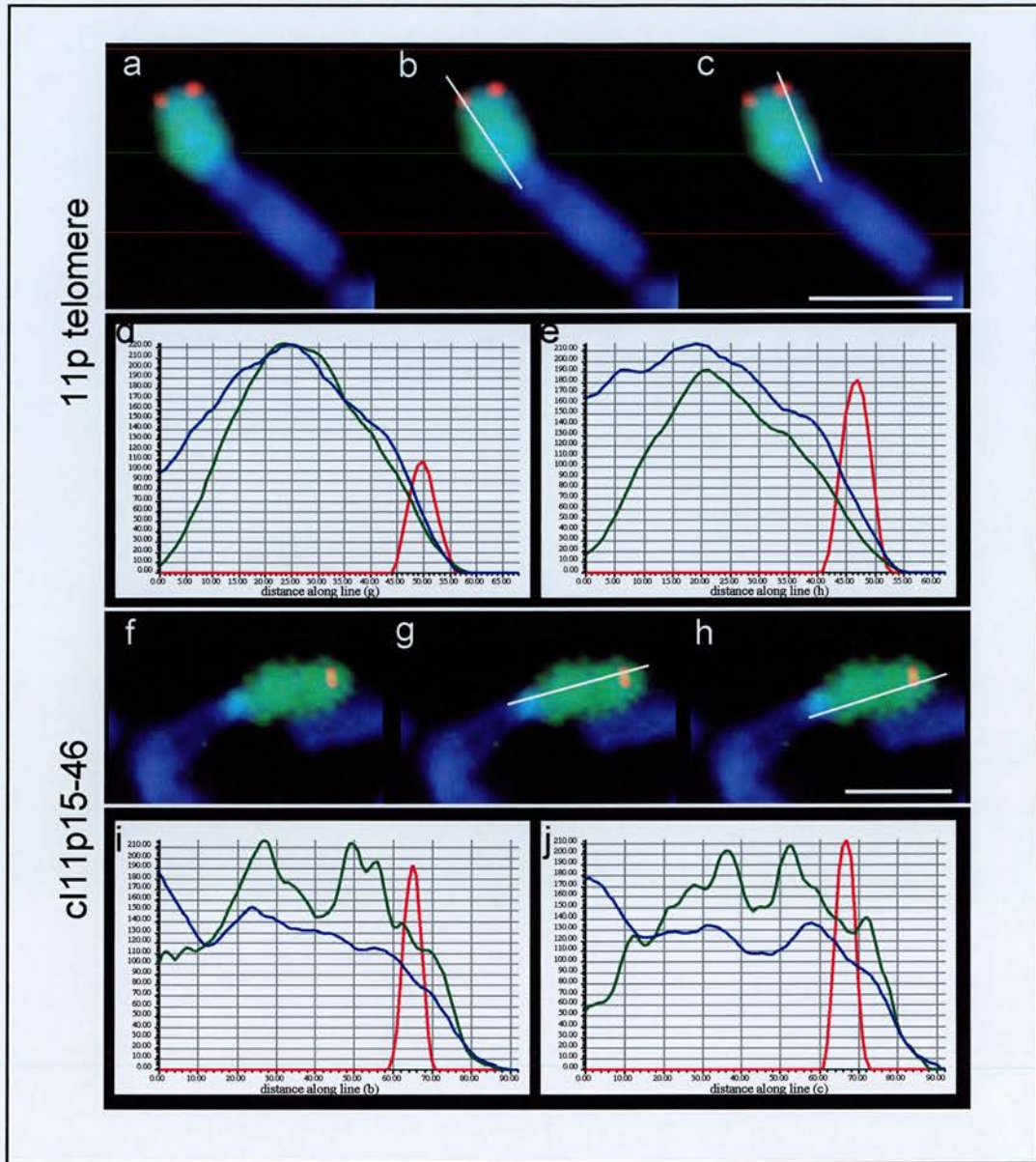


Figure 4.6 The HSA11p-specific paint hybridises to sequences spanning the entire length of the 11p chromosome arm.

HSA11 metaphase chromosomes after FISH with a HSA11p paint (green) and cosmid/PAC probes from 11p15.5 (red), counterstained with DAPI (blue) (a and f). Lines were drawn along the length of HSA11p through each probe signal (one on each sister chromatid of the homologue) (b,c g and h) and, using the IPLab Spectrum software extension 'polygon graph' (Digital Scientific), graphs of the relative intensities of the three colour signals were produced illustrating graphically the intensities of fluorescence for each colour along each sister chromatid (d, e, i and j). Scale bars represent 5 μm.

localisation of the telomere and more proximal sequences from 11p15.5 outside of the chromosome territory cannot be explained by incomplete coverage of the 11p paint.

4.7 THE POSITION OF OTHER TELOMERIC SEQUENCES WITHIN INTERPHASE CHROMOSOME DOMAINS

In yeast, telomeres are clustered together to form a compartment near the nuclear periphery (Gotta *et al.*, 1996). Human telomeres, however, form discrete nucleoprotein complexes occupying dispersed, prominently non-peripheral positions throughout the nuclear volume (Luderus *et al.*, 1996). There is also some evidence to suggest that telomere associations occur in interphase nuclei (Nagele *et al.*, 2001) (more predominantly in quiescent than in cycling cells) and perhaps play a role in maintaining chromosome position. It has also been suggested that subtelomeric chromosome regions that contain a high density of genes show an elevated incidence of somatic pairing in human nuclei (Stout *et al.*, 1999). There is no data in the literature describing the position of telomeres in relation to their chromosome domains in interphase human nuclei. Therefore a second possible explanation for the observation of distal sequences outside of the HSA11p domain was that all human telomeres adopt a position distinct from the main body of their native chromosome territory in the interphase nucleus.

To address this, the positions of probes corresponding to telomeres from the p and q arms of HSA18 (Fantes, 1997) were analysed in relation to the HSA18 chromosome territory (Table 4.5).

Table 4.5 Location of telomere sequences in relation to their respective interphase chromosome domains.

Fifty images were captured depicting nuclei after FISH using chromosome paints in combination with probes to the corresponding telomeres (S. Boyle). The scripts described in Chapter 3 (Boxes 3.1 and 3.2) were used to analyse the position of probe signals in relation to the chromosome arm domain. Mean normalised positions of the telomeres relative to the chromosome edge, and the corresponding distances in microns were calculated (\pm SEM) (Section 2.10).

Telomere	Mean normalised position of telomere relative to edge of chromosome domain	Mean distance of telomere from nearest chromosome edge (μm)
11p	$-0.750^{\pm 0.090}$	$-1.340^{\pm 0.168}$
18p	$0.069^{\pm 0.043}$	$0.129^{\pm 0.080}$
18q	$-0.100^{\pm 0.061}$	$-0.187^{\pm 0.114}$

From these data it was obvious that the telomere from HSA11p localised significantly further outside of its chromosome domain than the telomeres on chromosome 18 ($P < 0.001$ using a standard t test), which are positioned at the surface of the HSA18 territory. In the most extreme case, a HSA11p telomere signal was observed 6.219μ m outside of the 11p chromosome domain. The mean nuclear radius for FATO lymphoblastoid cells was $11.491^{\pm 0.032}\mu$ m ($n=3498$). Therefore the telomere localised up to approximately half the length of the nuclear radius (0.54) outside of the chromosome domain.

4.8 USING A CELL LINE WITH AN INVERSION WITHIN HSA11p TO FURTHER DEFINE PATTERNS OF ORGANISATION

To examine further the role of chromosome context in the intra-chromosomal positioning of loci, a lymphoblastoid cell line with an inversion of almost the entire p arm of HSA11 ($inv(11)(p11.2 p15.5)$) was obtained (CV581), from a patient with Beckwith-Weidemann syndrome (BWS) (Norman *et al.*, 1992). The distal breakpoint was below the

INS gene, around marker D11S724, and the proximal breakpoint was distal of the *RAPSN* gene 11p11.2 (Buckel *et al.*, 1996) (CV581 breakpoints characterised by J. Fantes, personal communication) (Figure 4.7). I used this cell line to see if there was any difference in the positioning of either the *INS/IGF2* genes, which remain in 11p15.5 on the inverted chromosome arm or the D11S470 marker, which is now adjacent to HSA11p11.2 on the inverted chromosome.

Table 4.6 Comparison of the position of loci from 11p15 in a cell line containing an inversion within HSA11p and a control cell line

Fifty images were captured of FATO of CV581 nuclei after FISH on MAA-fixed nuclei using a HSA11p chromosome paint in combination with probes mapped to 11p15.5. The scripts described in Boxes 3.1 and 3.2 were used to analyse the position of probe signals in relation to the HSA11p domain. Mean normalised positions for the position of the loci relative to the chromosome edge (\pm SEM) (Section 2.10) and also mean distances in microns were calculated.

Locus	FATO (control)		CV581 (inv(11)(p11.2 p15.5))	
	Mean position along HSA11p radius	Distance to chromosome edge (μ m)	Mean position along HSA11p radius	Distance to chromosome edge (μ m)
<i>INS/IGF2</i>	$-0.325^{\pm 0.069}$	$-0.607^{\pm 0.129}$	$-0.711^{\pm 0.102}$	$-1.165^{\pm 0.167}$
D11S470	$-0.307^{\pm 0.069}$	$-0.573^{\pm 0.129}$	$-0.193^{\pm 0.096}$	$-0.316^{\pm 0.157}$

In the control cell line, *INS/IGF2* and D11S470 had similar mean positions outside of the HSA11p domain ($P=0.854$) using a standard t test). However, in the inverted cell line the means were significantly different ($P<0.001$ using a standard t test). The *INS/IGF2* genes were more external to HSA11p in the inversion cell line, compared to the control and D11S470 had a mean position closer to the chromosome. Therefore, both loci were behaving differently in the inversion cell line, but it was not possible to attribute these differences to the inversion specifically, as there may be differences in interphase chromatin organisation between cell lines without chromosomal rearrangements.

To investigate this in a statistical manner, analysis of variance was performed on the data sets. In all previous analyses, the variation between chromosomes within a nucleus had been small enough to ignore. This was the case for both *INS/IGF2* and D11S470 loci in the

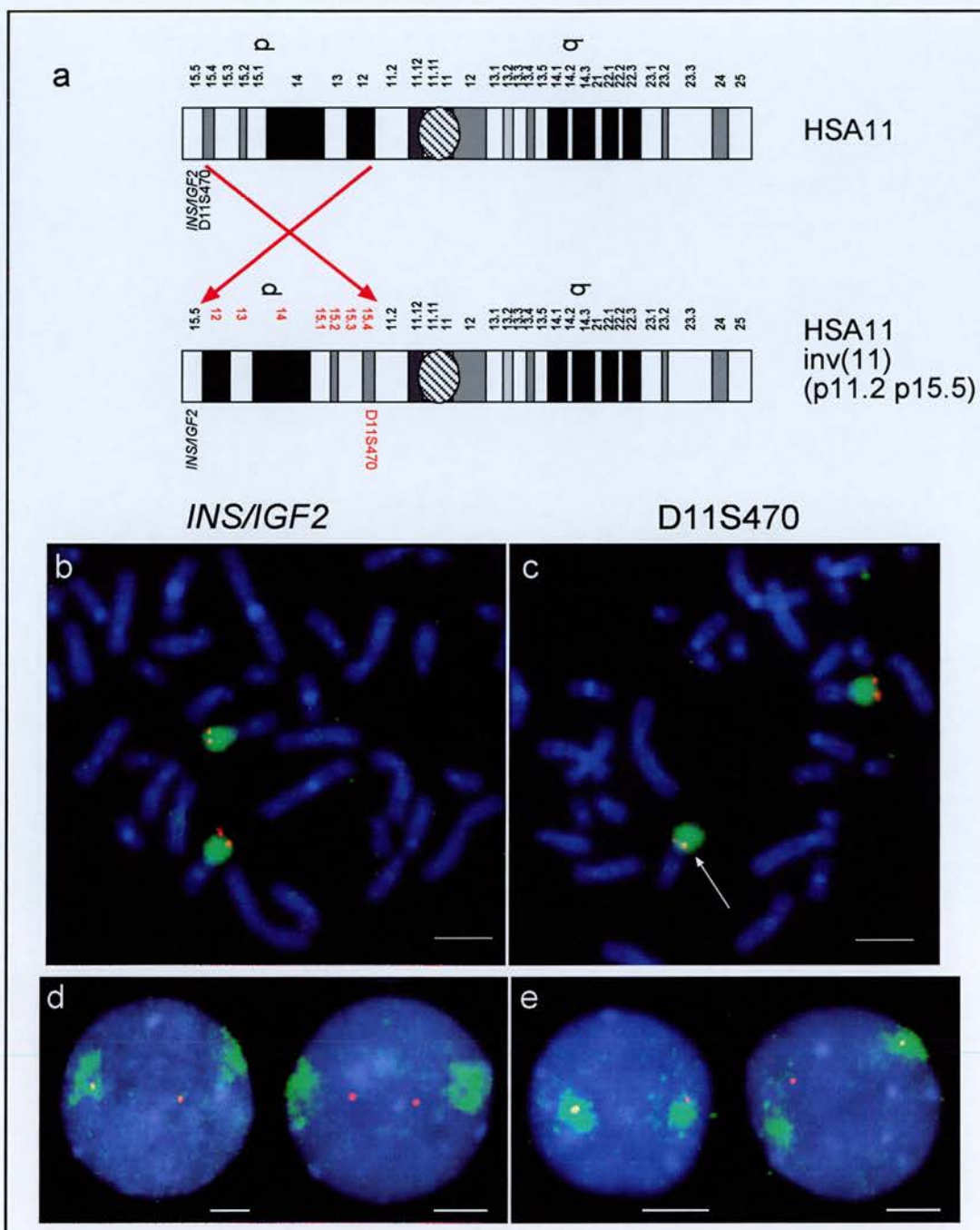


Figure 4.7 Analysis of a lymphoblastoid cell line, CV581, with an inversion within HSA11 (inv(11)(p11.2 p15.5)).

a) CV581 has one normal HSA11 homologue (top) and one (bottom) in which a large part of the p arm has been inverted: inv(11)(p11.2 p15.5).

b and c) Metaphase chromosome spreads counterstained with DAPI after FISH with a HSA11p paint (green) combined with a locus from HSA11p illustrate the resultant reorganisation of the inverted homologue. b) The position of the distal breakpoint is proximal to the *INS* gene: this locus had an identical (normal) position on both HSA11 homologues. c) The distal breakpoint is distal to the marker D11S470 as indicated by the more centromeric position of this locus on the inverted chromosome homologue, adjacent to 11p11.2 (white arrow). d and e) FISH analysis of MAA-fixed nuclei using a HSA11p paint (green) combined with probes to loci within 11p15.5 indicated that the interphase organisation of loci vary within CV581 nuclei (see Section 4.8). *INS/IGF2* was frequently positioned outside of the HSA11p territory, and in 56% of nuclei both alleles were external. However, this was true for D11S470 in only 20% of nuclei.

control FATO cell line, and also for *INS/IGF2* in the inverted cell line where the metaphase position of *INS/IGF2* was unaltered. In a cell line where the position of a locus differed between the two homologous chromosomes, such as D11S470 in CV581, variance might be greater if the organisation of the locus relative to the chromosome is affected by the rearrangement. This was found to be the case for D11S470 in the inverted cell line compared to the other three data sets ($P=0.045$, just significant at the 5 % level).

Variation in *INS/IGF2* position between homologous chromosome territories within nuclei of CV581 was not significantly greater than the variation in the control cell line. Therefore, the positioning of *INS/IGF2* significantly further outside of the chromosome territory of the CV581 inversion cell line compared to the FATO control cell line ($P=0.002$ using a standard t test), was common to both the inverted and the normal chromosome of CV581. This suggests that the differences in spatial organisation of *INS/IGF2* between CV581 and FATO were characteristic of the cell line rather than due to the inversion *per se*.

It is difficult to determine from this data what the cause of the variation in position of D11S470 between cell lines might be. There are many factors that must be taken into consideration, such as the gene density of the region to which D11S470 is juxtaposed in CV581. HSA11p11.2 is an R-band and detail on the gene richness around the *RAPSN* gene (distal of which the break-point is positioned) from NCBI ([http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/maps.cgi?org=hum&chr=11&MAPS=gene,est,loc\[46844840.02:46886201.33\]&query=uid\(5684\)&QSTR=rapsn](http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/maps.cgi?org=hum&chr=11&MAPS=gene,est,loc[46844840.02:46886201.33]&query=uid(5684)&QSTR=rapsn)) indicate that this region is gene-rich containing approximately 13 genes per Mb (also includes predicted genes). This gene density is similar to that of 11p15.5, an interesting fact in light of the observation that extrusion of chromatin containing *INS/IGF2* still occurs to a comparable extent from both HSA11p chromosome territories in CV581. In addition, the effect of position along the chromosome may have an effect, as I have shown that all telomeres and centromere sequences may be preferentially positioned at the surface of a chromosome domain.

A refinement to this experiment would be to identify the inverted chromosome homologue at interphase and analyse the relative positions of both D11S470 loci within the same nuclei. This would be possible if, instead of counterstaining the nuclei using DAPI, a cosmid or BAC probe spanning the inversion breakpoint was incorporated into the FISH probe mixture and detected using a blue fluorochrome (I do not have access to facilities allowing detection of more than three colours simultaneously). Therefore, the inverted chromosome would be

identifiable as the territory containing two blue locus signals, rather than the single signal in the normal homologue. Thus, an internal control to the experiment would be provided and comparison between cell lines would not be necessary. The fate of 11p12 material juxtaposed to 11p15.5 loci should also be characterised.

4.9 THE LOCALISATION OF SEQUENCES OUTSIDE OF CHROMOSOME DOMAINS IS NOT CHARACTERISTIC OF IMPRINTED REGIONS

11p15.5 contains genes that are subject to imprinting. The mechanistic basis of imprinting is yet to be fully defined, but aspects of higher order chromatin structure are known to be important (Section 1.3). Furthermore, in human T lymphocytes, a temporal and spatial association between chromosomes 15 at the Prader-Willi Syndrome and Angelman Syndrome (PWS/AS) locus at 15q11-q13, have been observed during late S phase (LaSalle and Lalande, 1996). Therefore, the conformation within the nucleus adopted by chromatin from 11p15.5 could be linked to the imprinted state of genes within this region. From data presented in Table 4.2, this looked unlikely because both alleles were found in the same conformation in 41% of nuclei. However, I decided to investigate the organisation of the PWS/AS locus on 15q11-13 (Figure 1.1). The most common cause of disruption of imprinting at this locus is a large chromosomal deletion; all deletions in PWS are paternal in origin and all deletions in AS are maternal in origin (Glenn *et al.*, 1997). 14 genes have been identified in a region spanning approximately 2.5 Mb of DNA (5.6 genes/Mb), and at least 7 of these are subject to imprinting (Gabriel *et al.*, 1998). The critical imprinting centre has also been identified (Figure 1.1).

I obtained two BAC probes from 15q12 (80H14 and 88P10, BACPAC Resources, <http://www.chori.org/bacpac/>) and used them in FISH on MAA-fixed lymphoblastoid nuclei in combination with a HSA15q paint (gift of Michael Bittner, Section 2.4.1) (Figure 4.8). The mean distance of each probe from the nearest chromosome edge is displayed in Table 4.7.

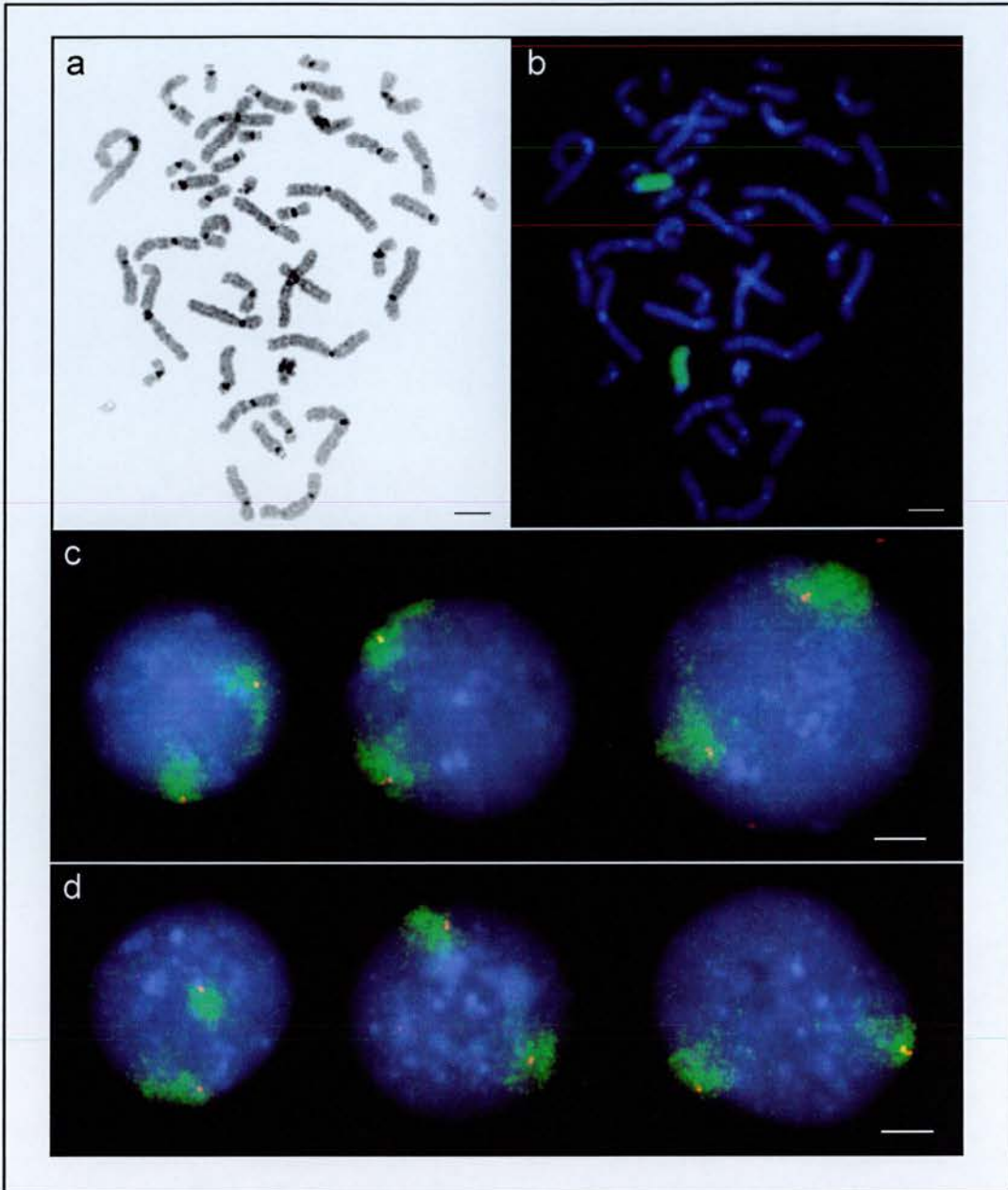


Figure 4.8 FISH analysis of the organisation of the imprinted gene cluster on HSA15q11-13 in lymphoblastoid nuclei .

a) Grey-scale representation of a metaphase chromosome spread from lymphoblastoid cells.
 b) The specificity of the HSA15q paint (green) is illustrated on the same chromosome spread counterstained with DAPI. Nuclei counterstained with DAPI, hybridised with HSA15q paint (green) in combination with BAC probes 80H14 (c) and 88P10 (d) (red). Across (c) and (d) cell cycle progression is illustrated by increasing nuclear size and the replication of the BAC probes (indicated by the presence of one or two red probe signals).
 Scale bars represent 5 μm .

Table 4.7 Mean positions of sequences from imprinted regions of the genome within their respective chromosome domains.

Fifty images were captured of nuclei after FISH using a biotin-labelled HSA15q paint in combination with probes mapped to 15q12. The scripts described in Boxes 3.1 and 3.2 were used to analyse the position of the probe signal in relation to the chromosome arm domain. Mean positions for the position of the probe relative to the chromosome edge were calculated \pm SEM (Section 2.10).

Probe from imprinted region	Mean normalised position of signal relative to chromosome domain	Mean distance of probe signal from edge of chromosome (μm)
11p15.5 (<i>INS/IGF2</i>)	$-0.324^{\pm 0.069}$	$-0.605^{\pm 0.129}$
15q12 (88P10)	$0.152^{\pm 0.026}$	$0.284^{\pm 0.049}$
15q12 (80H14)	$0.186^{\pm 0.032}$	$0.347^{\pm 0.060}$
Mean WAGR (11p13)	$0.315^{\pm 0.010}$	$0.588^{\pm 0.019}$

The mean positions of the two BAC probes from 15q12 within the 15q domain are statistically similar ($P=0.407$), as expected for two probes mapped to the same locus and they are located well within the 15q territory. Therefore, chromatin from the PWS-AS associated region on HSA15q11-13 is not subject to the same unusual organisation as HSA11p15.5. I conclude that the distinct organisation of 11p15.5 was unlikely to be a configuration adopted due to constraints imposed in achieving correct imprinting of genes in this region.

4.10 IS THERE A CONNECTION BETWEEN GENE-DENSITY AND THE POSITION ADOPTED BY CHROMATIN RELATIVE TO THE HSA11p TERRITORY?

11p15.5 is very gene-rich, containing at least 47 genes over approximately 4.5Mb (10.4 genes per Mb). 11p13 is a moderately gene-rich locus containing 4 genes within 1Mb of DNA (4.0 genes per Mb) and 15q12 has an intermediate gene density, with at least 14 genes spanning approximately 2.5Mb (5.6 genes per Mb). I have presented evidence discounting incomplete coverage of chromosome paints or a position effect as the cause for the positioning of probes mapped to 11p15.5 at a distance from the HSA11p territory. Therefore, data described above prompts speculation that there is a connection between the gene-richness of a particular region of DNA and the position adopted by these sequences relative to chromosome domains of the interphase nucleus. A high gene density may dictate a more peripheral positioning of sequences relative to an interphase chromosome domain.

To test this idea, I extended my use of HSA11p as a model system to determine patterns of organisation within an interphase chromosome to cytogenetic band 11p14. 11p14 is the cytogenetic band immediately distal to 11p13 (Figure 4.1) and is an extreme example of a gene-poor G-band. Only four genes are contained within approximately 8 Mb (Fantes *et al.*, 1995), therefore 11p14 includes large stretches of DNA where there are no genes. If there is a connection between gene density and the position adopted by DNA within an interphase chromosome, I would expect 11p14 to localise within the interior of HSA11p. I selected a cosmid mapped to cytogenetic marker D11S865 within 11p14 (cSRL7d7) (Fantes *et al.*, 1995) and determined its position within HSA11p in a lymphoblastoid cell line, a primary fibroblast cell line and a lens epithelium cell line in 2D (Table 4.8).

Table 4.8 Mean position of sequences from 11p14 relative to WAGR (11p13) in lymphoblastoid cells

A labelled cosmid mapped to D11S865 within 11p14 was used in FISH analysis of MAA fixed lymphoblastoid cell nuclei in combination with a HSA11p paint. Fifty images were captured for each cosmid and the scripts described in Chapter 3 (Boxes 3.1 and 3.2) were used to calculate the mean position of the locus within the interphase chromosome 11p domain from the data output (section 2.10).

Gene/locus probe (chromosomal location)	Lymphoblastoid cell line		Lens epithelium		Primary fibroblast cell line	
	Mean probe location as proportion of chromosome radius	Mean distance of probe from edge of chromosome (μm)	Mean probe location as proportion of chromosome radius	Mean distance of probe from edge of chromosome (μm)	Mean probe location as proportion of chromosome radius	Mean distance of probe from edge of chromosome (μm)
D11S85 (11p14)	0.342 \pm 0.025	0.636 \pm 0.047	0.419 \pm 0.017	0.660 \pm 0.027	0.386 \pm 0.020	0.700 \pm 0.036
Mean WAGR (11p13)	0.315 \pm 0.010	0.588 \pm 0.019	0.285 \pm 0.010	0.449 \pm 0.016	0.282 \pm 0.010	0.512 \pm 0.018

Using standard t tests I compared the mean position of D11S865 with that of a mean value describing the overall position of WAGR (Chapter 3) in each cell line. 11p14 was significantly more internal within HSA11p compared to the mean WAGR position in two dimensional analysis of a lens epithelium and a primary fibroblast cell line ($P < 0.001$ in each case). However, in a lymphoblastoid cell line the more internal localisation of 11p14 compared to 11p13 was not statistically significant ($P = 0.320$). The implications of this are not clear, but differences indicate that specific intra-chromosomal organisation varies between differentiated cell lines, and also between differentiated cell lines compared to primary cell lines.

4.11 THE ORGANISATION OF THE MURINE REGION OF SYNTENY TO 11p15.5 WITHIN MMU7

The murine region of conserved synteny to the BWS locus 11p15.5 (*NAP2* to *RL23mrp* centromere to telomere, Figure 4.2) on MMU7 is very closely conserved to the human sequence, but the entire region is inverted with respect to the telomere. Two BAC probes mapped within this region of MMU7 were obtained (gifts of S. Engemann, Berlin). BAC9 contains the 5' part of the *Kcnq1* gene (Paulsen *et al.*, 1998) and BAC 5 covers the region from *Obph1* to *Tssc5* (Engemann *et al.*, 2000) (Figure 4.2). The BACs were used in combination with a MMU7 chromosome paint (gift of I. Solovei) in 2D FISH on MAA-fixed ES cell nuclei (Figure 4.9). The mean position of each probe relative to the nearest edge of the MMU2 chromosome domain, normalised for the chromosome radii, was calculated. Positions were found to be $-0.202^{\pm 0.103}$ μm outside of the MMU7 territory for BAC5 and $0.184^{\pm 0.067}$ μm from the chromosome edge (within the territory) for BAC9. These probes are therefore more external than the mean murine WAGR locus, which was $0.871^{\pm 0.024}$ μm from the nearest MMU2 territory edge.

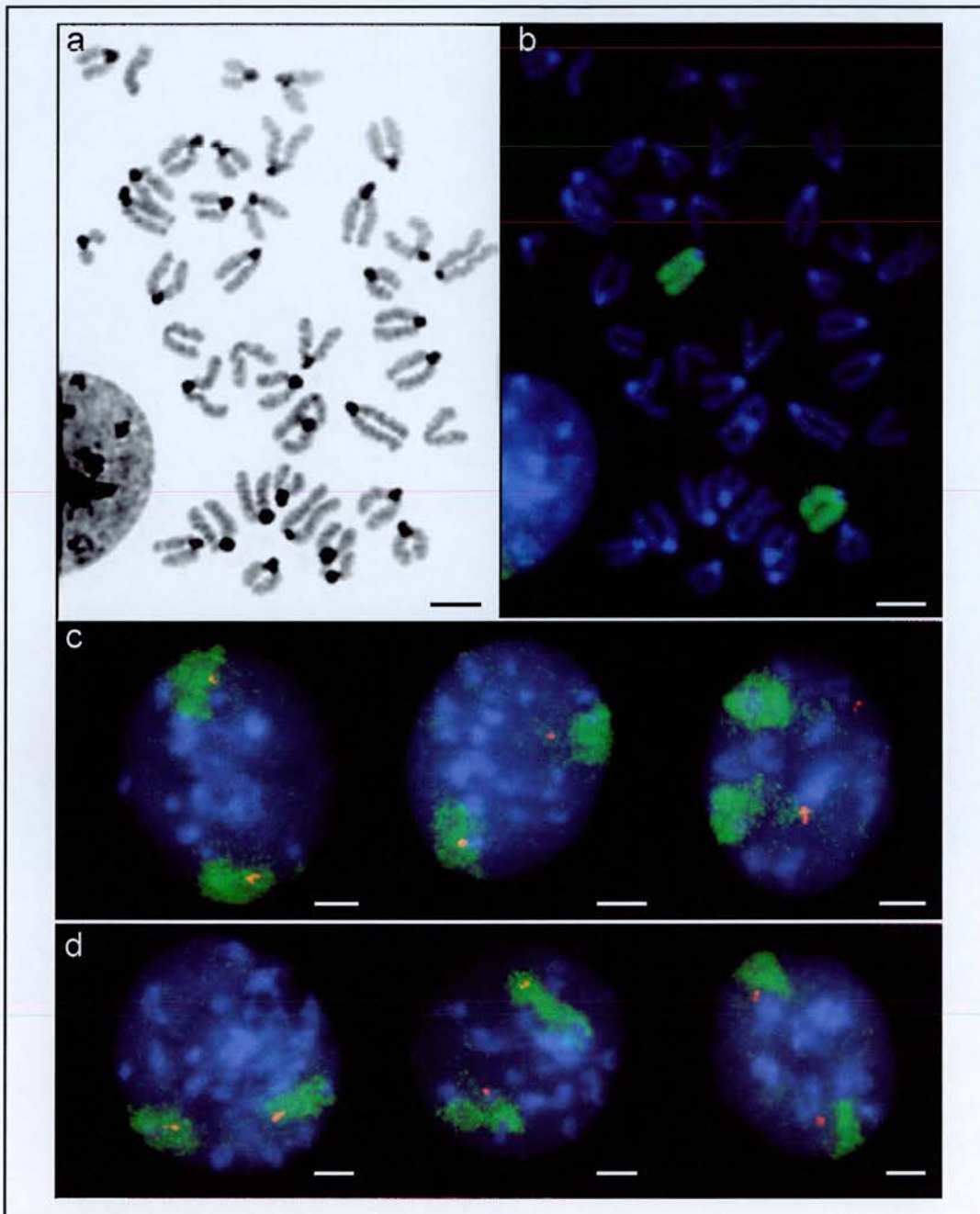


Figure 4.9 Analysis of the organisation of the murine BWS locus within MMU7 in embryonic stem cell nuclei.

a) Grey-scale representation of a metaphase chromosome spread from mouse embryonic stem cells.

b) The specificity of the MMU7 paint (green) is illustrated on the same chromosome spread counterstained with DAPI.

c and d) MAA-fixed nuclei were hybridised with biotin-labelled MMU7 paint (green) and dig-labelled BAC probes mapped to the BWS locus (red). Nuclei illustrating different configurations of BAC 9 (c) and BAC 5 (d) are shown. As apparent in the right hand nucleus of panels c and d, probes frequently localise outside of the visible MMU7 domain. Scale bars represent 5 μm.

Both probes were therefore positioned near the surface of the chromosome with a significantly different position to the murine WAGR locus ($P < 0.001$), which had a mean position $0.872^{\pm 0.024}$ μm from the chromosome edge. This difference is obvious when the distributions of loci are compared (Figure 4.10). The MMU2 WAGR loci were rarely positioned outside of the chromosome domain, but, like the syntenic human region in 11p15.5, MMU7 loci were frequently located at a distance from the MMU7 chromosome territory. This suggested that features of the organisation of a chromosome domain are conserved between mammalian species, and must therefore be functionally important.

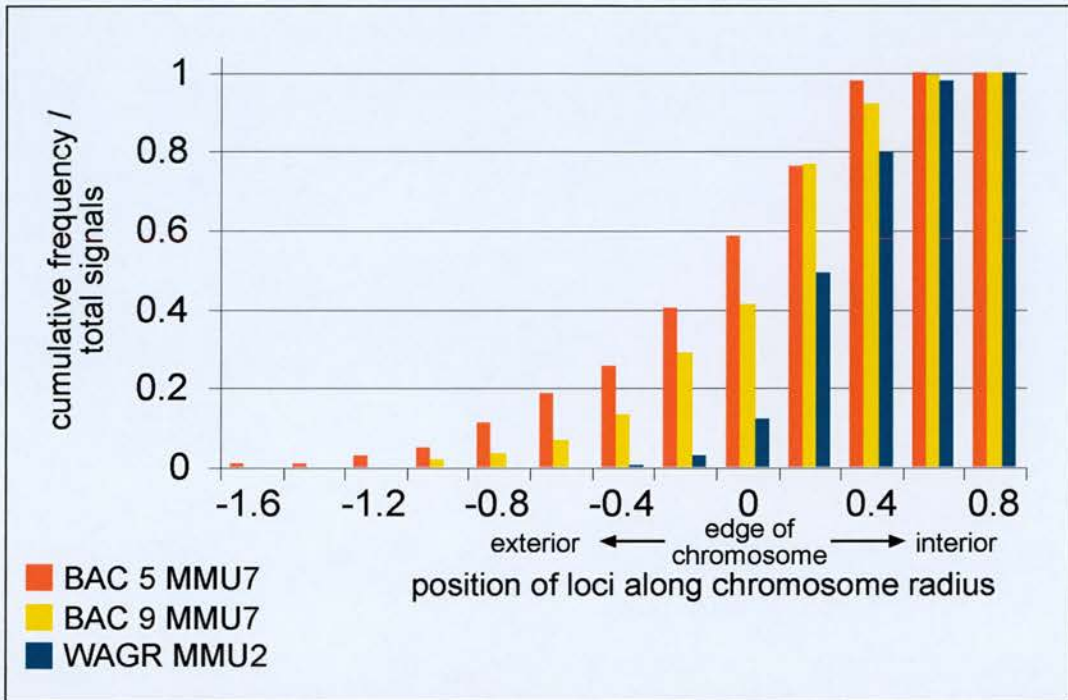


Figure 4.10 Comparing the distribution of loci relative to MMU2 and MMU7 chromosomes in ES cell nuclei.

The cumulative distributions along of loci from the WAGR locus on MMU2 (Chapter 3) and the BWS homology region on MMU7 are compared. Note that loci from WAGR were rarely positioned outside of the MMU2 territory. In contrast, loci from MMU7 were frequently outside of MMU7 and the distribution of the more distal locus, represented by BAC 5, appears to be more external to the chromosome domain than proximal BAC 9.

4.12 DISCUSSION

In this Chapter I have examined the organisation of more than 9Mb of DNA spanning HSA11p from the telomere into cytogenetic band 11p15.3. Part of this region, up to cytogenetic markers D11S682 and D11S12 in 11p15.5 (approximately 4Mb from the most telomeric mapped cytogenetic marker in 11p15.5, D11S483), appeared to be extruded from the main body of the chromosome territory when the mean positions of loci relative to the nearest chromosome edge, or the % signals outside of the chromosome domain, were considered. I suggested possible explanations for the unusual organisation of this region within human interphase nuclei and addressed each in turn. In addition, I have compared the organisation of the conserved syntenic region in the mouse.

4.12.1 The unusual configuration of chromatin from 11p15.5 is not just a position effect

The localisation of sub-telomeric sequences from 11p15.5 apparently at a distance from the HSA11p domain could be simply explained if the chromosome-specific paint used to define the chromosome territory did not contain sequences corresponding to the end of the chromosome. I have presented evidence showing that this was not the case, and that fluorescence signal completely coated the metaphase chromosome, right up to the telomere (Figure 4.6). Therefore, loci apparently at a distance from the chromosome territory must be located on a stretch of chromatin that cannot be visualised by FISH and light microscopy. In addition, a probe specific to alphoid DNA from the centromere of HSA11 had a mean position at the surface of the interphase chromosome domain, even though alphoid sequences are not represented in the chromosome paint and alphoid sequences are beyond the HSA11p paint coverage of a metaphase chromosome.

There has been speculation about the organisation of telomere sequences within the human interphase nucleus. Data so far indicates that telomeres have a dispersed nuclear distribution, with a lower density in the outer 25% of nuclear volume than expected for a random distribution (Luderus *et al.*, 1996). The apparent positioning of the HSA11p telomere at a significant distance from the main body of the chromosome domain could have been a configuration common to all human telomeres. I have presented evidence that telomeres from other chromosome arms have a mean position at the surface of their respective

chromosome domains and are not frequently positioned at a distance. Similarly, HSA11 alphoid sequences, representing proximal HSA11p, had a mean position at the surface of the territory, providing evidence that the extended configuration of chromatin from 11p15.5 is specific to this cytogenetic band and not common to all chromosome arm termini. However, my observation that an inversion of HSA11p influences territorial organisation and intranuclear position of chromatin demands further analysis.

4.12.2 The unusual configuration of chromatin from 11p15.5 is not common to other imprinted loci

In the human genome there are two main clusters of genes subject to imprinting. One, the BWS region is located in 11p15.5. I suggested that the unusual configuration of 11p15.5 in the interphase nucleus could be characteristic of an imprinted region of the genome and that this configuration is somehow required to achieve the epigenetic mark facilitating monoallelic expression of genes. A second cluster of imprinted genes, the PWS/AS region is located on chromosome 15q12. If extrusion of genes from the main chromosome domain in interphase was characteristic of an imprinted region, chromatin from 15q12 should have exhibited a similar configuration. My data suggested that this was not the case: loci from 15q12 had a mean position well within the HSA15q chromosome domain (Table 4.8), significantly more internal than the imprinted *INS* gene from 11p15.5.

4.12.3 Organisation of chromatin within an interphase chromosome may be influenced by gene density

Lastly, the unusual organisation of 11p15.5 could have been attributed to the fact that this region is extremely gene rich. In the sub-telomeric 4.5 Mb of sequence from HSA11p, more than 45 genes have so far been mapped and finished sequence with novel and predicted genes accurately annotated will probably increase this density further. If more gene-dense regions localise towards the periphery, or at a distance from the surface of a chromosome domain, this suggests that gene-poor DNA might localise more towards the interior of a chromosome territory. This prediction was implied by data from a previous study (Kurz *et al.*, 1996), where a non-coding sequence was consistently found to be more internally positioned within an interphase chromosome territory as compared to gene loci.

To test this idea I examined the position of a locus from the gene-poor cytogenetic G-band 11p14, which contains only approximately 4 genes over 8Mb DNA (Fantes *et al.*, 1995). 2D data from two cell lines, a lens epithelium transformed cell line (CD5a) and a primary fibroblast cell line (1HD) suggested that the gene-poor locus was indeed more internal within HSA11p than loci from both 11p13 and 11p15.5. However, in a lymphoblastoid cell line (FATO), 11p14 was not located significantly more internally within HSA11p than the WAGR locus from 11p13. This data was therefore not conclusive, but it is perhaps important to look at the wider context of the locus. WAGR is located in distal 11p13, on the border between a moderately gene dense region and a gene poor region (Gawin *et al.*, 1999). Within the 4 Mb of sequence proximal to WAGR there are at least 13 genes and 20 ESTs (an average of 8 genes per megabase). However, the WAGR locus is only approximately 550kb from the start of 11p14 which has an average of 0.5 genes per megabase (Gawin *et al.*, 1999; Fantes *et al.*, 1995). In contrast, 11p15 contains at least 47 genes over 4.5 Mb (more than 10 genes per Mb). Further evidence that increasing gene density may result in a more peripheral positioning of chromatin within a chromosome territory came from comparing data from analysis of 15q12. 15q12 has a gene density intermediate to that of 11p13 and 11p15.5 and correspondingly has an intermediate position within a chromosome domain with respect to distance from the chromosome territory surface.

It is interesting to note here that two of the loci included in my analysis of HSA11p actually featured in a previous analysis of chromosome territory organisation (Kurz *et al.*, 1996). In the previous study, the *HBB* gene (represented by pLCR in my analysis) was found at the periphery of the HSA11 territory, and an anonymous, non-coding sequence from 11p14 was found randomly positioned within the HSA11 territory, or preferentially positioned within the territory interior, in different cell types. In the analysis of Kurz *et al.*, the 'periphery' of a chromosome territory actually represented between 40 and 50% of the territory volume, therefore my analysis, which found pLCR to be positioned 0.160 of a HSA11p radius from the edge, is consistent with this data. Similarly, I found a locus from 11p14 to be consistently more interior within the chromosome territory than the pLCR locus in all three cell lines examined ($p < 0.001$). Therefore, data presented here support the conclusions of Kurz and co-workers that sequences from 11p15 and 11p14 are differentially positioned within the HSA11 chromosome territory. However, I speculate that this difference is due to the contrasting gene densities of these two regions, rather than whether the probe analysed, represented a gene or a non-transcribed sequence.

4.12.4 The organisation of the syntenic region in the mouse

The differences in the organisation of gene-poor/gene-rich regions in the human were common to syntenic regions of conservation in the mouse. The syntenic region in the mouse to the entire HSA11p15 region is located on MMU7 (<http://www.ncbi.nlm.nih.gov/Omim/Homology/mouse7.html>), but is split into two sections in bands B2-B4 and F4-F5. In this analysis of the organisation of murine MMU7, I considered only the BWS-associated region in MMU7F4-F5, which has been extensively studied and is closely conserved to the human in terms of gene order, although the entire locus is inverted with respect to the telomere (Onyango *et al.*, 2000). The murine BWS locus contains 22 genes in 1.1Mb (Engemann *et al.*, 2000) spanning from *Obph1* near the telomere to *Rpl231*, similar to the gene density for the corresponding human sequence. Mouse sequencing projects have not yet extended far beyond the BWS region and it remains to be seen how comparable over large distances the gene density of this region in the mouse is to HSA11p15. It has previously been inferred that the distribution of genes across the murine karyotype is much more even than that of the human (Cross *et al.*, 1997). This conclusion was drawn from CpG island mapping where no individual mouse autosome stood out as being either very island-rich or very island-poor. Intrachromosomal variation in CpG island distribution was also less dramatic than that seen in the human. However, the gene density of this region in the mouse is high at 20 genes per Mb (Engemann *et al.*, 2000), and it is intriguing that there is an incidence of chromatin decondensation away from the chromosome territory at this locus, albeit to a smaller degree than at the human locus. This perhaps indicates that there is no need for such an extreme decondensation of chromatin from surface of a chromosome domain in the mouse, due to the more even distribution of genes in the genome.

4.12.5 Is interplay between gene density and transcription activity important in organising chromosome territories?

If gene density influences the interphase position and perhaps degree of condensation, of chromatin relative to a human chromosome domain, the distance over which these influences take effect is open to debate. In addition, the driving force behind the localisation of a gene-rich region at a distance from a chromosome domain remains to be determined. Is it purely the number of genes in a stretch of DNA or is the number of genes that are actively transcribed an important consideration? In the previous Chapter I have showed that changes

in transcription status do not necessarily alter the position of genes within a chromosome domain (Chapter 3). Therefore, number of genes in a region of DNA sequence may dictate its organisation relative to an interphase chromosome domain, rather than the number of genes that are actively transcribed *per se*.

A recent study examined the large-scale organisation of approximately 4Mb of DNA from the MHC locus on chromosome HSA6p (Volpi *et al.*, 2000). This region is extremely gene rich, containing an average of 62.5 genes per megabase (including pseudogenes) (The MHC Sequencing Consortium, 1999), however, there are distinct regional variations across the locus. In this study, large chromatin loops containing several megabases of DNA corresponding to the MHC locus (HSA6p21) were observed extending outwards from the surface of the domains defined by a chromosome HSA6 or 6p paint. The frequency with which loci from this region were observed on an external chromatin loop was cell type dependent and appeared to be related to the transcriptional status of MHCII genes. Up-regulation of transcription by IFN- γ increased the frequency with which the MHC gene cluster was found on an external loop.

This high gene density of the MHC locus can, in part, be explained by the fact that the region contains many ‘families’ of genes clustered together, which arise from duplications of one original gene and have similar functions. The EDC locus on HSA1q21 also consists of families of genes with the same function and therefore the same regulation kinetics (referred to in Section 3.8) and is also subject to a similar ‘chromatin looping’ from the surface of HSA1 on transcription up-regulation (R. Williams, personal communication). Therefore, it appears that both gene density and large-scale alteration of transcriptional activity genes may both play a role in the spatial organisation of chromatin.

It is important to note here that the MHC locus was found to adopt a looping structure, and sequences either side of the MHC genes appeared to localise within the chromosome 6 domain. 11p15.5 however, appears to be extruded as a tail of DNA, the limit of which is the telomere. Distal sequences do not loop back into the chromosome 11p domain.

4.12.6 Extension of analysis to three dimensions is required

Observations reported in this Chapter and in Chapter 3 imply that gene density may be important in determining the organisation of chromatin within an interphase chromosome

territory. From 2D analysis I was unable to draw conclusions as to the absolute position of genes within the HSA11p domain, due to the flattening of specimens. To address directly the issue of whether all genes are at the surface of a chromosome domain as previously suggested (Zirbel *et al.*, 1993; Kurz *et al.*, 1996), analysis of three-dimensionally preserved nuclei is required. Extension of this quantitative analysis to 3D will enable precise definition of the position of genes within chromosome domains. I proposed to develop an analysis of three-dimensionally preserved nuclei, and this forms the subject of Chapter 5.

CHAPTER 5

THREE DIMENSIONAL ANALYSIS OF THE ORGANISATION OF HSA11p

5.1 INTRODUCTION

Analysing the spatial relationships of nuclear entities is technically demanding. The procedure for the visualisation of nuclear components should cause as little damage to the nuclei as possible and not disrupt the relative position of the components. The demands of preserving nuclear architecture yet allowing hybridisation to denatured DNA are conflicting. However, analysis indicates that the FISH protocol has no major effect on the structure of interphase chromosome territories at the light microscopy resolution level. It has been shown that in cells permeabilised and then processed for FISH (Section 2.6), nuclear structures e.g. centromeres (Cremer *et al.*, 1993; Kurz *et al.*, 1996; Verschure *et al.*, 1999), nuclear speckles (Zirbel *et al.*, 1993) and PML bodies (Verschure *et al.*, 1999) remain in the same spatial positions before and after processing. The spatial distribution of acetylated histones in interphase nuclei has also been used as a sensitive marker to show that, at the level of light microscopy, the FISH procedure does not alter the distribution of acetylated histone H4 in fibroblast nuclei (Hendzel and Bazett-Jones, 1997; Verschure *et al.*, 1999).

Processing of two-dimensional images is efficient, as large numbers of nuclei can be stored and analysed (Croft *et al.*, 1999; Volpi *et al.*, 2000; Boyle *et al.*, 2001). Extrapolation of data from 2D preparations assumes that the relative organisation of the intact nucleus is not significantly perturbed. However, signals that appear to be in the centre of flattened chromosome domains in two-dimensional analysis might, in reality, be located close to the periphery on either the top or bottom surface of the territory. Plane projections can be used to make inferences about three-dimensional objects (Carothers, 2000) and, in previous studies, conclusions drawn from 2D analysis have been confirmed in 3D (Croft *et al.*, 1999; Boyle *et al.*, 2001). In a third study, which considered the organisation of a single chromosome

domain, the trend of organisation was similar in 2D as compared to 3D, however, the chromosome domains appeared to be more decondensed in 2D MAA-fixed nuclei than in 3D pFA-fixed nuclei (Volpi *et al.*, 2000). A comparison of data obtained in nuclei prepared using different fixation techniques found formaldehyde fixation to produce images that most closely resembled those observed in living cells (Tsukamoto *et al.*, 2000).

Therefore, I decided it was important to assess the organisation of the HSA11p chromosome domain in three-dimensionally preserved pFA-fixed nuclei to confirm patterns of organisation suggested by 2D analysis of MAA-fixed nuclei. This would also enable me to comment on the localisation of genes and intergenic sequences in relation to the edge of the chromosome domain without having to consider the problems introduced by use of flattened specimens.

5.2 DEVELOPMENT OF A 3D ANALYSIS TECHNIQUE

FISH was performed on three-dimensionally preserved nuclei (Section 2.6) using the same probes described in Chapters 3 and 4. Biotin-labelled chromosome arm paints were detected using FITC-conjugated antibodies and dig-labelled locus probes were detected with TR-conjugated antibodies. Nuclei were counterstained using DAPI. When considering 3D analysis, the optical detection system should allow resolution of different nuclear signals in 3D in order to reveal their spatial relationships. I used a light microscope fitted with a motorised stage to capture 3D stacks of images at 0.5 μm intervals (Section 2.9.1). The three colour image stacks captured contained approximately 20 image planes, so as to include the whole of the nucleus. The average diameter of a fibroblast nucleus in 2D analysis was $23.1^{\pm 0.2} \mu\text{m}$ ($\pm\text{SEM}$) ($n=1800$), however, fibroblast nuclei have an ellipsoid shape and are compressed along the z -axis when mounted under a coverslip. Therefore, stacks containing 25-30 image planes at 0.5 μm intervals included the whole nucleus.

The segmentation of each 3D chromosome and locus-signal domain was performed manually using the program MAPaint, developed at the MRC Human Genetics Unit for 3D-image segmentation (Section 2.9.1) (Figure 5.1). Files corresponding to matching chromosome and locus domain pairs were then passed to a program, (written by R. Baldock, MRC Human Genetics Unit) that calculated the volume of the chromosome signal domain

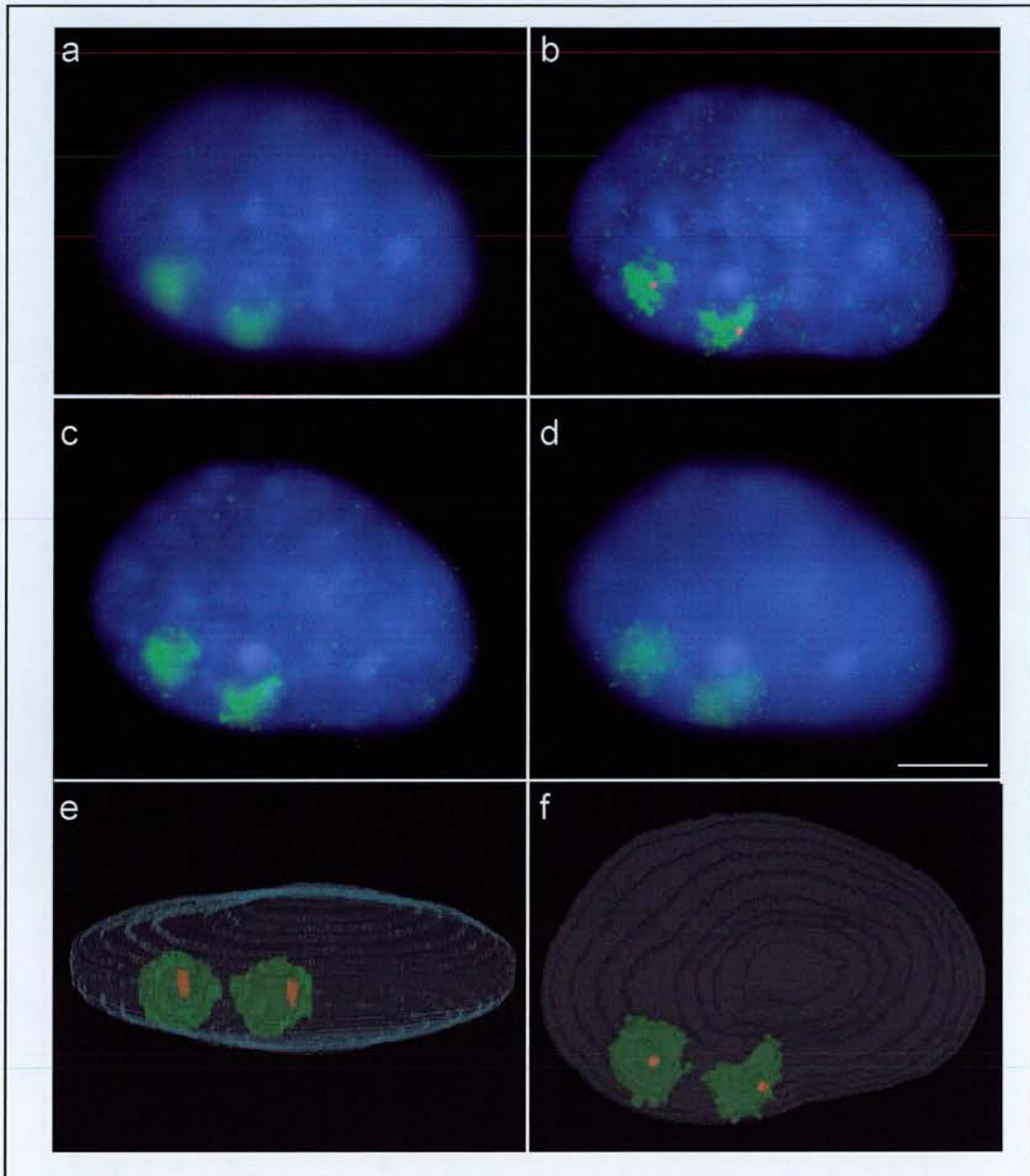


Figure 5.1 Reconstructing 3D fibroblast nuclei from image plane stacks.

Stacks of image planes at $0.5\ \mu\text{m}$ intervals were captured through 3D-preserved nuclei after FISH using a chromosome paint and a locus probe (section 2.9.1). Illustrated here is HSA11p (green) and the *RCN* gene (red) within a fibroblast nucleus counterstained with DAPI (blue). The planes shown are $1\ \mu\text{m}$ apart along the z -axis (a-d). The scale bar represents $5\ \mu\text{m}$. Image stacks were imported into MAPaint (Mouse Atlas Group, MRC Human Genetics Unit) and domains of interest 'painted' separately in each colour stack (section 2.9.1). Nucleus reconstruction was achieved using the AVS/Express visualisation system (section 2.9.3) and MPEG movie files rotating nuclei about the x -axis were created from the resultant images. Two such movie stills are illustrated here (e and f).

and the distance of the centre of mass of the locus signal domain to the nearest edge of that chromosome domain (Section 2.9.2). Thus, the radius of the chromosome territory could be calculated from the volume (Section 2.10) enabling normalisation of data in a manner analogous to the analysis procedure used for 2D images. Again, data outputs representing loci apparently outside of the chromosome territory were assigned negative values.

In 2D analysis, a degree of interpretation of data is required due to the flattening of nuclei. However, statistically significant conclusions could be drawn from this data as large numbers of nuclei were examined: fifty nuclei were analysed in 2D for each locus and data for different loci were compared within the 2D analysis employed. Preservation of the architecture of the nucleus is maintained to a greater extent in 3D than in 2D, but due to the time-consuming nature of the 3D-analysis technique developed, only 15-20 nuclei were considered for each locus and chromosome paint combination.

5.3 THE ORGANISATION OF WAGR IN 3D PRESERVED NUCLEI CONFIRMS 2D DATA

A value describing the mean probe position within the HSA11p chromosome domain along a calculated territory radius, \pm standard error of the mean (SEM), was calculated (Section 2.10) for each of the seven cosmids spanning the human WAGR locus (Figure 3.2) in 3D primary fibroblast nuclei. The corresponding distance to the chromosome edge in microns (calculated from the mean chromosome radius, Table 7.2) was also measured (Table 5.1, Figure 5.2).

Table 5.1 The position of sequences from WAGR relative to the HSA11p chromosome periphery of primary fibroblasts in 2D and 3D

Sets of fifteen to eighteen 3D image stacks were processed to delineate chromosome and locus domains (Section 2.9.1) and the position of the locus assessed by script analysis (section 2.9.2). Mean values were calculated for the normalised position of loci relative to the nearest edge of the chromosome domain, \pm SEM (Section 2.10). Equivalent 2D data is presented for comparison (Section 3.5.2).

Gene/ locus from WAGR	2D analysis		3D analysis	
	Mean locus position as proportion of chromosome radius	Mean distance of probe from edge of chromosome (μ m)	Mean locus position as proportion of chromosome radius	Mean distance of probe from edge of chromosome (μ m)
WT1	0.335 \pm 0.019	0.608 \pm 0.034	0.282 \pm 0.021	0.390 \pm 0.029
RCN	0.336 \pm 0.018	0.610 \pm 0.033	0.320 \pm 0.024	0.442 \pm 0.033
D11S324	0.240 \pm 0.030	0.436 \pm 0.054	0.269 \pm 0.021	0.372 \pm 0.029
D11S323	0.317 \pm 0.019	0.576 \pm 0.034	0.258 \pm 0.021	0.357 \pm 0.029
PAX6	0.291 \pm 0.023	0.528 \pm 0.042	0.290 \pm 0.029	0.401 \pm 0.040
PAXNEB 3'	0.289 \pm 0.024	0.525 \pm 0.044	0.217 \pm 0.021	0.300 \pm 0.029
PAXNEB 5'	0.343 \pm 0.020	0.623 \pm 0.036	0.335 \pm 0.030	0.463 \pm 0.041
Mean WAGR	0.282 \pm 0.010	0.512 \pm 0.018	0.282 \pm 0.010	0.390 \pm 0.014

Using analysis of variance as described in Section 2.10, the probability for all loci belonging in the same data population was 0.037 ($P=0.037$). This was marginally significant and on examination of Figure 5.2, the variation across the locus appeared to lie with probe cG0453. This probe represented the 3' end of the *PAXNEB* gene, and the adjacent probes mapped to the *PAX6* gene and the 5' end of the *PAXNEB* gene do not have such an outlying position. A regression line fitted to the raw data had a slope of only -0.0007; therefore the trend in variation across the locus was almost 0.

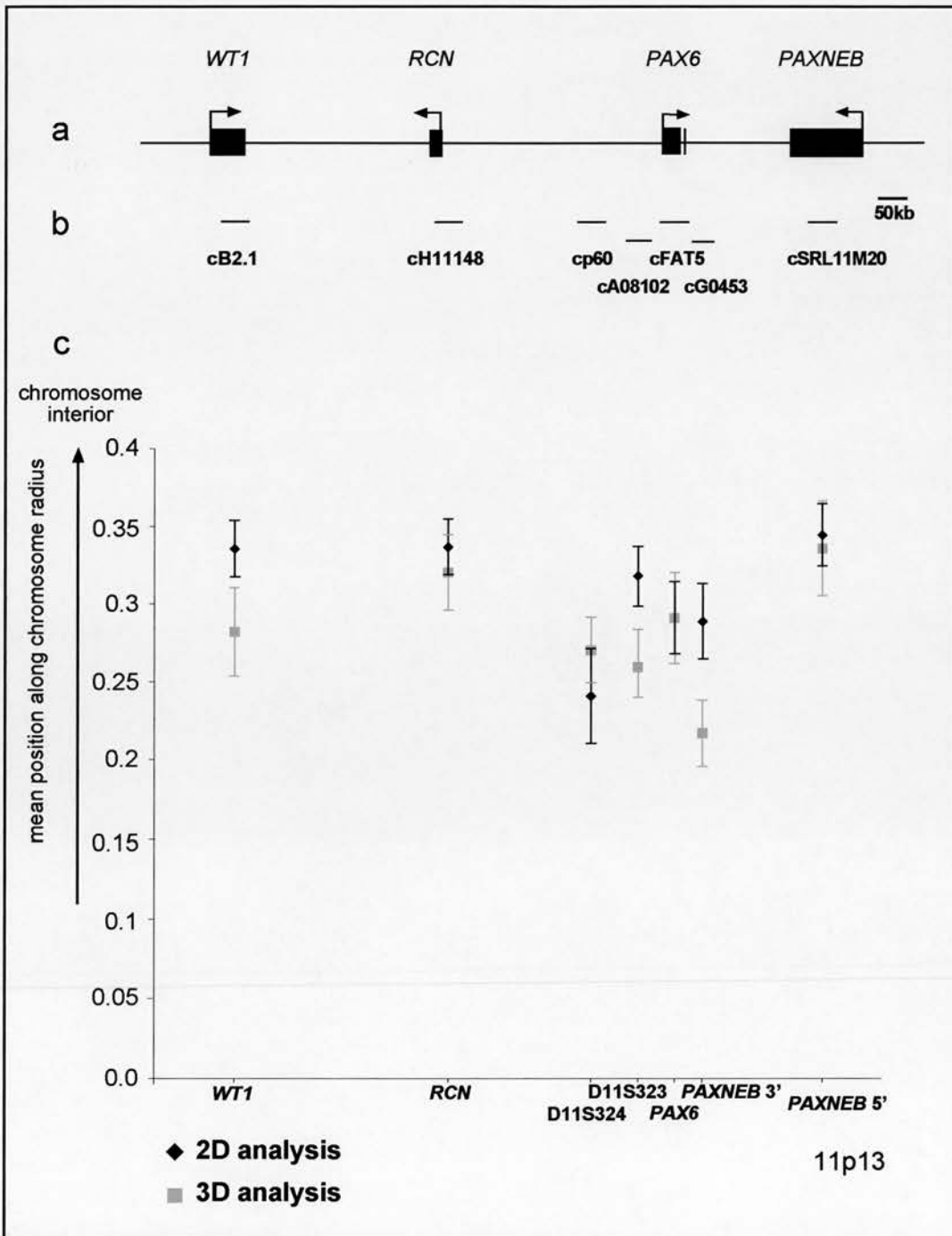


Figure 5.2 Comparing 2D and 3D analysis of the organisation of the WAGR locus within interphase HSA11p in primary fibroblast nuclei.

The linear organisation of the human WAGR locus is illustrated to scale (a) and the relative positions of the cosmids used in analysis of the organisation of this region within HSA11p are indicated (b). The mean normalised position of each probe along the radius of the interphase 11p domain is graphically depicted (c). Error bars indicate the standard error of the mean (SEM). A value of 0 represents the edge of the chromosome and 1.0 the chromosome centre. Labelled across the *x* axis are the names of the genes or intergenic cytogenetic markers to which the plots correspond.

Variation in distances relative to the nearest chromosome edge could be linked to sequence identity: whether probes represented gene or intergenic sequence, as previously suggested (Kurz *et al.*, 1996). The analysis of 3D preserved nuclei presented here generally supported conclusions drawn from 2D analysis (Chapter 3): there was no preference for gene sequences to be located more towards the periphery of the chromosome as compared to intergenic, non-transcribed sequences within the megabase of DNA studied. In addition transcription status of a gene did not affect its position within the interphase chromosome. Actively transcribed genes (*RCN* and *PAXNEB*) are not differentially positioned within the chromosome territory as compared to non-transcribed genes (*WT1* and *PAX6*) or intergenic sequence (D11S324 and D11S323).

A previous 3D study suggested that all gene sequences were located at the surface of their respective chromosome domains regardless of transcriptional status (Kurz *et al.*, 1996). As in 2D nuclei, WAGR was positioned approximately one third of the way along the HSA11p territory radius in 3D-preserved nuclei, where the flattening of nuclei is not a consideration. Therefore, 3D data has supported the idea that genes can be transcribed from deep within the territory interior, as suggested by 2D data (Chapter 3). The ubiquitously expressed *RCN* gene can be positioned deep within the HSA11p chromosome territory of primary fibroblasts, as visualised on 3D reconstruction of nuclei (Figure 5.1). This data shows conclusively that transcription does not just occur at the chromosome periphery.

While the mean values for the overall position of WAGR in 2D and 3D analysis were identical in terms of the normalised position along the chromosome radius ($0.282^{\pm 0.010}$), the size of the chromosome domain varied between analyses in terms of the length of the average territory radius in microns. In 2D the average HSA11p radius was $1.816^{\pm 0.017}$ μm while in 3D it was $1.382^{\pm 0.015}$ μm . This probably reflected the different fixation procedures used. MAA fixation dehydrates nuclei effectively to render proteins insoluble and some proteins, including histones, are stripped out of chromatin. This may serve to decondense chromatin, resulting in a more open structure, as I observed. pFA fixation, however, covalently crosslinks proteins, but is a reaction reversible using NaOH at an increased temperature. It has previously been noted that the two fixation procedures quantitatively produce slightly different results, but the trends observed were conserved between the two fixation procedures. Similarly, I found that the radius of the chromosome domain was larger within MAA-fixed nuclei than within 3D preserved pFA fixed nuclei, using the same

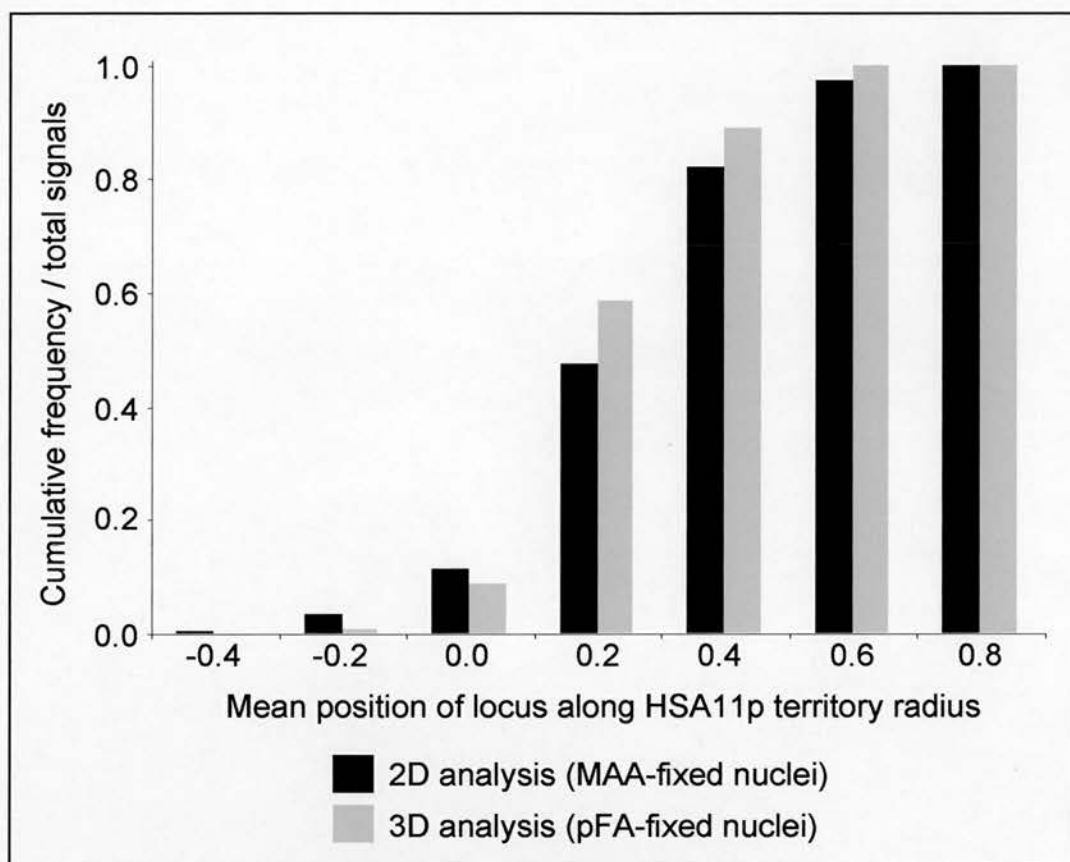


Figure 5.3 Comparing the distributions of signals from WAGR within HSA11p in 2D and 3D analysis of primary fibroblasts.

The cumulative frequency distributions of all WAGR locus signals along the radius of the HSA11p domain were calculated for 2D analysis of MAA-fixed and 3D analysis of pFA-fixed primary fibroblast nuclei. Frequencies are expressed as fractions of the total number of signals (therefore the frequency range is 0-1.0). Note the similar shape of the distributions for both analysis techniques. However, the spread of signals is slightly more peripheral in MAA fixed nuclei: chromatin appears more decondensed.

chromosome paint. Together these observations suggested that MAA fixation may distend chromatin within the nucleus more than pFA fixation.

To assess whether the trends in probe location within the HSA11p domain are the same between 2D and 3D I decided to compare the distributions of signals. In a similar analysis to that performed previously (Dietzel *et al.*, 1999), I plotted a cumulative distribution histogram for all WAGR locus signals in MAA-fixed and 3D-preserved fibroblasts (Figure 5.3). From this assessment of data it again appears that chromatin is decondensed by MAA fixation, but the distribution of loci relative to the chromosome territory is very similar between the two analyses.

5.4 THE UNUSUAL CONFIGURATION OF 11p15.5 IS APPARENT IN 3D BUT IS LESS PRONOUNCED

Analysis of the organisation of 11p15 was assessed from images produced after FISH on 3D-preserved nuclei, using a selection of the locus probes described in Chapter 4 (Section 4.4, Figure 4.2), in combination with a HSA11p paint. The nucleus, chromosome and signal domains were manually delineated (Section 2.9.1) (Figure 5.4). The 3D analysis script (Section 2.9.2) was used to assess the intra-chromosomal position of loci and a mean position was calculated for each locus (Table 5.2).

Table 5.2 The position of sequences from 11p15 relative to the HSA11p chromosome periphery in three dimensions

Sets of fifteen to eighteen 3D image stacks were processed to delineate chromosome and locus domains (Section 2.9.1) and the position of the locus assessed by script analysis (Section 2.9.2) Mean values were calculated for the normalised position of probes relative to the nearest edge of the chromosome domain \pm SEM (Section 2.10). The percentage of locus signals appearing outside the chromosome domains was also calculated for each locus. The identical data set from 2D analysis is also presented for comparison.

Gene/ locus from HSA11p	2D primary fibroblasts			3D primary fibroblasts		
	Mean locus position as proportion of chromosome radius	Mean distance of locus from edge of chromosome (μ m)	% signals within chromosome domain	Mean locus position as proportion of chromosome radius	Mean distance of locus from edge of chromosome (μ m)	% signals within chromosome domain
WAGR (p13)	0.282 \pm 0.010	0.512 \pm 0.018	89.3	0.282 \pm 0.010	0.390 \pm 0.014	98.6
D11S685 (p14)	0.386 \pm 0.020	0.700 \pm 0.036	96.0	0.327 \pm 0.025	0.452 \pm 0.035	100.0
D11S12 (p15.5)	0.125 \pm 0.033	0.227 \pm 0.060	57.4	0.091 \pm 0.045	0.146 \pm 0.072	69.7
D11S679 (p15.5)	-0.465 \pm 0.076	-0.844 \pm 0.134	32.5	0.101 \pm 0.068	0.160 \pm 0.109	69.0
INS (p15.5)	-0.403 \pm 0.079	-0.732 \pm 0.143	27.8	0.103 \pm 0.040	0.165 \pm 0.084	71.8
D11S483 (p15.5)	-0.809 \pm 0.084	-1.469 \pm 0.153	12.4	-0.077 \pm 0.070	-0.123 \pm 0.112	56.8
11p tel (p15.5)	-0.744 \pm 0.086	-1.351 \pm 0.157	13.1	-0.181 \pm 0.063	-0.290 \pm 0.101	41.9

Data in Table 5.2 indicated that chromatin extending from the surface of HSA11p in 2D analysis was also present in 3D-preserved nuclei. At the most distal end of the extruded chromatin, the telomere was positioned $-0.181^{+0.063}$ chromosome radii from the territory edge. However, as in analysis of WAGR, chromatin appeared to be less extended by pFA fixation of 3D-preserved nuclei: the telomere was positioned at $-0.744^{+0.086}$ chromosome radii away from the HSA11p domain in 2D analysis of MAA fixed primary fibroblast nuclei (Chapter 3).

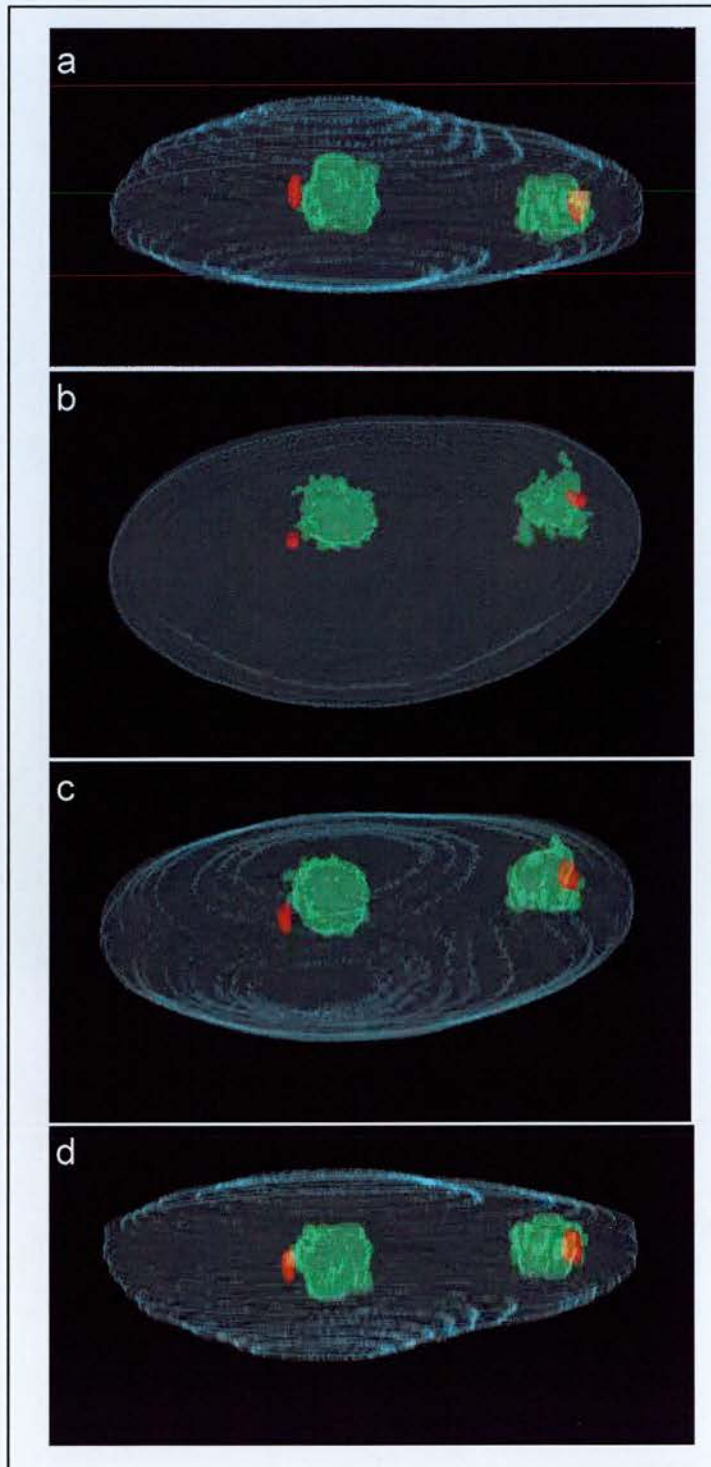


Figure 5.4 Reconstruction of a 3D fibroblast nucleus showing *INS/IGF2* at the surface of HSA11p.

Stacks of image planes at 0.5 μm intervals were captured through 3D-preserved nuclei after FISH using a HSA11p chromosome paint (green) and the *INS/IGF2* locus probe (red) (section 2.9.1) within a fibroblast nucleus (blue). Image stacks were imported into MAPInt (Mouse Atlas Group, MRC Human Genetics Unit) and domains of interest 'painted' separately in each colour stack (Section 2.9.1 and 2.9.2). Nucleus reconstruction was achieved using the AVS/Express visualisation system (Section 2.9.3) and an MPEG movie file, showing a nucleus rotating about the x-axis, was created from the resultant images. Four movie still illustrating rotation from 0 to 180° about the x-axis are shown (a-d). Note that the *INS/IGF2* genes are close to the surface of the HSA11p territory. One allele is outside of the chromosome territory (left), and the other is just within the territory (right).

The % of signals within the chromosome domain differed considerably between the 2D and 3D data sets in the analysis of 11p15. However, this was also true for WAGR when 2D and 3D data were compared (Table 5.2). The organisation of 11p15 in 3D was still noticeably different to that of WAGR from 11p13. To directly compare the spread of data from 2D (Chapter 4) and 3D analysis of 11p15, I plotted cumulative distribution histograms for selected loci from both 2D and 3D data sets. When these distributions were compared, it was apparent that the trends observed in two- and three-dimensions were very similar, even though the absolute mean values are different. These differences were again probably due to the different fixation procedures used and the problems incurred when comparing data obtained in 2D and 3D.

Expressing the data in terms of the distribution of locus positions relative to the chromosome edge is perhaps more meaningful than calculating mean values for each locus: values were spread over a wide range. Comparison of mean values from 2D data facilitated interpretation of a tail of chromatin extruded from the surface of HSA11p, with D11S12 representing the start of this chromatin tail. There were still a significant proportion of D11S12 signals positioned outside of the chromosome domain in 3D (Figure 5.5) and the trend, identified in 2D, for more distal sequences to be more external than proximal loci from 11p15 was also apparent in 3D. Furthermore, comparing the distributions of 11p15 loci (Figure 5.5) to WAGR loci (Figure 5.3), revealed that in 3D, as in 2D, 11p15 is subject to contrasting constraints of organisation to compared to 11p13.

5.5 GENE-POOR 11p14 IS MORE INTERNAL WITHIN HSA11p IN 3D ANALYSIS

In Chapter 4 I speculated that the unusual configuration adopted by chromatin from 11p15.5 was due to the fact that this region is very gene dense. I proceeded to extend use of HSA11p as a model system for studying interphase chromosome organisation and examined the organisation of gene-poor G-band 11p14. This analysis was inconclusive as to whether 11p14 was more internally positioned within HSA11p in the 3 cell lines examined in 2D. However, as discussed (Section 4.11.3), WAGR is only approximately 550kb from the start of 11p14 and may itself be positioned relatively internally within the chromosome domain, due to its proximity to a very gene-poor G-band. I decided to repeat the comparison of the

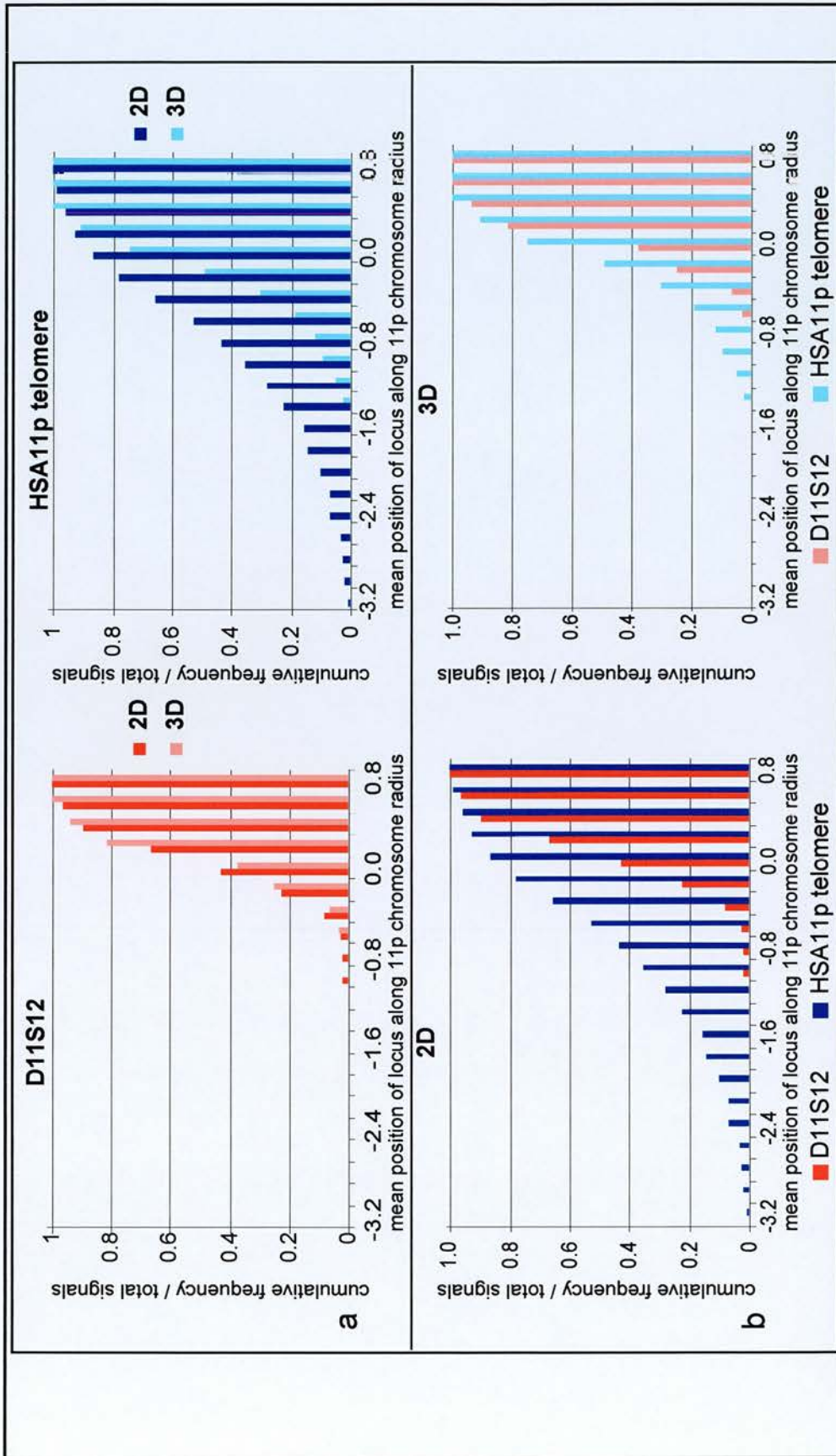


Figure 5.5 The distributions of signals from 11p15 within HSA11p in 2D and 3D analysis of primary fibroblasts.

The cumulative distributions of the positions relative to HSA11p of D11S12 (the start of the tail of chromatin extruded from the surface of HSA11p identified in 2D analysis, Chapter 4) and the 11p telomere in 2D and 3D are compared. The chromatin appears more distended in 2D analysis (a), compared to 3D data for the same locus, but the trend of organisation is the same: the telomere is much more peripheral than D11S12 in both 2D and 3D (b).

position of a locus from 11p14 with 11p13 loci in 3D preserved nuclei. As in Chapter 4 (Section 4.9), the marker D11S865, mapped to 11p14 (Fantes *et al.*, 1995) was combined with a HSA11p paint and used in FISH on 3D preserved fibroblast nuclei. (Table 5.3).

2D analysis of MAA-fixed primary fibroblast nuclei showed that 11p14 was significantly more internal within the HSA11p territory of primary fibroblasts than mean WAGR, ($P < 0.001$). This was also true in three dimensionally preserved (pFA-fixed) fibroblasts taking a 10% confidence interval ($P = 0.095$).

5.6 DISCUSSION

5.6.1 Genes can be transcribed from within the interior of a chromosome territory

I found the entire WAGR locus to be positioned approximately one third of the way along the radius of the HSA11p chromosome territory (Chapter 3). However, due to the spatial problems incurred in 2D analysis due to the flattening of nuclei, I could not comment on the absolute position of WAGR genes with respect to the chromosome territory surface. In this Chapter, I have confirmed this localisation in 3D analysis of pFA-fixed nuclei. Contrary to previous reports that indicated that genes are positioned at the surface of a chromosome domain (Kurz *et al.*, 1996), this study conclusively shows that genes can be positioned within the territory interior. Moreover, I have shown that genes undergoing active transcription can also be buried within the chromosome interior, indicating that there is no exclusion of transcription factors from the interior of a chromosome territory; large-scale chromatin remodelling to position genes on the surface of a chromosome domain is not required to facilitate transcription. Both ubiquitously expressed genes and genes with a tissue-restricted expression pattern can be transcribed from within the interior of an interphase chromosome.

In the analysis of the position of genes within interphase chromosome territories previously performed by Kurz *et al.*, the area taken to represent the periphery of the chromosome depended on the diameter of signals representing the genes (Kurz *et al.*, 1996). The size of these locus signals would have varied with the intensity of fluorescence and the integration

time used in image capture. Therefore, the volume taken as the ‘periphery’ of the chromosome domain in that analysis may actually have included a large proportion of the total chromosome volume; half the volume of a sphere is located in a shell with a radius 1/5 of the radius of the whole sphere.

The idea of genes being exposed on the surface of interphase chromosomes correlated with reports of accumulations of specific RNAs at the border of the chromosome territory from which they originated, and components of the splicing machinery localised outside of chromosome territories (Zirbel *et al.*, 1993). Taken together, these observations provided evidence for the existence of an inter-chromosomal domain (ICD) compartment, a space between chromosomes linked to the nuclear pores, where transcription, RNA processing and RNA transport were suggested to occur (Cremer *et al.*, 1993; Zirbel *et al.*, 1993).

This data questions the extent to which large-scale packaging of chromatin within the chromosome territory can actually contribute to control of gene expression via accessibility of transcription factors to genes. It was thought that large-scale chromatin organisation, above the 30nm fibre, might affect transcription by presenting an accessibility barrier to large protein complexes (Zirbel *et al.*, 1993; Kurz *et al.*, 1996). The idea of a compact chromosome territory that might limit accessibility is contradicted however, by studies of actual chromatin distribution at higher resolution by both light and electron microscopy (Belmont *et al.*, 1999). A chromosome painting study that combined confocal microscopy and deconvolution also revealed considerable internal substructure within chromosome territories (Verschure *et al.*, 1999). If there are accessibility limitations these are perhaps more likely to map to the periphery of large-scale chromatin fibres or larger chromosome subdomains formed by the folding of these fibres. TEM localisation of uridine or BrUTP incorporation has shown heavy labelling at the edge of condensed, large-scale chromatin domains rather than at the surface of chromosomes *per se* (Fakan and Nobis, 1978; Wansink *et al.*, 1996). These studies again suggest that locally compacted and unfolded regions within an interphase chromosome form distinct subdomains, and that chromatin folding is organised in such a way that transcriptionally active loci are at the surface of large-scale chromatin fibres throughout a chromosome territory (Verschure *et al.*, 1999). The fact that the WAGR locus is positioned inside the HSA11p chromosome territory, rather than at the ‘periphery’, is reconcilable with these ideas.

The process of replication has been shown to take place in foci throughout the entire chromosome territory volume, in early as well as in late S-phase (Visser *et al.*, 1998). This demonstrates that activity of macromolecular enzyme complexes takes place throughout chromosome territories and is not confined to the territory surface as suggested previously (Zirbel *et al.*, 1993; Kurz *et al.*, 1996). Nuclear proteins with diverse functions and distinct distribution patterns have been shown to move rapidly throughout the entire interphase nucleus by diffusion using FRAP (fluorescence recovery after photobleaching) (Phair and Misteli, 2000). From these data, proteins are obviously not excluded from chromosome territories and Phair and Misteli suggest that nuclear compartments are the reflection of the steady-state association/dissociation of its 'residents' with the nucleoplasmic space.

To accommodate such observations, the concept of the ICD has been updated. The inter-chromatin domain (also ICD) now includes channels within chromatin domains (Cremer *et al.*, 2000), perhaps ending in branches between 1Mb and 100kb chromatin-loop domains. However, only one further description of gene positioning within chromosomes has been attempted in light of this update (Volpi *et al.*, 2000). This study observed a loop of chromatin extending from the surface of HSA6, suggesting that genes may not be at the 'periphery' of a chromosome territory, but can in fact localise at a distance from their native chromosome domains, as suggested by my data from analysis of 11p15. Volpi *et al.*, performed a qualitative study, so could not comment on the position of genes within a chromosome domain as I have been able to do. Experiments based on fluorescence microscopy at present lack the resolution to examine chromosome territory sub-structure in order find evidence to support or disprove the 'inter-chromatin' concept.

5.6.2 Extrusion of gene-rich chromatin from the surface of the HSA11p chromosome territory is apparent in 3D analysis.

The mean positions of probes spanning more than 4 Mb of distal 11p15 in 3D were different to values obtained in 2D analysis. However, when the distributions of data sets for each locus from 2D and 3D analysis were compared, the shapes of cumulative frequency histograms were very similar. 3D data therefore supported the idea that gene-rich 11p15 is subject to organisational constraints different to those operating on chromatin from 11p13.

Comparing 2D and 3D data, it appears that MAA fixation may result in a preferential loosening of decondensed chromatin. Differences in the compaction of chromatin between

pFA and MAA fixed nuclei have been reported on previous occasions. In one study, analysis of the organisation of the MHC within HSA6 found that a higher percentage of class II region loci were external to the chromosome domain in MAA-fixed cells compared to pFA-fixed cells (Volpi *et al.*, 2000). Studies attempting to delineate DNA organisation within interphase chromosomes found similar hallmarks of random walk/giant loop organisation in three alternatively fixed nuclear preparations (Yokota *et al.*, 1995). Nuclear dimensions of pFA-fixed nuclei were ~0.75 times shorter than those of MAA-fixed nuclei, but relative probe to probe distances were similar within each analysis. This study suggested that, while chromatin may be more loosely packaged in MAA fixed nuclei than pFA fixed nuclei, the organisation of chromatin is not dramatically altered: random walk folding was apparent in both preparations. As noted above (Section 5.2), the radius of the chromosome domain in my analysis was larger within MAA-fixed nuclei than within 3D preserved pFA fixed nuclei, using the same chromosome paint. However, overall trends in organisation observed in 2D and 3D were similar when the data from nuclei subject to the two different fixation procedures were directly compared. A second study, which described regional differences in large-scale chromatin structure during interphase corresponding with more openly configured chromatin of R-bands and a more condensed structure of G-bands, found that R-band and G-band differences were common to both MAA- and pFA-fixed nuclei (Yokota *et al.*, 1997).

5.6.3 The internal position of 11p14 within HSA11p in 3D analysis provides further evidence for a role for gene density in chromatin organisation within interphase chromosomes

In 3D-preserved pFA-fixed primary fibroblast nuclei, a locus from gene-poor G-band 11p14 was significantly more internal within HSA11p than WAGR from the R-band 11p13 (10% confidence interval). This result could be taken to support the more significant difference in relative positions determined by analysis of MAA-fixed nuclei from the same cell line. As described above (Section 5.2), the chromosome domain was more distended by MAA fixation, which is perhaps why the two regions were more significantly displaced in the 2D analysis. The WAGR locus is located in distal 11p13, on the border between a moderately gene dense region (average of 8 genes per megabase over 4 Mb) and a gene poor region in 11p14 (average of 0.5 genes per megabase over 8 Mb) (Gawin *et al.*, 1999; Fantes *et al.*, 1995) (Section 4.11.3). In 2D and 3D, both 11p13 and 11p14 are significantly more internal within the HSA11p territory compared to loci from very gene-rich 11p15 ($P < 0.001$), which

contains at least 47 genes over 4.5 Mb (>10 genes per Mb). These data indicate a possible dependency of chromatin position within a chromosome domain on gene density.

I have shown that the looped out chromatin from gene rich 11p15.5 is also present in three-dimensionally preserved primary fibroblast nuclei and is not a fixation artefact only present in nuclei of MAA-fixed 2D cells. However, chromatin is obviously more decondensed by this method of fixation. As described in Chapter 4, other telomeric sequences, or the HSA11 centromere, are not frequently positioned at a distance from native chromosome territories in 2D, nor are other imprinted human genes. I suggest that this unusual configuration of chromatin could be due to the high gene density of a significant stretch of sequence from 11p15.5. In the next Chapter I will investigate whether other human interphase chromosome domains exemplify patterns of organisation suggested by my analyses of HSA11p.

CHAPTER 6

DO OTHER REGIONS OF THE HUMAN GENOME EXEMPLIFY PATTERNS OF INTERPHASE ORGANISATION SUGGESTED BY HSA11p?

6.1. INTRODUCTION

Analysis of chromosome 11p (Chapters 3-5) provided evidence for a correlation between gene density and the organisation of chromatin within chromosome territories. To explore these ideas further I decided to investigate the intra-chromosomal position of other gene-rich regions of the human genome. Gene-dense regions have previously been predicted from CpG island mapping (Craig and Bickmore, 1994) (Figure 6.1) and these results are supported by emerging data from the Human Genome Project. However, the extent to which the current human genome sequence is accurately assembled and annotated varies considerably from one chromosome to another, so published sequences cannot yet be used in isolation to calculate gene-densities of different regions of the human genome. Most of the sequence still exists only in draft form (<http://www.ncbi.nlm.nih.gov/genome/seq/HsHome.shtml>) with only 47.1% of the human genome completely finished (September 2001). Mapping projects published to date, have been restricted to fine-detail analysis of small stretches of DNA, so are hard to assimilate over extended distances. For my analysis it was necessary to consider extensively mapped regions of the genome, which was a limiting factor when choosing loci to investigate.

6.1.1 Identifying gene-rich human loci

CpG island mapping data (Craig and Bickmore, 1994) (Figure 6.1) provided me with a starting point to investigate gene maps of corresponding regions in the literature. In this way, I identified 11q13, a T-band on HSA11, and the 16p13.3 subtelomeric band from chromosome 16, as well-mapped gene-rich regions (Figure 6.2). Several human diseases have been localised to 11q13 and transcript mapping and EST mapping have long identified this as a gene-rich region of the human genome. It is suggested that there are at least 31

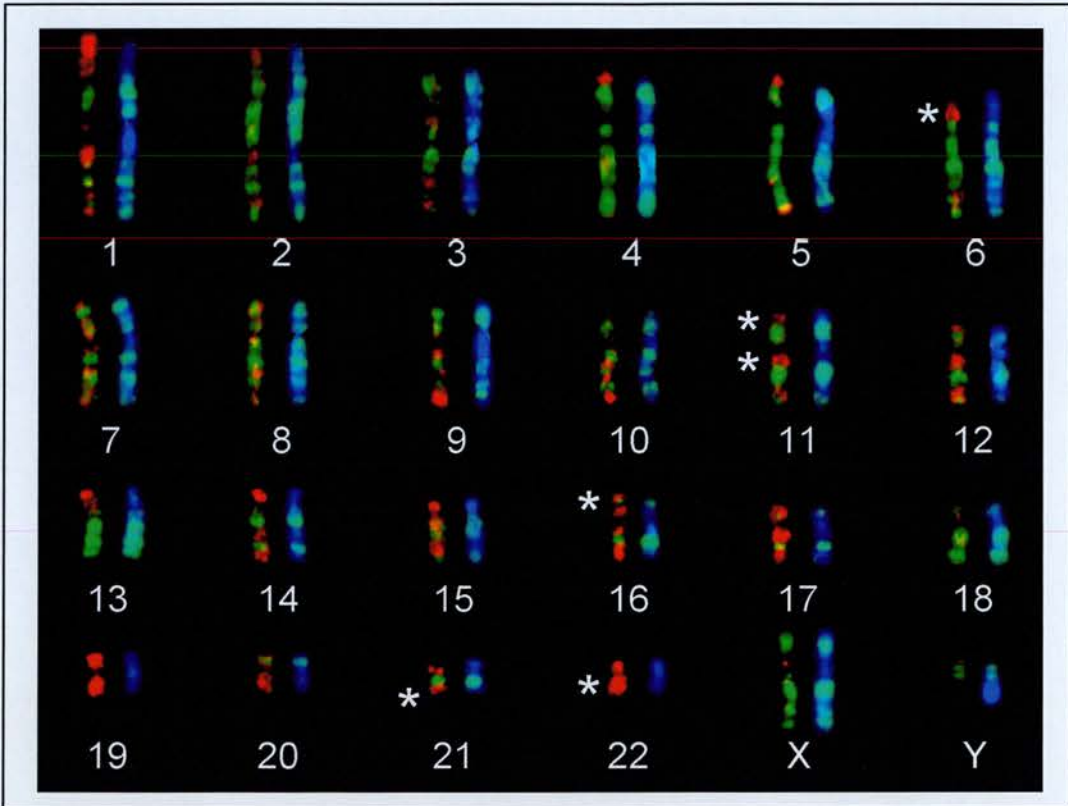


Figure 6.1 The distribution of CpG islands in the human genome.

Karyotype analysis after FISH with a CpG island probe on human metaphase chromosomes from male lymphocytes, with BrdU incorporated into late replicating DNA. The left hand chromosome of each set shows CpG islands in red and inactive DNA in green. The right hand chromosome shows the sites of BrdU incorporation (green) on DAPI-stained chromosomes (blue) (adapted from Craig and Bickmore, 1994).

The distribution of CpG islands predicted regions of high gene density in the human genome, and these predictions are being confirmed by the emerging finished sequence of the human genome. Note the intense red signal at the termini of the short arms of chromosomes HSA11 (11p15) and HSA16 (16p13.3), over distal HSA6p (6p21.3, position of the MHCII locus), proximal HSA11q (11q13) and also over the entirety of chromosome 22. Chromosome 21 is characterised by high gene intensity on the short arm, which contains rDNA, and also in the sub-telomeric region, whereas the rest of the q arm consists of inactive DNA (green). All gene-rich regions considered or discussed in this chapter are indicated by white asterisks.

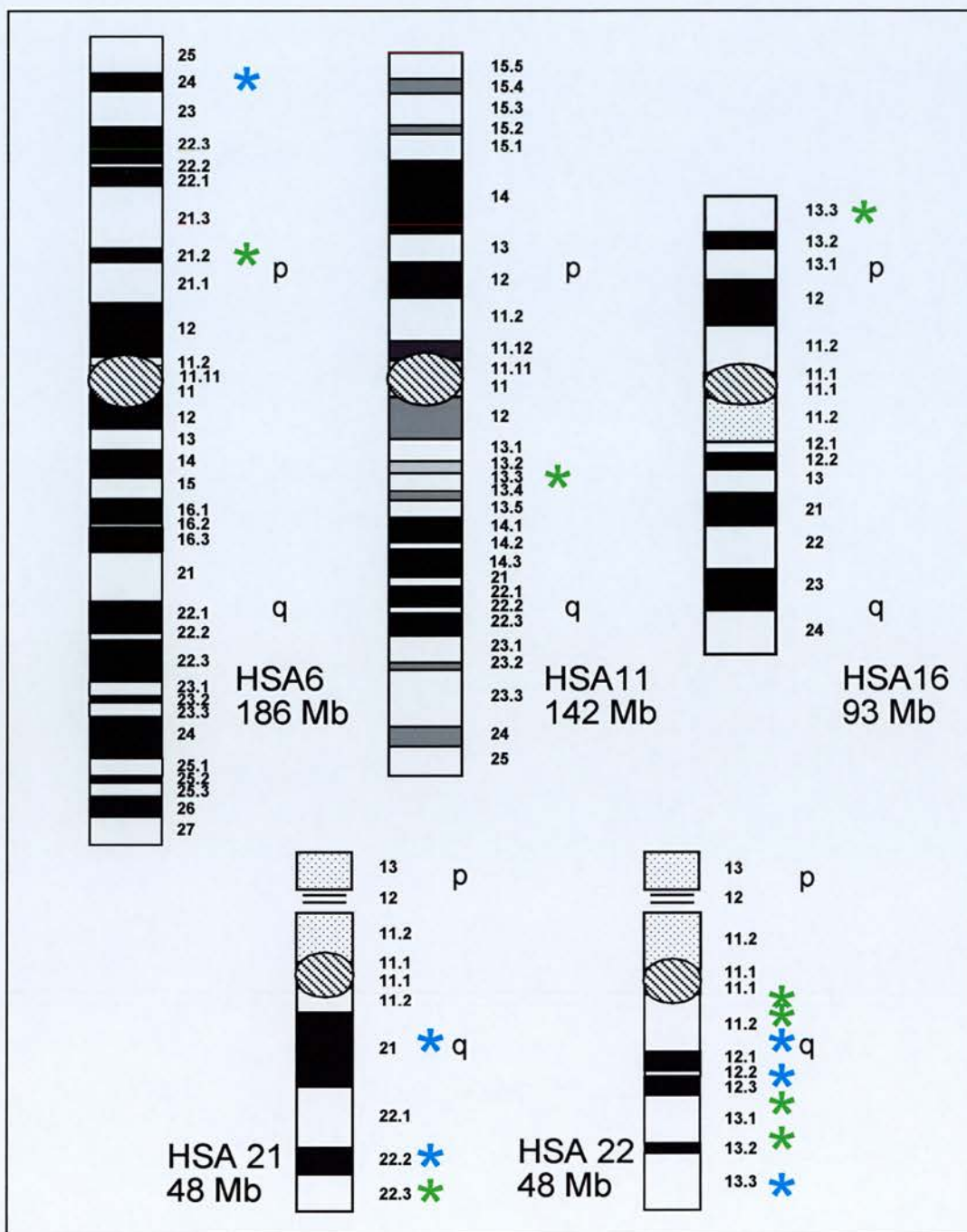


Figure 6.2 G-banded ideograms of human chromosomes 6, 11, 16, 21 and 22.

Standard G-banded ideograms of human chromosomes 6, 11, 16, 21 and 22 (ISCN 1995). Diagrams are to-scale and the position of particular cytogenetic bands, sequences from which have been considered in this chapter, are indicated by asterisks (green represents a gene-rich region, blue a gene-poor region).

transcripts present in a 1.65 Mb contig centred around the *FEN1* gene (Cooper *et al.*, 1997; Cooper *et al.*, 1998), which is equivalent to at least 20 genes/Mb. The terminal 1.95 Mb of chromosome 16p has been sequenced (Daniels *et al.*, 2001), confirming that this region is extremely gene rich, containing at least 100 genes and a further 20 predicted genes (approximately 50 genes/Mb). The α -globin genes are contained within 16p13.3 and span about 25kb, lying approximately 130kb from the subtelomeric region with another four widely expressed genes between the α -globin cluster and the telomere (Vyas *et al.*, 1992).

A previous study examined intra-chromosomal position of sequences derived from the short arm of chromosome 6 (Figure 6.2) (Volpi *et al.*, 2000). Large chromatin loops were observed extending outwards from the surface of the domain, the frequency of which was cell type dependent and correlated with the number of active genes in that region. The MHC locus is also gene-rich and contains an average of 62.5 genes per megabase (including pseudogenes) (The MHC Sequencing Consortium, 1999).

6.1.2 Using finished sequence from the human genome: HSA21q and HSA22q

After determining the interphase organisation of these gene-rich regions from HSA11q and HSA16p relative to their native chromosome domains, I decided that further extension of this investigation of chromatin from scattered chromosomal locations would be unsatisfactory. This analysis was more informative, than previous studies of isolated genes (Kurz *et al.*, 1996; Dietzel *et al.*, 1999), as I was able to put genes studied into context of the surrounding sequence. However, as sequence composition account beyond the extent of published maps might influence the position adopted by a locus in relation to its chromosome territory; a more complete DNA sequence was required to facilitate a comprehensive analysis of interphase intra-chromosome organisation.

Chromosomes 21 and 22 represent two of the five acrocentric human chromosomes, each of which share sequence similarity in their short arms, where the tandemly repeated ribosomal RNA genes and a series of other tandem repeat sequences are located. Part way through my research project, the sequences of the euchromatic parts of chromosomes HSA21 and HSA22 were published (Dunham *et al.*, 1999; The chromosome 21 mapping and sequencing consortium, 2000). Chromosomes 21 and 22 have a similar physical size, consisting of approximately 48 Mb each (http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/hum_srch?chr=hum_chr.inf&query) (Figure 6.2), and together represent around 2% of the human genome (Table 7.3). However, they have strikingly different characteristics: chromosome 22

is predicted to contain at least 545 genes spanning 33.4 Mb, in contrast, there is evidence for only 225 genes embedded in the 33.8 Mb of genomic DNA of chromosome 21. Therefore, as was predicted by CpG island mapping by the absence of late replicating DNA on HSA22 and the large portion of late replicating DNA constituting proximal HSA21q, chromosome 22 is gene-rich whereas chromosome 21 is gene-poor, (Figure 6.1). Correspondingly, trisomy HSA21 is associated with Down syndrome, the most frequent genetic cause of significant mental retardation, and is one of the few trisomies viable in humans. The completed sequences of chromosomes 21 and 22 thus provided the ideal tool, enabling exploration of the distribution of gene-rich and gene-poor chromatin in the context of contrasting environments comprising more than 30 Mb of characterised linear sequence and whole chromosome arms.

6.2 LOCI FROM GENE-RICH 11q13 AND 16p13 ARE FREQUENTLY POSITIONED OUTSIDE OF THEIR CHROMOSOME TERRITORIES

6.2.1 11q13

To locate 11q13 within its chromosome territory, I used a BAC probe, 80N22 (BACPAC Resources, <http://www.chori.org/bacpac/>), in combination with a HSA11q-specific paint (gift of M. Bittner, produced as described in Section 2.4.1) in 2D and 3D FISH analysis. Fifty 2D images and fifteen 3D image stacks were captured and the position of the BAC analysed relative to the interphase 11q chromosome territory, using scripts described above (Boxes 3.1, 3.2 and Section 2.9.2) (Figure 6.3). In 2D analysis the 11q13 probe (80N22) had a mean position $-0.325^{\pm 0.062}$ of a chromosome radius outside of the chromosome domain ($-0.606^{\pm 0.116}$ μm). Similarly, in 3D analysis, 80N22 had a mean position $0.300^{\pm 0.106}$ times the chromosome radius outside of the chromosome domain, equivalent to $0.397^{\pm 0.140}$ μm . Therefore, as suggested by my analysis of HSA11p (Chapters 3 and 4), chromatin from gene-rich 11q13 localised outside of the main chromosome 11q domain, as defined by a chromosome arm paint. In 3D analysis, 42.9% of signals were within the chromosome territory, whereas only 29.9% within the chromosome in 2D, again indicating a greater compaction in 3D pFa-fixed nuclei than 2D. This was also apparent from a smaller territory radius in 3D ($1.979^{\pm 0.028}$ μm) compared to 2D ($1.324^{\pm 0.049}$ μm), although the distribution of signals is very similar when 2D and 3D data sets are compared (Figure 6.3e).

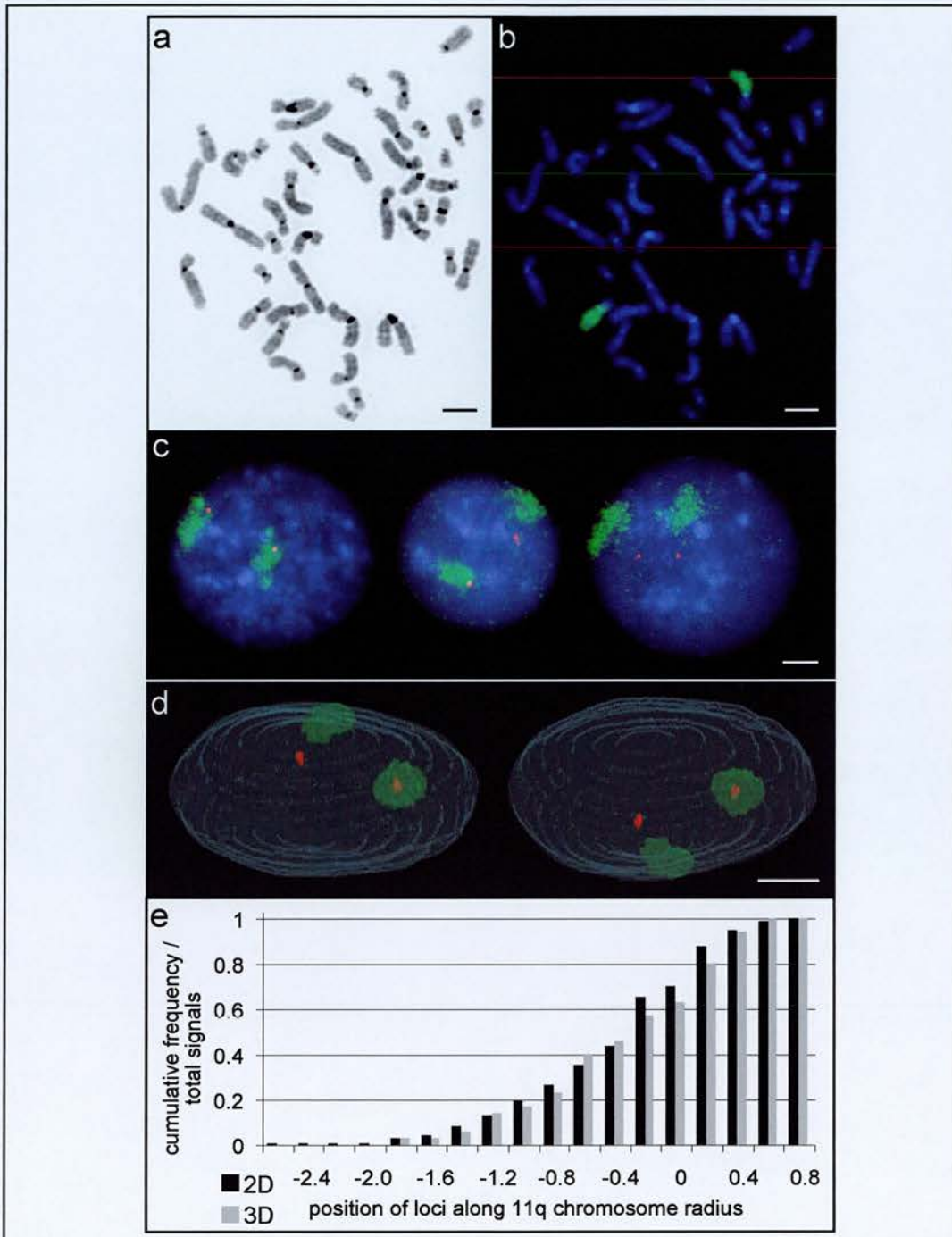


Figure 6.3 The spatial organisation of 11q13 in two and three dimensions.

a) Grey-scale representation of a DAPI-stained metaphase chromosome spread from lymphoblastoid cells.

b) The specificity of the HSA11q paint (green) is illustrated on the same chromosome spread.

c) MAA-fixed nuclei counterstained with DAPI, hybridised in FISH with a HSA11q paint (green) in combination with the BAC probe 80N22 (red). 80N22 maps to cytogenetic band 11q13 which is gene-rich, containing at least 31 transcripts within 1.65Mb. Note the widely varying positions adopted by the cosmid probe in relation to the HSA11q domain, usually on the edge of or external to the chromosome territory.

d) Views of a reconstruction of a three-dimensionally preserved fibroblast nucleus after FISH using identical probes, rotated about the x -axis. The BAC probe appears distant from the territory of one homologue, but is within the interior of the other chromosome.

e) The cumulative distributions of the locus along the 11q territory radius are compared in 2D and 3D. Note the similar distributions, but that the 11q territory appears more decondensed in 2D. 0 represents the edge of the chromosome and negative values are outside of the territory domain as defined in FISH. Scale bars represent $5\mu\text{m}$.

6.2.2 16p13.3

The second gene-rich region outside of HSA11 that I chose to analyse was located within 16p13.3. I obtained a cosmid probe, cCOS12, which mapped approximately 175kb from the telomere of chromosome 16p, just centromeric of the α globin genes (D. Higgs, Oxford) (Flint *et al.*, 1997). I combined this cosmid, labelled with Dig, with a biotin-labelled 16p paint (gift of Michael Bittner) and hybridised the probes to MAA-fixed FATO nuclei. Analysis of fifty 2D images (Figure 6.4), produced a mean position of $-0.840^{\pm 0.102}$ of the chromosome radius outside of the chromosome domain (equivalent to $-1.567^{\pm 0.190}$ μm outside of the territory). Therefore, as found for 11q13, a locus from gene-rich 16p13.3 was frequently positioned outside of the HSA16p territory. Only 16.5 % of α -globin locus signals were located within the HSA16p domain, in stark contrast with the organisation of the β -globin locus in HSA11p15.4, which was positioned within 87.2% of HSA11p territories analysed (Chapter 4).

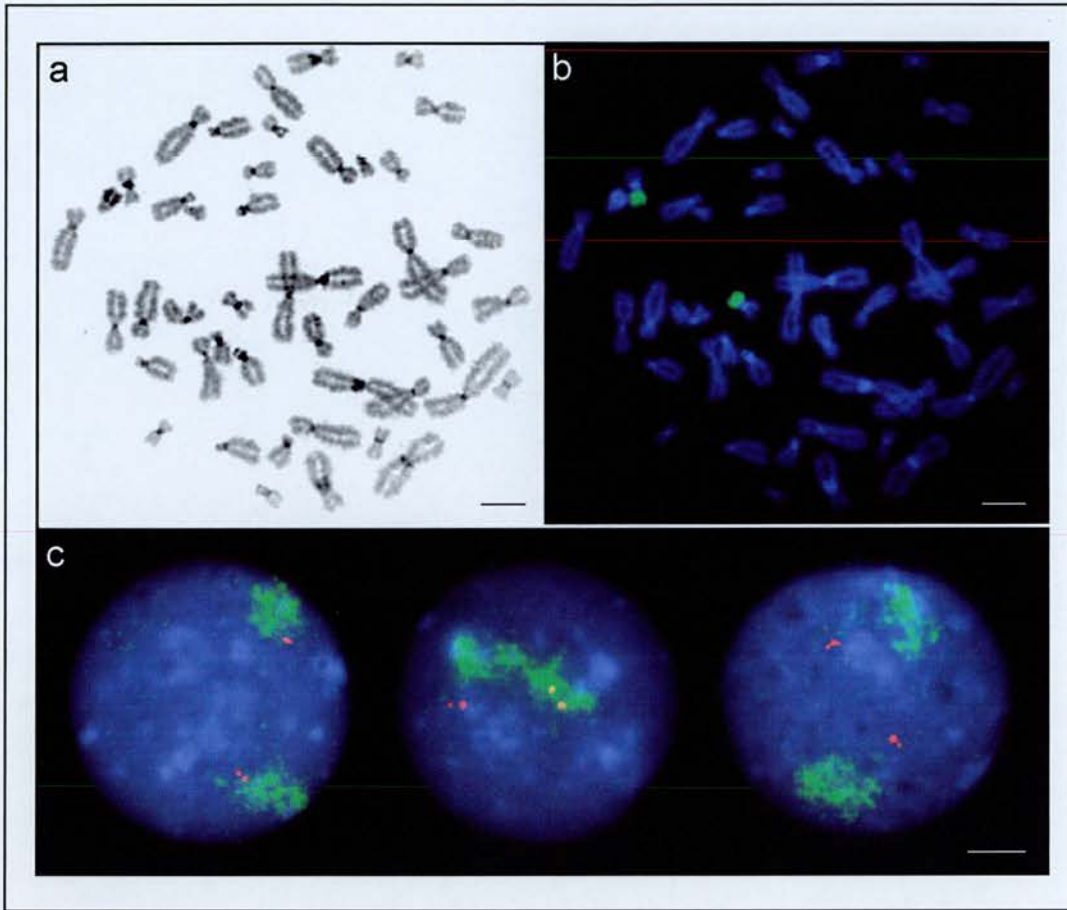


Figure 6.4 The spatial organisation of 16p13.3.

- a) Grey-scale representation of a DAPI-stained metaphase chromosome spread from lymphoblastoid cells.
- b) The specificity of the HSA16p paint (green) is illustrated on the same chromosome spread
- c) MAA-fixed nuclei counterstained with DAPI, hybridised in FISH with this HSA16p paint in combination with the c12 cosmid probe (red) are depicted. cCOS12 maps to a position 175 kb from the telomere of HSA16. Note the varying positions adopted by the cosmid probe in relation to the HSA16p domain, usually on the edge of, or at a distance from, the chromosome domain, but sometimes also within the chromosome interior.

6.3 UTILISING THE FINISHED SEQUENCE OF HSA22q TO INVESTIGATE PATTERNS OF INTERPHASE CHROMOSOME ORGANISATION

The genomic sequence of 22q was the first complete sequence of an entire chromosome arm to be published (Dunham *et al.*, 1999). I decided to utilise these data to test my hypothesis that gene rich regions are more peripherally positioned relative to interphase chromosome domains than gene-poor regions. I used the online resource that allowed long-range analysis of chromosome 22q to calculate the number of genes per megabase along the entire chromosome arm (<http://www.sanger.ac.uk/cgi-bin/humace/SuperMap22>) (Figure 6.5a). When calculating the number of genes in each megabase region, I included genes, related genes and predicted genes but not pseudogenes. A set of BACs had been selected, spaced along the long arm of chromosome 22 at approximately 2 megabase intervals, and made available from the Sanger Centre clone resources group (<http://www.sanger.ac.uk/HGP/Cytogenetics/Bacset.shtml>). I chose four gene-dense and three gene-poor megabase regions of chromosome 22 to which a corresponding BAC was available (Table 6.1) (<http://www.sanger.ac.uk/cgi-bin/humace/SuperMap22>).

2D FISH experiments were performed combining each labelled HSA22 probe with a chromosome 22q paint (gift of Michael Bittner). Fifty images were captured for each probe (Figure 6.6), and used to calculate the mean position of each locus relative to the HSA22q interphase chromosome territory (Table 6.1 and Figure 6.5b). As the HSA22q paint cross-hybridised with the centromeres of HSA14, care was taken not to include centromere fluorescence in the area defined as the HSA22q territory. The percentages of signals positioned within the HSA22 territory were also calculated for each locus (Table 6.1 and Figure 6.5c).

Results (Figure 6.5) showed that there was a correlation between the mean position of a probe relative to the interphase chromosome 22q territory and the gene-richness of the corresponding region of the chromosome. As identified in analysis of HSA11p15 (Section 4.5), there was also a close correlation between quantitative comparisons of mean positions of loci relative to the edge of the HSA22q chromosome and the percentage of locus signals positioned within the chromosome territory (qualitative analysis, see Section 4.5).

Table 6.1 Positions of BACs corresponding to gene-rich and gene-poor regions of chromosome 22, relative to the interphase HSA22q territory.

Gene-rich and gene-poor Mb regions of HSA22q were identified using the on-line resource provided by the Sanger centre for long-range analysis of HSA22 (<http://www.sanger.ac.uk/cgi-bin/humace/SuperMap22>). Probes corresponding to these regions were obtained (<http://www.sanger.ac.uk/HGP/Cytogenetics/Bacset.shtml>), labelled with digoxigenin and used in FISH on MAA fixed nuclei from lymphoblastoid cell line (FATO) in combination with HSA11p-biotin chromosome arm paint. Fifty images were captured for each probe and the positions of locus signals within chromosome domains were analysed using the scripts described in Chapter 3. Mean positions of loci relative to the nearest chromosome edge were calculated (\pm SEM) (Section 2.10). The percentages of signals positioned within and outside of the chromosome territory were also calculated for each locus.

Name of BAC	Cytogenetic location	Mb along HSA22 sequence (exact BAC location, kb)	Gene rich/ gene poor (genes/ Mb)	Mean probe location as proportion of chromosome radius	Mean distance of probe from edge of chromosome (μ m)	% signals within chromosome	% signals outside of chromosome
154H4	22q11.2	3-4 (3211)	Rich (22)	0.085 \pm 0.033	0.132 \pm 0.051	75.8	24.2
433F6	22q11.2	4-5 (4467)	Rich (15)	-0.040 \pm 0.049	-0.062 \pm 0.076	73.2	26.8
992D9	22q11.2	11-12 (11076)	Poor (3)	0.333 \pm 0.049	0.518 \pm 0.076	82.7	17.3
221H1	22q12.2-q12.3	18-19 (18082)	Poor (3)	0.139 \pm 0.031	0.216 \pm 0.020	82.3	17.7
390B3	22q12.3-q13.1	21-22 (21264)	Rich (27)	-0.048 \pm 0.057	-0.075 \pm 0.089	69.6	30.4
250D10	22q13.1-13.2	25-26 (25784)	Rich (25)	-0.028 \pm 0.040	-0.044 \pm 0.062	65.8	34.2
1109B9	22q13.3	31-32 (31245)	Poor (2)	0.217 \pm 0.031	0.338 \pm 0.048	83.0	17.0

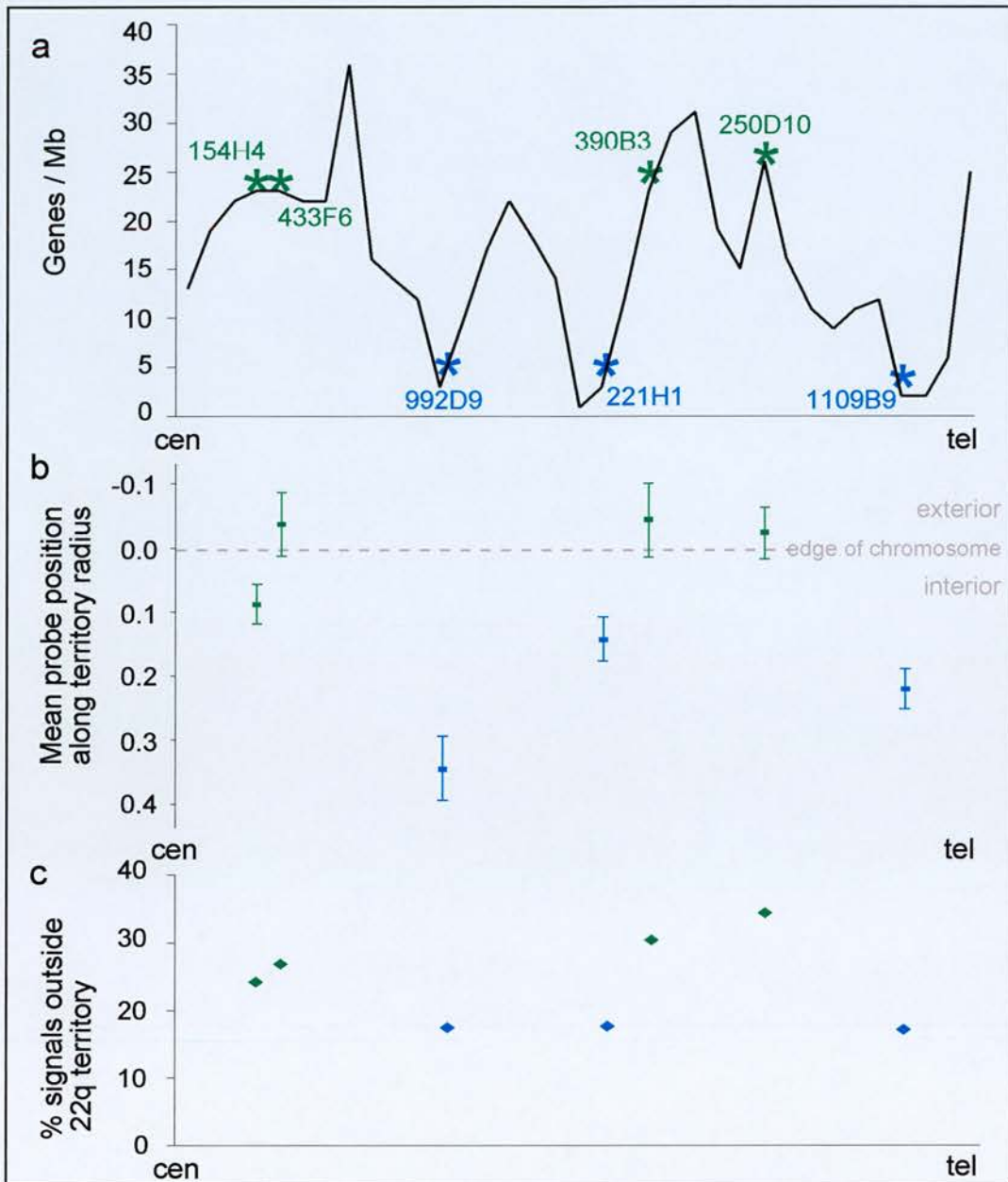


Figure 6.5 Sequence location and spatial organisation at interphase of gene-rich and gene-poor regions from chromosome HSA22q.

a) Gene density per megabase of DNA across the long arm of human chromosome 22 is depicted (as calculated from the published sequence at <http://www.sanger.ac.uk/cgi-bin/humace/SuperMap22>). The number of genes in each megabase includes known genes, predicted genes and related genes but does not include pseudogenes. The positions of probes corresponding to gene-rich and gene-poor megabase regions across the HSA22q sequence are indicated by asterisks (gene-rich in green, gene-poor in blue). The published sequence (Dunham *et al.*, 1999) spans 34 Mb from the centromere (0 Mb) to the telomere.

b) The mean positions of each probe in relation to the nearest edge of the HSA22q chromosome territory are plotted (normalised for chromosome radii), according to their location along the 22q sequence. Note the strong correlation between gene density and the probe position relative to the chromosome domain: probes corresponding to gene-rich regions (green) are peripherally located or outside of the chromosome limits, whereas gene-poor probes (blue) are all positioned within the chromosome.

c) There is also a strong correlation between the % of loci positioned within the chromosome domain and gene density.

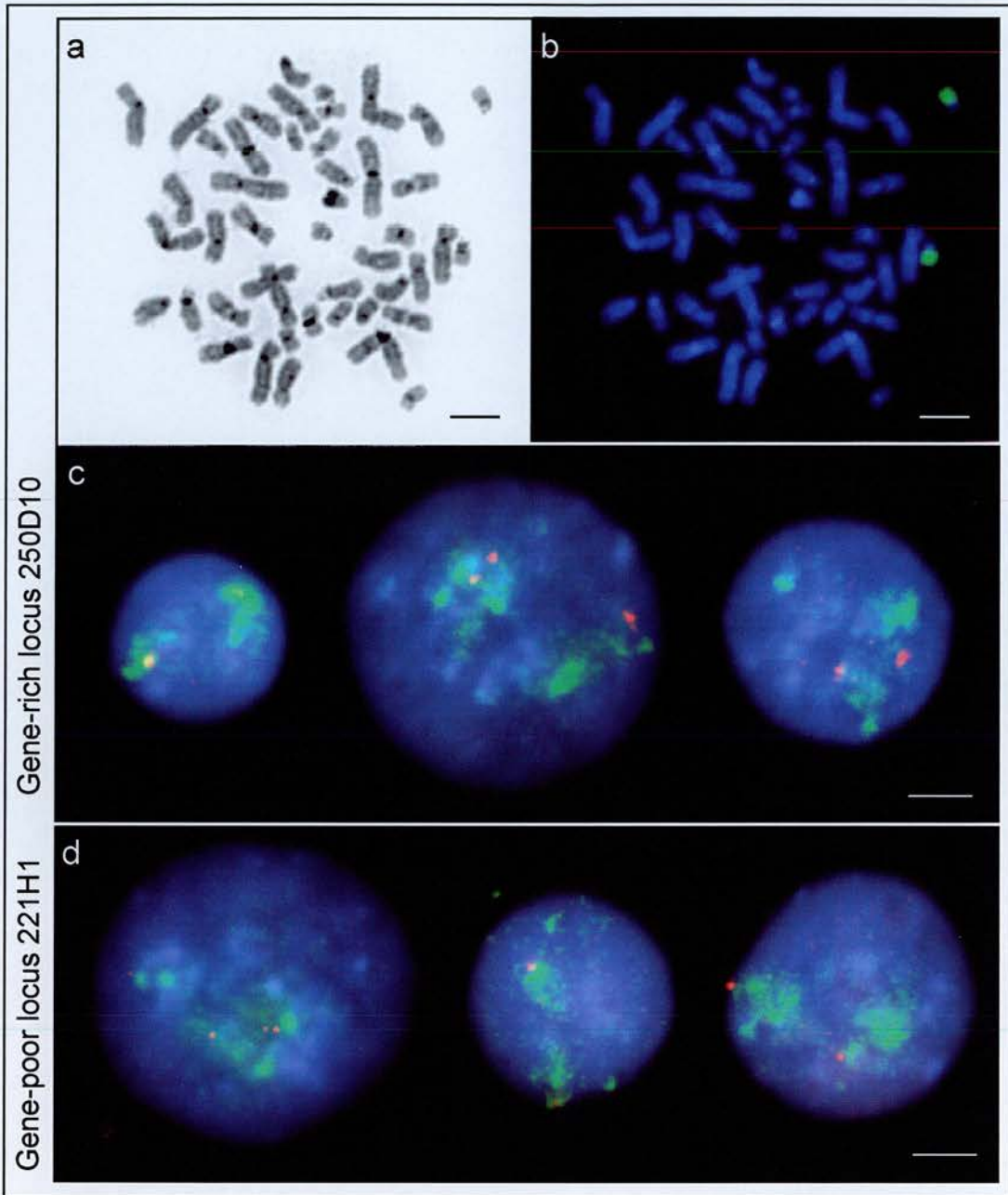


Figure 6.6 The spatial distribution of gene-rich and gene-poor regions of HSA22q within lymphoblastoid cell nuclei.

a) Grey-scale representation of a DAPI stained metaphase chromosome spread from lymphoblastoid cells.

b) The specificity of the HSA22q paint (green) is illustrated on the same chromosome spread (note the cross reactivity with sequences from the centromere of chromosome 14).

c and d) MAA-fixed nuclei counterstained with DAPI and hybridised in FISH with this HSA22q paint in combination with BAC probes (red) to a gene-rich region (250D10) (c) and a gene-poor region (221H1) (d) are shown. Note the very open structure of HSA22q and the internal position occupied by the chromosome domain within the nucleus. Both BAC probes adopt widely varying positions in relation to the HSA22q domain, both within and external to the chromosome, however, probes corresponding to gene-rich loci are generally more external to the chromosome domain than those representing gene-poor loci. Scale bars represent 5 μm .

Probes that represented gene-rich regions all had a mean location near the edge of the chromosome domain, and three of the four (433F6, 390B3 and 250D10) were all similarly positioned, outside of the chromosome. In contrast, gene-poor loci, localised within the territory interior. When the positions of the four gene-rich loci were compared to three gene-poor loci, the difference was highly significant ($P < 0.001$ using a standard *t* test). When the data are examined in more detail, a correlation between mean positions of loci and the proximity of other gene-poor and gene-rich regions of sequence can perhaps be made. For example, gene-rich loci 154H4 and 433F6 represent adjacent megabase regions of HSA22q, but the mean position of 154H4 is within the chromosome whereas the mean position of 433F6 is external (Figure 6.5). These differences can be understood better when the locations of these loci are put into the context of the chromosome as a whole and the gene density of adjacent regions of chromatin considered. 154H4 is next to the gene-poor centromeric region, and 433F6 is near to the highest peak of gene density across the entire chromosome arm. Therefore, if the interplay between forces exerted by the gene density of chromatin act to organise chromatin within an interphase chromosome territory, a long-range analysis such as this may be required to appreciate patterns of organisation.

A similar analysis in 3D-preserved primary fibroblast nuclei was undertaken, but the diffuse nature of the HSA22q territory meant that the chromosome domain was not sufficiently distinguishable from background signals, which are often present at high levels in 3D FISH experiments.

6.4 USING THE FINISHED SEQUENCE OF HSA21q TO INVESTIGATE INTERPHASE CHROMOSOME ORGANISATION

In May 2000, the complete chromosome sequence of HSA21 was published (The chromosome 21 mapping and sequencing consortium, 2000), comprising 33.5 Mb from the long arm and 0.3 Mb from the short arm. Distinguishing gene-rich and gene-poor regions and obtaining clones corresponding to these regions was not as simple as for HSA22 (Section 6.3), as on-line resources enabling long range analysis of the sequence were not set up by the centres that collaborated to sequence chromosome 21, nor was there an extensive set of mapped clones easily available.

The published article (The chromosome 21 mapping and sequencing consortium, 2000) stated that HSA21 has a 7 Mb region (5-12 Mb) that contains only two genes, and that there

are three other 1 Mb regions virtually devoid of genes. Together, these gene-poor regions comprise almost 10 Mb, one third of HSA21q, and correspond to a large late-replicating CpG island-poor region previously identified on the proximal part of the 21q arm (Craig and Bickmore, 1994b) (Figure 6.1). Using data presented in this paper, I attempted to locate gene-poor and gene-rich regions by scanning the figure of the sequence map annotated with transcript predictions. In addition, I also used the NCBI website (<http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/maps.cgi?org=hum&chr=21>), which contains annotations of genes within the published sequence. It was difficult to relate the two sequences as the paper (The chromosome 21 mapping and sequencing consortium, 2000) began kilobase numbering at the centromeric end of the q arm, and NCBI uses zero to denote the telomere of the p arm, the start of the entire chromosome. However, using both resources to assess the distribution of genes across the long arm of chromosome 21 as accurately as possible, I identified gene-rich and gene-poor megabase regions (Figure 6.7a). Three probes were obtained from the consortium involved in the HSA21 sequencing project. One was mapped to the centre of the 7 Mb gene-poor region between 5 Mb and 12 Mb on the published sequence map of HSA21q. A second probe mapped to a gene-poor megabase between 21.9 and 22.9 Mb and a probe to the gene-rich subtelomeric region between 31.3 and 32.3 Mb were also obtained (gifts of M-L Yaspo, Berlin) (Table 6.2). Images were captured after combining these probes with a 21q paint (gift of Michael Bittner) (Figure 6.8). The HSA21q cross-reacts with the centromeres of HSA13, so care was taken not to include fluorescence from the centromeres in the area manually defined to contain the HSA21q territory signal in script analysis. Mean positions for each locus, relative to the HSA21q chromosome territory were calculated (Table 6.2 and Figure 6.7b), as was the % of signals within the chromosome territory (Figure 6.7c) (qualitative analysis, see Section 4.5).

Only two genes (*PRSS7* and *NCAM2*) and 5 hypothetical genes are contained within 7 Mb, from 5-12Mb on the published sequence map of HSA21 (The chromosome 21 mapping and sequencing consortium, 2000). A probe mapped within this region, approximately 7.7Mb from the beginning of HSA21, had a mean location $0.392^{\pm 0.047}$ of the chromosome radius from the nearest edge, well within the territory. The first 13Mb of HSA21q are all, in fact, very gene-poor and contain only 21 genes (NCBI). This is equivalent to an average of 1.6 genes per megabase. In contrast, the subtelomeric region of HSA21q (represented by P31N7), is very gene rich. There are 73 genes in the distal 5.2 Mb of DNA (NCBI), an average of 14.6 genes per megabase. The mean position of this locus was $-0.895^{\pm 0.127}$ of a chromosome radius outside of the chromosome domain, significantly more peripheral than the gene-poor probe ($P < 0.001$). Therefore, once again, gene-rich loci localised outside of the HSA21q territory, whilst gene-poor loci were internal.

Table 6.2 Positions of BACs corresponding to gene-rich and gene-poor regions of chromosome 21, relative to the interphase HSA21q territory.

Gene-rich and gene-poor Mb regions of HSA21q were identified using the published paper describing the sequence of HSA21 (Hattori *et. al.*, 2000). Probes corresponding to these regions were obtained (M-L Yaspo, Berlin), labelled with digoxigenin and used in FISH on MAA fixed nuclei from lymphoblastoid cell line (FATO) in combination with HSA11p-biotin chromosome arm paint. Fifty images were captured for each probe and the positions of locus signals within chromosome domains were analysed using the scripts described in Chapter 3. Mean positions of loci relative to the nearest chromosome edge were calculated (\pm SEM) (Section 2.10). The percentages of signals positioned within and outside of the chromosome territory were also calculated for each locus.

Name of probe	Cytogenetic location	Mb along HSA21 sequence (exact probe location, kb)	Gene rich/ gene poor (genes/ Mb)	Mean probe location as proportion of chromosome radius	Mean distance of probe from edge of chromosome (μ m)	% signals within chromosome	% signals outside of chromosome
BAC R47C12	21q21-q21	5000-12000 (7700)	Poor (<1)	0.392 \pm 0.047	0.298 \pm 0.036	85.4	14.6
Cosmid Q98A3	21q22.2	21900-22900 (22800)	Poor (2)	0.254 \pm 0.048	0.193 \pm 0.036	81.6	18.4
PAC 314N7	21q22.3	31300-32300 (31550)	Rich (24)	-0.895 \pm 0.127	-0.680 \pm 0.097	41.5	58.5

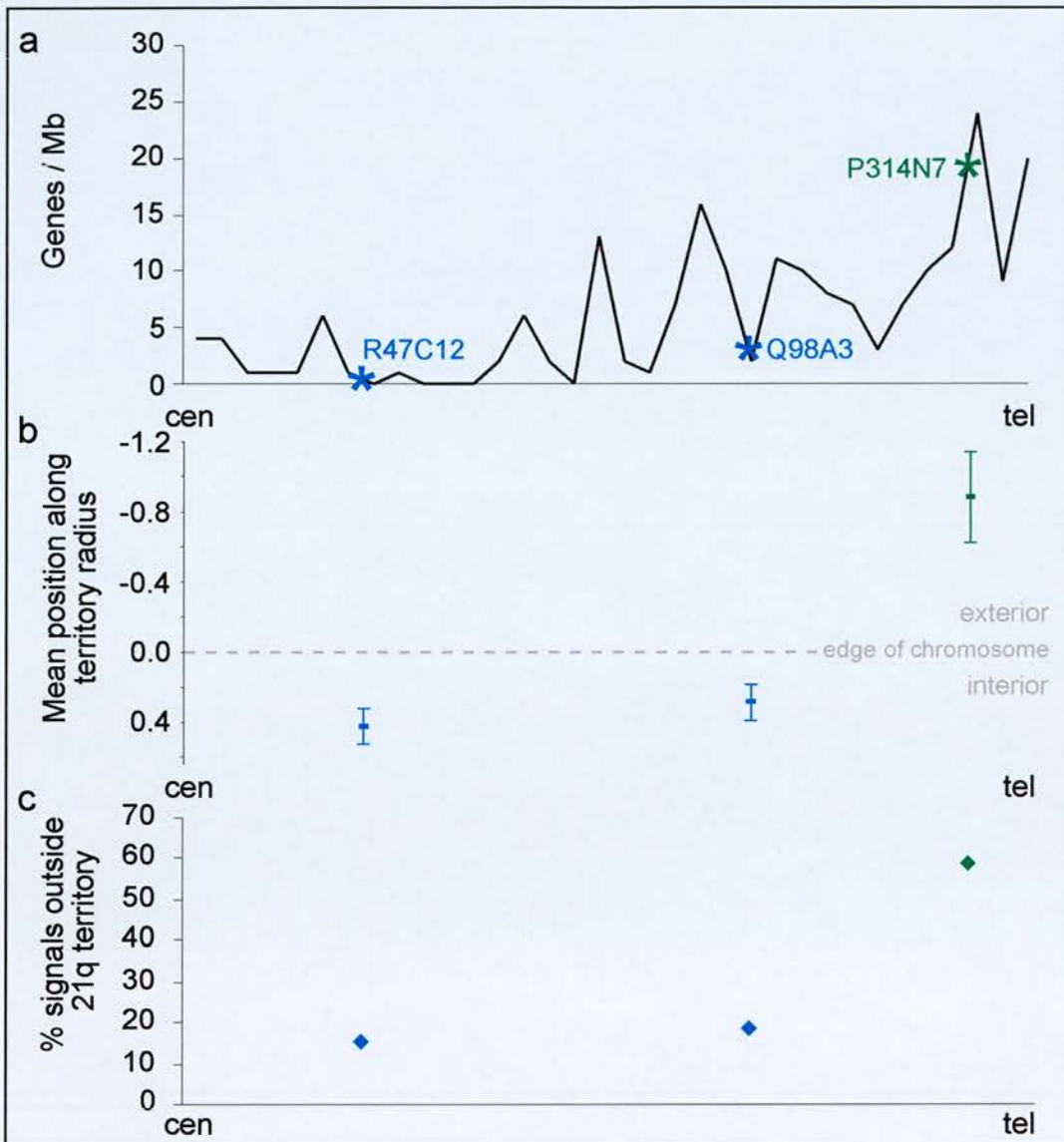


Figure 6.7 Sequence location and spatial organisation at interphase of gene-rich and gene-poor regions from chromosome HSA21q.

a) The gene density per megabase of DNA across the long arm of human chromosome 21 is depicted (calculated from the published sequence at <http://www.ncbi.nlm.gov/cgi-bin/Entrez/maps.cgi?org=hum&chr=21>). These estimates are approximate as the sequence is still incompletely annotated, and data detailing the exact position of genes across the chromosome is not yet available. The positions of probes corresponding to gene-rich and gene-poor megabase regions are indicated by asterisks (gene-rich in green, gene-poor in blue). The published sequence (Hattori *et al.*, 2000) spans 35kb from the centromere (0 Mb) to the telomere.

b) The mean positions of each locus in relation to the nearest edge of the HSA21q chromosome territory are plotted (normalised for chromosome radius), according to their location along the 21q sequence. Note the strong correlation between gene density and the probe position relative to the chromosome domain. The gene-poor loci (blue) are both positioned within the chromosome, however, the locus representing the 7 Mb region which contains only two genes (R47C12) is significantly more internal than Q98A3 which is close to a gene rich region. In contrast, the mean position of the very gene-dense region is outside of the main HSA21q territory.

c) There is also a strong correlation between the % of loci positioned within the chromosome domain and the gene density (see section 4.5).

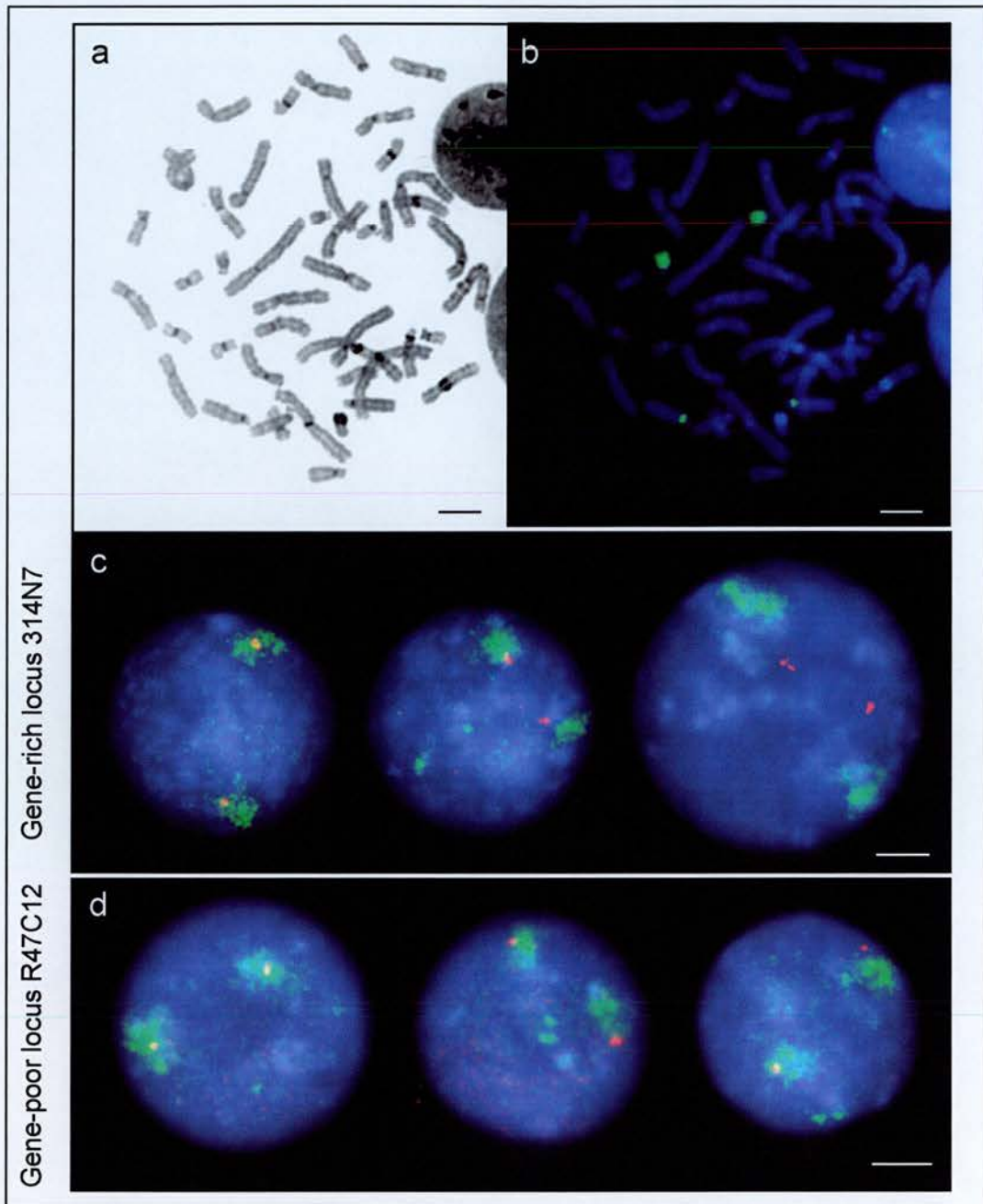


Figure 6.8 The spatial distribution of gene-rich and gene-poor regions of HSA21q within lymphoblastoid cell nuclei.

a) Grey-scale representation of a DAPI stained metaphase chromosome spread from lymphoblastoid cells is shown.
 b) The specificity of the HSA21q paint (green) illustrated on the same chromosome spread (note the cross reactivity with sequences from the centromere of chromosome 13).
 c and d) MAA-fixed nuclei counterstained with DAPI, hybridised in FISH with this HSA21q paint in combination with 21q probes (red) are depicted (c and d). PAC 314N7 corresponds to the sub-telomeric gene-rich region between 31300 and 32300 kb (c) and BAC R47C12 maps to the gene-poor 7 Mb domain between 5000 and 12000 kb (d). Note the different positions adopted by the probes relative to the HSA21q domain, especially the distant localisation of the gene-rich sequence (c). The gene-poor probe was never observed at such a distance, and rarely appeared outside of the chromosome domain.
 Scale bars represent 5 μm .

R47C12 is mapped within an extensive (7 Mb) gene-poor region and is more internally positioned within HSA11p than cQ98A3, which is mapped very close to a gene-rich region (see Figure 6.10). Continuing consideration of loci in context of the gene density of adjacent chromatin (discussed for HSA22q, Section 6.3), cQ98A3 was perhaps drawn more towards the chromosome periphery by adjacent gene-rich chromatin, which I would predict to localise towards the chromosome periphery.

As for HSA22q, analysis of the percentage of locus signals positioned within the chromosome correlated strongly with the distribution of probes along the chromosome radius. 52% of signals were positioned more than 0.4 of the way along the chromosome radius for R47C12, compared to 36.8% for the other gene-poor locus (adjacent to two gene-rich regions) and only 21.8% for the gene rich locus represented by P314N7 (Figure 6.7).

6.5 GENE DENSITIES OF REGIONS OF THE GENOME MAY CORRELATE WITH THE SIZE OF CHROMATIN LOOPS EXTENDING FROM THE SURFACE OF CHROMOSOME DOMAINS

My hypothesis anticipated that all gene-rich regions (11q, 16p13.3 and loci within HSA21q and HSA22q) would localise outside of their respective chromosome domains, and the data presented in this Chapter and Chapter 4 supported this prediction. 16p13.3 contains at least 100 genes and 20 predicted genes within approximately 2 megabases, equivalent to about 50 genes per megabase. This is the highest gene density of all the loci I investigated and, interestingly, the mean localisation of cCOS12, 175 kb from the telomere of 16p, is further away from the nearest edge of the chromosome territory than loci from 11p15.5 and 11q13. This suggested that chromatin associated with increasingly gene-rich regions of DNA is progressively more decondensed. To investigate this further, I plotted the mean position of a locus along a chromosome territory radius against the gene density per Mb around that locus (Figure 6.9) for all loci studied. The scattered data points indicated a correlation and a trend line was fitted to these data using Microsoft Excel. Obviously the relationship will not be simply defined by a straight line, as there are many factors acting to determine the interphase position of chromatin, such as transcription activity, cell-cycle stage, differentiation state etc. However, there does appear to be a connection between gene density and the extent of extrusion of chromatin. Note that this graph confirms that gene-poor loci from all chromosomal locations studied (Figure 6.9, blue data points) are clustered with the chromosome interior. 11p13, previously classified as a moderately gene-rich region, is

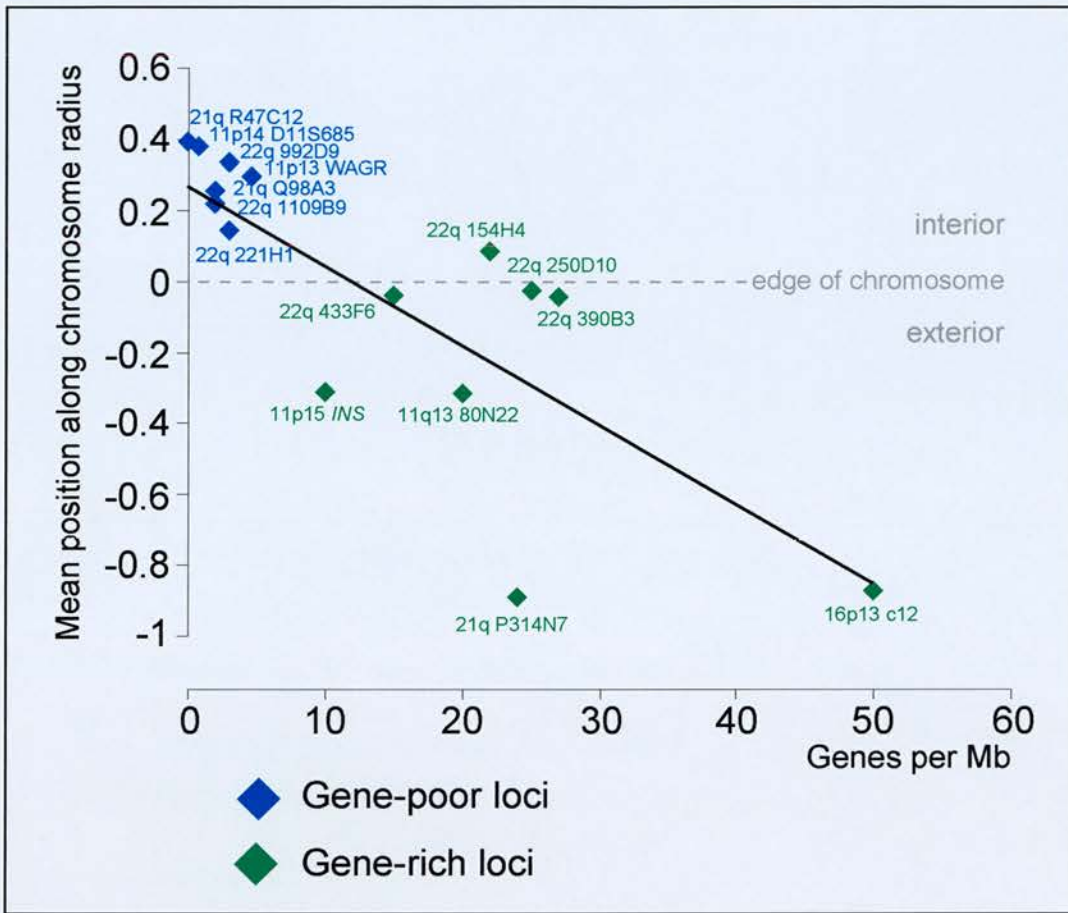


Figure 6.9 Correlation between gene density and intra-chromosome locus position.

The gene density per megabase of all the regions considered in this thesis were calculated. Mean locus positions relative to the nearest edge of native chromosome domains, normalised for the radius of the chromosome territory, were plotted against these calculated gene densities. A value of 0 on the y-axis represents the edge of the chromosome and negative values indicated that the mean locus position was outside of the chromosome territory, as defined by a chromosome paint in FISH. Loci corresponding to gene-poor domains (blue) were clustered within the chromosome interior, whereas gene-poor loci (green) were positioned at the chromosome periphery or outside of the territory. The best-fit line was determined using Microsoft Excel. It appears that there is a strong correlation: the more gene rich a region of interest, the further outside of the chromosome territory a locus mapped to that region was positioned.

obviously clustered with the gene-poor loci from HSA21q and HSA22q and also from the G-band 11p14. This indicates that inference that sequence from R-bands of metaphase chromosomes will be gene-rich is not necessarily true. In the future, a complete human genome resource will enable accurate classification of gene density direct from annotated sequence.

6.6 DISCUSSION

6.6.1 Gene density contributes to the organisation of chromatin within a chromosome territory

In this Chapter I have presented evidence in support of my hypothesis that gene density plays a role in the organisation of DNA within interphase chromosome territories across the human genome. Patterns of organisation suggested by the study of HSA11p (Chapters 3-5) were confirmed for other human chromosomes. My data also suggest a correlation between the gene density of a region of DNA and the mean localisation of a probe relative to its native chromosome territory: gene-poor loci are obviously clustered within the interior of the chromosome domain and gene-rich loci are more peripheral. In addition, I have presented data suggesting that the more gene-rich a significant stretch of sequence, the further outside of the chromosome domain it is (Figure 6.9). This relationship will obviously depend on the location of the probe within the loop or tail of chromatin extruded from the surface of the chromosome domain, especially where the loop is in the middle of a chromosome arm. The sub-telomeric gene-rich locus from HSA21q appears to be more external than might be expected from its similar gene density to loci from HSA22q. However, HSA21q is generally a gene-poor chromosome and this locus represents the major peak in gene density across the whole chromosome arm (Figure 6.7a), which may contribute to its extreme position.

It remains to be seen how frequently sequences localise at a distance from a chromosome domain, as defined by a chromosome paint, within the interphase nucleus, but I believe that it is not uncommon. A study of images of interphase nuclei showing juxtaposed chromosomes painted in different colours using in situ hybridisation procedures did not find evidence for intermingling of territories (Cremer *et al.*, 1996). However, cross hybridisation to other chromosomes and incomplete painting of chromosome domains due to suppression of repetitive sequences and the complex nature of chromosome paints, makes it difficult to accurately identify surfaces of chromosomes using this technique. I observed chromatin from

chromosomes 11, 16, 21 and 22 being extruded from the surface of the main body of interphase chromosome and this phenomenon has also been observed in study of HSA6 (Volpi *et al.*, 2000). Therefore, there may be some intermingling of chromosomes within the context of the interphase nucleus, indicated by the fact that large chromatin loops from one chromosome could well enter the 'territory' of another chromosome.

A previous light microscopy study of *in vivo* labelled chromatin has demonstrated that, while some borders between chromosome domains are well defined, fibre-like structures can be observed embedded in other chromosome territories (Visser and Aten, 1999). Therefore, small amounts of dispersed chromatin extending from different domains may be in contact, and this is perhaps important for the formation of chromosome exchanges in interphase (Chen *et al.*, 1996; Cremer *et al.*, 1996; Sachs *et al.*, 2000). A second study, using EM analysis of the distribution of *in vivo* labelled chromatin, observed that chromatin of adjacent chromosomes may be in contact in limited regions, thus indirectly implying chromosome-chromosome interactions (Visser *et al.*, 2000).

I believe that interphase chromosomes are not mutually exclusive and the true 'territorial' nature of their structure is not as defined as has been suggested in previous studies e.g. (Zirbel *et al.*, 1993; Kurz *et al.*, 1996). Loci positioned outside of bulk chromosome territories must be visualised independently of the chromosome domain using smaller specific probes, as the fine strands of chromatin on which they are situated are not discernible using a chromosome paint. This is one reason why intermingling between chromosomes has not been specifically reported. Such fine strands of chromatin are not visualised using chromosome paints; EM techniques using silver staining can only be used to suggest patterns of organisation, as there will always be background signals which cannot definitively be interpreted as true signal if they are at a distance from other signal-dense regions. The concept of a 'chromosome territory' is still applicable in that chromatin is not completely decondensed within the nucleus to the extent that fibres from all chromosomes intermingle. Contact between material from different chromosomes will depend on proximity of the territories in the nucleus, determined by their nuclear addresses (Boyle *et al.*, 2001), as few fibres will extend great distances across a cell nucleus. However, I believe that there is significantly more contact between 'halos' of fibres at the surface of adjacent chromosome domains than suggested by either ICD model (Cremer *et al.*, 1993; Cremer *et al.*, 2000).

Before we can understand structure-function relationships within chromosomes, the spatial organisation of chromatin within territories needs to be investigated. The concept of the

periphery of a territory or of a chromosomal subdomain makes sense only when the surface of that territory or subdomain can be defined with sufficient accuracy. Studies such as this are important in helping to determine the nature of chromosome territories, however, extensive analysis of the localisation of individual sequences relative to chromosome domains would be very labour-intensive and is approaching the limits of resolution of light microscopy.

The idea that gene-rich chromatin is preferentially localised towards the surface of a chromosome territory is in accordance with the observed converse distribution of AT-rich (gene-poor) isochores preferentially located within the interior of a chromosome territory (Tajbakhsh *et al.*, 2000). Tajbakhsh and co-workers also investigated gene-rich chromatin (represented by GC-rich DNA) and found that it to be distributed throughout chromosome territories, with no preference for positioning at the periphery. This observation appears to be in direct conflict with my data. However, as discussed in the publication, the H3 isochores used to represent gene-dense DNA consist of much more than the GC-richest DNA. Therefore, the very gene-richest DNA at the surface of a chromosome domain, as highlighted by my data, would not have been specifically detected using an H3 isochore probe. My study has specifically identified gene-dense and gene-poor regions using annotated sequence, rather than using indicators that approximate or infer gene-rich or gene-poor sequence. Using this more complete information, I have observed contrasting distributions of gene-rich and gene-poor chromatin within a chromosome territory.

My data suggests that gene-rich chromatin is preferentially positioned at the surface of a chromosome domain and gene-poor chromatin is more likely to be located within the territory interior. The study of individual genes is insufficient to determine patterns of organisation within interphase chromosomes; context of DNA sequence over many hundreds of kilobases of sequence is important in determining position within an interphase chromosome, rather than the nature of a sequence *per se*.

6.6.2 What mechanisms might cause decondensation of chromatin from the surface of a chromosome territory?

My data and previously published observations (Volpi *et al.*, 2000) suggest that a moderate level of transcriptional activity on a stretch of DNA is compatible with the sequences remaining within or close to the main body of a chromosome domain. Over and above that threshold, where many genes are being actively transcribed, chromatin may be extruded from the surface of a chromosome in order to achieve appropriate regulation of transcription

of all genes in that region. This might be an active extrusion regulated in an energy-dependent manner, or it might be a passive consequence of high levels of transcriptional activity perhaps loosening attachments of chromatin as part of the normal transcription process.

‘Constrained diffusional motion’ of the chromatin fibre was found to be independent of ATP hydrolysis (Marshall *et al.*, 1997). Similarly, the decondensation of chromatin with a high gene density might be a passive process. Passive chromatin decondensation may result from an increase in the acetylation or phosphorylation levels of chromatin, within a confined region of the chromosome territory. Therefore increased electrostatic repulsion between highly negatively charged histone tails might cause decondensation. Histone acetylases (Section 1.3.2) play a role in transcription regulation via increasing access of transcription factors to DNA through structural changes in nucleosomes or nucleosomal arrays. High levels of histone acetylation are also known to be inhibitory to binding of proteins that act to package chromatin in a condensed manner to render it inaccessible to transcription (Section 1.3.3). Decondensation may also be facilitated by loss of linker histones.

It is possible that the molecular-level activity of other enzymatic complexes may play a role in large-scale decondensation of gene-rich chromatin from the territory surface. Evidence suggests that histone acetylases act synergistically with ATP-dependent chromatin remodelling proteins (Section 1.3.6) and nucleosome remodelling by these complexes may be an upstream contributory factor facilitating large-scale decondensation of a chromosome territory. Similarly, the molecular-level process of transcription may contribute to large-scale alteration of chromatin structure. The ability of RNA polymerases to act as motors to move DNA has also been demonstrated and active positioning of chromatin by locally applied forces may thus turn out to be a widespread phenomenon. A direct role for transcription in large-scale chromatin structure was indicated by the observation that inhibition of RNA Polymerase II by DRB treatment of nuclei caused compaction of HSA19 (Croft *et al.*, 1999). Other candidates for a role in large-scale chromatin decondensation include motor proteins, such as actin and myosin, both of which are found in the nucleus (Milankov and De Boni, 1993), and also the SMC family of ATPases that have been proposed to act as motors in chromatin condensation (Section 1.8.1).

6.6.3 The decondensed nature of 16p13 correlates with previous descriptions of an open chromatin structure for α -globin genes

Chromatin from 16p13, where the α -globin genes are located, was a mean distance of $1.567^{\pm 0.190}$ μm outside of the HSA16p territory, whereas the β -globin locus was found within the HSA11p territory (Section 4.4), $0.299^{\pm 0.065}$ μm from the chromosome edge even though the β -globin locus itself is very gene-rich (5 genes in <50 kb, equivalent to over 100 genes per Mb). These data extend studies that have previously observed that the α - and β -globin clusters lie in radically different chromatin environments (Vyas *et al.*, 1992; Smith and Higgs, 1999). In non-erythroid cells, the β -globin cluster is contained within a late replicating, DNaseI insensitive, transcriptionally inactive chromatin domain but in erythroid cells becomes more decondensed (DNaseI sensitive), early replicating, and transcriptionally active. In contrast, the chromatin structure of the α globin cluster appears to be relatively open in all cells, contained within an early replicating, GC-rich isochore, each gene being associated with a methylation free CpG-rich island (Bird *et al.*, 1987). However, this is the first time that the loci have been observed to behave differently in terms of positioning relative to their chromosome territories

Similarly, the nuclear localisation of α - and β -globin genes in cycling lymphocytes, which do not express globin genes, has been compared to erythroblasts where both gene clusters are expressed (Brown *et al.*, 2001). In lymphoblasts, the sub-nuclear distribution of α - and β -globin genes was markedly different; β -globin loci, in common with several control inactive genes, were associated with pericentric heterochromatin in a high proportion of cycling lymphocytes, an association that has previously been correlated with repression of transcription (Brown *et al.*, 1997; Brown *et al.*, 1999). In contrast, α -globin genes were not associated with centromeric heterochromatin in lymphocytes, even though transcription of α -globin genes was repressed (Brown *et al.*, 2001), which correlates well with the open structure of the α -globin locus found in all cell types (Vyas *et al.*, 1992; Smith and Higgs, 1999) and the contrasting organisational features I found at the two loci. This suggests that the decondensed structure of chromatin from HSA16p13.3 is not related to active transcription from the α -globin locus and may be characteristic of the locus in all cell types. In erythroblasts, where α - and β -globin gene clusters are both expressed, both loci were positioned in areas of the nucleus discrete from heterochromatin (Brown *et al.*, 2001). It would be interesting therefore, to compare the position of α - and β -globin loci relative to chromosome territories in erythroblasts to see if there was any increased incidence in the positioning of the β -globin locus outside of the HSA11p territory.

CHAPTER 7

DISSECTING PATTERNS OF CHROMATIN ORGANISATION WITHIN THE NUCLEUS

7.1 INTRODUCTION

Previous Chapters have examined the organisation of chromatin within interphase chromosome territories. In this Chapter I will put this analysis into the wider context of the interphase nucleus: there is much evidence that larger-scale spatial distribution of chromatin within the nucleus can impinge on functional processes such as replication and transcription.

7.1.1 The nuclear interior has been associated with gene-rich chromatin

Within the nucleus, gene-rich chromosomes are preferentially positioned within the interior and gene-poor chromosomes are located towards the nuclear periphery (Croft *et al.*, 1999; Boyle *et al.*, 2001), suggesting that the nuclear interior is perhaps associated with facilitation of gene expression. In support of this idea, there is evidence suggesting that higher-order compartments of chromatin within the nucleus are built up by the alignment of polar chromosome territories (Ferreira *et al.*, 1997; Sadoni *et al.*, 1999). R band sequences (gene-rich) have been suggested to cluster within the nuclear interior forming an early replicating compartment that is transcriptionally competent and active (Sadoni *et al.*, 1999).

However, while this may be a trend it is not always true, as I have shown that probes mapped to the R-band 11p13 tend to be positioned near to the nuclear periphery within the bulk of the HSA11p territory. In addition, nascent transcripts and transcription factors have been visualised throughout the entire nucleus (Wansink *et al.*, 1993; Jackson *et al.*, 1993; Grande *et al.*, 1997; reviewed Szentirmay and Sawadogo, 2000). In wheat nuclei, genes are concentrated towards the telomeres, which are positioned at one side of the nucleus due to the adoption of a Rab1 configuration by chromosomes at interphase. However, localisation of

sites of transcription were not correlated with the positioning of coding sequence: the distribution of foci of BrUTP incorporation was found to be uniform across the nucleus (Abranches *et al.*, 1998). This suggests that either the BrUTP-containing transcripts can disperse through the nucleus very quickly, or that there is probably transcription occurring throughout the nucleus, not necessarily closely correlated with sites of high gene density.

7.1.2 Nuclear compartments associated with transcriptional repression

One region of the nucleus associated with repression of genes is the nuclear periphery. In *S. cerevisiae*, heritable inactivation of genes occurs at the silent mating type loci and at telomeres, which cluster near the nuclear periphery. Although heterochromatin is often found juxtaposed to the nuclear envelope in higher eukaryotes, it is not known whether this association actively facilitates repression of associated genes. Indirect evidence from *Drosophila* suggests that it may (Ye *et al.*, 1997; Gerasimova and Corces, 1998; Gerasimova *et al.*, 2000).

In *Drosophila*, association of the variegated brown eye colour gene locus with a heterochromatic environment accompanies position effect variegation (Csink and Henikoff, 1996; Dernburg *et al.*, 1996). These data suggest that transcriptional repression can be accomplished by relocating genes to heterochromatic late-replicating compartments of the nucleus (Talbert *et al.*, 1994; Csink and Henikoff, 1996; Dernburg *et al.*, 1996). Studies in human and mouse cells have shown that the silencing of certain genes is accompanied by their association with centromeric heterochromatin (Brown *et al.*, 1997; Brown *et al.*, 1999; Francastel *et al.*, 1999; Schubeler *et al.*, 2000). However, a recent study has shown that the Ikaros proteins initiate gene silencing through a direct effect on the promoter with localisation to pericentromeric heterochromatin likely to affect the action of Ikaros on regulatory sequences rather than causing the heterochromatinisation of the gene (Sabbattini *et al.*, 2001).

These studies underscore the importance of nuclear context for silencing, but how commonly is relocation of sequences within the nucleus directly related to transcriptional silencing and activation? Are there differences in the location of genes in terminally differentiated cell types where they are differentially expressed? Topological constraints on chromatin structure would prevent all silenced genes from localising to pericentric heterochromatin. Therefore, which genes are sequestered in these specific transcriptionally repressed domains and how

do the silencing mechanisms used for these genes differ to those used to silence genes not sequestered at heterochromatic sites?

In light of data presented for the organisation of chromatin within chromosome territories in previous Chapters and the evidence for larger-scale nuclear order described above, I decided to put my data in the wider context of the compartmentalised nucleus. One obvious question raised by data in the previous Chapters concerns whether the loops of chromatin extending from chromosome territories are directed a particular region or to specific sites within the nucleus. In addition, localisation of probes at a distance from their native chromosome domains may suggest intermingling between chromosomes.

7.2 ORIENTATION OF THE WAGR LOCUS WITHIN THE NUCLEUS

To examine whether the WAGR locus was specifically orientated with respect to the nuclear periphery, PAC probes spanning 1150kb were obtained (Figure 3.2, (Gawin *et al.*, 1999). Combining several biotin- and dig-labelled PACs in different combinations, in the absence of a chromosome paint, enabled orientation of an extended stretch of sequence to be examined within the nucleus (Figure 7.1). The results of manually scoring the relative positions of PAC signals with respect to the nuclear periphery are recorded in Table 7.1.

Table 7.1 Using PACs spanning the locus in 11p13 to determine whether WAGR is specifically orientated within the nucleus.

Two independent observers scored the fifty images for each PAC combination illustrated in Figure 7.1. Each nucleus contained two sets of PAC signals corresponding to the two homologous chromosome territories. The average of the two data sets is presented, expressed as percentage signals answering the set criteria. If an order of signal colours could not be discerned, or if both colour signals were equidistant from the nuclear periphery then they were scored as indistinguishable.

PAC combination (Figure 7.1)	% signals with discernible order		% signals indistinguishable	Total number of signals scored (2 per nucleus)
	Nuclear periphery-green-red	Nuclear periphery-red-green		
259N9 685F3 61M11 17K7	34.0	27.4	35.8	106
259N9 685F3 17K7	44.4	16.6	39.1	169
259N9 685F3 61M11 17K7	35.2	26.9	38.0	108

The analysis of PAC signal orientation with respect to the nuclear envelope suggested that the WAGR locus was consistently orientated within the nucleus. Distal sequence containing the *PAX6* gene was closer to the edge of the nucleus than the proximal end of the locus containing the *WT1* gene, at least in lymphoblastoid cells. The functional significance of this is not clear.

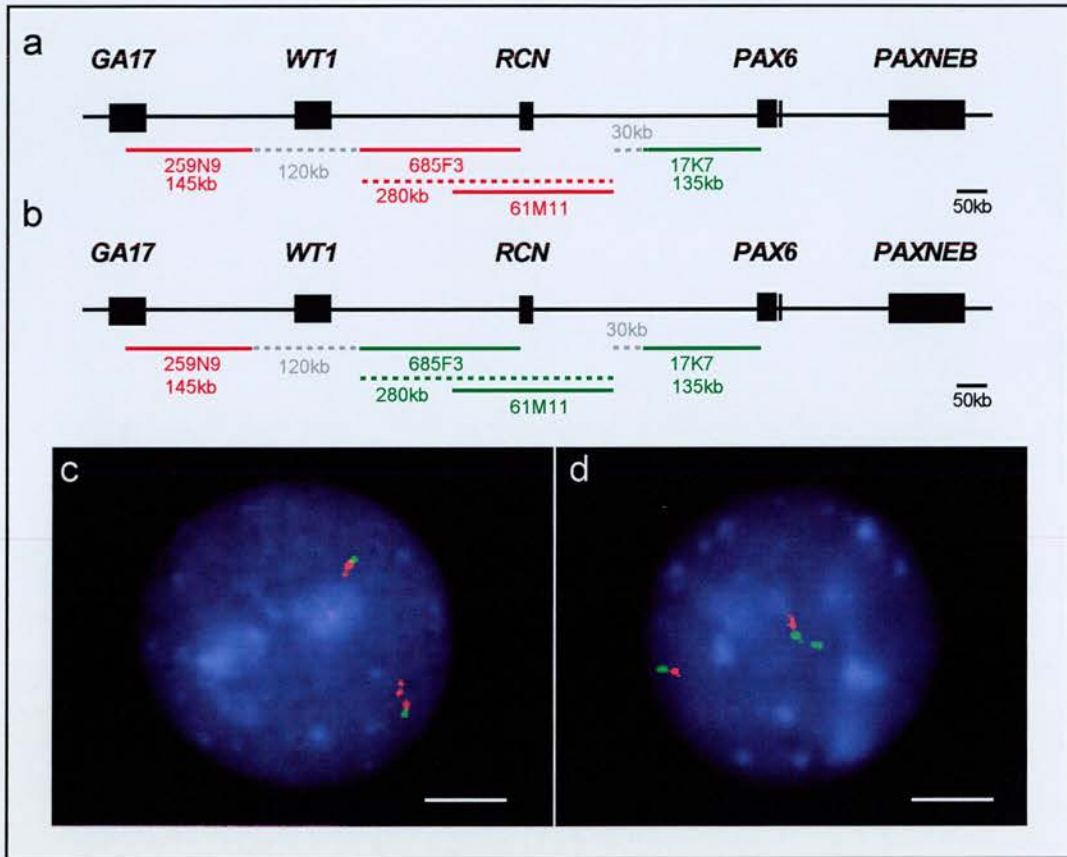


Figure 7.1 Determining the orientation of WAGR within the nucleus.

PACs mapped to WAGR were combined in FISH on lymphoblastoid nuclei in the absence of a chromosome paint. Differential labelling of the probes with biotin and Dig enabled the orientation of the locus to be examined within the nucleus. The positions of PACs used, and the colour in which they are visible, is depicted (a and b) and corresponding nuclear images are shown (c and d respectively). Analysis of fifty similar images confirmed that the *PAX6* gene appears to be orientated towards the periphery of the nucleus as compared to more proximal sequences from WAGR (Table 7.1). Scale bars represent 5µm.

7.3 DETERMINING THE NUCLEAR ORIENTATION OF GENE RICH CHROMATIN EXTRUDED FROM HSA11

To address whether there was a relationship between the relative nuclear positioning, of a locus and its native chromosome territory could be attributed to gene density of the locus, a script was written. This script directly compared the positions of the chromosome territory and the locus probe within concentric shells of the nucleus, moving from the nuclear periphery to the interior (Paul Perry) (Figure 7.2, Box 7.1) and was based on erosion analysis scripts previously described (Croft *et al.*, 1999; Bridger *et al.*, 2000; Boyle *et al.*, 2001).

Box 7.1 'Relative erosion' script for determining the relative 2D positions of a locus and its chromosome domain within the nucleus.

'Relative erosion' script: (Figure 7.2)

- i) The DAPI signal was automatically segmented from background (Figure 7.2b).
- ii) The area of this nuclear segment was determined, as was the centroid of the nucleus. The amount of FITC signal in the entire segment was calculated.
- iii) The FITC signal was segmented and a region of interest was manually defined around one of the chromosome territories (Figure 7.2c).
- iv) All FITC signal outside of the region manually defined in (iii) was removed. (Figure 7.2d).
- v) The grey value threshold for the chromosome signal was manually determined to a level which best represented the chromosome domain in the original image (Figure e and f).
- vi) The area of the segmented nucleus was eroded until the interior segment had an area four fifths that of the entire nucleus using the nucleus centroid co-ordinates calculated in (ii) as the centre of erosion. A line was drawn to define the outermost segment which therefore had an area one fifth that of the entire nucleus. The amount of FITC signal in the internal four fifths of the nucleus was calculated and subtracted from the total FITC signal calculated in (ii). This produced a value for the amount of FITC signal in the outermost shell and the percentage FITC signal in this shell was calculated as a proportion of the total FITC fluorescence in the nucleus. This step was repeated until the entire nucleus was divided into 5 segments of equal area and the % FITC signal in each shell was computed. Lines defining the boundaries of each segment were superimposed on the original image (Figure 7.2g).
- vii) The segmented nucleus was examined to determine in which shell(s) the locus signal (red) from each territory was positioned and the output data was exported to Microsoft Excel for manipulation (Figure 7.2h).

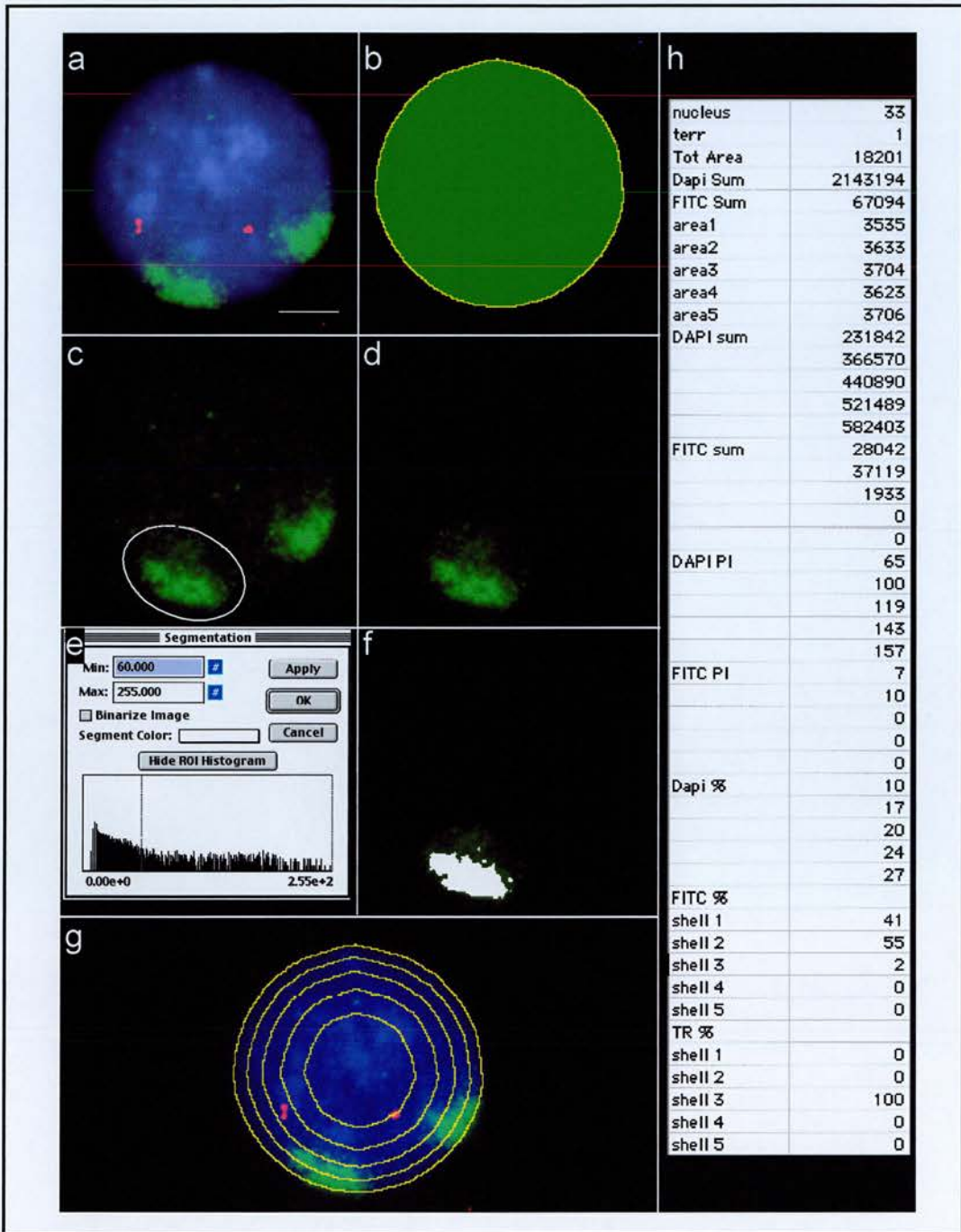


Figure 7.2 'Relative erosion' script for assessing the relative positions of the chromosome territory and a locus within the interphase nucleus.

a) A MAA-fixed lymphoblastoid nucleus after FISH with a HSA11p chromosome paint (green) and the *INS/IGF2* cosmid (red). The scale bar represents 5 μ m. The erosion script (Box 7.1) segments the DAPI signal (b) and records the area and centroid coordinates. The FITC signal is then segmented and a region of interest defined manually around the FITC signal representing the chromosome domain (c). The chromosome fluorescence (FITC) within this region is segmented and the intensity manually determined (e), to represent the original image (f). The area of the nucleus is then divided into concentric shells (1-5), of equal area, from the periphery of the nucleus to the centre and the proportion of FITC signal and DAPI fluorescence within each shell is calculated. The original image with the shell boundaries overlaid is then used to determine in which shell or shells the locus probe (TR signal) appears (g). The data output is exported to Microsoft Excel (h) and the mean % of each signal appearing in each shell is calculated after the script has been run on fifty such images. See Box 7.1 for more details.

Using the erosion script (Box 7.1), the relative nuclear positions of FISH signals representing the 11p or 11q chromosome domain (FITC) and loci from HSA11 (TR) could be directly compared. As a control, the erosion script was used to assess the relative positions of the HSA11p domain and the *WT1* gene from WAGR in 11p13 which was positioned one third of the way along the chromosome domain radius, using the same sets of fifty images used in Chapter 3 (Figure 7.3a). Consistent with previous analysis, HSA11p had a relatively peripheral distribution within the interphase nucleus (Boyle *et al.*, 2001), with almost 70% of chromosome signals located in the outermost three shells of the nucleus. As expected, the profiles for distribution of both the *WT1* locus signal and the chromosome domain within the nucleus were very similar (Figure 7.3b).

This script was then used to assess the nuclear distribution of loci from 11p15.5, using same images described in Chapter 4. Loci from 11p15.5 considered in this analysis were chosen so as to span the tail of chromatin determined to be extruded from the surface of the HSA11p chromosome territory (Figure 4.5). D11S12 represented the region at which the tail of chromatin appeared to re-enter the bulk 11p chromosome domain, the 11p telomere probe was at the distal end of the tail and D11S724 was in the middle. D11S12, had a nuclear distribution profile very similar to that of the chromosome domain, and very similar to those of the WAGR locus *WT1* (Figure 7.3b), as expected from its mean position at the surface of the 11p chromosome domain (Table 4.3). 47% of locus signals were positioned in the outermost nuclear shells, (shells 1 and 2). D11S724 and the telomere probe, however, had obviously different profiles with only 15% and 7% of signals respectively appearing in the two outermost nuclear shells (Figure 7.3b). Moreover, almost 70% of telomere signals and 60% of D11S724 signals appeared in the innermost two shells (shells 4 and 5). From this analysis it appeared that the tail of chromatin extruded from the surface of the interphase HSA11p domain was orientated specifically towards the centre of the nucleus.

To account for variation in the intra-nuclear position of the HSA11p chromosome territory between data sets, the mean % locus fluorescence signal per shell was normalised for the mean chromosome signal in that shell and a histogram plotted for each locus considered in erosion analysis (Figure 7.3c). For loci positioned within the chromosome domain a normalised value of approximately 1.0 would be expected, as locus signals would be coincident with chromosome territory signal. This was true for the 11p13 locus *WT1* and D11S12 from 11p15.5. However, both D11S724 and the 11p telomere loci were found 1.5 to 2.5 times more frequently in each of the innermost nuclear shells (shells 4 and 5) than would

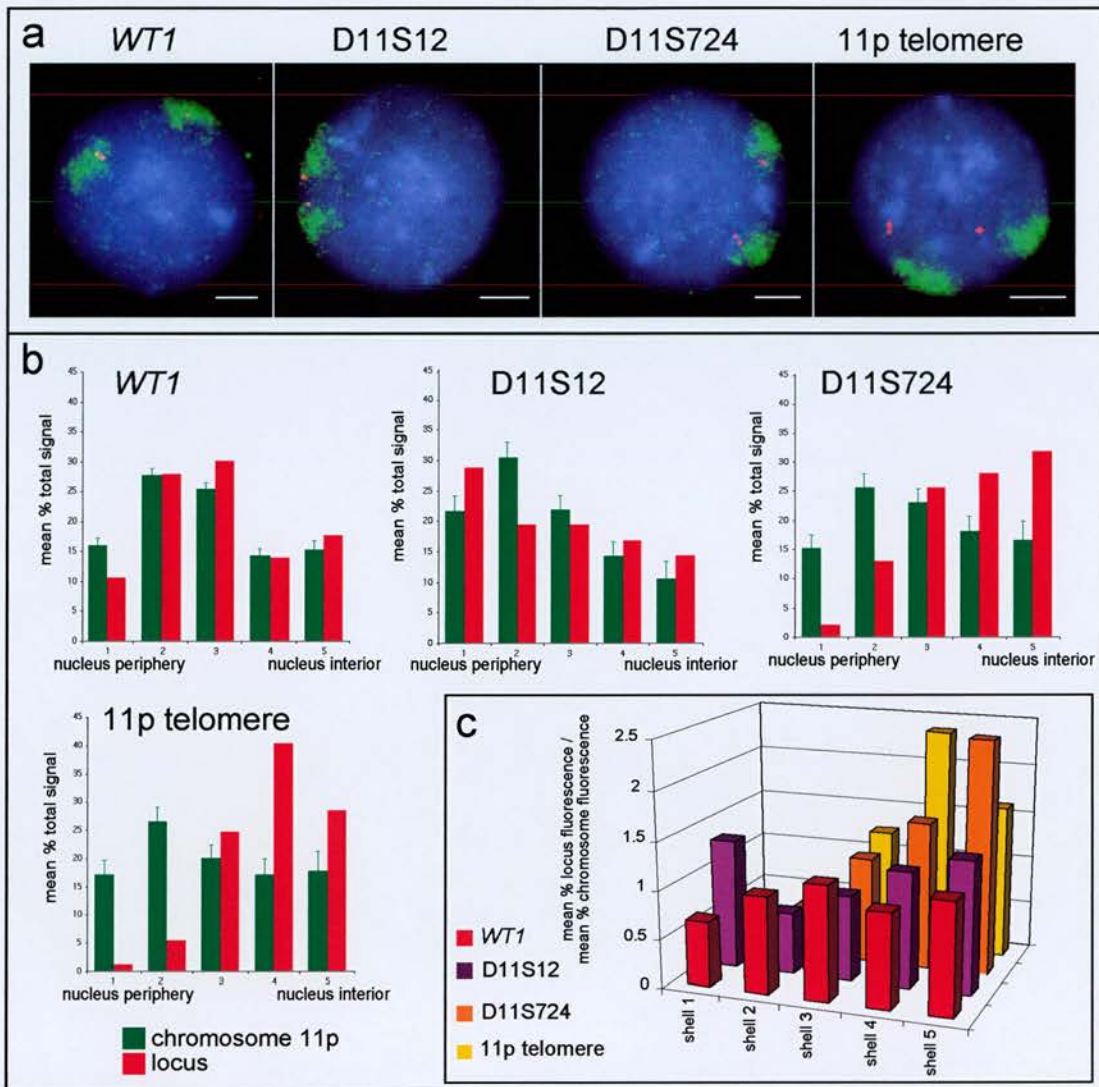


Figure 7.3 Erosion analysis shows that the tail of chromatin extruded from the surface of the HSA11p territory orientates towards the centre of the nucleus.

a) Lymphoblastoid nuclei representing the four image sets chosen for erosion analysis. Cosmid probes (red) were combined with a HSA11p paint (green) in FISH on MAA-fixed nuclei. Scale bars represent 5 μ m.

b) To determine the nuclear orientation of chromatin extruded from HSA11p (Table 4.3), three loci from 11p15 were chosen: D11S12 localised at the surface of the chromosome domain, D11S724 was midway along the tail of chromatin, and the telomere was the most distal locus. As a comparison, a WAGR locus data set (*WT1* from 11p13) was also analysed by erosion. Histograms were constructed for each locus depicting the mean % of chromosome (green) and locus (red) fluorescence in each of five concentric shells of equal area (1-5, outside to inside), into which the nucleus was divided. From these histograms it was obvious that *WT1* had a similar localisation pattern to the 11p territory, and a similar distribution was also found for D11S12. However, the tail of chromatin containing D11S724 and the telomere, was obviously displaced towards the centre of the nucleus compared to the chromosome territory. Error bars represent SEM.

c) Mean % locus fluorescence was normalised for mean % chromosome fluorescence in each shell, therefore the expected proportion of locus signal in each shell would be 1 if loci were colocalised with chromosome territories. Note, however, that both D11S724 and 11p telomere loci were found 1.5 to 2.5 times more frequently in the innermost shells (4 and 5) than would be expected from the distribution of the HSA11p territory.

be expected from the position of the HSA11p domain. In addition, these loci were rarely positioned in the outermost shell (shell 1) where up to 20% of the HSA11p chromosome is located. This confirmed that the negative mean positions found for D11S724 and the 11p telomere relative to the chromosome domain were due to an extrusion of the loci into the interior of the nucleus.

The erosion script (Box 7.1) was also applied to images depicting nuclei after FISH with probes to a gene-rich locus from 11q13 and the HSA11q territory (Figure 6.3). Erosion data from all three regions of HSA11 studied, 11p15.5, 11p13 and 11q13, were directly compared (Figure 7.4). Erosion analysis shows that a locus from gene-rich 11q13, like gene-rich 11p15.5, has a nuclear distribution which contrasts with that of its native chromosome arm domain (Figure 7.4b). In turn, the nuclear distributions of both gene-rich regions contrast with that of *WT1* from 11p13, which colocalises with the HSA11p territory. Normalising for mean chromosome fluorescence in each nuclear shell indicates that the 11q13 locus signal is orientated towards the centre of the nucleus compared to the HSA11q territory (Figure 7.4c). This data is consistent with that of 11p15.5, which is also orientated towards the centre of the nucleus compared to the HSA11p territory.

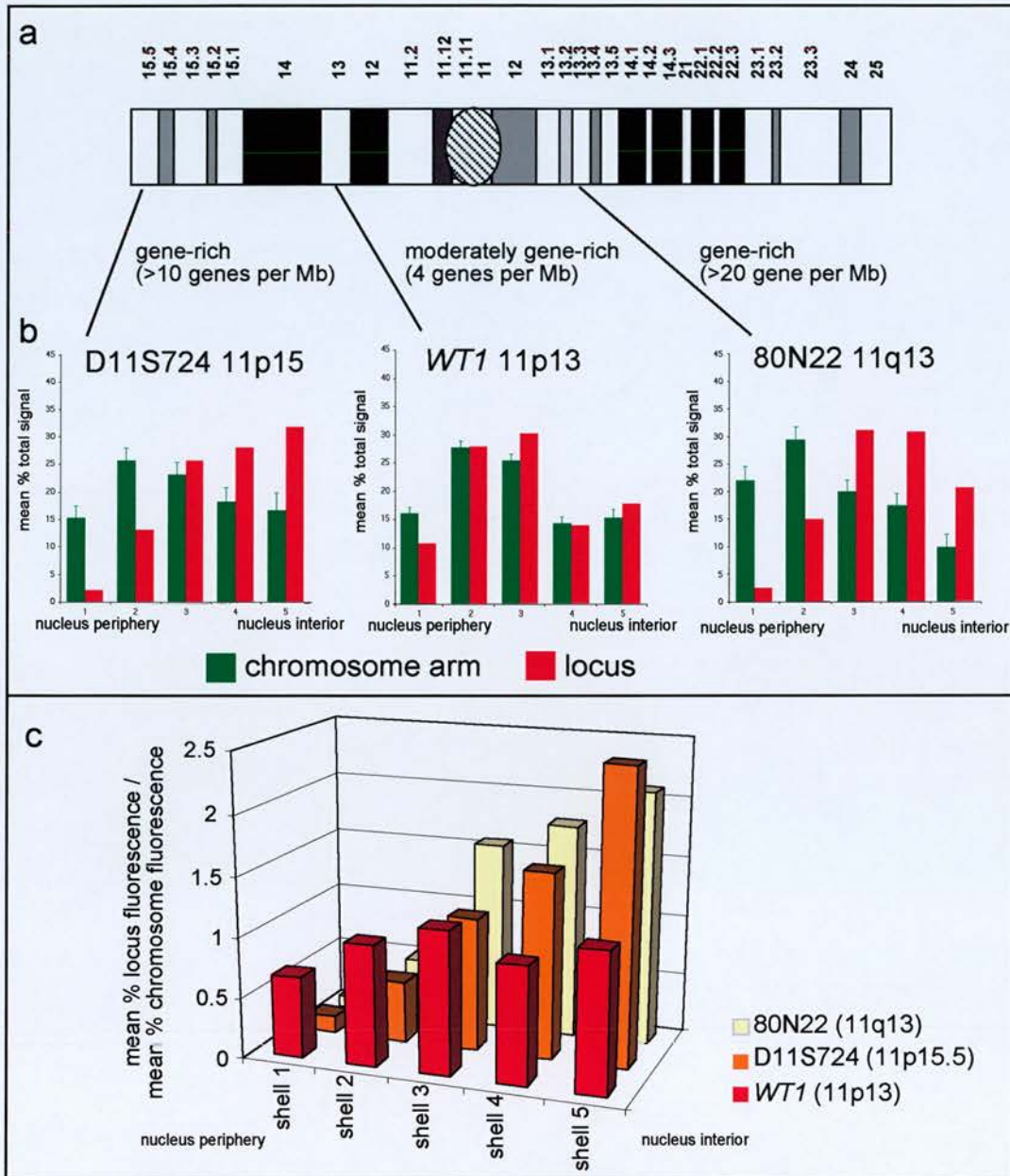


Figure 7.4 Erosion analysis comparing loci from three regions of HSA11p.

a) G-banded ideogram of HSA11.

b) Erosion analysis of three loci from HSA11p comparing the mean % fluorescence from the locus signal (red) and the chromosome arm (green) (11p or 11q) in five concentric shells of the nucleus (1-5, outside to inside). *WT1* from 11p13 has a similar distribution to the HSA11p chromosome territory. The gene-rich loci 11p15.5 and 11q13 however, have contrasting distributions compared to the chromosome domain, and are positioned preferentially towards the nuclear interior. Error bars represent SEM.

c) Mean % locus fluorescence was normalised for mean % chromosome fluorescence in each shell, therefore, if the locus was colocalised with the chromosome domain, the expected proportion of locus signal in each shell would be approximately 1. This is true for *WT1*, but for the gene-rich loci from 11p15.5 and 11q13, the locus signals were found 1.5-2.5 times more frequently within the innermost shells than would be expected from the distribution of the chromosome territories.

7.4 CHROMATIN FROM THE HSA11p15.5 CONSERVED REGION OF SYNTENY ON MMU7 IS ALSO ORIENTATED TOWARDS THE CENTRE OF THE NUCLEUS

I have shown that decondensation of chromatin containing sequence from 11p15.5 from the surface of the HSA11p chromosome territory is also characteristic of the spatial organisation of the murine region of conserved synteny in MMU7 (Section 4.11). Conservation of features of organisation of a chromosome domain between mammalian species must be functionally important and if the specific orientation of this tail of chromatin from HSA11p15.5 towards the centre of the nucleus is functionally significant I would expect chromatin from MMU7 to also be orientated towards the nuclear interior.

To investigate this, I performed erosion analysis on images depicting nuclei after FISH with probes to MMU7 and a locus syntenic with HSA11p15.5 (Figure 4.10). As a comparison, the erosion script (Box 7.1) was also applied to images depicting nuclei after FISH with an MMU2 paint and a probe to the *Wt1* gene from murine WAGR, which was rarely found at a distance from the MMU2 territory (Section 3.6, Figure 3.13). Erosion analysis shows that the *Wt1* gene from MMU2 has a nuclear distribution very similar to that of the chromosome territory (Figure 7.5a). In contrast, the nuclear distribution of the locus from MMU7 differs from the nuclear profile of the MMU7 chromosome territory itself (Figure 7.5b). Normalising for mean chromosome fluorescence in each nuclear shell indicates that the MMU7 locus signal is orientated towards the centre of the nucleus compared to the MMU7 territory (Figure 7.5c): the locus is more frequently found in interior nuclear shells than would be expected from the nuclear distribution of the MMU7 chromosome. This distribution is consistent with that of the region of conserved synteny in HSA11p15.5 (Figure 7.3), which is also orientated towards the centre of the nucleus compared to the HSA11p territory. In contrast, the nuclear distribution of the MMU2 locus is as would be expected from the distribution of the MMU2 territory (approximately 1.0 in each nuclear shell), similar to the erosion profile of the human region of conserved synteny in HSA11p13 (Figures 7.3 and 7.4).

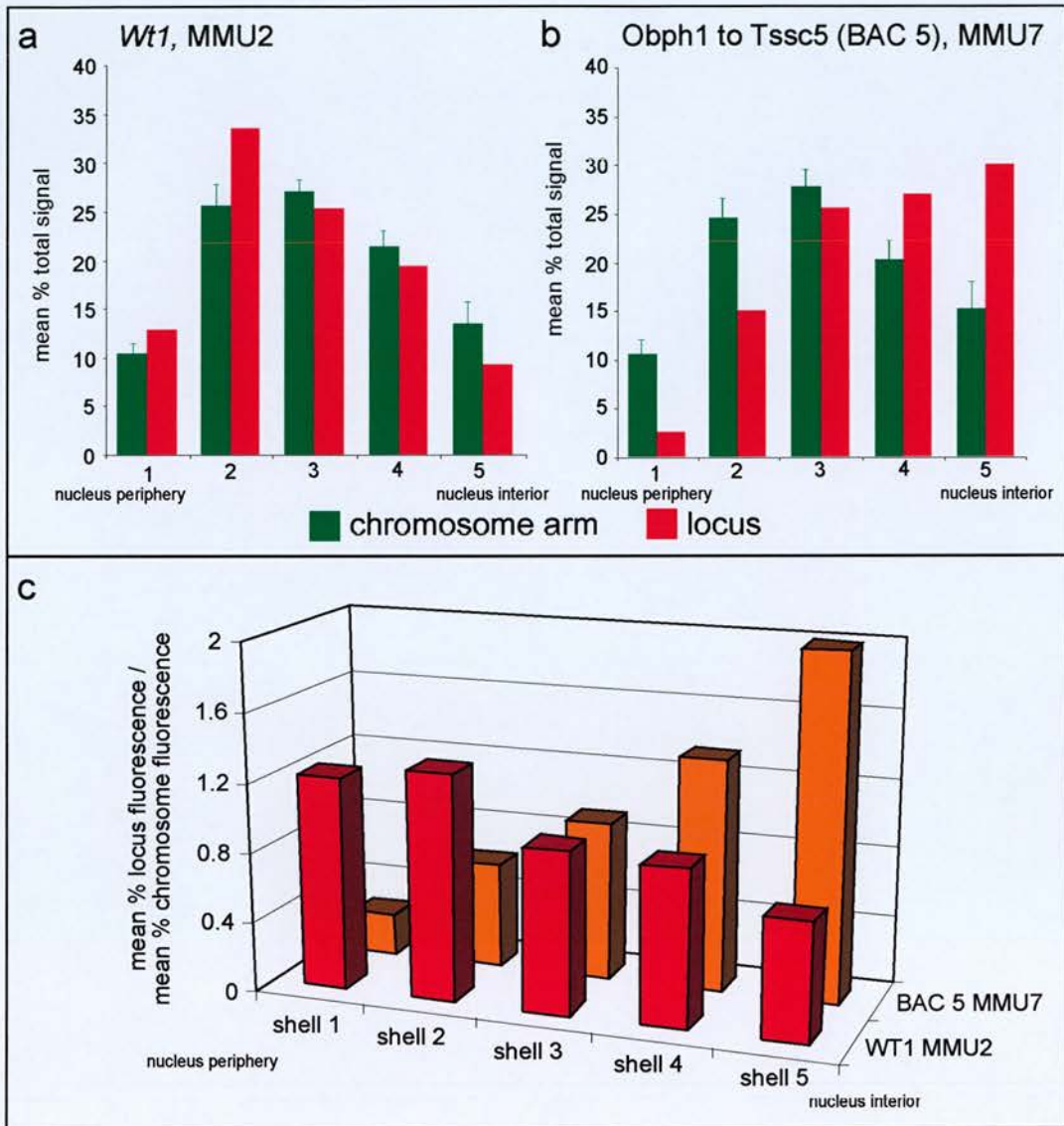


Figure 7.5 Erosion analysis of mouse chromosomes MMU2 (WT1) and MMU7 (BWS region).

a and b) Comparative erosion analysis of mouse loci. The *Wt1* locus is compared to MMU2 (a) and BAC 5 (containing genes from *Obph1* to *Tssc5* (see Figure 4.2)) is compared to MMU7. The mean % fluorescence from the locus signal (red) and the chromosome (green) in five concentric shells of the nucleus (1-5, outside to inside) are compared. Note the similar nuclear distributions of the chromosome territories, but the contrasting profiles of the two loci analysed.

c) The mean % locus fluorescence was normalised for mean chromosome fluorescence in each shell. If the locus colocalised with the chromosome domain, the expected proportion of locus signal in each shell would be approximately 1. By this method of analysis, the difference in nuclear distribution of the two loci is even more apparent. BAC 5 appears in the innermost nuclear shell almost twice as frequently than expected from the distribution of the MMU7 chromosome territory. Therefore, this locus is orientated towards the centre of the nucleus. The *Wt1* gene, however, is colocalised with the MMU2 chromosome territory, as demonstrated by a more even normalised locus distribution, approximately 1, in all five shells.

7.5 GENE-POOR REGIONS FROM CHROMOSOMES HSA21 AND HSA22 ARE ORIENTATED TOWARDS THE NUCLEAR PERIPHERY

HSA11 is peripherally positioned within the nucleus (Boyle *et al.*, 2001, and erosion analysis profiles of HSA11p and HSA11q, Figure 7.3) and I have observed gene-rich chromatin from this peripheral chromosome to be orientated towards the centre of the nucleus. In contrast, the entire territories of gene-rich chromosomes, such as HSA22, have been observed to be preferentially positioned within the nuclear interior (Croft *et al.*, 1999; Boyle *et al.*, 2001). So, how is such gene-rich chromatin, which I also observed to frequently be extruded from the main body of the HSA22q chromosome territory, organised within the nucleus?

The erosion script (Box 7.1) was also used to compare the intra-nuclear positions of gene-rich and gene-poor loci from HSA21q and HSA22q relative to the edge of the chromosome territories. I analysed the nuclear positions of one gene-rich and one gene-poor probe from each chromosome. The erosion script was run on the same sets of fifty used in Chapter 6. Data from these erosion script analyses are depicted in Figure 7.6. As previously shown (Boyle *et al.*, 2001), I found HSA22q and HSA21q to be positioned within the nucleus interior (HSA21 is not as peripherally positioned within the nucleus as would be expected from its low gene density, and this has been attributed to it being an acrocentric chromosome and therefore constrained by the nucleolus (Boyle *et al.*, 2001) (HSA22 is also acrocentric)).

A relative difference in the position of the gene rich and gene-poor loci could be discerned for each chromosome. The gene-rich loci (390B3 and P314N7) (Figure 7.6a and c respectively) appeared to be more internally positioned within the nucleus compared to gene-poor loci, yet the gene-rich loci still had negative mean values relative to the territories ($-0.075^{\pm 0.089}$ μm and $-0.680^{\pm 0.097}$ μm respectively). Therefore gene-rich chromatin extruded from the surface of HSA21 and HSA22 remained within the nucleus interior. This data continues the trend that gene-rich loci from 11p15.5 and 11q13, extruded from the peripherally positioned HSA11 territory, were orientated towards the centre of the nucleus (Section 7.3).

Normalisation of mean % locus signal for mean chromosome fluorescence in each nuclear shell shows that the distributions of the gene-rich loci are coincident with the territories for both chromosomes, perhaps with a tendency for the outermost shell (shell 1) to be under-

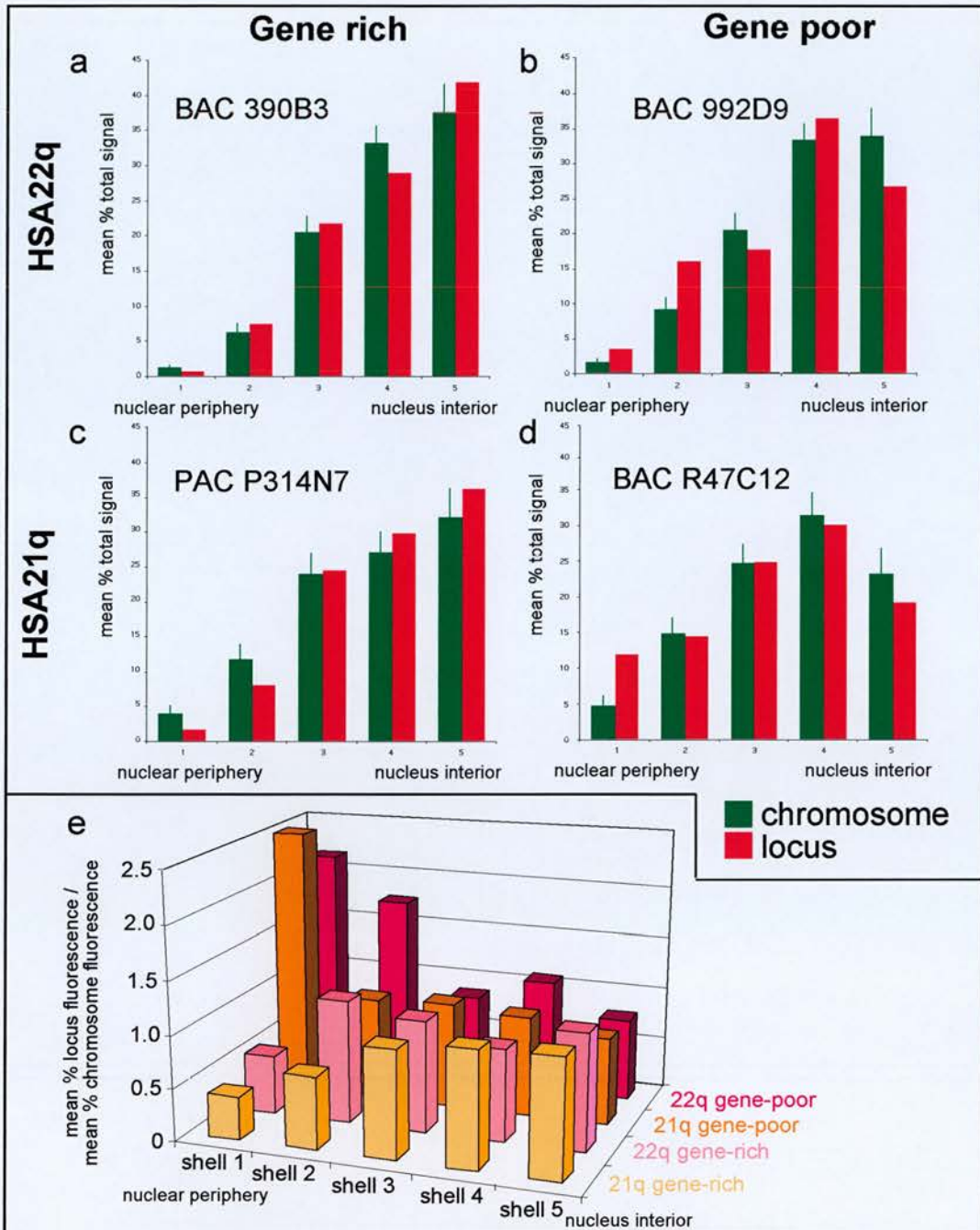


Figure 7.6 Erosion analysis shows that gene-rich loops of chromatin from HSA21q and HSA 22q orientate towards the centre of the nucleus.

a-d) Sets of fifty 2D FISH images depicting gene-rich (a and c) and gene-poor (b and d) locus probes combined with either HSA22q (a and b) or HSA21q (c and d) paints were subjected to erosion analysis (Box 7.1, Figure 7.2). Histograms were constructed showing the mean % distributions of signal corresponding to chromosome territories (green) and locus probes (red) within concentric shells of the nucleus (1-5, outside to inside). Probes corresponding to gene-rich regions are positioned more internally than the bulk of the chromosome territories to which they belong (a and c) and also more internally than probes corresponding to gene-poor regions (b and d). Error bars represent \pm SEM.

e) The mean % locus fluorescence in each shell was normalised for mean % chromosome fluorescence. The expected proportion of locus signal in each shell would be 1.0 if loci were coincident with chromosome domains in the nucleus. Note that both gene-poor loci studied from HSA21 and HSA22 were more often found in the outermost shells than would be expected from the nuclear distribution of the chromosome domains. Gene-rich loci colocalise with chromosome territories within the nuclear interior.

represented (Figure 7.6e). Interestingly, the gene-rich locus from HSA21 (P314N7), which has a dramatic localisation outside of the chromosome domain (Section 6.5), does not stand out in this normalised analysis. This is further confirmation that the extruded chromatin on which gene-rich loci frequently lie (resulting in negative mean values) must remain in the centre of the nucleus where the bulk of the HSA21q and HSA22q territories are positioned.

In contrast, gene-poor loci 992D9 and R47C12 (Figure 7.6b and d respectively) appear to have a peripheral position within the nucleus compared to the interior HSA21q and HSA22q chromosome territories respectively. This is again confirmed by normalising the mean % locus fluorescence signal per shell for the mean % fluorescence corresponding to the chromosome territories in that shell for both data sets (Figure 7.6e). Gene-poor loci from both HSA21q and HSA22q were positioned in the outer shell more than twice as frequently than would be expected from the nuclear distribution of the chromosomes (although the percentage of chromosome signal in this shell was very low in each case). This suggests that the nuclear periphery might be associated with gene-poor chromatin (Section 1.7.3.1).

Erosion analysis of HSA21q and HSA22q therefore supported the trend observed for gene-rich 11p15.5: that suggested that gene-rich chromatin was more internally positioned within the nucleus than gene poor chromatin. This trend was previously suggested by the generally position of GC-rich chromatin towards the centre of the nucleus (Sadoni *et al.*, 1999) and by the preference for gene-rich chromosomes to be positioned within the nuclear interior (Croft *et al.*, 1999; Boyle *et al.*, 2001).

7.6 CORRELATING THE AREA OCCUPIED BY INTERPHASE CHROMOSOMES WITH THEIR GENE RICHNESS

It has previously been noted that chromosomes of similar size but contrasting gene content occupy significantly different areas in the interphase nucleus (Croft *et al.*, 1999). I decided to see if this correlation between gene density and chromosome territory area, observed for chromosomes of similar physical size, extended to the rest of the human karyotype. Script analysis data calculates the areas and volumes of chromosome territories from which values for territory radii can be derived (Section 2.10). The mean chromosome territory radius was calculated for all chromosomes considered in this thesis (Table 7.2).

Table 7.2 Calculating a mean chromosome territory radius

A mean chromosome territory radius was calculated for each chromosome in different cell lines. Script output describes the area (2D) or volume (3D) of the territory in pixels or voxels respectively, from which a territory radius can be calculated (Section 2.10). From this, a value for each territory radius was estimated in microns, using calculated pixel to micron relationships: at 100x magnification and bin 2 capture, 100 pixels is equal to 13.4 μ m and at 100x magnification and bin 1 capture, 100 pixels is equal to 6.7 microns. These data were used to calculate mean distances of loci from the nearest edge of a chromosome territory in microns from normalised values throughout this thesis.

Cell line	Chromosome	Territory radius (pixels)	Territory radius (μ m)
2D			
FATO (Human)	HSA11p	13.928 \pm 0.043	1.866 \pm 0.006
CV581 (Human)	HSA11p	12.232 \pm 0.138	1.639 \pm 0.018
1HD (Human)	HSA11p	13.549 \pm 0.127	1.816 \pm 0.017
COV434 (Human)	HSA11p	13.217 \pm 0.114	1.771 \pm 0.015
CD5a (Human)	HSA11p	11.748 \pm 0.077	1.574 \pm 0.010
FATO (Human)	HSA11q	14.767 \pm 0.210	1.979 \pm 0.028
FATO (Human)	HSA15q	16.066 \pm 0.183	2.153 \pm 0.025
FATO (Human)	HSA18	14.633 \pm 0.156	1.961 \pm 0.021
FATO (Human)	HSA21q	5.670 \pm 0.052	0.760 \pm 0.007
FATO (Human)	HSA22q	11.616 \pm 0.114	1.557 \pm 0.015
E14 (Mouse)	MMU2	19.743 \pm 0.117	2.646 \pm 0.016
E14 (Mouse)	MMU7	17.139 \pm 0.189	2.297 \pm 0.025
3D			
1HD (Human)	HSA11p	20.623 \pm 0.220	1.382 \pm 0.015
1HD (Human)	HSA11q	19.762 \pm 0.733	1.324 \pm 0.049

Examination of data in Table 7.2 indicated that there was significant variation in size of chromosome domains. This obviously is to some extent determined by the physical size of chromosomes, but HSA21q and HSA22q are of a similar physical size (approximately 34 Mb) yet the radius of the HSA22q territory was twice that of HSA21q. The disparate areas of HSA21q and 22q in interphase were also apparent on examination of FISH images (Figures 6.8 and 6.11).

To directly examine if this variation in chromosome size correlated with gene densities of chromosomes, I compared the % of nuclear area occupied by each chromosome domain (in lymphoblastoid cells) to its gene density and normalised each measurement for the physical size of the chromosome (Table 7.3 and Figure 7.7).

Table 7.3 Correlation between the gene density and area of a chromosome territory.

For each chromosome considered in this study, the size (Mb) and number of genes present on each was taken from the NCBI human genome sequence database (http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/hum_srch?chr=hum_chr.inf&query) and the equivalent NCBI mouse genome database (http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/hum_srch?chr=mouse_chr.inf). Using these values, the number of genes per Mb of sequence was calculated for each chromosome domain, as was the % of the genome represented by each chromosome or chromosome arm, taking the size of the human genome to be 3200 Mb and the mouse genome to be 3100 Mb (<http://www.ncbi.nlm.nih.gov/genome/seq/MmProgress.shtml>). Using all script analysis data generated in previous Chapters, the % of nuclear area occupied by each chromosome domain in lymphoblastoid nuclei (human) and ES cell nuclei (mouse) was calculated and from this, the proportion of the nucleus occupied per Mb for each domain was computed.

Chromosome	Size (Mb) (NCBI)	% genome represented	Number of genes	% nuclear area	Genes per Mb (NCBI)	% nuclear area/ Mb (x100)
Human						
11p	56	1.8	328	2.7	5.9	4.8
11q	86	2.7	660	2.8	7.7	3.3
15q	80	2.5	545	3.5	6.8	4.4
16p	38	1.2	361	2.4	9.5	6.3
18	82	2.6	273	3.5	3.3	4.3
21q	34	1.1	225	1.9	6.6	5.6
22q	34	1.1	545	3.0	16.0	8.8
Murine						
2	209	6.7	750	6.6	3.6	3.2
7	156	5.0	789	5.2	5.1	3.3

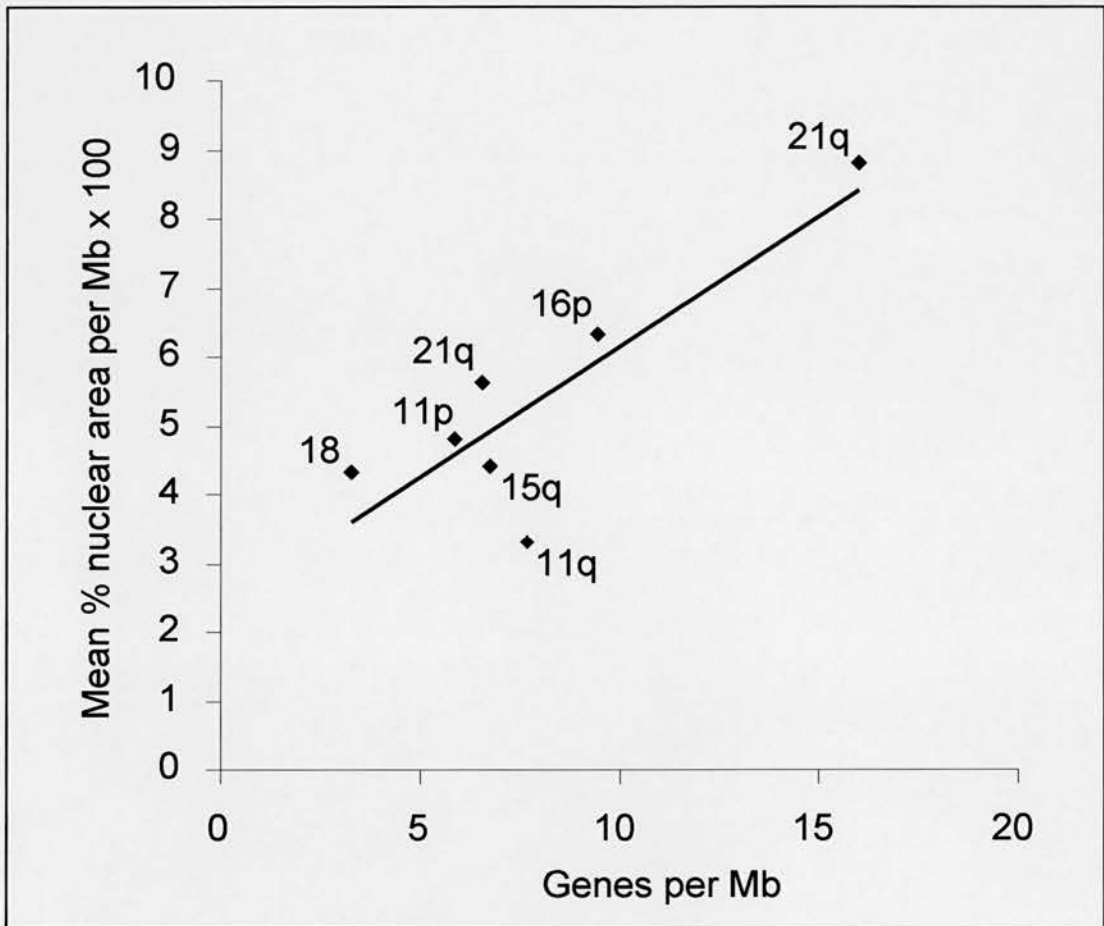


Figure 7.7 Correlation of gene density with the area occupied by the chromosome territory within the nucleus

The number of genes across each chromosome arm (taken from NCBI, (http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/hum_srch?chr=hum_chr.inf&query)) and the mean percentage of nuclear area occupied by each chromosome calculated by script analysis (Boxes 3.1 and 3.2) were normalised for the physical size of the chromosome (Mb) (same NCBI website). These two variables are strongly related and a straight line can be fitted to the data when presented graphically: the more genes per Mb a chromosome contains, the greater the proportion of the nuclear area is occupied by this chromosome.

This provided further evidence for the unusually open structure of chromosome 22q. I found it occupies 3% of the interphase nucleus, an area much greater than that expected for a chromosome representing only 1.1% of genome (Table 7.3). In contrast, the 34Mb (1.1% of the genome) of chromosome 21q only occupied 1.9% of the interphase nucleus.

My data suggest a strong correlation between gene-richness and the area occupied by a chromosome domain within the nucleus (Figure 7.7). The correlation is approximately linear and the equation of the best-fit line (Microsoft Excel) is $y=0.38x+2.33$ (where y is the % nuclear area per Mb x 100 and x is the number of genes per Mb). I noted that there was one obvious out-lying point in Figure 7.7, which represents HSA11q. This could perhaps be explained if the complexity of the paints used to distinguish the chromosome domains is considered. HSA11q is physically the largest chromosome domain included in this analysis (86 Mb) and, per megabase, occupies a small nuclear area relative to its gene density. The complexity of the HSA11q paint is therefore higher than that of paints for smaller chromosome domains and therefore may be less efficient at detecting the entirety of the HSA11q territory. Paints specific to smaller chromosomes with less sequence complexity may allow visualisation of a greater proportion of chromosome sequences in the paint.

These data provided an extension to observations previously published (Croft *et al.*, 1999), where gene-rich chromosome 19 was noted to occupy a significantly greater proportion of the interphase nucleus than gene-poor chromosome 18. The values obtained for the % of nuclear area occupied by HSA18 in the two independent studies differ. Croft *et al.*, found HSA18 to occupy 5.3% of the two-dimensional nuclear area whereas I found the area of HSA18 to be 3.5% of the nucleus. However, different chromosome paints were used in the different analyses and also different segmentation scripts were utilised. In the context of this study, the relationship described is based on direct comparison of values (Table 7.3) obtained using the same analysis scripts and paints from the same source (created by PCR amplification of micro-dissected chromosome arms (Guan *et al.*, 1996), Section 2.4.1). Therefore the data presented here are directly comparable.

It has previously been inferred that the distribution of genes across the murine karyotype is much more even than that of the human (Cross *et al.*, 1997). This conclusion was drawn from CpG island hybridisation in which no individual mouse autosome stood out as being either very island-rich or island-poor. Intrachromosomal variation in CpG island distribution was also less dramatic than that seen in the human. I compared the % of nuclear area

occupied by mouse chromosomes 2 and 7 (in embryonic stem cells), relative to its gene density, measured in genes per Mb, and normalised each measurement for the physical size of the chromosome (Table 7.3). It appears that the area of murine chromosome territories may be linked more closely to physical size than gene density. Both MMU2 and MMU7 occupy a similar % of nuclear area per Mb (3.2 and 3.3% respectively), even though MMU7 has a higher number of genes per Mb than MMU2 (5.1 compared to 3.6 for MMU2, calculated from data at NCBI: http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/hum_srch?chr=mouse_chr.inf). There is also a close correlation between the proportion of the genome represented by MMU2 and MMU7 (6.7% and 5.0% respectively, calculated from http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/hum_srch?chr=mouse_chr.inf) and the nuclear area occupied (6.6% and 5.2% respectively). I only have two values to comment on, but this data is intriguing and warrants further investigation.

7.7 DISCUSSION

7.7.1 The nuclear distribution of chromatin with respect to gene density

I found no difference in the localisation of silent genes and transcribed genes from the 1 Mb WAGR locus within the HSA11p chromosome territory (Chapter 3). Extension of the analysis of WAGR to look at its organisation within the nucleus suggested that in lymphoblastoid nuclei, WAGR was orientated with respect to the nuclear envelope: the *PAX6* gene is closer to the edge of the nucleus than the proximal end of the locus containing the *WT1* gene. The functional significance of this observation is not clear. The orientation does not appear to be connected with the transcriptional activity of the locus as the *RCN* gene is ubiquitously transcribed, and is found closer to the nuclear envelope (a region often associated with gene silencing) in lymphoblastoid cells, than genes not transcribed in this cell line, e.g. *WT1*. Distal 11p13 however, is only 550kb from very gene-poor 11p14. Therefore, perhaps this orientation of distal WAGR towards the nuclear periphery is due to a tendency for 11p14 to be positioned towards the nuclear envelope if the nuclear periphery does have an association with gene-poor chromatin. However, the *RCN* and *PAXNEB* genes are actively transcribed from within the HSA11p territory, which is positioned near the nuclear periphery. This suggests that localisation at the nuclear periphery is not a general mechanism to elicit transcriptional repression of individual genes. Chromosomes occupy

specific locations within the nucleus, with gene-rich chromosomes preferentially positioned towards the nucleus interior (Croft *et al.*, 1999; Boyle *et al.*, 2001) and chromosome 11 has a more peripheral nuclear position than might be guessed from its gene density (Boyle *et al.*, 2001). Interestingly, loci from 11p15.5 and 11q13, frequently observed outside of the HSA11 domain (Chapters 4-6), are more internal within the nucleus than the overall nuclear position of the chromosome territory (Figure 7.4). Therefore, gene-rich chromatin observed to be extruded from the surface of the HSA11 domain is specifically orientated towards the centre of the nucleus.

Chromosomes HSA21 and HSA22 both have a more internal localisation within the nucleus than HSA11. This is due, in part, to the constraints placed on acrocentric chromosome positioning within the nucleus by nucleolar association, and also perhaps to the fact that HSA22 is gene-rich (Boyle *et al.*, 2001). In further support of the idea that the nuclear interior is associated with gene rich chromatin, gene-rich loci extruded from the surface of HSA21q and HSA22q remain within the centre of the nucleus. Correspondingly, gene-poor loci from HSA21q and HSA22q were found more frequently at the nuclear periphery (in shell 1) (Figure 7.6), than would be expected from the internal nuclear distribution of the chromosome territories.

The internal orientation of gene-rich DNA from HSA11 may indicate a global orientation, perhaps suggesting that the nuclear interior is an area associated with active transcription in human cells. This conclusion correlates with previous studies of the general distribution of gene-rich and gene-poor chromatin within the nucleus (Croft *et al.*, 1999; Sadoni *et al.*, 1999; Boyle *et al.*, 2001) and the clustering of early replicating DNA within the nuclear interior and later replicating DNA at the nuclear or nucleolar peripheries (O'Keefe *et al.*, 1992; Sadoni *et al.*, 1999). Data presented in this thesis represents the first analysis of the orientation of large-scale subchromosomal domains within the nucleus, with regard to absolute gene density, rather than inferred gene density from sequence composition or replication timing. However, the fact that a probe mapped to the very gene-poor 7 Mb region within chromosome HSA21q (R47C12) frequently co-localised with the HSA21q territory in the interior of the nucleus suggests that gene-poor DNA is not excluded from the nuclear interior.

MMU7 is 1.4x as gene dense as MMU2, but there appears to be no preference for the more gene-rich MMU7 to be localised towards the nuclear interior within ES cells. As discussed

(Section 4.11.4), gene density of mouse chromosomes, inferred by CpG island mapping, is much more even throughout the whole karyotype. However, a locus from gene-rich distal MMU7 still appears to be orientated towards the centre of the nucleus compared to the overall nuclear distribution of MMU7 (Figure 7.5c), suggesting that the nucleus interior may be associated with regions of chromatin that are especially transcriptionally active or compared to the rest of the genome. Alternatively, there may be a different mechanism acting to position chromatin from HSA11p15.5 and the murine region of conserved synteny in the nucleus interior. For example, association with particular genes from this region with a particular nuclear body that could be concentrated in the nuclear interior.

7.7.2 Is there intermingling between chromatin from different chromosomes?

Observation of loci positioned apparently at a distance from their native chromosome territories suggests that regions of intermingling of chromatin from different chromosomes may exist. Previous studies using FISH have not found direct evidence for extensive intermingling (Cremer *et al.*, 1996). However, only a few contact regions can be analysed using chromosome paints, and cross hybridisation to other chromosomes and incomplete staining by suppression of repetitive sequences make it difficult to accurately identify surface areas between chromosomes. Higher resolution analysis using incorporation of BrUTP and EM analysis combined with immuno-gold labelling of BrUTP showed that chromosome domains are local regions of mostly condensed chromatin with some dispersed chromatin extending into an interchromatin space almost devoid of DNA (Visser *et al.*, 2000). But Visser *et al.*, also found that chromosome domains are not necessarily bordered at all sides by the interchromatin space: and two domains may touch each other without a visible morphological separation. This implies that adjacent chromosomes may be in contact in limited regions. Furthermore, my data suggests that it might not just be chromatin from adjacent chromosomes that might intermingle.

7.7.3 The nuclear area occupied by a chromosome territory is related to gene density

The nuclear volume or area occupied by a few interphase chromosomes has previously been examined (Eils *et al.*, 1996; Croft *et al.*, 1999). In one study, a connection between the gene content of a chromosome and the area/volume occupied by that chromosome within the

interphase nucleus was suggested (Croft *et al.*, 1999). Gene-dense chromosome 19 occupied a significantly greater area of the interphase nucleus than gene-poor chromosome 18, even though the DNA content of the two chromosomes is very similar (77 Mb and 79 Mb respectively). Data presented in this Chapter extend this observed correlation between gene density and area of interphase chromosomes to other chromosomes of the human karyotype. These data correlate with a previous study that demonstrated regional differences in large-scale chromosome structure during interphase, where more openly configured chromatin corresponded to R-bands (Yokota *et al.*, 1997). Data presented here (Chapters 4-6) suggest that gene-rich chromatin often is frequently positioned outside of a chromosome territory, in strong correlation with the published observation that gene-rich chromosome arms have a more diffuse nature and adopt more open, looping configurations.

A study of the organisation of active and inactive X chromosomes revealed that volumes occupied by each differed by less than a factor of two within a given cell (Eils *et al.*, 1996). However, the two X territories differed significantly in shape and surface structure. This suggests that, while the DNA sequence (and by inference, gene density) itself might dictate the proportion of the nucleus occupied by a chromosome domain, epigenetic or active processes such as transcription may alter the nature of this volume, perhaps causing extrusion of sequences as I observed to be common to gene-rich regions of DNA (previously noted in (Volpi *et al.*, 2000). Further evidence for this idea was provided by a second study of X_a and X_i chromosome territories, which found that, although the volumes of both X territories were similar, 3D distances between specific gene loci were significantly smaller in X_i compared to X_a territories reflecting different territory shapes (Dietzel *et al.*, 1999).

Inhibition of RNA polymerase II activity caused a decrease in nuclear size of gene-rich HSA19 (Croft *et al.*, 1999), indicating that transcription has a role in the nuclear organisation of chromosome domains. It is also interesting to note that two adjacent homologous genes were more often involved in recombination events when they were actively transcribed (Nickoloff, 1992), indicating that there may have been some alteration in the chromatin architecture on induction of transcription. I cannot comment on the role of transcription played in determining the degree of condensation of a chromosome domain but this would be an interesting extension to the project.

In contrast to data obtained for human chromosomes, I found the nuclear areas of mouse chromosomes to more closely reflect their physical size, rather than their gene density (Table

7.3). I have only data for two chromosomes to compare, but the correlation for the two studied so far is compelling. It is interesting to note in light of this that loci from the syntenic region to gene-rich HSA11p15.5 on MMU7 were found to frequently localise at a distance from the MMU7 domain, in a manner comparable to the behaviour of the human loci. While mouse CpG islands were found to be more uniformly distributed than those of the human, there was evidence for enrichment in R-band compartments of the mouse genome (Cross *et al.*, 1997). Examination of known regions of extensive conserved synteny between human and mouse showed that CpG island density was also conserved. Therefore, even though gross organisation of murine chromosome territories within the interphase nucleus does not appear to be conserved compared to organisation within human nuclei the fine detail organisation of syntenic, unusually gene-rich loci (rarer in the mouse than in the human) does appear to be conserved. This conservation points to the unusual structure adopted by chromatin from 11p15.5 as being of functional significance.

CHAPTER 8

THESIS OVERVIEW AND DISCUSSION OF FUTURE DIRECTIONS

8.1 THE HUMAN GENOME—FROM PRIMARY SEQUENCE TO 3D ORGANISATION

With the recent completion of the Human Genome Sequencing project draft sequence (Lander *et al.*, 2001; Venter *et al.*, 2001), we are better placed to examine genes in the context of their surrounding DNA in terms of well-characterised linear sequences. However, we know that the packaging of these sequences into chromatin, and subsequent recruitment of chromatin modifying proteins that modulate chromatin assembly plays a direct role in the control of gene expression, as does the spatial positioning of these sequences in the nucleus. Despite the rewards of intensive efforts to decode the linear structure of complex genomes, still relatively little is known about the three-dimensional genome organisation within the cell nuclei of higher organisms.

Many studies have focussed on the large-scale organisation of the compartmentalised nucleus. Chromosomes, and chromosome arms, have been described as occupying discrete territories in the nucleus (Cremer *et al.*, 1993; Dietzel *et al.*, 1998), and evidence is emerging that these domains are positioned in a non-random manner. Gene-rich chromosomes are more centrally located within the nucleus compared to gene-poor chromosomes that localise more towards the nuclear periphery (Croft *et al.*, 1999; Boyle *et al.*, 2001). However, the spatial organisation of the genome within chromosome territories is not well documented; a major question regarding spatial organisation of sequence within the nucleus is whether large-scale chromatin organisation above the level of the 30 nm solenoid fibre facilitates selective transcription, gene silencing, mRNA processing, or transport of RNA.

The relationship between the positioning of individual genes and their transcriptional activity has been studied within the context of the nucleus as a whole. In such studies, the nuclear

periphery has been associated with repression of genes. In *Saccharomyces cerevisiae*, heritable inactivation of genes occurs at the silent mating type loci and at telomeres, which cluster near the nuclear periphery. Although heterochromatin is often found juxtaposed to the nuclear envelope in higher eukaryotes, it is not known whether this association actively facilitates repression of associated genes. Indirect evidence from *Drosophila* suggests that it may (Ye *et al.*, 1997; Gerasimova and Corces, 1998; Gerasimova *et al.*, 2000).

Other studies in *Drosophila* suggest that association of a locus with a heterochromatic environment accompanies position effect variegation; transcriptional repression can be accomplished by relocating genes to heterochromatic late-replicating compartments of the nucleus (Talbert *et al.*, 1994; Csink and Henikoff, 1996; Dernburg *et al.*, 1996; Platero *et al.*, 1998). In human and mouse cells, the silencing of certain genes is accompanied by their association with centromeric heterochromatin (Brown *et al.*, 1997; Brown *et al.*, 1999; Francastel *et al.*, 1999; Lundgren *et al.*, 2000; Schubeler *et al.*, 2000).

These studies underscore the importance of nuclear context for gene silencing, but it is not clear how general the mechanism of relocation of sequences within the nucleus in relation to transcriptional activity actually is. Topological constraints on chromatin structure would prevent all silenced genes from localising to pericentric heterochromatin and it has recently been found that regulated expression of α - and β -globin gene clusters occurs in different nuclear environments in primary haemopoietic cells (Brown *et al.*, 2001). The localisation of silent genes with centromeric heterochromatin had previously found to be important in lymphocyte differentiation and, as expected, in cycling lymphoblasts, the non-expressed β -globin loci were found associated with pericentric heterochromatin in a high proportion of nuclei. However, in the same cells, the non-transcribed α -globin genes were not associated with centromeric heterochromatin. Therefore, there must be other silencing mechanisms used for genes not sequestered at heterochromatic sites, and it is not known if nuclear compartmentalisation also plays a role in these alternative mechanisms of transcription control.

8.2 DISCUSSION OF MY OBSERVATIONS IN RELATION TO CURRENT MODELS OF NUCLEAR ORGANISATION

Models of nuclear and chromosome territory organisation with respect to the positioning of active and inactive genes have implications for the spatial organisation of genomic DNA in the nucleus both at the large-scale (whole chromosome and chromosome band) level and at a more local (megabase) level. To address these questions I have used FISH to analyse the intra-nuclear and intra-territorial organisation of contiguous stretches of the human genome that differ in gene-content and activity, and the syntenic regions in the mouse.

8.2.1 All genes are not preferentially positioned at the surface of a chromosome territory

It has been proposed that transcription, RNA processing and RNA transport might occur in a space between territories called the interchromosomal domain (ICD) compartment (Zirbel *et al.*, 1993; Cremer *et al.*, 1993; Bridger *et al.*, 1998) (Figure 1.7). Accumulations of specific RNAs have been seen at the border of the chromosome territory from which they originated, and components of the splicing machinery have also been reported to be outside of chromosome territories (Zirbel *et al.*, 1993; Clemson and Lawrence, 1996). Therefore, genes were proposed to be preferentially located at the surface of interphase chromosomes (Kurz *et al.*, 1996). On the X chromosome, an active gene has been found in a more peripheral location within the chromosome territory than its inactive counterpart (Dietzel *et al.*, 1999). Over larger regions, it has been shown that the gene-rich MHC complex lies on large chromatin loops that extended away from the surface of the chromosome 6 territory that is detectable with a chromosome paint (Volpi *et al.*, 2000). An extreme interpretation of these data is that most (active) genes lie on the surface of chromosome territories.

However, both early- (gene-rich) and late-replicating (gene-poor) DNA are distributed throughout chromosome territories, implying that gene-rich DNA can be located within the chromosome interior (Visser *et al.*, 1998), and GC-rich (gene-rich) DNA sequences show no preferential localisation within chromosome territories (Tajbakhsh *et al.*, 2000). As sites of transcription also localise throughout chromosome territories and poly(A) RNA is not excluded from chromatin domains (Abranches *et al.*, 1998; Verschure *et al.*, 1999), there seems to be little requirement for genes to locate to the surface of a chromosome to facilitate transcription.

Furthermore, TEM localisation of uridine or BrUTP incorporation has shown heavy labelling at the edge of condensed, large-scale chromatin domains rather than at the surface of chromosomes *per se* (Fakan and Nobis, 1978; Wansink *et al.*, 1996). These studies suggest that locally compacted and unfolded regions within an interphase chromosome form distinct subdomains, and that chromatin folding is organised in such a way that transcriptionally active DNA is at the surface of large-scale chromatin fibres. The ICD compartment has thus been extended to include space between surfaces of subchromosomal compartments, the inter-chromatin domain (Cremer and Cremer, 2001) (Figure 1.8).

I have shown that the distribution of chromatin within a chromosome territory is not dependent on whether the DNA sequence examined represents a gene or not. Data presented in this thesis suggest that study of individual genes is insufficient to determine patterns of organisation; the context of a sequence is important and the overall nature of the surrounding DNA should be considered when examining trends of organisation within an interphase chromosome. Using 2D and 3D analysis of the WAGR locus from HSA11p13 I found that there is no difference in the overall intra-chromosomal position of probes spanning the locus, whether they represent active genes, non-transcribed genes, or intergenic sequences (Chapters 3 and 5). In addition, I have shown that ubiquitously expressed *RCN* and *PAXNEB* genes are frequently positioned well within the HSA11p chromosome territory in the four cell lines examined (Chapters 3-5). Therefore, genes can be transcribed from within a chromosome domain and implies that proteins, e.g. RNA Pol II and transcription factors, are not excluded from within the interior of chromosome territories as was suggested by the original ICD model of nuclear organisation (Zirbel *et al.*, 1993). My data are consistent with FRAP analysis, suggesting that proteins can move readily and rapidly throughout the nuclear volume (Phair and Misteli, 2000). In addition, transcription factors have been demonstrated to be distributed more or less homogeneously throughout the nucleoplasm, occupying numerous small domains with no obvious exclusion zones (Grande *et al.*, 1997). However, since many of the transcription factor-rich nuclear domains were not actively involved in transcription (Grande *et al.*, 1997), this still leaves open the question of whether active transcription is compartmentalised within the nucleus, perhaps within the nuclear interior (Section 8.1.4).

8.2.2 Is the chromosomal or nuclear position of a gene dependent on transcription status?

When the WAGR locus was compared between cell lines in which transcription from tissue-specific genes *PAX6* and *WT1* was up-regulated, intra-chromosomal organisation was unaltered (Chapter 3). Hence, I have concluded that small-scale changes in transcription do not alter the intra-chromosomal position of the WAGR locus; rather than large-scale repositioning within the nucleus, molecular-level chromatin remodelling is likely to control small-scale changes in gene expression (Section 1.3).

Transcription status has been implicated in the control of large-scale chromosome architecture. However, most gross alterations in spatial organisation previously observed were in the context of large-scale changes in transcription, not just of a single gene, but of many genes within the same region of sequence or entire chromosome domains (Park and De Boni, 1998; Tumber *et al.*, 1999; Dietzel *et al.*, 1999; Volpi *et al.*, 2000). Protrusions of chromatin sequences from the HSA17 chromosome territory containing amplified *ERBB-2* genes were observed in human breast cancer cells (Park and De Boni, 1998). The incidence of similar protrusions from the HSA6 territory containing the MHC locus genes increased after IFN- γ treatment of primary fibroblast cells which results in specific activation of transcription of many genes in that region (Volpi *et al.*, 2000). Thirdly, Dietzel *et al.*, observed a preference for a particular gene on the X chromosome to be more peripherally located within the territory of the active X than the inactive X. Finally, transcriptional activation of a 90 Mb heterochromatic region containing 256 copies of the lac operator sequences in CHO cells was associated with large-scale chromatin decondensation (Tumber *et al.*, 1999). However, this decondensation was seen in the presence of a transcription-activating protein, even if transcription itself was inhibited.

One example where the position of genes with respect to specific sites in the nucleus has been observed to correlate with transcription activity is in the differentiation of B cells (Brown *et al.*, 1997; Brown *et al.*, 1999). Genes differentially expressed during different stages of B cell maturation show a correlation of expression status with Ikaros protein association in pericentromeric heterochromatin-containing regions of the nucleus. Ikaros association is evident only in cells in which those genes are not transcribed. However, it is not known how the movement of such loci to pericentromeric heterochromatin is achieved within territorial constraints of chromatin organisation, and whether the intra-chromosomal position of the

developmentally regulated loci actually changes to achieve this mode of transcriptional control.

In addition, the transcriptional activity of adjacent genes to those differentially transcribed has not been investigated. It would be interesting to know over how great a linear sequence distance association with pericentric heterochromatin is able to effect transcriptional repression. It has been suggested that juxtaposition to repetitive DNA is not incompatible with expression, but that a strong activator is required to overcome the repressed state (Lundgren *et al.*, 2000). These observations raise the questions of which genes are sequestered in these transcriptionally repressed domains, and how such loci are configured with respect to their native chromosome territories; topographical constraints on chromatin structure would prevent all silenced genes from localising to pericentric heterochromatin.

Investigation into whether a particular immunoglobulin (Ig) gene occupies a predetermined position within the nucleus found that heavy- and light-chain genes in transformed human B-lymphoid cells and fibroblasts are differentially and non-randomly distributed in different nuclear sub-volumes independent of the activity of the genes or the cell type (Parreira *et al.*, 1997). Both κ and λ light chain genes are located on the acrocentric chromosome HSA22, the short arms of which are integral components of the nucleolus. Therefore, the fact that genes examined in this study were found in similar nuclear sub-volumes within nuclei from cell types where they were differentially expressed, could be attributed to the location of these genes on chromosomes constrained by nucleolar association. However, I have observed differences in the nuclear organisation of loci from the acrocentric chromosomes, HSA21 and HSA22, so if major differences in κ and λ gene positioning had existed between cell lines, they should have been perceptible.

In contrast, study of the intra-nuclear position of Ig genes in primary mouse splenic B cells showed that endogenous IgH, and the light-chain genes κ , and λ , localise to different subnuclear environments after activation, which results in mono-allelic expression of a single IgH and a single IgL gene (Skok *et al.*, 2001). 95% of a population of B cells were found to express Ig κ -derived light chains, and only a minority (5%) expressed Ig λ -derived light chains. Corresponding closely with expression status, in approximately 90% of these cells, one κ allele was found closely associated with centromeric heterochromatin, whereas both λ alleles were associated with centromeric. This suggests that genes differentially expressed in different cells occupy different nuclear environments; association with heterochromatin correlates with silencing of alleles. These two studies were undertaken in cells from different

mammalian species. I have shown that patterns of organisational constraints in mouse and human nuclei might be similar by comparing two regions of conserved synteny between mouse and human (Sections 3.6, 4.11 and 7.4). Therefore, interpreting the data from these two studies together might indicate that association with centromeric heterochromatin may occur without significant movement outside of a nuclear 'sub-volume'. While changes in transcription status may alter nuclear position in relation to nuclear landmarks, such as concentrations of centromeric heterochromatin, these rearrangements might still occur within the constraints of the territorial organisation of chromosomes.

I have presented evidence of a contrasting intra-chromosomal organisation of the α - and β -globin loci. Chromatin from gene-rich 16p13 (>50 genes/ Mb over 2 Mb), where the α -globin genes are located, was located a mean distance of $1.567^{+0.190}$ μm outside of the HSA16p territory, whereas the β -globin locus was found within the HSA11p territory (Section 4.4), $0.299^{+0.065}$ μm from the chromosome edge even though the β -globin locus itself is very gene-rich (5 genes in <50 kb, equivalent to over 100 genes per Mb). The distribution relative to sites of heterochromatin of α - and β -globin genes has previously been found to be markedly different in lymphoblasts. β -globin loci, were associated with pericentric heterochromatin in a high proportion of cycling lymphocytes, an association that has previously been correlated with repression of transcription (Brown *et al.*, 1997; Brown *et al.*, 1999). In contrast, α -globin genes were not associated with centromeric heterochromatin in lymphocytes, even though transcription of α -globin genes was repressed (Brown *et al.*, 2001). This contrasting nuclear organisation appears to mirror differences between these two loci at the molecular level (Vyas *et al.*, 1992; Smith and Higgs, 1999) (Section 6.3.3) and suggest that the decondensed structure of chromatin from HSA16p13.3 is not related to active transcription from the α -globin locus specifically, and may be characteristic of the locus in all cell types. The extremely gene-dense nature of 16p13.3 (>50 genes/Mb) may prevent it from being located adjacent to heterochromatin, thus preventing possible silencing of adjacent genes which are actively transcribed in cell types that do not express the tissue-specific α -globin gene cluster.

However, we still lack experimental paradigms that show unequivocally whether changes in higher-order chromatin architecture are the cause or consequence of an orchestrated gene expression. It remains to be proven that the sub nuclear localisation of a gene contributes actively to the establishment and/or mitotic inheritance of a particular chromatin state and hence transcriptional activity. The spatial positioning of genes may merely be the consequence of regulatory events that lead to the activated or the repressed state. The

clustering of similar chromatin states (active/inactive) is likely to provide a nuclear context, beyond that provided by flanking sequences, which can influence gene expression.

8.2.3 Gene-rich chromatin is preferentially located at the surface of a chromosome 'territory', and sometimes appears outside of the chromosome domain

Although individual genes are not necessarily at the periphery of a chromosome territory, I have presented data suggesting a trend for gene-rich chromatin to be positioned not only near the surface of an interphase chromosome, but often at a distance from the main chromosome domain as visualised using a chromosome paint in FISH (Chapters 4-6). The exact extent of decondensation depends on the fixation and analysis techniques used (Chapter 5), but the trend remains the same: gene-rich regions are preferentially located at the periphery of chromosome territories. This was found for a selection of gene rich regions throughout the human genome, selected for sites of high CpG island density (Craig and Bickmore, 1994), and confirmed using data from the Human Genome Mapping project. Complete sequences of HSA21q (The chromosome 21 mapping and sequencing consortium, 2000) and HSA22q (Dunham *et al.*, 1999) were analysed for gene density across the entire chromosome arms, and the distributions of gene-rich and gene-poor regions relative to respective chromosome territories were examined in two dimensions (Chapter 6). For both chromosomes, the predicted peripheral distribution of gene-rich loci was supported. In addition, a significant proportion of signals was positioned outside of the chromosome domain, and probes mapped to gene-poor regions were distributed more internally. Thus, I have begun to unravel patterns of organisation of specific regions of contiguous linear sequence within interphase chromosome domains, rather than inferring sequence identity from replication timing or GC-content.

Data suggest that a moderate level of transcriptional activity on a stretch of DNA is compatible with the sequences remaining within or close to the main body of a chromosome domain. Over and above that threshold, where many genes are being actively transcribed, chromatin may be extruded from the surface of a chromosome in order to achieve appropriate regulation of transcription of all genes in that region (Volpi *et al.*, 2000, and Chapter 6). Extrusion of chromatin might be an active, directional, energy-dependent process, or a passive consequence of high levels of transcriptional activity that may loosen chromatin structure, perhaps facilitated by increased electrostatic repulsion within highly acetylated transcriptionally active chromatin (discussed in 6.6.2). Alternatively, it could be a combination of the two: molecular level chromatin modification associated with

transcriptional activation, such as histone acetylation and ATP-dependent nucleosome remodelling, may affect higher-order levels of organisation and serve to disrupt higher-order chromosome architecture in a more passive manner. A direct role for transcription in large-scale chromatin structure was indicated by the observation that inhibition of RNA Polymerase II by DRB treatment of nuclei caused compaction of the HSA19 territory to an area similar to that adopted by HSA18, which is of a similar physical size (Croft *et al.*, 1999). This effect was reversed on removal of the block to transcription. Similarly, it would be interesting to assess the effect of DRB treatment on the nuclear configuration of gene-rich regions of the genome. This would enable appreciation of the effect that the mechanism of transcription *per se*, has on large-scale chromatin structure of individual chromatin fibres and allow separation of this effect from the influence of chromatin molecular structure, which serves to facilitate initiation and progression of transcription.

Measurements relating linear sequence distance between two loci and their spatial separation in MAA-fixed nuclei suggest that in extreme cases, higher-order chromatin structure may be decondensed to the level of the 30 nm fibre (Section 4.5). I suggest that increased transcriptional activity, facilitated by molecular-level chromatin modifications, render a looser configuration of large-scale chromatin structure, perhaps with fewer, or weaker, 'attachments' in the chromosome substructure. When transcriptional activity within a confined region of chromatin reaches a critical level, perhaps repulsion forces or diminished constraining attachments result in extrusion of chromatin from the surface of a chromosome territory. This effect is probably compounded by the MAA fixation procedure, where weaker or more infrequent sites of attachment within actively transcribed regions may be preferentially detached, resulting in the greater degree of chromatin decondensation induced by MAA treatment compared to cross-linking by pFA (Chapter 5). Increased chromatin extrusion from a territory surface was similarly observed in MAA-fixed nuclei, as compared to pFA-fixed when the organisation of gene-rich 6p21.3 was compared between (Volpi *et al.*, 2000). Extrusion of chromatin could also be facilitated by exclusion of higher-order packaging protein complexes from gene-rich regions. This would again probably be modulated by molecular-level chromatin modifications, such as high levels of histone acetylation. More densely packaged repressed chromatin, where there is little transcription activity and hypoacetylated histones, would maintain a condensed structure.

Projections of chromatin containing MHC genes have been found to associate preferentially with PML bodies (C. Shiels, personal communication). Directed chromatin extrusion could be the result of an active process driving chromatin to a particular nuclear subcompartment, or could be the result of passive extrusion of chromatin and subsequent association with a distant

nuclear body, such as a transcription factory, or indeed a PML body. Candidates for motor proteins inside the nucleus that could move chromatin in this way include actin and myosin (Milankov and De Boni, 1993), and the SMC family of ATPases, which have been proposed to act as motors in the process of chromatin condensation (Section 1.8.1). The ability of RNA polymerases to act as motors to move DNA has also been demonstrated; active positioning of chromatin by locally applied forces may thus turn out to be a widespread phenomenon.

8.2.4 Gene-rich chromatin is preferentially positioned within the nuclear interior

R-band sequences have been predicted to cluster within the nucleus interior forming an early replicating compartment that is transcriptionally competent and active (Sadoni *et al.*, 1999). In support of this, gene-rich chromosomes were found to be preferentially positioned within the nuclear interior (Croft *et al.*, 1999; Boyle *et al.*, 2001). However, I have shown that this is an over-simplification: the R-band 11p13 was found to co-localise with HSA11p, a chromosome arm which had an unusually peripheral position within the nucleus when its overall gene content was considered (Boyle *et al.*, 2001). Erosion analysis showed that almost 70% of loci from 11p13 were found in the outer three shells of the nucleus, mirroring the 70% of chromosome fluorescence in the outer three shells (Chapter 7). Therefore R-band 11p13 is not preferentially positioned within the nucleus interior; position of this locus is dependent on the gross organisation of the HSA11p chromosome domain within the nucleus.

When examined more closely, most chromosomes found in the centre of the nucleus (Boyle *et al.*, 2001) are enriched in T bands, a class of R-bands with the highest gene and GC-density (Craig and Bickmore, 1994). Loci from T-bands on HSA11p15.5 and 11q13 were preferentially positioned within the nuclear interior, at a distance from the main HSA11 territory, which is near the nuclear periphery. For a locus from distal 11p15.5 (D11S724), only 40% of signals were positioned in the outer three shells (Figure 7.4), compared to almost 70% of the chromosome territory fluorescence, as expected from analysis of 11p13. Similarly, erosion analysis of a sequence from HSA11q13, found less than less than 50% of signals in the outer three shells compared to more than 70% of the chromosome fluorescence in the outer three-fifths of the nucleus.

To extend this analysis, I identified gene-rich and gene-poor regions from chromosomes HSA21 and HSA22 (Chapter 6). Both HSA21 and HSA22 are generally internally positioned within the nucleus. As discussed in Section 8.1.3, gene-rich chromatin was positioned near

the surface, or at a distance, from these interphase chromosome territories,. However, examination of the nuclear position of these loops of gene-rich chromatin using erosion analysis (Chapter 7) found that they persist within the centre of the nucleus. Therefore, although these gene-rich loci are extruded from the main HSA21q and HSA22q domains, they remain orientated towards the nucleus interior. Gene-poor loci remained colocalised with the chromosome territories and were not extruded to the nuclear periphery as might be predicted by the hypothesis of nuclear organisation proposed by Sadoni *et al.*. However, normalisation of mean % locus signal in each shell by mean % chromosome fluorescence in each shell did suggest that the outer shell might be enriched for gene-poor loci, compared to gene-rich loci (Chapter 7).

8.2.5 How accurate is the concept of a chromosome ‘territory’?

Interphase chromosomes were suggested to be unravelled and intertwined like spaghetti on a plate (Comings, 1968; Vogel and Schroeder, 1974). Over the subsequent 20 years, however, evidence suggested a defined territorial organisation of chromosomes in the nucleus and levels of sub-chromosome organisation are also gradually being ascertained (Section 1.4.1 and 1.6). However, the concept of a chromosome territory, or for that matter of a chromosomal subdomain, makes sense only when the surface of that territory or subdomain can be defined with sufficient accuracy. Recent studies suggest that the models of Comings, and Vogel and Schroeder, might not be as inaccurate as previously thought. The organisation of chromosomes within the nucleus is probably somewhere in between the complete decondensation suggested 30 years ago (Comings, 1968; Vogel and Schroeder, 1974), and the model of discrete territorial organisation favoured in the 1990’s (Zirbel *et al.*, 1993; Cremer *et al.*, 1993; Kurz *et al.*, 1996). A very recent model suggests an ‘inter-chromatin domain’, a network of channels that permeates the entire nucleus, between sub-domains of chromosome territories through which protein complexes, nascent mRNA etc. can move, and that these spaces are linked to the nuclear pores to facilitate processes such as shuttling of proteins and export of mRNA. Loops or tails of chromatin extruded from the surface of chromosome ‘territories’ have been observed (Park and De Boni, 1998; Volpi *et al.*, 2000) and data presented here indicate that this configuration of chromatin in relation to its native chromosome ‘territory’ may not be uncommon. Furthermore, decondensed chromatin may be embedded in the territory of another chromosome (Visser and Aten, 1999).

Observations described in this thesis were made on fixed material. Such studies therefore provide only snapshots of the genome during the cell cycle and during development. Recent advances in live cell analysis are enabling questions of chromatin and nuclear body dynamics

to be investigated in real time. Active movement of chromatin against a background of passive diffusion has been suggested, and these processes, and the mechanisms that govern them, remain to be distinguished. Specific visualisation of pre-selected genomic regions is not currently possible, and only randomly tagged genome sites can be observed (Robinett *et al.*, 1996; Belmont and Straight, 1998). Similarly, visualisation of specific chromosome territories *in vivo* relies on incorporation of BrdU label and random segregation over subsequent cell divisions, until a few, or single chromosome territories are labelled in a single nucleus (Zink *et al.*, 1998; Visser *et al.*, 1998). When developments in live-cell analysis techniques allow, it will be interesting to monitor *in vivo* movement of gene-rich sites that have been observed at a distance from their native chromosome territories. Cell cycle-dependent motion and movement in response to factors that up- or down-regulate transcription of specific genes within such loci, in relation to chromosome territories and other nuclear bodies, will be of particular interest. Such chromatin extrusions as observed in this study and previously (Park and De Boni, 1998; Volpi *et al.*, 2000) must be retracted and condensed during prophase, and it would be interesting to monitor these regions as cells enter and exit mitosis. Live cell analysis would also eliminate fixation artefacts; as I have shown, different fixation procedures can have different effects on chromatin structure (Chapter 5).

We are becoming increasingly aware of chromatin existing outside of chromosome boundaries definable using fluorescent technologies (FISH, or *in vivo* labelling of chromatin and light microscopy detection). To further delineate chromosome organisation, fine detail organisation of small-scale chromatin fibres in the nucleus is necessary. A recent study combining *in vivo* labelling of chromatin with electron microscopy techniques attempted to define the structure of chromosome domains (Visser *et al.*, 2000), however, the *in vivo* labelling technique relies on random segregation of labelled chromatin, so specific regions cannot be investigated. Similarly, observation of the behaviour of specific regions of the genome within the nucleus, relative to nuclear bodies and other landmarks such as sites of heterochromatin or the nuclear lamina, must be correlated with molecular-level study of these same regions. As discussed (Section 8.2.2), the contrasting large-scale intra-chromosomal organisation of the α - and β -globin loci reflects differences between the chromatin structure of these two loci at the molecular level. It is therefore important to draw parallels between research on these different levels. Advances in our understanding of the nuclear organisation of specific loci will be enabled by the emerging finished human genome sequence, it will become possible to relate characteristics discerned from the linear sequence to behavioural characteristics of chromatin in the nucleus.

8.3 IMPLICATIONS OF THIS RESEARCH

A major goal in studying control of nuclear processes is to integrate our knowledge of nuclear organisation, chromosome structure and primary DNA sequence with epigenetic modifications e.g. DNA methylation and histone acetylation, and the key nuclear processes of replication, transcription, recombination and repair. Genome sequencing projects will have a significant impact on this area of research: the availability of finely mapped probes to all regions of genomes and access to large, contiguous stretches of well-characterised sequence will facilitate the correlation of linear sequence to higher-order chromatin structure and 3D spatial distribution within the nucleus. Regulation of large-scale chromatin structure is believed to be mediated by *cis*-acting sequences that have been defined largely through biochemical and functional assays. In the future it may be possible to directly identify regulatory sequences that mediate an effect through higher-order chromatin structure by examination of sequence databases.

Failure of correct gene expression underlies many human genetic disorders. The spatially, temporally and quantitatively correct expression of a gene requires not only intact coding sequence, but also functional regulatory control. In addition to the linear DNA sequence in terms of the promoter and enhancer/silencer elements, the local chromatin environment of a gene-locus is also important. Modification of DNA and chromatin is integral to the correct control of gene expression in mammals and several human diseases have been found to be due to mutations in genes encoding proteins involved in maintaining or modifying chromatin structure (Hendrich and Bickmore, 2001). In addition, the cause of several genetic diseases have been related to chromatin rearrangements outside the transcription and promoter regions, categorised as position effects (Kleinjan and van Heyningen, 1998). The *PAX6* gene was one of the first to provide a compelling argument that position effect could play a role in human eye disease. Long range control elements were implicated as there were several cases of aniridia (a disease resulting from homozygous loss of *PAX6* function in humans) in individuals with chromosomal breakpoints more than 100 kb from the gene (Kleinjan *et al.* 2001b, and references therein). A specific region more than 150 kb distal to the major *PAX6* promoter has been identified to contain regulatory elements required for controlled expression of the *PAX6* gene, using evolutionary sequence comparison, DNaseI hypersensitivity analysis and transgenic reporter (Kleinjan *et al.* 2001b). This is just one example of how functional gene domains can extend far beyond the transcription unit.

Greater understanding of position effects in mammalian systems has been facilitated study of gene regulation in transgenic mice. Typically, only a proportion of transgenic lines with promoter and enhancer sequences linked to a reporter gene will express the reporter, and expression level is not usually linked to transgene copy number. This variability is a function of the integration site of the reporter, for example little or no transgene expression is observed when the transgene integrates into or near a heterochromatic region and it has been shown that inclusion of an appropriate LCR (Section 1.3.7.2) in the construct can overcome this expression level variability (Francastel *et al.*, 1999, and references therein). Importantly, the study of Francastel *et al.*, showed that inclusion of an enhancer in a transgene construct specifically prevented association of transgenes with sites of heterochromatin in the nucleus (Section 1.3.7.2). Thus, it is not necessarily site of integration *per se* that affects transgene expression, perhaps by the spreading of heterochromatin over the transgene at the site of integration; large-scale intra-nuclear positioning also plays a significant role in the control of gene expression. Perhaps the ultimate chromatin structure and nuclear context of transgene, which may or may not be dictated by the site of integration in the genome, should be considered when designing transgenic or therapeutic vectors where high level of expression of the transgene is desired. It appears that, as one example, the ability to ensure that a transgene is located away from, or at the surface of a region of heterochromatin is important to facilitate expression.

We are only just beginning to appreciate the impact that perturbing chromatin organisation has on human genome function, and to understand the chromatin-based mechanisms that mammalian cells use to silence or activate gene expression. My data indicates that the nuclear interior might preferentially contain gene-rich chromatin, and one could infer from this that the nuclear interior might be associated with elevated levels of transcription, and perhaps increased concentrations of transcription factors and molecular-level chromatin modifying complexes. Therefore, one could speculate that targeting transgenes to the nuclear interior, perhaps by insertion into a gene-rich region of chromatin, might facilitate high levels of transcription. This study has presented largely correlative data, illustrating the fact that we still have only limited understanding of how large-scale nuclear organisation can impinge on the control of gene expression. The completion of the human genome sequence represents a landmark in modern-day science; however, contours and landmarks must be assigned to this linear sequence if we are to understand further how the fundamental properties of a cell are controlled. This is of importance for the development of many gene-based therapies; we cannot expect to be able to manipulate the genome effectively until we more thoroughly understand chromatin structure, both at the molecular level and in the context of the nucleus.

REFERENCES

- Aagaard, L., Laible, G., Selenko, P., Schmid, M., Dorn, R., Schotta, G., Kuhfittig, S., Wolf, A., Lebersorger, A., Singh, P. B., Reuter, G., and Jenuwein, T.** (1999). Functional mammalian homologues of the *Drosophila* PEV-modifier Su(var)3-9 encode centromere-associated proteins which complex with the heterochromatin component M31. *EMBO J.* **18**, 1923-1938.
- Abranches, R., Beven, A. F., Aragon-Alcaide, L., and Shaw, P. J.** (1998). Transcription sites are not correlated with chromosome territories in wheat nuclei. *J. Cell Biol.* **143**, 5-12.
- Akasaka, T., van Lohuizen, M., van der, L. N., Mizutani-Koseki, Y., Kanno, M., Taniguchi, M., Vidal, M., Alkema, M., Berns, A., and Koseki, H.** (2001). Mice doubly deficient for the Polycomb Group genes *Mel18* and *Bmi1* reveal synergy and requirement for maintenance but not initiation of Hox gene expression. *Development* **128**, 1587-1597.
- Alcobia, I., Dilao, R., and Parreira L.** (2000). Spatial associations of centromeres in the nuclei of hematopoietic cells: evidence for cell-type-specific organizational patterns. *Blood* **95**, 1608-1615.
- Alders, M., Hodges, M., Hadjantonakis, A. K., Postmus, J., van, W., I, Blik, J., de Meulemeester, M., Westerveld, A., Guillemot, F., Oudejans, C., Little, P., and Mannens, M.** (1997). The human Achaete-Scute homologue 2 (*ASCL2*, *HASH2*) maps to chromosome 11p15.5, close to *IGF2* and is expressed in extravillous trophoblasts. *Hum. Mol. Genet.* **6**, 859-867.
- Allfrey VG, Faulkner R, and Mirsky AE.** (1964) Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc. Natl. Acad. Sci. U.S.A* **51**, 786-794.
- Allshire, R. C., Javerzat, J. P., Redhead, N. J., and Cranston, G.** (1994). Position effect variegation at fission yeast centromeres. *Cell* **76**, 157-169.
- Andrulis, E. D., Neiman, A. M., Zappulla, D. C., and Sternglanz, R.** (1998). Perinuclear localization of chromatin facilitates transcriptional silencing. *Nature* **394**, 592-595.
- Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., and Kouzarides, T.** (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120-124.
- Bell, A. C., West, A. G., and Felsenfeld, G.** (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* **98**, 387-396.
- Bell, A. C., West, A. G., and Felsenfeld, G.** (2001). Insulators and boundaries: versatile regulatory elements in the eukaryotic genome. *Science* **291**, 447-450.

- Belmont, A. S., Bignone, F., and Ts'o, P. O.** (1986). The relative intranuclear positions of Barr bodies in XXX non-transformed human fibroblasts. *Exp Cell Res.* **165**, 165-179.
- Belmont, A. S. and Bruce, K.** (1994). Visualization of G1 chromosomes: a folded, twisted, supercoiled chromonema model of interphase chromatid structure. *J. Cell Biol.* **127**, 287-302.
- Belmont, A. S., Dietzel, S., Nye, A. C., Strukov, Y. G., and Tumber, T.** (1999). Large-scale chromatin structure and function. *Curr. Opin. Cell Biol.* **11**, 307-311.
- Belmont, A. S. and Straight, A. F.** (1998). In vivo visualization of chromosomes using lac operator-repressor binding. *Trends Cell Biol.* **8**, 121-124.
- Bepler, G., O'Briant, K. C., Kim, Y. C., Schreiber, G., and Pitterle, D. M.** (1999). A 1.4-Mb high-resolution physical map and contig of chromosome segment 11p15.5 and genes in the LOH11A metastasis suppressor region. *Genomics* **55**, 164-175.
- Berezney, R., Mortillaro, M. J., Ma, H., Wei, X., and Samarabandu, J.** (1995). The nuclear matrix: a structural milieu for genomic function. *Int. Rev. Cytol.* **162A**, 1-65.
- Berezney, R. and Wei, X.** (1998). The new paradigm: integrating genomic function and nuclear architecture. *J. Cell Biochem. Suppl* **30**, 238-242.
- Berg-Bakker, C. A., Hagemeijer, A., Franken-Postma, E. M., Smit, V. T., Kuppen, P. J., Ravenswaay Claasen, H. H., Cornelisse, C. J., and Schrier, P. I.** (1993). Establishment and characterization of 7 ovarian carcinoma cell lines and one granulosa tumor cell line: growth features and cytogenetics. *Int. J. Cancer* **53**, 613-620.
- Bestor, T., Laudano, A., Mattaliano, R., and Ingram, V.** (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *J. Mol. Biol.* **203**, 971-983.
- Bickmore, W. A. and Carothers, A. D.** (1995). Factors affecting the timing and imprinting of replication on a mammalian chromosome. *J. Cell Sci.* **108**, 2801-2809.
- Biggar, S. R. and Crabtree, G. R.** (1999). Continuous and widespread roles for the Swi-Snf complex in transcription. *EMBO J.* **18**, 2254-2264.
- Billia, F. and De Boni, U.** (1991). Localization of centromeric satellite and telomeric DNA sequences in dorsal root ganglion neurons, in vitro. *J. Cell Sci.* **100**, 219-226.
- Bird, A. P., Taggart, M. H., Nicholls, R. D., and Higgs, D. R.** (1987). Non-methylated CpG-rich islands at the human alpha-globin locus: implications for evolution of the alpha-globin pseudogene. *EMBO J.* **6**, 999-1004.
- Bird, A. P. and Wolffe, A. P.** (1999). Methylation-induced repression--belts, braces, and chromatin. *Cell* **99**, 451-454.
- Bobrow, M. and Heritage, J.** (1980). Nonrandom segregation of nucleolar organizing chromosomes at mitosis? *Nature* **288**, 79-81.

- Bone, J. R., Lavender, J., Richman, R., Palmer, M. J., Turner, B. M., and Kuroda, M. I.** (1994). Acetylated histone H4 on the male X chromosome is associated with dosage compensation in *Drosophila*. *Genes Dev.* **8**, 96-104.
- Bornfleth, H., Edelmann, P., Zink, D., Cremer, T., and Cremer, C.** (1999). Quantitative motion analysis of subchromosomal foci in living cells using four-dimensional microscopy. *Biophys.J.* **77**, 2871-2886.
- Bourns, B. D., Alexander, M. K., Smith, A. M., and Zakian, V. A.** (1998). Sir proteins, Rif proteins, and Cdc13p bind *Saccharomyces* telomeres in vivo. *Mol.Cell Biol.* **18**, 5600-5608.
- Boveri T.** Die Blastomerenkerne von *Ascaris megalocephala* und die Theorie der Chromosomenindividualitat (1909). *Arch Zellforschung* **3**, 181-286. 1909.
- Boyle, S., Gilchrist, S., Bridger, J. M., Mahy, N. L., Ellis, J. A., and Bickmore, W. A.** (2001). The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Hum.Mol.Genet.* **10**, 211-219.
- Bregman, D. B., Du, L., van der, Z. S., and Warren, S. L.** (1995). Transcription-dependent redistribution of the large subunit of RNA polymerase II to discrete nuclear domains. *J.Cell Biol.* **129**, 287-298.
- Brenner, S., Elgar, G., Sandford, R., Macrae, A., Venkatesh, B., and Aparicio, S.** (1993). Characterization of the pufferfish (*Fugu*) genome as a compact model vertebrate genome. *Nature* **366**, 265-268.
- Bridger, J. M. and Bickmore, W. A.** (1998). Putting the genome on the map. *Trends Genet.* **14**, 403-409.
- Bridger, J. M., Herrmann, H., Munkel, C., and Lichter, P.** (1998). Identification of an interchromosomal compartment by polymerization of nuclear-targeted vimentin. *J.Cell Sci.* **111**, 1241-1253.
- Bridger, J. M., Kill, I. R., and Lichter, P.** (1998). Association of pKi-67 with satellite DNA of the human genome in early G1 cells. *Chromosome.Res.* **6**, 13-24.
- Bridger, J. M., Boyle, S., Kill, I. R., and Bickmore, W. A.** (2000). Re-modelling of nuclear architecture in quiescent and senescent human fibroblasts. *Curr.Biol.* **10**, 149-152.
- Brown, K. E., Guest, S. S., Smale, S. T., Hahn, K., Merkenschlager, M., and Fisher, A. G.** (1997). Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell* **91**, 845-854.
- Brown, K. E., Baxter, J., Graf, D., Merkenschlager, M., and Fisher, A. G.** (1999). Dynamic repositioning of genes in the nucleus of lymphocytes preparing for cell division. *Mol.Cell* **3**, 207-217.
- Brown, K. E., Amoils, S., Horn, J. M., Buckle, V. J., Higgs, D. R., Merkenschlager, M., and Fisher, A. G.** (2001). Expression of alpha- and beta-globin genes occurs within different nuclear domains in haemopoietic cells. *Nat.Cell Biol.* **3**, 602-606.

- Brownell, J. E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D. G., Roth, S. Y., and Allis, C. D.** (1996). Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84**, 843-851.
- Buckel, A., Beeson, D., James, M., and Vincent, A.** (1996). Cloning of cDNA encoding human rapsyn and mapping of the RAPSN gene locus to chromosome 11p11.2-p11.1. *Genomics* **35**, 613-616.
- Buettner, J. A., Glusman, G., Ben Arie, N., Ramos, P., Lancet, D., and Evans, G. A.** (1998). Organization and evolution of olfactory receptor genes on human chromosome 11. *Genomics* **53**, 56-68.
- Byvoet, P., Shepherd, G. R., Hardin, J. M., and Noland, B. J.** (1972). The distribution and turnover of labeled methyl groups in histone fractions of cultured mammalian cells. *Arch.Biochem.Biophys.* **148**, 558-567.
- Callan, H. G.** (1982). The Croonian Lecture, 1981. Lampbrush chromosomes. *Proc.R.Soc.Lond B Biol.Sci.* **214**, 417-448.
- Carothers, A. D.** Projective sterology in biological microscopy. *Image Processing and Analysis, a Practical Approach*, Oxford University Press. 2000.
- Carruthers, L. M., Bednar, J., Woodcock, C. L., and Hansen, J. C.** (1998). Linker histones stabilize the intrinsic salt-dependent folding of nucleosomal arrays: mechanistic ramifications for higher-order chromatin folding. *Biochemistry* **37**, 14776-14787.
- Carruthers, L. M. and Hansen, J. C.** (2000). The core histone N termini function independently of linker histones during chromatin condensation. *J.Biol.Chem.* **275**, 37285-37290.
- Cassidy, S. B., Dykens, E., and Williams, C. A.** (2000). Prader-Willi and Angelman syndromes: sister imprinted disorders. *Am.J.Med.Genet.* **97**, 136-146.
- Cavalli, G. and Paro, R.** (1998). The Drosophila Fab-7 chromosomal element conveys epigenetic inheritance during mitosis and meiosis. *Cell* **93**, 505-518.
- Cavalli, G. and Paro, R.** (1999). Epigenetic inheritance of active chromatin after removal of the main transactivator. *Science* **286**, 955-958.
- Chadwick, B. P. and Willard, H. F.** (2001). Histone H2A variants and the inactive X chromosome: identification of a second macroH2A variant. *Hum.Mol.Genet.* **10**, 1101-1113.
- Chai, L. S. and Sandberg, A. A.** (1988). Chromosomes and their relationship to nuclear components during the cell cycle in Chinese hamster cells. *Cell Tissue Res.* **251**, 197-204.
- Chaly, N. and Munro, S. B.** (1996). Centromeres reposition to the nuclear periphery during L6E9 myogenesis in vitro. *Exp Cell Res.* **223**, 274-278.

- Chen, A. M., Lucas, J. N., Hill, F. S., Brenner, D. J., and Sachs, R. K.** (1996). Proximity effects for chromosome aberrations measured by FISH. *Int.J.Radiat.Biol.* **69**, 411-420.
- Chen, D., Ma, H., Hong, H., Koh, S. S., Huang, S. M., Schurter, B. T., Aswad, D. W., and Stallcup, M. R.** (1999). Regulation of transcription by a protein methyltransferase. *Science* **284**, 2174-2177.
- Cheung, P., Allis, C. D., and Sassone-Corsi, P.** (2000). Signaling to chromatin through histone modifications. *Cell* **103**, 263-271.
- Chevret, E., Volpi, E. V., and Sheer, D.** (2000). Mini review: form and function in the human interphase chromosome. *Cytogenet.Cell Genet.* **90**, 13-21.
- Choh, V. and De Boni, U.** (1996). Spatial repositioning of centromeric domains during regrowth of axons in nuclei of murine dorsal root ganglion neurons in vitro. *J.Neurobiol.* **31**, 325-332.
- Chuang, P. T., Lieb, J. D., and Meyer, B. J.** (1996). Sex-specific assembly of a dosage compensation complex on the nematode X chromosome. *Science* **274**, 1736-1739.
- Chung, H. M., Shea, C., Fields, S., Taub, R. N., Van der Ploeg, L. H., and Tse, D. B.** (1990). Architectural organization in the interphase nucleus of the protozoan *Trypanosoma brucei*: location of telomeres and mini-chromosomes. *EMBO J.* **9**, 2611-2619.
- Clayton, A. L., Rose, S., Barratt, M. J., and Mahadevan, L. C.** (2000). Phosphoacetylation of histone H3 on c-fos- and c-jun-associated nucleosomes upon gene activation. *EMBO J.* **19**, 3714-3726.
- Clemson, C. M. and Lawrence, J. B.** (1996). Multifunctional compartments in the nucleus: insights from DNA and RNA localization. *J.Cell Biochem.* **62**, 181-190.
- Cmarko, D., Verschure, P. J., Martin, T. E., Dahmus, M. E., Krause, S., Fu, X. D., van Driel, R., and Fakan, S.** (1999). Ultrastructural analysis of transcription and splicing in the cell nucleus after bromo-UTP microinjection. *Mol.Biol.Cell* **10**, 211-223.
- Cockell, M. and Gasser, S. M.** (1999). Nuclear compartments and gene regulation. *Curr.Opin.Genet.Dev.* **9**, 199-205.
- Comings, D. E.** (1968). The rationale for an ordered arrangement of chromatin in the interphase nucleus. *Am.J.Hum.Genet.* **20**, 440-460.
- Compton, D. A., Weil, M. M., Bonetta, L., Huang, A., Jones, C., Yeger, H., Williams, B. R., Strong, L. C., and Saunders, G. F.** (1990). Definition of the limits of the Wilms tumor locus on human chromosome 11p13. *Genomics* **6**, 309-315.
- Cooper, P. R., Nowak, N. J., Higgins, M. J., Simpson, S. A., Marquardt, A., Stoehr, H., Weber, B. H., Gerhard, D. S., de Jong, P. J., and Shows, T. B.** (1997). A sequence-ready high-resolution physical map of the best macular dystrophy gene region in 11q12-q13. *Genomics* **41**, 185-192.

- Cooper, P. R., Nowak, N. J., Higgins, M. J., Church, D. M., and Shows, T. B.** (1998). Transcript mapping of the human chromosome 11q12-q13.1 gene-rich region identifies several newly described conserved genes. *Genomics* **49**, 419-429.
- Cosma, M. P., Tanaka, T., and Nasmyth, K.** (1999). Ordered recruitment of transcription and chromatin remodeling factors to a cell. *Cell* **97**, 299-311.
- Costanzi, C. and Pehrson, J. R.** (1998). Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature* **393**, 599-601.
- Craig, J. M. and Bickmore, W. A.** (1994). The distribution of CpG islands in mammalian chromosomes. *Nat.Genet.* **7**, 376-382.
- Cremer, C., Munkel, C., Granzow, M., Jauch, A., Dietzel, S., Eils, R., Guan, X. Y., Meltzer, P. S., Trent, J. M., Langowski, J., and Cremer, T.** (1996). Nuclear architecture and the induction of chromosomal aberrations. *Mutat.Res.* **366**, 97-116.
- Craig, J. M., Boyle, S., Perry, P., and Bickmore, W. A.** (1997). Scaffold attachments within the human genome. *J.Cell Sci* **110**, 2673-2682.
- Cremer, T., Cremer, C., Schneider, T., Baumann, H., Hens, L., and Kirsch-Volders, M.** (1982). Analysis of chromosome positions in the interphase nucleus of Chinese hamster cells by laser-UV-microirradiation experiments. *Hum.Genet.* **62**, 201-209.
- Cremer, T., Kurz, A., Zirbel, R., Dietzel, S., Rinke, B., Schrock, E., Speicher, M. R., Mathieu, U., Jauch, A., and Emmerich, P.** (1993). Role of chromosome territories in the functional compartmentalization of the cell nucleus. *Cold Spring Harb.Symp.Quant.Biol.* **58**, 777-792.
- Cremer, T., Kreth, G., Koester, H., Fink, R. H., Heintzmann, R., Cremer, M., Solovei, I., Zink, D., and Cremer, C.** (2000). Chromosome territories, interchromatin domain compartment, and nuclear matrix: an integrated view of the functional nuclear architecture. *Crit Rev.Eukaryot.Gene Expr.* **10**, 179-212.
- Cremer, T. and Cremer, C.** (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat.Rev.Genet.* **2**, 292-301.
- Croft, J. A., Bridger, J. M., Boyle, S., Perry, P., Teague, P., and Bickmore, W. A.** (1999). Differences in the localization and morphology of chromosomes in the human nucleus. *J.Cell Biol.* **145**, 1119-1131.
- Cross, S. H., Lee, M., Clark, V. H., Craig, J. M., Bird, A. P., and Bickmore, W. A.** (1997). The chromosomal distribution of CpG islands in the mouse: evidence for genome scrambling in the rodent lineage. *Genomics* **40**, 454-461.
- Cryderman, D. E., Tang, H., Bell, C., Gilmour, D. S., and Wallrath, L. L.** (1999). Heterochromatic silencing of *Drosophila* heat shock genes acts at the level of promoter potentiation. *Nucleic Acids Res.* **27**, 3364-3370.
- Csink, A. K. and Henikoff, S.** (1996). Genetic modification of heterochromatic association and nuclear organization in *Drosophila*. *Nature* **381**, 529-531.

- Csink, A. K. and Henikoff, S.** (1998). Large-scale chromosomal movements during interphase progression in *Drosophila*. *J. Cell Biol.* **143**, 13-22.
- Daniels, R. J., Peden, J. F., Lloyd, C., Horsley, S. W., Clark, K., Tufarelli, C., Kearney, L., Buckle, V. J., Doggett, N. A., Flint, J., and Higgs, D. R.** (2001). Sequence, structure and pathology of the fully annotated terminal 2 Mb of the short arm of human chromosome 16. *Hum. Mol. Genet.* **10**, 339-352.
- Dasso, M., Dimitrov, S., and Wolffe, A. P.** (1994). Nuclear assembly is independent of linker histones. *Proc. Natl. Acad. Sci U.S.A* **91**, 12477-12481.
- De Boni, U.** (1994). The interphase nucleus as a dynamic structure. *Int. Rev. Cytol.* **150**, 149-171.
- de la Barre, A. E., Gerson, V., Gout, S., Creaven, M., Allis, C. D., and Dimitrov, S.** (2000). Core histone N-termini play an essential role in mitotic chromosome condensation. *EMBO J.* **19**, 379-391.
- Dernburg, A. F., Broman, K. W., Fung, J. C., Marshall, W. F., Philips, J., Agard, D. A., and Sedat, J. W.** (1996). Perturbation of nuclear architecture by long-distance chromosome interactions. *Cell* **85**, 745-759.
- Dhalluin, C., Carlson, J. E., Zeng, L., He, C., Aggarwal, A. K., and Zhou, M. M.** (1999). Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399**, 491-496.
- Dietzel, S., Eils, R., Satzler, K., Bornfleth, H., Jauch, A., Cremer, C., and Cremer, T.** (1998). Evidence against a looped structure of the inactive human X-chromosome territory. *Exp. Cell Res.* **240**, 187-196.
- Dietzel, S., Jauch, A., Kienle, D., Qu, G., Holtgreve-Grez, H., Eils, R., Munkel, C., Bittner, M., Meltzer, P. S., Trent, J. M., and Cremer, T.** (1998). Separate and variably shaped chromosome arm domains are disclosed by chromosome arm painting in human cell nuclei. *Chromosome. Res.* **6**, 25-33.
- Dietzel, S., Schiebel, K., Little, G., Edelmann, P., Rappold, G. A., Eils, R., Cremer, C., and Cremer, T.** (1999). The 3D positioning of ANT2 and ANT3 genes within female X chromosome territories correlates with gene activity. *Exp Cell Res.* **252**, 363-375.
- Dimitrova, D. S. and Gilbert, D. M.** (1999). The spatial position and replication timing of chromosomal domains are both established in early G1 phase. *Mol. Cell* **4**, 983-993.
- Dong, F. and Jiang, J.** (1998). Non-Rabl patterns of centromere and telomere distribution in the interphase nuclei of plant cells. *Chromosome. Res.* **6**, 551-558.
- Donze, D., Adams, C. R., Rine, J., and Kamakaka, R. T.** (1999). The boundaries of the silenced HMR domain in *Saccharomyces cerevisiae*. *Genes Dev.* **13**, 698-708.
- Dundr, M. and Misteli, T.** (2001). Functional architecture in the cell nucleus. *Biochem. J.* **356**, 297-310.

- Dunham, I. et al.**, (1999). The DNA sequence of human chromosome 22. *Nature* **402**, 489-495.
- Eils, R., Dietzel, S., Bertin, E., Schrock, E., Speicher, M. R., Ried, T., Robert-Nicoud, M., Cremer, C., and Cremer, T.** (1996). Three-dimensional reconstruction of painted human interphase chromosomes: active and inactive X chromosome territories have similar volumes but differ in shape and surface structure. *J. Cell Biol.* **135**, 1427-1440.
- Eissenberg, J. C., James, T. C., Foster-Hartnett, D. M., Hartnett, T., Ngan, V., and Elgin, S. C.** (1990). Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A* **87**, 9923-9927.
- Eissenberg, J. C. and Elgin, S. C.** (2000). The HP1 protein family: getting a grip on chromatin. *Curr. Opin. Genet. Dev.* **10**, 204-210.
- Ekwall, K., Nimmo, E. R., Javerzat, J. P., Borgstrom, B., Egel, R., Cranston, G., and Allshire, R.** (1996). Mutations in the fission yeast silencing factors *clr4+* and *rik1+* disrupt the localisation of the chromo domain protein Swi6p and impair centromere function. *J. Cell Sci* **109**, 2637-2648.
- Ekwall, K., Olsson, T., Turner, B. M., Cranston, G., and Allshire, R. C.** (1997). Transient inhibition of histone deacetylation alters the structural and functional imprint at fission yeast centromeres. *Cell* **91**, 1021-1032.
- Elgar, G., Sandford, R., Aparicio, S., Macrae, A., Venkatesh, B., and Brenner, S.** (1996). Small is beautiful: comparative genomics with the pufferfish (*Fugu rubripes*). *Trends Genet.* **12**, 145-150.
- Engemann, S., Stroedicke, M., Paulsen, M., Franck, O., Reinhardt, R., Lane, N., Reik, W., and Walter, J.** (2000). Sequence and functional comparison in the Beckwith-Wiedemann region: implications for a novel imprinting centre and extended imprinting. *Hum. Mol. Genet.* **9**, 2691-2706.
- Ericsson, C., Mehlin, H., Bjorkroth, B., Lamb, M. M., and Daneholt, B.** (1989). The ultrastructure of upstream and downstream regions of an active Balbiani ring gene. *Cell* **56**, 631-639.
- Fakan, S. and Nobis, P.** (1978). Ultrastructural localization of transcription sites and of RNA distribution during the cell cycle of synchronized CHO cells. *Exp Cell Res.* **113**, 327-337.
- Fantes, J. A.** (1997). Molecular cytogenetics of 11p. PhD Thesis, University of Edinburgh.
- Fantes, J. A., Bickmore, W. A., Fletcher, J. M., Ballesta, F., Hanson, I. M., and van, H., V** (1992). Submicroscopic deletions at the WAGR locus, revealed by nonradioactive in situ hybridization. *Am. J. Hum. Genet.* **51**, 1286-1294.
- Fantes, J. A., Oghene, K., Boyle, S., Danes, S., Fletcher, J. M., Bruford, E. A., Williamson, K., Seawright, A., Schedl, A., Hanson, I., and .** (1995). A high-resolution integrated physical, cytogenetic, and genetic map of human chromosome 11: distal p13 to proximal p15.1. *Genomics* **25**, 447-461.

- Ferreira, J., Paoella, G., Ramos, C., and Lamond, A. I.** (1997). Spatial organization of large-scale chromatin domains in the nucleus: a magnified view of single chromosome territories. *J.Cell Biol.* **139**, 1597-1610.
- Finch, J. T., Lutter, L. C., Rhodes, D., Brown, R. S., Rushton, B., Levitt, M., and Klug, A.** (1977). Structure of nucleosome core particles of chromatin. *Nature* **269**, 29-36.
- Flaus, A. and Owen-Hughes, T.** (2001). Mechanisms for ATP-dependent chromatin remodelling. *Curr.Opin.Genet.Dev.* **11**, 148-154.
- Flint, J., Thomas, K., Micklem, G., Raynham, H., Clark, K., Doggett, N. A., King, A., and Higgs, D. R.** (1997). The relationship between chromosome structure and function at a human telomeric region. *Nat.Genet.* **15**, 252-257.
- Francastel, C., Walters, M. C., Groudine, M., and Martin, D. I.** (1999). A functional enhancer suppresses silencing of a transgene and prevents its localization close to centromeric heterochromatin. *Cell* **99**, 259-269.
- Francastel, C., Schubeler, D., Martin, D. I., and Groudine, M.** (2000). Nuclear compartmentalization and gene activity. *Nat.Rev.Mol.Cell Biol.* **1**, 137-143.
- Francis, N. J. and Kingston, R. E.** (2001). Mechanisms of transcriptional memory. *Nat.Rev.Mol.Cell Biol.* **2**, 409-421.
- Frey, M. R., Bailey, A. D., Weiner, A. M., and Matera, A. G.** (1999). Association of snRNA genes with coiled bodies is mediated by nascent snRNA transcripts. *Curr.Biol.* **9**, 126-135.
- Fry, C. J. and Peterson, C. L.** (2001). Chromatin remodeling enzymes: who's on first? *Curr.Biol.* **11**, R185-R197.
- Funabiki, H., Hagan, I., Uzawa, S., and Yanagida, M.** (1993). Cell cycle-dependent specific positioning and clustering of centromeres and telomeres in fission yeast. *J.Cell Biol.* **121**, 961-976.
- Gabriel, J. M., Higgins, M. J., Gebuhr, T. C., Shows, T. B., Saitoh, S., and Nicholls, R. D.** (1998). A model system to study genomic imprinting of human genes. *Proc.Natl.Acad.Sci U.S.A* **95**, 14857-14862.
- Galy, V., Olivo-Marin, J. C., Scherthan, H., Doye, V., Rascalou, N., and Nehrass, U.** (2000). Nuclear pore complexes in the organization of silent telomeric chromatin. *Nature* **403**, 108-112.
- Gasser, S. M.** (2001). Positions of potential: nuclear organization and gene expression. *Cell* **104**, 639-642.
- Gawin, B., Niederfuhr, A., Schumacher, N., Hummerich, H., Little, P. F., and Gessler, M.** (1999). A 7.5 Mb sequence-ready PAC contig and gene expression map of human chromosome 11p13-p14.1. *Genome Res.* **9**, 1074-1086.
- Gemkow, M. J., Verveer, P. J., and Arndt-Jovin, D. J.** (1998). Homologous association of the Bithorax-Complex during embryogenesis: consequences for transvection in *Drosophila melanogaster*. *Development* **125**, 4541-4552.

- Gerasimova, T. I., Byrd, K., and Corces, V. G.** (2000). A chromatin insulator determines the nuclear localization of DNA. *Mol. Cell* **6**, 1025-1035.
- Gerasimova, T. I. and Corces, V. G.** (1998). Polycomb and trithorax group proteins mediate the function of a chromatin insulator. *Cell* **92**, 511-521.
- Ghosh, D., Gerasimova, T. I., and Corces, V. G.** (2001). Interactions between the Su(Hw) and Mod(mdg4) proteins required for gypsy insulator function. *EMBO J.* **20**, 2518-2527.
- Gibbons, R. J., Bachoo, S., Picketts, D. J., Aftimos, S., Asenbauer, B., Bergoffen, J., Berry, S. A., Dahl, N., Fryer, A., Keppler, K., Kurosawa, K., Levin, M. L., Masuno, M., Neri, G., Pierpont, M. E., Slaney, S. F., and Higgs, D. R.** (1997). Mutations in transcriptional regulator ATRX establish the functional significance of a PHD-like domain. *Nat. Genet.* **17**, 146-148.
- Gibbons, R. J., Picketts, D. J., Villard, L., and Higgs, D. R.** (1995). Mutations in a putative global transcriptional regulator cause X-linked mental retardation with alpha-thalassemia (ATR-X syndrome). *Cell* **80**, 837-845.
- Gibbons, R. J., McDowell, T. L., Raman, S., O'Rourke, D. M., Garrick, D., Ayyub, H., and Higgs, D. R.** (2000). Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation. *Nat. Genet.* **24**, 368-371.
- Glenn, C. C., Driscoll, D. J., Yang, T. P., and Nicholls, R. D.** (1997). Genomic imprinting: potential function and mechanisms revealed by the Prader-Willi and Angelman syndromes. *Mol. Hum. Reprod.* **3**, 321-332.
- Gobbi, P., Falconi, M., Vitale, M., Galanzi, A., Artico, M., Martelli, A. M., and Mazzotti, G.** (1999). Scanning electron microscopic detection of nuclear structures involved in DNA replication. *Arch. Histol. Cytol.* **62**, 317-326.
- Gorman, M., Kuroda, M. I., and Baker, B. S.** (1993). Regulation of the sex-specific binding of the maleless dosage compensation protein to the male X chromosome in *Drosophila*. *Cell* **72**, 39-49.
- Gotta, M., Laroche, T., Formenton, A., Maillet, L., Scherthan, H., and Gasser, S. M.** (1996). The clustering of telomeres and colocalization with Rap1, Sir3, and Sir4 proteins in wild-type *Saccharomyces cerevisiae*. *J. Cell Biol.* **134**, 1349-1363.
- Grande, M. A., van, d. K., I, van Steensel, B., Schul, W., de The, H., van der Voort, H. T., de Jong, L., and van Driel, R.** (1996). PML-containing nuclear bodies: their spatial distribution in relation to other nuclear components. *J. Cell Biochem.* **63**, 280-291.
- Grande, M. A., van, D. K., I, de Jong, L., and van Driel, R.** (1997). Nuclear distribution of transcription factors in relation to sites of transcription and RNA polymerase II. *J. Cell Sci.* **110**, 1781-1791.
- Grant, P. A.** (2001). A tale of histone modifications. *Genome Biol.* **2**
- Gravel, S., Larrivee, M., Labrecque, P., and Wellinger, R. J.** (1998). Yeast Ku as a regulator of chromosomal DNA end structure. *Science* **280**, 741-744.

- Grewal, S. I. and Klar, A. J.** (1996). Chromosomal inheritance of epigenetic states in fission yeast during mitosis and meiosis. *Cell* **86**, 95-101.
- Grigoryev, S. A., Bednar, J., and Woodcock, C. L.** (1999). MENT, a heterochromatin protein that mediates higher order chromatin folding, is a new serpin family member. *J.Biol.Chem.* **274**, 5626-5636.
- Grosveld, F., van Assendelft, G. B., Greaves, D. R., and Kollias, G.** (1987). Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell* **51**, 975-985.
- Guan, X. Y., Zhang, H., Bittner, M., Jiang, Y., Meltzer, P., and Trent, J.** (1996). Chromosome arm painting probes. *Nat.Genet.* **12**, 10-11.
- Haaf, T. and Schmid, M.** (1991). Chromosome topology in mammalian interphase nuclei. *Exp.Cell Res.* **192**, 325-332.
- Hansen, J. C., Tse, C., and Wolffe, A. P.** (1998). Structure and function of the core histone N-termini: more than meets the eye. *Biochemistry* **37**, 17637-17641.
- Hanson, I. M., Seawright, A., Hardman, K., Hodgson, S., Zaletayev, D., Fekete, G., and van, Heyningen, V** (1993). PAX6 mutations in aniridia. *Hum.Mol.Genet.* **2**, 915-920.
- Hayes, J. J. and Hansen, J. C.** (2001). Nucleosomes and the chromatin fiber. *Curr.Opin.Genet.Dev.* **11**, 124-129.
- He, D. and Brinkley, B. R.** (1996). Structure and dynamic organization of centromeres/prekinetochores in the nucleus of mammalian cells. *J.Cell Sci.* **109**, 2693-2704.
- Henderson, A. S., Warburton, D., and Atwood, K. C.** (1972). Location of ribosomal DNA in the human chromosome complement. *Proc.Natl.Acad.Sci.U.S.A* **69**, 3394-3398.
- Hendrich, B. and Bird, A.** (1998). Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol.Cell Biol.* **18**, 6538-6547.
- Hendrich, B. and Bickmore, W. A.** Human diseases with underlying defects in chromatin structure and modification. *Hum.Mol.Genet.* **10**, 1-10. 2001.
- Henzel, M. J. and Bazett-Jones, D. P.** (1997). Fixation-dependent organization of core histones following DNA fluorescent in situ hybridization. *Chromosoma* **106**, 114-123.
- Henikoff, S.** (1997). Nuclear organization and gene expression: homologous pairing and long- range interactions. *Curr.Opin.Cell Biol.* **9**, 388-395.
- Heslop-Harrison, J. S. and Bennett, M. D.** (1990). Nuclear architecture in plants. *Trends Genet.* **6**, 401-405.

- Heslop-Harrison, J. S., Leitch, A. R., and Schwarzacher, T.** 1993. The physical organization of interphase nuclei. *The Chromosomes*. 221-232. Oxford: Bios Scientific Publishers,
- Hirano, T.** (1998). SMC protein complexes and higher-order chromosome dynamics. *Curr. Opin. Cell Biol.* **10**, 317-322.
- Hochstrasser, M., Mathog, D., Gruenbaum, Y., Saumweber, H., and Sedat, J. W.** (1986). Spatial organization of chromosomes in the salivary gland nuclei of *Drosophila melanogaster*. *J. Cell Biol.* **102**, 112-123.
- Hooper, M., Hardy, K., Handyside, A., Hunter, S., and Monk, M.** (1987). HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. *Nature* **326**, 292-295.
- Hu, R. J., Lee, M. P., Connors, T. D., Johnson, L. A., Burn, T. C., Su, K., Landes, G. M., and Feinberg, A. P.** (1997). A 2.5-Mb transcript map of a tumor-suppressing subchromosomal transferable fragment from 11p15.5, and isolation and sequence analysis of three novel genes. *Genomics* **46**, 9-17.
- Huang, S. and Spector, D. L.** (1991). Nascent pre-mRNA transcripts are associated with nuclear regions enriched in splicing factors. *Genes Dev.* **5**, 2288-2302.
- Iborra, F. J., Jackson, D. A., and Cook, P. R.** (2001). Coupled transcription and translation within nuclei of mammalian cells. *Science* **293**, 1139-1142.
- Imai, S., Armstrong, C. M., Kaeberlein, M., and Guarente, L.** (2000). Transcriptional silencing and longevity protein Sir2 is an NAD- dependent histone deacetylase. *Nature* **403**, 795-800.
- Jackson, D. A. and Cook, P. R.** (1995). The structural basis of nuclear function. *Int. Rev. Cytol.* **162A**, 125-149.
- Jackson, D. A., Dolle, A., Robertson, G., and Cook, P. R.** (1992). The attachments of chromatin loops to the nucleoskeleton. *Cell Biol. Int. Rep.* **16**, 687-696.
- Jackson, D. A., Hassan, A. B., Errington, R. J., and Cook, P. R.** (1993). Visualization of focal sites of transcription within human nuclei. *EMBO J.* **12**, 1059-1065.
- Jackson, D. A., Iborra, F. J., Manders, E. M., and Cook, P. R.** (1998). Numbers and organization of RNA polymerases, nascent transcripts, and transcription units in HeLa nuclei. *Mol. Biol. Cell* **9**, 1523-1536.
- Jackson, D. A. and Pombo, A.** (1998). Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells. *J. Cell Biol.* **140**, 1285-1295.
- Jacobson, R. H., Ladurner, A. G., King, D. S., and Tjian, R.** (2000). Structure and function of a human TAFII250 double bromodomain module. *Science* **288**, 1422-1425.

- Janevski, J., Park, P. C., and De Boni, U.** (1997). Changes in morphology and spatial position of coiled bodies during NGF- induced neuronal differentiation of PC12 cells. *J.Histochem.Cytochem.* **45**, 1523-1531.
- Jenuwein, T. and Muller, R.** (1987). Structure-function analysis of fos protein: a single amino acid change activates the immortalizing potential of v-fos. *Cell* **48**, 647-657.
- 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.
- Jones, D. O., Cowell, I. G., and Singh, P. B.** (2000). Mammalian chromodomain proteins: their role in genome organisation and expression. *Bioessays* **22**, 124-137.
- Jones, P. A. and Takai, D.** (2001). The role of DNA methylation in mammalian epigenetics. *Science* **293**, 1068-1070.
- Jones, P. L., Veenstra, G. J., Wade, P. A., Vermaak, D., Kass, S. U., Landsberger, N., Strouboulis, J., and Wolffe, A. P.** (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat.Genet.* **19**, 187-191.
- Kal, A. J., Mahmoudi, T., Zak, N. B., and Verrijzer, C. P.** (2000). The Drosophila brahma complex is an essential coactivator for the trithorax group protein zeste. *Genes Dev.* **14**, 1058-1071.
- Kato, R., Shirohzu, H., Yokomine, T., Mizuno, S., Mukai, T., and Sasaki, H.** (1999). Sequence-ready 1-Mb YAC, BAC and cosmid contigs covering the distal imprinted region of mouse chromosome 7. *DNA Res.* **6**, 401-405.
- Kelley, R. L., Solovyeva, I., Lyman, L. M., Richman, R., Solovyev, V., and Kuroda, M. I.** (1995). Expression of msl-2 causes assembly of dosage compensation regulators on the X chromosomes and female lethality in Drosophila. *Cell* **81**, 867-877.
- Kellum, R. and Schedl, P.** (1992). A group of scs elements function as domain boundaries in an enhancer- blocking assay. *Mol.Cell Biol.* **12**, 2424-2431.
- Kendall, F., Swenson, R., Borun, T., Rowinski, J., and Nicolini, C.** (1977). Nuclear morphometry during the cell cycle. *Science* **196**, 1106-1109.
- Kent, J., Lee, M., Schedl, A., Boyle, S., Fantès, J., Powell, M., Rushmere, N., Abbott, C., van, H., V, and Bickmore, W. A.** (1997). The reticulocalbin gene maps to the WAGR region in human and to the Small eye Harwell deletion in mouse. *Genomics* **42**, 260-267.
- Keohane, A. M., Lavender, J. S., O'Neill, L. P., and Turner, B. M.** (1998). Histone acetylation and X inactivation. *Dev.Genet.* **22**, 65-73.
- Kim, J., Sif, S., Jones, B., Jackson, A., Koipally, J., Heller, E., Winandy, S., Viel, A., Sawyer, A., Ikeda, T., Kingston, R., and Georgopoulos, K.** (1999). Ikaros DNA-binding proteins direct formation of chromatin remodeling complexes in lymphocytes. *Immunity.* **10**, 345-355.

- Kingston, R. E. and Narlikar, G. J.** (1999). ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. *Genes Dev.* **13**, 2339-2352.
- Kioussis, D. and Festenstein, R.** (1997). Locus control regions: overcoming heterochromatin-induced gene inactivation in mammals. *Curr. Opin. Genet. Dev.* **7**, 614-619.
- Klein, C., Cheutin, T., O'Donohue, M. F., Rothblum, L., Kaplan, H., Beorchia, A., Lucas, L., Heliot, L., and Ploton, D.** (1998). The three-dimensional study of chromosomes and upstream binding factor- immunolabeled nucleolar organizer regions demonstrates their nonrandom spatial arrangement during mitosis. *Mol. Biol. Cell* **9**, 3147-3159.
- Kleinjan, D. J. and van, H., V** (1998). Position effect in human genetic disease. *Hum. Mol. Genet.* **7**, 1611-1618.
- Kleinjan, D. A., Seawright, A., Elgar, G., and van Heyningen, V.** (2001a). Characterisation of a novel gene adjacent to PAX6, revealing synteny conservation with functional significance. *Mamm. Genome.* In Press.
- Kleinjan, D.A., Seawright, A., Schedl, A., Quinlan, R.A., Danes, S., and van Heyningen, V.** (2001b). Aniridia-associated translocations, DNase hypersensitivity, sequence comparison, and transgenic analysis redefine the functional domain of PAX6. *Hum. Mol. Genet.* **10**, 2049-2059.
- Kornberg, R. D. and Lorch, Y.** (1999). Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell* **98**, 285-294.
- Krebs, J. E., Kuo, M. H., Allis, C. D., and Peterson, C. L.** (1999). Cell cycle-regulated histone acetylation required for expression of the yeast HO gene. *Genes Dev.* **13**, 1412-1421.
- Krebs, J. E., Fry, C. J., Samuels, M. L., and Peterson, C. L.** (2000). Global role for chromatin remodeling enzymes in mitotic gene expression. *Cell* **102**, 587-598.
- Kurz, A., Lampel, S., Nickolenko, J. E., Bradl, J., Benner, A., Zirbel, R. M., Cremer, T., and Lichter, P.** (1996). Active and inactive genes localize preferentially in the periphery of chromosome territories. *J. Cell Biol.* **135**, 1195-1205.
- Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T.** (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116-120.
- Lamond, A. I. and Earnshaw, W. C.** (1998). Structure and function in the nucleus. *Science* **280**, 547-553.
- Lampel, S., Bridger, J. M., Zirbel, R. M., Mathieu, U. R., and Lichter, P.** (1997). Nuclear RNA accumulations contain released transcripts and exhibit specific distributions with respect to Sm antigen foci. *DNA Cell Biol.* **16**, 1133-1142.
- Lander, E. S., et al.,** (2001). Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921.

- Landry, J., Sutton, A., Tafrov, S. T., Heller, R. C., Stebbins, J., Pillus, L., and Sternglanz, R.** (2000). The silencing protein SIR2 and its homologs are NAD-dependent protein deacetylases. *Proc. Natl. Acad. Sci U.S.A* **97**, 5807-5811.
- Langer-Safer, P. R., Levine, M., and Ward, D. C.** (1982). Immunological method for mapping genes on Drosophila polytene chromosomes. *Proc. Natl. Acad. Sci U.S.A.* **79**, 4381-4385.
- LaSalle, J. M. and Lalande, M.** (1996). Homologous association of oppositely imprinted chromosomal domains. *Science* **272**, 725-728.
- Lattanzi, G., Galanzi, A., Gobbi, P., Falconi, M., Matteucci, A., Breschi, L., Vitale, M., and Mazzotti, G.** (1998). Ultrastructural aspects of the DNA polymerase alpha distribution during the cell cycle. *J.Histochem.Cytochem.* **46**, 1435-1442.
- Lawrence, J. B., Singer, R. H. , and Marselle, L. M.** (1989). Highly localized tracks of specific transcripts within interphase nuclei visualized by in situ hybridization. *Cell* **57**, 493-502.
- Lawrence, J. B., Carter, K. C. , and Xing, X.** (1993). Probing functional organization within the nucleus: is genome structure integrated with RNA metabolism? *Cold Spring Harb.Symp.Quant.Biol.* **58**, 807-818.
- Leblond, C. P. and El Alfy, M.** (1998). The eleven stages of the cell cycle, with emphasis on the changes in chromosomes and nucleoli during interphase and mitosis. *Anat.Rec.* **252**, 426-443.
- Lee, M. P., Brandenburg, S., Landes, G. M., Adams, M., Miller, G., and Feinberg, A. P.** (1999). Two novel genes in the center of the 11p15 imprinted domain escape genomic imprinting. *Hum.Mol.Genet.* **8**, 683-690.
- Leitch, A. R., Brown, J. K., Mosgoller, W., Schwarzacher, T., and Heslop-Harrison, J. S.** (1994). The spatial localization of homologous chromosomes in human fibroblasts at mitosis. *Hum.Genet.* **93**, 275-280.
- Li, E., Bestor, T. H., and Jaenisch, R.** (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915-926.
- Li, G., Sudlow, G., and Belmont, A. S.** (1998). Interphase cell cycle dynamics of a late-replicating, heterochromatic homogeneously staining region: precise choreography of condensation/decondensation and nuclear positioning . *J.Cell Biol.* **140**, 975-989.
- Li, Q., Harju, S., and Peterson, K. R.** (1999). Locus control regions: coming of age at a decade plus. *Trends Genet.* **15**, 403-408.
- Lichter, P., Cremer, T., Borden, J., Manuelidis, L., and Ward, D. C.** (1988). Delineation of individual human chromosomes in metaphase and interphase cells by in situ suppression hybridization using recombinant DNA libraries. *Hum.Genet.* **80**, 224-234.
- Lichter, P., Boyle, A. L., Cremer, T., and Ward, D. C.** (1991). Analysis of genes and chromosomes by nonisotopic in situ hybridization. *Genet.Anal.Tech.Appl.* **8**, 24-35.

- Lieb, J. D., Capowski, E. E., Meneely, P., and Meyer, B. J.** (1996). DPY-26, a link between dosage compensation and meiotic chromosome segregation in the nematode. *Science* **274**, 1732-1736.
- Little, M., Holmes, G., and Walsh, P.** (1999). WT1: what has the last decade told us? *Bioessays* **21**, 191-202.
- Lo, W. S., Duggan, L., Tolga, N. C., Emre, Belotserkovskya, R., Lane, W. S., Shiekhattar, R., and Berger, S. L.** (2001). Snf1--a histone kinase that works in concert with the histone acetyltransferase Gcn5 to regulate transcription. *Science* **293**, 1142-1146.
- Lopez-Velazquez, G., Marquez, J., Ubaldo, E., Corkidi, G., Echeverria, O., and Vazquez Nin, G. H.** (1996). Three-dimensional analysis of the arrangement of compact chromatin in the nucleus of G0 rat lymphocytes. *Histochem. Cell Biol.* **105**, 153-161.
- Lorch, Y., Zhang, M., and Kornberg, R. D.** (2001). RSC unravels the nucleosome. *Mol. Cell* **7**, 89-95.
- Lu, B. Y., Bishop, C. P., and Eisenberg, J. C.** (1996). Developmental timing and tissue specificity of heterochromatin-mediated silencing. *EMBO J.* **15**, 1323-1332.
- Luderus, M. E., van Steensel, B., Chong, L., Sibon, O. C., Cremers, F. F., and de Lange, T.** (1996). Structure, subnuclear distribution, and nuclear matrix association of the mammalian telomeric complex. *J. Cell Biol.* **135**, 867-881.
- Lundgren, M., Chow, C. M., Sabbattini, P., Georgiou, A., Minaee, S., and Dillon, N.** (2000). Transcription factor dosage affects changes in higher order chromatin structure associated with activation of a heterochromatic gene. *Cell* **103**, 733-743.
- Ma, H., Siegel, A. J., and Berezney, R.** (1999). Association of chromosome territories with the nuclear matrix. Disruption of human chromosome territories correlates with the release of a subset of nuclear matrix proteins. *J. Cell Biol.* **146**, 531-542.
- Madisen, L., Krumm, A., Hebbes, T. R., and Groudine, M.** (1998). The immunoglobulin heavy chain locus control region increases histone acetylation along linked c-myc genes. *Mol. Cell Biol.* **18**, 6281-6292.
- Maher, E. R. and Reik, W.** (2000). Beckwith-Wiedemann syndrome: imprinting in clusters revisited. *J. Clin. Invest* **105**, 247-252.
- Mahy, N. L., Bickmore, W. A., Tumbar, T., and Belmont, A. S.** (2000). Linking large-scale chromatin structure with nuclear function. *Chromatin Structure and Gene Expression* Edited by Elgin, S.C.R and Workman, J.L, 300-321.
- Maillet, L., Boscheron, C., Gotta, M., Marcand, S., Gilson, E., and Gasser, S. M.** (1996). Evidence for silencing compartments within the yeast nucleus: a role for telomere proximity and Sir protein concentration in silencer-mediated repression. *Genes Dev.* **10**, 1796-1811.

- Manuelidis, L.** (1984). Different central nervous system cell types display distinct and nonrandom arrangements of satellite DNA sequences. *Proc.Natl.Acad.Sci U.S.A* **81**, 3123-3127.
- Manuelidis, L.** (1985). Individual interphase chromosome domains revealed by in situ hybridization. *Hum.Genet.* **71**, 288-293.
- Manuelidis, L. and Borden, J.** (1988). Reproducible compartmentalization of individual chromosome domains in human CNS cells revealed by in situ hybridization and three-dimensional reconstruction. *Chromosoma* **96**, 397-410.
- Manuelidis, L.** (1990). A view of interphase chromosomes. *Science* **250** , 1533-1540.
- Manuelidis, L. and Chen, T. L.** (1990). A unified model of eukaryotic chromosomes. *Cytometry* **11**, 8-25.
- Marmorstein, R.** (2001). Protein modules that manipulate histone tails for chromatin regulation. *Nat.Rev.Mol.Cell Biol.* **2**, 422-432.
- Marshall, W. F., Dernburg, A. F., Harmon, B., Agard, D. A. , and Sedat, J. W.** (1996). Specific interactions of chromatin with the nuclear envelope: positional determination within the nucleus in *Drosophila melanogaster*. *Mol.Biol.Cell* **7**, 825-842.
- Marshall, W. F., Straight, A., Marko, J. F., Swedlow, J., Dernburg, A., Belmont, A., Murray, A. W., Agard, D. A., and Sedat, J. W.** (1997). Interphase chromosomes undergo constrained diffusional motion in living cells. *Curr.Biol.* **7**, 930-939.
- Mattson, B. A. and Albertini, D. F.** (1990). Oogenesis: chromatin and microtubule dynamics during meiotic prophase. *Mol.Reprod.Dev.* **25**, 374-383.
- McDowell, T. L., Gibbons, R. J., Sutherland, H., O'Rourke, D. M., Bickmore, W. A., Pombo, A., Turley, H., Gatter, K., Picketts, D. J., Buckle, V. J., Chapman, L., Rhodes, D., and Higgs, D. R.** (1999). Localization of a putative transcriptional regulator (ATRX) at pericentromeric heterochromatin and the short arms of acrocentric chromosomes. *Proc.Natl.Acad.Sci U.S.A* **96**, 13983-13988.
- Meller, V. H.** (2000). Dosage compensation: making 1X equal 2X. *Trends Cell Biol.* **10**, 54-59.
- Mermoud, J. E., Costanzi, C., Pehrson, J. R., and Brockdorff, N.** (1999). Histone macroH2A1.2 relocates to the inactive X chromosome after initiation and propagation of X-inactivation. *J.Cell Biol.* **147**, 1399-1408.
- Metz, C.** Chromosome studies on the Diptera II. The paired association of chromosomes in the Diptera and its significance (1916). *J.Exp Zoology* **21**, 213-279.
- Miklos, G. L. and John, B.** (1979). Heterochromatin and satellite DNA in man: properties and prospects. *Am.J.Hum.Genet.* **31**, 264-280.
- Milankov, K. and De Boni, U.** (1993). Cytochemical localization of actin and myosin aggregates in interphase nuclei in situ. *Exp Cell Res.* **209**, 189-199.

- Miles, C., Elgar, G., Coles, E., Kleinjan, D. J., van, H., V, and Hastie, N.** (1998). Complete sequencing of the Fugu WAGR region from WT1 to PAX6: dramatic compaction and conservation of synteny with human chromosome 11p13. *Proc.Natl.Acad.Sci U.S.A* **95**, 13068-13072.
- Morales, V. and Richard-Foy, H.** (2000). Role of histone N-terminal tails and their acetylation in nucleosome dynamics. *Mol.Cell Biol.* **20**, 7230-7237.
- Munkel, C., Eils, R., Dietzel, S., Zink, D., Mehring, C., Wedemann, G., Cremer, T., and Langowski, J.** (1999). Compartmentalization of interphase chromosomes observed in simulation and experiment. *J.Mol.Biol.* **285**, 1053-1065.
- Nagele, R. G., Velasco, A. Q., Anderson, W. J., McMahon, D. J., Thomson, Z., Fazekas, J., Wind, K., and Lee, H.** (2001). Telomere associations in interphase nuclei: possible role in maintenance of interphase chromosome topology. *J.Cell Sci.* **114**, 377-388.
- Nan, X., Ng, H. H., Johnson, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N., and Bird, A.** (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393**, 386-389.
- Neumann, B., Kubicka, P., and Barlow, D. P.** (1995). Characteristics of imprinted genes. *Nat.Genet.* **9**, 12-13.
- Ng, H. H. and Bird, A.** (1999). DNA methylation and chromatin modification. *Curr.Opin.Genet.Dev.* **9**, 158-163.
- Ng, H. H., Zhang, Y., Hendrich, B., Johnson, C. A., Turner, B. M., Erdjument-Bromage, H., Tempst, P., Reinberg, D., and Bird, A.** (1999). MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nat.Genet.* **23**, 58-61.
- Ng, H. H. and Bird, A.** (2000). Histone deacetylases: silencers for hire. *Trends Biochem.Sci* **25**, 121-126.
- Ng, H. H., Jeppesen, P., and Bird, A.** (2000). Active repression of methylated genes by the chromosomal protein MBD1. *Mol.Cell Biol.* **20**, 1394-1406.
- Nickoloff, J. A.** (1992). Transcription enhances intrachromosomal homologous recombination in mammalian cells. *Mol.Cell Biol.* **12**, 5311-5318.
- Niederfuhr, A., Hummerich, H., Gawin, B., Boyle, S., Little, P. F., and Gessler, M.** (1998). A sequence-ready 3-Mb PAC contig covering 16 breakpoints of the Wilms tumor/anirida region of human chromosome 11p13. *Genomics* **53**, 155-163.
- Nielsen, S. J., Schneider, R., Bauer, U. M., Bannister, A. J., Morrison, A., O'Carroll, D., Firestein, R., Cleary, M., Jenuwein, T., Herrera, R. E., and Kouzarides, T.** (2001). Rb targets histone H3 methylation and HP1 to promoters. *Nature* **412**, 561-565.
- Nishina, S., Kohsaka, S., Yamaguchi, Y., Handa, H., Kawakami, A., Fujisawa, H., and Azuma, N.** (1999). PAX6 expression in the developing human eye. *Br.J.Ophthalmol.* **83**, 723-727.

- Norman, A. M., Read, A. P., Clayton-Smith, J., Andrews, T. , and Donnai, D.** (1992). Recurrent Wiedemann-Beckwith syndrome with inversion of chromosome (11)(p11.2p15.5). *Am.J.Med.Genet.* **42**, 638-641.
- O'Keefe, R. T., Henderson, S. C., and Spector, D. L.** (1992). Dynamic organization of DNA replication in mammalian cell nuclei: spatially and temporally defined replication of chromosome-specific alpha-satellite DNA sequences . *J.Cell Biol.* **116**, 1095-1110.
- O'Neill, D. W., Schoetz, S. S. , Lopez, R. A., Castle, M., Rabinowitz, L., Shor, E., Krawchuk, D., Goll, M. G., Renz, M., Seelig, H. P., Han, S., Seong, R. H., Park, S. D., Agalioti, T., Munshi, N., Thanos, D., Erdjument-Bromage, H., Tempst, P., and Bank, A.** (2000). An ikaros-containing chromatin-remodeling complex in adult-type erythroid cells. *Mol. Cell Biol.* **20**, 7572-7582.
- O'Neill, L. P. and Turner, B. M.** (1995). Histone H4 acetylation distinguishes coding regions of the human genome from heterochromatin in a differentiation-dependent but transcription- independent manner. *EMBO J.* **14**, 3946-3957.
- O'Neill, L. P., Keohane, A. M., Lavender, J. S., McCabe, V., Heard, E., Avner, P., Brockdorff, N., and Turner, B. M.** (1999). A developmental switch in H4 acetylation upstream of Xist plays a role in X chromosome inactivation. *EMBO J.* **18**, 2897-2907.
- Ohsumi, K., Katagiri, C., and Kishimoto, T.** (1993). Chromosome condensation in *Xenopus* mitotic extracts without histone H1. *Science* **262**, 2033-2035.
- Okano, M., Bell, D. W., Haber, D. A., and Li, E.** (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247-257.
- Okano, M., Xie, S., and Li, E.** (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat.Genet.* **19**, 219-220.
- Onyango, P., Miller, W., Lehoczky, J., Leung, C. T., Birren, B., Wheelan, S., Dewar, K., and Feinberg, A. P.** (2000). Sequence and comparative analysis of the mouse 1-megabase region orthologous to the human 11p15 imprinted domain. *Genome Res.* **10**, 1697-1710.
- Owen, D. J., Ornaghi, P., Yang, J. C., Lowe, N., Evans, P. R., Ballario, P., Neuhaus, D., Filetici, P., and Travers, A. A.** (2000). The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase gcn5p. *EMBO J.* **19**, 6141-6149.
- Pagter-Holthuizen, P., Jansen, M., van Schaik, F. M., van der, K. R., Oosterwijk, C., Van den Brande, J. L., and Sussenbach, J. S.** (1987). The human insulin-like growth factor II gene contains two development- specific promoters. *FEBS Lett.* **214**, 259-264.
- Park, P. C. and De Boni, U.** (1998a). A specific conformation of the territory of chromosome 17 locates ERBB- 2 sequences to a DNase-hypersensitive domain at the nuclear periphery. *Chromosoma* **107**, 87-95.

- Park, P. C. and De Boni, U.** (1998b). A specific conformation of the territory of chromosome 17 locates ERBB-2 sequences to a DNase-hypersensitive domain at the nuclear periphery. *Chromosoma* **107**, 87-95.
- Parreira, L., Telhada, M., Ramos, C., Hernandez, R., Neves, H., and Carmo-Fonseca, M.** (1997). The spatial distribution of human immunoglobulin genes within the nucleus: evidence for gene topography independent of cell type and transcriptional activity. *Hum.Genet.* **100**, 588-594.
- Paulsen, M., Davies, K. R., Bowden, L. M., Villar, A. J., Franck, O., Fuermann, M., Dean, W. L., Moore, T. F., Rodrigues, N., Davies, K. E., Hu, R. J., Feinberg, A. P., Maher, E. R., Reik, W., and Walter, J.** (1998). Syntenic organization of the mouse distal chromosome 7 imprinting cluster and the Beckwith-Wiedemann syndrome region in chromosome 11p15.5. *Hum.Mol.Genet.* **7**, 1149-1159.
- Paulsen, M., El Maarri, O., Engemann, S., Strodicke, M., Franck, O., Davies, K., Reinhardt, R., Reik, W., and Walter, J.** (2000). Sequence conservation and variability of imprinting in the beckwith- wiedemann syndrome gene cluster in human and mouse. *Hum.Mol.Genet.* **9**, 1829-1841.
- Paulson, J. R. and Laemmli, U. K.** (1977). The structure of histone-depleted metaphase chromosomes. *Cell* **12**, 817-828.
- Pederson, T.** (2000). Half a century of "the nuclear matrix". *Mol.Biol.Cell* **11**, 799-805.
- Peterson, C. L. and Workman, J. L.** (2000). Promoter targeting and chromatin remodeling by the SWI/SNF complex. *Curr.Opin.Genet.Dev.* **10**, 187-192.
- Phair, R. D. and Misteli, T.** (2000). High mobility of proteins in the mammalian cell nucleus. *Nature* **404**, 604-609.
- Pinkel, D., Straume, T., and Gray, J. W.** (1986). Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. *Proc.Natl.Acad.Sci U.S.A* **83**, 2934-2938.
- Pirrotta, V.** (1998). Polycomb the genome: PcG, trxG, and chromatin silencing. *Cell* **93**, 333-336.
- Platero, J. S., Csink, A. K., Quintanilla, A., and Henikoff, S.** (1998b). Changes in chromosomal localization of heterochromatin-binding proteins during the cell cycle in *Drosophila*. *J.Cell Biol.* **140**, 1297-1306.
- Platero, J. S., Csink, A. K., Quintanilla, A., and Henikoff, S.** (1998a). Changes in chromosomal localization of heterochromatin-binding proteins during the cell cycle in *Drosophila*. *J.Cell Biol.* **140**, 1297-1306.
- Preumont, A. M., Capone, B., and van Gansen, P.** (1983). Replicative activity and actinomycin binding in mouse diploid fibroblasts (in vitro ageing). *Mech.Ageing Dev.* **22**, 167-177.
- Rabl, C.** (1885). Über Zelltheilung. *Morphologisches Jahrbuch* **10**, 214-330. 1885.

- Rappold, G. A., Cremer, T., Hager, H. D., Davies, K. E., Muller, C. R., and Yang, T.** (1984). Sex chromosome positions in human interphase nuclei as studied by in situ hybridization with chromosome specific DNA probes. *Hum.Genet.* **67**, 317-325.
- Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z. W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C. P., Allis, C. D., and Jenuwein, T.** (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**, 593-599.
- Redeker, E., Hoovers, J. M., Alders, M., van Moorsel, C. J., Ivens, A. C., Gregory, S., Kalikin, L., Blik, J., de Galan, L., and van den, B. R.** (1994). An integrated physical map of 210 markers assigned to the short arm of human chromosome 11. *Genomics* **21**, 538-550.
- Reichenzeller, M., Burzlaff, A., Lichter, P., and Herrmann, H.** (2000). In vivo observation of a nuclear channel-like system: evidence for a distinct interchromosomal domain compartment in interphase cells. *J.Struct.Biol.* **129**, 175-185.
- Reid, L. H., Davies, C., Cooper, P. R., Crider-Miller, S. J., Sait, S. N., Nowak, N. J., Evans, G., Stanbridge, E. J., deJong, P., Shows, T. B., Weissman, B. E., and Higgins, M. J.** (1997). A 1-Mb physical map and PAC contig of the imprinted domain in 11p15.5 that contains TAPA1 and the BWSCR1/WT2 region. *Genomics* **43**, 366-375.
- Reik, W. and Walter, J.** (2001). Genomic imprinting: parental influence on the genome. *Nat.Rev.Genet.* **2**, 21-32.
- Reuter, G. and Spierer, P.** (1992). Position effect variegation and chromatin proteins. *Bioessays* **14**, 605-612.
- Rice, J. C. and Allis, C. D.** (2001). Histone methylation versus histone acetylation: new insights into epigenetic regulation. *Curr.Opin.Cell Biol.* **13**, 263-273.
- Riesselmann, L. and Haaf, T.** (1999). Preferential S-phase pairing of the imprinted region on distal mouse chromosome 7. *Cytogenet.Cell Genet.* **86**, 39-42.
- Robinett, C. C., Straight, A., Li, G., Wilhelm, C., Sudlow, G., Murray, A., and Belmont, A. S.** (1996). In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. *J.Cell Biol.* **135**, 1685-1700.
- Roeder, G. S.** (1995). Sex and the single cell: meiosis in yeast. *Proc.Natl.Acad.Sci.U.S.A* **92**, 10450-10456.
- Rose, E. A., Glaser, T., Jones, C., Smith, C. L., Lewis, W. H., Call, K. M., Minden, M., Champagne, E., Bonetta, L., Yeger, H., and .** (1990). Complete physical map of the WAGR region of 11p13 localizes a candidate Wilms' tumor gene. *Cell* **60**, 495-508.
- Roth, S. Y., Denu, J. M., and Allis, C. D.** (2001). Histone Acetyltransferases. *Annu.Rev.Biochem.* **70**, 81-120.

- Sabbattini, P., Lundgren, M., Georgiou, A., Chow, C., Warnes, G., and Dillon, N.** (2001). Binding of Ikaros to the lambda5 promoter silences transcription through a mechanism that does not require heterochromatin formation. *EMBO J.* **20**, 2812-2822.
- Sachs, R. K., van den, E. G., Trask, B., Yokota, H., and Hearst, J. E.** (1995). A random-walk/giant-loop model for interphase chromosomes. *Proc.Natl.Acad.Sci.U.S.A* **92**, 2710-2714.
- Sachs, R. K., Levy, D., Chen, A. M., Simpson, P. J., Cornforth, M. N., Ingerman, E. A., Hahnfeldt, P., and Hlatky, L. R.** (2000). Random breakage and reunion chromosome aberration formation model; an interaction-distance version based on chromatin geometry. *Int.J.Radiat.Biol.* **76**, 1579-1588.
- Sadoni, N., Langer, S., Fauth, C., Bernardi, G., Cremer, T., Turner, B. M., and Zink, D.** (1999). Nuclear organization of mammalian genomes. Polar chromosome territories build up functionally distinct higher order compartments. *J.Cell Biol.* **146**, 1211-1226.
- Saurin, A. J., Shiels, C., Williamson, J., Satijn, D. P., Otte, A. P., Sheer, D., and Freemont, P. S.** (1998). The human polycomb group complex associates with pericentromeric heterochromatin to form a novel nuclear domain. *J.Cell Biol.* **142**, 887-898.
- Sawyer, J. R., Moore, M. M., and Hozier, J. C.** (1987). High resolution G-banded chromosomes of the mouse. *Chromosoma* **95**, 350-358.
- Schardin, M., Cremer, T., Hager, H. D., and Lang, M.** (1985). Specific staining of human chromosomes in Chinese hamster x man hybrid cell lines demonstrates interphase chromosome territories. *Hum.Genet.* **71**, 281-287.
- Scherthan, H., Eils, R., Trelles-Sticken, E., Dietzel, S., Cremer, T., Walt, H., and Jauch, A.** (1998). Aspects of three-dimensional chromosome reorganization during the onset of human male meiotic prophase. *J.Cell Sci.* **111**, 2337-2351.
- Schmidt, J. V., Matteson, P. G., Jones, B. K., Guan, X. J., and Tilghman, S. M.** (2000). The Dlk1 and Gtl2 genes are linked and reciprocally imprinted. *Genes Dev.* **14**, 1997-2002.
- Schubeler, D., Francastel, C., Cimborra, D. M., Reik, A., Martin, D. I., and Groudine, M.** (2000). Nuclear localization and histone acetylation: a pathway for chromatin opening and transcriptional activation of the human beta-globin locus. *Genes Dev.* **14**, 940-950.
- Schul, W., Groenhout, B., Koberna, K., Takagaki, Y., Jenny, A., Manders, E. M., Raska, I., van Driel, R., and de Jong, L.** (1996). The RNA 3' cleavage factors CstF 64 kDa and CPSF 100 kDa are concentrated in nuclear domains closely associated with coiled bodies and newly synthesized RNA. *EMBO J.* **15**, 2883-2892.
- Schul, W., van, D. K., I, Matera, A. G., van Driel, R., and de Jong, L.** (1999). Nuclear domains enriched in RNA 3'-processing factors associate with coiled bodies and histone genes in a cell cycle-dependent manner. *Mol.Biol.Cell* **10**, 3815-3824.

- Sedat, J. and Manuelidis, L.** (1978). A direct approach to the structure of eukaryotic chromosomes. *Cold Spring Harb. Symp. Quant. Biol.* **42 Pt 1**, 331-350.
- Shelby, R. D., Hahn, K. M., and Sullivan, K. F.** (1996). Dynamic elastic behavior of alpha-satellite DNA domains visualized in situ in living human cells. *J. Cell Biol.* **135**, 545-557.
- Shen, X., Yu, L., Weir, J. W., and Gorovsky, M. A.** (1995). Linker histones are not essential and affect chromatin condensation in vivo. *Cell* **82**, 47-56.
- Skok, J. A., Brown, K. E., Azuara, V., Caparros, M. L., Baxter, J., Takacs, K., Dillon, N., Gray, D., Perry, R. P., Merckenschlager, M., and Fisher, A. G.** (2001). Nonequivalent nuclear location of immunoglobulin alleles in B lymphocytes. *Nat. Immunol.* **2**, 848-854.
- Smith, J. S., Brachmann, C. B., Celic, I., Kenna, M. A., Muhammad, S., Starai, V. J., Avalos, J. L., Escalante-Semerena, J. C., Grubmeyer, C., Wolberger, C., and Boeke, J. D.** (2000). A phylogenetically conserved NAD⁺-dependent protein deacetylase activity in the Sir2 protein family. *Proc. Natl. Acad. Sci. U.S.A* **97**, 6658-6663.
- Smith, K. P., Carter, K. C., Johnson, C. V., and Lawrence, J. B.** (1995). U2 and U1 snRNA gene loci associate with coiled bodies. *J. Cell Biochem.* **59**, 473-485.
- Smith, Z. E. and Higgs, D. R.** (1999). The pattern of replication at a human telomeric region (16p13.3): its relationship to chromosome structure and gene expression. *Hum. Mol. Genet.* **8**, 1373-1386.
- Sparvoli, E., Levi, M., and Rossi, E.** (1994). Replicon clusters may form structurally stable complexes of chromatin and chromosomes. *J. Cell Sci* **107**, 3097-3103.
- Spector, D. L.** (1996). Nuclear organization and gene expression. *Exp. Cell Res.* **229**, 189-197.
- Sterner, D. E. and Berger, S. L.** (2000). Acetylation of histones and transcription-related factors. *Microbiol. Mol. Biol. Rev.* **64**, 435-459.
- Stout, K., van der, M. S., Frants, R. R., Padberg, G. W., Ropers, H. H., and Haaf, T.** (1999). Somatic pairing between subtelomeric chromosome regions: implications for human genetic disease? *Chromosome. Res.* **7**, 323-329.
- Strachan, T. and Read, A. P.** (1994). PAX genes. *Curr. Opin. Genet. Dev.* **4**, 427-438.
- Strahl, B. D. and Allis, C. D.** (2000). The language of covalent histone modifications. *Nature* **403**, 41-45.
- Strahl, B. D., Ohba, R., Cook, R. G., and Allis, C. D.** (1999). Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in Tetrahymena. *Proc. Natl. Acad. Sci. U.S.A* **96**, 14967-14972.
- Sullivan, G. J., Bridger, J. M., Cuthbert, A. P., Newbold, R. F., Bickmore, W. A., and McStay, B.** (2001). Human acrocentric chromosomes with transcriptionally silent nucleolar organizer regions associate with nucleoli. *EMBO J.* **20**, 2867-2877.

- Sun, H. B. and Yokota, H.** (1999). Correlated positioning of homologous chromosomes in daughter fibroblast cells. *Chromosome Res.* **7**, 603-610.
- Sun, H. B., Shen, J., and Yokota, H.** (2000). Size-dependent positioning of human chromosomes in interphase nuclei. *Biophys.J.* **79**, 184-190.
- Szentirmay, M. N. and Sawadogo, M.** (2000). Spatial organization of RNA polymerase II transcription in the nucleus. *Nucleic Acids Res.* **28**, 2019-2025.
- Tagawa, Y., Nanashima, A., Yasutake, T., Hatano, K., Nishizawa-Takano, J. E., and Ayabe, H.** (1997). Differences in spatial localization and chromatin pattern during different phases of cell cycle between normal and cancer cells. *Cytometry* **27**, 327-335.
- Tajbakhsh, J., Luz, H., Bornfleth, H., Lampel, S., Cremer, C., and Lichter, P.** (2000). Spatial distribution of GC- and AT-rich DNA sequences within human chromosome territories. *Exp Cell Res.* **255**, 229-237.
- Talbert, P. B., LeCiel, C. D., and Henikoff, S.** (1994). Modification of the *Drosophila* heterochromatic mutation brownDominant by linkage alterations. *Genetics* **136**, 559-571.
- Tartof, K. D. and Henikoff, S.** (1991). Trans-sensing effects from *Drosophila* to humans. *Cell* **65**, 201-203.
- Taunton, J., Hassig, C. A., and Schreiber, S. L.** (1996). A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* **272**, 408-411.
- The chromosome 21 mapping and sequencing consortium** (2000). The DNA sequence of human chromosome 21. *Nature* **405**, 311-320.
- The MHC sequencing consortium** (1999). Complete sequence and gene map of a human major histocompatibility complex. *Nature* **401**, 921-923.
- Thoma, F., Koller, T., and Klug, A.** (1979). Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *J. Cell Biol.* **83**, 403-427.
- Thomas, J. O.** (1999). Histone H1: location and role. *Curr. Opin. Cell Biol.* **11**, 312-317.
- Thomson, S., Clayton, A. L., Hazzalin, C. A., Rose, S., Barratt, M. J., and Mahadevan, L. C.** (1999). The nucleosomal response associated with immediate-early gene induction is mediated via alternative MAP kinase cascades: MSK1 as a potential histone H3/HMG-14 kinase. *EMBO J.* **18**, 4779-4793.
- Thon, G. and Klar, A. J.** (1992). The *clr1* locus regulates the expression of the cryptic mating-type loci of fission yeast. *Genetics* **131**, 287-296.
- Tilghman, S. M.** (1999). The sins of the fathers and mothers: genomic imprinting in mammalian development. *Cell* **96**, 185-193.

- Trask, B., Pinkel, D., and van den, E. G.** (1989). The proximity of DNA sequences in interphase cell nuclei is correlated to genomic distance and permits ordering of cosmids spanning 250 kilobase pairs. *Genomics* **5**, 710-717.
- Tsukamoto, T., Hashiguchi, N., Janicki, S. M., Tumber, T., Belmont, A. S., and Spector, D. L.** (2000). Visualization of gene activity in living cells. *Nat. Cell Biol.* **2**, 871-878.
- Tsukiyama, T. and Wu, C.** (1997). Chromatin remodeling and transcription. *Curr. Opin. Genet. Dev.* **7**, 182-191.
- Tumber, T. and Belmont, A. S.** (2001). Interphase movements of a DNA chromosome region modulated by VP16 transcriptional activator. *Nat. Cell Biol.* **3**, 134-139.
- Tumber, T., Sudlow, G., and Belmont, A. S.** (1999). Large-scale chromatin unfolding and remodeling induced by VP16 acidic activation domain. *J. Cell Biol.* **145**, 1341-1354.
- Turner, B. M.** (1998). Histone acetylation as an epigenetic determinant of long-term transcriptional competence. *Cell Mol. Life Sci* **54**, 21-31.
- Turner, B. M.** (2000). Histone acetylation and an epigenetic code. *Bioessays* **22**, 836-845.
- Tyler, J. K. and Kadonaga, J. T.** (1999). The "dark side" of chromatin remodeling: repressive effects on transcription. *Cell* **99**, 443-446.
- van den, E. G., Sachs, R., and Trask, B. J.** (1992). Estimating genomic distance from DNA sequence location in cell nuclei by a random walk model. *Science* **257**, 1410-1412.
- van Driel, R., Wansink, D. G., van Steensel, B., Grande, M. A., Schul, W., and de Jong, L.** (1995). Nuclear domains and the nuclear matrix. *Int. Rev. Cytol.* **162A**, 151-189.
- Varga-Weisz, P. D. and Becker, P. B.** (1998). Chromatin-remodeling factors: machines that regulate? *Curr. Opin. Cell Biol.* **10**, 346-353.
- Venter, J. C., et al.,** (2001). The sequence of the human genome. *Science* **291**, 1304-1351.
- Verschure, P. J., van, D. K., I, Manders, E. M., and van Driel, R.** (1999). Spatial relationship between transcription sites and chromosome territories. *J. Cell Biol.* **147**, 13-24.
- Visser, A. E., Eils, R., Jauch, A., Little, G., Bakker, P. J., Cremer, T., and Aten, J. A.** (1998). Spatial distributions of early and late replicating chromatin in interphase chromosome territories. *Exp. Cell Res.* **243**, 398-407.
- Visser, A. E. and Aten, J. A.** (1999). Chromosomes as well as chromosomal subdomains constitute distinct units in interphase nuclei. *J. Cell Sci.* **112**, 3353-3360.
- Visser, A. E., Jaunin, F., Fakan, S., and Aten, J. A.** (2000). High resolution analysis of interphase chromosome domains. *J. Cell Sci.* **113**, 2585-2593.
- Vogel, F. and Schroeder, T. M.** (1974). The internal order of the interphase nucleus. *Humangenetik.* **25**, 265-297.

- Volpi, E. V., Chevret, E., Jones, T., Vatcheva, R., Williamson, J., Beck, S., Campbell, R. D., Goldsworthy, M., Powis, S. H., Ragoussis, J., Trowsdale, J., and Sheer, D.** (2000). Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei. *J. Cell Sci.* **113**, 1565-1576.
- Vyas, P., Vickers, M. A., Simmons, D. L., Ayyub, H., Craddock, C. F., and Higgs, D. R.** (1992). Cis-acting sequences regulating expression of the human alpha-globin cluster lie within constitutively open chromatin. *Cell* **69**, 781-793.
- Wang, J., Mager, J., Chen, Y., Schneider, E., Cross, J. C., Nagy, A., and Magnuson, T.** (2001). Imprinted X inactivation maintained by a mouse Polycomb group gene. *Nat. Genet.* **28**, 371-375.
- Wansink, D. G., Schul, W., van, D. K., I, van Steensel, B., van Driel, R., and de Jong, L.** (1993). Fluorescent labeling of nascent RNA reveals transcription by RNA polymerase II in domains scattered throughout the nucleus. *J. Cell Biol.* **122**, 283-293.
- Wansink, D. G., Sibon, O. C., Cremers, F. F., van Driel, R., and de Jong, L.** (1996). Ultrastructural localization of active genes in nuclei of A431 cells. *J. Cell Biochem.* **62**, 10-18.
- 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.
- Wei, X., Samarabandu, J., Devdhar, R. S., Siegel, A. J., Acharya, R., and Berezney, R.** (1998). Segregation of transcription and replication sites into higher order domains. *Science* **281**, 1502-1506.
- Weksberg, R., Teshima, I., Williams, B. R., Greenberg, C. R., Pueschel, S. M., Chernos, J. E., Fowlow, S. B., Hoyme, E., Anderson, I. J., and Whiteman, D. A.** (1993). Molecular characterization of cytogenetic alterations associated with the Beckwith-Wiedemann syndrome (BWS) phenotype refines the localization and suggests the gene for BWS is imprinted. *Hum. Mol. Genet.* **2**, 549-556.
- Widom, J.** (1998). Chromatin structure: linking structure to function with histone H1. *Curr. Biol.* **8**, R788-R791.
- Willard, H. F. and Salz, H. K.** (1997). Remodelling chromatin with RNA [news; comment]. *Nature* **386**, 228-229.
- Williamson, K. A.** Involvement of the WT1 and PAX6 genes in Wilms' Tumour and Human Malignant Mesothelioma (1996). *PhD Thesis*. Edinburgh University.
- Winston, F. and Allis, C. D.** (1999). The bromodomain: a chromatin-targeting module? *Nat. Struct. Biol.* **6**, 601-604.
- Wolffe, A. P. and Hayes, J. J.** (1999). Chromatin disruption and modification. *Nucleic Acids Res.* **27**, 711-720.
- Wolffe, A. P. and Guschin, D.** (2000). Review: chromatin structural features and targets that regulate transcription. *J. Struct. Biol.* **129**, 102-122.

- Woodcock, C. L. and Horowitz, R. A.** (1995). Chromatin organization re-viewed. *Trends Cell Biol.* **5**, 272-277.
- Woodcock, C. L. and Dimitrov, S.** (2001). Higher-order structure of chromatin and chromosomes. *Curr.Opin.Genet.Dev.* **11**, 130-135.
- Workman, J. L. and Kingston, R. E.** (1998). Alteration of nucleosome structure as a mechanism of transcriptional regulation. *Annu.Rev.Biochem.* **67**, 545-579.
- Wu, C. T. and Morris, J. R.** (1999). Transvection and other homology effects. *Curr.Opin.Genet.Dev.* **9**, 237-246.
- Wu, J. and Grunstein, M.** (2000). 25 years after the nucleosome model: chromatin modifications. *Trends Biochem.Sci* **25**, 619-623.
- Xing, Y., Johnson, C. V., Dobner, P. R., and Lawrence, J. B.** (1993). Higher level organization of individual gene transcription and RNA splicing [see comments]. *Science* **259**, 1326-1330.
- Xing, Y., Johnson, C. V., Moen, P. T., Jr., McNeil, J. A., and Lawrence, J.** (1995). Nonrandom gene organization: structural arrangements of specific pre- mRNA transcription and splicing with SC-35 domains. *J.Cell Biol.* **131**, 1635-1647.
- Xu, L., Glass, C. K., and Rosenfeld, M. G.** (1999). Coactivator and corepressor complexes in nuclear receptor function. *Curr.Opin.Genet.Dev.* **9**, 140-147.
- Ye, Q., Callebaut, I., Pezhman, A., Courvalin, J. C., and Worman, H. J.** (1997). Domain-specific interactions of human HP1-type chromodomain proteins and inner nuclear membrane protein LBR. *J.Biol.Chem.* **272**, 14983-14989.
- Yokota, H., Singer, M. J., van den Engh, G. J., and Trask, B. J.** (1997). Regional differences in the compaction of chromatin in human G0/G1 interphase nuclei. *Chromosome.Res.* **5**, 157-166.
- Yokota, H., van den, E. G., Hearst, J. E., Sachs, R. K., and Trask, B. J.** (1995). Evidence for the organization of chromatin in megabase pair-sized loops arranged along a random walk path in the human G0/G1 interphase nucleus. *J.Cell Biol.* **130**, 1239-1249.
- Yu, B. D., Hanson, R. D., Hess, J. L., Horning, S. E., and Korsmeyer, S. J.** (1998). MLL, a mammalian trithorax-group gene, functions as a transcriptional maintenance factor in morphogenesis. *Proc.Natl.Acad.Sci U.S.A* **95**, 10632-10636.
- Zeng, C., Kim, E., Warren, S. L., and Berget, S. M.** (1997). Dynamic relocation of transcription and splicing factors dependent upon transcriptional activity. *EMBO J.* **16**, 1401-1412.
- Zhang, Y., Sun, Z. W., Iratni, R., Erdjument-Bromage, H., Tempst, P., Hampsey, M., and Reinberg, D.** (1998). SAP30, a novel protein conserved between human and yeast, is a component of a histone deacetylase complex. *Mol.Cell* **1**, 1021-1031.

- Zhang, Y. and Reinberg, D.** (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev.* **15**, 2343-2360.
- Zhong, S., Salomoni, P., and Pandolfi, P. P.** (2000). The transcriptional role of PML and the nuclear body. *Nat. Cell Biol.* **2**, E85-E90.
- Zink, D., Cremer, T., Saffrich, R., Fischer, R., Trendelenburg, M. F., Ansorge, W., and Stelzer, E. H.** (1998). Structure and dynamics of human interphase chromosome territories in vivo. *Hum. Genet.* **102**, 241-251.
- Zink, D. and Cremer, T.** (1998). Cell nucleus: chromosome dynamics in nuclei of living cells. *Curr. Biol.* **8**, R321-R324.
- Zink, D., Bornfleth, H., Visser, A., Cremer, C., and Cremer, T.** (1999). Organization of early and late replicating DNA in human chromosome territories. *Exp. Cell Res.* **247**, 176-188.
- Zirbel, R. M., Mathieu, U. R., Kurz, A., Cremer, T., and Lichter, P.** (1993). Evidence for a nuclear compartment of transcription and splicing located at chromosome domain boundaries. *Chromosome. Res.* **1**, 93-106.
- Zlatanova, J., Leuba, S. H., and van Holde, K.** (1999). Chromatin structure revisited. *Crit Rev. Eukaryot. Gene Expr.* **9**, 245-255.