

GENETIC PATHOLOGY OF NON-SMALL CELL LUNG CANCER

BY

STEPHEN K. SIKKINK

**Submitted in fulfilment of the requirements for the degree of Doctor of
Philosophy in the Department of Clinical Dental Sciences and Department
of Pathology, University of Liverpool.**

September 2001

ABSTRACT

Lung cancer is the most widespread human malignancy in the world, and current figures suggests that the incidence is increasing. Lung tumourigenesis is a multistage process in which multiple genetic abnormalities lead to malignant transformation. Evidence suggests that genetic alterations in oncogenes and tumour suppressor genes combined with allelic instability (AI) play a major part in tumour development. The work presented in this thesis is involved with identifying some of the genetic changes that occur during the development of human non-small cell lung cancer (NSCLC).

Bronchial carcinogenesis in squamous lesions is thought to proceed through various metaplastic and dysplastic stages. This work has used laser capture microdissection to isolate and enrich cell subpopulations from the bronchial airway and tumour of multiple consecutive tissue samples within three lungs containing squamous carcinomas. Allelic imbalance analysis using fluorescent multiplex PCR at multiple chromosomal locations revealed AI in common with the tumour within histologically normal and metaplastic bronchial epithelium at multiple regions throughout the airway. The highest frequencies of AI were detected on chromosomes 3p and 5q with AI being more common in airway proximal to the tumour than distal to it. Each lung displayed different patterns of AI within 1cm intervals and separate cell populations from the same topographical region of the lung displayed different patterns of AI, providing further evidence for field carcinogenesis occurring within NSCLC development. This study was extended to include AI analysis of 31 consecutive samples from a whole lung containing a squamous carcinoma. Analysis revealed widespread allelic imbalance for several loci in the primary tumour as well as in the

histologically normal airway of the upper and lower lobes. Surprisingly, allelic imbalance at seven microsatellite loci was found in histologically normal epithelium tissue into the peripheral lung at a distance from the tumour. Immunocytochemical analysis also revealed overexpression of p53 and Cyclin D1 protein within histologically normal bronchial epithelium, thus suggesting early genetic changes in this distal tissue. Allele specific allelic imbalance at D3S1263 was detected in every lobe tested as well as in the tumour thus it maybe suggested that this represents one of the initial genetic alterations in the progression of this individuals tumour due to smoking related damage. Overall this has provided an insight into the initial genetic alterations that occurred in each lung, and also the subsequent genetic pathways followed by cell subpopulations from each region in the airway.

AI of chromosomes 5q, 9p22-p23 and 17 is a frequent and early event in NSCLC. This work has focused on the development of a high throughput fluorescent microsatellite multiplex approach used to analyse chromosome 17 allelic imbalance in a panel of NSCLC tumours based on a previously defined fluorescent microsatellite analysis sensitivities and limits. A high frequency of AI was detected in NSCLC tumours on both 17p and 17q and included three minimal regions of AI located at 17p12-p13 within a distance of 5cM. The frequency of AI on chromosome 17p was significantly higher than 17q and also higher in squamous cell carcinomas when compared to adenocarcinomas. The frequency of AI detected on both chromosome 17 arms was also significantly higher than previously documented studies in NSCLC. This has provided further evidence of the existence of tumour suppressor genes on these chromosome arms involved in the development of NSCLC.

This work also describes AI analysis of a small region of deletion in NSCLC, using 5 polymorphic microsatellite loci located in and around 5q11-q13. Using AI analysis the minimal region of AI/LOH located around the microsatellite loci D5S431 was further reduced to a 1cM genetic distance. NIX and other database analysis mapped multiple EST's and 3 candidate genes within the region.

This study has also focused on efforts to minimise the deleted region between D9S259-D9S1778 and includes the region which was detected by Neville *et al* (1995). The minimal region of AI on chromosome 9p22-p23 was refined to a distance of 2-2.5 Mb in human NSCLC, a region which was further defined using YAC and BAC physical mapping strategies. Analysis of a target gene (MLLT3) located within this region revealed infrequent conformation abnormalities indicative of mutation, suggesting that MLLT3 was not the target gene involved in this tumour type. NIX analysis of the minimal region of AI revealed the position of multiple EST's and two putative target genes indicating that further screening of this region was required to eliminate these as candidate target genes. Several other minimal regions of deletion were also identified within 9p22-p23, especially near the microsatellite loci D9S157, suggesting the presence of a putative target gene near this loci which plays a role in NSCLC.

This work may aid the identification of molecular biomarkers such as specific patterns of AI and genetic abnormalities for the screening of high risk individuals for NSCLC.

DECLARATION

I declare that the work presented in this thesis is my own and has not been submitted, in whole or in part, for any other qualifications, at this institution or any other.

TABLE OF CONTENTS

ABSTRACT 2

DECLARATION 5

TABLE OF CONTENTS 6

LIST OF FIGURES 15

LIST OF TABLES 21

LIST OF ABBREVIATIONS 24

ACKNOWLEDGMENTS 29

CHAPTER 1 : GENERAL INTRODUCTION 30

1.1 The structure of human lungs 31

1.2 Human Lung Cancer 36

1.2.1 Squamous cell carcinoma 36

1.2.2 Small cell lung carcinoma 36

1.2.3 Adenocarcinoma 37

1.2.4 Large cell carcinoma 37

1.2.5 Preneoplastic lesions of the lung 37

1.2.6 Metastasis of lung cancer 38

1.2.7 Incidence and aetiology of human lung cancer 44

1.2.8 Genetic predisposition to lung cancer 46

1.3 Molecular changes associated with cancer: The genetic basis of cancer 46

1.3.1 Introduction 46

1.3.2 Tumour Suppressor genes 48

1.3.3 The p53 tumour suppressor gene 51

1.3.4 Oncogenes 53

1.3.5 DNA repair and mutator genes 53

1.4 Detecting genetic abnormalities in human cancers 57

1.4.1 Cytogenetic analysis 57

1.4.2 Linkage analysis 58

1.4.3 Loss of heterozygosity/Allelic imbalance analysis 58

1.4.4 Physical mapping strategies 59

1.4.5 Isolation, identification and analysis of the FHIT gene in the 3p14.2 region 63

1.5 Lung Carcinogenesis 66

1.5.1 Allelic imbalance in human lung cancer 66

1.5.2 Allelic imbalance in preneoplastic lesions 67

1.5.3 Microsatellite instability in lung tumours 72

1.5.4 Chromosome 3p deletions in lung carcinogenesis 76

1.5.5 Chromosome 9 deletions, p15INK4B and p16INK4 abnormalities in NSCLC carcinogenesis 77

1.5.6 p53 mutation in lung carcinogenesis 80

1.6 Field cancerization and clonal evolution in NSCLC tumourigenesis 81

1.7 Comparison of NSCLC with other epithelial malignancies 84

1.8 Aims of project 85

CHAPTER 2: MATERIALS AND METHODS 87

2.1 Tissue preparation, lysis and DNA extraction 88

2.1.1 Ethical Approval 88

2.1.2 Bronchial Sample Collection 88

2.1.3 Paraffin Embedding of Bronchial samples 88

2.1.4 Sectioning of paraffin embedded blocks 89

2.1.5 Haematoxylin and Eosin staining of tissue sections	89
2.1.6 Microdissection of tissue using micromanipulators	89
2.1.7 Microdissection of tissue using Laser Capture Microdissection (LCM)	90
2.1.8 DNA isolation and purification from paraffin tissue	93
2.1.9 DNA isolation from frozen lung tissue	93
2.1.10 DNA isolation from venous blood samples	93
2.2 Non-fluorescent PCR amplification and PAGE	94
2.2.1 Primer design and selection	94
2.2.2 Whole Genome Amplification and Degenerate Oligonucleotide PCR (DOP-PCR)	94
2.2.3 Polymerase Chain Reaction and PAGE	95
2.2.4 Silver Staining of Polyacrylamide gels	96
2.2.5 Methods and criteria for assessing allelic instability using PAGE followed by silver staining	97
2.3 Multiplex analysis using the ABI-310™ and ABI-377™	106
2.3.1 Panel S+ Multiplex amplification	106
2.3.2 Panel F (Chromosome 3) Multiplex amplification	107
2.3.3 Panel P (Chromosome 17) Multiplex amplification	108
2.3.4 Analysis of ABI-310 and ABI-377 results using Genescan™ and Genotyper™ software	112
2.4 Determination of sensitivity limits for scoring AI	128
2.5 Chromosome 9p and 5q analysis in NSCLC	131
2.5.1 Deletion mapping of NSCLC tumours using PAGE and silver staining	131
2.5.2 Contig assembly and microsatellite marker positioning	131

- 2.5.3 YAC culture and DNA extraction 132
- 2.5.4 YAC DNA microsatellite PCR 132
- 2.5.5 BAC culture and DNA extraction 133
- 2.5.6 BAC PCR from human artificial chromosome DNA pools 133
- 2.5.7 BAC PCR 134
- 2.5.8 Candidate Gene and EST identification 135
- 2.5.9 Sequence searching, alignment and BLAST sequence comparison 135
- 2.5.10 NIX analysis 135
- 2.5.11 Single Strand Conformation Polymorphism (SSCP) analysis of Lung DNA 136
- 2.5.12 Sequence analysis of MMLT3 137

2.6 Statistical analysis of data 137

CHAPTER 3: HIGH THROUGHPUT FLUORESCENT MICROSATELLITE ANALYSIS OF CHROMOSOME 17 IN HUMAN NON-SMALL CELL LUNG TUMOURS 138

3.1 Introduction 139

3.2 Fluorescent multiplex analysis of chromosome 17 in NSCLC. 140

- 3.2.1 Methods and sample selection 140
- 3.2.2 Chromosome 17 marker selection 140
- 3.2.3 Analysis of chromosome 17 in NSCLC 144
- 3.2.4 Multiple regions of allelic imbalance of chromosome 17q 156

3.3 Discussion 156

- 3.3.1 High throughput fluorescent analysis revealed high frequencies of allelic instability on chromosome 17 156

3.3.2 Identification of three regions of AI on chromosome 17p provides further evidence for the involvement of 17p12-p13 in NSCLC 161

3.4 Conclusion 162

CHAPTER 4: MOLECULAR ANALYSIS OF RESPIRATORY EPITHELIUM IN TUMOUR-BEARING LUNGS 164

4.1 Introduction 165

4.2 Analysis of cell subpopulations from multiple regions within tumour-bearing lungs 166

4.2.1 Methods, sample selection and sample description 166

4.2.2 Microdissection of cell subpopulations from bronchial tissue 167

4.2.3 Accurate amplification of microsatellite sequences from paraffin-embedded lung tissue after laser capture microdissection 173

4.2.4 Analysis of AI in laser microdissected bronchial tissue using fluorescent microsatellite multiplex PCR 179

4.2.5 Tumour bearing lungs display frequent AI in both bronchial epithelium and tumour 182

4.2.6 Different patterns of AI in tumour bearing lungs 198

4.3 Discussion 202

4.3.1 Accurate isolation of cell subpopulations from paraffin embedded bronchial tissue using LCM and amplification of DNA for AI analysis 202

4.3.2 Analysis of bronchial epithelium and tumour tissue in tumour-bearing lungs revealed multiple regions of AI 205

4.3.3 Patterns of AI in tumour-bearing lungs 206

4.3.4 Individual regions of field cancerization in tumour bearing lungs 208

4.4 Conclusion 215

CHAPTER 5: IN-DEPTH ANALYSIS OF MOLECULAR ALTERATIONS WITHIN NORMAL AND TUMOUR TISSUE FROM AN ENTIRE BRONCHIAL TREE 216

5.1 Introduction and methods 217

5.2 Allelic Imbalance in the tumour and normal bronchial epithelium 220

5.3 Allelic specific imbalance patterns and outgrowth clonality 226

5.4 Expression of p53, Cyclin D1 and PCNA in tumour tissue, normal and metaplastic bronchial epithelium 229

5.5 Discussion 232

5.6 Conclusion 235

CHAPTER 6: ANALYSIS OF CHROMOSOME 9p22-p23 IN HUMAN NON-SMALL CELL LUNG CANCER 237

6.1 Introduction 238

6.2 Deletion mapping of chromosome 9p22-p23 in human NSCLC 239

6.2.1 Methods and sample selection 239

6.2.2 Frequency of AI on chromosome 9p23 in NSCLC 243

6.2.3 Multiple partial deletions in NSCLC on chromosome 9p22-p23 254

6.3 Search for a putative tumour suppressor gene involved in NSCLC tumourigenesis on chromosome 9p22-p23 259

6.3.1 Incomplete YAC positioning of microsatellite loci within the deleted region using the WI and CEPH – Genethon databases 259

6.3.2 Positioning of BACs and microsatellite loci within the SRD using the Sanger ACEDB database and BAC PCR 260

- 6.3.3 Candidate gene and EST search within SRD 262
- 6.3.4 Database analysis of the MLLT3 gene 262
- 6.3.5 SSCP analysis of the MLLT3 gene in NSCLC 274
- 6.3.6 Sequencing of SSCP products 274
- 6.3.7 cDNA analysis of MLLT3 281
- 6.3.8 HGMP NIX analysis of deleted region 286

6.4 Discussion 293

- 6.4.1 AI analysis around 9p22-p23 revealed multiple regions of AI and further reduced the affected chromosomal region 293
- 6.4.2 Accurate positioning of microsatellite marker order within the 9p22-p23 region and selection of a putative target gene for analysis within the minimal region of AI 296
- 6.4.3 Analysis of the MLLT3 gene in NSCLC 298
- 6.4.4 NIX analysis of genomic sequence between D9S1684 and D9S1778 reveals multiple EST's and several candidate genes for analysis 302

6.5 Conclusion 305

CHAPTER 7: ANALYSIS OF CHROMOSOME 5q11-q13 IN HUMAN NON-SMALL CELL LUNG CANCER 306

7.1 Introduction 307

7.2 Analysis of chromosome 5q11-q13 in human NSCLC 308

- 7.2.1 Methods and sample selection 308
- 7.2.2 AI analysis of chromosome 5q11-q13 in human NSCLC 310
- 7.2.3 Multiple minimal regions of AI at 5q11-q13 in NSCLC 316
- 7.2.4 YAC/BAC mapping within the 5q11-q13 region 316

7.2.5 NIX database analysis of genomic DNA sequences within the

SRAI 324

7.2.6 Search for putative candidate genes within the SRAI 324

7.3 Discussion 330

7.3.1 AI analysis of 5q11-q13 further reduced the minimal region of deletion around D5S431 in NSCLC 330

7.3.2 Physical mapping and NIX analysis within the minimal region of AI 332

7.3.3 Candidate genes within the region 5q11-q13 333

7.4 Conclusion 335

CHAPTER 8: CONCLUDING REMARKS AND FUTURE

CONSIDERATIONS 336

CHAPTER 9: REFERENCES 354

APPENDIX A : METHODS AND SOLUTIONS 412

A1 Parrafin embedding of bronchial samples 413

A2 Hematoxylin and Eosin staining of bronchial samples 414

A3 Arcturus LCM Hematoxylin and Eosin staining protocol for LCM 416

A4 Toluidine blue staining protocol for LCM 418

A5 Solution And Buffer Preparation 419

APPENDIX B : CHROMOSOME 9p22-p23 YAC, BAC AND EST

MAPPING DATA 423

APPENDIX C : BLAST ANALYSIS AND SEQUENCE ALIGNMENT OF MLLT3 GENE 434

C1 MLLT3 BLAST search results 435

C2 Sequence comparison for MLLT3 versus Cosmid 34a5 487

C3 Sequence comparison for MLLT3 versus Cosmid 92f5 490

C4 Sequence comparison for MLLT3 versus Cosmid C48 493

C5 Sequence comparison for MLLT3 versus Cosmid 213a7 497

LIST OF FIGURES

CHAPTER 1

Figure 1.1. Figure showing the bronchopulmonary segments in human lungs 32

Figure 1.2.. Diagrammatic representation of the lower respiratory tract 34

Figure 1.3. Photomicrographs displaying the histological features of bronchial tissue in various stages of bronchial carcinogenesis in squamous lesions 40

Figure 1.4. A summary of the cellular and morphological changes that occur during the various stages of bronchial carcinogenesis in squamous lesions 42

CHAPTER 2

Figure 2.1. Overview of Laser Capture Microdissection 91

Figure 2.1. Screen capture from UVIBand software analysis program displaying PAGE image manipulation prior to O.D. analysis 98

Figure 2.2. Screen capture displaying band identification and separation post O. D. analysis 100

Figure 2.3. Screen capture displaying volume reference recalculation for normal (A) and tumour (B) allele peaks 102

Figure 2.4. Screen capture displaying histogrammatic representation of the recalculated % volume of each allele for lung normal (A) and tumour (B) 104

Figure 2.5. Screen capture displaying sample sheet and run control on the ABI-377 prior to sample electrophoresis 114

Figure 2.6. Screen capture displaying gel file (magnified) from ABI-377 analysis 116

Figure 2.7. Screen capture displaying lane tracking for analysis of fluorescent primers after gel electrophoresis on the ABI-377 118

Figure 2.8. Screen capture displaying imported sample file data for size standard (A) and analysis parameter (B) assignment 120

Figure 2.9. Screen capture displaying peak data after Genescan™ analysis and size standard alignment 122

Figure 2.10. Screen capture displaying predefined range in Genotyper™ analysis software for fluorescent dye labelled microsatellite markers in panel P (chromosome 17) 124

Figure 2.11. Screen capture displaying heterozygous DNA sample amplified with D17S1829 after initial (A) and corrected (B) analysis using genotyper macro 126

Figure 2.12. Distribution of the X_{ij} values $[(A_2:A_1 \text{ repeat } i) / (A_2:A_1 \text{ repeat } j)]$ produced from repetition experiments on normal DNA samples to assess interassay variation 129

CHAPTER 3

Figure 3.1. Diagrammatic representation of the current cytogenetic position of microsatellite markers on chromosome 17 used in fluorescent panel 143

Figure 3.2. Examples of heterozygosity, homozygosity, allelic imbalance, and microsatellite instability on chromosome 17 150

Figure 3.3. Histogram displaying the frequency of AI detected for each microsatellite marker on chromosome 17 in 64 NSCLC tumours 152

Figure 3.4. Histogram comparing the mean FRL between adenocarcinoma and squamous cell carcinoma of the lung for 17p and 17q 154

CHAPTER 4

Figure 4.1. Diagrammatic representation of methods used to dissect bronchial airway and tumour samples from multiple locations in surgically resected lung tissue 169

Figure 4.2. Photomicrographs demonstrating the accuracy of laser capture microdissection of normal bronchial epithelial cells and tumour cells from paraffin embedded tissue 171

Figure 4.3. Photomicrograph displaying PCR amplification of microsatellite sequences from laser capture microdissected bronchial epithelium after direct cell lysis 175

Figure 4.4. Comparison of PCR products amplified from a range of blood DNA concentrations using WGA-PCR followed by microsatellite amplification and conventional microsatellite amplification 177

Figure 4.5. Screen capture of ABI-377 gel picture displaying blue 'flare' in bronchial tissue samples created by hematoxylin and eosin staining of bronchial tissue prior to LCM 180

Figure 4.6. Examples of AI at microsatellite loci D3S1263 and D17S1828 for L257 186

Figure 4.7. Examples of AI at microsatellite locus D3S1263 for S12 190

Figure 4.8. Examples of AI at microsatellite locus D5S644 in RC134 194

Figure 4.9. Overall incidence of AI at informative loci in 3 tumour bearing lungs 196

Figure 4.10. Histograms showing comparison of mean FAL for tumour and proximal and distal airways in tumour bearing lungs 200

Figure 4.11. Diagrammatic representation of the distribution of AI detected in the tumour and proximal and distal airways for L257 and S12 211

Figure 4.12. Diagrammatic representation of the distribution of AI detected in the tumour and proximal and distal airways for RC134 213

CHAPTER 5

Figure 5.1. Allelic imbalance analysis of DNA from multi-sampled lung at multiple locations in a patient with lung cancer 222

Figure 5.2. Comparison of allelic imbalance within the main bronchus and upper and lower lobes of the left lung at informative microsatellite loci 224

Figure 5.3. Allele specific AI occurring throughout the whole lung at microsatellite locus D13S171 227

Figure 5.4. Positive p53, Cyclin D1 and PCNA expression within multi-sampled lung 230

CHAPTER 6

Figure 6.1. Diagram displaying the position and order of polymorphic microsatellite markers used for deletion mapping studies on chromosome 9p22-p23 241

Figure 6.2. Figure showing retention of the region 9p22-p23 in 13/56 informative lung tumours from DM1 246

Figure 6.3. Figure showing AI of the region 9p22-p23 in 16/56 informative lung tumours from DM1 248

Figure 6.4. Figure showing AI of the region 9p22-p23 in 13/56 informative lung tumours from DM1 250

Figure 6.5. Figure showing partial deletions of the region 9p22-p23 in 14/56 lung tumours from DM1 and DM2 252

- Figure 6.6.** Figure displaying partially deleted regions detected between the microsatellite loci D9S269 – D9S1778 255
- Figure 6.7.** Photomicrographs displaying minimal regions of deletion between D9S925 and D9S1778 257
- Figure 6.8.** BAC physical map for the region D9S1778 – D9S925 on chromosome 9p23 264
- Figure 6.9.** Photomicrograph displaying agarose gel electrophoresis of BAC DNA PCR products to confirm the position of microsatellite markers and the MLLT3 gene within the region D9S925 – D9S1778 266
- Figure 6.10.** Primary sequence of the MLLT3 gene 270
- Figure 6.11.** Photomicrograph displaying abnormal MLLT3 PCR products after SSCP analysis 275
- Figure 6.12.** Screen captures from Chromas™ displaying nucleotide substitution at 148bp of MLLT3 exon 1 in SCC#H4 277
- Figure 6.13.** Screen captures from Chromas™ displaying nucleotide substitution at 663bp of MLLT3 exon 4 in SCC#J3 279
- Figure 6.14.** Expression of full length cDNA products from exons 4-10 of the MLLT3 gene in various human and foetal tissues 282
- Figure 6.15.** Expression of full length cDNA products from exons 4-10 of the MLLT3 gene in lung tumour and normal pairs 284
- Figure 6.16.** Diagrammatic representation of NIX analysis results for cosmid sequences within the SRD 287
- Figure 6.17.** Diagrammatic representation of NIX analysis results for BAC sequences within the SRD 290

CHAPTER 7

Figure 7.1. Diagrammatic representation of the cytogenetic position and genetic distance between microsatellite loci used in 5q11-q13 AI analysis in NSCLC 309

Figure 7.2. Histogram displaying the frequency of AI at 5q11-q13 in NSCLC 313

Figure 7.3. Diagrammatic representation of the minimal regions of AI detected on 5q11-q13 in seven lung tumours 318

Figure 7.4. Photomicrographs displaying minimal region of AI around D5S431 320

Figure 7.5. Diagrammatic representation of YAC contig WC-1412 322

Figure 7.6. Diagrammatic representation of NIX analysis of genomic sequences within the 5q11-q13 region 325

LIST OF TABLES

CHAPTER 1

Table 1.1. Examples of known tumour suppressor genes involved in human cancers 50

Table 1.2. The location and function of several human oncogenes implicated in lung cancer 56

Table 1.3. Examples of physical mapping strategies commonly used for analysing the human genome 61

Table 1.4. Cloning vectors used in physical mapping and the size of inserted DNA 62

Table 1.5. Summary of the frequencies of allelic imbalance (%AI) detected in various stages of NSCLC carcinogenesis 71

Table 1.6. Summary of microsatellite instability detected in lung cancer 75

CHAPTER 2

Table 2.1. Table showing fluorescent microsatellite panels selected from markers within the Linkage Mapping Set v2.0 110

CHAPTER 3

Table 3.1. Table displaying markers used in fluorescent multiplex panel for chromosome 17 analysis 142

Table 3.2. Allelic imbalance results from 64 NSCLC tumours analysed with 8 microsatellite markers on chromosome 17 146

Table 3.3. Frequency of AI detected on chromosome 17 in this and previous studies and a comparison between histological sub-types in NSCLC 159

CHAPTER 4

Table 4.1. Results of AI analysis for L257 184

Table 4.2. Results of AI analysis for S12 188

Table 4.3. Results of AI analysis for RC134 192

CHAPTER 6

Table 6.1. Table displaying frequency of tumour samples which displayed AI at informative loci tested 245

Table 6.2. Description of the introns and exons in the MLLT3 gene 268

CHAPTER 7

Table 7.1. AI analysis of chromosome 5q11-q13 in 35 NSCLC tumours 311

Table 7.2. Table displaying correlations between AI at two loci on chromosome 5q11-q13 315

Table 7.4. Table displaying candidate genes mapped within the 5q11-q13 region according to the genome database (GDB) 328

APPENDIX B

Table B1. Whitehead Institute database order and YAC contig location of microsatellites loci within the region D9S269 – D9S1778 on chromosome 9p22-p23 424

Table B2. Table displaying Whitehead Institute database YAC hits for microsatellite loci used in DM1 and DM2 426

Table B3. Table displaying CEPH-Genethon database YAC hits for microsatellite loci used for DM1 and DM2 428

Table B4. Table displaying a list of STS's and EST's within the region of D9S157 – D9S1778 according to Genemap '99 database 431

STS/EST ID	Gene/Feature
D9S157	...
D9S158	...
D9S159	...
D9S160	...
D9S161	...
D9S162	...
D9S163	...
D9S164	...
D9S165	...
D9S166	...
D9S167	...
D9S168	...
D9S169	...
D9S170	...
D9S171	...
D9S172	...
D9S173	...
D9S174	...
D9S175	...
D9S176	...
D9S177	...
D9S178	...

LIST OF ABBREVIATIONS

ABR	Active BCR-related gene
AceDB	<i>A. C. elegans</i> database
Adeno	Adenocarcinoma of the lung
AF9	ALL1 fused gene from chromosome 9
AI	Allelic imbalance
APC	Adenomatous polyposis of the colon
ARFD1	ADP-ribosylation factor domain protein 1
BAC	Bacterial artificial chromosome
BLAST	Basic Local Alignment Search Tool
BRCA1	Breast Cancer, type I
BRCA2	Breast Cancer, type II
cDNA	Copy DNA
CEPH	Centre d'Etude du Polymorphisme Humain
CGH	Comparative genomic hybridisation
CIS	Carcinoma <i>in situ</i>
CKN1	Classical Cockayne syndrome
CT	Computed tomography
DCC	Deleted in colorectal carcinoma
DM1	Deletion mapping region 1
DM2	Deletion mapping region 2
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dnc	<i>Drosophila dunce</i> gene
dNTP	deoxynucleoside triphosphate
DOP-PCR	Degenerate oligonucleotide primer – Polymerase chain

	reaction
ENL	Eleven-nineteen leukemia gene
ERCC	Excision repair cross complementation
EST	Expressed sequence tag
FAL	Fractional allelic loss
FAP	Familial adenomatous polyposis
FHIT	Fragile histidine triad
FISH	Fluorescent in-situ hybridisation
FRL	Fractional regional loss
GDB	Genome Data Base
GI	GenInfo Identifier
GRAIL	Gene Recognition Analysis Internet Link
H and E	Hematoxylin and Eosin
HGMP	Human Genome Mapping Project
HGP	Human Genome Project
HIC – 1	Hypermethylated in Cancer – 1
<i>hMLH1</i>	MutL, <i>E. coli</i> , homologue of, 1
<i>hMSH2</i>	<i>MutS</i> , <i>E. coli</i> , human homolog of, 2,
<i>HNPCC</i>	Hereditary nonpolyposis colorectal cancer
HNSCC	Head and neck squamous cell carcinoma
HUGO	Human Genome Organisation
ID	Identification
IFN	Interferon

<i>IFNα</i>	Interferon-alpha
<i>IL6ST</i>	Interleukin 6 signal transducer
<i>ITGA2</i>	Integrin CD49B, alpha 2 subunit
<i>KIF2</i>	Kinesin heavy chain member 2
<i>LCM</i>	Laser Capture Microdissection
<i>LIFE</i>	Laser-induced fluorescence endoscope
<i>LOD</i>	Logarithm of odds
<i>LOH</i>	Loss of heterozygosity
<i>LY64</i>	Lymphocyte antigen 64 (mouse) homolog
<i>MAP1A/MAP1B LC3</i>	Microtubule-associated proteins 1A/1B light chain 3
<i>MCC</i>	Mutated in Colorectal Cancers
<i>MDM2</i>	Mouse double minute 2 homologue
<i>MI</i>	Microsatellite instability
<i>MLL</i>	Myeloid/Lymphoid or Mixed-Lineage Leukemia
<i>MLLT3</i>	Myeloid/Lymphoid or Mixed-Lineage Leukemia (<i>Drosophila Trithorax</i> homolog); translocated to, 3
<i>MMR</i>	Mismatch repair
<i>MRAI</i>	Minimal region of allelic imbalance
<i>mRNA</i>	Messenger RNA
<i>MTAP</i>	Methylthioadenosine phosphorylase
<i>NCBI</i>	National Centre for Biotechnology Information
<i>NCKX2</i>	<i>Homo sapiens</i> cone sodium-calcium potassium exchanger
<i>NER</i>	Nucleotide excision repair
<i>NF1</i>	Neurofibromatosis, type I
<i>NF2</i>	Neurofibromatosis, type II
<i>NSCLC</i>	Non small cell lung cancer

O.D.	Optical density
OMIM	Online Mendelian Inheritance in Man
ORF	Open reading frame
OVCA1	Ovarian Cancer 1
OVCA2	Ovarian Cancer 2
PAC	P1 artificial chromosome
PAGE	Polyacrylamide gel electrophoresis
PCNA	Proliferating Cell Nuclear Antigen
PCR	Polymerase chain reaction
PTC	<i>Patched</i> , human homologue of <i>Drosophila ptc</i> gene
PTEN	Phosphatase and Tensin homologue deleted on chromosome 10
RAB-3B	Human <i>Ras</i> -related protein
RACE	Rapid amplification of cDNA ends
Rb	Retinoblastoma
Rb-1	Retinoblastoma-1
RCICLCR	Roy Castle International Centre for Lung Cancer Research
RNA	Ribonucleic acid
RNase A	Ribonuclease A
SCC	Squamous cell carcinoma
SCLC	Small cell lung cancer
Spiral CT	Spiral computed tomography
SqCCL	Squamous cell carcinoma of the lung
SRD	Smallest region of deletion
SSCP	Single strand conformation polymorphism
STS	Sequence tagged site

<i>TCEB1L</i>	Transcription elongation factor B, 1-like
T_m	Melting temperature of oligonucleotide
TNM	Tumour Node Metastasis
TSG	Tumour suppressor gene
UV	Ultraviolet
<i>VHL</i>	Von Hippel-Lindau syndrome
WGA-PCR	Whole genome amplification polymerase chain reaction
WI	Whitehead Institute
<i>WT1</i>	Wilms tumour 1
WWW	World Wide Web
XP	Xeroderma pigmentosum
XRCC	X-ray repair, complementing defective, in chinese hamster
YAC	Yeast artificial chromosome

ACKNOWLEDGMENTS

I would first and foremost like to thank my supervisors, Professor John K. Field and Dr. John R. Gosney, for their continued support and guidance throughout this PhD and also during the writing of this thesis.

Secondly, thanks to all who worked (at one time or another) in the Roy Castle International Centre for Lung Cancer Research, especially in labs during an enjoyable three years.

Specific thanks to Jill Callaghan (gone but not forgotten!) for technical support with all things histology and those 'Mark and Lard' afternoons; Paul Maloney for ABI-377 and ABI-310 technical assistance and endless debates on footy; Dr. Triantafilos Liloglou (Lakis) for assistance and advice with most of the PhD and facing endless questions even though he was a very busy man; Sarah Lake for technical assistance with the WGA/DOP PCR work; Anna Need for technical assistance with the chromosome 5 work; Dr. Janet Risk for advice during the early days of my YAC and BAC work; Drs. Jim Heighway and Margaret Laversha for answering questions on multiple occasions. And also big thanks to everyone else in the labs at the RCICLCR, including Dr. Teresa Knapp, Shelley Brennand, Joanne Fletcher, Sarah 'Cherries' Kearney, Penny Plater, Asaf Niaz, Simon Critchley and Suzanne Watson, who have helped me during the last four years.

Special thanks also to my parents, family and friends for their constant support during the last four years.

CHAPTER 1: GENERAL INTRODUCTION

1.1 The structure of human lungs

The human respiratory system extends from the nasal cavity to the periphery of the lung and includes the nasal passages and sinuses, the larynx and the lungs themselves. Human lungs are divided into several lobes, the right lung having three lobes (upper, middle and lower) and the left lung having two lobes (upper and lower). Each lung is composed of defined bronchopulmonary segments which are supplied with blood by a segmental artery and with air from the bronchus (figure 1.1). The structure of the lower respiratory tract is composed of the trachea which forms the main airway into both lungs, the main bronchi, bronchioles, alveolar ducts and alveoli (figure 1.2). The trachea is composed of anterior C-shaped plates of cartilage surrounded by smooth muscle which keeps the main airway open and is designed to transport air into the smaller distal airways. Bronchi have a more irregular cartilage and smooth muscle structure and supply clean air to each lung. Both the trachea and bronchi are lined with mucous glands. Bronchioles have no cartilage or submucosal mucous glands, but are lined with ciliated respiratory epithelium. The bronchioles branch until they become less than 2 mm in diameter where they form terminal bronchioles. Further distal to the bronchioles is the acinus or terminal respiratory unit where gas exchange occurs. This unit is composed of alveolar ducts and alveoli which are lined by flattened type I pneumocytes with occasional type II pneumocytes. Type II pneumocytes are more rounded with surface microvilli and secrete surfactant and replicate quickly after damage to the alveolar wall. Beneath the alveolar cells is a thick double layer of fibrous tissue lined by mesothelial cells called the pleura.

Figure 1.1 Figure showing the bronchopulmonary segments in human lungs. Diagrammatic representation of the main airways and segments in both the left and right lung. The trachea branches into the left and right bronchi which in turn branch into 19 pulmonary segments, three in the right upper lobe (RUL), four in the left upper lobe (LUL), two in the middle/lingual lobes (RML) and five in the right lower lobe (RLL). There are four segments in the left lower lobe (LLL). This lobe does not have medial segment since the heart occupies this area on the left side. Figure adapted from M. Sheppard, Practical Pulmonary Pathology.

Figure 1.1 Figure showing the bronchopulmonary segments in human lungs

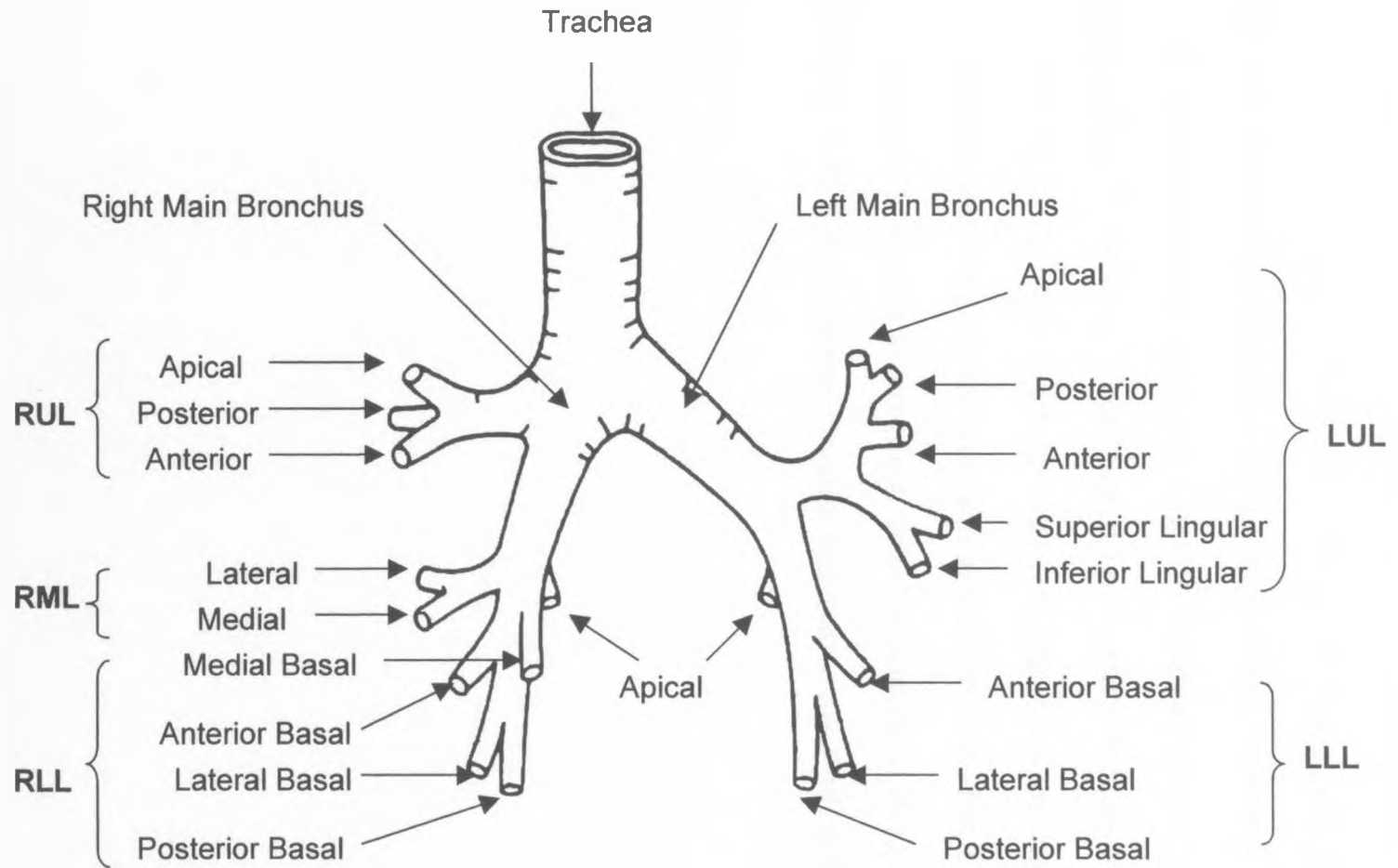
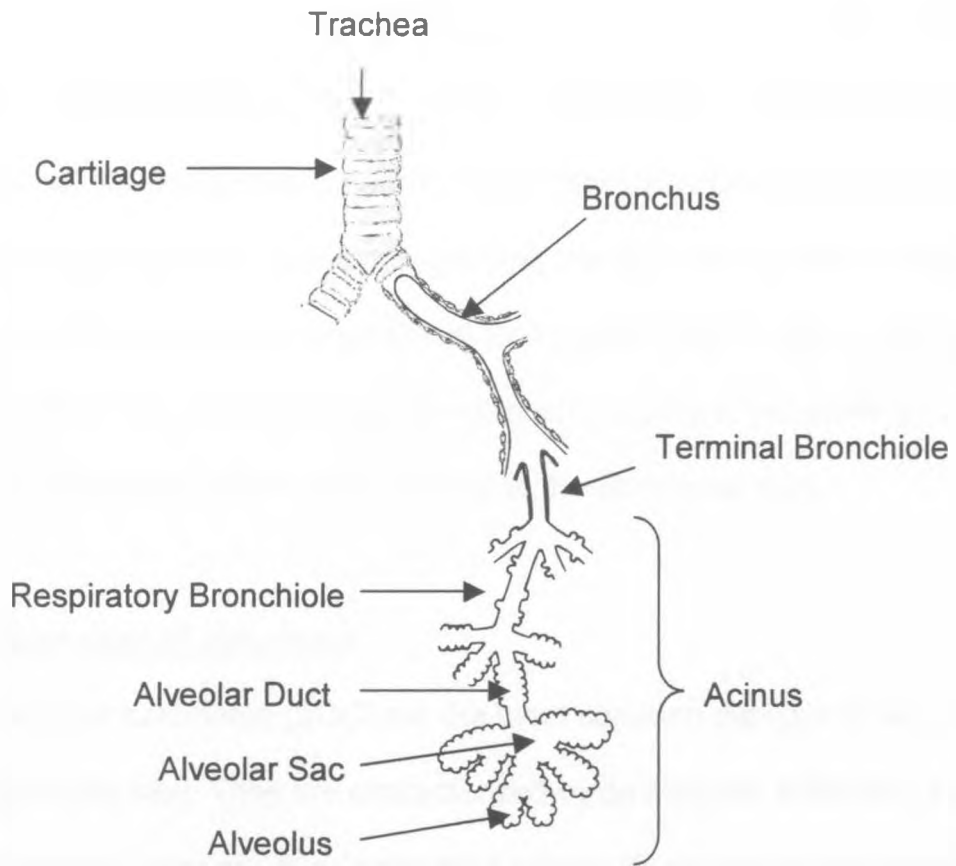


Figure 1.2. Diagrammatic representation of the lower respiratory tract.

Beyond the bronchi in each segment the airway continues to branch, with up to 23 divisions before the air-exchanging region of the peripheral lung. Travelling distal through the lung, cartilage gradually disappears and the airways become bronchioles (the earliest of which has a diameter of 2 mm) and eventually terminal bronchioles within a lobule. Posterior to the terminal bronchioles are the respiratory bronchioles which along with the alveolar ducts and alveoli form the acini. Each lobule contains between 3-10 acini. In turn, each acini contains up to 2000 alveoli. Figure adapted from J. C. E. Underwood. *General and Systematic Pathology* (1996).

Figure 1.2. Diagrammatic representation of the lower respiratory tract.



1.2 Human Lung Cancer

'Lung cancer' comprises four main types of tumour which account for 85% of all tumours which arise in the lungs. These are squamous cell carcinoma (50% of cases), small cell carcinoma (30% of cases), adenocarcinoma (15% of cases) and large cell carcinoma (5% of cases), although figures differ depending on the survey (Underwood, 1996). Studies generally divide these four tumour types into two main histopathological groups, non-small-cell lung cancer (NSCLC) which usually comprises 75% of cases, and small cell lung cancer (SCLC) which represents 25% of cases. Most lung tumours arise from bronchi, usually in the proximal airways, but some adenocarcinomas also develop in the peripheral lung.

1.2.1 Squamous cell carcinoma

Squamous cell carcinoma (SCC) are the most common subtype of lung carcinoma and are usually hilar. They are characterized by an irregular infiltrating edge and histologically by keratinisation (a deeply eosinophilic change in the cytoplasm of nucleated cells with keratin pearl formation and keratohyaline granules) and by 'prickles' (fine intercellular bridges between cells, spaced at regular intervals). Most SCCs show a varying degree of heterogeneity with well differentiated and poorly differentiated cells intermingled within the same region (Travis *et al.*, 1999).

1.2.2 Small cell lung carcinoma

Small cell lung carcinomas (SCLC) are highly aggressive tumours which usually have spread beyond the lung before diagnosis. The tumours are usually centrally located, arising in the hilar region. Histologically, this highly cellular tumour is composed of small ovoid cells (also called 'oat cells' due to their resemblance to

oat grains) with hyperchromatic and indistinct nuclei. The cells are very delicate and can appear smudged and, in some tumours can form interconnecting nests. SCLC is characterised by neuroendocrine differentiation (Travis *et al.*, 1999).

1.2.3 Adenocarcinoma

Adenocarcinomas of the lung are sometimes peripheral and are thought to arise from the peripheral bronchi, bronchioles and alveoli. The majority form glands and acini. Histologically, their cells can range from cuboidal to tall columnar and form a wide range of structural patterns and they often contain mucus (Travis *et al.*, 1999). As with squamous cell carcinoma heterogeneity is common.

1.2.4 Large cell carcinoma

Also called large cell undifferentiated carcinomas, these typically aggressive tumours are usually central, large, bulky masses with extensive areas of necrosis and haemorrhage. Histologically, their cells display a nondescript morphology and architecture with gross nuclear pleomorphism and numerous bizarre mitoses. Most of these tumours are very poorly differentiated SCCs or adenocarcinomas, in which no distinctive characteristics are evident (Travis *et al.*, 1999). Thus, no squamous or glandular differentiation can be seen using light microscopy, although ultrastructural evidence of it may be available (Mackay *et al.*, 1989) .

1.2.5 Preneoplastic lesions of the lung

Bronchial carcinogenesis in squamous lesions is thought to proceed through various hyperplastic (an increase in number of morphologically normal cells) and dysplastic (a distortion of the normal architectural and cytological features)

changes. This is typically seen in metaplastic tissue, squamous metaplasia developing at multiple sites throughout the respiratory tract with intervening 'normal' areas. Dysplasia of this metaplastic epithelium is evident as varying degrees of atypia, from mild to severe (Underwood, 1996). It is thought that these hyperplastic, metaplastic and dysplastic stages are reversible steps in lung carcinogenesis. The incidence of squamous metaplasia was reportedly high (83%) in males with squamous carcinoma and small cell lung cancer, as opposed to females and patients with adenocarcinoma (Tsuchiya *et al.*, 1987). Carcinoma 'in situ' (CIS) refers to a marked nuclear atypia and cellular changes throughout the entire thickness of epithelium in question with multiple mitoses. These abnormal cells express the nuclear and cytoplasmic features of malignant cells, but do not penetrate the subepithelial basement membrane, which remains intact. Once destruction of the basement membrane occurs and the cells invade the deeper tissues, the growth is by definition malignant (Travis *et al.*, 1999). A 50% risk of lung carcinoma has been associated to the presence of severely dysplastic cells in sputum, which is considered a warning signal for underlying malignancy (Risse *et al.*, 1988). A summary of the histological and cellular changes that occur during the various stages of bronchial carcinogenesis is displayed in figures 1.3 and 1.4.

1.2.6 Metastasis of lung cancer

Metastasis is a common event in lung cancer. Metastases can arise from blood, bone or lymphatic spread through the lymph nodes in the centre of the chest. Secondary tumours are often multiple and can occur in the same and contralateral lung. The most common sites for metastasis outside the lungs are the liver, bones, adrenal glands and brain. Even peripheral adenocarcinomas frequently spread to

regional lymph nodes, and tumours less than 3cm in diameter are accompanied by involvement of lymph nodes in upto 50% of cases. Up to two-thirds of patients with SCLC have clinically evident metastasis at the time of diagnosis (Coleman, 1999).

Figure 1.3. Photomicrographs displaying the histological features of bronchial tissue in various stages of bronchial carcinogenesis in squamous lesions. Photomicrograph 1 displays normal bronchial epithelium; 2, Metaplasia (no atypia); 3, Metaplasia (mild atypia); 4, Metaplasia (moderate atypia); 5, Metaplasia (severe atypia)/Carcinoma *in situ*; 6, squamous cell carcinoma. The distinction between severe atypia and carcinoma *in situ* is unclear, however both are involved in the pathogenesis and progression of neoplasia.

Figure 1.3. Photomicrographs displaying the histological features of the various stages of bronchial carcinogenesis in squamous lesions.

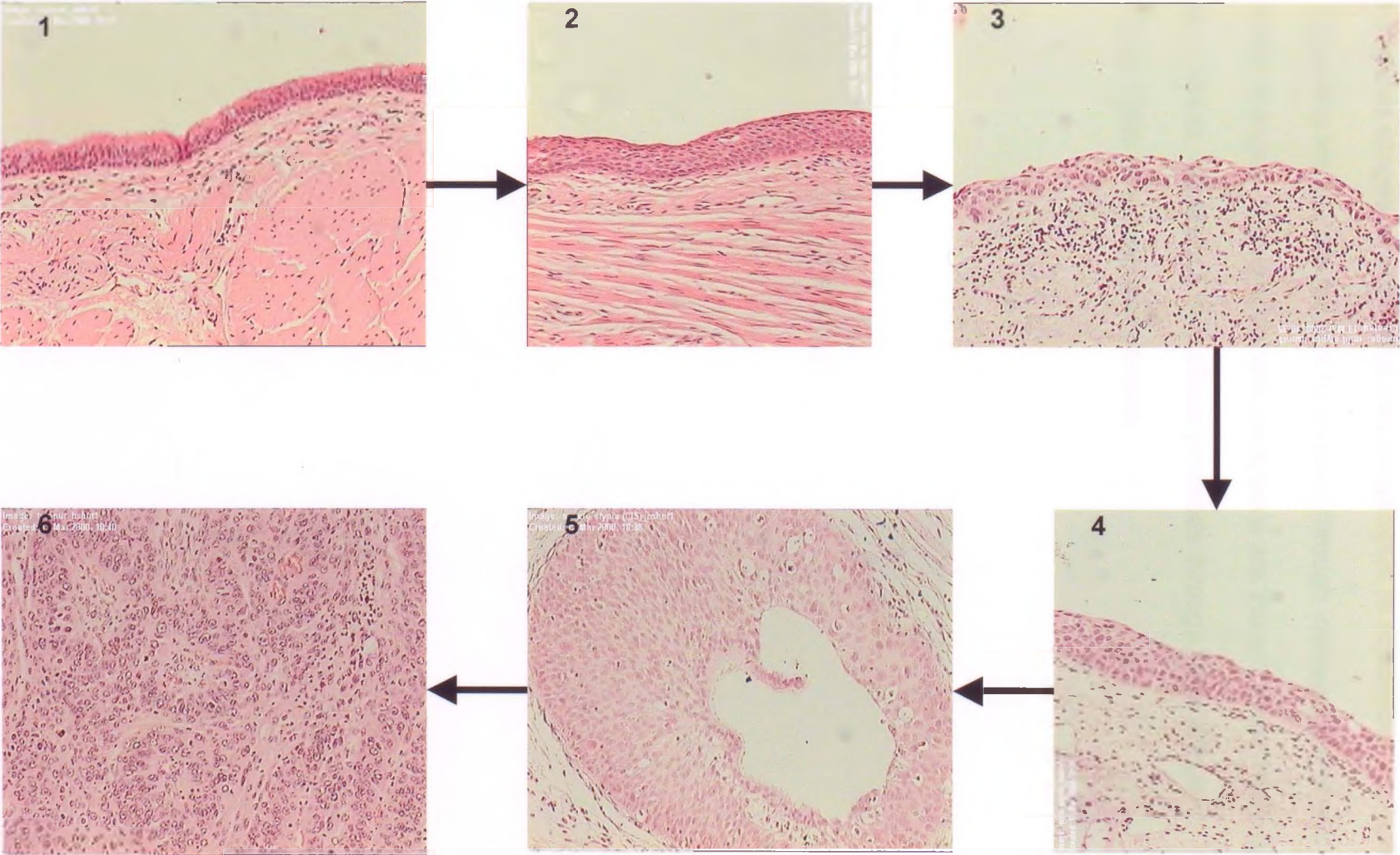
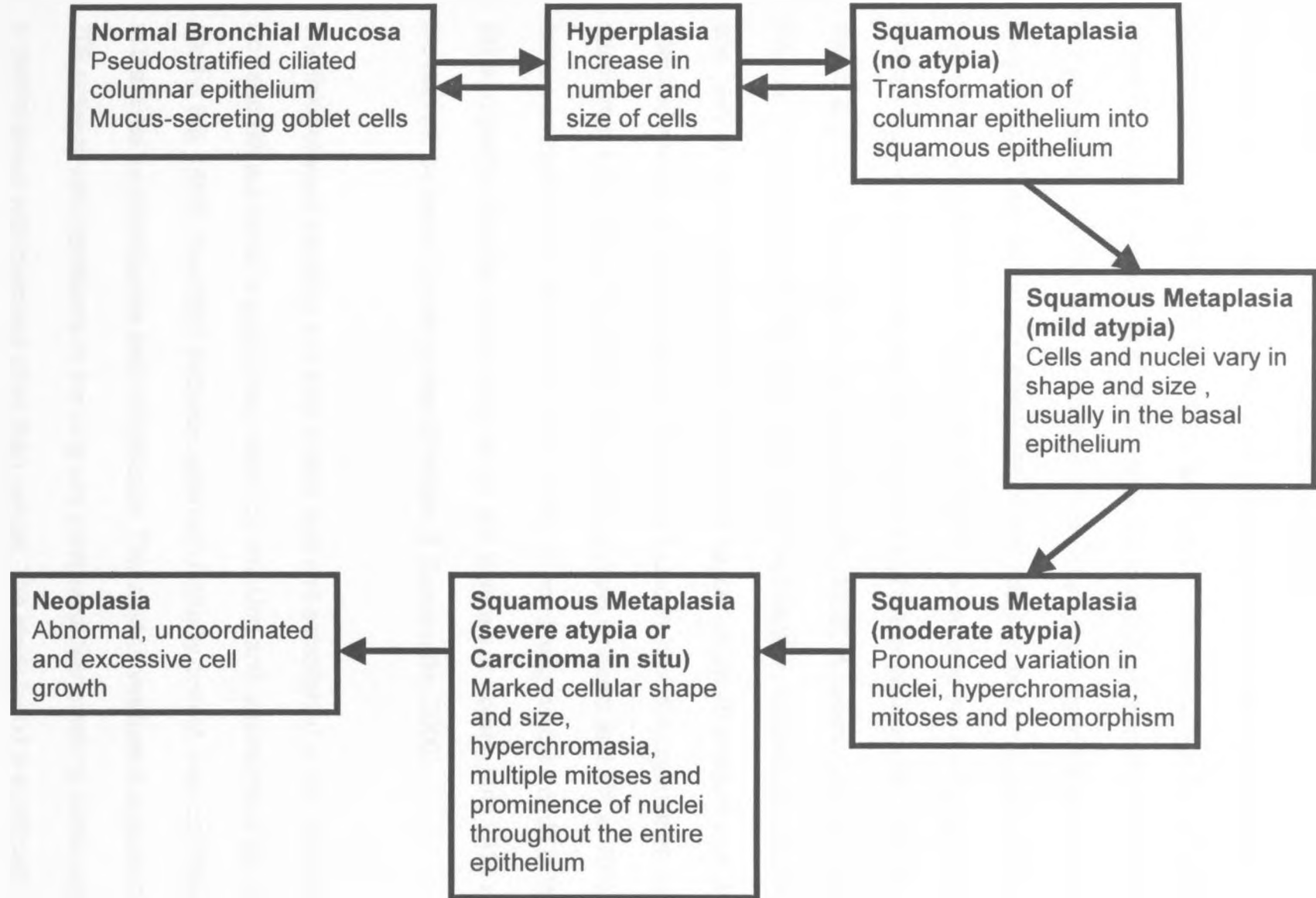


Figure 1.4 A summary of the cellular and morphological changes that occur during the various stages of bronchial carcinogenesis in squamous lesions. Bronchial squamous lesions are thought to proceed through various hyperplastic and metaplastic changes. Hyperplasia and metaplasia are thought to be reversible stages in the process. Squamous metaplasia proceeds through three different stages of atypia (mild, moderate and severe).

Figure 1.4 A summary of the cellular and morphological changes that occur during the various stages of bronchial carcinogenesis in squamous lesions.



1.2.7 Incidence and aetiology of human lung cancer

Lung cancer is the most common human malignancy in the world today and current figures suggest that the incidence of lung cancer is increasing. In western populations this is mostly attributed to cigarette smoking, accounting for over 80% of the disease. The five year survival rate (overall survival rate within five years of diagnosis) of lung cancer in England and Wales has also been estimated to be between 5-10% (Coleman, 1999). The incidence is perhaps most striking in the Mersey region where it is higher than England and Wales as a whole (Mersey Regional Cancer Registry- Lung Cancer Bulletin, 1993). A recent study compared the rate of lung cancer in two large case controlled studies, centred around 1950 and 1990 in hospital patients with and without cancer under 75 years of age, and a matched sample of the population. This study found in 1990 the cumulative risk of death from lung cancer rose from 6% to 16% in male smokers and 1% to 10% in female smokers from 1950 (Peto *et al.*, 2000). Current lung cancer trends in the UK also predict that the rate of lung cancer will increase in women until 2015 when the rate will be equal to that of men (Parsons & Somervaille, 2000).

The link between smoking and lung cancer was well established in the 1950s when Doll and Hill published a preliminary report on smoking and carcinoma of the lung (Doll & Hill, 1950). Pre-1950 evidence, although factually correct, was considered at that time as inconclusive and coincidental. This study investigated a group of 709 patients with carcinoma of the lung and compared their smoking habits against a control group with diseases other than cancer. The study found a significant increase in lung carcinomas in patients who had smoked from an early age and for 20+ years. Doll and Hill also showed a significant decrease in the rate of lung

carcinoma in patients who had stopped smoking for more than ten years. A transition effect between various histological subtypes and tumour location has been linked with a change in smoking habits. Auerbach *et al* (1991) showed a 11.3% increase (30.7% to 42%) in peripheral tumours from 1978-1989, which was also matched by a corresponding 12% decrease (69.3% to 57.3%) in centrally occurring bronchial carcinomas during the same time period. Several recent reports have also described an increase in the incidence of adenocarcinoma to overtake squamous cell carcinoma, especially in the United states (Thun *et al.*, 1997). This is thought to be due to the increased use of reduced tar and nicotine yields in filter cigarettes and smokers compensating with deeper and more frequent inhalations, hence increasing the dosage of carcinogens to the lungs (Djordjevic *et al.*, 1997; Hoffmann & Djordjevic, 1997; Stellman *et al.*, 1997a; Stellman *et al.*, 1997b). There is also increasing evidence that environmental factors such as passive smoking (especially through a smoking relative) and pollution may increase the risk of lung cancer especially amongst women (Cardenas *et al.*, 1997; Choi *et al.*, 1994; Engholm *et al.*, 1996).

Environmental factors that have a causative effect on lung cancer include radiation from atomic weapons used in nuclear attacks, radon gas and mining (especially in uranium miners who show high incidences of squamous and SCLC (Crowell *et al.*, 1996; Land *et al.*, 1993; Lechner *et al.*, 1998)), and asbestos exposure leading to asbestosis and increased risk of mesothelioma (Murthy & Testa, 1999; Xu *et al.*, 1999). Other causative effects may include exposure to organic chemicals and metals such as vinyl chloride, arsenic, cadmium, and nickel, diet and scarring of the lung.

1.2.8 Genetic predisposition to lung cancer.

When studying the causative effects of lung cancer it is impossible to eliminate genetic predisposition as a factor in lung carcinogenesis. Genetic predisposition to lung cancer may occur in three different ways which are not mutually exclusive and can interact with one another to form a predisposition to lung cancer. The first is predisposition caused by gene-environment interactions, including defective carcinogen metabolism due to polymorphism in specific genes (Egeli *et al.*, 1997; el-Zein *et al.*, 1997; Haugen *et al.*, 2000; Shields *et al.*, 1993; Shields & Harris, 2000; Spitz *et al.*, 1999). The second is a predisposition caused by ineffective DNA repair (Xinarianos *et al.*, 2000; Amos *et al.*, 1999; Benachenhou *et al.*, 1998; Spitz *et al.*, 1999; Cheng *et al.*, 2000; Zienolddiny *et al.*, 1999; Wei & Spitz, 1997) and the third is inherited predisposition due to a familial aggregation or positive family history (Bennett *et al.*, 1999a; Broman *et al.*, 2000; Huncharek *et al.*, 1996; Mayne *et al.*, 1999; Sellers *et al.*, 1987; Wood *et al.*, 2000; Yang *et al.*, 1999).

1.3 Molecular genetic changes associated with cancer: The genetic basis of cancer.

1.3.1 Introduction

Tumourigenesis is a multistep process in which multiple genetic abnormalities lead to malignant transformation (Kinzler & Vogelstein, 1996; Yokota & Sugimura, 1993). Characterization of these genetic changes will be useful in determining the biological features necessary for the development of neoplasia and may allow the identification of molecular markers which can provide a foundation for therapeutic approaches. When human cancers are considered, it is evident that the most prominent factor that determines susceptibility to cancer is age (Miller, 1980). This

is explained by the idea that tumours grow by a process of clonal evolution driven by mutation (Nowell, 1976). A single cell could receive a first 'hit' which would produce the clonal expansion of a limited group of cells. As subsequent daughter cells from this group acquire successive mutations and expand they eventually form benign, advanced and malignant states of progression. In 1990, Fearon and Vogelstein provided direct evidence that human colorectal tumours evolved through multiple stages. Colorectal tumours have distinct morphological stages from normal cells to metastasis, which make analysing the order of the mutations possible. For the first time these workers demonstrated the accumulation of 5 or 6 genetic changes such as oncogene activation, tumour suppressor gene inactivation and allele loss in the step-wise progression of human cancer. It should be noted that not all colorectal tumours follow this specific pathway of tumourigenesis (although most tumour types usually include a subset which follow a common order of genetic events). The critical factor is the accumulation of the common genetic changes rather than the order in which they occur. As a tumour progresses, genetic instability can influence the amount of chromosome loss and genetic damage in tumour cells, causing the deletion of influential genes on chromosome arms other than those predicted. This would suggest that only a subset of genetic changes can initiate tumourigenesis in a particular cell type and later mutations confer a selective growth advantage to the cells (reviewed in Vogelstein & Kinzler, 1993). It is generally accepted that 3 main classes of genes are involved in tumourigenesis (tumour suppressor genes, oncogenes and DNA repair genes), although distinctions between the particular genes, their properties, and their classification within a certain group is subject to controversy. The next part of this chapter will provide a brief description of these main classes of genes.

1.3.2 Tumour Suppressor genes

The existence of a distinct class of genes that prevents a cell from becoming tumourigenic was originally discovered as a result of cell fusion experiments between normal and tumour cells (Harris *et al.*, 1969). It was found that in most cell hybrids the transformed phenotype was repaired resulting in a normal cell, and whenever tumour cells did persist it was due to the loss of a specific chromosome(s). These particular genes are classed as tumour suppressor genes. Several examples of known tumour suppressor genes that are involved in human cancer are shown in table 1.1. In 1971 a retinoblastoma (Rb) study using age/incidence curves demonstrated for the first time the concept of tumour suppressor genes (Knudson, 1971). Knudson analysed the ages of children when diagnosed with bilateral (familial) retinoblastoma and compared them with that of unilateral (sporadic) retinoblastoma and formulated that retinoblastoma is induced by two mutations or two 'hits'. In familial Rb the mutated allele (*Rb-1*) is inherited from a parent and needs only one somatic mutation ('one hit') on the other normal allele to initiate tumourigenesis. In sporadic cases both alleles are normal, and the disease occurs through two separate somatic mutations ('two hits'). Knudson stated that if tumourigenesis occurred through a series of genetic changes, occasionally one of these changes could already be present in the germline through parental inheritance. This would give individuals with these changes an increased susceptibility for tumour formation. Subsequent cytogenetic studies associated a small percentage of these individuals with a small visible deletion present on chromosome 13q14 (Knudson *et al.*, 1976), as well as loss of expression of the polymorphic enzyme esterase D which displays close linkage to the *Rb-1* locus (Sparkes *et al.*, 1983; Sparkes & Sparkes, 1983; Sparkes *et al.*,

1980). In 1983 Cavenee et al used restriction fragment length polymorphism (RFLP) analysis to detect loss of the *Rb-1* locus (loss of heterozygosity analysis). This group reasoned that if Knudson was correct, markers which were heterozygous for a chromosome carrying a tumour suppressor gene would often become hemizygous, or even homozygous, in a tumour as a result of a second somatic change which Knudson predicted would be chromosome loss by non-disjunction or homozygous loss resulting from mitotic recombination. This study showed that the affected individuals (for both familial and sporadic retinoblastoma) were heterozygous for one or more of the markers in blood, while the tumour cells were homozygous for the markers, indicating loss of the remaining functional copy of *Rb-1* (the second somatic 'hit'). The next section will discuss *p53* as an example of a tumour suppressor gene.

Tumour Suppressor Gene	Chromosomal locus	Location/function	Inherited disease	Associated cancers
<i>p53</i>	17p13.1	Nuclear/transcription factor	Li-Fraumeni	Most human cancers
<i>Rb-1</i>	13q14	Nuclear/transcription modifier	Familial Retinoblastoma	Retinoblastoma, osteosarcoma, breast, bladder, lung carcinomas
<i>APC</i>	5q21	Cytoplasmic/exact function unknown but binds to β -catenin	Familial adenomatous polyposis (<i>FAP</i>)	Colon, stomach, pancreas and thyroid carcinomas
<i>WT1</i>	11p13	Nuclear/transcription factor	Wilms' tumour	Wilms' tumour
<i>DCC</i>	18q21	Membrane/possible role in cell adhesion	-	Colon carcinoma
<i>NF1</i>	17q11.2	Cytoplasmic/GTPase activating protein	Neurofibromatosis type 1	Neural tumours
<i>NF2</i>	22q12	Cytoplasmic/cytoskeletal membrane linkage	Neurofibromatosis type 2	Central schwannomas and meningiomas
<i>p16^{INK4}</i>	9p21	Nuclear/cyclin dependant kinase inhibitor	Familial melanoma	Melanoma, pancreas, glioblastoma, mesothelioma
<i>VHL</i>	3p25-26	Nuclear/adaptor	Von-Hippel Lindau	Hemangioblastoma, renal cell carcinoma
<i>BRCA1</i>	17q21	Nuclear/BRCA2 cofactor in DNA repair of double strand breaks. Also involved in transcription and recombination	Breast-ovarian cancer	Breast and ovarian cancer
<i>BRCA2</i>	13q12-q13	Nuclear/ cofactor in repair of DNA double-strand breaks	Early onset breast cancer	Breast and ovarian cancer
<i>PTC</i>	9q22.3	Transmembrane protein/ Cell signalling, cell fates, patterning and growth	Gorlin syndrome	Basal cell carcinomas
<i>PTEN</i>	10q23.3	Nuclear/phospholipid phosphatase involved in signal transduction	Cowden disease, Bannayan-Riley-Ruvalcaba syndrome	Endometrial carcinoma, glioblastoma, breast, brain, prostate and ovarian carcinomas

Table 1.1. Examples of known tumour suppressor genes involved in human cancers

1.3.3 The p53 tumour suppressor gene

The p53 tumour suppressor gene maps to the short arm of chromosome 17 and encodes a 53 kDa phosphoprotein that is produced in very low levels in normal cells. The protein was first identified through its ability to complex with the SV40 T-antigen (Lane & Crawford, 1979). Although isolated in its mutated form and originally thought to be an oncogene as overexpression of p53 in cells appeared to cause oncogenic transformation, subsequent studies have proved p53 to be anti-oncogenic (Lane, 1992). Loss of p53 function has been associated with a broad variety of sporadic tumours, and abnormalities in the gene are thought to be present in almost half of all human cancers (Hollstein *et al.*, 1999; Hollstein *et al.*, 1991; Nigro *et al.*, 1989). Individuals with Li-Fraumeni syndrome, caused by a germline mutation in the p53 gene, have a high frequency of different cancers including breast cancers, brain tumours, lung adenocarcinomas and osteosarcomas from an early age (Freboung *et al.*, 1995). In vivo studies have demonstrated the tumour suppressor function of the p53 gene. Transgenic murine models which are null for p53, or have both copies of p53 mutated, develop normally but display an increased frequency of tumours, especially on the skin when carcinogens are applied to induce tumours (Kemp *et al.*, 1994). The exact role of p53 in normal and abnormal cells is still being investigated but the main functions of the protein are well documented. In simplified terms p53 is expressed in cells at increased levels as a response to damaging or stressful events which could lead to abnormal proliferation. p53 avoids this by inducing arrest at cell cycle checkpoints (Kastan *et al.*, 1991; Kastan *et al.*, 1992; Lane, 1992), mediating differentiation of the cell (Rotter *et al.*, 1994) or triggering cell death/apoptosis by interaction with downstream genes (el-Deiry, 1998; el-Deiry *et al.*, 1994). p53

activity appears to be regulated in a cell cycle manner through both the level of expression (with levels being lowest post mitosis to those increasing in G1) and also the degree to which it is phosphorylated. One of the major proteins involved in the regulation of p53 is MDM2 (reviewed in detail in Woods & Vousden, 2000). MDM2 interacts with the N-terminal region of p53 and inhibits p53 transcriptional activity (Momand *et al.*, 1992). As MDM2 is also transcriptionally activated by p53, p53 can control the expression of its own negative regulator to establish a feedback loop, serving to suppress p53 activity when needed and regulating the expression of both proteins (Wu *et al.*, 1993). MDM2 is also thought to regulate p53 nuclear export to the cytoplasm where it undergoes degradation, hence controlling protein stability in unstressed cells (Haupt *et al.*, 1997; Kubbutat *et al.*, 1997; Kubbutat *et al.*, 1999; Roth *et al.*, 1998). Human papilloma virus 16 E6 (HPV16 E6) oncogenic protein is known to complex with p53 and enhance p53 degradation both *in vivo* and *in vitro*, resulting in decreased stability of the protein and disrupting the normal cellular response to DNA damage mediated by p53 (Hubbert *et al.*, 1992; Kessis *et al.*, 1993; Lechner *et al.*, 1992). p53 has been shown to be inactivated in a variety of ways including rearrangement and deletion of the gene, but by far the most common alteration observed is a missense mutation involving a single amino acid substitution (Bennett *et al.*, 1999b; Harris, 1993; Hollstein *et al.*, 1991; Hussain & Harris, 1998; Hussain & Harris, 1999). The importance of p53 in cell regulation and repair and its involvement in so many tumour types has led to the gene being referred to as the 'guardian of the genome' (Lane, 1992) and a cellular 'gatekeeper' for growth and division (Levine, 1997).

1.3.4 Oncogenes

Oncogenes originate from the discovery that certain elements from both DNA and RNA viruses had the ability to transform cells from being healthy to cancerous. In 1911 Rous demonstrated that the cell-free filtrates of chicken sarcomas gave rise to sarcomas when inoculated into normal birds. Further evidence was supplied in the late 1970s when Shih *et al* transfected DNA from cell lines which had been transformed with chemical carcinogens into normal cells and produced transformed phenotypes (Shih *et al.*, 1979). Following this work the rapid identification of oncogenes by the transfection of cells in vitro with DNA taken from human tumours occurred. Oncogenes classically act in a dominant manner in that an abnormality of only one allele is sufficient to cause transformation. These abnormalities can occur through a variety of mechanisms including amplification, point mutation, chromosomal translocation or transposition to an active chromosomal domain leading to increased expression. Oncogene gene products are often involved in the normal proliferation, signalling or differentiation pathways of the cell, and become oncogenic because either their protein product is abnormal or there is a quantitative defect in the transcription of the gene (reviewed in detail in Cooper, 1990; Todd & Wong, 1999). The location and function of several human oncogenes are shown in table 1.2.

1.3.5 DNA repair and mutator genes

DNA is highly susceptible to damage by environmental insults with carcinogens such as UV radiation and tobacco smoke. It is logical therefore, that the cell must have a DNA repair system to prevent or limit this damage before cellular dysfunction occurs or the damage is inherited by subsequent progeny cells. This

system is called nucleotide excision repair (NER). In prokaryotes NER is controlled by 3 specific proteins (uvrA, -B and -C) which bind to the DNA strand at the site of the lesion and cleave out the damaged part, before resynthesizing the new strand and sealing it. However in eukaryotes the process is more complex, involving at least 12 proteins which include the xeroderma pigmentosum (XP) and excision repair cross complementation (ERCC) genes. Mutations in these genes are responsible for the xeroderma pigmentosum disorder (OMIM 194400) in which patients are susceptible to skin cancer due to being unable to repair UV-damaged DNA (reviewed in Chung & Rustgi, 1995). Murine studies have demonstrated that strains defective in the XP gene (involved in NER) develop skin, lung and liver tumours when exposed to carcinogens (Friedberg *et al.*, 2000). Another method in which DNA can be altered is through mismatch of individual nucleotides during replication of the genome. These abnormalities are usually removed by DNA mismatch repair enzymes. Again this repair system has been well documented in prokaryotes which has aided the discovery of homologous mismatch repair enzymes in eukaryotes and humans. This is an important part of DNA repair in humans due to a large part of the genome being composed of repetitive DNA sequences (microsatellites) which are particularly susceptible to genetic infidelity. Cell extracts in humans have been demonstrated to contain accurate strand specific DNA mismatch repair (MMR) function (Holmes *et al.*, 1990; Thomas *et al.*, 1991a; Thomas *et al.*, 1991b) which allows the repair of mismatched nucleotides, in addition to 1-2bp loops which are thought to develop when multiple copies of repeat sequence are inserted or deleted. The importance of mismatch repair enzymes has been clearly shown in colorectal cancer studies and more specifically in hereditary nonpolyposis colorectal cancer (HNPCC) syndrome (OMIM *120436).

HNPCC syndrome is inherited in an autosomal manner and is thought to be responsible for between 3-6% of all colorectal tumours (Lynch *et al.*, 1993). These HNPCC kindreds inherit one mutated MMR allele and the subsequent tumours which arise generally have a 'second hit' in accordance with Knudson's hypothesis. Around 57% of the adenomas in HNPCC kindreds show microsatellite instability consistent with a MMR deficient state in that microsatellite DNA from the tumour varies in length from the normal unaffected DNA in the same patient due to the insertion or deletion of repeat units (Aaltonen & Peltomaki, 1994; Aaltonen *et al.*, 1994a; Jacoby *et al.*, 1995). Similar alterations have also been found in sporadic colorectal carcinomas (Leach *et al.*, 1993; Liu *et al.*, 1995; Papadopoulos *et al.*, 1994), adenomas (Shibata *et al.*, 1994) and colonic aberrant crypt foci (Augenlicht *et al.*, 1996).

Table 1.2 The location and function of several human oncogenes implicated in lung cancer.

Protooncogene or Oncogene	Location in cell	Function	Possible mode of action
<i>ras</i> family	p21 membrane associated protein	Signal transduction	Overexpression of normal protein, point mutation
<i>raf</i>	Plasma Membrane	Signal transduction kinase which interacts with <i>ras</i>	Chromosomal deletion, activation causes growth arrest
<i>jun</i>	Nucleus	Transcription factor which binds to AP-1 motif	Altered expression, mediator of growth factor signals and cell regulation
<i>myb</i>	Nucleus	Involved in cell proliferation and differentiation	Deletion, amplification, differential expression in tumours
<i>fms</i> (<i>CSF1R</i>)	Plasma membrane	Transmembrane protein	Chromosomal rearrangement
<i>fur</i>	Plasma membrane	Recognition function	Overexpression of protein
<i>c-erbB</i>	Plasma membrane	Glycoprotein and epidermal growth factor receptor	Altered transcription and overexpression
<i>myc</i> family	Nucleus	DNA binding and transcriptional regulation	Gene amplification

1.4 Detecting genetic abnormalities in human cancers.

1.4.1 Cytogenetic analysis

Cytogenetic research into the basis of human disease utilises various methods of investigating the structure and number of all or individual chromosomes within cells. Early experiments relied on karyotyping or the investigation of chromosome number. Early karyotyping techniques were rather basic and time consuming, only differentiating between abnormalities such as increases and decreases in individual chromosome number and gross chromosomal abnormalities. Recent advances in chromosome banding techniques have allowed accurate differentiation between chromosomes and more importantly, increased resolution within individual chromosomes and their structure, thus permitting more accurate definition of translocation breakpoints and subchromosomal deletions (Craig & Bickmore, 1993). The isolation of chromosome-specific DNA markers through physical mapping methods now allows techniques such as flow-karyotyping to analyse chromosome numbers within samples, collecting rapid information on specific sets of chromosomes or single chromosomes (Green, 1990). Low resolution physical mapping using somatic cell hybrids (experimentally produced fusions of cultured cells, usually human and rodent) produce DNA markers mapped to particular chromosomes (reviewed in Bentley *et al.*, 1988).

The development of fluorescent in-situ hybridisation (FISH) has allowed the rapid and accurate mapping of specific multiple DNA clones on individual chromosomes, a technique that can be applied to large genomic DNA deletions and translocations within metaphase and interphase cells (Ferguson-Smith, 1997). One application of FISH is comparative genomic hybridisation or CGH (Kallioniemi *et al.*, 1992) which

can highlight chromosomal regions of loss or amplification within a single experiment. CGH uses a mixture of DNA from normal and tumour cells labelled with different fluorochromes in a competitive FISH. Specific regions can be isolated using image analysis equipment which highlights deviation from the ratio of FISH signal expected signifying allele loss or amplification (Levin *et al.*, 1995; van Dekken *et al.*, 1999).

1.4.2 Linkage analysis

Pedigree-based linkage analysis and multilocus linkage analysis provides a useful method of mapping an unidentified disease locus by genetic linkage in hereditary or familial disease (reviewed in Morton, 1984). This approach utilises polymorphic DNA markers spread throughout the whole human genome to map the most probable location of a disease gene. The marker is tested for evidence of co-segregation with the disease locus at meiosis. LOD scores (logarithm of the odds) from several affected pedigrees are then used to prove positive linkage occurring at the gene locus (Morton, 1955). These methods have been reliably used to identify genetic loci involved in various human diseases including cancer related genes such as the breast cancer genes BRCA1 (Smith *et al.*, 1994) and BRCA2 (Thorlacius *et al.*, 1995; Wooster *et al.*, 1994).

1.4.3 Loss of heterozygosity / allelic imbalance analysis

The research carried out by Cavenee *et al* (1983) involving loss of heterozygosity in sporadic retinoblastoma provided a means of identifying chromosomal regions present in the DNA of normal tissue but deleted within the DNA from tumour cells. Loss of heterozygosity (LOH), also known as allelic imbalance (AI), is indicative of

genomic imbalance within the human genome. LOH or AI can be caused by multiple phenomena such as chromosomal loss or duplication, DNA amplification and aneuploidy, which commonly occur in human cancer (Mertens *et al.*, 1997; Mitelman *et al.*, 1997). Previous studies have used loss of heterozygosity (LOH) and deletion mapping as a tool to locate small chromosomal regions in which putative tumour suppressor genes exist. Initially a genome wide search (allelotype analysis) is carried out using microsatellite markers mapped to individual chromosome arms to identify specific chromosomes which are targeted for deletion in tumour cells (Dolan *et al.*, 1998; Mitra *et al.*, 1994; Nawroz *et al.*, 1994; Neville *et al.*, 1996; Quinn *et al.*, 1994; Shibagaki *et al.*, 1994; Tsuchiya *et al.*, 1992). Deletion mapping or fine deletion mapping is then used to establish the approximate size and position of the deletion on chromosomal arms with high LOH frequencies using high density microsatellite markers located on a single chromosomal arm (Huang *et al.*, 1994; Konishi *et al.*, 1998; Okami *et al.*, 1997; Sikkink *et al.*, 1997; Wu *et al.*, 1994). If the region is sufficiently small, physical and genetic mapping strategies can be applied to isolate and identify target genes which may exist within the area.

1.4.4 Physical mapping strategies

Multiple high and low density mapping strategies exist for the selective study of DNA sequences within the human genome (table 1.3). The acceleration of the Human Genome Project (HGP), which aims to produce high density genetic maps and eventually the complete sequence of the human genome, in addition to the genomes of associated model organisms such as *Drosophila*, the mouse and *Fugu rubripes* for comparison, has allowed the construction and analysis of high density maps for each human chromosome (Bouck *et al.*, 2000; Crnogorac-Jurcevic *et al.*,

1997; Gilley & Fried, 1999; Lander *et al.*, 2001; Roest Crolius *et al.*, 2000; Rubin, 2001). One of the most common techniques used for physical mapping is cell-based DNA cloning methods involving cloning vectors (foreign DNA fragments capable of independent replication). As the sequence of human DNA is very large and individual genes can also cover large genetic distances, cloning vectors are frequently used for organising and characterising large human DNA sequences (table 1.4). Recent physical mapping strategies have relied heavily on yeast artificial chromosome (YAC), bacterial artificial chromosome (BAC) and P1 artificial chromosome (PAC) overlapping maps to order short 'landmark' DNA sequences, such as expressed sequence tags (EST) and sequence tagged sites (STS) and gene sequences, which can be conveniently localised to individual chromosomes using PCR assays (Lander *et al.*, 2001). Once clones are ordered and localised to subchromosomal regions using STS/EST content mapping, sequencing and exon trapping of genomic inserts can identify potential genes within the mapped region. The Fragile Histidine Triad (FHIT) gene is a recent example of utilising LOH and physical mapping to locate candidate disease genes for analysis in regions of minimal deletion (Ohta *et al.*, 1996). This will be discussed in the next section.

Table 1.3 Examples of physical mapping strategies commonly used for analysing the human genome (adapted from Strachan and Read, 1997).

Mapping Strategy	Examples/Requirements	Resolution
Cytogenetic map	Chromosome banding map	Several Mb
Breakpoint map	Somatic cell hybrid panels derived from translocations	Several Mb
	Monochromosomal radiation hybrid maps	Several Mb
	Whole genome radiation hybrid maps	0.5 Mb
Restriction map	Rare cutter restriction digest mapping	Several hundred kb
Clone contigs	Overlapping YAC clones	Several hundred kb
	Overlapping BAC clones	~100-300 kb
	Overlapping cosmid clones	40 kb
STS map	Requires sequence information from ordered clones	100 kb spacing
EST map	Requires cDNA sequencing and verification from other physical maps	~40 kb
DNA sequence map	Complete nucleotide sequencing of DNA	1bp

Table 1.4. Cloning vectors used in physical mapping and the size of inserted DNA (adapted from Strachan and Read, 1997).

Cloning vector	Average size of insert
Standard plasmid	0-10 kb
Bacteriophage λ insertion	0-10 kb
Bacteriophage λ replacement	9-23 kb
Cosmid	30-44 kb
Bacteriophage P1	70-100 kb
PAC	130-150 kb
BAC	<300 kb
YAC	0.2-2.0 Mb

1.4.5 Isolation, identification and analysis of the FHIT gene in the 3p14.2 region

Allelotype studies have demonstrated that chromosome 3p is a frequent target for LOH/AI in multiple tumours including NSCLC, SCLC, skin, oesophageal, HNSCC, salivary gland, cervical and gastric tumours (Ah-See *et al.*, 1994; Dolan *et al.*, 1998; Johns *et al.*, 1996; Mitra *et al.*, 1994; Nawroz *et al.*, 1994; Neville *et al.*, 1996; Quinn *et al.*, 1994; Shibagaki *et al.*, 1994; Tsuchiya *et al.*, 1992; Virmani *et al.*, 1998; Yustein *et al.*, 1999). Further deletion mapping and fine mapping studies detected that a region at 3p14.2 was a specific target for deletion in several tumour types and that putative tumour suppressor genes could be located in this region (Fullwood *et al.*, 1999; Larson *et al.*, 1997; Shridhar *et al.*, 1997; Sikkink *et al.*, 1997; Sung *et al.*, 2000; Wistuba *et al.*, 2000b; Wistuba *et al.*, 1997b).

In 1996, Ohta *et al.* used a combination of homozygous deletion mapping and exon trapping combined with physical mapping knowledge to identify the FHIT gene, which in turn encompasses the FRA3B fragile site (Glover *et al.*, 1988) and also the renal cell carcinoma-associated t(3;8) breakpoint (Cohen *et al.*, 1979) within the 3p14.2 region. This study assembled a cosmid map using polymerase chain reaction of STS's mapped to the region and new STS markers derived from cosmid ends. Six cosmids were assembled into a map within the 3p14.2 region which spanned the homozygous deleted region detected in various carcinoma cell lines. Exon trapping experiments were then used to isolate expressed partial FHIT cDNA sequences within the region. Complete cDNA sequence was obtained from sequencing of RACE (rapid amplification of cDNA ends) products. Sequence analysis of the FHIT cDNA then established the open reading frame (ORF) and poly(A) tail.

Distributed over 500kb and composed of ten exons, the FHIT protein was a member of a highly conserved human histidine triad gene family and showed similarity to the *Schizosaccaromyces pombe* Ap4A asymmetrical hydrolase, which is proposed to have a role in DNA replication (Baxi *et al.*, 1994), dinucleotide oligophosphate metabolism (Barnes *et al.*, 1996) and in the cell's response to metabolic stress (Baker & Jacobson, 1986; Farr *et al.*, 1989). RNA transcripts of FHIT were demonstrated to have abnormalities in 40% of NSCLC and 80% of SCLC, with 76% of tumours showing loss of FHIT alleles (Sozzi *et al.*, 1996b). Studies have also detected aberrant FHIT cDNA transcripts in tumours which displayed LOH within this region including digestive tract cancers (Ohta *et al.*, 1996), small cell and non-small cell lung tumours (Sozzi *et al.*, 1996b), Merkel cell carcinomas (Sozzi *et al.*, 1996a), primary breast tumours and breast cancer cell lines (Negrini *et al.*, 1996), HNSCC (Mao *et al.*, 1996; Virgilio *et al.*, 1996) and benign proliferative breast disease associated with familial breast cancer (Panagopoulos *et al.*, 1996).

Thiagalingam *et al* (1996) evaluated the FHIT gene in a panel of colorectal cancer cell lines and xenografts and detected a normal cDNA transcript in 29/31 cancers. This study suggested that the aberrant cDNAs detected by using a nested PCR approach could be caused by alternatively spliced FHIT transcripts being over represented in the nested PCR. Similar results were also detected in a preliminary evaluation of the FHIT gene in non-melanoma skin cancer (Sikkink *et al.*, 1997). Sozzi *et al* (1998) investigated FHIT expression in NSCLC and preneoplastic lesions using immunocytochemistry and detected loss of FHIT protein in 73% of NSCLC and 95% of precancerous lesions. Loss of expression was also

significantly higher in smokers (75%) than non-smokers (39%). Similarly, Tokuchi *et al* (1999) investigated the frequency of aberrant FHIT transcripts in both lung cancer and normal lung tissue and again detected abnormal FHIT transcripts in both neoplastic and non-neoplastic tissues.

Aberrant transcripts or deletions of the FHIT gene detected in the above tumours may be due to the fragile nature of this region of chromosome 3p. A study positioned the amphidicolin-induced gaps at the FRA3B fragile site in FHIT intron five, less than 30 kb telomeric to the FHIT exon 5, a human papillomavirus 16 integration site in FHIT intron 4 and the t(3;8) break point within FHIT intron 3 (Zimonjic *et al.*, 1997). This suggests that the cancer-specific deletions involving exons 3, 4 and 5 which have been detected in several FHIT studies may have originated through breaks in the fragile sites. More detailed analysis of the 300kb sequence which encompasses the FHIT gene exon 5 and the FRA3B fragile site has provided insight into the mechanism of repair of the fragile site double strand breaks involving recombination between LINE 1 elements and deletion of the intervening sequence, a process which often deletes FHIT exons (Huebner *et al.*, 1998; Huebner *et al.*, 1997). There is also evidence that the fragility of the FRA3B fragile site is closely connected with allele-specific late replication which may contribute to breaks or gaps preferentially occurring on chromosome 3p within this region (Wang *et al.*, 1999).

1.5 Lung Carcinogenesis

1.5.1 Allelic imbalance in human lung cancer

The pattern of loss of heterozygosity in NSCLC and SCLC has revealed specific chromosome deletions which may be important in the development of these tumour types. In NSCLC several studies detected a high frequency of deletion on chromosome arms 3p, 9p and 17p with significant allelic imbalances on other chromosome arms including 4q, 5q, 8p, 13q and 17q (Girard *et al.*, 2000; Liloglou *et al.*, 2000; Neville *et al.*, 1996; Sato *et al.*, 1994; Tsuchiya *et al.*, 1992). In SCLC allelic loss occurred at a high frequency on chromosome arms 3p, 5q, 13q and 17p (Girard *et al.*, 2000; Miura *et al.*, 1992; Yokota *et al.*, 1987). Differences in the pattern of chromosome deletions have also been detected between squamous cell carcinomas and adenocarcinoma of the lung with loss of 3p, 9p and 17p being significantly less in adenocarcinoma than squamous cell carcinomas indicating different mechanisms of tumourigenesis between these two tumour types (Neville *et al.*, 1996; Sato *et al.*, 1994; Tsuchiya *et al.*, 1992).

Comparison between Fractional Allele Loss (FAL; the number of chromosome arms showing LOH/number of informative chromosome arms) scores based symmetrically around a median FAL value (i.e. low, medium and high FAL) and LOH on chromosomes 3p, 9p and 17p detected a correlation between a high frequency of LOH on these chromosomes and high FAL values. However, a distinct genetic population of NSCLC tumours which showed a low frequency of LOH on these chromosome arms and a low FAL value, showed LOH at other chromosome sites (5p, 5q, 13q, 16q and 19q) indicating a subset of tumours with different genetic events (Field *et al.*, 1996). Virmani *et al* (1998) demonstrated

non-random chromosomal losses in lung cancer cell lines involving multiple regions. This study showed that some losses were common to both NSCLC and SCLC subtypes, others subtype specific, in addition to genetic deletions at novel chromosomal lesions. Similar patterns of chromosome loss in lung tumours have also been demonstrated using Comparative Genomic Hybridisation (Petersen *et al.*, 1997).

Allelotype and mutational studies have also characterised specific molecular genetic changes which occur at certain stages of lung tumour development. Shiseki *et al* (1994 and 1996) carried out a comparative allelotype of both early and advanced stage NSCLC, including brain metastases and found that loss of heterozygosity of chromosomes 3p, 13q and 17p occurred at a high frequency in both the primary tumours and brain metastases (>60%). However a high incidence of allelic loss was detected on chromosomes 2q, 18q and 22q (>60%) more frequently in the brain metastases than the primary tumour, indicating that LOH on these chromosomes is a late event in the progression of lung carcinomas, possibly giving the tumour a more aggressive phenotype.

1.5.2 Allelic imbalance in preneoplastic lesions

In order to understand the molecular order of the genetic events involved in lung carcinogenesis several studies have investigated LOH in preneoplastic lesions (Sozzi *et al.*, 1992; Sozzi *et al.*, 1995; Hung *et al.*, 1995; Kishimoto *et al.*, 1995; Sundaresan *et al.*, 1995; Wistuba *et al.*, 1997; Kitaguchi *et al.*, 1998; Wistuba *et al.*, 1999; Kohno *et al.*, 1999; Wistuba *et al.*, 1999; Wistuba *et al.*, 2000). Sundaresan *et al* (1992) demonstrated allelic loss at 3p and mutation and allelic loss of the p53

gene in severe bronchial dysplasia (carcinoma in situ) from patients with and without lung cancer, indicating that genetic damage of these chromosomes precedes NSCLC tumour invasion. Further work by this group established that similar genetic changes also occurred in preneoplastic dysplasias and that changes in dysplastic lesions were often more discreet than those detected in the invasive tumours. Chromosome 3p allelic deletion also preceded damage to the p53 gene, indicating that molecular damage in lung cancer may occur in sequential steps (Chung *et al.*, 1995). Kishimoto *et al* (1995) detected allelic loss in both invasive tumours and preneoplastic lesion at the loci IFN α and D9S171 on chromosome 9p. Loss of 9p was detected in 38% of hyperplasia, 80% of dysplasia and 100% of carcinoma in situ lesions.

Witsuba *et al* (1997a) detected multiple genetic changes in the bronchial epithelium of current and former smokers without lung carcinomas. Multiple biopsy specimens were obtained from current smokers, former smokers and non-smokers and analysed for genetic instability using 15 microsatellite loci. Approximately half of the histologically normal specimens from smokers showed LOH and microsatellite alterations were detected in 64% of individuals. No genetic alterations were detected in non-smokers. Common microsatellite changes included LOH on 3p and 9p and less frequently 5q, 17p and 13q. Recent work by Sanchez-Cespedes *et al* (2001) also demonstrated significantly more allelic imbalance at chromosome arms 3p, 6q, 9p, 16p and 17p in the lung tumours from smokers when compared to non-smokers. Further evidence for allelic loss on 3p and 9p and p53 abnormalities was provided by Kohno *et al* (1999). This study detected genetic alterations in 8 metaplastic lesions and 3 alveolar hyperplastic

lesions (commonly thought to be preneoplastic lesions specific to adenocarcinoma of the lung).

Witsuba *et al* (1999a) investigated allele loss at 12 chromosomal regions frequently deleted in lung cancer in histologically normal bronchial epithelium, preneoplastic lesions, CIS lesions and invasive tumours. Genetic abnormalities were detected in 31% of histologically normal cells and 42% of mildly abnormal lesions (hyperplasia/metaplasia). Frequent deletions were detected at 3p21, 3p22-p24, 3p25 and 9p21. This group also investigated allelic losses on the short arm of chromosome 8, which is frequently deleted in both SCLC and NSCLC. Allelic deletions at 8p21-p23 were frequent in both tumour and in preneoplastic and hyperplastic cells. No deletions were detected in histologically normal bronchial epithelium, however loss of loci on 8p was found in 15% of mildly abnormal epithelium, 50% of dysplastic epithelium, 92% of CIS lesions and 91% of invasive tumours. The frequency and size of the deleted 8p21-p23 region increased with the severity of the lesion. When compared to allelic loss on 3p and 9p, deletions at 8p21-p23 always followed 3p deletions and usually followed 9p deletions (Wistuba *et al.*, 1999b).

A similar study by Park *et al* (1999) detected abnormalities in 32% of histologically normal or slightly abnormal epithelium and 52% of dysplastic epithelium. One of the most common genetic abnormalities throughout lung carcinogenesis appears to be allelic loss on chromosome 3p. Witsuba *et al* (2000b) performed a high resolution analysis of 3p LOH at 28 loci in 97 lung cancer samples and 54 preneoplastic/preinvasive bronchial lesions. A 600 kb region at 3p21.3 displayed LOH at 6 loci in 77% of lung tumours, 70% of normal or preneoplastic/preinvasive

lesions associated with lung cancer, 49% of normal or preneoplastic/preinvasive lesions associated smokers without lung cancer. No LOH was detected in 18 samples of epithelial tissue from patients who never smoked. This strongly implicates the deletion of the 600 kb region at 3p21.3 as one of the earliest events in lung tumourigenesis. Although there is little direct molecular evidence that preneoplastic lesions such as squamous metaplasia and dysplasia/atypia proceed to develop into lung tumours, molecular genetic studies have revealed that these lesions show distinct patterns of allelic imbalance with common genetic losses in both tumour and preneoplastic lesions (table 1.5).

Table 1.5. Summary of the frequencies of allelic imbalance (%AI) detected in various stages of NSCLC carcinogenesis (adapted from Wistuba *et al* (1999a) [Ref 1], Wistuba *et al* (1999b) [Ref 2] and Khono *et al* (1999) [Ref 3]).

Histology	% AI	Regions with high frequencies of AI	Reference
Normal tissue	4/13 (31%)	3p21-p25, 9p21	[1]
Squamous Metaplasia	13/33 (42%)	3p, 5q, 9p21, 13q14, 17p13	[1]
	5/34 (15%)	8p21-p23	[2]
	1/13 (8%, SCC only)	3p	[3]
Dysplasia/Atypia	17/21(81%)	3p, 5q, 9p21, 13q14, 17p13	[1]
	9/18 (50%)	8p21-p23	[2]
	7/22 (32%, SCC only)	3p, 9p, 17p (p53 mutation in 9%)	[3]
	3/20 (15%, Adeno only)	3p, 9p	[3]
CIS	15/15 (100%)	3p, 5q, 9p21, 13q14, 17p13	[1]
	11/12 (92%)	8p21-p23	[2]
Invasive tumour	12/12 (100%)	3p, 5q, 9p21, 13q14, 17p13	[1]
	42/46 (91%)	8p21-p23	[2]
	17/35 (48%, SCC only)	3p, 9p, 17p (p53 mutation in 26%)	[3]
	13/26 (50%, Adeno only)	3p, 9p, 17p (p53 mutation in 23%)	[3]

1.5.3 Microsatellite instability in lung tumours

Microsatellite alterations (often called microsatellite instability (MI)) are thought to reflect instability of the genome due to damage to the DNA mismatch repair genes which leads to the deletion or insertion of nucleotides during DNA repair in affected cells (Aaltonen *et al.*, 1994a; Aaltonen *et al.*, 1994b; Jacoby *et al.*, 1995). Although mutations in the DNA mismatch repair genes are infrequent in lung cancers (Benachenhou *et al.*, 1998; Xinarianos *et al.*, 2000), several studies have used microsatellite analysis to investigate the frequency of MI in lung tumours (table 1.6). Shridhar *et al* (1994) investigated the genetic instability of microsatellite sequences in 38 NSCLC and found 25/38 with no MI. A total of 13/38 (34%) NSCLC displayed MI, 5/38 displayed MI with one marker and 8/38 showed MI with more than one microsatellite marker. Adenocarcinomas were the most prominent tumour type which displayed MI in this group (7/21). In addition, microsatellite markers for chromosome 3 displayed the highest frequency of MI when compared with markers on other chromosomes, although this may be due to the greater number of microsatellite loci on chromosome 3 that were tested in this study.

In contrast with these results, Fong *et al* (1995b) showed that MI was infrequent in NSCLC (7 of 108 cases; 6.5%) and only occurred at specific loci (3/6 tested). This suggests either possible population differences between NSCLC tumours or a preference of MI to occur at specific loci (for example microsatellites located on chromosome 3p) and have different rates of instability. This study also detected frequent LOH on chromosomes 5q, 8p, 9p, 11p and 17p in addition to concurrent molecular changes such as K-ras and p53 mutation associated with MI.

Microsatellite instability was also found to be a frequent event in SCLC, occurring

in 45% and 50% of SCLC tumours respectively (Mao *et al.*, 1994; Merlo *et al.*, 1994b). Merlo *et al.* (1994b) detected MI in 8/15 of tumours and in more than 10% of all the alleles tested. A high frequency of allele loss was detected with or without MI, although high LOH was associated with markers flanking the unstable regions. Mao *et al.* (1994) detected microsatellite alterations in the sputum of several patients with SCLC who also showed MI in primary tumours. Field *et al.* (1999) detected microsatellite alterations in the bronchial lavage of 9/47 (19%) individuals with no clinical evidence of lung cancer. A study by Chen *et al.* (1996) also detected microsatellite alterations in 16/21 (76%) of SCLC tumours in addition to 15/21 (71%) plasma DNA samples from the same SCLC patients. In one case the microsatellite alteration was present only in the plasma DNA and not in the tumour. Miozzo *et al.* (1996) investigated samples of tumour, histologically normal bronchial mucosa and sputum and detected MI in 49% of samples. MI was also detected in the distant bronchus in 36% of samples tested and 29% of samples showed MI in the distant bronchus but not in the tumour. As microsatellite aberrations were detected in the histologically 'normal' bronchial mucosa and samples with low grade atypia, this suggests that MI may occur early in lung carcinogenesis.

The importance and incidence of microsatellite instability in both lung tumours and other neoplasms is subject to much controversy due to differences in the reported rates from various studies. This is considered to be related to problems in the methods used to analyse microsatellite instability, related to *Taq* polymerase slippage during PCR, enrichment of cell populations to provide monoclonal cell populations for analysis and control of the terminal deoxynucleotidyl transferase activity in *Taq* polymerase to control 1bp overhangs generated during the

amplification of microsatellite sequences. (Ginot *et al.*, 1996; Maehara *et al.*, 2001; Oda *et al.*, 1997). A recent study using a fluorescent detection system to analyse microsatellite alterations in lung tumours suggested that the rate of microsatellite instability in lung tumours (~1%) may be lower than previously reported, however this maybe partly attributed to the fluorescent detection technology (Liloglou *et al.*, 2000).

Table 1.6. Summary of microsatellite instability detected in lung cancer

<i>Tissue Type</i>	No. of Patients	Frequency of MI detected	Reference
NSCLC	38	13/38 (34%)	Shridhar <i>et al</i> , 1994
SCLC	33	15/33 (45%)	Merlo <i>et al</i> , 1994
NSCLC	108	7/108 (6%)	Fong <i>et al</i> , 1995
NSCLC/SCLC	51	17/51 (33%)	Miozzo <i>et al</i> , 1996
Normal Bronchial Mucosa		15/42 (35%)	
Sputum		3/5 (60%)	
Overall		25/51 (49%)	
SCLC	21	16/21 (76%)	Chen <i>et al</i> , 1996
Plasma		15/21 (71%)	
*Bronchial Lavage	26	16/26 (61%)	Field <i>et al</i> , 1999
Cytology positive		1/8 (12%)	
Cytology negative		14/19 (73%)	
NSCLC	85	1/85 (1%)	Liloglou <i>et al</i> , 2000

* Microsatellite alterations were detected in bronchial lavage from individuals with lung cancer diagnosed through cytological review (cytology positive) compared with individuals with no cytological evidence of disease (cytology negative). Seven individuals with negative cytology (five demonstrated microsatellite alterations in bronchial lavage) were subsequently diagnosed with lung cancer.

1.5.4 Chromosome 3p deletions in lung carcinogenesis

Perhaps the most common genetic change in lung cancer (both NSCLC and SCLC) is deletion of regions of chromosome arm 3p (Mitsudomi *et al.*, 1996; Mori *et al.*, 1989; Neville *et al.*, 1996; Sundaresan *et al.*, 1992; Weston *et al.*, 1989; Wistuba *et al.*, 2000a; Yamakawa *et al.*, 1993; Yokota *et al.*, 1987). Deletion mapping in a wide variety of tumour types has detected deletions on the short arm of chromosome 3p usually in the same distinct regions, located at 3p25-3pter, 3p21 and 3p14-3pcen (Chen *et al.*, 1994; Druck *et al.*, 1995; Ehlen & Dubeau, 1990; Hibi *et al.*, 1992; Hosoe *et al.*, 1994a; Maestro *et al.*, 1993; Sikkink *et al.*, 1997; Wistuba *et al.*, 2000a; Wu *et al.*, 1994). Not only is this change common but it has been shown to occur at the earliest stage (hyperplasia) of human lung carcinogenesis and involve all regions of the respiratory tract including histologically normal bronchial epithelium (Hung *et al.*, 1995; Kohno *et al.*, 1999; Park *et al.*, 1999; Park *et al.*, 2000; Sozzi *et al.*, 1995; Sundaresan *et al.*, 1992; Sundaresan *et al.*, 1995; Wistuba *et al.*, 1999a; Wistuba *et al.*, 2000a; Wistuba *et al.*, 1997a). Several genes have been identified in these regions, mutation of which may contribute to lung carcinogenesis.

The region 3p21, which is deleted at a high frequency in all major types of lung cancer, contains the DNA mismatch repair gene hMLH1 which has been shown to be mutated in sporadic colorectal cancers (Papadopoulos *et al.*, 1994), although as yet there is no evidence for mutations of hMLH1 in lung cancer (Xinarianos *et al.*, 2000). Cosmid cloning of 3p21 isolated 5 new genes, one of which was found mutated in several tumour types including renal cell carcinomas (Shridhar *et al.*, 1996b), lung, breast and prostate cancers and in HNSCC (Shridhar *et al.*, 1996a).

The main candidate gene in 3p25-3pter is the *Von Hippel-Lindau (VHL)* gene (Latif *et al.*, 1993) which displays tumour suppressor properties in certain cell lines (Iliopoulos *et al.*, 1995). Although mutations in the *VHL* gene have been found in familial renal cell carcinoma and hemangioblastomas (reviewed in Gnarr *et al.*, 1996), a study involving HNSCC detected no *VHL* gene mutations in tumours showing LOH at 3p25-p26 (Waber *et al.*, 1996). This suggests another gene of significance in this area for this particular tumour type.

The region of loss at 3p14-3pcen contains the recently cloned FHIT tumour suppressor gene which in turn encompasses the FRA3B fragile site and also the renal cell carcinoma-associated t(3;8) breakpoint (Ohta *et al.*, 1996). The involvement of the FHIT gene in lung cancer is discussed in section 1.4.5 of this chapter. A recent study by Witsuba *et al* (2000) performed a high resolution allelotyping of human lung cancer in addition to preneoplastic lesions. A high frequency of 3p loss was detected in SCLC (91%) and squamous cell carcinomas (95%) over a larger area of chromosome 3p whereas adenocarcinomas displayed a lower frequency of allele loss (71%) over a small area of chromosome 3p. This study also identified multiple regions of localised 3p loss and common independent sites of breakpoints included 3p21.3, 3p14.2 and 3p12.

1.5.5 Chromosome 9 deletions, p15^{INK4B} and p16^{INK4} abnormalities in NSCLC carcinogenesis.

As previously mentioned NSCLC have a high frequency of deletion on chromosome arm 9p. Further studies have attempted to narrow this region of loss using a panel of highly polymorphic markers for chromosome 9p. Neville *et al*

(1995) investigated LOH on 9p and detected a high frequency of LOH with 44% of tumours within the D9S156-D9S161 region. A minimal region of loss within this group of tumours was detected at 9p23. Mead *et al* (1997) also investigated LOH at 9p21, where the cyclin-dependant kinase inhibitors p15^{INK4B} and p16^{INK4} (which are highly likely to play roles as tumour suppressor genes) are located, in 32 NSCLC tumours and 3 cell lines using 15 microsatellite markers. A common region of LOH was detected around the marker D9S259 within a 1.7 Mb region for 52% of tumours. LOH at D9S292 (the closest marker to p16^{INK4}) was seen in only 17% of tumours. This suggests that there is a tumour suppressor gene located proximal to p16^{INK4} which may be involved in the pathogenesis of NSCLC.

Immunohistochemical staining for p16^{INK4} expression in 114 resected NSCLC revealed reduced expression in 30 (27%) tumours (Kinoshita *et al.*, 1996). Mutation analysis of p16^{INK4} in lung cancers has detected mutations in NSCLC but not in SCLC (Ohno, 1996). Mutation rates also differ considerably from various studies, ranging between 7% and 48% (de Vos *et al.*, 1995; Hayashi *et al.*, 1994; Marchetti *et al.*, 1997; Nakagawa *et al.*, 1995; Ohno, 1996; Rusin *et al.*, 1996; Shimizu & Sekiya, 1995).

Another common method of p16^{INK4} inactivation appears to be by homozygous deletion, especially in NSCLC cell lines derived from primary tumours (de Vos *et al.*, 1995; Nakagawa *et al.*, 1995; Ohno, 1996). Several studies have also linked inactivation of the p16^{INK4} with hypermethylation of G:C rich regions within the gene (Belinsky *et al.*, 1998; Esteller *et al.*, 1999; Kashiwabara *et al.*, 1998; Kersting *et al.*, 2000; Otterson *et al.*, 1995; Swafford *et al.*, 1997). Although mutation of p15^{INK4B} is infrequent in NSCLC (Rusin *et al.*, 1996; Shimizu & Sekiya, 1995), one

study has demonstrated co-deletion of p15^{INK4B} and p16^{INK4} to occur in over half of NSCLC (Xiao *et al.*, 1995). Schmid *et al* (1998) investigated homozygous deletions around the methylthioadenosine phosphorylase (MTAP) gene located at 9p21 proximal to the p16^{INK4} gene in 50 NSCLC tumours. Homozygous deletions were detected for MTAP exon 8 (located near p16) in 38% of tumours.

Adenocarcinomas displayed a higher frequency of deletion than squamous cell carcinomas (44% and 29% respectively). MTAP homozygous deletions were also more frequent than p16^{INK4} deletions (18%) in these tumours. Several tumours which had MTAP deletions also displayed loss of p16 protein expression. Frequent breakpoints were detected between exons 3 and 4 of the MTAP gene and homozygous deletion of MTAP were associated with deletions of p16^{INK4} in only half of cases. This indicates that either another target gene exists in this region which may play a role in lung tumourigenesis or that deletions of the 3'-untranslated region of p16^{INK4} can disrupt p16 expression or function.

Hamada *et al* (2000) fine mapped homozygous deletions in 149 lung cancer cell lines using 24 markers located in the region of 9p21. This study detected deletions in 26% of cell lines and found clusters of deletions at two independent regions. One region was located around the p16^{INK4} locus and deleted in 26% of cell lines and the other region located around the microsatellite loci D9S171 and deleted in 12% of cell lines. Several cell lines had identical deletions of a 17,036 bp sequence located 20 kb distal to D9S171. However, this deletion was also detected in a B-lymphoblastoid cell line and in 5 noncancerous tissue samples, suggesting a genetic polymorphism.

1.5.6 p53 mutation in lung carcinogenesis.

Allelotype studies have demonstrated that both NSCLC and SCLC display a high frequency of chromosome deletion on 17p (Miura *et al.*, 1992; Neville *et al.*, 1996; Sato *et al.*, 1994; Saxena *et al.*, 1992; Tsuchiya *et al.*, 1992; Tsuchiya *et al.*, 2000; Yokota *et al.*, 1987). The p53 tumour suppressor gene maps to the short arm of chromosome 17 (17p13.1) and encodes a 53 kDa phosphoprotein that is produced in very low levels in normal cells (Lane & Crawford, 1979; Lane & Crawford, 1980). The finding of a large proportion of p53 mutations in a wide range of tumours has led to many molecular epidemiological studies on the significance of these mutations (Bennett *et al.*, 1992; Hollstein *et al.*, 1991). This has led to the discovery that the mutation pattern of a tumour can differ depending on the site where the tumour is found and the specific carcinogenic risk can be linked to a specific 'signature' mutation (Bennett *et al.*, 1999b; Hussain & Harris, 1998; Hussain & Harris, 1999; Hussain & Harris, 2000). For example, a G→T mutation at p53 codon 249 is often common in aflatoxin-associated hepatocellular carcinoma (Hollstein *et al.*, 1993) and tandem base substitutions at dipyrimidine sites (C→T or CC→TT mutations) are associated with UV exposure in skin carcinogenesis (Brash *et al.*, 1991).

Mutations in the p53 gene have also been demonstrated to occur frequently in both NSCLC and SCLC with mutation rates of 45% and 80% (100% in SCLC cell lines) respectively (Chiba *et al.*, 1990; D'Amico *et al.*, 1992) with p53 mutations being more common in squamous cell carcinoma than any other tumour type in NSCLC. GC→TA transversions were found to occur more frequently in smokers than in

non-smokers which has led to this mutation being associated with specific carcinogens (such as free radicals and reactive oxygen species) which are present in tobacco smoke (Greenblatt *et al.*, 1994). In contrast with this, Liloglou *et al* (1997) detected a higher prevalence of GC→AT transitions than GC→TA transversions in a series of tumours from the Merseyside population suggesting a carcinogen other than tobacco smoke to be involved in the pathogenesis of these lung tumours.

1.6 Field cancerization and clonal evolution in NSCLC tumourigenesis.

The majority of current research is centred on documenting the genetic changes that occur in each stage of tumour development, especially the first event which initiates the abnormal growth of a cell. In 1953 Slaughter *et al* described a process called 'field cancerization' occurring in oral stratified squamous epithelium. Field cancerization is thought to occur by the prolonged exposure of a large area of tissue to carcinogens which render the cells 'cancer-prone', carrying genetic damage that imparts preferential growth and survival characteristics although they appear morphologically normal. Subsequent genetic insults to small groups of such preconditioned cells in this tissue could initiate the growth of abnormal clones which then develop into multiple tumours. This theory would seem to explain the independent development of multiple primary and secondary neoplasms which are presented in cancers such as HNSCC (Scholes *et al.*, 1998; van Oijen *et al.*, 2000) and lung carcinoma (Smith *et al.*, 1976; Nishiyama *et al.*, 1989; Carey, 1996; Hiroshima *et al.*, 1998). Work by Sozzi *et al* (1995) supported this theory of field cancerization by providing evidence that metachronously arising tumours had different genetic changes on chromosome 3p and within the p53 gene and hence

near complete independency with regard to the origin of the tumours. However research has shown that these multiple tumours can share common genetic changes such as identical small deletions of chromosome 3p and common p53 mutations in lesions arising at physically different locations which would seem to support the clonal development of tumours (Chung *et al.*, 1996; Hung *et al.*, 1995).

The use of improved technical methods such as laser capture microdissection for the enrichment and genetic analysis of small cell numbers in bronchial tissue has allowed a more comprehensive study of the specific genetic abnormalities that occur in the respiratory epithelium of lung cancer patients as well as smokers and non-smokers (Hittelman, 1999; Park *et al.*, 2000). Mao *et al* (1997) investigated the respiratory epithelium of both current and former smokers without lung tumours and identified LOH at 3p14 (75%), 9p21(57%) and 17p13 (18%). LOH at 3p14 was more frequent in current smokers than former smokers and a control group of non-smokers displayed LOH in only 1 patient at this chromosomal location. This indicates that in smokers the bronchial epithelium may be preconditioned for tumour formation through field cancerization by long term exposure to carcinogenesis in cigarette smoke. Kishimoto *et al* (1995a) investigated chromosomal loss on 9p in invasive and preinvasive lesion in the lungs of patients with NSCLC and detected identical patterns of loss affecting the same allele ('allele specific loss') indicative of either an early genetic abnormality or clonal changes. Similar findings were observed by Witsuba *et al* (1999a). This study analysed 94 individual microdissected foci of bronchial epithelial cells from 12 resected tumours at 10 chromosomal regions and detected clonally related changes in 28 pre- and non-invasive foci compared with 30 foci with independent clonal events. Allele

specific loss was detected in 90% of comparisons. Park *et al* (1999) examined the size and frequency of clonally related abnormalities in 218 multiple discontinuous foci of bronchial epithelium from 19 lung resections using twelve microsatellite markers. This study detected distinct molecular changes in adjacent patches of 200-1000 lung epithelial cells, suggesting separate evolving clones or subclones within the epithelium of patients with lung tumours. This study also suggested that most clonal patches may contain approximately 90,000 cells. The existence of multiple discontinuous foci of cells with differing patterns of molecular damage due to field cancerization may explain the high incidence of preneoplastic lesions in patients with lung tumours and also in the lungs of current and former smokers.

There has been much debate about the origin of lung tumours, whether they are clonal populations arising from a single cell or from the co-operation between multiple cells. Current opinion favours a series of mutations in a single cell which begins a process of preselection and clonal development leading to the development of the malignant tumour. However Shimizu *et al* (2000) have indicated that this phenomenon most likely does not account for multiple synchronous pulmonary tumours. This group investigated genetic abnormalities in patients with multiple synchronous lung tumours and detected clonal changes in 11/14 (79%) cases indicative of intrapulmonary metastasis. Only 1 case displayed different patterns of genetic damage indicative of multicentric lung cancer. In contrast, Hiroshima *et al* (1998) compared genetic abnormalities in 19 cases of synchronous multiple lung carcinomas and 11 cases of metachronous multiple lung carcinomas. This study detected different spectrums of p53 mutations and genetic loss in both groups suggesting that some multiple carcinomas have separate clonal

origins. One theory for the development of multiple tumours and preinvasive lesions is the expansion and migration of advanced clones of bronchial cells throughout the respiratory tree in a similar manner to the movement of mutated colonic crypts throughout the colorectal epithelium (Garcia *et al.*, 1999; Wright, 2000).

1.7 Comparison of NSCLC with other epithelial malignancies.

Research has shown that NSCLC displays common genetic abnormalities when compared with other epithelial malignancies such as HNSCC, oral SCC, non-melanoma skin cancer and Barrett's adenocarcinoma. Allelotyping studies have detected similar patterns of chromosome loss on 3p, 9p and 17q (Ah-See *et al.*, 1994; Dolan *et al.*, 1998; Johns *et al.*, 1996; Nawroz *et al.*, 1994; Neville *et al.*, 1996; Quinn *et al.*, 1994; Shibagaki *et al.*, 1994; Tsuchiya *et al.*, 1992; Field *et al.*, 1995). Deletion mapping studies have also identified loss of similar chromosomal regions (such as 3p25, 3p14.2, 3p21, 9p21-p23 and 17p13.3) on these three chromosomes which may reflect the inactivation of the same tumour suppressor gene (Dunn *et al.*, 1999; Dunn *et al.*, 2000a; Holland *et al.*, 1994; Hosoe *et al.*, 1994a; Huang *et al.*, 1994; Kisielewski *et al.*, 1998; Konishi *et al.*, 1998; Mao *et al.*, 1996; Mori *et al.*, 1994; Nakagawa *et al.*, 1999; Okami *et al.*, 1997; Pateromichelakis *et al.*, 2000; Sikkink *et al.*, 1997; Sozzi *et al.*, 1998; Sung *et al.*, 2000; Tanimoto *et al.*, 2000; Virgilio *et al.*, 1996; Wu *et al.*, 1994; Zou *et al.*, 1997). Barrett's oesophagus is a condition which causes the lining of the lower oesophagus to revert to columnar epithelium due to the continued reflux of stomach acid (Johnson *et al.*, 1978). This condition is classed as premalignant due to a 100 times higher risk of malignancy of patients with Barrett's oesophagus than

the general population (Duhaylongsod & Wolfe, 1991; Haggitt *et al.*, 1978; Hamilton & Smith, 1987; Sanfey *et al.*, 1985; Sarr *et al.*, 1985; Williamson *et al.*, 1991). As in lung cancer, patients with Barrett's oesophagus and also oesophageal adenocarcinoma display a range of metaplastic and atypical subtypes with intervening normal areas (Jankowski *et al.*, 1999; Mueller *et al.*, 2000; Schmidt *et al.*, 1985). Studies have also showed an increased risk of progression to adenocarcinoma due to alcohol and tobacco consumption (Brown *et al.*, 1994; Gray *et al.*, 1993; Levi *et al.*, 1990; Reynolds *et al.*, 1999; Zhang *et al.*, 1996). Because of the wide range of lesions produced in this condition and the predisposition to formation of oesophageal adenocarcinoma, the initiation of this tumour is postulated to occur through a process of field cancerization followed by the clonal expansion of geographically distinct but genetically related progeny to metaplastic lesions and tumours (Dunn *et al.*, 2000b; Galipeau *et al.*, 1999; van Dekken *et al.*, 1999; Zhuang *et al.*, 1996). Similar hypotheses have also been proposed for HNSCC, skin and oral tumourigenesis which report a high frequency of multiple tumours and second primary carcinomas (Califano *et al.*, 1996; Cense *et al.*, 1997; Hittelman *et al.*, 1993; Kanjilal *et al.*, 1995; Lydiatt *et al.*, 1998; Piccinin *et al.*, 1998; Roth & Sage, 1969; Scholes *et al.*, 1998; Slaughter *et al.*, 1953; Tian *et al.*, 1998).

1.8 Aims of project

The aim of the present study was to apply a molecular genetics approach to the investigation of chromosomal changes and other genetic alterations associated with the development of human non-small cell lung cancer. Specific aims included

1. In-depth analysis of tumour specimens in which there were multiblocks prepared for the whole lung resection and also partial lobe resection. The objective of this approach was to obtain a detailed analysis of allelic imbalance and altered expression of specific genes in the complete lung or lobe. This detail will provide evidence of clonality and tumour progression within the investigated individual. Furthermore, it will support or question the reliability of single specimen acquisition for a tumour and if this may be considered representative of the lung specimen.
2. Detailed allelic imbalance analysis of chromosomal region 9p22-23 in NSCLC tumours as this region is considered to contain a putative tumour suppressor gene. The objective of this approach was to fine map the region around the microsatellite marker D9S157 in a panel of NSCLC tumours. Small regions of deletion were then investigated further using contig mapping and candidate gene search techniques.
3. Allelic imbalance analysis of chromosome 5q in NSCLC tumours. The objective of this approach was to narrow the region of deletion on chromosome 5q in the region of 5q11-q13 in a panel of NSCLC tumours to enable contig mapping and candidate gene search. This would provide evidence for possible candidate genes within this chromosomal region involved in NSCLC.
4. High throughput allelic imbalance of analysis of chromosome 17. The objective was the development of a fluorescent microsatellite multiplex approach used to analyse chromosome 17 allelic imbalance in a panel of NSCLC tumours based on a previous fluorescent microsatellite analysis sensitivities and limits devised by Liloglou *et al* (2000).

CHAPTER 2. MATERIALS AND METHODS

2.1 Tissue preparation, lysis and DNA extraction

2.1.1 Ethical Approval

Ethical approval for this research was granted by the Liverpool Research Ethics Committee.

2.1.2 Bronchial Sample Collection

Bronchial samples were collected from tissue obtained at the time of surgical resection from patients with lung tumours. Frozen tissue was collected immediately after surgical resection and snap frozen in liquid nitrogen for transportation and storage. Whole lungs or lobes were inflated with 10% neutral buffered formalin post surgery and fixed for either 24 hours or overnight. Lung tissue was examined and sampled by Dr. John Gosney in the University of Liverpool Pathology Department. Samples were selected for multiblock analysis only if the primary tumour had arisen from or invaded the bronchial airway. Once selected for analysis, lung tissue was carefully sliced at intervals of approximately 1cm from the surgical resection margin, through the primary tumour and into the airway of the distal lung. At each interval a sample of the same bronchial airway and the tumour were taken for analysis. Bronchial samples were stored in 10% neutral buffered formalin until paraffin embedding.

2.1.3 Paraffin Embedding of Bronchial samples

Briefly bronchial samples were dehydrated through a series of ethanols and xylene then embedded in paraffin wax under vacuum using the Citadel 2000 autoprocessor and Histocentre 2 (Shandon). Full protocols for tissue processing may be seen in appendix A. Wax blocks were left to solidify overnight.

2.1.4 Sectioning of paraffin embedded blocks

Paraffin embedded blocks were placed on ice 10 minutes prior to sectioning.

Tissue sections were cut at a thickness of 4 μM for Haematoxylin and Eosin (H and E) staining and 8 μM for tissue microdissection using a microtome (Shandon).

Wax sections were then placed onto the surface of clean distilled water in a heated water bath at 45°C to eliminate creases within the section. Tissue was then floated onto clean non-coated glass slides and dried for 15-20 minutes at 65°C (Microdissection) or overnight at 37°C (H and E).

2.1.5 Haematoxylin and Eosin staining of tissue sections

All samples for H and E staining were sectioned at a thickness of 4 μM . H and E staining was carried out using the LinistainGL automated stainer (Shandon). Full protocols can be seen in appendix A.

2.1.6 Microdissection of tissue using micromanipulators

Prior to microdissection all lung samples were examined by the reporting pathologist (Dr. John Gosney) and any regions of interest marked on the slide for future reference. Manual microdissection was carried out using 6 X 8 μM sections floated onto uncoated slides. Sections were dried at 65°C for 15-20 minutes then dewaxed in xylene and left to air dry. Microdissection was carried out using a Eclipse TE300 inverted microscope with attached micromanipulators (Nikon). Small areas of tissue were dissected by scraping away the tissue around the area of interest using a syringe needle (Terumo 25G X 5/8" 0.5 X 16mm) fitted on the end of the micromanipulator arm. Larger areas were dissected in a similar fashion using a heat-stretched glass capillary tube or a small scalpel blade. Once tissue

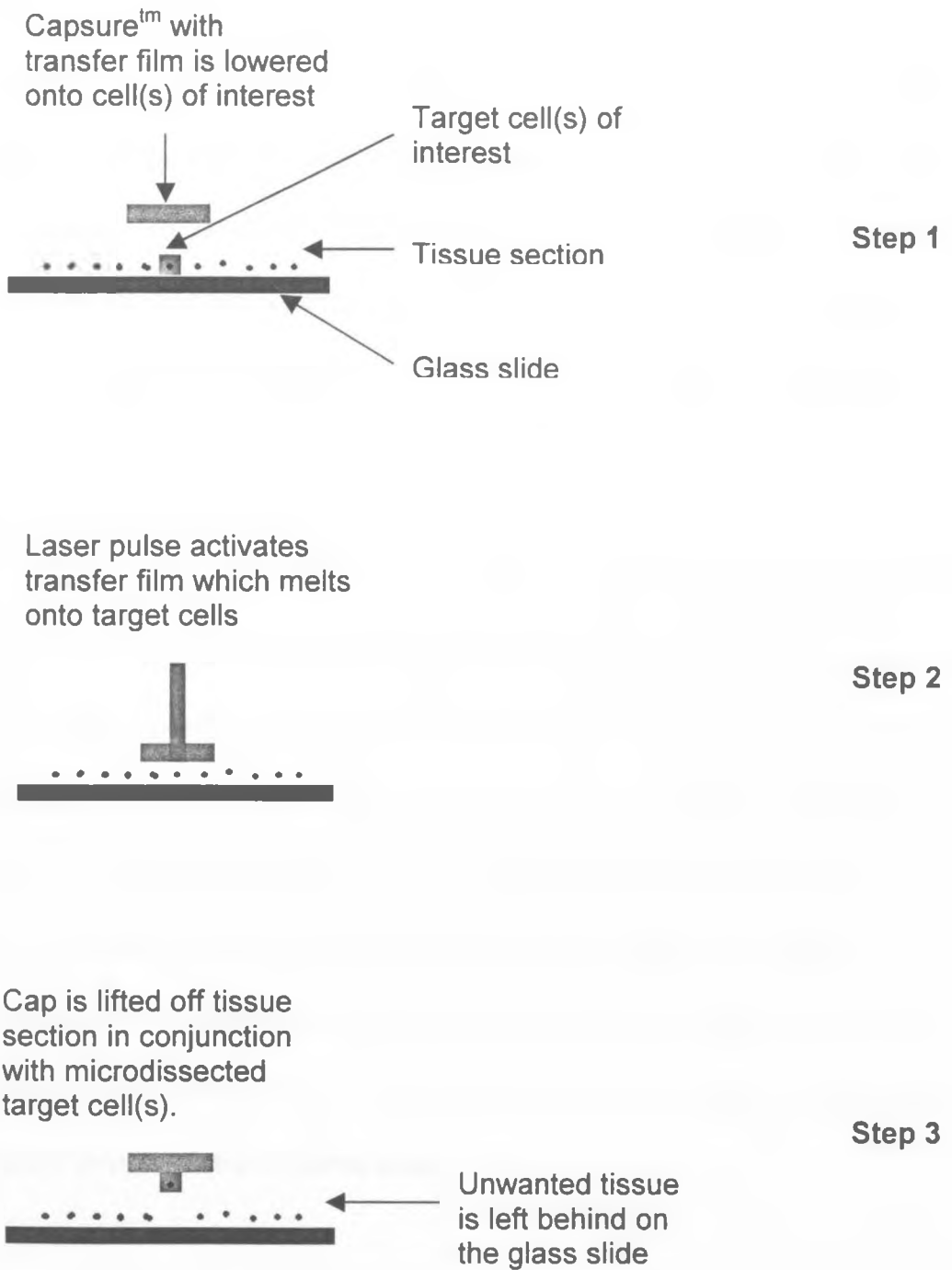
was completely isolated debris was cleaned off the slide using a paper tissue and microdissected cells were aspirated off slides into a 0.5 ml microfuge tube after suspension in a small volume of sterile water. After multiple collections samples were centrifuged at 12,000g for 2 minutes and the sterile water aspirated off the pellet. Tissue was lysed overnight in 50 μ l of 1 X P lysis buffer at 37-42°C. Following lysis proteinase K (Qiagen, UK) was inactivated by incubation at 95°C for 10 minutes. All samples were stored at -20°C.

2.1.7 Microdissection of tissue using Laser Capture Microdissection (LCM)

Five 8 μ m sections were cut from tissue blocks and floated onto uncoated slides. Slides were dried for 10-20 mins at 60°C. Sections were dewaxed in Xylene then stained according to a modified Arcturus H + E protocols (<http://www.arctur.com> ; Appendix A). Following a 10 minute incubation in xylene to ensure complete dehydration, samples were laser microdissected using the Arcturus Pixcell II™ system. An overview of LCM can be seen in figure 2.1. Polymer coated caps with laser captured cells were fitted into 0.5 ml microfuge tubes containing 50 μ l 1X P lysis buffer and incubated in an inverted position overnight. Following lysis proteinase K was inactivated by incubation at 95°C for 10 minutes. All samples were stored at -20°C.

Figure 2.1 Overview of Laser Capture Microdissection. Tissue sections were prepared according to recommended protocols then inspected prior to LCM to eliminate any creases within the tissue to ensure adequate contact of the LCM cap. For bronchial sections loose cartilage was removed prior to LCM. In step 1 the heat-activated polymer-coated cap is lowered onto the tissue slide onto the cells of interest. In step 2, once the laser is activated and fired at the cap (the laser does not contact the tissue slide at any time) the polymer coating melts and sticks to the target cells. In step 3 as the cap is removed from the slide the target cells which are attached to the cap are also removed, leaving behind any unwanted tissue on the slide.

Figure 2.1 Overview of Laser Capture Microdissection.



2.1.8 DNA isolation and purification from paraffin tissue

Paraffin tissue was lysed overnight in 50 µl of 1 X P lysis buffer at 37-42°C. Following lysis proteinase K was inactivated by incubation at 95°C for 10 minutes. After a brief centrifuge spin the volume was made upto 200 µl using sterile water. Following isopropanol extraction, DNA was precipitated in a quarter volume of 10M ammonium acetate and two volumes of absolute ethanol (-20°C), washed in 70% ethanol (-20°C), dried and resuspended in sterile double-distilled water (50 µl).

2.1.9 DNA isolation from frozen lung tissue

Five 10 micron sections were cut onto slides using a cryostat (Shandon, UK), then microdissected using a corresponding haematoxylin and eosin stained slides as a guide using sterile needles to ensure $\geq 70\%$ tumour cells. Sections were lysed in 400 mM Tris-HCl (pH 8.0), 150 mM NaCl, 60 mM EDTA , 1% SDS, 100 µg/ml Proteinase K and incubated at 42°C for 16hr in an orbital shaker. DNA was deproteinized by phenol/chloroform and chloroform extraction followed by precipitation using an equal volume of isopropanol. DNA was washed with 70% ethanol and resuspended in 200µl TE buffer. Working stocks were prepared by 5-fold dilution in double-distilled sterile water.

2.1.10 DNA isolation from venous blood samples

DNA was extracted from 3mls whole venous blood. Specimens were washed with 10 mM Tris-HCl (pH 8.0), 320 mM sucrose, , 1% Triton X-100 and white blood cells were collected by centrifugation. Pellets were lysed in 400 mM Tris-HCl (pH 8.0), 150 mM NaCl, 60 mM EDTA , 1% SDS, 100 µg/ml Proteinase K and incubated at

42°C for 16hr in an orbital shaker. DNA was deproteinized by phenol/chloroform and chloroform extraction followed by precipitation using an equal volume of isopropanol. DNA was washed with 70% ethanol and resuspended in 200µl TE buffer. Working stocks were prepared by 5-fold dilution in double-distilled sterile water.

2.2 Non-fluorescent PCR amplification and PAGE

2.2.1 Primer design and selection

All primers were designed by importing sequences from the GDB sequence database (<http://www.gdb.org>) for analysis using the Oligo™ primer design program. Primers were specifically designed whenever possible to limit large T_m differences. Hairpin structures and dimer formation caused by complementary sequence between sense and antisense primer pairs which could limit the efficiency of DNA amplification during PCR were also eliminated. Primers were obtained from MWG Biotech (Germany) and resuspended in sterile water to a stock concentration of 100 pmol/µl. All primers were stored at -20°C.

2.2.2 Whole Genome Amplification and Degenerate Oligonucleotide PCR (DOP-PCR)

DOP-PCR was performed in a 50µl final volume containing 2µM DOP primer (5'-CCGACTCGAGNNNNNNATGTGG-3') 200µM dNTPs, 10mM Tris.HCl (pH 9.0), 50mM KCl, 1.5mM Magnesium chloride, 0.1% Triton X-100, 0.01% gelatine and 2.5 Units of *Taq* polymerase overlaid with mineral oil. PCR consisted of initial denaturation at 96°C for 8 min followed by 8 cycles of 93°C for 1 min, 30°C for 1

min and 72°C for 3 min, then 28 cycles of 93°C for 1min, 60°C for 1 min and 72°C for 3 min. Whole genome amplification was performed in a 25µl final volume containing 2.5µl (1 µg/µl) of primer (5'-NNNNNN(G/C)(G/C)(G/C)-3'), 2.5µl of 10 x reaction buffer (50mM Tris.HCl (pH 8.4), 15mM Magnesium chloride, 0.1% gelatine, 1% Triton X-100), 4µl of 1.25mM dNTPs, 5 Units of *Taq* polymerase and 5 µl of template DNA (2ng) overlaid with mineral oil. PCR consisted of one step-cycle file of 60°C for 3min and 94°C for 5 min then 30 thermo-cycles of 2.75 min ramp to 94°C/ 94°C for 10 sec and 2 min ramp to 24°C/ 24°C for 10 sec. 10 µl of DOP-PCR or whole genome amplification products were used for a second PCR reaction containing microsatellite primers.

2.2.3 Polymerase Chain Reaction and PAGE

DNA or cell lysis products(1-2µl) were amplified for single marker microsatellite analysis in a 25µl reaction volume using between 5-10 pg of sense and antisense primer equimolar mix, 0.2 mM dNTP mix, 0.5 Units *Taq* polymerase (Roche) and 1 X *Taq* Polymerase Buffer containing 1.5 mM MgCl₂. PCR consisted of denaturation at 95°C for 5 minutes then 28 cycles of 95°C for 45 sec, 57°C for 45 sec and 72°C for 20 sec followed by an extended annealing step of 72°C for 10 minutes (Slight variations to the T_m and annealing steps of these cycles were made depending on the microsatellite markers used). Ten microliters of PCR products were electrophoresed through 10% polyacrylamide gels and visualised by silver staining.

2.2.4 Silver Staining of Polyacrylamide gels

After electrophoresis polyacrylamide gels were fixed for 10 minutes in 300mls gel fixative (10% Ethanol, 10% Acetic Acid). Following rinsing with distilled water, gels were stained by immersion in 300 mls of silver nitrate solution (0.1% w/v) with agitation. After rinsing with distilled water bands, gels were developed by immersion in 500 mls of developer solution (1.5% (w/v) Sodium Hydroxide, 0.06% (v/v) Formaldehyde) with agitation until DNA bands could be clearly defined. Excess sodium hydroxide was neutralised by immersion for 3-5 minutes with agitation in Sodium Carbonate solution (0.75%). Gels were immediately sealed in clear plastic covering and photographed over a white light source (UVIBand software version 97).

2.2.5 Methods and criteria for assessing allelic instability using PAGE followed by silver staining

Once photographed using the UVIBand, allelic instability was scored by measuring the optical density of large and small alleles within each sample using the UVIBand analysis software. Initially the gel pictures were inverted to display a negative image then both the large and small alleles for each sample (normal and tumour) were tagged for a comparative optical density analysis (figure 2.2). The optical density was measured for each allele and displayed as two peaks of the which the average area was calculated. Total volumes were then calculated for each peak (figure 2.3). The total volume of the peak for allele 2 (the smaller allele) was then given a reference value of 1.000, which was then compared to the larger allele to provide an allele ratio for each normal and tumour sample (figure 2.4). It is important to note that the main alleles were recalculated and validated against one

another only within each DNA sample and not between normal and tumour DNA samples from the same individual or DNA samples from different individuals. This ensured that no bias occurred within the calculated values due to differences in DNA quality and concentration between samples and also PCR reaction kinetics. The allele ratio for each normal and tumour pair was then compared to calculate the percentage of allelic instability within each sample using the following methods. The % volume for the smaller allele was divided by the larger allele in both the normal and tumour. Then the tumour value was divided by the normal value to give a value denominated by x .

$$x = \frac{\text{Tumour DNA (\% volume small allele / \% volume large allele)}}{\text{Normal DNA (\% volume small allele / \% volume large allele)}}$$

For the purposes of this study a 30% or greater difference in either the larger or smaller alleles within the tumour DNA, when compared to the normal DNA, was scored as allelic instability. Therefore if $x < 0.7$ or $x > 1.3$ then allelic instability was scored as positive in the tumour for that marker, otherwise the tumour was scored as heterozygous (normal) for that marker.

Figure 2.1. Screen capture from UVIBand software analysis program displaying PAGE image manipulation prior to O.D. analysis. The original image (A) was inverted to display a negative image (B). A rectangle was tagged to the image area selected to be analysed (C). In this case the area containing both the larger and smaller alleles for each sample (normal (N) and tumour (T) DNA) were enclosed within two purple rectangles for O.D. analysis.

Figure 2.1A Normal image

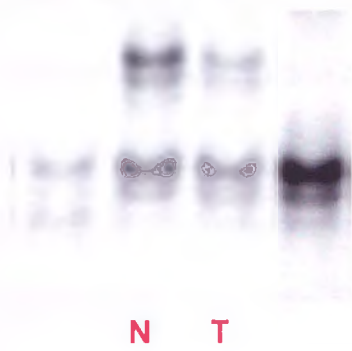


Figure 2.1B Inverted Image

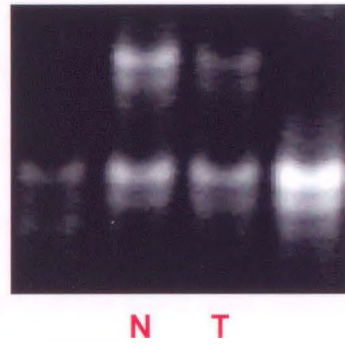


Figure 2.1C Selection of area for O. D. analysis

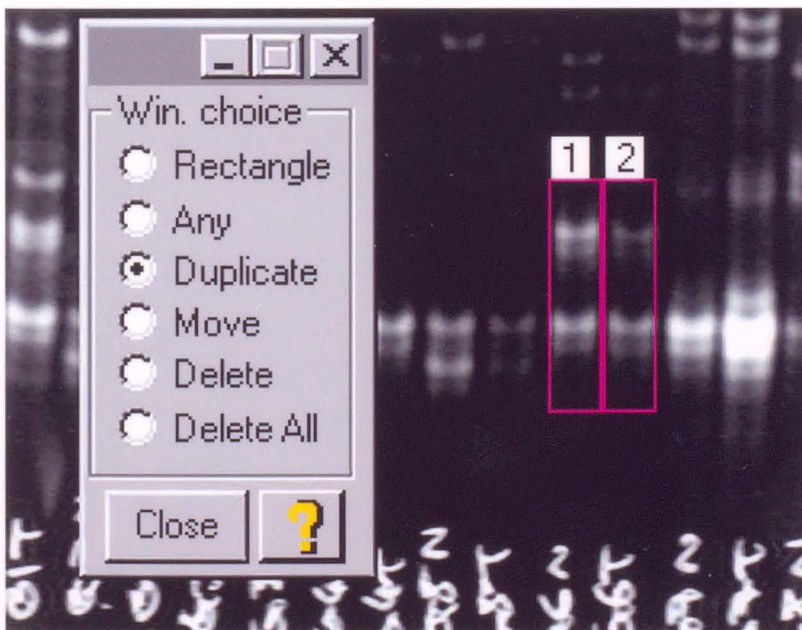


Figure 2.2. Screen capture displaying band identification and separation post O. D. analysis. Alleles were analysed using O. D analysis to produce an image profile usually represented by two peaks (A). Noise was eliminated by compiling an average representation of the image profile (B). Individual alleles for each sample were isolated by manual separation of each peak within the profile. This allowed the elimination of shadow bands and other artefacts which appeared when using microsatellite analysis. Peak volume, height and area were then calculated for each allele in both the normal and tumour DNA (C and D).

Figure 2.2A Complete image profile post O.D. analysis

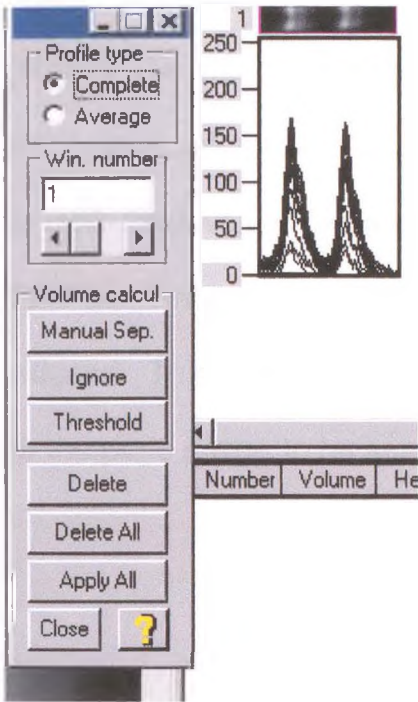


Figure 2.2B Average image profile post O.D. analysis (Noise processing)

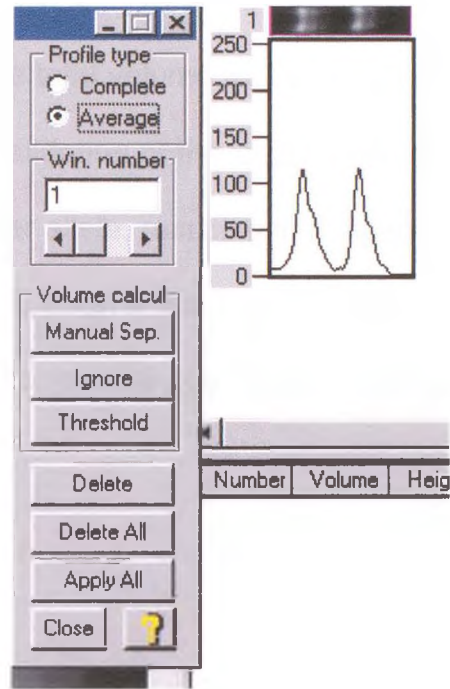


Figure 2.2C Manual Separation of allele peak areas for normal DNA sample

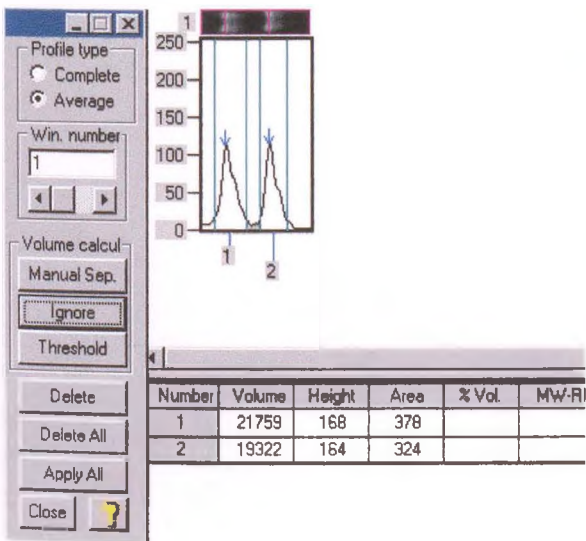


Figure 2.2D Manual Separation of allele peak areas for tumour DNA sample

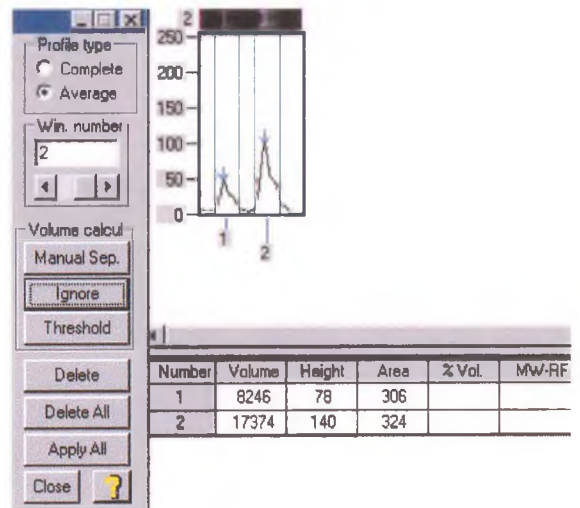


Figure 2.3. Screen capture displaying volume reference recalculation for normal (A) and tumour (B) allele peaks. Each normal and tumour sample was recalculated to express the peak volume of the smaller allele as a ratio of 1.000 (100%). The analysis software converted the total volume for the small allele into a percentage volume with a value of 1.000. Validation of the larger allele expressed its own peak volume as a ratio of the smaller allele, resulting in a higher % volume if the band intensity was greater or a lower % volume if the band intensity was reduced . This method was applied to all other normal and tumour DNA samples on the page image to be analysed.

Figure 2.3A Volume reference recalculation for normal DNA

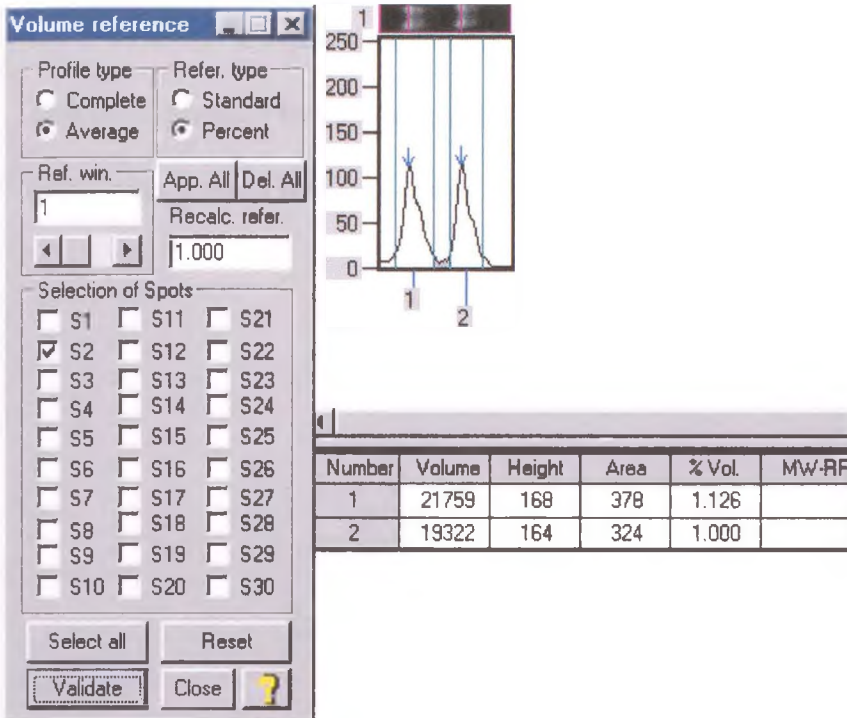


Figure 2.3B Volume reference recalculation for tumour DNA

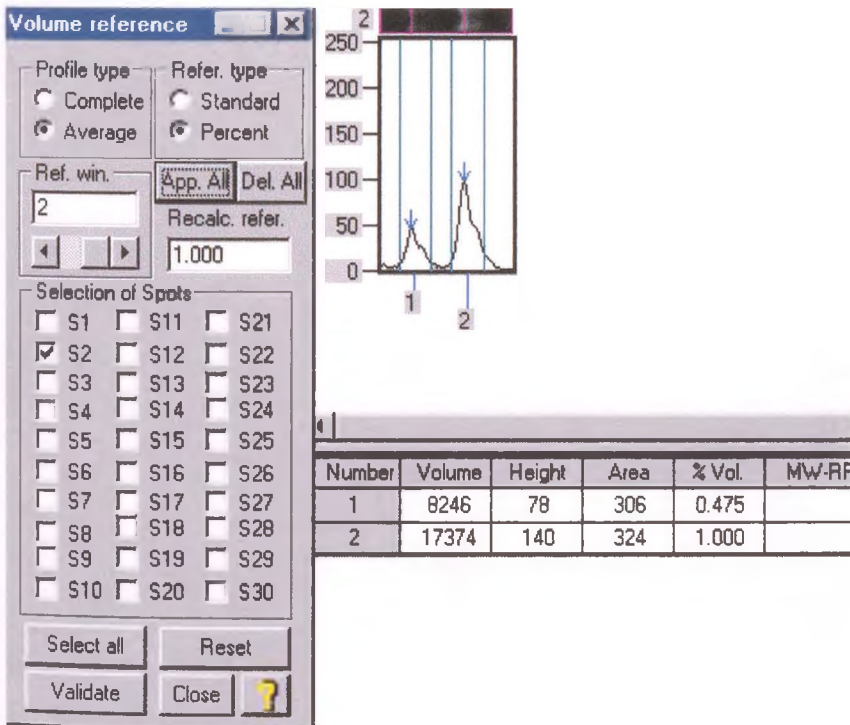


Figure 2.4. Screen capture displaying histogrammatic representation of the recalculated % volume of each allele for lung normal (A) and tumour (B) .
The difference in the allele intensity between the normal and tumour can be clearly seen with a reduction in allele 1 in the tumour compared to the normal.

Figure 2.4A Histogrammatic representation of recalculated % volume for normal DNA.

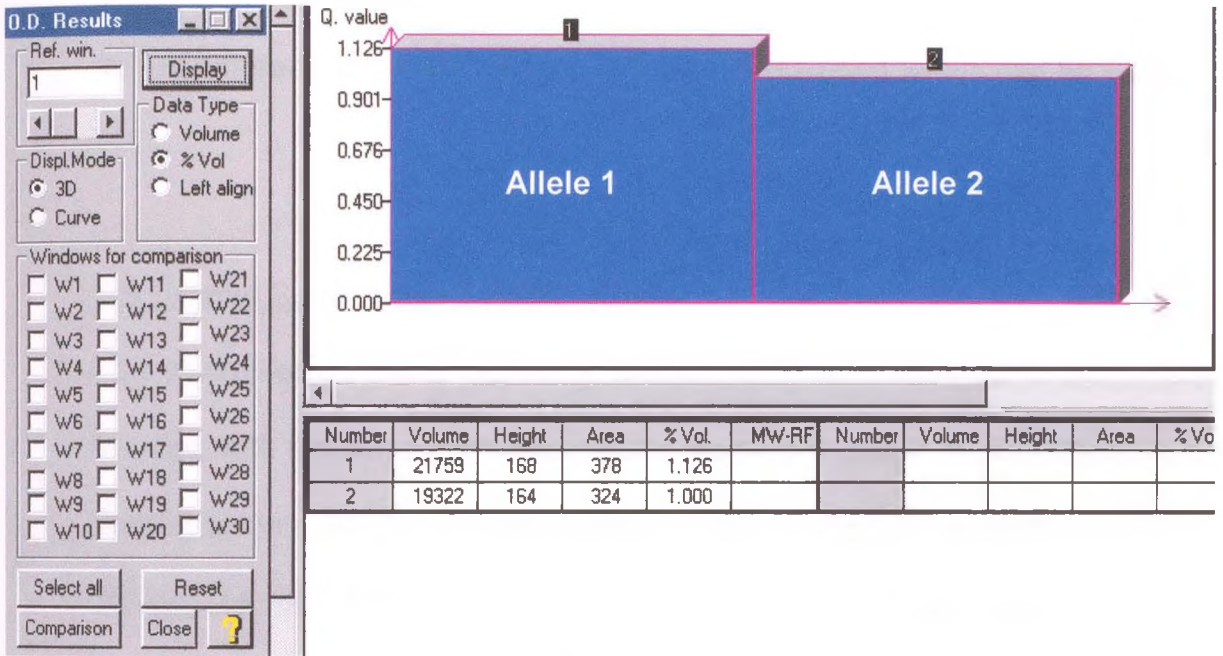
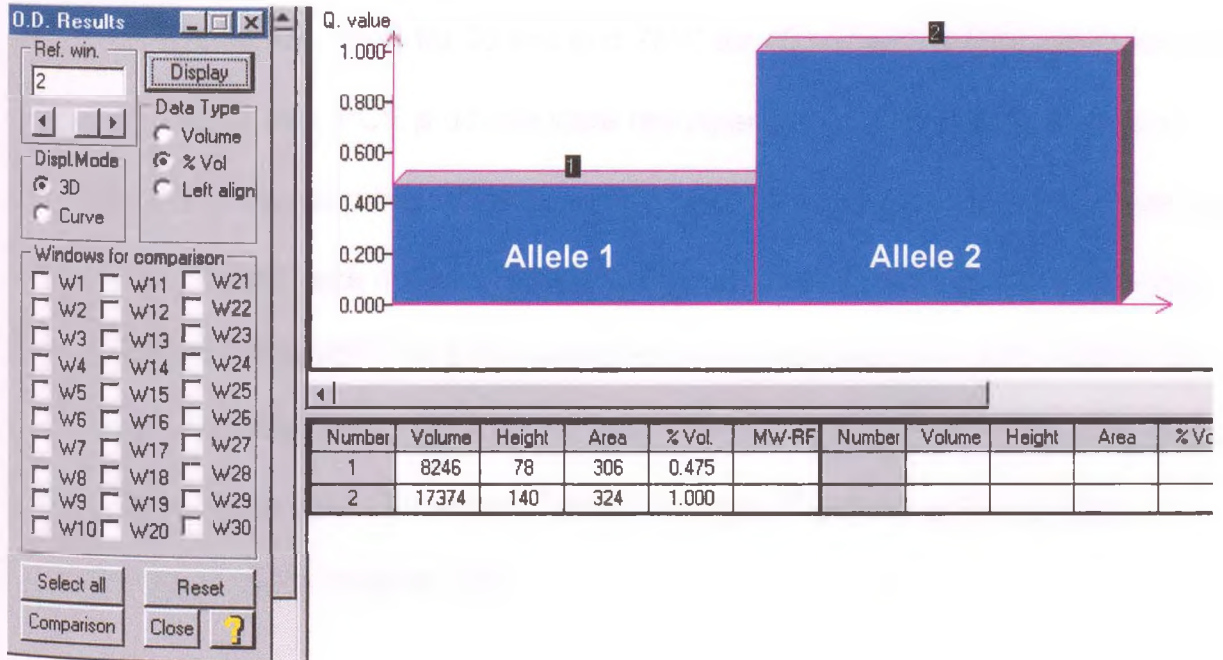


Figure 2.4B Histogrammatic representation of recalculated % volume for tumour DNA.



2.3 Multiplex analysis using the ABI-310™ and ABI-377™

Microsatellite markers for bronchial and HNSCC analysis were selected from the Linkage Mapping Set V2.0 (PE-Applied Biosystems, Warrington, UK). Information on the different panels of markers and their fluorescent labels can be seen in table 2.1.

2.3.1 Panel S+ Multiplex amplification

Four microliters of bronchial tissue lysis products were amplified in a multiplex reaction with the panel S+ microsatellite markers (Table 2.1). Amplification was carried out in a 10µl volume containing 1 x GeneAmp™ Gold Buffer, 0.5mM dNTPs, 2.5 mM MgCl₂ , and 0.034 U AmpliTaq Gold™ polymerase (PE-Applied Biosystems, Warrington, UK). PCR amplification consisted of an initial denaturation for 12 min at 95°C to activate Taq polymerase followed by 10 cycles consisting of 94°C for 30 sec, 55°C for 30 sec and 72°C for 30 sec, then 18 cycles consisting of 92°C for 30 sec, 55°C for 30 sec and 72°C for 45 sec with a final extension step at 72°C for 20 min. PCR products were resuspended in 60µl of 70% ethanol to remove unbound primers and to concentrate DNA samples then mixed with 4µl loading buffer: size marker ROX-350, dextran blue, formamide (1: 1: 5). After denaturation at 95°C for 5 min samples were analysed on an ABI-PRISM 377 automatic sequencer (PE-Applied Biosystems, Warrington, UK). Results were analysed using the Genescan™ and Genotyper™ software (PE-Applied Biosystems, Warrington, UK).

2.3.2 Panel F (Chromosome 3) Multiplex amplification

Four microliters of bronchial tissue lysis products were amplified in each of the two multiplex reactions with the panel F microsatellite markers (Table 2.1). Large and small markers were amplified in separate reactions using different reaction conditions and PCR profiles. For small markers amplification was carried out in a 10 μ l reaction volume containing 1 x GeneAmp™ Gold Buffer, 0.5mM dNTPs, 2.5 mM MgCl₂ , 0.026 U AmpliTaq Gold™ polymerase (PE-Applied Biosystems, Warrington, UK) and pre-optimised amounts of primers. PCR amplification consisted of an initial denaturation for 11 min at 95°C to activate Taq polymerase followed by 10 cycles consisting of 94°C for 30 sec, 55°C for 30 sec and 72°C for 25 sec, then 18 cycles consisting of 92°C for 30 sec, 55°C for 30 sec and 72°C for 35 sec with a final extension step at 72 for 20 min. For large markers amplification was carried out in a 10 μ l reaction volume containing 1 x GeneAmp™ Gold Buffer, 0.55mM dNTPs, 2.5 mM MgCl₂ , 0.03 U AmpliTaq Gold™ polymerase (PE-Applied Biosystems, Warrington, UK) and pre-optimised amounts of primers. PCR amplification consisted of an initial denaturation for 11 min at 95°C to activate Taq polymerase followed by 10 cycles consisting of 94°C for 30 sec, 55°C for 30 sec and 72°C for 40 sec, then 18 cycles consisting of 92°C for 30 sec, 55°C for 30 sec and 72°C for 45 sec with a final extension step at 72°C for 20 min. PCR products from both reactions were mixed together then resuspended in 120 μ l of 70% ethanol to remove unbound primers and to concentrate DNA samples then mixed with 4 μ l loading buffer: size marker ROX-350, dextran blue, formamide (1: 1: 5). After denaturation at 95°C for 5 min samples were analysed on an ABI-PRISM 377 automatic sequencer (PE-Applied Biosystems, Warrington, UK). Results were

analysed using the Genescan™ and Genotyper™ software (PE-Applied Biosystems, Warrington, UK).

2.3.3 Panel P (Chromosome 17) Multiplex amplification

Four microliters of bronchial tissue lysis products were amplified in each of the two multiplex reactions with the panel P microsatellite markers (Table 2.1). Large and small markers were amplified in separate reactions using different reaction conditions and PCR profiles. For small markers amplification was carried out in a 10µl reaction volume containing 1 x GeneAmp™ Gold Buffer, 0.5mM dNTPs, 2.5 mM MgCl₂ , 0.026 U AmpliTaq Gold™ polymerase (PE-Applied Biosystems, Warrington, UK) and pre-optimised amounts of primers. PCR amplification consisted of an initial denaturation for 11 min at 95°C to activate Taq polymerase followed by 10 cycles consisting of 94°C for 30 sec, 55°C for 30 sec and 72°C for 25 sec, then 18 cycles consisting of 92°C for 30 sec, 55°C for 30 sec and 72°C for 35 sec with a final extension step at 72 for 20 min. For large markers amplification was carried out in a 10µl reaction volume containing 1 x GeneAmp™ Gold Buffer, 0.55mM dNTPs, 2.5 mM MgCl₂ , 0.026 U AmpliTaq Gold™ polymerase (PE-Applied Biosystems, Warrington, UK) and pre-optimised amounts of primers. PCR amplification consisted of an initial denaturation for 11 min at 95°C to activate Taq polymerase followed by 10 cycles consisting of 94°C for 30 sec, 55°C for 30 sec and 72°C for 40 sec, then 18 cycles consisting of 92°C for 30 sec, 55°C for 30 sec and 72°C for 45 sec with a final extension step at 72°C for 20 min. PCR products from both reactions were mixed together then resuspended in 120µl of 70% ethanol to remove unbound primers and to concentrate DNA samples then mixed

with 4µl loading buffer: size marker ROX-350, dextran blue, formamide (1: 1: 5).

After denaturation at 95°C for 5 min samples were analysed on an ABI-PRISM 377 automatic sequencer (PE-Applied Biosystems, Warrington, UK). Results were analysed using the Genescan™ and Genotyper™ software (PE-Applied Biosystems, Warrington, UK).

Table 2.1 Table showing fluorescent microsatellite panels selected from markers within the Linkage Mapping Set V2.0 (PE-Applied Biosystems, Warrington, UK). The microsatellite locus, current cytogenetic locations and the marker size range is displayed for each marker. Also displayed is the concentration (in picograms) of each primer used within the multiplex PCR reaction and which subset reaction was used for PCR. Panel S+ markers were amplified within a single reaction, whereas Panel F and Panel P large and small markers were amplified within separate reactions. 'S' - denotes small marker subset reaction and 'L' - denotes large marker subset reaction.

Table 2.1 Table showing fluorescent microsatellite panels.

Microsatellite Locus	Current Cytogenetic location	Expected Size Range (bp)	Primer Concentration in PCR (pg)	Subset Reaction
Panel S				
D3S1300	3p14.2	235-267	1.35	
D9S161	9p21	126-144	0.5	
D17S2179E	17p13.1	141-171	0.9	
D5S644	5q15	86-116	1.05	
D13S153	13q14	94-126	0.65	
D13S171	13q12.3-q13	179-207	1.4	
D9S157	9p23	229-253	1.65	
Panel F				
D3S1580	3q29	220-240	0.85	L
D3S3521	3p21.3	268-302	0.7	L
D3S1285	3p14.1	230-256	1	L
D3S1277	3p22.1	292-314	1.2	L
D3S1263	3p26.1	194-214	0.9	L
D3S1600	3p14.2	189-205	0.85	L
D3S1276	3p11.1	102-118	1.1	S
D3S2338	3p25.2	94-114	1.1	S
D3S1614	3q26.1	107-131	1.1	S
D3S1309	3q22.2	138-156	0.75	S
D3S3630	3p11.1	108-196	0.85	S
D3S1566	3p13	158-180	0.75	S
Panel P				
D17S1862	17q23	205-235	1.4	L
D17S784	17q25	233-247	1	L
D17S921	17cen	196-214	0.6	L
D17S1876	17p12-p13	104-136	0.6	S
D17S1828	17p12-p13	101-121	0.7	S
D17S787	17q21.3	142-178	0.06	S
D17S831	17p13	112-134	1	S

2.3.4 Analysis of ABI-310 and ABI-377 results using Genescan™ and Genotyper™ software.

Prior to loading for analysis on the ABI-377 or ABI-310 all samples were tagged with a unique ID using the prepared 48 or 96-well sample sheet format to prevent sample mix-up (fig. 2.5). Post electrophoresis all lanes containing samples were tracked and analysed within a gel file to determine the intensity of the fluorescence for each microsatellite marker (fig. 2.6 and fig 2.7). A data file was created from the sample sheet ID and gel file containing all detected microsatellite markers for each individual sample. Sample files were exported to Genescan™ analysis software and assigned a pre-defined size standard matrix (in this case Rox size standard; fig 2.8A). Analysis parameters (analysis range) were fixed by eliminating primer flare (excess fluorescence caused by unbound primers; fig 2.8B) then all results analysed for the wavelength of each primer type detected. Results were checked post Genescan™ analysis to confirm correct size allocation of each peak within the Rox size standard. Once confirmed the Rox standard peaks within all samples were aligned by size (fig 2.9). Analysed sample files were exported to Genotyper™ software for allele measurement and tumour/normal ratio calculation. Expected size ranges for each microsatellite marker and label within the multiplex PCR panel (fig 2.10) were assigned to an initial macro (LOH initial). LOH initial scanned each analysed sample file for the correct size and fluorescent wavelength of microsatellite markers then created profiles for all microsatellite markers containing the peak size, intensity and area for parental alleles within each individual sample (fig 2.11A). Samples were corrected to remove shadow or non-specific peaks (fig. 2.11B) then analysed using a second macro (LOH corrected) to evaluate the allele ratio within every sample for each microsatellite marker. Once allele ratios were

determined, a tumour/normal ratio for each sample was calculated and used to score AI using previously defined limits (Liloglou *et al.*, 2000).

Figure 2.5 Screen capture displaying sample sheet and run control on the ABI-377 prior to sample electrophoresis. Samples were loaded in a 48- or 96-sample multiwell format. The sample sheet allowed accurate analysis of multiple multiplex PCR samples without errors occurring through mix-up by assigning a full ID to each DNA sample.

Figure 2.5 ABI-377 Sample Sheet

▶ Plate Check
▶ PreRun
▶ Run
|| Pause
■ Cancel

Plate Check Module <none>
PreRun Module <none>

Run Module <none>
Sample Sheet <none> ...
Gel's Matrix File <none>
Lanes 96
Run Mode 96 Lane ...

Collect time 3.0 hours
Well-to-Read distance 36 cm
Operator

Lane	Sample Number	Sample Name	Sample File Name	Matrix File	Auto Analyze	Analysis Parameters	Size Standard	Auto Print
1	1	97 22961 N 3p	97 22961 N 3pe	<Gel's Matrix>	<input type="checkbox"/>			<input type="checkbox"/>
2					<input type="checkbox"/>			<input type="checkbox"/>
3	3	97 22961 N Pen S	97 22961 N Pen Se	<Gel's Matrix>	<input type="checkbox"/>			<input type="checkbox"/>
4					<input type="checkbox"/>			<input type="checkbox"/>
5	5	97 22961 N Ch17	97 22961 N Ch17e	<Gel's Matrix>	<input type="checkbox"/>			<input type="checkbox"/>
6					<input type="checkbox"/>			<input type="checkbox"/>
7	7	97 22961 P3 3p	97 22961 P3 3pe	<Gel's Matrix>	<input type="checkbox"/>			<input type="checkbox"/>
8					<input type="checkbox"/>			<input type="checkbox"/>
9	9	97 22961 P3 pan s	97 22961 P3 pan se	<Gel's Matrix>	<input type="checkbox"/>			<input type="checkbox"/>
10					<input type="checkbox"/>			<input type="checkbox"/>
11	11	97 22961 P3 ch 17	97 22961 P3 ch 17e	<Gel's Matrix>	<input type="checkbox"/>			<input type="checkbox"/>
12					<input type="checkbox"/>			<input type="checkbox"/>
13	13	97 22961 P2 3p	97 22961 P2 3pe	<Gel's Matrix>	<input type="checkbox"/>			<input type="checkbox"/>
14					<input type="checkbox"/>			<input type="checkbox"/>
15	15	97 22961 P2 pan s	97 22961 P2 pan se	<Gel's Matrix>	<input type="checkbox"/>			<input type="checkbox"/>
16					<input type="checkbox"/>			<input type="checkbox"/>
17	17	97 22961 P2 ch 17	97 22961 P2 ch 17e	<Gel's Matrix>	<input type="checkbox"/>			<input type="checkbox"/>
18					<input type="checkbox"/>			<input type="checkbox"/>
19	19	97 22961 P2 3p	97 22961 P2 3pe-01	<Gel's Matrix>	<input type="checkbox"/>			<input type="checkbox"/>
20					<input type="checkbox"/>			<input type="checkbox"/>
21	21	97 22961 P2 pan s	97 22961 P2 pan se-0	<Gel's Matrix>	<input type="checkbox"/>			<input type="checkbox"/>
22					<input type="checkbox"/>			<input type="checkbox"/>
23	23	97 22961 P2 ch 17	97 22961 P2 ch 17e-1	<Gel's Matrix>	<input type="checkbox"/>			<input type="checkbox"/>
24					<input type="checkbox"/>			<input type="checkbox"/>

Figure 2.6 Screen capture displaying gel file (magnified) from ABI-377 analysis. Active lanes containing multiplex DNA samples are shown as blue diamonds at the top of the image. Unused lanes are displayed as grey diamonds. Fluorescent markers were represented by the blue, green and yellow bands within each lane (FP). The internal ROX size standard was displayed as a series of red bands within each lane (several ROX positions within the gel are indicated by black arrows at the right of the figure). At the base of each lane a mix of all four colours was detected due to unbound primers (FL). Lanes could be selected by highlighting each diamond individually which would display a series of peaks representing the intensity of each marker analysed within the gel.

Figure 2.6 Screen capture displaying gel file from ABI-377 analysis.

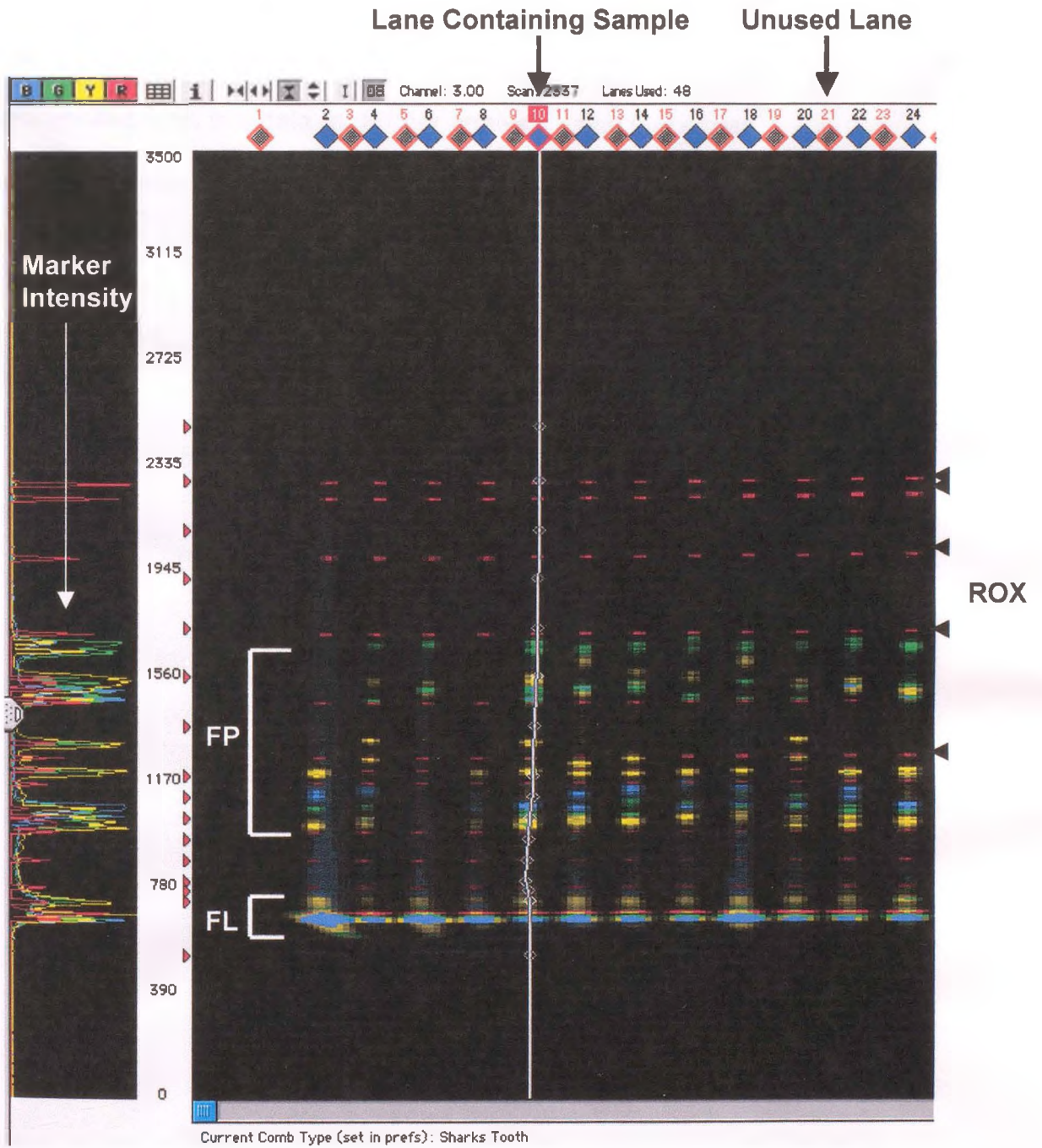


Figure 2.7 Screen capture displaying lane tracking for analysis of fluorescent primers after gel electrophoresis on the ABI-377. Lanes were automatically tracked by analysis software and the intensity of each fluorescent label in addition to the size standard was scanned in every sample. Samples files were created from this data for further analysis using Genescan™ and Genotyper™ software.

Figure 2.7 Screen capture displaying gel file after lane tracking on the ABI-377.

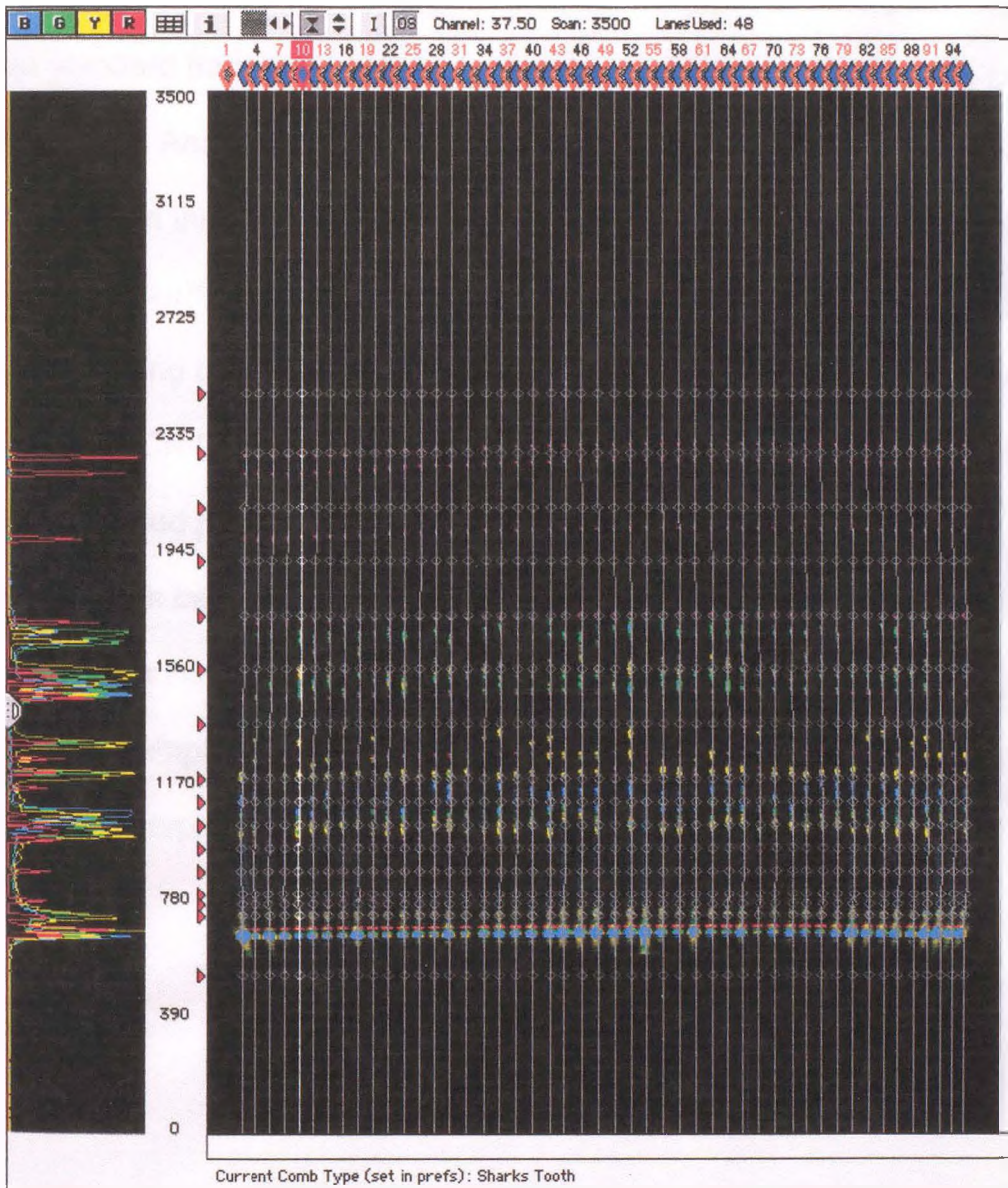


Figure 2.8 Screen capture displaying imported sample file data for size standard (A) and analysis parameter (B) assignment. Sample files were imported into Genescan™ to assign a predefined size in base pairs to every ROX size standard fragment (previously seen as a red band or peak in the gel file) as shown in A. Analysis parameters were also defined to eliminate erroneous results derived from the initial primer flare seen in the gel file as shown in B. This ensured the Genescan™ software only analysed DNA peaks within a specific size range corresponding to microsatellite DNA. In 2.34B an analysis file is displayed for 1 multiplex PCR DNA sample. Y-axis displays peak intensity and the X-axis peak size. Coloured peaks corresponding to each fluorescent label and ROX size standard can be seen (red – ROX size standard, blue – FAM, yellow – NED, green – HEX). The initial large peak highlighted in a green box represents primer flare within the sample. Subsequent peaks within the box marked range of analysis parameters represent actual microsatellite DNA.



Figure 2.8 Screen capture displaying imported sample file data for size standard (A) and analysis parameter (B) assignment.

A

Analyze Print Results Print Setup ...

	Sample File	Size Standard	Parameters
1	D1B	PHI_A000	<Analysis Parameters>
2	E1B	PHI_A000	<Analysis Parameters>
3	F1B	PHI_A000	<Analysis Parameters>
4	D2B	PHI_A000	<Analysis Parameters>
5	E2B	PHI_A000	<Analysis Parameters>
6	F2B	PHI_A000	<Analysis Parameters>
7	D3B	PHI_A000	<Analysis Parameters>
8	E3B	PHI_A000	<Analysis Parameters>
9	F3B	PHI_A000	<Analysis Parameters>
10	D4B	PHI_A000	<Analysis Parameters>
11	E4B	PHI_A000	<Analysis Parameters>
12	F4B	PHI_A000	<Analysis Parameters>
13	D5B	PHI_A000	<Analysis Parameters>
14	E5B	PHI_A000	<Analysis Parameters>
15	F5B	PHI_A000	<Analysis Parameters>
16	D6B	PHI_A000	<Analysis Parameters>
17	E6B	PHI_A000	<Analysis Parameters>
18	F6B	PHI_A000	<Analysis Parameters>
19	D7B	PHI_A000	<Analysis Parameters>
20	E7B	PHI_A000	<Analysis Parameters>
21	F7B	PHI_A000	<Analysis Parameters>
22	D8B	PHI_A000	<Analysis Parameters>
23	E8B	PHI_A000	<Analysis Parameters>
24	F8B	PHI_A000	<Analysis Parameters>
25	D1L	PHI_A000	<Analysis Parameters>
26	E1L	PHI_A000	<Analysis Parameters>
27	F1L	PHI_A000	<Analysis Parameters>
28	D2L	PHI_A000	<Analysis Parameters>
29	E2L	PHI_A000	<Analysis Parameters>
30	F2L	PHI_A000	<Analysis Parameters>
31	D3L	PHI_A000	<Analysis Parameters>
32	E3L	PHI_A000	<Analysis Parameters>
33	F3L	PHI_A000	<Analysis Parameters>
34	D4L	PHI_A000	<Analysis Parameters>
35	E4L	PHI_A000	<Analysis Parameters>
36	F4L	PHI_A000	<Analysis Parameters>
37	D5L	PHI_A000	<Analysis Parameters>
38	E5L	PHI_A000	<Analysis Parameters>
39	F5L	PHI_A000	<Analysis Parameters>
40	D6L	PHI_A000	<Analysis Parameters>
41	E6L	PHI_A000	<Analysis Parameters>
42	F6L	PHI_A000	<Analysis Parameters>
43	<Analysis Parameters>

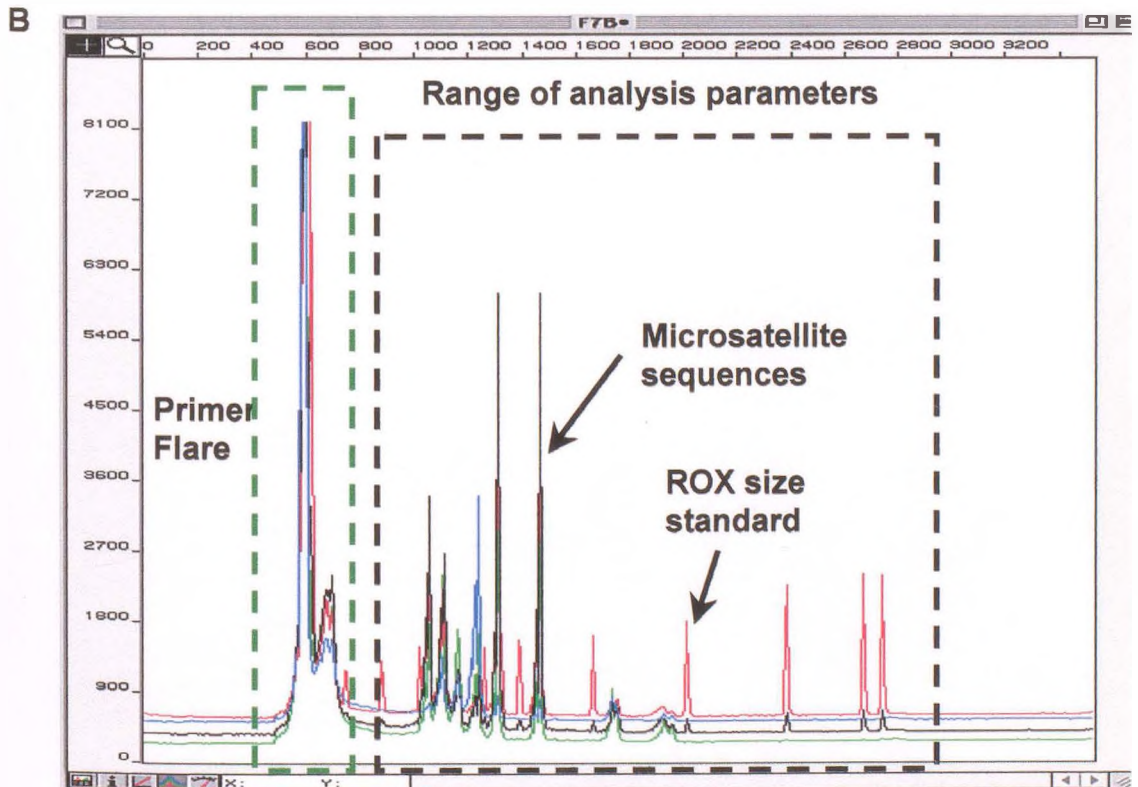


Figure 2.9 Screen capture displaying peak data after Genescan™ analysis and size standard alignment. Samples were analysed by Genescan™ after setting analysis parameters and a size standard range to each sample file. To compensate for slight differences in electrophoresis speeds, the ROX size standard was aligned by size in every genescan sample file. Allele sizes could then be determined for every microsatellite marker within the multiplex PCR reaction. Y-axis displays peak intensity and the X-axis peak size. Coloured peaks corresponding to each fluorescent label and ROX size standard can be seen (red – ROX size standard, blue – FAM, yellow – NED, green – HEX). Note that the original primer flare is now removed from the sample file. Also individual peaks for all fluorescent dye labels have now been assigned sizes relative to the internal ROX size standard.

Figure 2.9 Screen capture displaying peak data after Genescan™ analysis and size standard alignment.

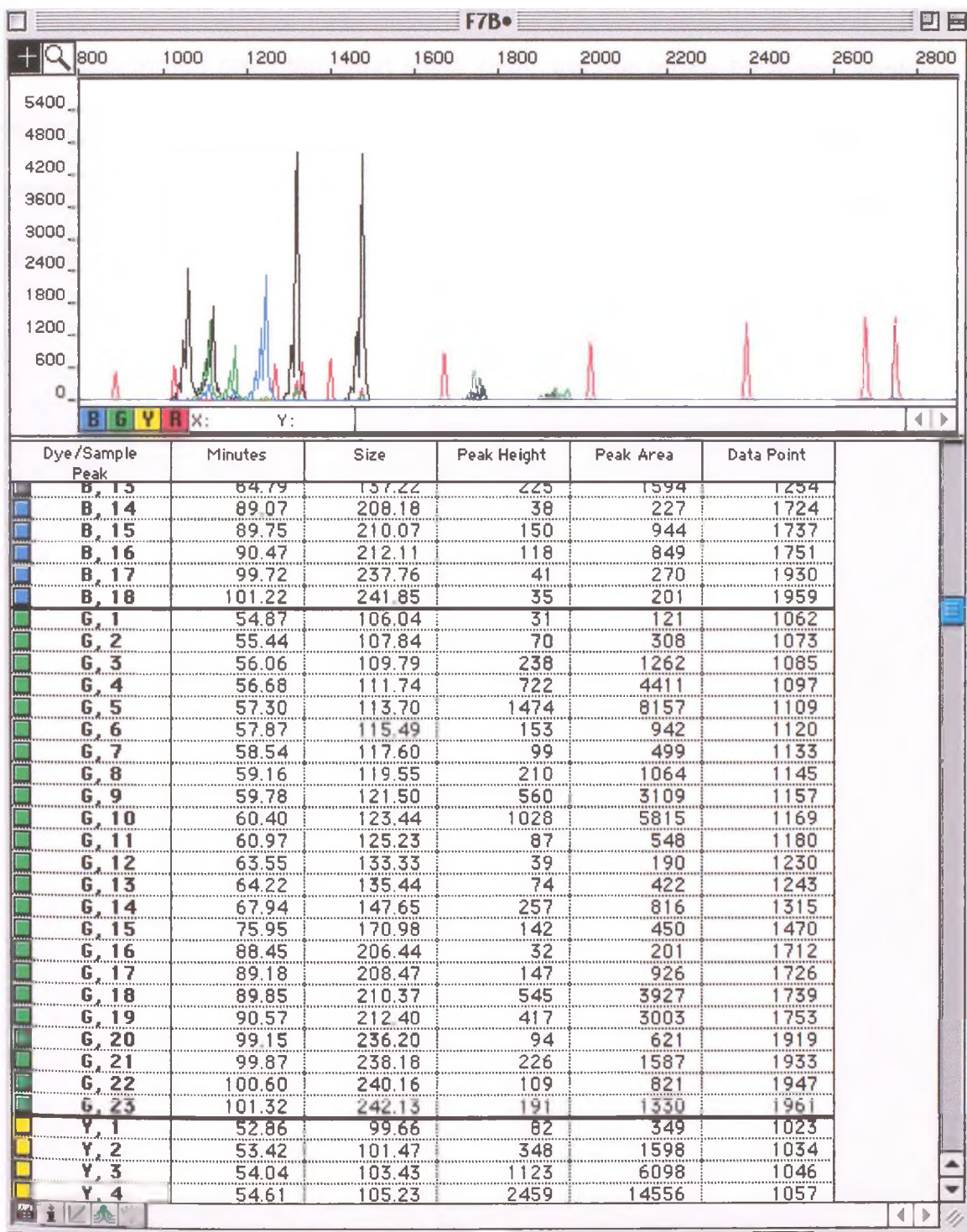


Figure 2.10 Screen capture displaying predefined range in Genotyper™ analysis software for fluorescent dye labelled microsatellite markers in panel P (chromosome 17). All microsatellite markers within the multiplex PCR panels were designated predetermined allele sizes and dye labels for macro analysis within the Genotyper™ software. This ensured that microsatellite markers with identical dye labels, yet different allele (or peak) sizes could be distinguished by the Genotyper™ macro during analysis.

Figure 2.10 Screen capture displaying predefined range in Genotyper™ analysis software for fluorescent dye labelled microsatellite markers in panel P (chromosome 17).

to
 Min height
 Max height

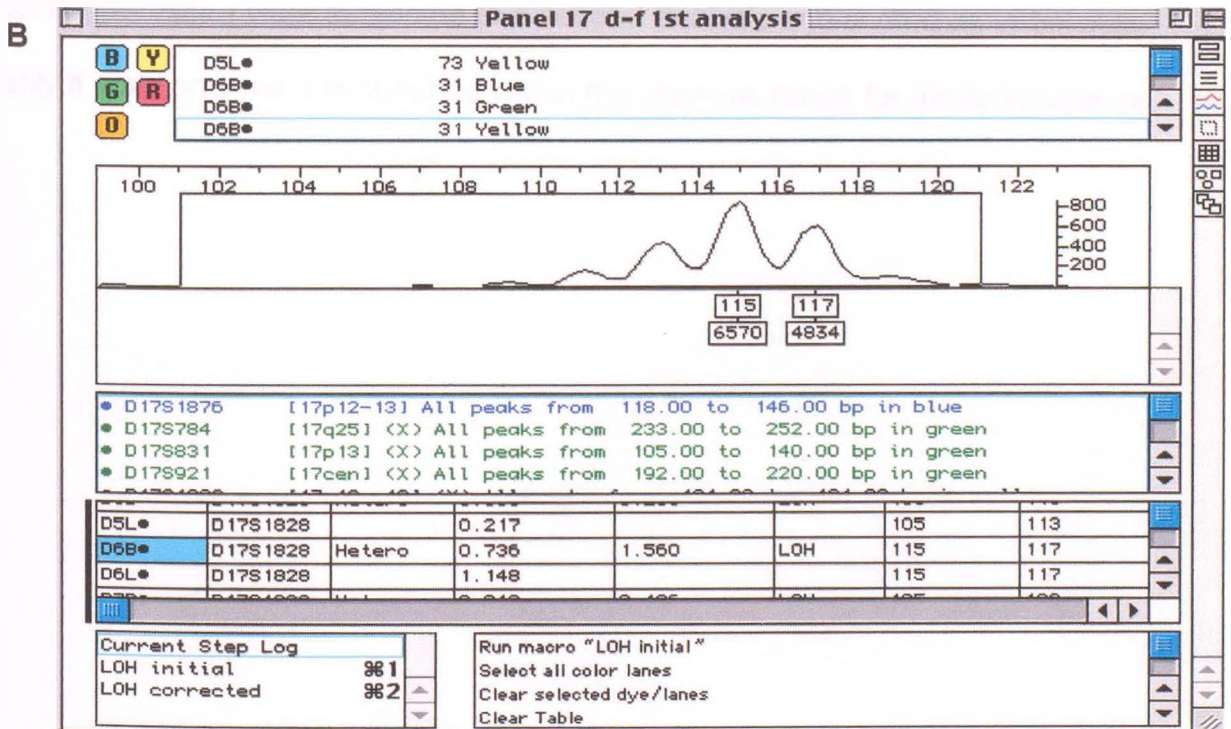
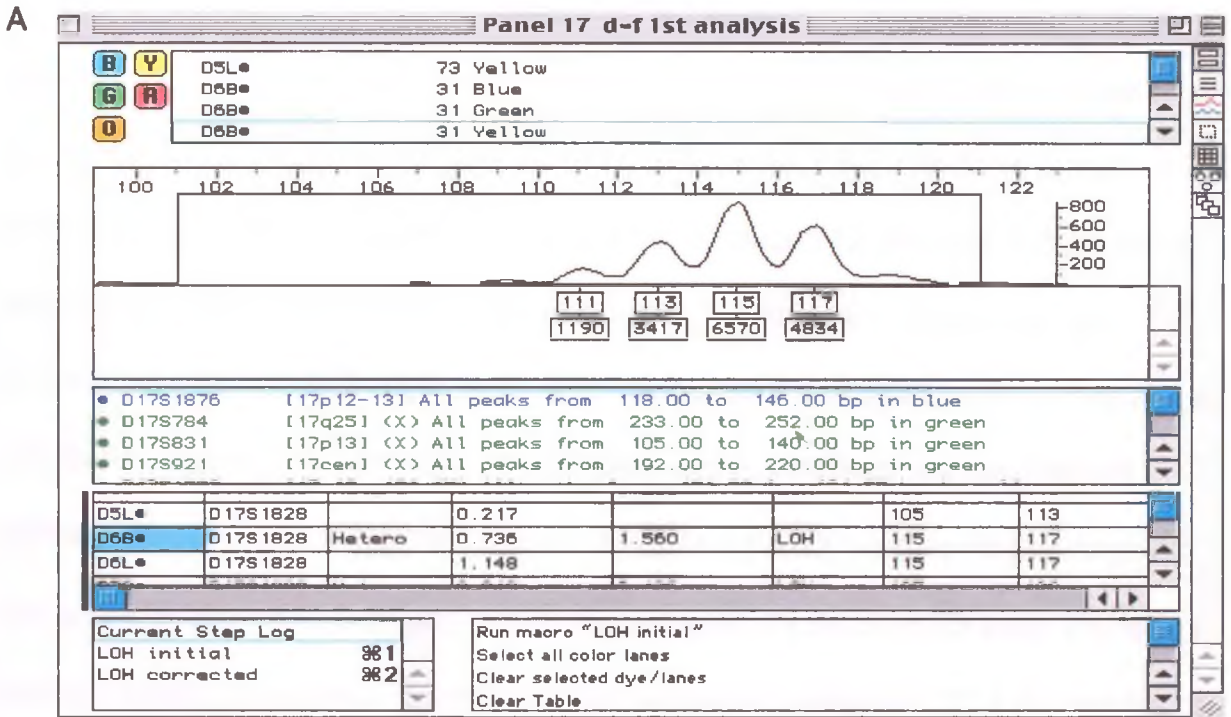
Group
 All Highest Highest
 Leftmost Rightmost
 Exclusive B G Y
 R O

• D17S1876	[17p12-13]	All peaks from 118.00 to 146.00 bp in blue
• D17S784	[17q25] (X)	All peaks from 233.00 to 252.00 bp in green
• D17S831	[17p13] (X)	All peaks from 105.00 to 140.00 bp in green
• D17S921	[17cen] (X)	All peaks from 192.00 to 220.00 bp in green
• D17S1828	[17p12-p13] (X)	All peaks from 101.00 to 121.00 bp in yellow
• D17S1862	[17q23] (X)	All peaks from 203.00 to 238.00 bp in yellow
• D17S787	[17q21.3] (X)	All peaks from 142.00 to 178.00 bp in yellow

Figure 2.11 Screen capture displaying heterozygous DNA sample amplified with D17S1829 after initial (A) and corrected (B) analysis using Genotyper™ macro. Sample files from genescan were imported into Genotyper™ software for further analysis. Using the predetermined limits for marker range and dye label each sample file was analysed for peaks (representing microsatellite alleles) within the specified range and dye colour. Using the 'LOH initial' macro all peaks or bands were labelled for size and intensity including non-specific PCR products. In figure A both main alleles can be seen labelled as 115bp and 117bp with an allele intensity of 6570 and 4834 respectively displayed beneath. However shadow peaks or stutter artefacts were also labelled (peaks at 111bp and 113bp) which could produce an error or even a false AI score when samples are analysed for AI using the second Genotyper™ macro designed to calculate tumour/normal allele ratios for AI analysis. Subsequently labels representing non-specific PCR products or shadow peaks were removed to leave only the main allele labels (B), then a second analysis carried out using the 'LOH corrected' macro to evaluate AI within normal and test DNA.



Figure 2.11 Screen capture displaying heterozygous DNA sample amplified with D17S1829 after initial (A) and corrected (B) analysis using Genotyper™ macro.

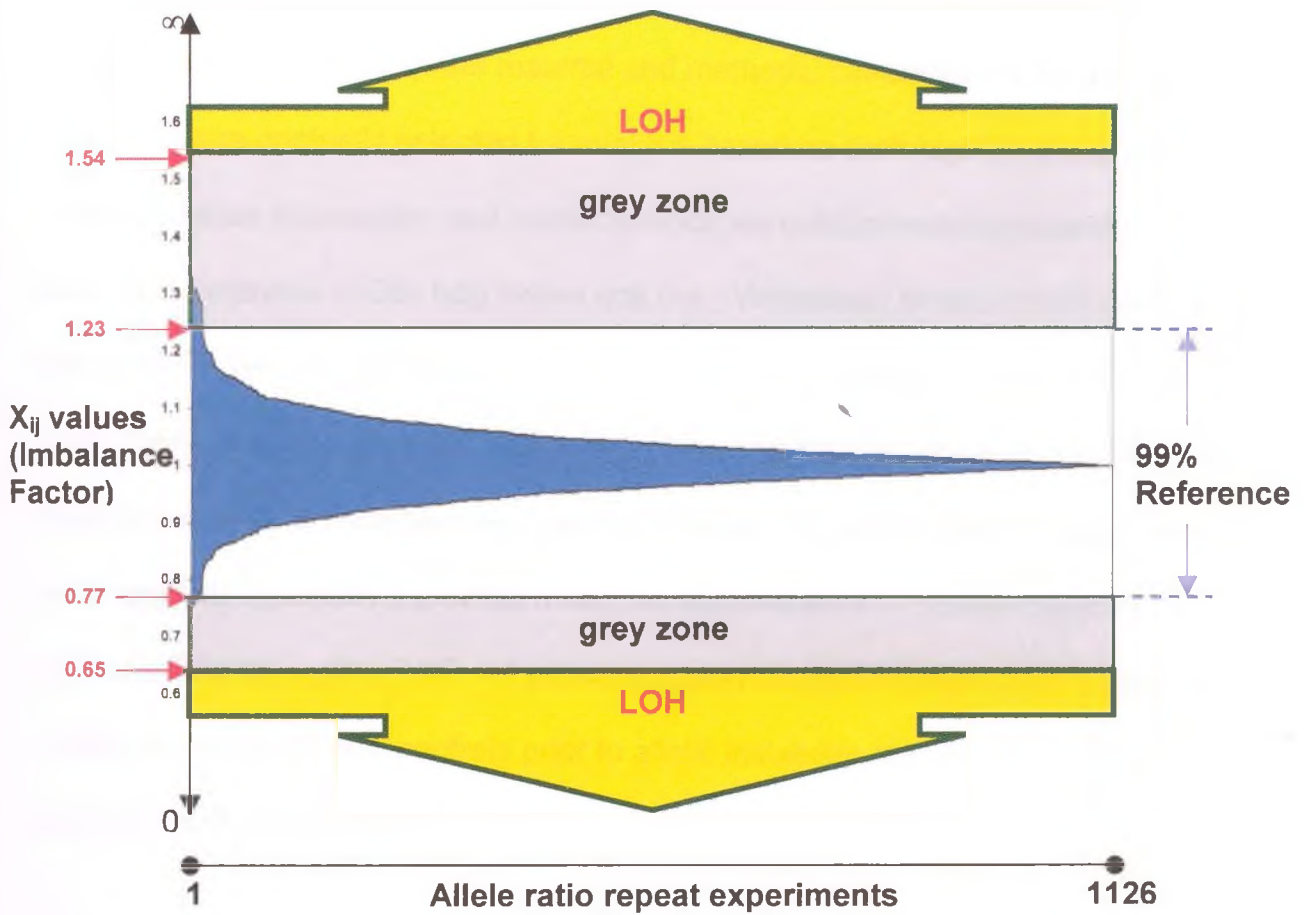


2.4 Determination of sensitivity limits for scoring AI

The sensitivity limits for scoring AI using the fluorescent assay were defined from previous experiments performed at the Roy Castle Centre upon manually micro-dissected tumour/ normal tissue (Liloglou *et al.*, 2000). Briefly reaction variability and reproducibility was assessed by subjecting 24 normal blood DNA samples to 6-plex assays, repeating each sample four times. For each informative sample the four allele ratios ($A_2:A_1$) taken for each marker produced 12 different comparisons which were tabulated (resulting in a table with 1126 entries, half of which are reciprocals). Reference ranges were calculated from these values for scoring allelic imbalance (figure 2.12). In summary samples which produced tumour / normal ratios (Tumour $[A_2:A_1]$ / Normal $[A_2:A_1]$) of $0.77 \leq x \leq 1.23$ were scored as negative (no allelic imbalance). Samples producing ratios of ≥ 1.54 or ≤ 0.65 were scored as positive (allelic imbalance). Samples which produced ratios of 1.23-1.54 or 0.65-0.77 (grey zone) were subjected to a second analysis and scored as imbalance only if a second value is obtained within the previous range for allelic imbalance.

Figure 2.12. Distribution of the X_{ij} values $[(A_2:A_1 \text{ repeat } i) / (A_2:A_1 \text{ repeat } j)]$ produced from repetition experiments on normal DNA samples to assess interassay variation. X_{ij} values represent the false imbalance observed because of the normal reaction variation. Horizontal axis displays the individual 1126 comparisons sorted by their X_{ij} value around 1. The 99% reference range was constructed conservatively as $\text{mean} \pm 3\text{SD}$. Imbalance factor cut-off value for conclusive allelic imbalance was calculated to 23%. A grey zone was introduced to reduce the chance of false positive. Both figure and legend are reproduced with the kind permission of authors from Liloglou *et al*, 2000

Figure 2.12 Distribution of the X_{ij} values produced from repetition experiments on normal DNA samples to assess interassay variation.



2.5 Chromosome 9p and 5q analysis in NSCLC

2.5.1 Deletion mapping of NSCLC tumours using PAGE and silver staining

DNA from lung tumours along with the corresponding control DNA were amplified in a single microsatellite marker PCR reaction and analysed for allelic instability as described in section 2.2 of the material and methods. Chromosome 9p and 5q markers were originally selected for analysis based on their high heterozygosity values. Marker information and characteristics were determined by searching the Genome Database (GDB; <http://www.gdb.org>) Whitehead Institute (WI) Database (<http://www-genome.wi.mit.edu/>) and NCBI Entrez Nucleotide database (<http://www.ncbi.nlm.nih.gov/>) resources. Once selected for analysis microsatellite sequences were downloaded and analysed using the oligo primer design program as described in section 2.2 of the materials and methods to enable maximum primer efficiency during PCR. All primer reaction conditions were optimised on a range of 8 normal DNA controls prior to allelic instability analysis to eliminate non-specific PCR amplification.

2.5.2 Contig assembly and microsatellite marker positioning

Initial searches were carried out to identify yeast artificial chromosomes (YAC) containing microsatellite sequences corresponding to polymorphic repeat sequences on chromosomes 5q and 9p23 using the Whitehead Institute YAC web database (<http://www-genome.wi.mit.edu/>) . BAC sequences also corresponding to microsatellite markers within the same region were searched for using the Sanger Centre ACEDB web database (<http://webace.sanger.ac.uk/cgi-bin/ace/simple/9ace>) to build contig maps containing the deleted region.

2.5.3 YAC culture and DNA extraction

YAC cultures from the CEPH Human Mega YAC library (Chumakov *et al.*, 1992) were obtained from the Human Genome Mapping Project (HGMP) Resource Centre (Cambridge, UK) and individual colonies were streaked onto solid SD medium then incubated overnight at 30°C. Single colonies were picked using a sterile loop then streaked again onto solid SD medium and incubated overnight at 30°C, then single yeast colonies were selected for DNA extraction and grown overnight in 10 ml sterile liquid SD media at 30°C with agitation. Following incubation 1 ml of YAC culture was transferred to a sterile eppendorf tube then centrifuged at 10,000 rpm for 1 minute and the supernatant discarded. YAC DNA was extracted from the remaining pellet using the Yeast Y2 DNA extraction kit (Helena Biosciences, UK) according to manufacturers instructions. DNA was eluted in 50µl sterile distilled water and incubated with 5µg/ml RNase (Roche) for 30 minutes at 37°C. Samples were stored at -20°C.

2.5.4 YAC DNA microsatellite PCR

YAC DNA or alternatively YAC culture (1-2µl) was amplified for single marker microsatellite analysis in a 25µl reaction volume using between 5-10 pg of sense and antisense primer equimolar mix, 0.2 mM dNTP mix, 0.5 Units *Taq* polymerase (Roche) and 1 X *Taq* Polymerase Buffer containing 1.5 mM MgCl₂. PCR consisted of denaturation at 95°C for 5 minutes then 28 cycles of 95°C for 45 sec, 57°C for 45 sec and 72°C for 20 sec followed by an extended annealing step of 72°C for 10 minutes. PCR products were electrophoresed through 3% agarose gels containing 10 mg/ml ethidium bromide and visualised under UV light.

2.5.5 BAC culture and DNA extraction

BAC cultures for the region D9S925-D9S1778 were donated as a kind gift from Dr. Sean Humphrey (Sanger Centre, UK). A sterile loop was used to streak BAC stab cultures onto LB solid media plates, which were incubated overnight at 37°C.

Single colonies from the plates were used to inoculate 5 ml LB cultures which were grown overnight at 37°C with agitation. DNA was extracted by small scale alkaline lysis preparation of plasmid DNA according to a modification of the methods of Birnbiom and Doly (1979) and Ish-Horowicz and Burke (1981). Overnight cultures of colonies (1.5 ml) were centrifuged at 12,000 rpm for 30 seconds in a microfuge and medium removed by aspiration to dry the bacterial pellet as much as possible. The pellet was resuspended in 100 µl of ice-cold solution I by vigorous vortexing. 200 µl of freshly prepared solution II was added and mixed by rapidly inverting the tube five times, then stored on ice. The lysed pellets were gently vortexed in the inverted position for 10 seconds with 150 µl of ice-cold solution III then centrifuged at 12,000 rpm for 5 minutes and the supernatant transferred to a fresh tube.

Following phenol-chloroform extraction, DNA was precipitated in two volumes of absolute ethanol (-20°C), washed in 70% ethanol (-20°C), dried and resuspended in sterile double-distilled water (50 µl). RNA was removed by incubation at 37°C for 30 minutes with 5µg/ml RNase A (Roche). Samples were stored at -20°C until use.

2.5.6 BAC PCR from human artificial chromosome DNA pools

A human bacterial artificial chromosome DNA pool release IV (Research Genetics, AL) was screened for microsatellite sequences using PCR testing. Two microliters each of the initial 'superpool' of 96 PCR reactions was amplified in a 25µl reaction

volume using between 5-10 pg of sense and antisense primer equimolar mix, 0.2 mM dNTP mix, 0.5 Units *Taq* polymerase (Roche) and 1 X *Taq* Polymerase Buffer containing 1.5 mM MgCl₂. PCR consisted of denaturation at 95°C for 5 minutes then 28 cycles of 95°C for 45 sec, 57°C for 45 sec and 72°C for 20 sec followed by an extended annealing step of 72°C for 10 minutes. PCR products were electrophoresed through 3% agarose gels containing 10 mg/ml ethidium bromide and visualised under UV light. The presence of a PCR product determined which plate pool and column to test in a second group of PCR reactions. Positive PCR products in this reaction test revealed the identity of a library plate containing Row and Column pools. A final PCR reaction revealed two positives for each BAC clone containing the microsatellite sequence tested, corresponding to a plate-row-column format (or final clone address). Subsequent positive BAC clones were ordered from Research Genetics and grown in culture as described in section 2.5.5.

2.5.7 BAC PCR

BAC DNA or alternatively BAC culture (1-2µl) was amplified for single marker microsatellite analysis in a 25µl reaction volume using between 5-10 pg of sense and antisense primer equimolar mix, 0.2 mM dNTP mix, 0.5 Units *Taq* polymerase (Roche) and 1 X *Taq* Polymerase Buffer containing 1.5 mM MgCl₂. PCR consisted of denaturation at 95°C for 5 minutes then 28 cycles of 95°C for 45 sec, 57°C for 45 sec and 72°C for 20 sec followed by an extended annealing step of 72°C for 10 minutes. PCR products were electrophoresed through 3% agarose gels containing 10 mg/ml ethidium bromide and visualised under UV light.

2.5.8 Candidate Gene and EST identification

Candidate genes and EST's (Expressed Sequence Tags) within the region D9S254-D9S161 were investigated by searching the NCBI Genemap '99 database (<http://www.ncbi.nlm.nih.gov/genemap/map.cgi?CHR=9>) using the microsatellite marker identities as a search parameter.

2.5.9 Sequence searching, alignment and BLAST sequence comparison

Gene sequences were searched for by name using the NCBI Entrez Nucleotide database (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Nucleotide>) and saved in FASTA format for BLAST analysis. Sequences were checked for isoforms and complementary gene sequences by performing a basic BLAST search of the FASTA sequence format using the NCBI basic *blastn* program and *nr* database (<http://www.ncbi.nlm.nih.gov/blast/blast.cgi?Jform=0>). Sequence alignment for intron/exon boundaries and microsatellite sequences was determined by comparing two FASTA sequences using the NCBI *Blast2* database (<http://www.ncbi.nlm.nih.gov/gorf/bl2.html>; Tatusova & Madden, 1999).

2.5.10 NIX analysis

Clone sequences were searched for by name using the NCBI Entrez Nucleotide database (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Nucleotide>) and saved in FASTA format for NIX analysis. Genomic clone DNA sequences were split previously to sequence submission, into 50kb sections with 200 bp overlaps for NIX analysis using the HGMP resources NIX database (<http://www.hgmp.mrc.ac.uk/Registered/Webapp/nix/>).

2.5.11 Single Strand Conformation Polymorphism (SSCP) analysis of Lung DNA.

DNA (1-2 μ l) was amplified for single marker SSCP analysis in a 25 μ l reaction volume using between 5-10 pg of sense and antisense primer equimolar mix, 0.2 mm dNTP mix, 0.5 units *Taq* polymerase (Roche) and 1 x *Taq* polymerase buffer containing 1.5 mm MgCl₂. PCR consisted of denaturation at 95°C for 5 minutes then 28 cycles of 95°C for 45 sec, 57°C for 45 sec and 72°C for 20 sec followed by an extended annealing step of 72°C for 10 minutes (slight variations to the T_m and annealing steps of these cycles were made depending on the microsatellite markers used). Prior to loading on acrylamide gels, 5 μ l of 1 x SSCP denaturing buffer was added to 25 μ l of PCR product, then denatured at 95°C for 10 minutes. Denatured samples were rapidly cooled on ice for 1-2 minutes then 6 μ l of PCR products were electrophoresed through 10% SSCP polyacrylamide gels and visualised by silver staining.

2.5.12 Sequence analysis of MMLT3

Abnormal SSCP products were reamplified using the original primer in a 25 μ l reaction volume using between 5-10 pg of sense and antisense primer equimolar mix, 0.2 mm dNTP mix, 0.5 units *Taq* polymerase (Roche) and 1 x *Taq* polymerase buffer containing 1.5 mm MgCl₂. PCR consisted of denaturation at 95°C for 5 minutes then 28 cycles of 95°C for 45 sec, 57°C for 45 sec and 72°C for 20 sec followed by an extended annealing step of 72°C for 10 minutes. 5 μ l of PCR product was electrophoresed through 3% agarose gels stained with ethidium bromide (10 mg/ml) to confirm amplification of DNA. The remaining PCR product was cleaned to remove excess dNTP's and unbound primers using the Wizard

DNA Clean-Up kit (Qiagen, UK) according to manufacturers instructions and resuspended in 10 μ l of sterile distilled water. DNA (3-10 ng) was sequenced in both the sense and antisense direction in a 10 μ l reaction volume containing 10 pg primer, 0.75 μ l Big Dye™ reaction mix and 1 x Big Dye sequencing buffer. PCR consisted of denaturation at 95°C for 60 seconds then 25 cycles of 96°C for 10 seconds, 55°C for 10 seconds, 70°C for 30 seconds and 60°C for 4 minutes. Sequencing products were resuspended in 60 μ l of 70% ethanol at room temperature for 30 minutes. Ethanol was removed and the pellet dried under vacuum for 5 minutes then mixed with 15 μ l of template suppression agent (Applied Biosystems). Prior to loading samples were heated at 95°C for 2 minutes and rapidly cooled on ice then analysed on an ABI-PRISM 310™ automatic sequencer (PE-Applied Biosystems, Warrington, UK). Results were analysed using Chromas™ sequence alignment software (v1.43).

2.6 Statistical analysis of data

AI data was analysed for correlations with clinicopathological data using Pearson's Rank correlation and two tailed t-test. Comparison of means for FAL and FRL data was employed using independent samples t-test or paired-samples t-test where stated. SPSS for Windows™ (v10.0.7) was used for statistical analysis.

**CHAPTER 3: HIGH THROUGHPUT FLUORESCENT
MICROSATELLITE ANALYSIS OF CHROMOSOME 17 IN HUMAN
NON-SMALL CELL LUNG TUMOURS**

3.1 Introduction

PCR based microsatellite analysis of normal versus tumour DNA is a common method of measuring allelic imbalance (AI) in multiple tumour types. Because of their abundance and high frequency within the human genome, microsatellites are ideal markers for the analysis of allelic imbalance. Allelotype and deletion mapping studies have used AI as a tool to locate small chromosomal regions in which putative tumour suppressor genes exist through a means of identifying chromosomal regions present in the DNA of normal tissue but deleted or amplified within the DNA from tumour cells (Kastury *et al.*, 1996; Lundberg *et al.*, 1987; section 1.4.5). AI of chromosome 17 is a frequent and early event in the development of NSCLC (Tsuchiya *et al.*, 1992; Wistuba *et al.*, 1997a), and previous studies have relied on single marker microsatellite analysis to detect AI on chromosome 17 in NSCLC (Fong *et al.*, 1995a; Fontanini *et al.*, 1993; Froudarakis *et al.*, 1998). Fluorescent microsatellite analysis differs from conventional microsatellite analysis using ^{32}P or silver staining, as it allows accurate determination of AI through better band resolution and sizing in addition to more accurate quantification of band intensity through the attached fluorophore. High throughput analysis can also be achieved due to multiplex PCR of multiple microsatellite loci from the same DNA sample using different fluorescent labels. This chapter focuses on the development of a fluorescent microsatellite multiplex approach used to analyse chromosome 17 allelic imbalance in a panel of NSCLC tumours based on a previous fluorescent microsatellite analysis sensitivities and limits devised by Liloglou *et al.*, (2000).

3.2 Fluorescent multiplex analysis of chromosome 17 in NSCLC.

3.2.1 Methods and sample selection

Tumour and normal DNA was isolated from the resected lung tissue of 64 lung cancer patients as described in the materials and methods section.

Microsatellite markers for NSCLC analysis were selected from the Linkage Mapping Set V2.0, chromosome 17 high density panel (Applied Biosystems, Warrington, UK). DNA was amplified using multiplex PCR as described in the material and methods section and analysed on an ABI-PRISM 377 automatic sequencer (Applied Biosystems, Warrington, UK). Results were analysed using the Genescan™ and Genotyper™ software (PE-Applied Biosystems, Warrington, UK). AI was scored as described in chapter 2 and in Liloglou *et al*, (2000).

3.2.2 Chromosome 17 marker selection

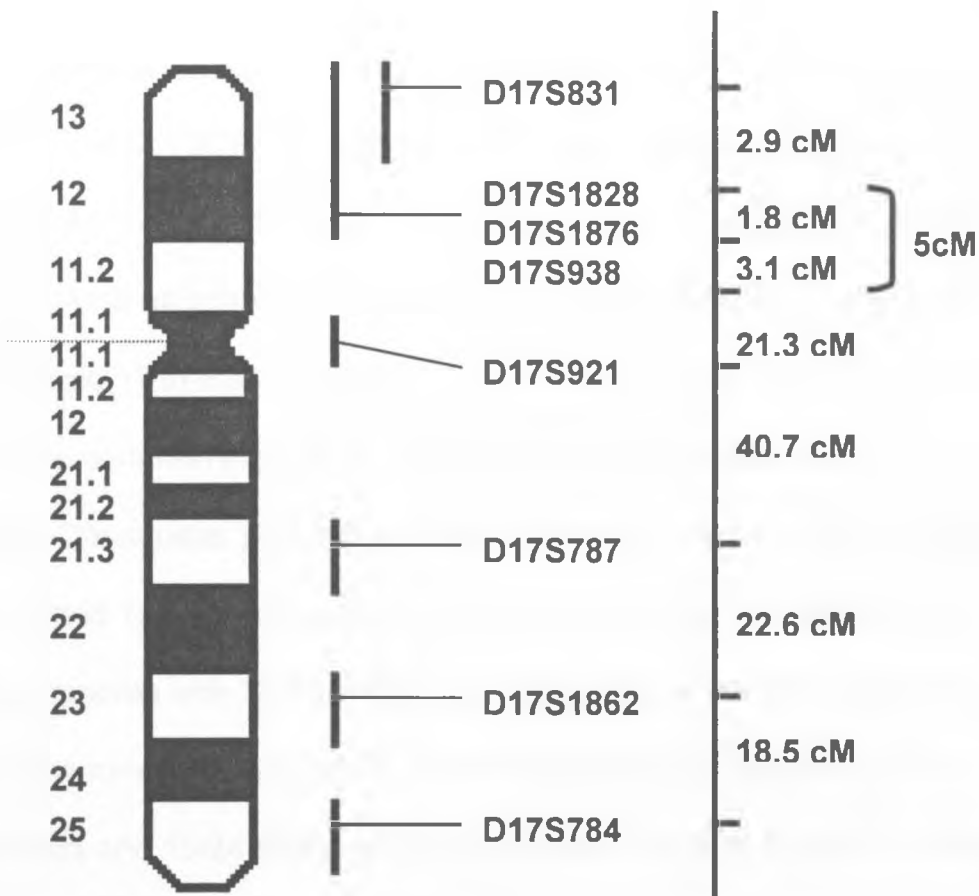
Markers were selected for analysis based on their position on chromosome 17, high heterozygosity, allele size range and fluorescent label. Eight microsatellite markers were from the Linkage Mapping Set V2.0 high density panel derived from chromosome 17 were chosen for analysis. These are displayed in table 3.1. Markers were selected to represent chromosome 17 on both arms. Figure 3.1 displays the position of selected microsatellite markers on chromosome 17. Reaction conditions for the panel were initially optimised by amplification of DNA from 8 paired normal and tumour samples then analysed on the ABI-377 automated sequencer using Genescan™ and Genotyper™ software. Primer concentrations within the multiplex reaction were modified accordingly to reduce non-specific products and equalise signal intensity from markers in the chromosome 17 multiplex panel. Fluorescent markers were also divided into

two smaller chromosome 17 panel sub-sets containing large and small allele sizes for PCR amplification to increase reaction specificity. After precipitation to remove excess dNTP's and non-incorporated primers prior to analysis, PCR products from the subset reactions were remixed and analysed on the ABI-377 automated sequencer.

Table 3.1. Table displaying markers used in fluorescent multiplex panel for chromosome 17 analysis. Marker information was obtained from ABI PRISM® Linkage Mapping Set-HD5 web page (<http://www.appliedbiosystems.com/molecularbiology/apply/dr/lmshd5/chrom17.html>).

Marker Name	Cytogenetic location	Average Heterozygosity	Allele Size Range (bp)	Fluorescent Label
D17S831	17p13	0.82	112-134	HEX
D17S938	17p12-p13	0.76	242-262	FAM
D17S1876	17p12-p13	0.82	104-136	FAM
D17S1828	17p12-p13	0.79	101-121	NED
D17S921	17cen	0.72	196-214	HEX
D17S787	17q21.3	0.81	142-178	NED
D17S1862	17q23	0.83	205-235	NED
D17S784	17q25	0.77	233-247	HEX

Figure 3.1 Diagrammatic representation of the current cytogenetic position of microsatellite markers on chromosome 17 used in fluorescent panel. Distances between markers are displayed at the right of the page in centimorgans.



3.2.3 Analysis of chromosome 17 in NSCLC

After multiplex conditions were optimised, a panel of 64 paired NSCLC normal and tumour samples were analysed using 8 fluorescent markers for AI on chromosome 17 (Table 3.2). Examples of allelic imbalance, retention of heterozygosity and microsatellite instability are displayed in figure 3.2. Fifty-nine out of 64 (92%) tumours displayed AI at one or more loci on chromosome 17. The highest frequencies of AI were detected at D17S831 (79%) and D17S1828 (75%) located at 17p12-p13. The lowest frequencies of AI were detected at D17S784 and D17S787, located at 17q25 and 17q21.3 respectively, which displayed AI in 46% of tumours (figure 3.3). Fifteen tumours (15/64, 23%) displayed AI at all informative loci indicating either deletion or amplification of chromosome 17 in these tumours. AI at D17S1862 and D17S1828 was found to be more common in squamous cell carcinoma than adenocarcinoma (Pearson Chi-square, $p=0.048$ and $p=0.022$ respectively). Allelic instability on both 17p and 17q was more frequent in squamous cell carcinoma than adenocarcinoma with 35/39 (89%) squamous cell carcinomas and 20/25 (80%) adenocarcinomas showing AI on 17p and 28/39 (72%) squamous cell carcinomas and 15/24 (63%) adenocarcinomas showing AI on 17q. Fractional regional loss for chromosome 17 (FRL17) values were calculated for each tumour as $[\text{total informative loci on chromosome 17 scored with AI}] / [\text{total informative chromosome 17 loci tested}]$. In addition, FRL values were also calculated for each chromosome 17 arm. FRL17p was calculated as $[\text{loci on the 17p scored with AI}] / [\text{total informative loci on 17p}]$ and FRL17q as $[\text{loci on the 17q scored with AI}] / [\text{total informative loci on 17q}]$. The mean FRL17p (0.721 ± 0.046) was greater than the mean FRL17q (0.542 ± 0.054) indicating a greater frequency of deletion of markers on chromosome arm 17p when compared to

17q (Paired T-test, $p=0.003$). The mean FRL17 was greater in squamous cell carcinomas than adenocarcinomas (0.711 ± 0.045 and 0.570 ± 0.069 respectively) however this was not found to be significant ($p=0.080$). This was also reflected in the FRL17p (0.776 ± 0.055 and 0.647 ± 0.078 respectively, $p=0.169$) and FRL17q (0.611 ± 0.069 and 0.431 ± 0.083 respectively, $p=0.103$). Figure 3.4 shows a comparison of mean FRL for 17p and 17q between squamous cell carcinomas and adenocarcinomas. No significant differences were observed between allelic imbalance on chromosome 17, FRL17, FRL17p, FRL17q and clinicopathological parameters such as morphology, differentiation, TNM stage and smoking.

Table 3.2. Allelic imbalance results from 64 NSCLC tumours analysed with 8 microsatellite markers on chromosome 17. The first column displays the sample code and the second column morphology (SqCCL, squamous cell carcinoma of the lung; Adeno, adenocarcinoma). Column 3-10 represents AI results from each microsatellite marker. Microsatellites are arranged in the table from left to right depending on their position from 17pter-17qter. 17cen is located between the microsatellites D17S921 and D17S787. Black cells with AI, allelic imbalance; white cells with H, heterozygous (no allelic imbalance); red cells with MI, microsatellite instability; blank cells, not determined or homozygous. Column 11 represents fractional regional loss (FRL 17). Column 12 represents the fractional regional loss on chromosome 17p (FRL 17p). Column 13 represents the fractional regional AI on chromosome 17q (FRL 17q)]. The FRL 17q for sample L030 could not be determined (nd) as all markers on 17q were homozygous for this tumour. Table was ordered according to FRL17 values.

Table 3.2. Allelic imbalance results from 64 NSCLC tumours analysed with 8 microsatellite markers on chromosome 17.

Sample #	Morphology	D17S831	D17S1828	D17S1876	D17S938	D17S921	D17S787	D17S1862	D17S784	FRL 17	FRL 17p	FRL 17q
L012	SqCCL	AI	AI	AI		AI	AI		AI	1.000	1.000	1.000
L041	SqCCL		AI	AI	AI	AI	AI	AI	AI	1.000	1.000	1.000
L043	SqCCL	AI	AI	AI	AI	AI	AI	AI	AI	1.000	1.000	1.000
L093	SqCCL	AI	AI	AI	AI	AI	AI	AI		1.000	1.000	1.000
L112	SqCCL	AI	AI	AI	AI	AI	AI	AI		1.000	1.000	1.000
L127	SqCCL	AI	AI	AI		AI	AI	AI	AI	1.000	1.000	1.000
L129	SqCCL	AI	AI	AI	AI			AI	AI	1.000	1.000	1.000
L146	SqCCL	AI	AI	AI	AI	AI	AI	AI	AI	1.000	1.000	1.000
L153	SqCCL	AI	AI				AI	AI	AI	1.000	1.000	1.000
L136	SqCCL	AI	AI	AI	AI			AI	AI	1.000	1.000	1.000
L164	SqCCL	AI			AI	AI		AI	AI	1.000	1.000	1.000
L171	SqCCL	AI	AI	AI	AI	AI	AI	AI		1.000	1.000	1.000
L167	SqCCL	AI	AI	AI	AI	AI	AI	AI	H	0.875	1.000	0.667
L107	SqCCL	AI	AI		AI	H	AI	AI	AI	0.857	1.000	1.000
L144	SqCCL		AI	AI	AI	AI	H	AI	AI	0.857	1.000	0.667
L179	SqCCL	H	AI	AI	AI	AI		AI	AI	0.857	0.750	1.000
L007	SqCCL	AI	AI	AI		AI	H	AI	MI	0.833	1.000	0.500
L126	SqCCL	AI	AI	H	AI			AI	AI	0.833	0.750	1.000
L150	SqCCL		H	AI		AI		AI	AI	0.800	0.500	1.000
L014	SqCCL	AI	AI	AI	AI		H	H	AI	0.714	1.000	0.333
L044	SqCCL	AI	AI	AI	AI	AI		H	H	0.714	1.000	0.000
L145	SqCCL	AI	AI	AI	AI	AI	H		H	0.714	1.000	0.000
L154	SqCCL	AI	AI	AI	AI	AI	H	H		0.714	1.000	0.000
L163	SqCCL	AI	AI		H	AI	AI	AI	H	0.714	0.667	0.667
L048	SqCCL	AI	AI	AI	H	H		AI		0.667	0.750	1.000
L147	SqCCL	AI	AI			H	AI	AI	H	0.667	1.000	0.667
L151	SqCCL	AI	AI	AI	AI	AI	H	H	H	0.625	1.000	0.000

Sample #	Morphology	D17S831	D17S1828	D17S1876	D17S938	D17S921	D17S787	D17S1862	D17S784	FRL 17	FRL 17p	FRL 17q
L177	SqCCL	AI		H	AI	AI	H	H	AI	0.571	0.667	0.333
L066	SqCCL	H	AI	AI		AI		H	H	0.500	0.667	0.000
L096	SqCCL				H	H	AI	AI		0.500	0.000	1.000
L152	SqCCL	AI	AI	H	H	H	H	AI	AI	0.500	0.500	0.667
L172	SqCCL	AI		AI		AI	H	H	H	0.500	1.000	0.000
L037	SqCCL	AI		AI			H	H	H	0.400	1.000	0.000
L027	SqCCL	AI			H				H	0.333	0.500	0.000
L165	SqCCL	H		H		H	AI	AI	H	0.333	0.000	0.667
L182	SqCCL	H	H	H	H		H	AI	AI	0.286	0.000	0.667
L180	SqCCL	AI	H	H	H			H		0.200	0.250	0.000
L138	SqCCL	H	AI	H	H		H	H	H	0.143	0.250	0.000
L033	SqCCL		H	H	H	H	H	H	H	0.000	0.000	0.000
L030	Adenocarcinoma	AI	AI		AI	AI				1.000	1.000	nd
L113	Adenocarcinoma	AI	AI	AI		AI	AI	AI		1.000	1.000	1.000
L143	Adenocarcinoma	AI	AI	AI	AI	AI	AI	AI	AI	1.000	1.000	1.000
L158	Adenocarcinoma	AI	AI	AI	H	AI	AI	AI	AI	0.875	0.750	1.000
L045	Adenocarcinoma	AI	H	AI	AI	AI		AI	AI	0.857	0.750	1.000
L086	Adenocarcinoma	AI	AI		AI	AI	AI	AI	H	0.857	1.000	0.667
L159	Adenocarcinoma	AI	AI	AI	AI	AI		H	AI	0.857	1.000	0.500
L087	Adenocarcinoma	AI	AI	AI	AI			AI	H	0.833	1.000	0.500
L123	Adenocarcinoma	AI	H	AI	AI			AI	AI	0.833	0.750	1.000
L091	Adenocarcinoma		AI	H	AI			AI		0.750	0.667	1.000
L049	Adenocarcinoma	AI	AI	AI	AI	AI	H		H	0.714	1.000	0.000
L124	Adenocarcinoma	AI	AI		AI	AI	H	H	AI	0.714	1.000	0.333
L047	Adenocarcinoma	AI	AI	H	AI		AI	H		0.667	0.750	0.500
L156	Adenocarcinoma	AI	AI	H	AI	H	AI	AI	H	0.625	0.750	0.667
L005	Adenocarcinoma	AI		AI			H	AI	H	0.600	1.000	0.333
L095	Adenocarcinoma	AI	AI	AI				H	H	0.600	1.000	0.000
L031	Adenocarcinoma	AI	H	AI		AI		H	H	0.500	0.667	0.000

Sample #	Morphology	D17S831	D17S1828	D17S1876	D17S938	D17S921	D17S787	D17S1862	D17S784	FRL 17	FRL 17p	FRL 17q
L157	Adenocarcinoma	H	AI		H	H	AI	H		0.333	0.333	0.500
L063	Adenocarcinoma	H	H	AI	H	AI	H	H	H	0.250	0.250	0.000
L155	Adenocarcinoma	AI	H	H	AI	H	H	H	H	0.250	0.500	0.000
L058	Adenocarcinoma	H	H	H	H		H	H	AI	0.143	0.000	0.333
L026	Adenocarcinoma		H	H	H		H	H	H	0.000	0.000	0.000
L040	Adenocarcinoma	H	H	H	H	H	H	H	H	0.000	0.000	0.000
L075	Adenocarcinoma	H		H	H	H		H	H	0.000	0.000	0.000
L106	Adenocarcinoma	H	H	H	H	H	H	H	H	0.000	0.000	0.000

Figure 3.2 Examples of heterozygosity, homozygosity, allelic imbalance, and microsatellite instability on chromosome 17. Screen captures from Genotyper™ demonstrating retention of heterozygosity in L026 at D17S1828, homozygosity in L027 at D17S1828 and allelic imbalance in L007, L012 and L014 at D17S1828. Microsatellite instability for L007 at D17S784 is also shown. Key to figure is as follows: N, normal DNA; T, tumour DNA; black arrow, indicates reduction of allele intensity; red arrow, indicates additional allele caused by microsatellite alteration.

Figure 3.2 Examples of heterozygosity, homozygosity, allelic imbalance, and microsatellite instability on chromosome 17

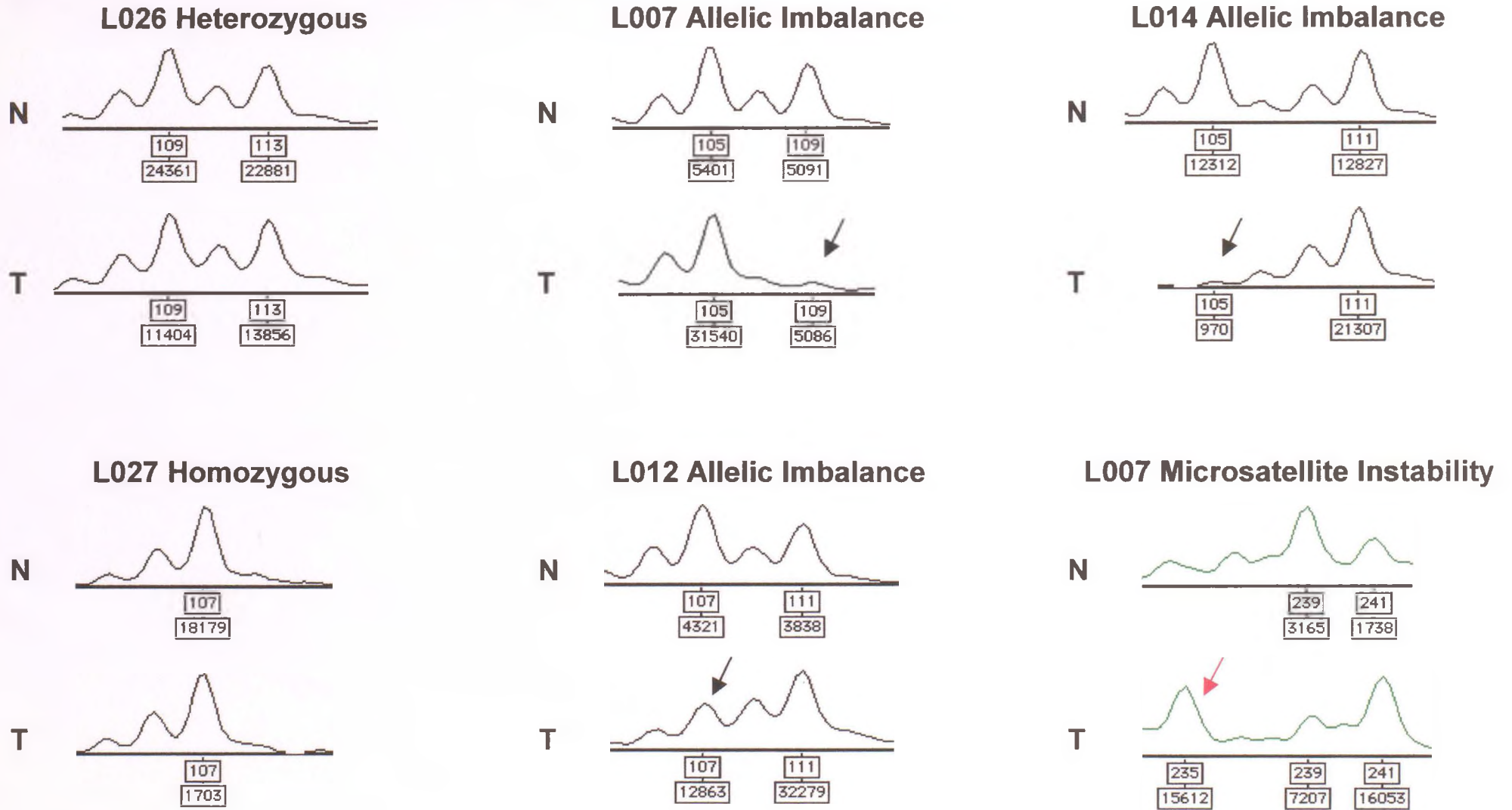


Figure 3.3 Histogram displaying the frequency of AI detected for each microsatellite marker on chromosome 17 in 64 NSCLC tumours. X-axis displays the microsatellite loci tested from 17pter – 17pcen – 17qter. Y-axis displays the frequency of AI detected for each loci calculated by $[\text{total number of samples that displayed AI}] / [\text{total number of informative samples}]$.

Figure 3.3 Histogram displaying the frequency of AI detected for each microsatellite marker on chromosome 17 in 64 NSCLC tumours.

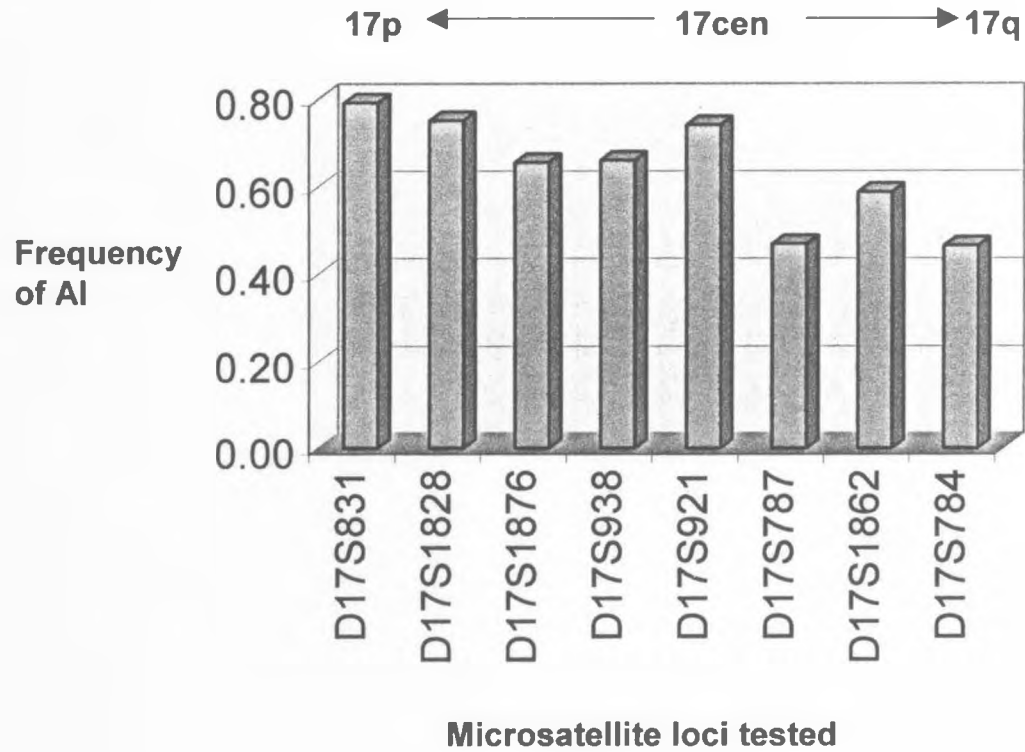
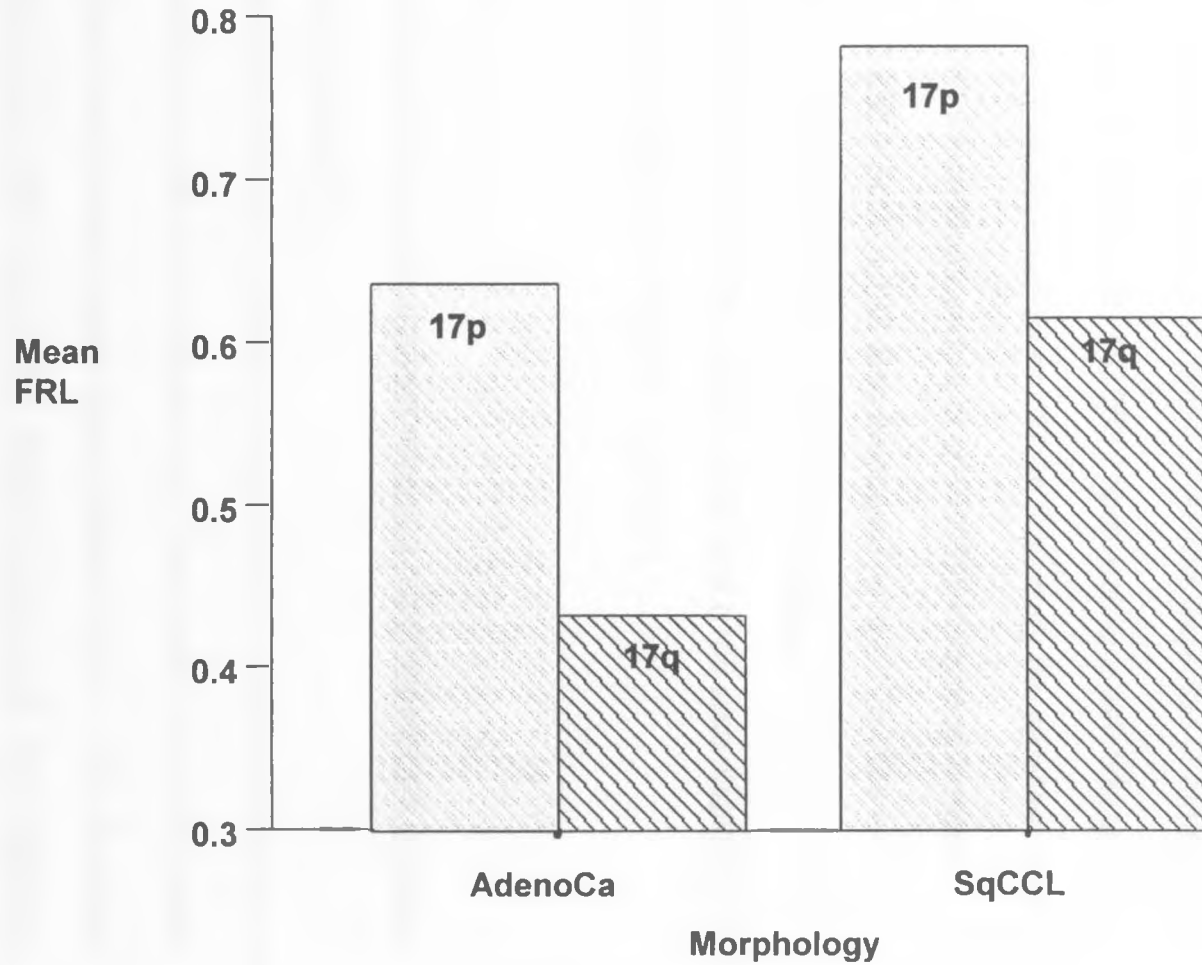


Figure 3.4. Histogram comparing the mean FRL between adenocarcinoma and squamous cell carcinoma of the lung for 17p and 17q.

Morphology is displayed on the x-axis (AdenoCa, adenocarcinoma; SqCCL, squamous cell carcinoma of the lung). FRL17p is represented by grey bars with white hatching and FRL17q by white bars with black hatching. The mean FRL for 17p and 17q is displayed on the y-axis.

Figure 3.4. Histogram comparing the mean FRL between adenocarcinoma and squamous cell carcinoma of the lung for 17p and 17q.



3.2.4 Multiple regions of allelic imbalance of chromosome 17q

Multiple regions of AI were detected on chromosome 17 which ranged from large regions involving almost all of chromosome 17 (L167) to smaller regions involving one cytogenetic band or marker (e.g. L138 and L063). Nine tumours displayed one minimal region of AI and three tumours displayed two regions of AI on chromosome 17. The smallest region of AI was detected in tumour L063 which extended approximately 5cM from D17S1828 – D17S938 around the loci D17S1876 at 17p12-p13. The largest region of AI was detected in tumour L167 which extended from D17S831 – D17S1862, representing a deletion or amplification of nearly all of chromosome 17. In this study the smallest region of AI was found in L063 which displayed AI at D17S1876 extending between D17S1828 – D17S938 (<5cM). L138 also displayed AI at D17S1828 extending between D17S831 – D17S1876 (<5cm) and L157 displayed AI at D17S1828 with no AI at D17S831 and D9S938 (<8cM). L155 and L156 also displayed AI at D17S938 extending between D17S1876 – D17S921.

3.3 Discussion

3.3.1 High throughput fluorescent analysis revealed high frequencies of allelic instability on chromosome 17.

Allelic instability (loss of heterozygosity) of chromosome 17 occurs in tumours of the breast, ovaries, liver, stomach, prostate, lung, head and neck and Barrett's oesophagus in addition to other tumour types (Adamson *et al.*, 1994; Eccles *et al.*, 1990; Fujimori *et al.*, 1991; Kato *et al.*, 1996; Mackay *et al.*, 1988; Macoska *et al.*, 1992; Mori *et al.*, 1994; Russell *et al.*, 1990; Sano *et al.*, 1991; Shibagaki *et al.*, 1994; Shimizu *et al.*, 1993; Swift *et al.*, 1995; Dunn *et al.*, 1999). The high frequency of AI in these tumour types has implicated chromosome 17 as

harbouring gene(s) important in the development of multiple cancers. This work has described the analysis of chromosome 17 using an accurate high throughput method to detect AI on both chromosome arms. The high frequency of AI detected is in keeping with other studies implicating chromosome 17 as a frequent target for deletion in NSCLC tumours (Fong *et al.*, 1995a; Girard *et al.*, 2000; Konishi *et al.*, 1998; Liloglou *et al.*, 2000; Mao *et al.*, 1997; Tsuchiya *et al.*, 2000; Weston *et al.*, 1989; Yoshioka *et al.*, 1998). Fong *et al.* (1995) detected allelic loss on chromosome 17 at one or more of six loci in 58% of NSCLC tumours with loss occurring at a frequency of 54% on 17p and 42% on 17q. Using highly sensitive methods this study has demonstrated that the frequency of AI in NSCLC was much greater than that detected in Fong *et al.*, as 92% of NSCLC tumours displayed AI at one or more loci on chromosome 17. The frequency of AI on each chromosome arm was also greater, with 53/62 (85%) of tumours displaying AI on 17p and 42/62 (68%) displaying AI on 17q. The increase in frequency of AI detected in this study may reflect the selection of different microsatellite loci for analysis of lung tumours or the analysis of a different tumour population. However, it is more likely that the increased sensitivity and a lower cut-off value used for scoring AI in this study has allowed greater detection of AI on chromosome 17 in NSCLC tumour cells, when compared to traditional methods such as PAGE followed by silver staining or radiolabelling of PCR products for AI analysis. This study has shown that chromosome 17 AI was more common in squamous cell carcinoma than adenocarcinomas on both chromosome arms. In addition the mean FAL for each chromosome 17 arm also displayed differences between these two histological types, although this difference was not statistically significant. Previous studies have compared the frequency of AI or allelic loss on

chromosome 17 between squamous cell carcinomas of the lung and adenocarcinomas (Sato *et al.*, 1994; Tsuchiya *et al.*, 1992). A comparison between the AI frequencies detected in this study and previous studies, and also between squamous cell carcinomas and adenocarcinomas is shown in table 3.3. As previously mentioned, this study demonstrated a high frequency of AI on both arms of chromosome 17 when compared to previous publications. However this study has provided further evidence that 17p is a more frequent target for AI or deletion than 17q in NSCLC. The mean FRL17p was also significantly greater than the mean FRL17q. In agreement with previous work, this study also detected a higher frequency of 17p AI in squamous cell carcinomas than adenocarcinomas. In addition to providing strong evidence that the frequency of AI on 17q is also greater in squamous cell carcinomas, as reflected in the differences between the mean FRL for both chromosome arms (figure 3.4). In comparison with this study, Girard *et al* (2000) recently completed a genome wide allelotyping of NSCLC lung cancer cell lines and detected a >60% frequency of allelic loss with 7 markers on chromosome 17p using the ABI-377 automated genotyper. This group also detected a high frequency of deletion at the microsatellite loci D17S831 and D17S1828 at 17p12-p13 (79% and 75%) and deletion of the 17p12-p13 region in 86% of tumours suggesting that this chromosomal region may play an important role in the development of NSCLC. The result presented in this chapter and in Girard *et al* shows that the increased sensitivity afforded by AI analysis using the ABI-377 automated genotyper greatly enhances AI detection in NSCLC.

Table 3.3. Frequency of AI detected on chromosome 17 in this and previous studies and a comparison between histological sub-types in NSCLC. References and the year of publication are displayed in the first column. The second column displays the method of analysis used (RL, radiolabelling; SS, silver staining; FL, fluorescent labelling). The third and fourth columns display the overall frequency of AI detected on chromosome arms 17p and 17q respectively and also the number of microsatellite markers used for AI analysis on each chromosome arm. The fifth and sixth columns display the frequency of AI detected on 17p in SqCCL and AdenoCa respectively. The seventh and eighth columns display the frequency of AI detected on 17q in SqCCL and AdenoCa respectively. The frequency of AI for 17q between the two histological types could not be determined (nd) from Tsuchiya *et al*, 1992.

Table 3.3. Frequency of AI detected on chromosome 17 in this study and previous studies and a comparison between histological sub-types in NSCLC.

Study	Methods	Total loss on 17p [# markers used]	Total loss on 17q [# markers used]	SqCCL 17p loss	AdenoCa 17p loss	SqCCL 17q loss	AdenoCa 17q loss
Tsuchiya <i>et al</i> , 1992	RL	23/37 (62%) [1]	8/39 (21%) [1]	9/9 (100%)	11/24 (46%)	nd	nd
Sato <i>et al</i> , 1994	RL	62/104 (58%) [1]	14/84 (17%) [1]	21/24 (88%)	41/80 (51%)	~21%	~15%
Neville <i>et al</i> , 1994	SS	17/45 (38%) [5]	8/45 (18%) [4]	6/15 (40%)	5/18 (28%)	1/15 (7%)	7/18 (39%)
This study	FL	53/62 (85%) [5]	42/62 (68%) [3]	35/39 (89%)	20/25 (80%)	28/39 (72%)	15/24 (63%)

3.3.2 Identification of three regions of AI on chromosome 17p provides further evidence for the involvement of 17p12-p13 in NSCLC

AI analysis revealed multiple regions of AI on the small arm of chromosome 17 in NSCLC. The long arm of chromosome 17 displayed several regions, however these involved all or most of the microsatellite loci on 17q, reflecting amplification or deletion of the whole chromosome arm. Five tumours displayed smaller regions of AI on 17p, which were centred around 3 microsatellite loci. The first locus was D17S938, the second locus was D17S1828 and the third locus was D17S1876. All three microsatellite loci have been mapped to 17p12-p13 on the short arm of chromosome 17 within a 5cM distance. Previous studies have mapped minimal deleted regions to 17p12-p13 in NSCLC and other tumour types (Kato *et al.*, 1996; Konishi *et al.*, 1998; Mackay *et al.*, 1988; Schultz *et al.*, 1996; Tsao *et al.*, 1991; Tsuchiya *et al.*, 2000). The p53 gene is located at 17p13.1 and abnormalities such as allelic loss at 17p13.1, gene deletions, mutations and aberrant protein expression have been well documented in NSCLC (Fong *et al.*, 1995a; Kishimoto *et al.*, 1992; Liloglou *et al.*, 1997; Marchetti *et al.*, 1993; Passlick *et al.*, 1994; Tammemagi *et al.*, 2000; Wistuba *et al.*, 1999c). However deletion mapping studies have demonstrated that a site distal to the p53 gene may contain a gene involved in the pathogenesis of multiple tumour types including NSCLC (Konishi *et al.*, 1998; McHale *et al.*, 2000; Schultz *et al.*, 1996; Tsuchiya *et al.*, 2000; Wales *et al.*, 1995). Wales *et al.* (1995) identified HIC – 1 (hypermethylated in cancer), a putative candidate gene located at 17p13.3. HIC – 1 is a zinc-finger transcription factor gene which is underexpressed in tumour cells when hypermethylated. This gene also has a p53 binding site in the 5' flanking region and expression is activated by wild type p53 (Wales *et al.*, 1995; Guerardel *et*

al., 2000). Eguchi *et al* (1997) investigated the methylation status of the HIC-1 gene (located at the D17S5 loci) in 51 lung normal and paired tumour tissues and detected DNA hypermethylation at D17S5 in 33% of lung tumours and in 31% of non tumourous tissues. Hypermethylation correlated with differentiation and allelic loss at this loci and was greater in tissue from smokers, suggesting that this mechanism may play a role in the development of NSCLC in cigarette smokers (Eguchi *et al.*, 1997). Schultz *et al* (1996) mapped two candidate genes (OVCA1 and OVCA2) to 17p13.3 whose mRNA expression levels were significantly reduced in ovarian tumours and tumour cell lines (Schultz *et al.*, 1996), however no evidence could be found of abnormalities of these two genes in NSCLC. Tsuchiya *et al* (2000) detected three minimal regions of deletion at 17p13.3 distal to the p53 gene centred around the loci ENO - D17S1566, D17S379 to D17S1574 and distal to ABR with frequencies of allelic loss similar to that of the p53 region (Tsuchiya *et al.*, 2000). Although the results presented in this chapter provide further evidence for the involvement of gene(s) within the 17p12-p13 region of chromosome 17 which may be involved in NSCLC, the regions of AI detected in this study were large (5-15 cM), thus high density mapping is required to target this region.

3.4 Conclusion

This chapter has described the selection of a high throughput chromosome 17 microsatellite panel for accurate AI detection in NSCLC. A high frequency of AI was detected in a panel of 64 NSCLC tumours on both 17p and 17q. The frequency of AI on chromosome 17p was significantly higher than 17q and also higher in squamous cell carcinomas when compared to adenocarcinomas. This was also reflected in the differences in FAL values between both chromosome

arms and also each morphological subset. The frequency of AI detected on both chromosome 17 arms was also significantly higher than previously documented studies in NSCLC. Several regions of AI were detected on chromosome 17 including three minimal regions located at 17p12-p13 within a genetic distance of 5cM . This study therefore provides further evidence of the presence of putative tumour suppressor gene(s) on this arm which may be involved in the pathogenesis of NSCLC.

CHAPTER 4: MOLECULAR ANALYSIS OF RESPIRATORY EPITHELIUM IN TUMOUR-BEARING LUNGS

4.1 Introduction

Squamous carcinoma is thought to proceed through various hyperplastic, metaplastic and dysplastic stages prior to development of a carcinoma and, in most cases, squamous metaplasia / dysplasia can occur at multiple sites throughout the respiratory tract with intervening 'normal' areas (Saccomanno *et al.*, 1974). Previous studies have identified specific chromosomes such as 3p, 9p and 17q as frequent targets for deletion in NSCLC (Field *et al.*, 1996; Girard *et al.*, 2000; Hung *et al.*, 1995; Neville *et al.*, 1995; Neville *et al.*, 1996; Wistuba *et al.*, 1999a; Wistuba *et al.*, 2000b; Wistuba *et al.*, 1999b; Wistuba *et al.*, 1997a; Yokota *et al.*, 1987) and also documented abnormalities in genes such as p53 (Liloglou *et al.*, 1997; Sozzi *et al.*, 1992), FHIT (Sozzi *et al.*, 1997a; Sozzi *et al.*, 1997b; Sozzi *et al.*, 1996b), p16 (Gazzeri *et al.*, 1998) and Rb (Gorgoulis *et al.*, 1998; Tanaka *et al.*, 1998). Several studies have also detected similar molecular abnormalities involving preneoplastic and histologically normal regions of the bronchial epithelium (Field *et al.*, 1999; Miozzo *et al.*, 1996; Park *et al.*, 1999; Park *et al.*, 2000; Waber *et al.*, 1996; Walker *et al.*, 1994; Wistuba *et al.*, 1999a; Wistuba *et al.*, 2000b; Wistuba *et al.*, 1999b; Wistuba *et al.*, 1997a). Recent research has concentrated on the molecular damage that occurs in the bronchial epithelium of smokers and patients with lung tumours in order to document the earliest genetic changes that may lead to lung cancer. Wistuba *et al.* (1999) identified multiple, sequential allele specific molecular changes in the bronchial epithelium of patients with lung tumours. These molecular changes occurred in widely dispersed, clonally independent foci in the lung providing further evidence that genetic alterations accumulate early in the pathogenesis of smoking altered bronchial epithelium. In contrast, recent work by Park *et al.* (1999) detected multiple clonal or subclonal patches within

the bronchial epithelium of patients with lung cancer and also determined the size, frequency and pattern of the molecular damage within the bronchial epithelium. Early detection of the genetic changes within respiratory epithelium, using techniques such as allelic imbalance analysis, may aid early detection of impending malignancy. Distinguishing specifically deleted loci from background allelic imbalance can allow identification of the specific genetic changes which cause tumour progression. This can be approached in two ways. First, by the analysis of large numbers of different tumours to find common underlying genetic changes within the tissue, a method which is widely used in current molecular genetic studies. An alternative method, is the detailed analysis of a single tumour and the surrounding tissue (Yatabe *et al.*, 2000). This chapter describes the enrichment and isolation of cell subpopulations from multiple sites proximal and distal to 3 bronchial carcinomas. Both tumour and bronchial epithelial cells were analysed for AI to determine the frequency of molecular damage in tumour-bearing lungs and establish the relationship between molecular damage in the airway and that in the tumour itself.

4.2 Analysis of cell subpopulations from multiple regions within tumour-bearing lungs.

4.2.1 Methods. sample selection and sample description

Samples were obtained shortly after resection of the tumour. Tissue was selected for this analysis only if the primary tumour had clearly arisen in an airway. Tissue was carefully dissected by a pathologist who sampled the tumour tissue and then the airway in which the tumour had arisen, at 1cm intervals proximal to and distal from the site of the tumour. Tissue sampled immediately proximal to the tumour was designated block P1. Airway was then

sampled at 1cm intervals away from the tumour and designated blocks P2 and P3. This method was repeated when sampling airway distal to the tumour when blocks were designated D1 (immediately distal to the tumour) then 1 cm intervals for blocks D2 and D3. Tumour tissue blocks were designated 'T' (figure 5.1). Tissue was immediately formalin fixed, then paraffin embedded as described in section 2.1.3 of chapter 2. Paraffin-embedded lung tissue was processed for microdissection as described in sections 2.1.4-2.1.7 of chapter 2. Tissues from three lungs were analysed for AI at multiple sites using fluorescent microsatellite markers. Lung tissue was sampled and examined by a reporting pathologist (Dr. John Gosney) prior to analysis. L257 was a moderately differentiated squamous carcinoma completely excised from the left upper lobe bronchus of a 71 year old male. S12 was a moderately differentiated squamous carcinoma completely excised from the left upper lobe bronchus of a 67 year old male. RC134 was a moderately differentiated squamous carcinoma with carcinoma in situ at the margin from the left lower lobe bronchus of a 56 year old male.

4.2.2 Microdissection of cell subpopulations from bronchial tissue

Target regions of bronchial epithelial cells were initially enriched for DNA analysis using manual microdissection with sterile syringe needles. Although relatively accurate, the syringe needles were cumbersome to direct and left debris attached to the target cells required for analysis. In addition, multiple tissue sections (5-10) were required for AI analysis as target regions of bronchial epithelium were often small with very few cells. Transferring dissected cells from the tissue slide for lysis and DNA analysis was also problematic and on several occasions tissue was lost during this process, making manual

microdissection often difficult and time consuming. Isolation using laser capture microdissection (LCM) provided an accurate and specific method for isolation of cell subpopulations within bronchial tissue. This method greatly reduced contamination of target tissue with non-target cells and surrounding stroma and provided a large degree of control for microdissection. LCM proved fast and reliable and, as dissected cells were attached to the tissue cap after dissection, problems which occurred when harvesting cells from tissue sections were eliminated. As LCM could be repeated using different tissue sections with the same cap, regions of lung tissue could be dissected from multiple identical sections. Technical problems were encountered when adhering lung tissue to glass slides for LCM, due to large amounts of cartilage within the tissue section (especially from bronchial tissue in the proximal airway) lifting off the slide during preparation for LCM. This was reduced by optimising the length of time slides were dried after tissue sectioning and also increasing the incubation temperature from 45°C to 60°C when drying slides for LCM. LCM efficiency was also severely affected by the presence of water in tissue sections and solutions used for slide preparation (dewaxing and staining) required regular replacement to prevent the build up of water during use. An example of LCM of epithelial and tumour cells from paraffin embedded lung tissue is shown in figure 5.2.

Figure 4.1 Diagrammatic representation of methods used to dissect bronchial airway and tumour samples from multiple locations in surgically resected lung tissue. Resected lung was carefully dissected to reveal the site of tumour growth. Tissue was sampled from the tumour (T) and from the airway in which the tumour had arisen, at 1cm intervals proximal (P1, P2 and P3 near the bronchial margin) and distal (D1, D2 and D3) to the site of the tumour. Control tissue (N) was sampled from the lung as distant from the tumour as possible. Key to figure is as follows: Red rectangle, resected lung tissue; Red Circle, Tumour; Yellow cylinder, bronchial airway; Black box, location of tissue sampling.

Figure 4.1 Diagrammatic representation of methods used to dissect bronchial airway and tumour samples from multiple locations in surgically resected lung tissue.

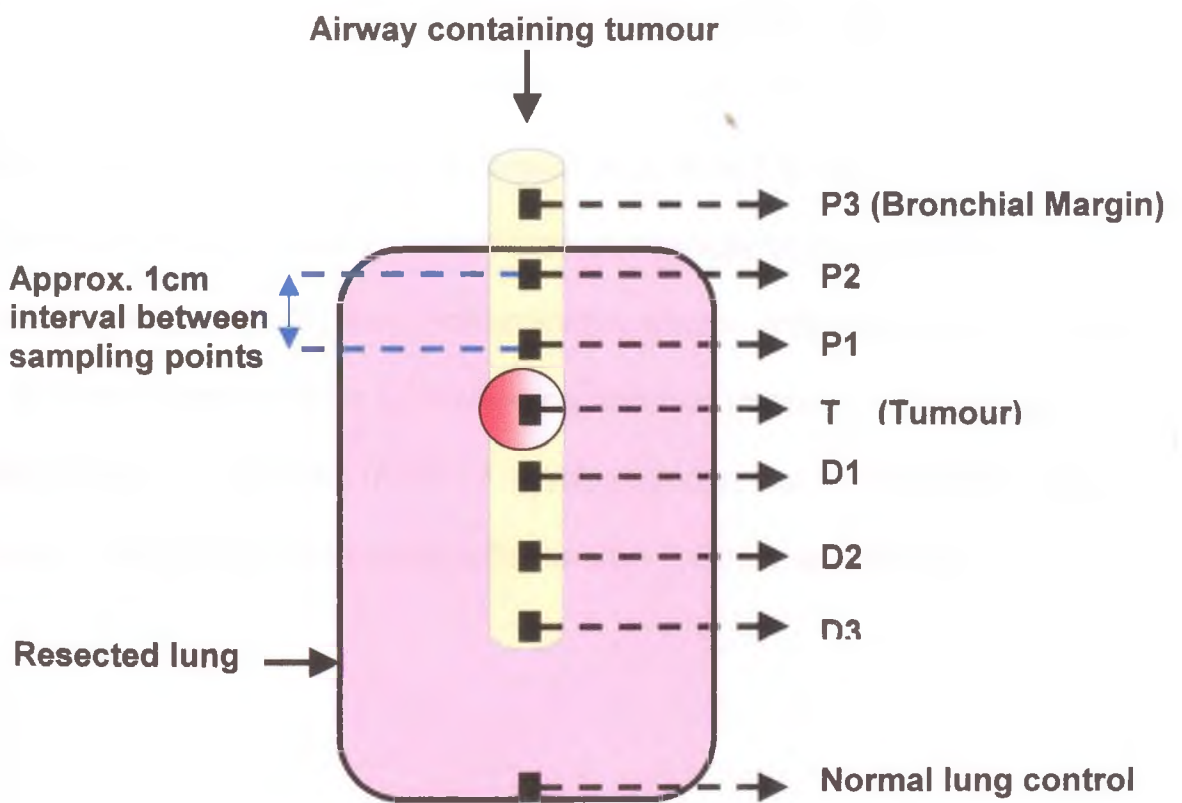
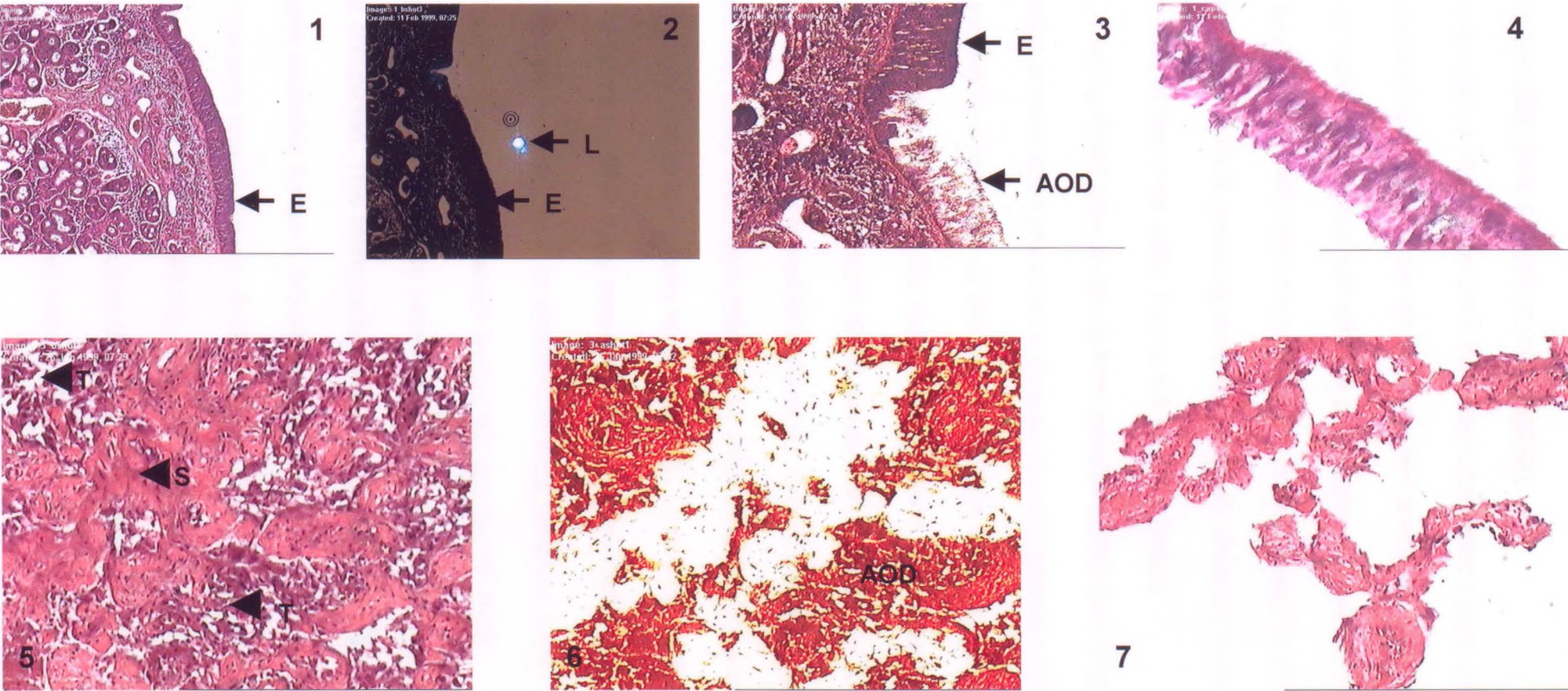


Figure 4.2. Photomicrographs demonstrating the accuracy of laser capture microdissection of normal bronchial epithelial cells and tumour cells from paraffin embedded tissue. Photomicrographs 1-4 display LCM for bronchial epithelium cells. Photomicrographs 5-7 display LCM of tumour cells. Key to figure is as follows: 1, Photomicrograph of H and E stained bronchial section prior to LCM of bronchial epithelium (E); 2, Reduced illumination of slide to show laser (L) as bright white spot in the centre of the picture prior to LCM of bronchial epithelium (E); 3, Tissue section after LCM showing area of dissection (AOD) where bronchial epithelium has been removed and adjacent unremoved bronchial epithelium (E); 4, Photomicrograph of the LCM cap after LCM showing accurately microdissected bronchial epithelial cells; 5, Photomicrograph of H and E stained bronchial tumour section prior to LCM to separate tumour cells (T) from contaminating stroma and inflammatory infiltrate (S); 6, Tumour section after LCM showing accurate removal of tumour cells from the area of dissection (AOD); 7, Photomicrograph of the slide after removal of tumour cells and the remaining inflammatory infiltrate and stroma.

Figure 4.2. Laser Capture Microdissection of normal bronchial epithelial cells (1-4) and tumour cells (5-7) from paraffin embedded tissue.



4.2.3 Accurate amplification of microsatellite sequences from paraffin-embedded lung tissue after laser capture microdissection.

Initial attempts to amplify microsatellite DNA from microdissected tissue using PCR were unsuccessful due to losing an already low yield and quality of DNA after cleaning, isopropanol extraction and ethanol precipitation. However, using a direct lysis protocol to free DNA from microdissected cells followed by PCR amplification, microsatellite PCR products were successfully obtained from small numbers (<200) of bronchial cells. Figure 4.3 shows a typical microsatellite PCR using direct lysis products from paraffin embedded lung tissue. Microsatellite amplification produced a diverse intensity between DNA samples which was highly dependent on the number of cells microdissected and the region of the lung from which the tissue was obtained. For example, a greater number of cells was required from distal airways for successful PCR due to the decreased thickness of bronchial epithelium. One of the technical problems involved with using direct lysis followed by PCR amplification was reduced PCR efficiency due to low quality template DNA. To overcome this greater volumes of template DNA were required for PCR amplifications (5-10 μ l template DNA obtained by direct lysis as opposed to 1-2 μ l template DNA obtained from blood or frozen tissue) which limited the number of PCR reactions for individual DNA samples to between 5-10 for every batch of microdissections completed. This technically hampered the amount of bulk molecular analyses available for each bronchial sample. Two random priming methods (whole genome amplification – PCR (WGA-PCR) and degenerative oligonucleotide priming – PCR (DOP-PCR)) were tested to increase DNA yields prior to molecular analysis using microsatellite PCR. Enrichment of DNA extracted from blood or bronchial direct lysis products using DOP-PCR was

unsuccessful after multiple repeated attempts. DNA extracted from whole blood was successfully amplified using the WGA method of random priming. A range of DNA concentrations were created using diluted blood DNA and WGA followed by microsatellite PCR (2 rounds of amplification) was compared with conventional microsatellite PCR (1 round of amplification). Small variations were observed between PCR reactions and diluted DNA amplified with equal efficiency in both reactions (figure 4.4). Random genomic amplification of bronchial lysis products was unsuccessful using WGA primers due to inconsistent WGA-PCR and an increase in non-specific PCR products during microsatellite PCR.

Figure 4.3 Photomicrograph displaying PCR amplification of microsatellite sequences from laser capture microdissected bronchial epithelium after direct cell lysis. Bronchial epithelial cells were microdissected using LCM from paraffin embedded tissue obtained from resected lung tissue (normal [N], tumour [T], proximal airway [P1, P2 and P3] and distal airway [D1 and D2]) then lysed overnight using proteinase K. Note that multiple individual regions of epithelium could be microdissected from the same tissue block in both lung samples reflecting metaplastic epithelium with intervening normal areas. Five microliters of direct lysis products were amplified using PCR for the microsatellite D3S1270 and electrophoresed through polyacrylamide gels. PCR products were visualised by silver staining. The photomicrograph shows PCR products from 2 multi-sampled lungs (L257 and S12) indicated in the figure. Lane numbers are displayed at the top of the photomicrograph and are as follows: Lane 1, 100 bp ladder; Lanes 2-15, amplified lysis products from L257; Lane 16, blank; Lanes 17-25, amplified lysis products from S12; Lane 26, blood DNA positive PCR control (C); Lane 27, negative PCR control; Lane 28, pBR322/HaeIII DNA size standard. Blue arrows indicate microsatellite alleles for L257 and red arrows indicate microsatellite alleles from S12.

Figure 4.3 Photomicrograph displaying PCR amplification of microsatellite sequences from laser capture microdissected bronchial epithelium after direct cell lysis.

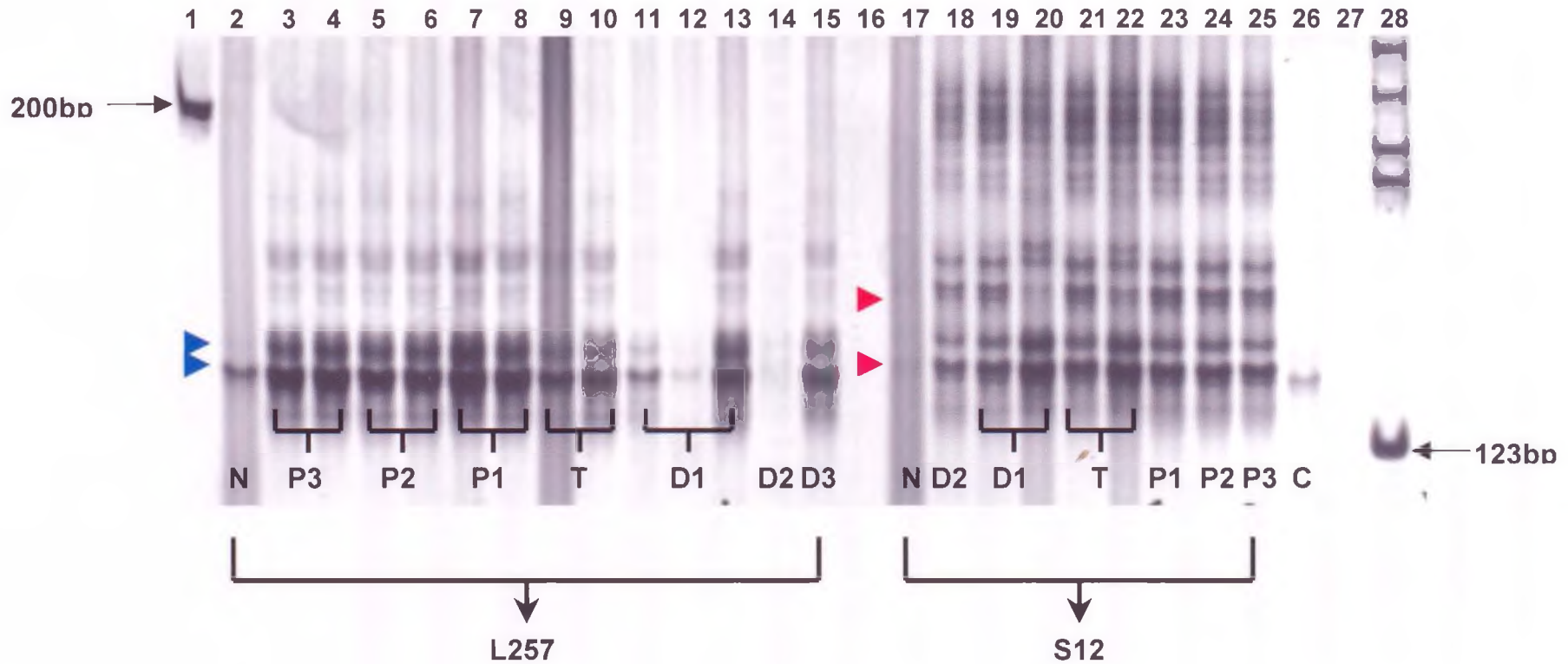
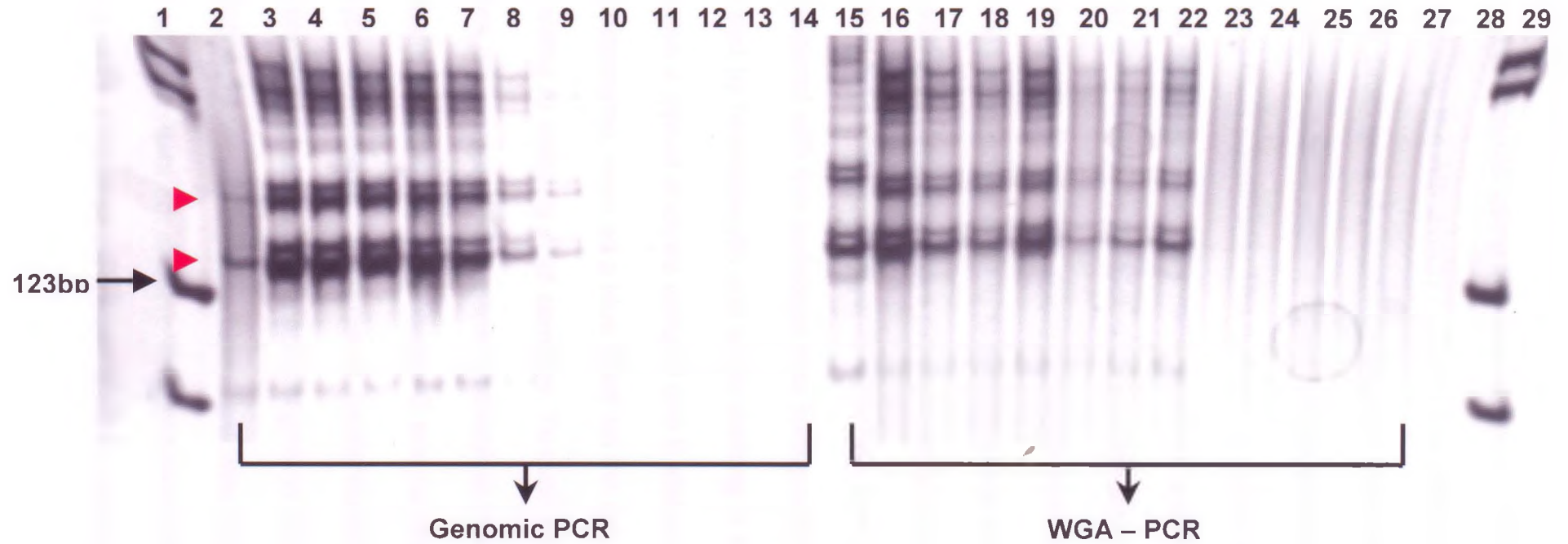


Figure 4.4. Comparison of PCR products amplified from a range of blood DNA concentrations using WGA-PCR followed by microsatellite amplification and conventional microsatellite amplification. DNA from whole blood was diluted with sterile water to form a range of DNA concentrations. One microliter of DNA from each concentration was amplified using WGA primers (WGA-PCR). One microliter of PCR product from WGA-PCR was used as a template in a second round of PCR amplification with the microsatellite primer D3S1300. In tandem the original diluted DNA was amplified in a conventional microsatellite PCR using only D3S1300 primers to provide a comparison with WGA-PCR. Products were electrophoresed through polyacrylamide gels and visualised using silver staining. Photomicrograph displays the products from both microsatellite PCR reactions. Lane orders are displayed at the top of the photomicrograph and are as follows: Lane 1 and 29, pBR322/HaeIII DNA size standard; Lane 2 and 15, Undiluted DNA; Lane 3 and 16, 1:2 dilution; Lane 4 and 17, 1:3 dilution; Lane 5 and 18, 1:4 dilution; Lane 6 and 19, 1:5 dilution; Lane 7 and 20, 1:6 dilution; Lane 8 and 21, 1:7 dilution; Lane 9 and 22, 1:8 dilution; Lane 10 and 23, 1:9 dilution; Lane 11 and 24, 1:10 dilution; Lane 12 and 25, 1:15 dilution; Lane 13 and 26, 1:20 dilution; Lane 14 and 27, 1:50 dilution; Lane 28 negative PCR control (sterile water). Conventional microsatellite PCR products are displayed in lanes 2-14 and WGA-PCR products in lanes 15-27 as indicated in the figure. Microsatellite alleles are indicated in the figure using red arrows.

Figure 4.4. Comparison of PCR products amplified from a range of blood DNA concentrations using WGA-PCR followed by microsatellite amplification and conventional microsatellite amplification.

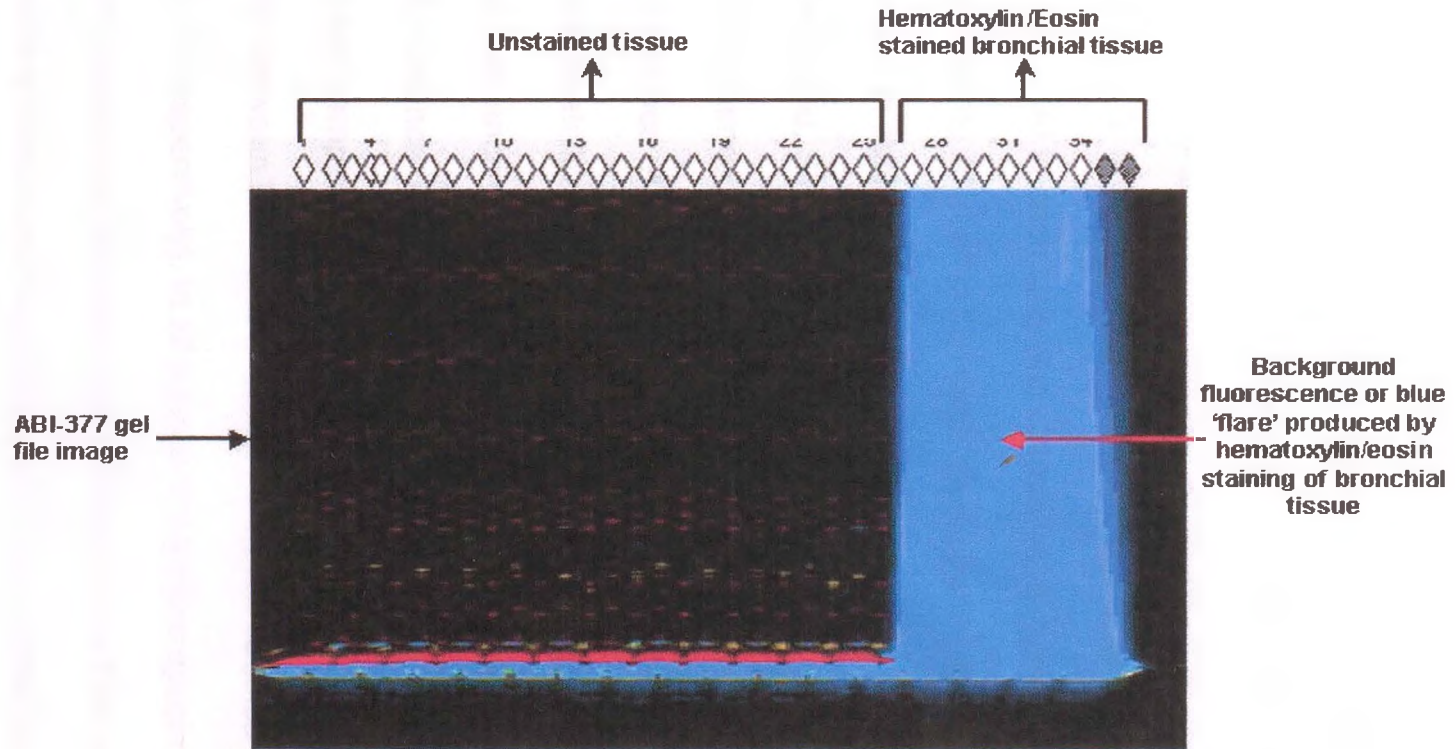


4.2.4 Analysis of AI in laser microdissected bronchial tissue using fluorescent microsatellite multiplex PCR.

Bronchial DNA was analysed using fluorescent microsatellite primers for AI in a multiplex PCR utilising the ABI 310/377. Due to low yield and quality of DNA obtained from the direct lysis of paraffin embedded bronchial tissue, fluorescent multiplex microsatellite panels were adapted for the amplification of bronchial DNA by optimising PCR conditions. The number of cycles during PCR amplification was increased from 26 cycles, used for amplification of DNA from blood and frozen lung tissue, to 28 cycles for paraffin embedded material to increase PCR products. In addition, primer concentrations were optimised within the multiplex reaction for each lung sample using multiple repeat amplifications with bronchial direct lysis products as template DNA, as variation between individual lung samples frequently occurred. One of the technical problems encountered with this technique was background fluorescence during AI analysis caused by haematoxylin and eosin staining of tissue prior to LCM. Figure 4.5 displays a typical analysis using H and E stained bronchial tissue. Background fluorescence, seen as a blue 'flare' on the gel, obscured all PCR products, preventing AI analysis in all samples. To overcome this problem, a range of dilutions were created for multiple histological stains including hematoxylin, eosin, toluidine blue and orange G and run through a gel on the ABI-377 genotyper to measure the degree of background 'flare' produced by each stain. All histological stains produced background 'flare' of varied intensity using a 1% solution. However, toluidine blue 'flare' was virtually undetectable using a 0.01% solution (data not shown). Toluidine blue (0.01%) was subsequently used as a histological stain to identify lung tissue during LCM.

Figure 4.5. Screen capture of ABI-377 gel picture displaying blue 'flare' in bronchial tissue samples created by hematoxylin and eosin staining of bronchial tissue prior to LCM. PCR products from fluorescent multiplex reactions were electrophoresed through polyacrylamide gels and analysed using the ABI-377. The figure shows the blue 'flare' present in PCR products caused by hematoxylin and eosin staining of bronchial tissue prior to LCM (red arrow), compared with PCR products from unstained tissue.

Figure 4.5. Screen capture of ABI-377 gel picture displaying blue 'flare' in bronchial tissue samples produced by hematoxylin and eosin staining of bronchial tissue prior to LCM.



4.2.5 Tumour bearing lungs display frequent AI in both bronchial epithelium and tumour.

Bronchial epithelial and tumour cells were dissected from lung tumours and the surrounding airway of 3 resected lungs (L257, S12 and RC134) containing NSCLC and analysed for AI at multiple sites using three panels of pre-optimised fluorescent microsatellite markers. All samples were examined by the reporting pathologist (Dr. John Gosney) and regions of metaplasia and dysplasia marked on the slide for reference during microdissection. The results of AI analysis are displayed in tables 4.1 – 4.3 and examples of AI can be seen in figures 4.6 – 4.8. The total incidence of AI in each lung was calculated as [total incidence of informative samples showing AI at all loci in each lung] / [total informative samples tested at all loci in each lung]. L257 displayed an incidence of AI in 30/87 (34%) samples tested, S12 an incidence of AI in 42/140 (30%) samples tested and RC134 an incidence of AI in 32/43 (74%) samples tested. Figure 5.9 displays the overall incidence of AI at each informative locus in the 3 tumour bearing lungs. The microsatellite loci D5S644 at 5q15 and D3S1600 at 3p14.2 displayed the highest incidence of AI (70%) and the locus D3S1580 at 3q29 displayed the lowest incidence of AI (10%). Fractional allelic loss (FAL) values were calculated for each sample as [loci scored with AI] / [total informative loci]. Mean FAL was similar between L257 and S12 but more than doubled in RC134 (0.339, 0.300 and 0.733 respectively). In all 3 lungs the tumour specimen displayed the highest frequency of AI which was reflected in the FAL value. In addition, severe dysplasia (carcinoma in situ) from S12 and normal bronchial epithelium, mild atypia and moderate atypia from RC134 also displayed a similar frequency of AI. Overall, histologically normal bronchial epithelium in all three tumours displayed a lower mean FAL than abnormal bronchial epithelium

(metaplasia and mild, moderate, severe atypia's) and tumour (0.263 ± 0.075 and 0.689 ± 0.108 respectively; t-test, $p=0.003$).

Table 4.1. Results of AI analysis for L257. Figure displaying the results of AI in microdissected tissue from tumour-bearing L257. The first column displays from where tissue was microdissected within the lung (proximal (P), distal (D) or tumour (T)). The second column displays the histology (N, histologically normal bronchial epithelium; SCC, tumour) of the microdissected cells. Microsatellite loci are displayed at the top of each column of blocks representing the results of AI analysis within each tissue sample. The last column displays the fractional allelic loss (FAL) for each microdissected region of the lung. Black boxes – allelic imbalance, white boxes – heterozygous (no AI). Blank spaces represent samples where AI could not be determined after repeated attempts.

Table 4.1. Results of AI analysis for L257

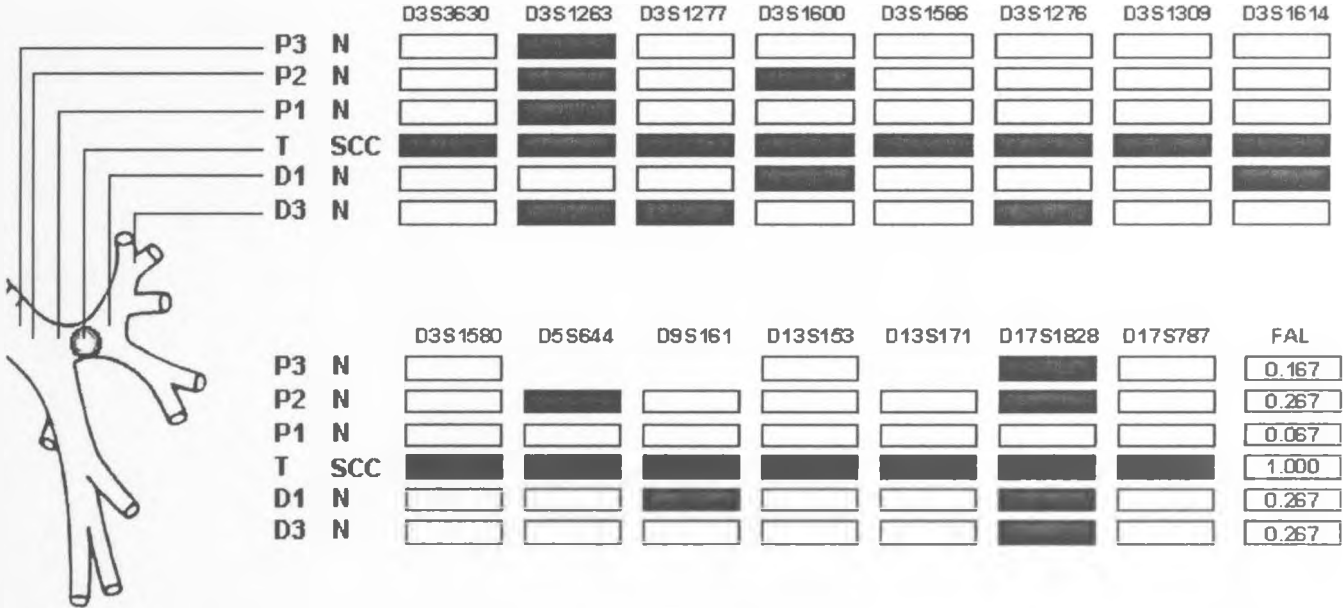


Figure 4.6. Examples of AI at microsatellite loci D3S1263 and D17S1828 for L257. Screen capture displaying examples of AI for the microsatellite loci D3S1263 (left column) and D17S1828 (right column) after Genescan™ analysis of ABI-377 data for L257. Examples are shown for each region of the lung tested from proximal airway, tumour and distal airway. Black arrows indicate AI in the sample. Note that epithelium from proximal and distal airways is histologically normal (P3, P2, P1, D1 and D3). Control tissue was dissected from the pleura.

Figure 4.6. Examples of AI at microsatellite loci D3S1263 and D17S1828 for L257.

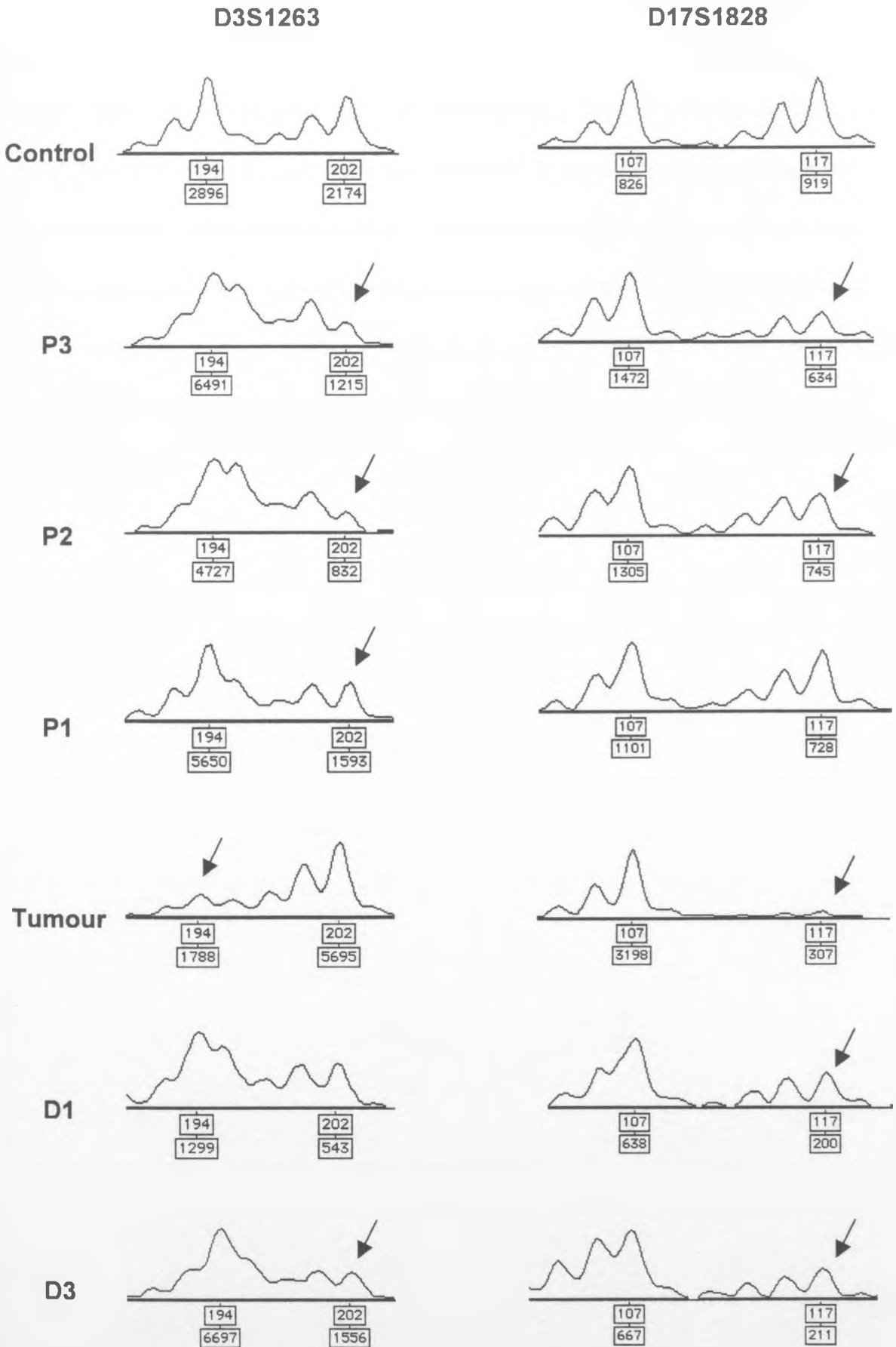


Table 4.2. Results of AI analysis for S12. Figure displaying the results of AI in microdissected tissue from tumour-bearing lung S12. The first column displays where tissue was microdissected from within the lung (proximal (P), distal (D) or tumour (T)). The second column displays the histology (N, histologically normal bronchial epithelium; SCC, tumour; Met, Metaplasia with no atypia; Mild, Mild atypia; Sev, Severe atypia) of the microdissected cells. Microsatellite loci are displayed at the top of each column of blocks representing the results of AI analysis within each tissue sample. The last column displays the fractional allelic loss (FAL) for each microdissected region of the lung. Black boxes – allelic imbalance, white boxes – heterozygous (no AI). Blank spaces represent samples where AI could not be determined after repeated attempts.

Table 4.2. Results of AI analysis for S12.

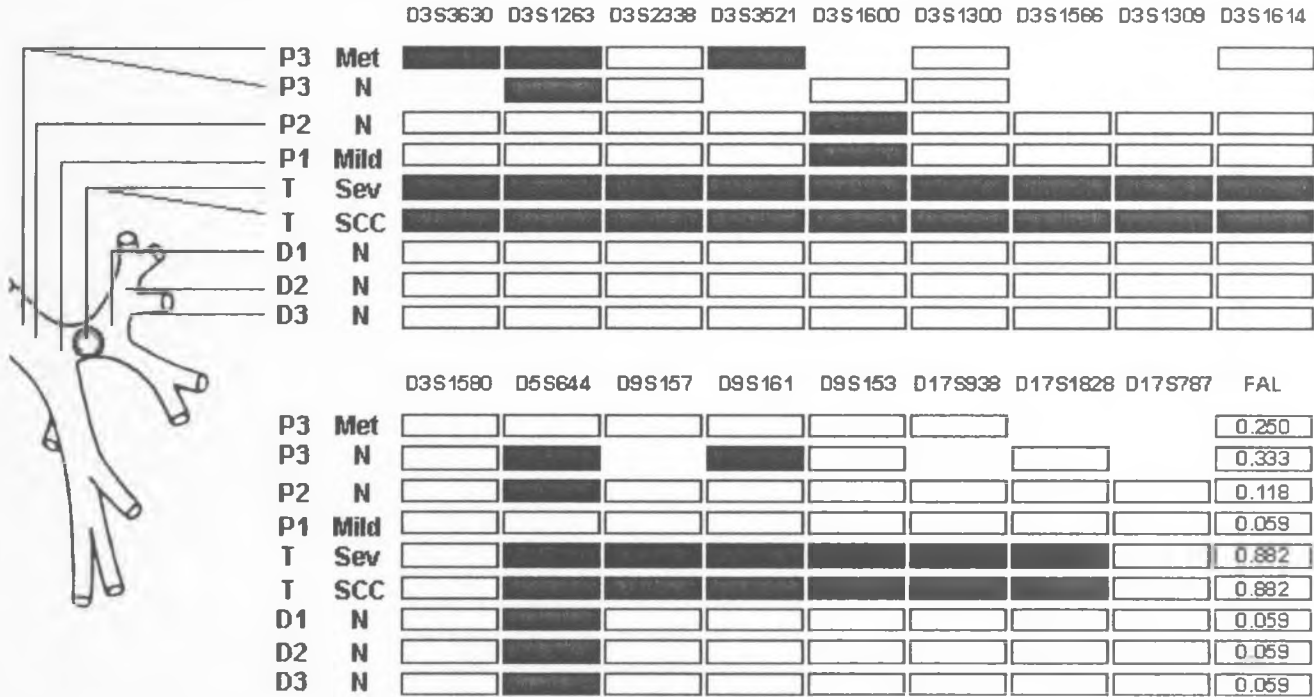
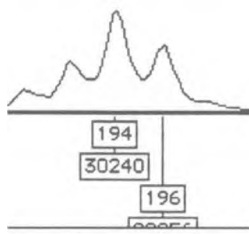
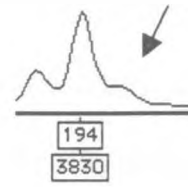


Figure 4.7. Examples of AI at microsatellite locus D3S1263 for S12. Screen capture displaying examples of AI for the microsatellite locus D3S1263 after Genescan™ analysis of ABI-377 data for S12. Examples are shown for each region of the lung tested from proximal airway, tumour and distal airway. Black arrows indicate AI in the sample. Metaplasia, mild and severe indicates different histological grades of metaplastic bronchial epithelium from regions P3, P1 and Tumour in the lung. Normal indicates histologically normal epithelium from regions P3, P2, D1, D2 and D3 in the lung. Control tissue was dissected from the pleura.

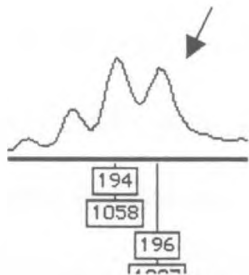
Figure 4.7. Examples of AI at microsatellite locus D3S1263 for S12.



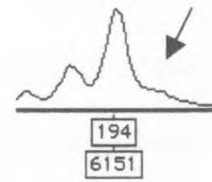
Control



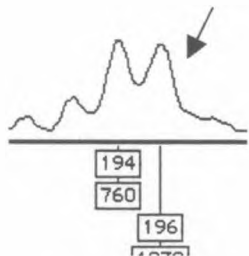
T Severe Atypia



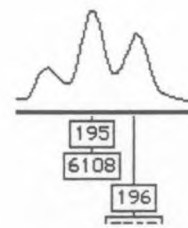
P3 Metaplasia



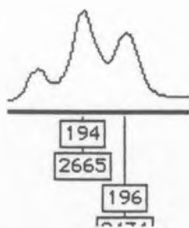
T Tumour



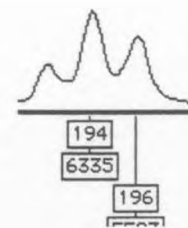
P3 Normal



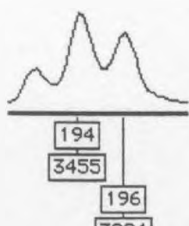
D1 Normal



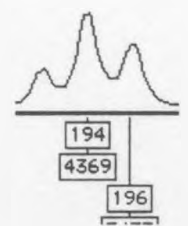
P2 Normal



D2 Normal



P1 Mild



D3 Normal

Table 4.3. Results of AI analysis for RC134. Figure displaying the results of AI in microdissected tissue from tumour-bearing lung RC134. The first column displays where tissue was microdissected from within the lung (proximal (P), distal (D) or tumour (T)). The second column displays the histology (N, histologically normal bronchial epithelium; SCC, tumour; Met, Metaplasia with no atypia; Mild, Mild atypia; Mod, Moderate atypia) of the microdissected cells. Microsatellite loci are displayed at the top of each column of blocks representing the results of AI analysis within each tissue sample. The last column displays the fractional allelic loss (FAL) for each microdissected region of the lung. Black boxes – allelic imbalance, white boxes – heterozygous (no AI). Blank spaces represent samples where AI could not be determined after repeated attempts.

Table 4.3. Results of AI analysis for RC134.

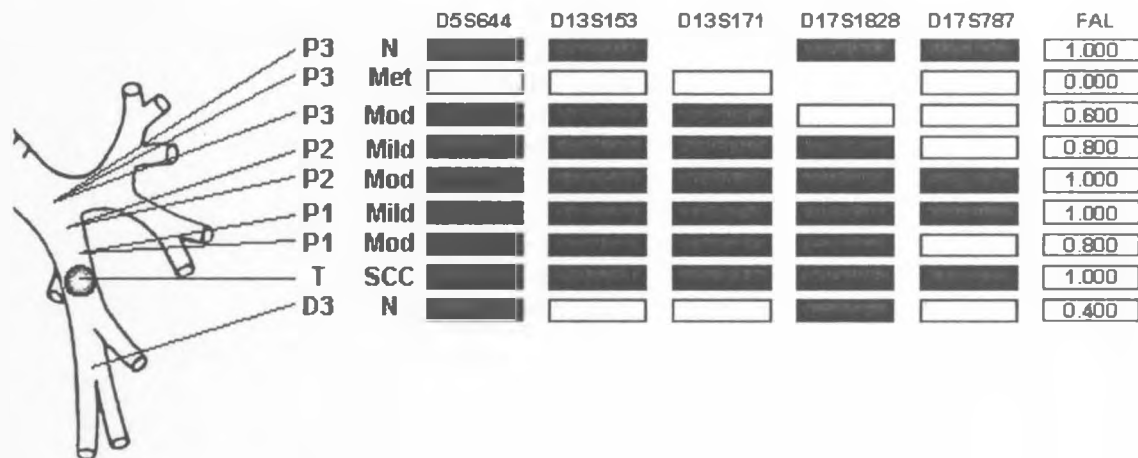


Figure 4.8. Examples of AI at microsatellite locus D5S644 in RC134. Screen capture displaying examples of AI for the microsatellite locus D5S644 after Genescan™ analysis of ABI-377 data for RC134. Examples are shown for each region of the lung tested from proximal airway, tumour and distal airway. Black arrows indicate AI in the sample. Metaplasia, mild and moderate indicates different histological grades of metaplastic bronchial epithelium from regions P3, P2 and P1 in the lung. Normal indicates histologically normal epithelium from regions P3 and D3 in the lung. Control tissue was dissected from the pleura.

Figure 4.8. Examples of AI at microsatellite locus D5S644 in RC134.

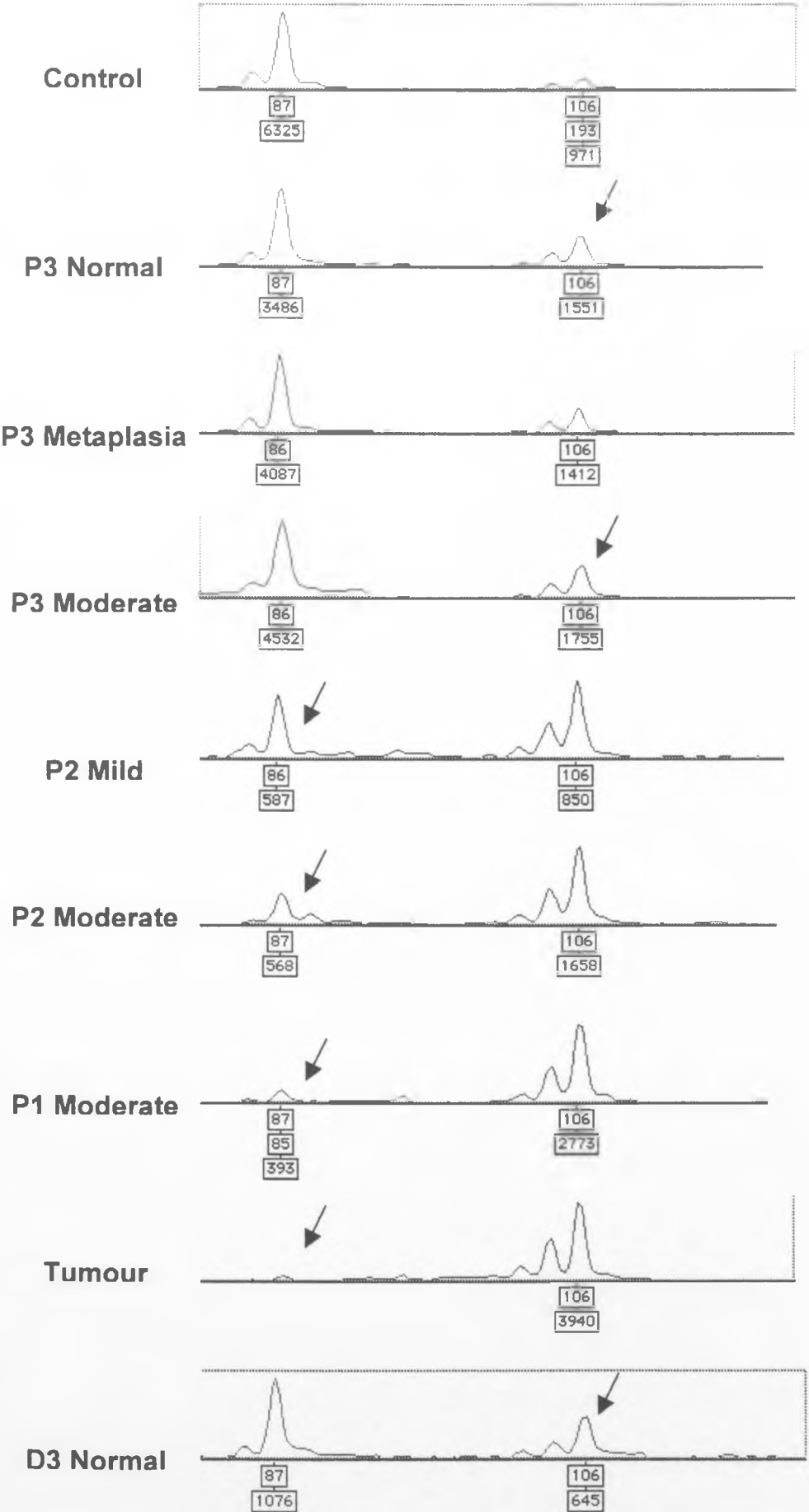
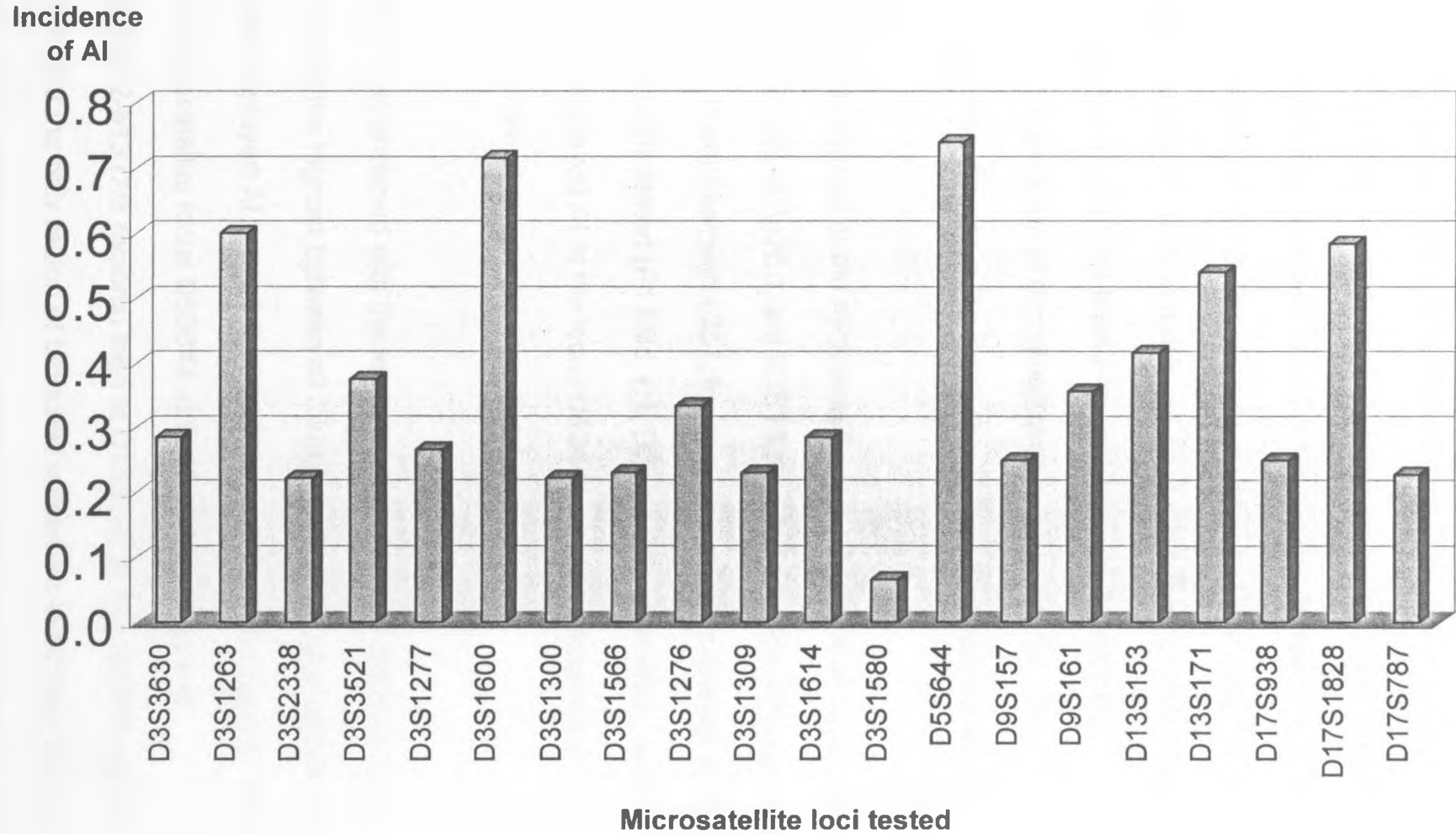


Figure 4.9. Overall incidence of AI at informative loci in 3 tumour bearing lungs. Histogram displaying the total incidence of AI detected in 3 tumour bearing lungs at informative loci. Horizontal axis displays the microsatellite loci tested and the vertical axis displays the total incidence of AI. Incidence of AI was calculated by [total number of samples displaying AI at loci] / [total number of informative loci] for the three combined tumour specimens.

Figure 4.9. Overall incidence of AI at informative loci in 3 tumour bearing lungs.



4.2.6 Different patterns of AI in tumour bearing lungs

In all three lungs tested, AI was detected in every sample at one or more loci. In L257, AI was more frequent at the microsatellite locus D3S1263 at 3p26.1 and D17S1828 at 17q12-p13 with AI detected in 5/6 regions tested in the lung. AI at D3S1600 at 3p14.2 was also detected in half of the regions tested (P2, T and D1). Bronchial epithelium from P2 displayed AI at only one locus (D3S1263) whereas all other regions of the lung displayed AI at more than one loci. The tumour L257 displayed AI at all loci tested on chromosome 3 indicating complete deletion or duplication of this chromosome in the tumour. The frequency of AI was also greater in the distal airway than the proximal airway.

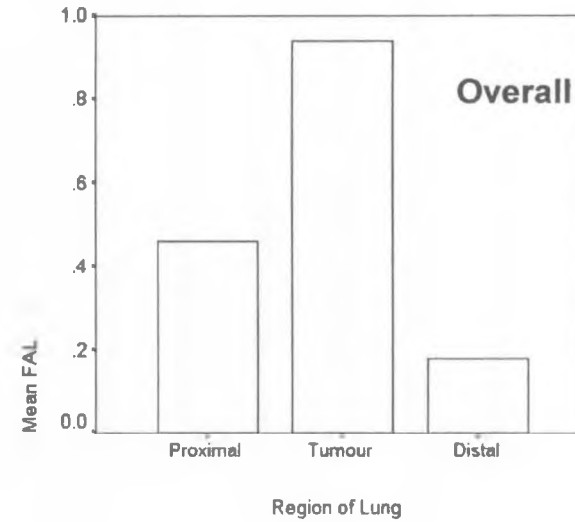
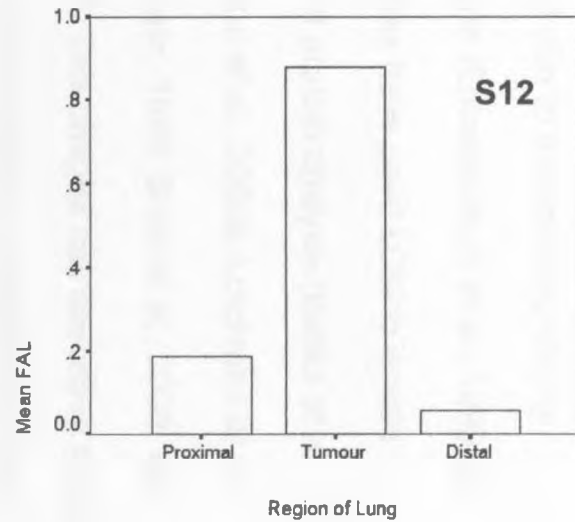
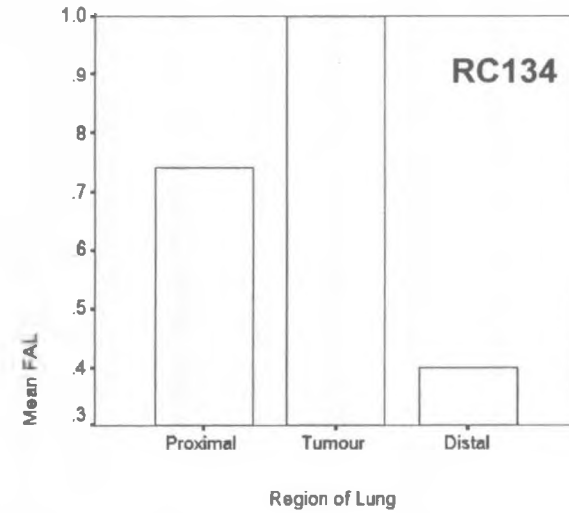
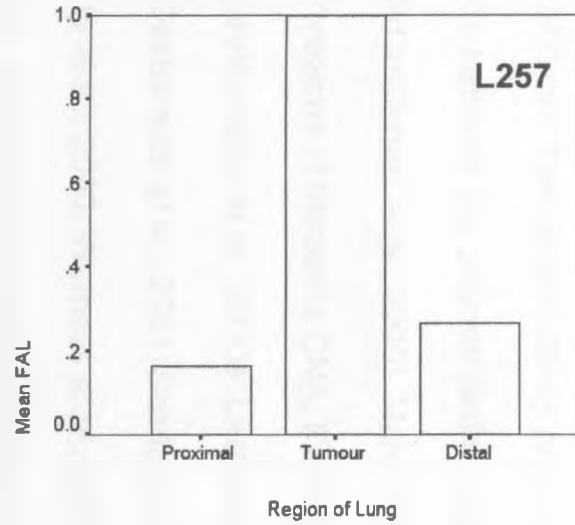
In S12, AI was most frequent at the microsatellite locus D5S644 at 5q13 (7/9 regions) then at D3S1263 at 3p26.1 and D3S1600 at 3p14.2 (both 4/9 regions) on chromosome 3p. In contrast with L257, four regions of S12 displayed AI at only 1 microsatellite locus tested (P1 Mild, D1, D2 and D3). In addition, distal airway in S12 only displayed AI at the locus D5S644 and the frequency of AI was greater in the proximal airway than the distal airway.

Although, RC134 was analysed with the least number of microsatellite loci, the frequency of AI was the highest between all 3 lungs tested. All microsatellite loci except D17S787 displayed AI in more than half the regions of the lung. AI was greatest at the microsatellite locus D5S644 at 5q13 (8/9 regions) and D17S1828 at 17p12-p13 (7/8 regions) then at D13S153 at 13q14 (7/9 regions). RC134 also provided the only region of bronchial epithelium in all three lungs which displayed no AI at all informative loci tested. In RC134 the frequency of AI was greater in proximal airway than distal airway. Figure 4.10 displays the

mean FAL obtained from each lung in tumour and also proximal and distal airway.

Figure 4.10. Histograms showing comparison of mean FAL for tumour and proximal and distal airways in tumour bearing lungs. Mean FAL was calculated in the tumour and also the proximal and distal airways for each lung and displayed as a histogram. Figure shows four histograms, one for each lung tested. The fourth histogram displays an overall mean FAL comparison for the tumour and proximal and distal airways based on FAL data for individual samples in all three lungs. In all 4 histograms the horizontal axis shows the region of the lung for which mean FAL was calculated (Tumour, proximal airway and distal airway) and the vertical axis displays Mean FAL value.

Figure 4.10. Comparison of mean FAL for tumour and proximal and distal airways in tumour bearing lungs.



4.3 Discussion

4.3.1 Accurate isolation of cell subpopulations from paraffin embedded bronchial tissue using LCM and amplification of DNA for AI analysis.

Molecular genetic techniques such as AI and LOH analysis in small cell populations relies heavily on the accurate isolation and enrichment of cell subpopulations, especially from paraffin-embedded formalin fixed tissue which yields low quality, short fragment size DNA (Shibata, 1994). Manual microdissection, using syringe needles or drawn-out capillary tubes attached to manipulators, to scrape away tissue surrounding target cells is a widely used technique to isolate small regions of tissue (Going & Lamb, 1996; Moskaluk & Kern, 1997) for DNA analysis. However, due to reasons described earlier in this chapter, this technique is often technically difficult and time consuming. The advent of laser capture microdissection (LCM) has provided a quick and reliable method for isolating small regions of target tissue. LCM works by overlaying the region of interest within a tissue section with a thermoplastic film (ethylene vinyl acetate polymer). A carbon dioxide laser melts the transparent film onto target cells which can then be selectively procured due to the resulting strong adhesion between the polymer and the tissue (Emmert-Buck *et al.*, 1996 and reviewed in Curran *et al.*, 2000). Many studies have used LCM to isolate specific regions of tissue for DNA, RNA and protein analysis (Banks *et al.*, 1999; Leethanakul *et al.*, 2000a; Leethanakul *et al.*, 2000b; Lutchman *et al.*, 1999; Oesterreich *et al.*, 2001; Schutze & Lahr, 1998; Shen *et al.*, 2000; von Eggeling *et al.*, 2000; Werness *et al.*, 2000). This chapter has described the use of LCM to successfully isolate small regions of bronchial epithelial tissue and tumour for AI analysis in tumour bearing lungs. Specific problems were encountered when manually microdissecting bronchial tissue due to the

accurate isolation and collection of small cell subpopulations for AI analysis. Multiple, small distinct regions of metaplastic epithelium and bronchial epithelium from distal airway were particularly difficult to microdissect due to the low cell numbers and reduced epithelial layer thickness. LCM eliminated many of the problems associated with manual microdissection involving specific cell isolation and collection for analysis. Direct lysis products from LCM tissue were reliably amplified for microsatellite sequences using PCR.

Multiple protocols exist for the total amplification of the human genome including whole genome amplification , primer extension preamplification and DOP-PCR (Barrett *et al.*, 1995; Cheung & Nelson, 1996; Dietmaier *et al.*, 1999; Hubert *et al.*, 1992; Klein *et al.*, 1999; Paulson *et al.*, 1999; Sermon *et al.*, 1996; Sun *et al.*, 1995; Telenius *et al.*, 1992a; Telenius *et al.*, 1992b; Wells *et al.*, 1999; Zhang *et al.*, 1992). Whole genome amplification (WGA) contrasts from conventional PCR in that a random oligonucleotide is used to amplify the total genomic DNA sequence in a single sample as opposed to single sequences. It can also be used to increase DNA yield in small samples and even single cells or molecules (Hubert *et al.*, 1992). It has two targets, one to increase the number of sequences available for analysis (i.e. DNA yield) and the other to amplify all the sequences in the sample so that one sequence is not over-represented. It has been estimated that starting with a single haploid cell, 50 primer extensions produce an average of 60 DNA copies, with approximately 78% of the genome represented at least 30 times (Zhang *et al.*, 1992). DOP-PCR is based on a similar method using a different oligonucleotide to random amplify the genome and studies have shown that DOP-PCR allows an unbiased, 100 fold amplification of the human DNA suitable for molecular

genetic analysis of the genome (Cheung & Nelson, 1996; Sanchez-Cespedes *et al.*, 1998; Telenius *et al.*, 1992a). This chapter has described an attempt to enrich template DNA concentrations and volumes using both WGA-PCR and DOP-PCR. DOP-PCR however proved to be unsuccessful after multiple repeated attempts to amplify both blood and bronchial DNA. WGA amplification was successfully used to amplify a range of DNA concentrations. However when compared to conventional microsatellite PCR products, no real difference was observed between DNA concentrations for each technique. Amplification of bronchial DNA using WGA methods followed by microsatellite PCR also proved unsuccessful after repeated attempts.

Amplification of DNA from LCM bronchial lysis products using fluorescent primers was initially hampered by dyeing lung tissue with hematoxylin and eosin. Hematoxylin and eosin staining was used prior to LCM to distinguish histological features in bronchial section and proved extremely useful in discriminating between normal and adjacent metaplastic region of bronchial epithelium. However these dyes also produced an unexpected blue 'flare' in LCM bronchial samples when analysed using the ABI-377 genotyper. Burton *et al* (1998) and Murase *et al* (2000) both examined the effects of a range of histological dyes on DNA amplification during PCR and concluded that haematoxylin should be avoided when analysing small microdissected regions of formalin fixed paraffin embedded tissue as it inhibits successful PCR amplification. As previously described in this chapter several histological stains were tested to compare their effects on tissue distinction for LCM and analysis using the ABI-377 genotyper. Toluidine blue (at a 0.01% concentration) was considered to be a reliable stain for tissue distinction during LCM. In addition,

the efficiency of PCR was not noticeably reduced and blue 'flare' that hindered genotyping analysis was eliminated from the gel.

4.3.2 Analysis of bronchial epithelium and tumour tissue in tumour-bearing lungs revealed multiple regions of AI.

Molecular genetic analysis of bronchial epithelium and tumour tissue from 3 tumour-bearing lungs revealed multiple regions of AI in histologically normal epithelium, metaplastic epithelium and tumour at several chromosomal locations. These results confirm the findings of previous studies which have demonstrated that molecular aberrations occur early in the pathogenesis of NSCLC (Chung *et al.*, 1995; Fong *et al.*, 1997; Hung *et al.*, 1995; Kishimoto *et al.*, 1995b; Kohno *et al.*, 1999; Sozzi *et al.*, 1992; Sozzi *et al.*, 1995; Sundaresan *et al.*, 1995; Wistuba *et al.*, 1999a; Wistuba *et al.*, 2000a; Wistuba *et al.*, 1997a; Park *et al.*, 1999; Park *et al.*, 2000; Sundaresan *et al.*, 1992). Work by Sozzi *et al* (1992) provided evidence for deletions of 17p and p53 abnormalities in preneoplastic lesions of the lung. Further work by Sundaresan *et al* (1992) identified LOH on chromosome 3p and abnormalities of the p53 gene, detected through LOH or immunohistochemical analysis. Kishimoto *et al* (1995b) demonstrated that LOH on chromosome 9p also appears in preneoplastic lesions of NSCLC. This study also detected abnormalities in all regions of the respiratory tract including bronchi, bronchioles and alveoli. Witsuba *et al* (1997a) detected multiple genetic changes in the bronchial epithelium of current and former smokers without lung carcinomas. In this study, approximately half of the histologically normal specimens from smokers showed LOH and microsatellite alterations were detected in 64% of individuals. Common microsatellite changes included LOH on 3p and 9p and, less

frequently, 5q, 17p and 13q. The work presented in this chapter has focused on examining molecular abnormalities in the respiratory airway at 1cm intervals both proximal and distal to the site of the tumour. AI was detected on chromosomes 3p, 5q, 9p, 13q and 17p in histologically normal or preneoplastic bronchial epithelium and tumour. The highest frequency of AI was detected at D3S1600 at 3p14.2 and D5S644 at 5q15. Chromosome 3p has been specifically identified as a common target for deletion in both NSCLC and SCLC (Neville *et al.*, 1996; Tsuchiya *et al.*, 1992; Virmani *et al.*, 1998) and preneoplastic lesions (Wistuba *et al.*, 2000a). The FHIT gene is located at chromosome 3p14.2 in close proximity to the microsatellite locus D3S1600, and encompasses the FRA3B fragile site and also the renal cell carcinoma associated t(3;8) breakpoint (Cohen *et al.*, 1979; Glover *et al.*, 1988; Ohta *et al.*, 1996). Abnormalities of the FHIT gene have been frequently demonstrated in both NSCLC and SCLC (Sozzi *et al.*, 1996b). Loss of FHIT expression also occurs in a high percentage of primary NSCLC lung tumours and preneoplastic lesions (Sozzi *et al.*, 1998). LOH on 5q has been demonstrated to commonly occur in NSCLC (Liloglou *et al.*, 2000; Neville *et al.*, 1996) although little evidence exists for the involvement of 5q in preneoplastic lesions (Wistuba *et al.*, 1997a) and the role of this chromosome arm in NSCLC remains unclear.

4.3.3 Patterns of AI in tumour-bearing lungs

L257 and S12 displayed similar overall frequencies of AI which was reflected in the total frequency of AI in each lung (34% and 30% respectively) and also the mean FAL value (0.339 and 0.300 respectively). Both lung samples displayed a similar distribution of molecular damage with concentration of AI detected in the tumour with patches of AI distributed through the proximal and distal

airways. AI common with the tumour was also detected in the proximal and distal airway in each case. In L257 the microsatellite D3S1263 displayed AI in the tumour and also all of the proximal airway tested. AI was also detected at this loci in distal airway at D3 but not in D1. Similarly AI at D17S1828 was detected in the tumour and also throughout all the distal airway tested and in regions P3 and P2 of the proximal airway. This indicates that AI at D3S1262 and D17S1828 may represent one of the initial genetic alterations in L257. In S12 the microsatellite locus D5S644 was detected in the tumour, the distal airway and in P3 and P2 in the proximal airways. This indicates that AI at D5S644 locus may represent an initial genetic alteration in S12. RC134 displayed a high frequency of AI and mean FAL was more than double the FAL value calculated in the other two lungs. Statistical analysis revealed that histologically normal epithelium was associated with a lower mean FAL (0.263 ± 0.075) than metaplastic bronchial epithelium and tumour (0.689 ± 0.108).

Witsuba *et al* (1999a) investigated the genetic damage in small foci of normal and metaplastic epithelium and tumour tissue using a panel of 31 microsatellites and detected a progressive increase in the overall frequency of LOH as the severity of the histopathological changes increased. Smith *et al* (1996) also investigated aneuploidy in the lungs of patients with NSCLC and detected an increase in the degree and incidence of aneuploidy with the progressive severity of morphological change in multiple preneoplastic lesions.

Comparisons of overall mean FAL in the proximal and distal airways for all three lungs indicated that mean FAL was greater in the proximal airway than the distal airways, indicative of increased AI in this region of the lung. However, when mean FAL was calculated individually, L257 displayed a higher mean FAL in the distal airway than the proximal airway. This may indicate that the

distribution of genetic damage in the airway proximal and distal to lung tumours varies between individuals and may reflect different environmental factors in lung tumour development such as the time duration of carcinogen exposure and smoking habits.

4.3.4 Individual regions of field cancerization in tumour bearing lungs.

In 1953 Slaughter *et al* described a process called 'field cancerization' occurring in oral stratified squamous epithelium. Field cancerization is thought to occur by the prolonged exposure of a large area of tissue to carcinogens which render the cells 'cancer-prone', although they appear morphologically normal.

Subsequent genetic insults to small groups of preconditioned cells in this tissue could initiate the growth of abnormal clones which then develop into multiple tumours. This theory would seem to explain the independent development of multiple primary and secondary neoplasm's which are presented in epithelial malignancies such as HNSCC (Califano *et al.*, 1999; Carey, 1996; Scholes *et al.*, 1998; van Oijen *et al.*, 2000; van Rees *et al.*, 2000), prostate (Cheng *et al.*, 1998), lung carcinoma (Hiroshima *et al.*, 1998; Shimizu *et al.*, 2000) (Smith *et al.*, 1976; Uyama *et al.*, 1989) and gastric cancer (Kang *et al.*, 1997). Work by Sozzi *et al* (1995) supported the theory of field cancerization in lung tumourigenesis by providing evidence that metachronously arising lung tumours had different genetic changes on chromosome 3p and within the p53 gene and hence near complete independency with regard to the genotype of the tumours. However research has shown that multiple lung tumours can also share common genetic changes such as identical small deletions of chromosome 3p and common p53 mutations in lesions arising at physically different locations

which would seem to support the clonal development of these tumours (Chung *et al.*, 1995; Hung *et al.*, 1995).

Figures 4.11 and 4.12 shows how distribution of AI in the proximal and distal airways related to AI in the tumour and the relationship between different 'fields' of AI (field cancerization). In L257, all airway analysed proximal and distal to the tumour, and the tumour itself, displayed AI of loci on 3p (in particular D3S1263) which probably represented one of the initial genetic abnormalities in the airway. This is followed by AI of 17p (D17S1828) which appeared not to preclude AI at D3S1263 as region P1 in the proximal airway did not display AI at this locus. Several regions of the lung then appeared to follow separate genetic pathways (reflecting different fields of cancerization) as epithelium from P2 displayed AI at D5S644 on 5q and in the tumour. Both tumour and epithelium from the adjacent airway distal to the tumour (D1) displayed AI at loci on 9p. Therefore the sequence of molecular damage in L257 could be considered to follow the order 3p→17p→9p/5q then 3q/13q/17q which was associated with tumour development in this lung. In contrast, S12 displayed a much simpler pathway. The primary genetic change in this lung was detected at the locus D5S644 which displayed AI at all tested regions of the airway, including the tumour, except for mild metaplastic epithelium in P1. The second most common abnormality was AI at 3p which only affected the proximal airway and the tumour. Cells from regions P3 and the tumour then followed separate genetic pathways with AI at 9p. A region of metaplasia with no atypical features in the proximal region at P3 did not display any AI at any loci tested other than 3p even though adjacent histologically normal epithelium displayed AI on chromosomes 5q and 9p. This supports the hypothesis of benign metaplasia

being a protective mechanism in the lung in response to exposure to carcinogens and a reversible lesion (Boers *et al.*, 1996; Thiberville *et al.*, 1995; Tsuchiya *et al.*, 1987; Wakamatsu *et al.*, 1999). A region of severe atypical metaplasia, located immediately adjacent to the tumour, was analysed for AI and displayed the same pattern of AI as the tumour. This suggests that AI at different chromosomal loci other than those analysed in this study most likely is responsible for tumour invasion in this lung. The sequence of molecular damage in S12 could be considered to follow the order of 5q→3p→9p then 3q/13q/17p which was associated with tumour development in this lung.

In RC134 the primary genetic change in this lung was detected at the locus D5S644 at 5q which displayed AI at all regions of the airway proximal and distal to the tumour, including the tumour itself. The second most common abnormality was AI at either 13q or 17p. Chromosome 13q AI was detected in all proximal regions of the airway except benign metaplasia at P3. No 13q AI was detected in the distal airway. Chromosome 17p AI was detected in the distal airway at D3 and also the tumour, and in the proximal airway at P2 and P3. No 17p AI was detected in benign metaplasia or moderate metaplasia at P3, however 17p AI was detected in histologically normal epithelium in this region. The sequence of molecular damage in RC134 could be considered to follow the order of 5q→13q/17p then 17q which was associated with tumour development in this lung.

Figure 4.11. Diagrammatic representation of the distribution of AI detected in the tumour and proximal and distal airways for L257 and S12. Each airway is represented as a yellow cylinder with the tumour denoted as a shaded red circle located centrally. In both figures P1, P2 and P3 denote the different regions of epithelium dissected from the proximal airway and D1, D2 (only in S12 as D2 was not analysed in L257) and D3 denote different regions of epithelium dissected from the distal airway. Note that all tissue is histologically normal bronchial epithelium unless otherwise stated (In S12 all dissected epithelium from region P2 was mild atypical metaplasia). The AI detected on each chromosomal arm in the tumours of L257 and S12 is indicated within the red shaded circle representing the tumour. The distribution of AI in the proximal and distal airways is represented as rectangles. The chromosomal arms affected in each region of the airway are displayed inside the rectangles. AI is colour coded as follows: Blue, 3p AI; Black, 5q AI; Green, 17p AI; Plum, 9p AI. In S12, AI detected in individual regions of metaplasia microdissected from the airway in addition to normal bronchial epithelium are represented by red rectangles.

Figure 4.11. Diagrammatic representation of the distribution of AI in L257 and S12.

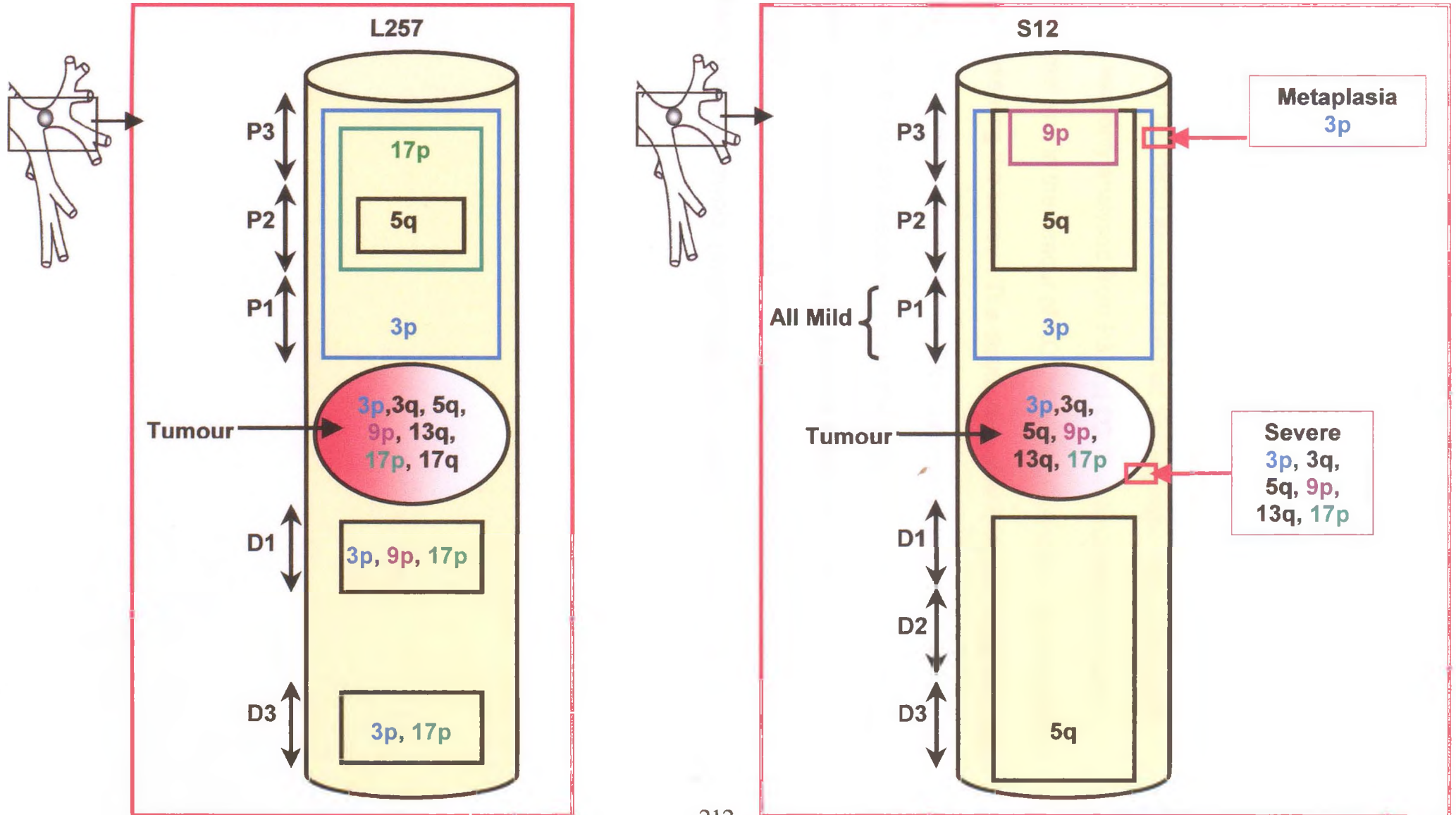
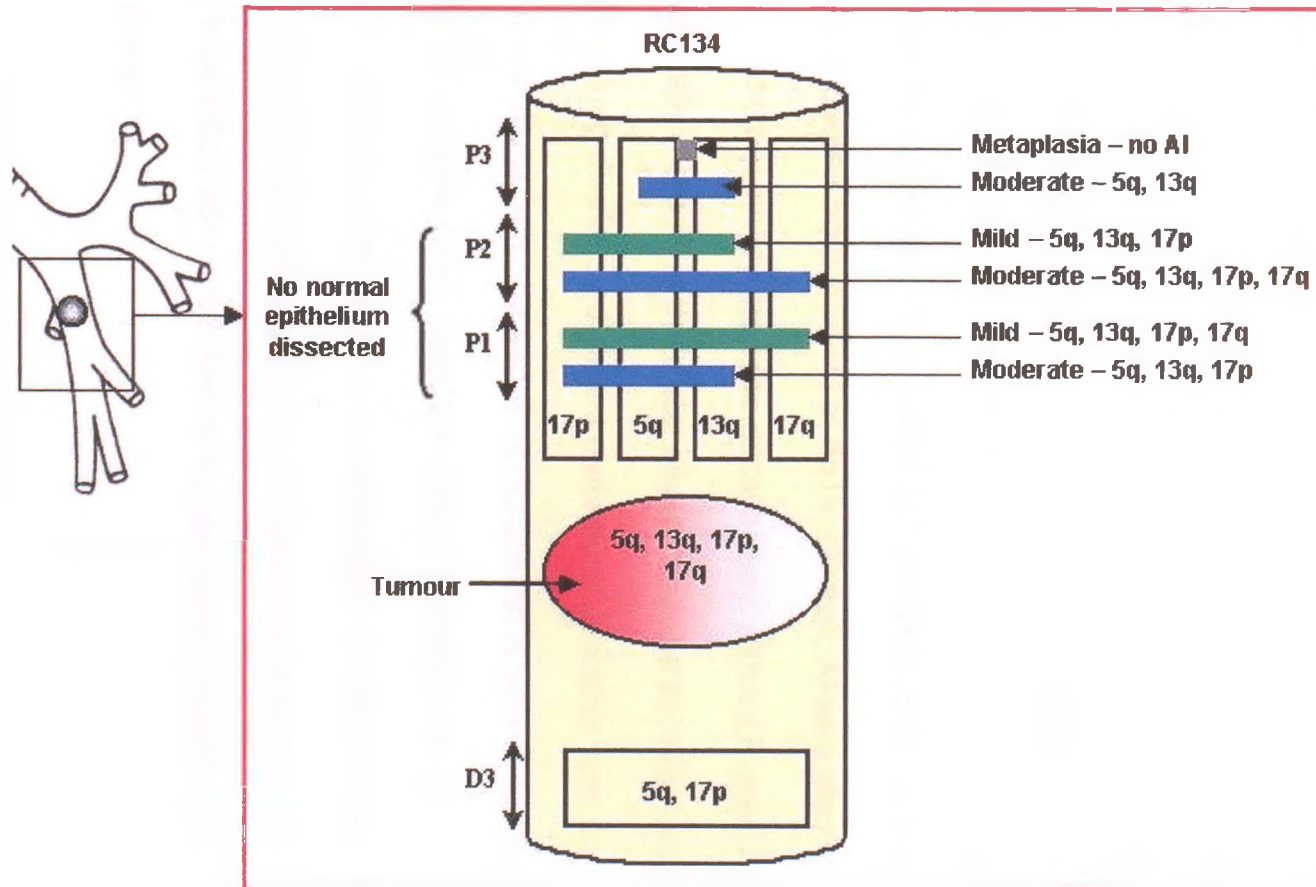


Figure 4.12. Diagrammatic representation of the distribution of AI detected in the tumour and proximal and distal airways for RC134. The airway is represented as a yellow cylinder with the tumour denoted as a shaded red circle located centrally. P1, P2 and P3 denote the different regions of epithelium dissected from the proximal airway and D3 denotes the region of epithelium dissected from the distal airway. Note that histologically normal bronchial epithelium was only analysed from P3 and D3. The AI detected on each chromosomal arm in the tumour of RC134 is indicated within the red shaded circle representing the tumour. The distribution of AI in the proximal and distal airways is represented as rectangles. The chromosomal arms affected in each region of the airway are displayed inside the rectangles. Metaplastic lesions dissected from the proximal airways and the results of AI analysis are indicated on the right of the figure. Key is as follows: Grey box, metaplasia (no atypia); blue bars, moderate atypia; green bars, mild atypia.

Figure 4.12. Diagrammatic representation of the distribution of AI in RC134.



4.4 Conclusion

This chapter has described methods for isolating small subpopulations of cells from bronchial epithelial tissue using laser capture microdissection and direct lysis of cells, followed by AI analysis at multiple microsatellite locations using fluorescent multiplex PCR. AI analysis of sequential regions of bronchial epithelium and tumour in 3 tumour bearing lungs revealed AI in common with the tumour within histologically normal and metaplastic bronchial epithelium at multiple regions throughout the airway. Although individual lungs displayed different patterns of AI, overall AI was more frequent in metaplastic and tumour cells than normal bronchial cells. The highest frequencies of AI was detected at microsatellite loci on chromosome 3p (D3S1263 and D3S1600) at 3p26 and 3p14.2 and chromosome 5q (D5S644) at 5q15. Overall, AI was also more common in the proximal airway than the distal airway, although one lung displayed fractionally more AI in the distal airway than the proximal airway. Individually, each lung displayed different patterns of AI within 1cm intervals and separate cell populations from the same topographical region of the lung displayed different patterns of AI, providing further evidence for field carcinogenesis occurring within NSCLC development. AI analysis also provided an insight into the initial genetic alterations that occurred in each lung, and also the subsequent genetic pathways followed by cell subpopulations from each region in the airway.

**CHAPTER 5: IN-DEPTH ANALYSIS OF MOLECULAR
ALTERATIONS WITHIN NORMAL AND TUMOUR TISSUE FROM
AN ENTIRE BRONCHIAL TREE.**

5.1 Introduction and methods

Previous studies have identified specific chromosomes such as 3p , 9p and 17q as frequent targets for deletion in NSCLC (Field *et al.*, 1996; Fong *et al.*, 1995a; Girard *et al.*, 2000; Hung *et al.*, 1995; Neville *et al.*, 1995; Wistuba *et al.*, 2000a; Wistuba *et al.*, 1999b; Yokota *et al.*, 1987) and also documented abnormalities in genes such as p53 (Liloglou *et al.*, 1997), FHIT (Sozzi *et al.*, 1996b), p16 (Gazzeri *et al.*, 1998) and Rb (Gorgoulis *et al.*, 1998; Tanaka *et al.*, 1998). Several studies have also detected similar molecular abnormalities involving preneoplastic and histologically normal regions of the bronchial epithelium (Gorgoulis *et al.*, 1998; Miozzo *et al.*, 1996; Park *et al.*, 1999; Tanaka *et al.*, 1998; Walker *et al.*, 1994; Wistuba *et al.*, 1999a; Wistuba *et al.*, 2000b; Wistuba *et al.*, 1999b; Wistuba *et al.*, 1997a). Recent research has concentrated on the molecular damage that occurs in the bronchial epithelium of smokers and patients with lung tumours in order to document the earliest genetic changes that may lead to lung cancer. Witsuba *et al* (1999) identified multiple, sequential allele specific molecular changes in the bronchial epithelium of patients with lung tumours. These molecular changes occurred in widely dispersed , clonally independent foci in the lung providing further evidence that genetic alterations accumulate early in the pathogenesis of smoking altered bronchial epithelium. In contrast, recent work by Park *et al* (1999) detected multiple clonal or subclonal patches within the bronchial epithelium of patients with lung cancer and also determined the size, frequency and pattern of the molecular damage within the bronchial epithelium. Early detection of the genetic changes within respiratory epithelium, using techniques such as allelic imbalance analysis, may aid early detection of impending malignancy. Distinguishing specifically deleted loci from background allelic imbalance can allow identification of the specific

genetic changes which cause tumour progression. This can be approached in two ways. First, by the analysis of large numbers of different tumours to find common underlying genetic changes within the tissue, a method which is widely used in current molecular genetic studies. An alternative method, is the analysis of a single tumour and the surrounding tissue in great detail. This chapter describes a combination of allelic imbalance analysis and immunohistochemistry used to construct a detailed topographical molecular map of the bronchial epithelium of a lung from a single patient with a primary bronchial squamous carcinoma, in order to determine the frequency of molecular damage within the entire whole bronchial tree, and to establish the relationship between the molecular damage within the airway and that in the tumour itself.

Samples were taken from a fresh, surgically-resected left lung from a 79 year old man. The lung contained a primary, moderately differentiated, squamous carcinoma, arising from, and invading the wall of, the main bronchus. Thirty-one consecutive slices of the bronchial tree were taken at 1cm intervals starting from the resection margin and following the respiratory tree through both lobes. The tumour itself was also sampled. The tissue blocks were formalin fixed and paraffin wax embedded, and serial 5 μ M sections were cut and stained with haematoxylin and eosin. Microscopic examination by Dr. J. Gosney revealed generally unremarkable bronchial epithelium of respiratory type apart from one small region of squamous metaplasia without atypical features in the anterior airway and one small region of mildly atypical squamous metaplasia within the superior lingular airway of the superior lobe. Hyperplastic type II pneumocytes were also detected in the distal posterior basal lobe.

Samples were laser microdissected as described in chapter 2 using a previously prepared haematoxylin and eosin slide as a guide for microdissection. To optimise each multiplex reaction, different concentrations of each marker primer set were analysed for amplifying DNA from formalin-fixed paraffin embedded tissue. Microsatellites selected for analysis were D3S1300 (3p14.2; FHIT gene), D3S1263 (3p26.1 VHL gene), D5S644 (5q21), D5S2027 (5q21-22), D9S157 (9p23), D9S161 (9p21, p16 gene), D13S153 (13q14, Rb-1 gene), D13S171 (13q12.3-q13, BRCA2 gene), D17S2179E (17p13.1, p53 pentanucleotide repeat), D17S1828 (17p12-p13), D17S1876 (17p12-p13), D17S787 (17q21.3), D17S831 (17p13, p53 gene). Results were analysed using the Genescan™ and Genotyper™ software (PE-Applied Biosystems, Warrington, UK). The sensitivity limits for scoring AI using the fluorescent assay were defined from previous experiments (Liloglou *et al.*, 2000) and are described in chapter 2. Tissue sections were also stained for p53, Cyclin D1 or PCNA protein expression.

5.2 Allelic Imbalance in the tumour and normal bronchial epithelium.

Fluorescent multiplex analysis revealed seven informative loci (D3S1276, D9S157, D9S161, D13S171, D13S153, D17S1828, D17S1876), while the patient was homozygous for the remaining 6 loci. Overall, microsatellite analysis revealed widespread AI in the main bronchus but also in the airways of both upper and lower lobes. In particular AI was detected in both the tumour tissue samples and the adjacent normal tissue, all 3-metaplastic/hyperplastic samples and 23/27 (85%) sites of normal epithelium distal to the tumour. AI was detected at D3S1263 in 16/32 (50%) examined areas, D9S157 in 6/23 (26%), D9S161 in 12/31 (39%), D13S171 in 18/31 (58%), D13S153 in 5/31 (16%), D17S1876 in 10/32 (31%) and D17S1828 in 14/30 (47%) (figure 5.1). In addition, microsatellite alterations (band shifts) were observed in 2 sites for D9S157 and 3 sites for D13S171. AI in at least one microsatellite loci was discovered in 25/30 (83%) normal epithelium sites examined deep into the peripheral lung away from the tumour in both lobes tested. Tumour tissue displayed AI at all loci except D17S1828. AI at D3S1263 was detected at similar frequencies in the region of the main bronchus and tumour and also the upper and lower lobes while D13S17, D9S161, D9S157 and D17S1876 displayed a lower frequency of AI in the distal sites from the tumour (Figure 5.2). AI at locus D13S153 was detected in the tumour and only in two sites of normal epithelium. Interestingly, AI at D17S1828 was detected in normal bronchial epithelium directly adjacent to the tumour and also in the bronchial epithelium of the upper and lower lobes, but not in the tumour itself. AI at D17S1828 was detected at nearly 3 times the frequency in the upper lobe than at the site of the tumour and the lower lobe (Figure 5.2). Fractional allele loss (FAL) values were calculated for each examined site as: $FAL = \frac{\text{no of markers displaying AI}}{\text{total informative}}$

markers]. FAL was statistically higher (0.75 ± 0.13) in the tumour site than in the distal site of the upper (0.42 ± 0.09) and lower lobes (0.31 ± 0.08), indicating gradual accumulation of genetic damage in the tumour area.

Figure 5.1. Allelic imbalance analysis of DNA from multi-sampled lung at multiple locations in a patient with lung cancer. Allelic imbalance pattern in the main bronchus and tumour, the upper lobe and the lower lobe at seven informative loci. The first column indicates the segment/lobe from which tissue was microdissected. 'MB' – main bronchus, 'P' – Posterior, 'IL' – Inferior Lingular, 'SL' – Superior Lingular, 'ANT' – Anterior, 'AB' – Anterior Basal, 'SA' – Superior Apical, 'PB' – Posterior Basal. The numbers 1-5 identify the serial position of the sample within each lobe, with '1' representing the proximal airway and '5' the distal airway. In the second column 'Normal' denotes histologically normal epithelium, 'SCC' - tumour cells, 'M' – metaplasia and 'H' – hyperplasia. Columns 3-5 display the results of immunohistochemical analysis in each tissue section. **Blue fill (+)** – positive staining, **White fill (-)** – negative staining. Columns 6-12 display the results of allelic imbalance analysis in each tissue section. Immunohistochemical antibodies / microsatellite loci are displayed at the top of each column. **Black boxes** – allelic imbalance, **hatched boxes** – microsatellite instability, **white boxes** – heterozygous (no AI). Allele-specific AI is represented by **1** (reduction of smaller allele / amplification of larger allele) or **2** (reduction of larger allele / amplification of smaller allele) in each sample. Blank spaces represent samples where AI could not be determined after repeated attempts.

Figure 5.1. Allelic imbalance analysis of DNA from multi-sampled lung at multiple locations in a patient with lung cancer.

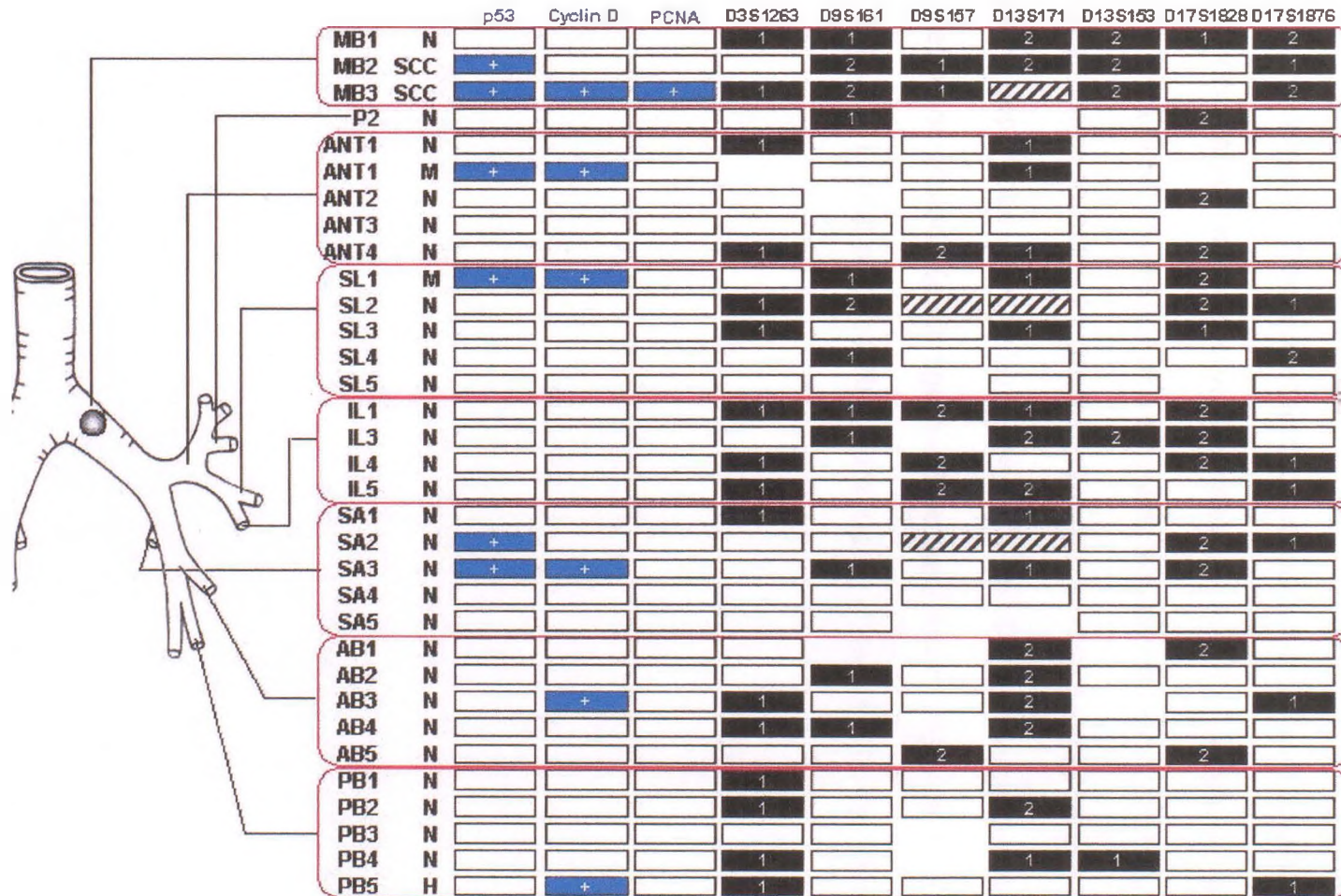
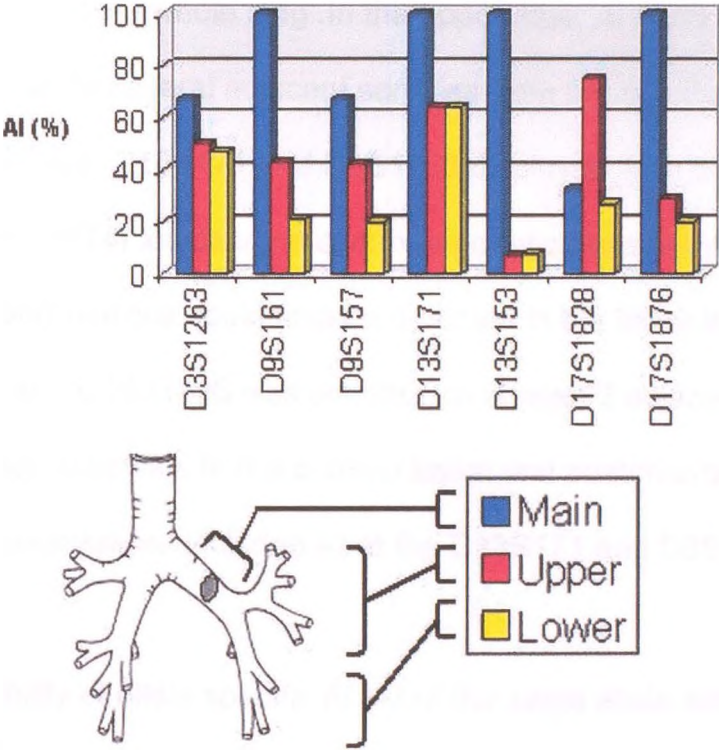


Figure 5.2. Comparison of allelic imbalance within the main bronchus and upper and lower lobes of the left lung at informative microsatellite loci.

Histogram displaying the incidence of AI within the main bronchus at the site of the tumour and in the airways of the upper and lower lobes for each informative microsatellite loci. Vertical axis displays the incidence of allelic imbalance detected at informative loci (allelic imbalance at a single microsatellite loci divided by the number of informative samples for the microsatellite loci within each region i.e. main bronchus, upper lobe, lower lobe of the lung) and the horizontal axis displays the microsatellite loci tested. Key to figure - **Blue bars**: main bronchus and tumour (shown as red circle in lung picture), **Red bars**: upper lobe and **Yellow bars**: lower lobe.

Figure 5.2. Comparison of allelic imbalance within the main bronchus and upper and lower lobes of the left lung at informative microsatellite loci.



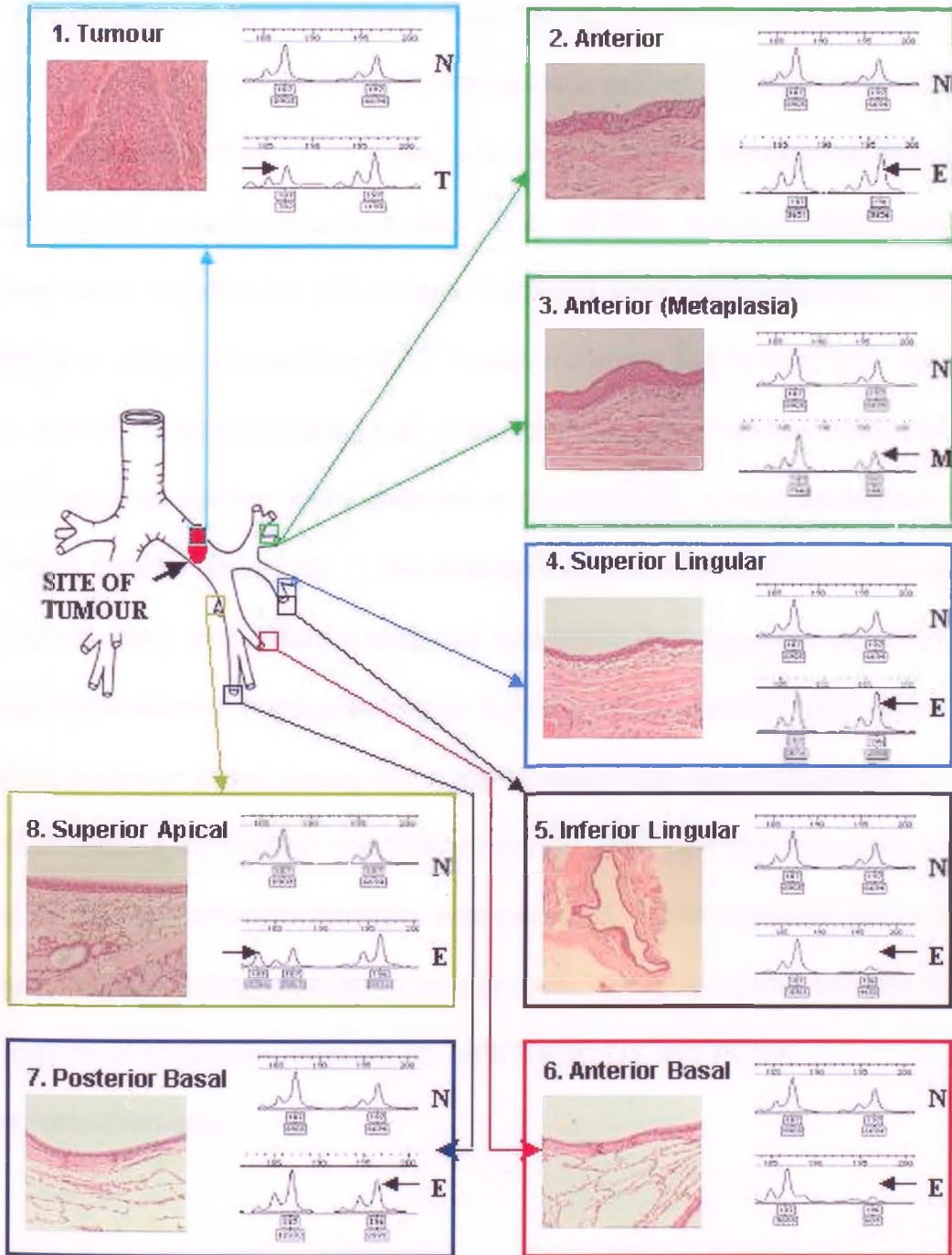
5.3 Allelic specific imbalance patterns and outgrowth clonality.

AI was detected at multiple locations in cells that had been laser microdissected from tissue deep into the lung and distant from the tumour. All areas examined displayed differing patterns of AI at multiple chromosomal locations, although AI at several microsatellite loci was reasonably similar to adjacent samples within the same lobe. On examining allele specificity, only D3S1263 had the same allele (A1) lost throughout the whole lung. In the upper lobe, AI at D17S1828 and D13S171 was common to several adjacent samples from the inferior lingular and superior lingular airways. D13S171 and D3S1263 displayed AI in the anterior airways (ANT1 and ANT4) indicating a common genetic event within this region. Common genetic abnormalities could also be detected in the lower lobe of the lung. AI at D13S171 and D17S1828 was common to at least 2 adjacent samples from the superior apical airway. In the anterior basal and posterior basal airways, common genetic abnormalities included AI at the D13S171 and D3S1263 (Figure 5.3).

The possibility of allele specific AI (AI of the same allele within individual lung segments) was investigated to define whether detected damage in the airways had originated as a clonal event or by individual genetic events. All microsatellite markers except D3S1263 displayed different patterns of allele specific allelic imbalance throughout the lung (Figure 5.1). Two separate tumour samples from the main bronchus also displayed different patterns of AI at D9S161 and D17S1876. Cells dissected from tissue from the two different tumour sites also displayed differing patterns of AI at D3S1263.

Figure 5.3. Allele specific AI occurring throughout the whole lung at microsatellite locus D13S171. A diagrammatic representation of the main airways within the left lung is displayed left of centre. The site of the tumour in the main bronchus is indicated in the figure. A small coloured box indicates in the diagram where each airway was sampled from within the lung. Each box is expanded through an arrow to display a larger box (1-8) of the same colour enclosing the name of the airway in the lung where tissue was laser microdissected. Each box also contains a haematoxylin and eosin stained photomicrograph displaying the histology of the bronchial epithelium microdissected in addition to a Genescan™ analysis of normal control DNA and the matched test DNA for microsatellite locus D13S171. 'N', represents normal; 'T', tumour; 'E', histologically normal bronchial epithelium and 'M', metaplasia with no atypical features. Arrows indicate allele change.

Figure 5.3. Allele specific AI occurring throughout the whole lung at microsatellite locus D13S171.

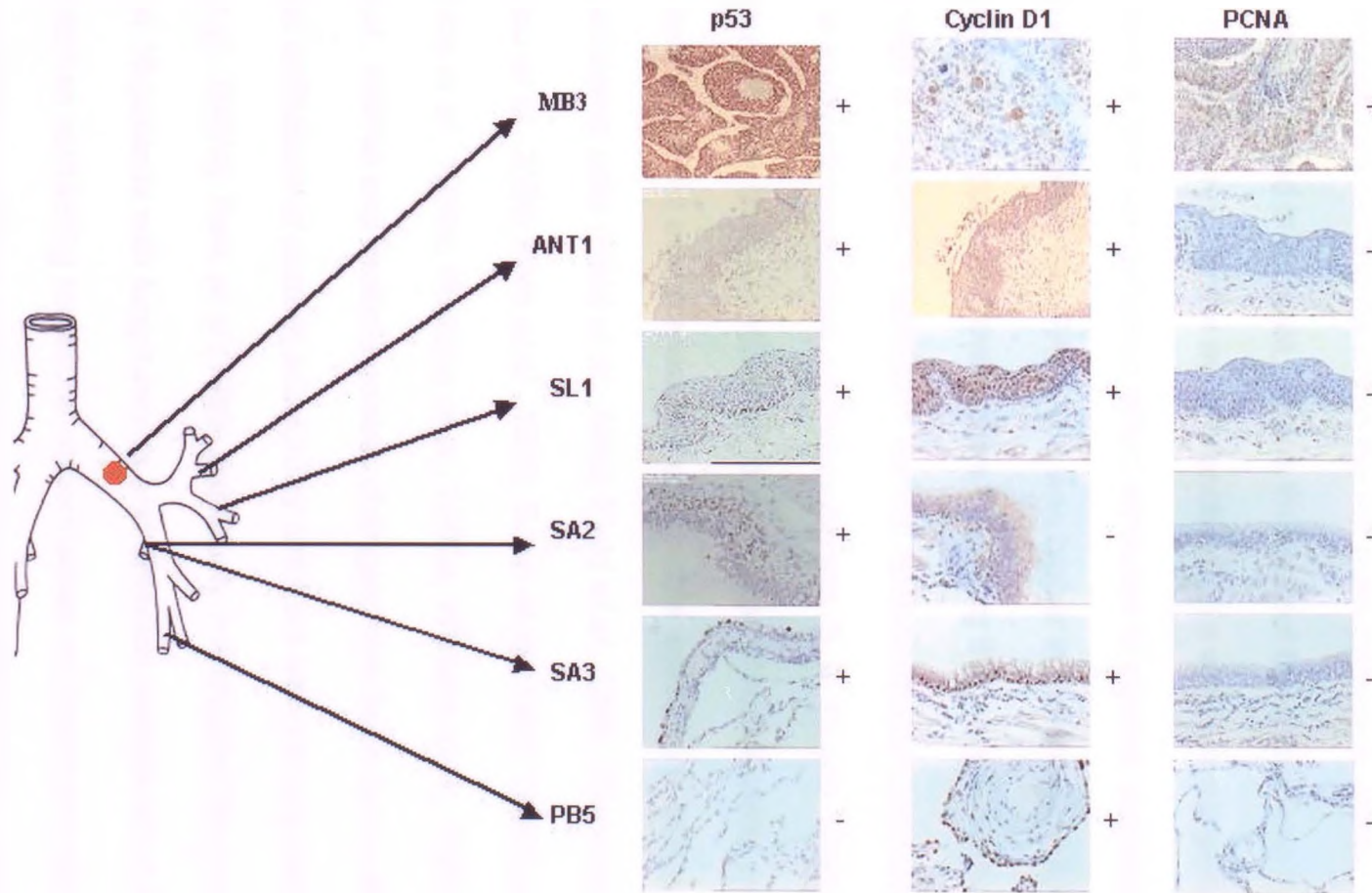


5.4 Expression of p53, Cyclin D1 and PCNA in tumour tissue, normal and metaplastic bronchial epithelium.

p53 immunohistochemistry revealed strong nuclear staining in both tumour samples (MB2 and MB3; figure 5.4). Other regions of positive nuclear immunocytochemistry were detected within normal bronchial epithelium of the superior apical airway (SA2 and SA3), metaplastic epithelium with no atypical features in anterior airway of the upper lobe (ANT1) and mildly atypical bronchial epithelium of the superior lingular airway (SL2). All other samples of bronchial epithelium were negative for p53 protein. CyclinD1 immunohistochemistry revealed no staining in tumour tissue from MB2, however tumour tissue from MB3 displayed nuclear staining in approximately half of the total number of tumour cells. Other regions of positive staining were detected in histologically normal epithelium from the superior lingular airway (SL1), the anterior basal airway (AB3) and the superior apical airway (SA3, in a different region of epithelium that displayed positive p53 immunocytochemistry), hyperplastic type II pneumocytes located within the distal alveoli the posterior basal airway of the lower lobe (PB5) and metaplastic epithelium without atypia from the anterior airway (ANT1), within the same region as positive p53 immunocytochemistry was detected. PCNA immunocytochemistry detected positive nuclear staining in tumour cells from only tumour sample MB3. All other regions of lung were negative for both Cyclin D1 and PCNA immunohistochemistry.

Figure 5.4. Positive p53, Cyclin D1 and PCNA expression within multi-sampled lung. Examples of immunocytochemical detection of p53, Cyclin D1 and PCNA at multiple sections throughout the lung. Arrows indicate the position of the tissue analysed. “**MB**” – main bronchus, “**SL**” – Superior Lingular, “**SA**” – Superior Apical, “**ANT**” – Anterior, “**PB**” – Posterior basal, “**+**”, positive staining; “**-**”, negative staining.

Figure 5.4. Positive p53, cyclin D1 and PCNA expression within multi-sampled lung.



5.5 Discussion.

Multiple adjacent regions of histologically normal bronchial epithelium, metaplasia, metaplasia with mild atypia and tumour tissue were analysed for allelic imbalance at multiple microsatellite loci in the lung of a patient with a primary squamous cell carcinoma within the left main bronchus. Allelic imbalance was detected in multiple regions throughout the entire lung in areas distal to the primary tumour in the main bronchus. This widespread genetic damage reflects a field cancerization effect attributed to chronic carcinogen exposure. It was of interest that FAL value was higher in the tumour site than in the distal sites of the upper and lower lobes indicating that accumulation of genetic damage is required to transform normal epithelium into malignant tissue. In this study microsatellites on chromosomes 3, 9, 13 and 17 were used. Previous studies have identified all four of these chromosomes as targets for genetic loss in normal and preneoplastic bronchial epithelium, and in non-small cell carcinoma cells (Field *et al.*, 1999; Field *et al.*, 1996; Girard *et al.*, 2000; Liloglou *et al.*, 2000; Park *et al.*, 1999; Sozzi *et al.*, 1995; Tsuchiya *et al.*, 1992; Wistuba *et al.*, 1999a; Wistuba *et al.*, 2000a; Wistuba *et al.*, 1999c; Wistuba *et al.*, 1997a) and multiple genetic changes have been demonstrated in the bronchial epithelium of current and former smokers without lung carcinomas (Wistuba *et al.*, 1997a). Park *et al.* (1999) previously investigated the bronchial epithelium in 16 patients with lung tumours and detected multiple small clonal or subclonal patches containing molecular abnormalities on chromosomes 9p, 17q and 3p. In agreement with this study, the results presented in this chapter also indicate multiple small, sometimes clonal patches throughout a whole lung containing a squamous carcinoma, on these three chromosomes in addition to chromosome 13q. Allele-specific AI has previously been detected in lung cancer

and is thought to represent the wide distribution of individual altered clones of cells (Hung *et al.*, 1995; Sidransky, 1995). AI at D3S1263 was detected in every lobe of the lung and AI was allele specific, indicating that genetic damage at this locus may be due to a single initiating event, probably reflecting smoking-related DNA damage. In comparison, AI at the rest of markers examined was also detected in every lobe of the lung tested apart from the posterior airway, however, was not allele specific. In the case of D17S1828, AI was detected in the airways distal and proximal to the tumour, but not the tumour itself. This indicates that AI at this locus may reflect smoking related genetic damage not necessarily related to the tumour progression. In this study AI in both the upper and lower lobes appeared to be concentrated in those segments, which were the first to branch off the main bronchus, and also within the proximal airways of each individual segment. This is reflected in the fact that AI appears to be more common within the initial bronchial airway of the superior lingular segment of the upper lobe and the superior apical and anterior basal segments of the lower lobe.

The mechanism of field cancerization due to the long-term exposure of tissues to carcinogens such as tobacco smoke, both within the bronchial tree and other tissue sites has been proposed by several previous studies as a possible step prior to tumour development (Cense *et al.*, 1997; Garcia *et al.*, 1999; Gazdar, 1994; Scholes *et al.*, 1998; Tian *et al.*, 1998). Further evidence for a field cancerization effect can be seen by studying the distribution of AI at certain microsatellite loci. Allelic imbalance at the microsatellite loci D17S1828 occurred at almost three times the frequency in the bronchial epithelium of the upper lobe than in the epithelium of the lower lobe (Figures 5.1 and 5.2). Cells

microdissected from two regions within the tumour did not display any AI for this loci suggesting that for this specific patient, genetic damage at D17S1828 was limited to the bronchial epithelium and was not involved in the development of this particular tumour.

Immunocytochemical analysis revealed alterations in protein expression of both p53 and Cyclin D1 in bronchial epithelium within the upper and lower lobes of the lung and also in tumour cells. One region of the tumour also displayed overexpression of proliferating cell nuclear antigen (PCNA) protein. p53 and Cyclin D1 have previously been implicated as a frequent target for genetic damage in lung tumours and preneoplastic lung lesions using both immunocytochemical and mutational screening methods for their analysis (Bennett *et al.*, 1993; Betticher *et al.*, 1997a; Betticher *et al.*, 1997b; Boers *et al.*, 1996; Brambilla *et al.*, 1999a; Brambilla *et al.*, 1999b; Caputi *et al.*, 1997; Dosaka-Akita *et al.*, 1994; Hirano *et al.*, 1994; Kurasono *et al.*, 1998; Liloglou *et al.*, 1997; Malusecka *et al.*, 1999; Wakamatsu *et al.*, 1999) although the exact timing and frequency of damage in multistage bronchial carcinogenesis is unclear. This study has shown that both p53 and Cyclin D1 protein are overexpressed in histologically normal bronchial epithelium and in squamous metaplastic epithelium within the same tumour-bearing lung. p53 overexpression was also found in mild atypia and tumour cells while Cyclin D1 overexpression was detected in hyperplastic type II pneumocytes located within the distal alveoli of the lower lobe. This supports evidence that p53 and Cyclin D1 abnormalities can occur early within the multistage bronchial carcinogenesis sequence. Cyclin D1 and PCNA immunocytochemistry also detected overexpression in the tissue section T2 but not in section T1 from a different

region of the tumour. Cyclin D1 has previously been associated with a role in tumour differentiation, displaying overexpression in poorly differentiated tumours whilst being absent in well-differentiated cells (Mate *et al.*, 1996).

Analysis of both tumour sections revealed a heterogeneous population of poorly differentiated (MB3) and well differentiated (MB2) cells, which may account for this difference in expression within the same tumour.

5.6 Conclusion

By taking closely adjacent slices of each portion of the whole lung and using precise laser capture microdissection to isolate small foci of bronchial epithelium, this study has shown that multiple patches of AI occurred in every lobe tested in a patient with lung cancer at sites distant from the tumour itself, possibly due a field carcinogenesis effect within the lung. Multiple clonal cell populations were detected within each lobe of the lung tested and also within the tumour. Allele specific AI at D3S1263 was also detected in every lobe tested in addition to the tumour in the main bronchus indicating that AI at this loci may be an possible initial step in lung carcinogenesis due to smoking-related damage. Building up a detailed topographical picture by the use of large numbers of individual samples of the genetic damage that occurs not only in the tumour itself, but also in the surrounding tissues, is an extremely useful tool in helping us to understand the progress of neoplasia. These techniques may be of relevance to study clonal changes, field cancerization and tumour progression in oral, head and neck, lung and oesophageal carcinogenesis. The implications of such studies have a direct bearing on the interpretation of genetic screening studies using AI, in individuals who are at risk of developing lung cancer and emphasises the necessity for the scientific community to

identify cancer specific genetic instability markers (Liloglou *et al.*, 2001) for early lung cancer detection.

**CHAPTER 6: ANALYSIS OF CHROMOSOME 9p22-p23 IN
HUMAN NON-SMALL CELL LUNG CANCER**

6.1 Introduction

Previous research has already determined that chromosome 9p is a frequent and early target for allelic instability in not only NSCLC and its precursor lesions but also in HNSCC and a wide range of other human neoplasms (Chapter 1). Further studies have attempted to narrow the region of loss on chromosome 9p in human NSCLC using fine mapping techniques. Neville *et al* (1995) investigated LOH on 9p and detected a high frequency of LOH with 44% of tumours within the D9S156-D9S161 region. A minimal region of loss within this group of tumours was detected at 9p23. Mead *et al* (1997) investigated LOH at 9p21, where the cyclin-dependant kinase inhibitors p15^{INK4B} and p16^{INK4} are located, in 32 NSCLC tumours and 3 cell lines using 15 microsatellite markers. A common area of LOH was detected around the marker D9S259 within a 1.7 Mb region for 52% of tumours. LOH at D9S292 (the closest marker to p16^{INK4}) was seen in only 17% of tumours. This suggests that there may be tumour suppressor gene(s) located proximal to p16^{INK4} which may be involved in the pathogenesis of NSCLC. Recent work by Hamada *et al* (2000) again focused on homozygous deletions within the p16^{INK4} region in 149 lung cancer cell lines using 24 DNA markers. This comprehensive study detected homozygous deletions clustered at two independent regions in a quarter of the cell lines. One region contained the p16 tumour suppressor gene and the other, a new target locus, was centred around the microsatellite marker D9S171. This chapter focuses on efforts to minimise the deleted region between D9S259-D9S1778 and includes the region which was detected by Neville *et al* (1995). In this investigation 96 lung tumours were initially screened for allelic instability using 5 microsatellite markers located between D9S269-D9S162. Lung tumours showing partial deletions within this region were then subjected to a more

detailed analysis with further microsatellite markers with an aim to identifying the smallest deleted region. The order of microsatellite markers was confirmed by BAC physical mapping using data available from web-based databases (Sanger ACEDB and Whitehead Institute) and also direct PCR analysis. A list of candidate genes and EST's within the minimal region were obtained and a possible target gene (MLLT3) was chosen for analysis using SSCP.

6.2 Deletion mapping of chromosome 9p22-p23 in human NSCLC

6.2.1 Methods and sample selection

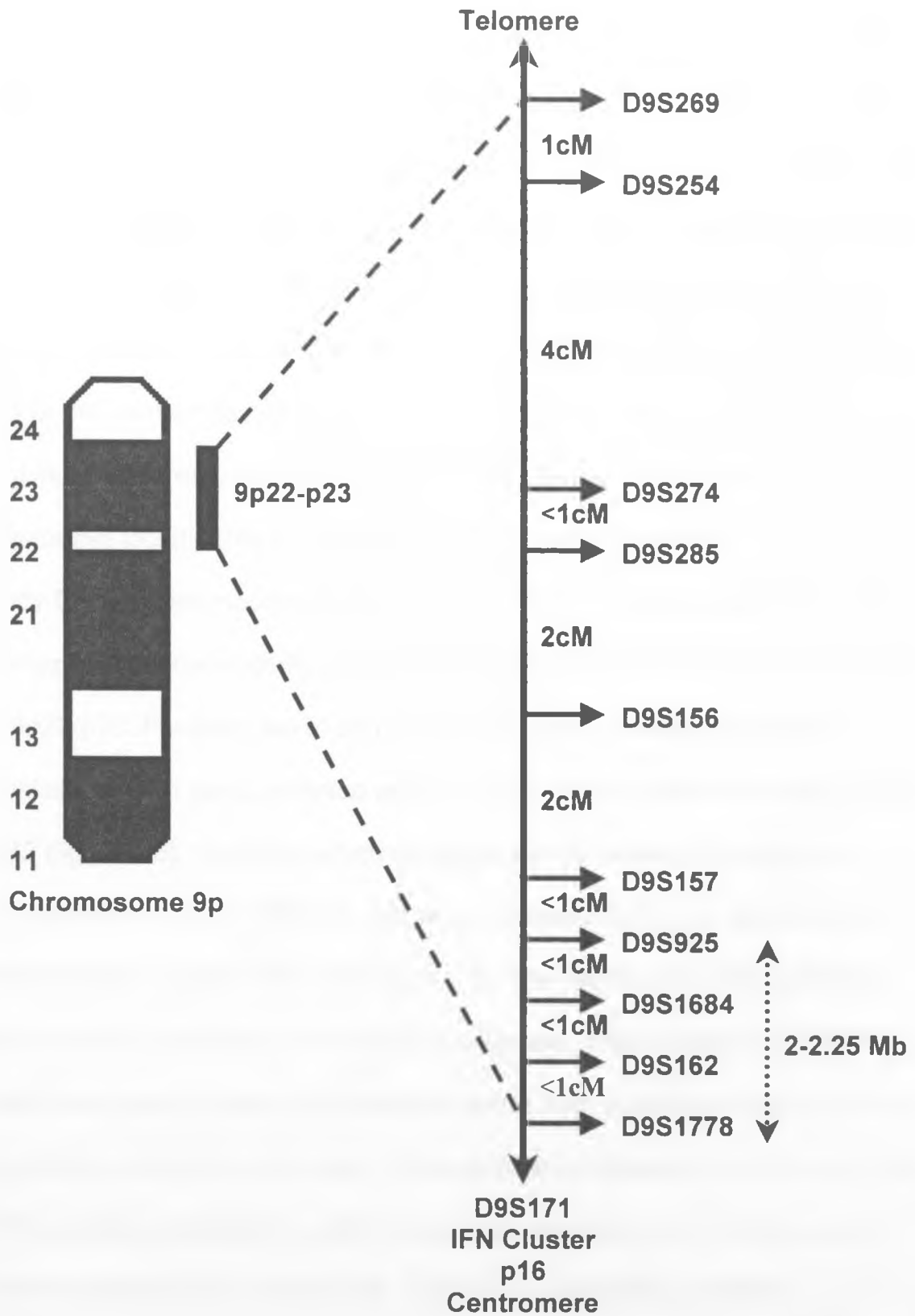
Tumour and normal DNA was isolated from the resected lung tissue of 96 lung cancer patients as described in the materials and methods section.

Microsatellite sequences situated within the region D9S269-D9S162 were investigated using both the Genome Database (GDB) and Whitehead Institute (WI) STS search engines. Suitable microsatellites within and around the region of D9S157 were chosen for analysis based on the availability of the microsatellite sequence and also the average heterozygosity of the microsatellite sequence if available. Once selected for analysis, primer sequences flanking the microsatellite repeat were redesigned using the Oligo™ program as described in Chapter 2 (2.2.1). Forty tumour and normal DNA pairs were excluded from the study because of histology and poor DNA quality leaving a panel of 56 NSCLC tumours. The panel of 56 tumours were analysed for allelic instability at 5 polymorphic repeat sequence locations spanning chromosome 9p22-23 using PCR followed by PAGE analysis as described in Chapter 2 (2.2.3). This original deletion map using 5 polymorphic microsatellite markers was designated deletion mapping region 1 (DM1). The study was expanded to analyse 9 microsatellite loci in 14 tumours which displayed partial

deletions within this region and designated deletion mapping region 2 (DM2). Eight tumours which displayed partial deletions in and around the microsatellite markers D9S1684 and D9S162 were also analysed using the marker D9S1778. The position and order of microsatellite markers on chromosome 9p23 is displayed in figure 6.1.

Figure 6.1. Diagram displaying the position and order of polymorphic microsatellite markers used for deletion mapping studies on chromosome 9p22-p23. The position and order of 10 polymorphic microsatellite markers used for deletion mapping 9p22-p23 from D9S269 – D9S1778 is displayed on the right of the figure with the approximate genetic distance in centimorgans (cM) displayed between each marker (genetic distances obtained from Genemap '99 and Whitehead Institute databases). The distance between D9S1778 and D9S925 was estimated from *HindIII* fingerprinting across the region (personal communication with Dr. Sean Humphrey, Sanger Centre, UK. Oct 2000).

Figure 6.1. Diagram displaying the position and order of polymorphic microsatellite markers on chromosome 9p22-p23.



6.2.2 Frequency of AI on chromosome 9p23 in NSCLC

Fifty-six tumours were informative for 1 or more microsatellite markers tested on chromosome 9p22-p23 (table 6.1; figures 6.2, 6.3, 6.4 and 6.5). Overall microsatellite analysis revealed frequent AI (121/225, 53%) for 10 informative microsatellite loci in 56 NSCLC's in deletion mapping studies DM1 and DM2. Table 6.1 displays the frequency of AI detected at individual microsatellite loci. In the initial deletion mapping experiment using 5 microsatellite markers (DM1) the highest frequency of deletion was detected at loci D9S156 (71%) and D9S157 (68%). Thirteen out of 56 (24%) tumours displayed no detectable AI at any of the markers tested (figure 6.2) indicating retention of the 9p22-p23 region. Forty-three out of 56 (76%) tumours displayed AI at one or more informative loci (figures 6.3, 6.4 and 6.5). Of the tumours which displayed AI in study DM1 on chromosome 9p23, over half (29/56; 51%) displayed AI or were homozygous at the majority of loci tested indicating deletion of the whole region of 9p22-p23. Fourteen out of 56 (25%) tumours which displayed partial deletions in DM1 were analysed with a further 4 microsatellite markers in study DM2 (figure 6.5). Samples which displayed partial deletions around the microsatellite marker D9S162 (n=8) were analysed further for AI using the microsatellite marker D9S1778 (figure 6.5). The relationship between AI at chromosome 9p22-p23, morphological diagnosis, differentiation and TNM status was investigated using details obtained from pathology reports and no significant correlation was found. The relationship between AI at 9p22-p23 and other genetic abnormalities (p53 immunocytochemistry and mutations, p21 immunocytochemistry, nuclear Rb-1 immunocytochemistry, hMLH1 immunocytochemistry and hMSH2 immunocytochemistry. p53, p21, Rb-1, hMLH1 and hMSH2 immunocytochemistry data was obtained from previous

experiments performed at the RCICLCR (Liloglou *et al.*, 1997; Xinarianos *et al.*, 2000)) . These analyses revealed a weak association between locus D9S254 and hMLH1 (n=35) and hMSH2 (n=35) positive immunocytochemistry (p=0.037 and p=0.017 respectively) and between locus D9S285 and hMLH1 positive immunocytochemistry (p=0.016, n=15).

Table 6.1. Table displaying frequency of tumour samples which displayed AI at informative loci tested. The first column displays the microsatellite loci tested and the second column the percentage of informative tumour samples which displayed AI. Note that the microsatellite markers D9S269, D9S274, D9S285 and D9S156 were only used to analyse 14 tumours for AI in study DM2 and D9S1778 used to analyse 8 tumours which displayed a partial deletion around D9S162.

Marker Name	Informative samples displaying AI
D9S269	2/6 (33%)
D9S254	9/26 (34%)
D9S274	5/11 (45%)
D9S285	5/10 (50%)
D9S156	5/7 (71%)
D9S157	35/51 (68%)
D9S925	20/39 (51%)
D9S1684	22/36 (61%)
D9S162	16/32 (50%)
D9S1778	2/7 (28%)



Figure 6.2. Figure showing retention of the region 9p22-p23 in 13/56 informative lung tumours from DM1. Paired normal and tumour DNA was analysed for allelic instability using 5 polymorphic microsatellite markers. Sample codes are displayed at the top of the figure and the microsatellite markers used for allelic instability analysis are displayed at the left side of the figure. Key for figure is as follows: Blue circle  heterozygous (No AI); Grey circle  homozygous.

Figure 6.2. Figure showing retention of the region 9p22-p23 in 13/56 informative lung tumours from DM1.

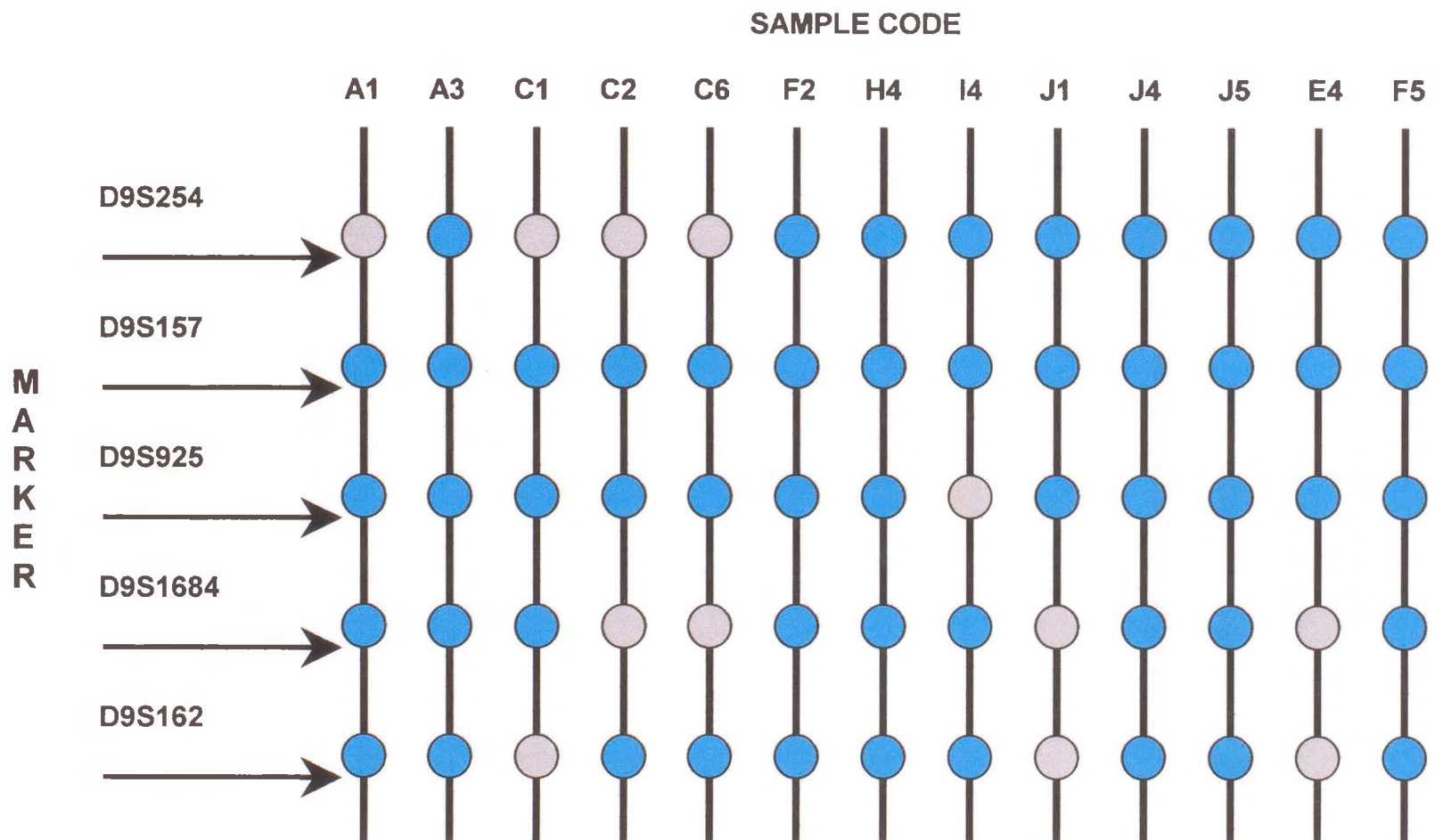


Figure 6.3. Figure showing AI of the region 9p22-p23 in 16/56

informative lung tumours from DM1. Paired normal and tumour DNA was analysed for allelic instability using 5 polymorphic microsatellite markers.

Sample codes are displayed at the top of the figure and the microsatellite markers used for AI analysis are displayed at the left side of the figure. Key

for figure is as follows: Red circle ● allelic instability; Grey circle ○ homozygous; Yellow circle ○ microsatellite instability (allele shift).



Figure 6.3. Figure showing AI of the region 9p22-p23 in 16/56 informative lung tumours from DM1.

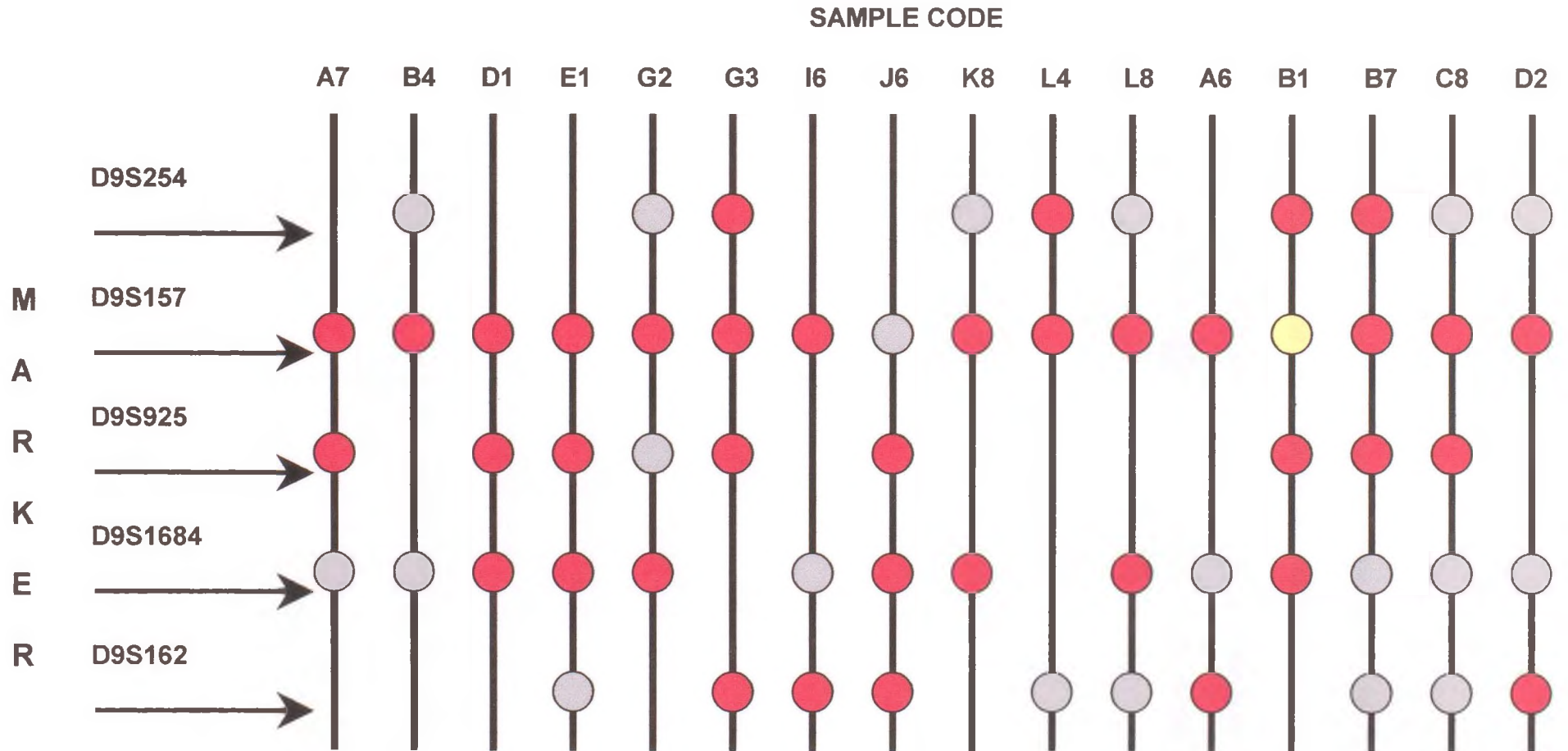


Figure 6.4. Figure showing AI of the region 9p22-p23 in 13/56

informative lung tumours from DM1. Paired normal and tumour DNA was analysed for allelic instability using 5 polymorphic microsatellite markers.

Sample codes are displayed at the top of the figure and the microsatellite markers used for AI analysis are displayed at the left side of the figure. Key

for figure is as follows: Red circle ● allelic instability; Grey circle ○ homozygous;

Figure 6.4. Figure showing AI of the region 9p22-p23 in 13/56 informative lung tumours from DM1.

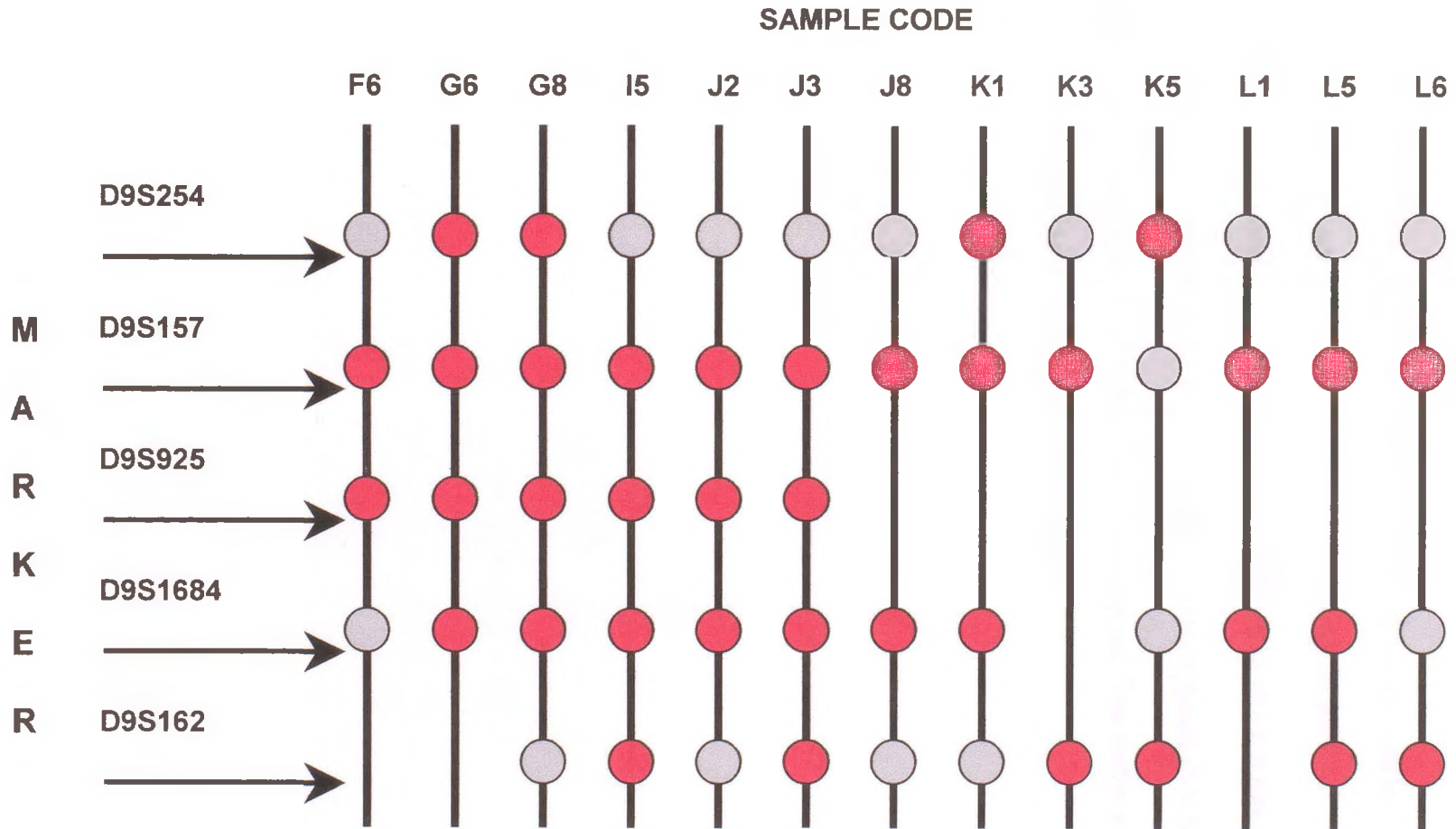





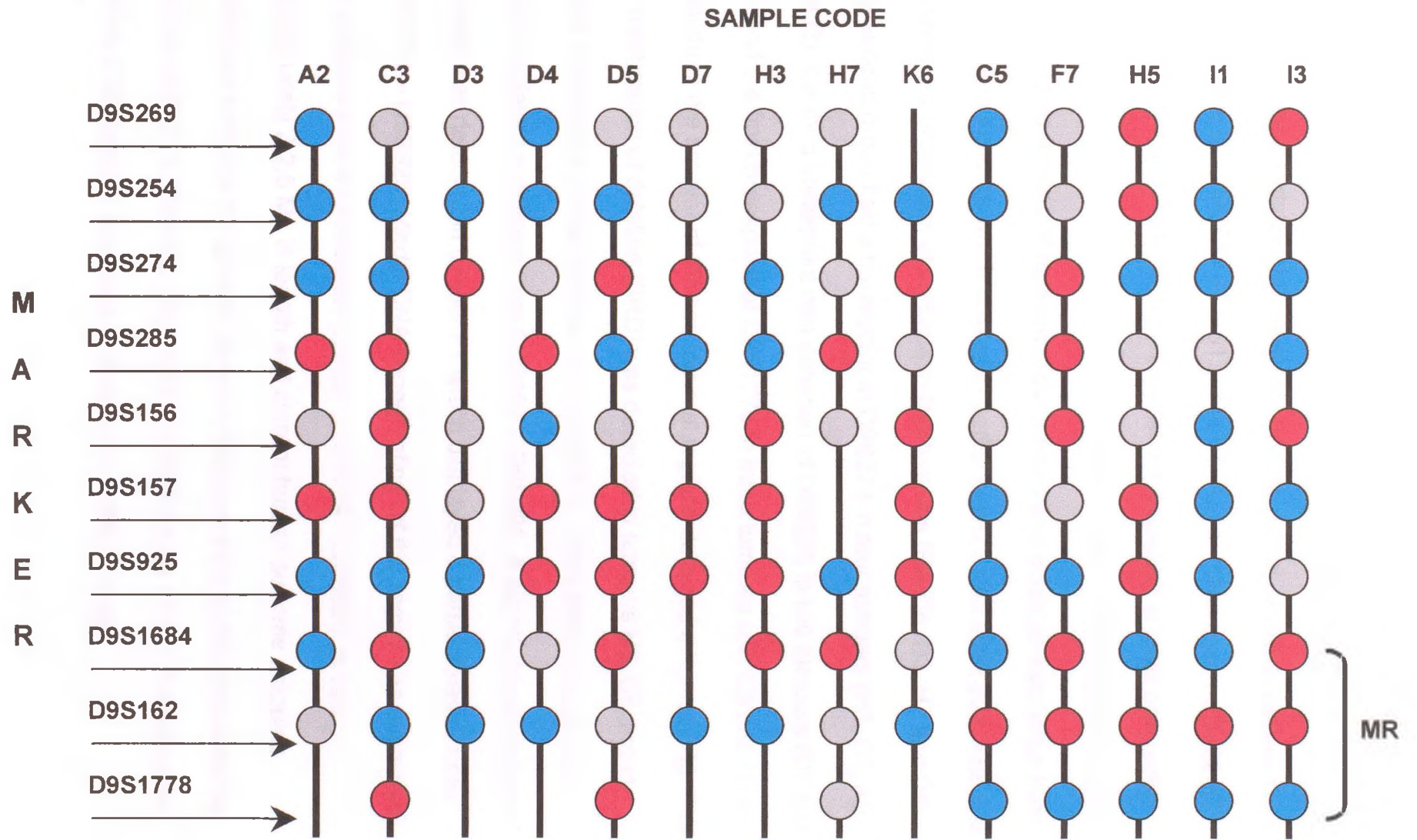
Figure 6.5. Figure showing partial deletions of the region 9p22-p23 in 14/56 lung tumours from DM1 and DM2. Paired normal and tumour DNA was analysed for allelic instability using 10 polymorphic microsatellite markers. Sample codes are displayed at the top of the figure and the microsatellite markers used for AI analysis are displayed at the left side of the figure. Key for figure is as follows: Blue circle  heterozygous (No AI); Red circle  allelic instability; Grey circle  homozygous. The minimal region of AI (MR) detected is indicated in the figure.

Figure 6.5. Figure showing partial deletions of the region 9p22-p23 in 14/56 lung tumours from DM1 and DM2.



6.2.3 Multiple partial deletions in NSCLC on chromosome 9p22-p23

Deletion mapping study DM2 revealed multiple partial deletions between the microsatellite loci D9S269 – D9S1778 in NSCLC on chromosome 9p22-p23 (figure 6.5, 6.6 and 6.7). All fourteen tumours displayed at least one partial region of deletion of varying length. Four tumours (C3, D4, H5 and I3) also displayed two regions of deletion. Four breakpoints could be estimated from the fourteen tumours which displayed partial deletions on 9p22-p23. The first group had a breakpoint at D9S254 in five tumours (K6, D4, D3, H7 and D5). The second group had a breakpoint at D9S274 in four tumours (H5, C3, A2 and I3), the third breakpoint was detected at D9S285 in two tumours (D7 and H3) and the final breakpoint at D9S1778 in three tumours (F7, C5 and I1) in addition to the smaller of the two minimal regions detected in tumour H5. The smallest region of deletion (SRD) was detected in tumours H5, C5 and I1 which displayed a partial deletion from D9S1778 – D9S1684 and also C3 which displayed a deletion from D9S162 – D9S925. A slightly larger region of deletion was detected in tumour F7 which displayed a partial deletion from D9S1778 – D9S925. *HindIII* DNA fingerprinting of this region has estimated the distance from microsatellite marker D9S1778 – D9S925 as being approximately 2-2.5 Mb in length and current human genome mapping techniques estimate the genetic distance between each of the microsatellite markers D9S1778, D9S162, D9S1684 and D9S925 as <1cM. This indicates that the SRD found in this study is likely to be <2Mb in length.

Figure 6.6. Figure displaying partially deleted regions detected between the microsatellite loci D9S269 – D9S1778. Figure displaying minimal regions of deletion detected in 14 NSCLC tumours in deletion mapping studies DM1 and DM2. The location of microsatellite loci on chromosome 9p22-p23 is displayed on the left of the figure and extended into the figure using hashed lines. Deletions are represented by solid black bars. Note that tumours D4, H5 and C3 are represented by two solid black lines representing two regions of deletion. Tumour D4 displays retention at D9S156, H5 at D9S1684 and C3 at D9S925.

Figure 6.6. Figure displaying partially deleted regions detected in 14 NSCLC tumours between the microsatellite loci

D9S269 – D9S1778.

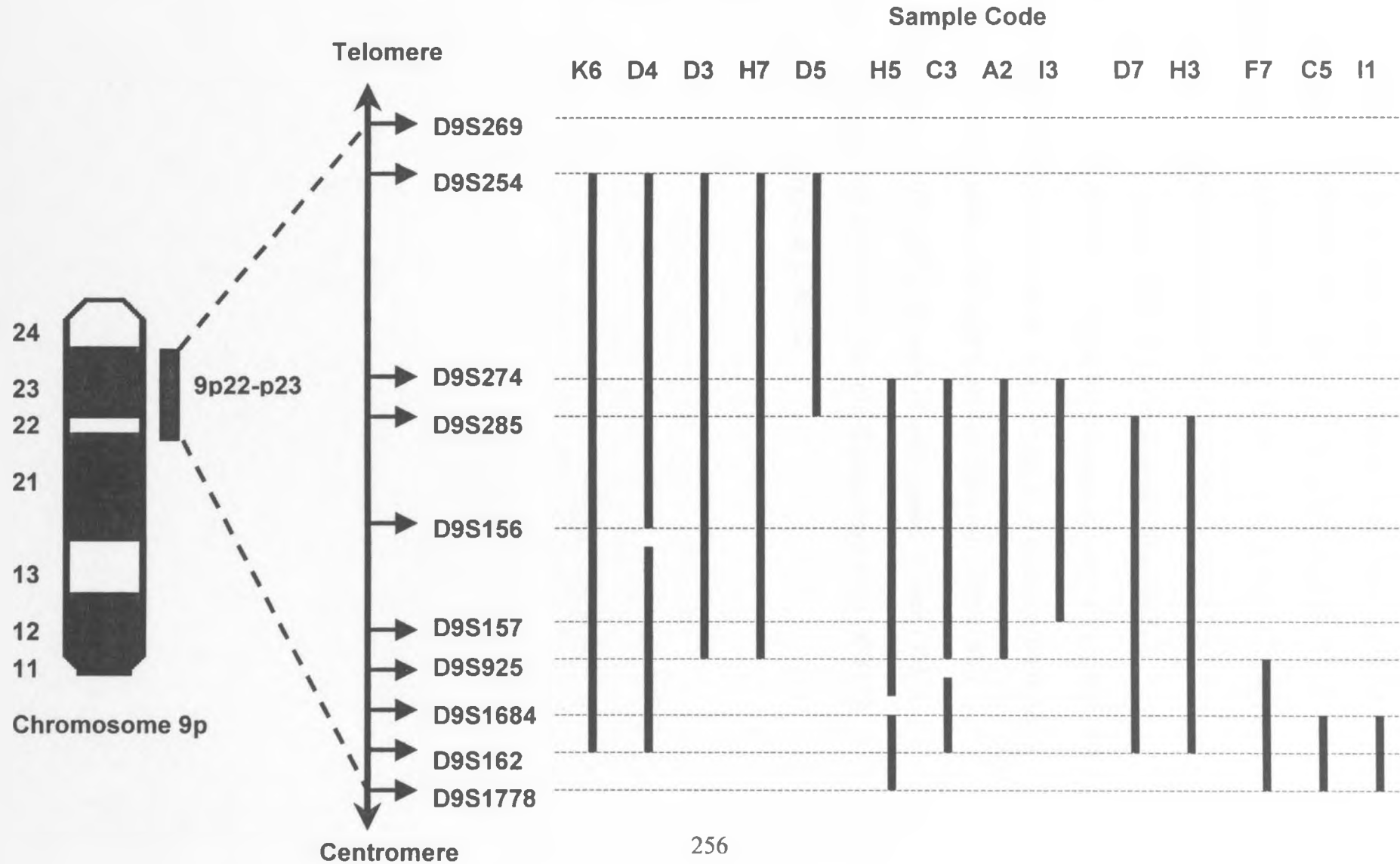
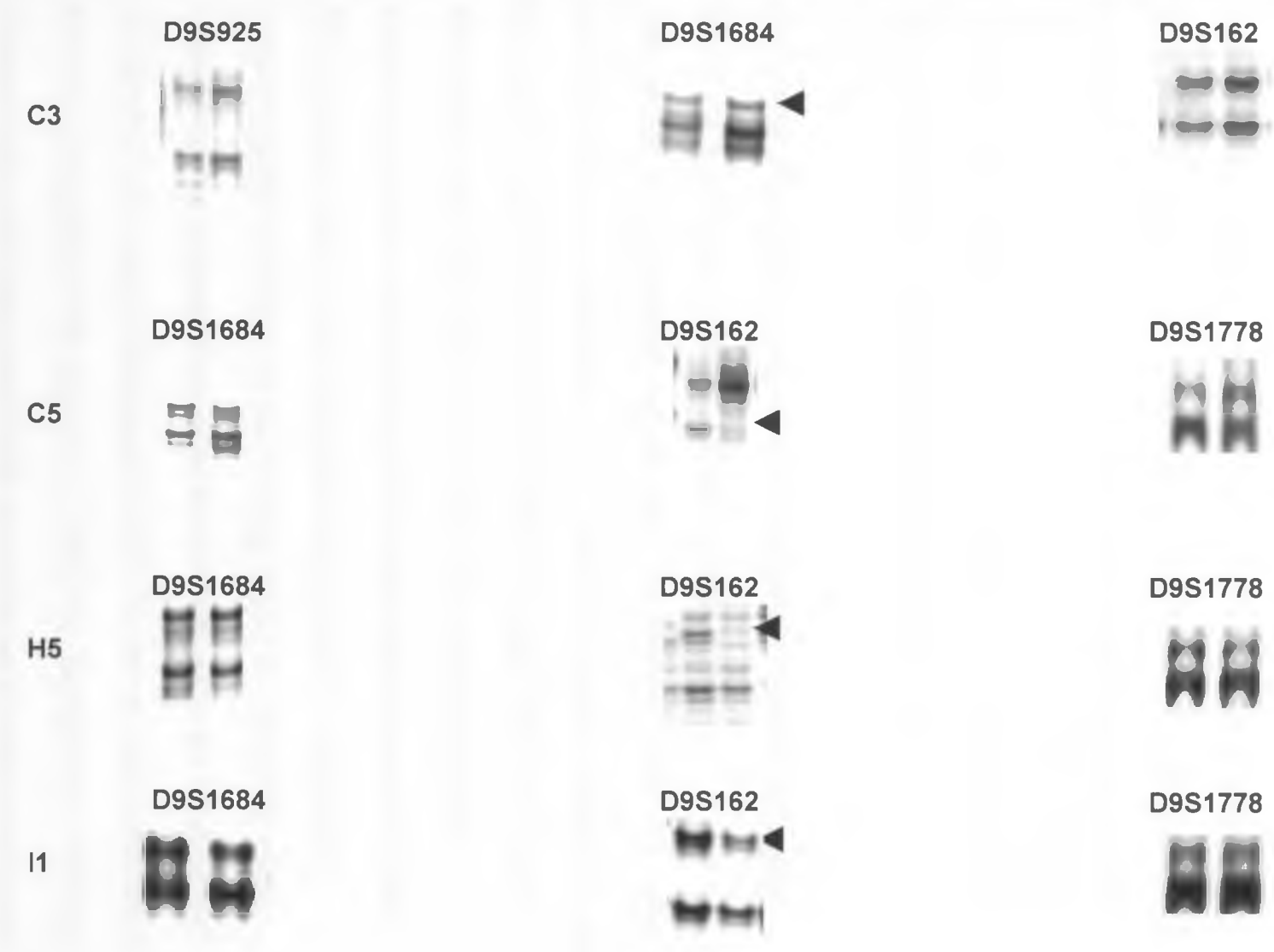


Figure 6.7. Photomicrographs displaying minimal regions of deletion between D9S925 and D9S1778. The results of deletion mapping at four microsatellite loci (displayed at the top of the figure) in 4 tumours (displayed at the left of the figure) which displayed partial deletions. Normal DNA is shown in the left lane and tumour DNA in the right lane within each photomicrograph. Arrows indicate a decrease/increase in band intensity in the tumour DNA reflecting allelic imbalance. Allelic imbalance was scored positive if a 30% difference in the allele ratio was detected between normal and tumour DNA by scanning the optical density of each allele. Tumour C3 displays AI at D9S1684 and D9S1778. Tumour C5 displays AI at D9S162. Tumours H5 and I1 display AI at D9S162 .

Figure 6.7. Photomicrographs displaying minimal regions of deletion between D9S925 and D9S1778.



6.3 Search for a putative tumour suppressor gene involved in NSCLC tumourigenesis on chromosome 9p22-p23

Several minimal regions of deletion were identified within the microsatellite loci D9S269 – D9S1778 by AI analysis in 14/56 NSCLC. This indicated the presence of putative tumour suppressor genes within these regions which play a role in the development of NSCLC. In this study the minimal region between microsatellite loci D9S1778 – D9S1684 was selected for further detailed analysis due to the small size of the region (<1 cm).

6.3.1 Incomplete YAC positioning of microsatellite loci within the deleted region using the WI and CEPH – Genethon databases

The position of microsatellite markers within the D9S269 – D9S1778 region was confirmed by a combination of YAC and BAC analysis. Initially YAC data was used to confirm the microsatellite order according to the WI database. The WI database was queried in December 1999 for microsatellite information and their corresponding YAC mapping data (appendix C, table C1). Several problems were encountered when using this database to correctly position microsatellite loci involving absent or incomplete marker information. The microsatellite marker D9S254 was not in the WI database and the final YAC colony addresses for D9S1778 had not been confirmed at the time of December 1999 (appendix C, table C2). Several YAC addresses containing both D9S1684 and D9S162 were detected as both unambiguous (confirmed) and disambiguous (unconfirmed) hits, however no hits could be

confirmed for D9S1778. The CEPH-Genethon database was also queried for all the microsatellites within the D9S269 – D9S1778 region (appendix C, table C3). As with the WI database, several microsatellite loci were also absent from the CEPH -Genethon database (D9S254, D9S925, D9S1684 and D9S1778). YAC addresses were confirmed for all other loci including a YAC containing the 3 microsatellite loci D9S274, D9S285 and D9S156 (952_g_4). PCR amplification of YAC DNA was undertaken to confirm the marker order however, results proved unreliable. YAC addresses containing two or more microsatellite loci or other STS markers were not represented in both the CEPH and WI databases for the D9S1684-D9S1778 region, preventing mapping of overlapping YAC vectors within this region. As D9S1778 flanked the minimal region of deletion detected in this study, and was not represented within both queried databases this would present difficulties in mapping the deleted region (SRD) and confirming correct microsatellite marker order. A BAC mapping approach was subsequently chosen to confirm microsatellite order.

6.3.2 Positioning of BACs and microsatellite loci within the SRD using the Sanger ACEDB database and BAC PCR.

Multiple BAC hits were detected for all microsatellite loci within the region D9S269 – D9S1778 using the Sanger Centre ACEDB database. BAC hits corresponding to microsatellite markers within the region of D9S1684 – D9S1778 were arranged into a physical map of the SRD in February 2000

(figure 6.8). Initially marker order within the region of D9S1684 - D9S1778, according to the WI database which was originally analysed for correct loci positioning, displayed the microsatellite order as D9S925 – D9S162 – D9S1684. ACEDB BAC mapping data within the region conflicted with the WI and CEPH-Genethon databases and positioned the microsatellite loci as D9S925 – D9S1684 – D9S162. It was evident from both YAC and BAC data that D9S1684 and D9S162 lie in close proximity to one another, however their particular order within the region was disputed by recent databases. As YAC data from this region was insufficient to construct a physical map from D9S925 – D9S1684/D9S162 using both the WI and CEPH-Genethon databases, and multiple BAC contigs were available for the region D9S925 – D9S162, the ACEDB BAC data was used for correct positioning. Additional sequencing data of the BAC containing D9S1684 and D9S162 (bA252O2) was also available from the Sanger ACEDB which confirmed microsatellite loci order within this region using the BLAST2 sequence alignment database (D9S1684 was positioned from nucleotide 8638 – 8859 bp and D9S162 was positioned from nucleotide 72670 – 72887 bp on BAC bA252O2. Additional evidence for the correct positioning of microsatellite loci was confirmed by microsatellite PCR analysis using BAC DNA from the SRD as a template (figure 6.9).

6.3.3 Candidate gene and EST search within SRD

The existence of a SRD between the microsatellite loci D9S1684 – D9S1778 indicated that putative genes(s) involved in the initiation or progression of NSCLC may reside within this region. A list of candidate genes and EST's was compiled from the Genemap '99 database in February 2000 (appendix B, table B4). Although the exact positioning of the EST's and candidate genes within the region could be questioned (for example the p16 and MTAP genes were originally incorrectly placed within the region when they have been mapped several centimorgans distal to the SRD (Nobori *et al.*, 1996)), BAC mapping data had located several EST's and candidate genes within the SRD. One such candidate gene selected for analysis was the Myeloid/lymphoid or Mixed-Lineage Leukaemia (*Drosophila Trithorax* homologue); translocated to, 3 (*MLLT3*) gene which was situated distal to D9S162.

6.3.4 Database analysis of the MLLT3 gene

The 3376bp MLLT3 mRNA sequence (GI:4758719) was retrieved from the NCBI Entrez Nucleotide Search database and analysed using a Basic BLAST sequence alignment database. BLAST search revealed homology with 11 database entries including *MLLT3* mRNA (full length and C-terminal), Human *AF-9* mRNA, Chromosome 9p22 Cosmid Clones 213a7, 92f5, 34a5 and c48, *Rattus norvegicus* mRNA for *AF-9*, *Homo sapiens* myeloid/lymphoid or mixed-lineage leukaemia (*trithorax (Drosophila)* homolog); translocated to, 1

(*MLLT1*) mRNA, human germline *ENL* mRNA and Human mRNA for *LTG19*. Intron/ Exon boundaries for *MLLT3* were confirmed by comparing the full length *MLLT3* mRNA sequence with the genomic DNA sequence contained in cosmid clones 213a7, 92f5, 34a5 and c48 using the BLAST2 database (Appendix C) with the exception of exon 3 which was not detected within any of the cosmid clones. In total 10 exons were detected and their positions are summarised in table 6.6 and figure 6.10. The open reading frame (ORF) of the *MLLT3* gene was determined by analysis using the NCBI ORF finder database (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>) . The initiation and stop codons were detected at nucleotide positions 196 and 1900 respectively and a poly(A) signal (AATAAA) at nucleotide 2032 within *MLLT3* mRNA. The *MLLT3* gene translates into a 568aa protein sequence and contains a TF2F conserved region from nucleotides 208-513bp.

Figure 6.8 BAC physical map for the region D9S1778 – D9S925 on chromosome 9p23. A figure representing the order of microsatellite loci and BAC constructs within the region of D9S1778 – D9S925. The double headed blue arrow represents the genomic sequence of chromosome 9p22-23 from centromere to telomere. Microsatellite loci and STS sequences are displayed in blue also. Hatched lines indicate the location of certain microsatellite markers within BAC colony addresses mapped to the region by computer (ePCR) or 'real' PCR analysis. BAC constructs are shown as black lines in the figure, adjacent to the BAC ID. The STS's stSG43 and D9S1778 are intronic sequences of the MLLT3 gene on chromosome 9p22-23 within the SRD.

Figure 6.8. BAC physical map for the region D9S1778 – D9S925 on chromosome 9p23

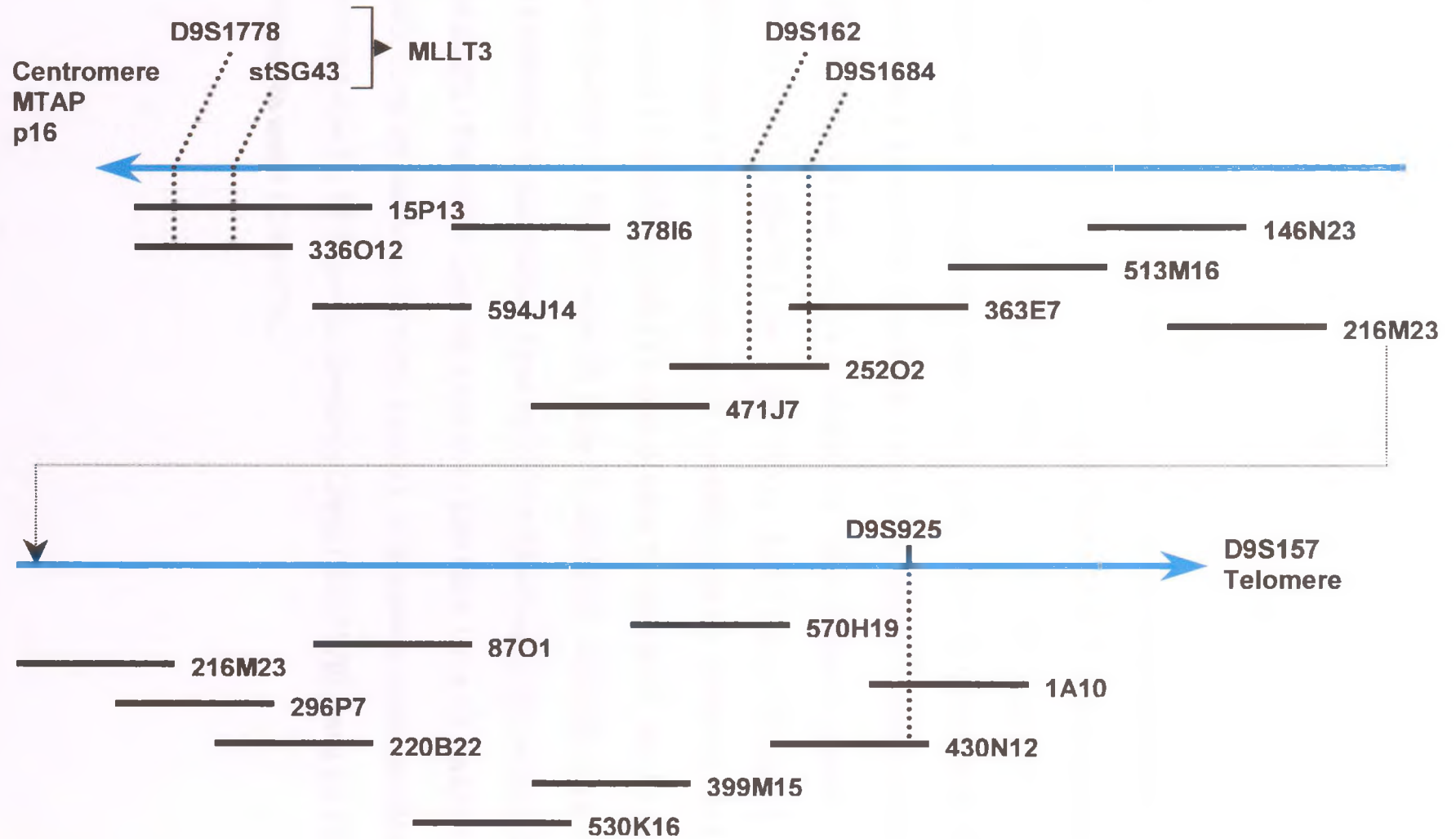


Figure 6.9. Photomicrograph displaying agarose gel electrophoresis of BAC DNA PCR products to confirm the position of microsatellite markers and the MLLT3 gene within the region D9S925 – D9S1778. BAC colonies mapping to the region D9S925 – D9S1778 were grown overnight in liquid cultures then DNA extracted by alkaline lysis methods. The order of microsatellites within this region and the exact location of the MLLT3 was confirmed by PCR analysis followed by electrophoresis through a 3% agarose gel. Lane order is numbered at the top of each well on both the upper and lower rows. Lane 1: ϕ X174 DNA/*Hae* III size marker; Lane 2: bA336O12 / D9S1778; Lane 3: bA336O12 / stSG43; Lane 4: bA336O12 / MLLT3 exon 4; Lane 5: bA336O12 / MLLT3 exon 6; Lane 6: bA336O12 / MLLT3 exon 9; Lane 7: bA336O12 / MLLT3 exon 10; Lane 8: bA15P13 / D9S1778; Lane 9: bA15P13 / stSG43; Lane 10: bA15P13 / MLLT3 exon 4; Lane 11: bA15P13 / MLLT3 exon 6; Lane 12: bA15P13 / MLLT3 exon 9; Lane 13: bA15P13 / MLLT3 exon 10; Lane 14: bA252O2 / D9S162; Lane 15: ϕ X174 DNA/*Hae* III size marker; Lane 16: ϕ X174 DNA/*Hae* III size marker; Lane 17: bA252O2 / D9S1684; Lane 18: bA363E7 / D9S1684; Lane 19: bA430N12 / D9S925; Lane 20: bA1A10 / D9S925; Lane 21: PCR control (Genomic DNA1 / D9S1778); Lane 22: PCR control (Genomic DNA2 / D9S1778); Lane 23: PCR control (sterile water / D9S1778).

Figure 6.9. Photomicrograph displaying agarose gel electrophoresis of BAC DNA PCR products to confirm the position of microsatellite markers and the MLLT3 gene within the region D9S925 – D9S1778.

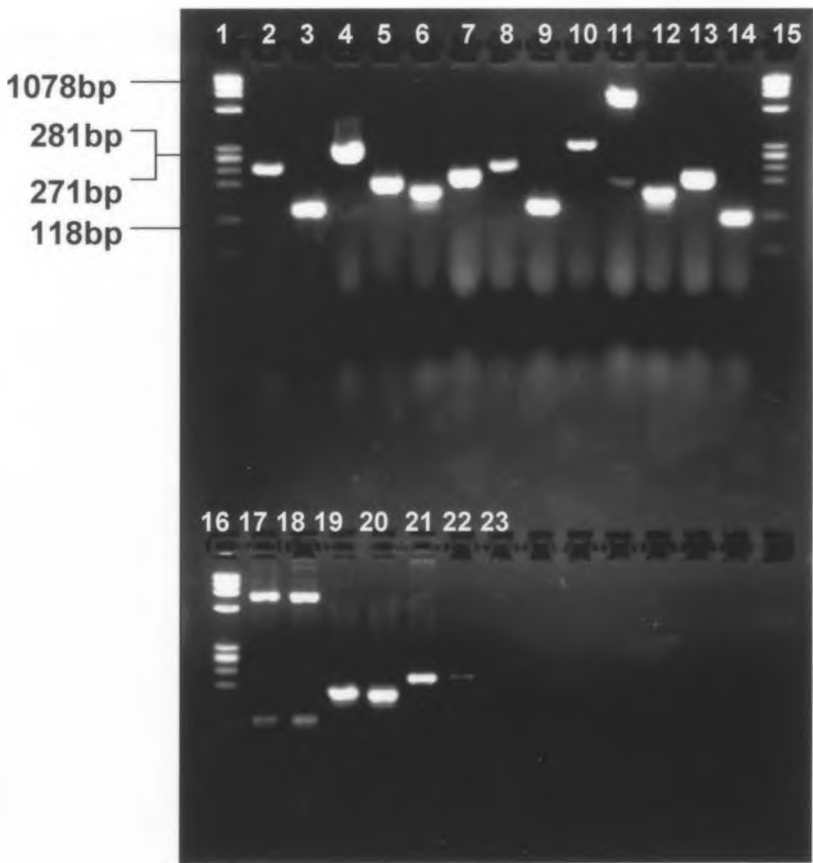


Table 6.2. Description of the introns and exons in the MLLT3 gene (adapted from Strissel *et al*, 2000). MLLT3 exons and introns are displayed by position and length in base pairs within the MLLT3 mRNA sequence and also genomic sequence respectively (columns 1-5). The sixth column displays the ID of the cosmid vector containing the MLLT3 genomic DNA and the seventh column the exon position within the cosmid vector in base pairs. The seventh and final column displays the accession number for the cosmid sequence. Note that the genomic DNA sequence containing exon 3 could not be determined through BLAST analysis.

Table 6.2. Description of the introns and exons in the MLLT3 gene

Exon	Position	Length (bp)	Intron	Length (bp)	Cosmid	Position of exon in cosmid	Accession no.
1	1 – 206	206	1	17788	34a5	13476 – 13682	002052
2	207 – 388	181	2	-	34a5	15092 – 15273	002052
3	389 – 614	225	3	-	-	-	-
4	615 – 1320	705	4	47893	92f5/C48	23303 – 24008 (92f5)	002053
5	1321 – 1398	77	5	2058	C48	14872 – 14949	000007
6	1399 – 1526	127	6	2630	C48	12681 – 12812	000007
7	1530 – 1626	96	7	5858	C48	9947 – 10049	000007
8	1627 – 1698	71	8	1211	C48	4012 – 4086	000007
9	1699 – 1770	71	9	5302	C48	2730 – 2803	000007
10	1771 – 3376	1605	-	-	213a7	38032 – 39638	002050

Figure 6.10 Primary sequence of the MLLT3 gene. Figure displaying the primary sequence of MLLT3 mRNA. The start of exons 1-10 are indicated in the sequence in blue. The initiation (start) codon is highlighted within a blue box (ATG) and the termination (stop) codon within a red box (TGA). The poly (A) signal (AATAAA) is underlined and highlighted within a green box. The TF2F conserved region from nucleotides 208-513bp is underlined.

Figure 6.10 Primary sequence of the MLLT3 gene.

1 1tttggggctg agttaataa gcgagcgagc gagcaagcga gcgcgggggg aaaaaggcag
61 agaatgtccg ccatctaccc tccgctcctg ggcgcgctct cattcatagc agcctttca
121 tgaattacag ctgagggggg gcgaggagg ggggggtacc acacaacacc ccagcaaacc
181 tccgggcccc caggcatggc tagctcg2tat tccatacaaa taaaactgaa actaagacac
241 cqcgcccagg tgaagaaaaa acccaccata gaggactca cccacgacta gatgatattc
301 gtacgcggtc cgaagcacag taacatacag cactttatg aaaaagtcat ctccactg
361 cacgaaagct ttctaggcc aaaaagag3tg tgcaaagatc cacctacaa agtagaagaa
421 tctgggtatg ctggttcat ttgccatt gaagttat taaaaacaa gaaagaacct
481 agaaaagtcc gctttgatta tgacttattc ctgcatctg aaggcatcc accagtgaat
541 cacctccgt gtgaaaagct aacttcaac aaccacacag aggacttag gagaaagtg
601 ctgaaggcag gagg4ggacc taataggagt attcatacca gcagcagcag cagcagcagc
661 agtagcagca gcagcagcag cagcagcagc agcagtagca gcagcagcag cagcagcagc
721 agcagcagta gcagcagcag tagcagcagc agcagcagca gtagtaccag ttttcaaag
781 cctcacaat taatgaagga gcacaaggaa aaaccttcta aagactccag agaacataaa
841 agtgccttca aagaacctc caggatcac acaaatctt ccaagaatc ctctaagaaa
901 ccaaagaaa ataaacct gaaagaagag aaaatagttc ctaagatggc ctcaaggaa
961 ctaaaccba tgcaaaaaga gccaaaacca gatagtaact tactacat caccagtga
1021 caagataaga aggctcctag taaaaggccg ccatttcag attctgaaga actctcagcc
1081 aaaaaagga aaaagagtag ctgagaggct ttattaaaa gttttctag cgaccacca
1141 ctgatactca ctgttctgc tgacaaaaa cagataaaag ataaatctca tgtcaagatg
1201 ggaaaggtca aaattgaaag tgagacatca gagaagaaga aatcaacgtt accgccatt
1261 gatgatattg tggatccaa tgattcagat gtggaggaga atatatctc taaatctgat
1321 5tctgaacaac ccagtctgc cagctccagc tccagctcca gctccagctt cacaccatcc

1381 cagaccaggc aaca6aggtcc ttgaggtct ataatgaaag atctgcattc tgatgacaat
1441 gaggaggaat cagatgaagt ggaggataac gacaatgact ctgaaatgga gaggcctgta
1501 aatagaggag gcagccgaag tcg7cagagtt agcttaagtg atggcagcga tagtgaaagc
1561 agttctgctt cttcaccctt acatcacgaa cctccaccac ccttactaaa aaccaacaac
1621 aacca8gattc ttgaagtga aagtccaata aagcaaagca aatcagataa gcaaataaag
1681 aatggtgaat gtgacaagg9c atacctagat gaactggtag agcttcacag aaggttaatg
1741 acattgagag aaagacacat tctgcagcag a10tcgtgaacc ttatagaaga aactggacac
1801 tttcatatca caaacacaac attgatttt gatctttgct cgctggacia aaccacagtc
1861 cgtaaactac agagttacct ggaacatct ggaacatcc ████ggatataa caactggatg
1921 catcaagaac tattgtgttt tttttttg gttttttt ttttgggtg tgatttttg
1981 ttctgtgtg ttatatgaa aactcaaaa tgatgcaacc aaaagggaaa aaataaaaaat
2041 caaacaacct tcagctttat tttctttaa agccagtcac catctctga taaaggagag
2101 gttaaagcaa accagcctca gcgaccact cttctctca aggaaatccc cggaagagt
2161 tagcctggat agcctgaaa acaacaaat caaacacaac acaagaaaac tcaaagaatg
2221 tgatggtat catgtatctc tctgtggtg ttattccac aggacgaatg catattcaac
2281 aactgcctt attacataac tgatctattt attatcgcat acagatattc taagtcgtg
2341 agggaatgac accatcagac attataagta ctgggtcccg tggatgctct tcaatgcag
2401 caccctgcc atccaagcc cagtgcctt actcgtatac cgtgccactt tccaccaact
2461 tttccaagt ctttaactc gttgcagtct gtattttca cttttgttt tccagttcc
2521 aggacacaga ttatcaactg gggggaccaa atagccacct tgattttctt cttgtggtc
2581 ttttctga aagtggggc ccagtcctg gctgtatcca tgtaatgac ttggaccatg
2641 gtagaaaatg caccaaatag gatcatatga atgctgtct agccttagtc aataaactg
2701 taggactttt aaacaaaagt gtacctgtaa atgtcctgaa tccagcattg ttgagctgc
2761 atcaacattc ttgtgtctg ttactgtta caatattagg tgaatatgga agtaaaggca
2821 ttccacagga tcatattta aaaaaaaga attctggtcc tgttttctaa aaaaaaaaaa

2881 ctgttgtaga aattcctaatttggatctat ttattagtca gagtttcagc tttcttcagc
2941 tgccagtgtg ttactcatct ttatcctaaa aatctggaat cagagatttt tgttgttca
3001 catatgattc tcttagacac ttttatattt gaaaaaatta aaatcttct ttggggaaaa
3061 attcttggtt attctgcat aacagattat gtattaactt gtagattcag tggttcaata
3121 cctgttagt tgcttgctaa tatttcaga aggatttctt gtattggtga aagacggtg
3181 gggatggggg gattttttg ttctgttgt accctgttt tgaaactaga aatctgtct
3241 gtggcatgca aaagaaagca aattattttt aaaagaaaaa aaccaaagta ctttggtgt
3301 cattattcca tcttccat aagtggagaa atgaaaagta agaacagctc atctcaaag
3361 ttttactag aaattc

6.3.5 SSCP analysis of the MLLT3 gene in NSCLC

Forty-eight paired NSCLC normal and tumour DNA samples were analysed for conformational changes caused by mutation using SSCP. Primers were designed flanking all of MLLT3 exons for SSCP analysis (except exon 3) using the genomic DNA sequence of MLLT3 contained in cosmids 213a7, 92f5, 34a5 and c48. As the genomic sequence surrounding exon 3 was unknown, primers were designed to amplify within the exon itself. SSCP exon analysis revealed conformation shifts for SCC#H4 and SCC#J1 in exon 1, SCC#C2 in exon 9. A shift in conformation was also detected in both the normal and tumour of SCC#J3 within exon 4 and Adeno#J5 within exon 7 (figure 6.11). No SSCP products were detected using primers designed for amplification within exon 3. Conformation shifts were not detected for any of the tumours that displayed a minimal region of AI around the region D9S1684 – D9S1778.

6.3.6 Sequencing of SSCP products

Sequencing of abnormal MLLT3 PCR products from genomic DNA revealed a 148(G→A) transition in exon 1 for SCC#H4 (figure 6.12), a 627(G→T) transversion and a 663(T→C) transition within a trinucleotide repeat sequence in exon 4 for SCC#J3 (figure 6.13). Normal MLLT3 sequence was obtained from all other tumours which displayed conformation shifts after SSCP analysis.

Figure 6.11 Photomicrograph displaying abnormal MLLT3 PCR products after SSCP analysis. SSCP exon analysis for SCC#H4 and SCC#J1 in exon 1, SCC#J3 in exon 4, Adeno#J5 in exon 7 and SCC#C2 in exon 9. Normal (N) and tumour (T) DNA is indicated above each photomicrograph. Arrows indicate conformation shifts.

Figure 6.11 Photomicrograph displaying abnormal MLLT3 PCR products after SSCP analysis.

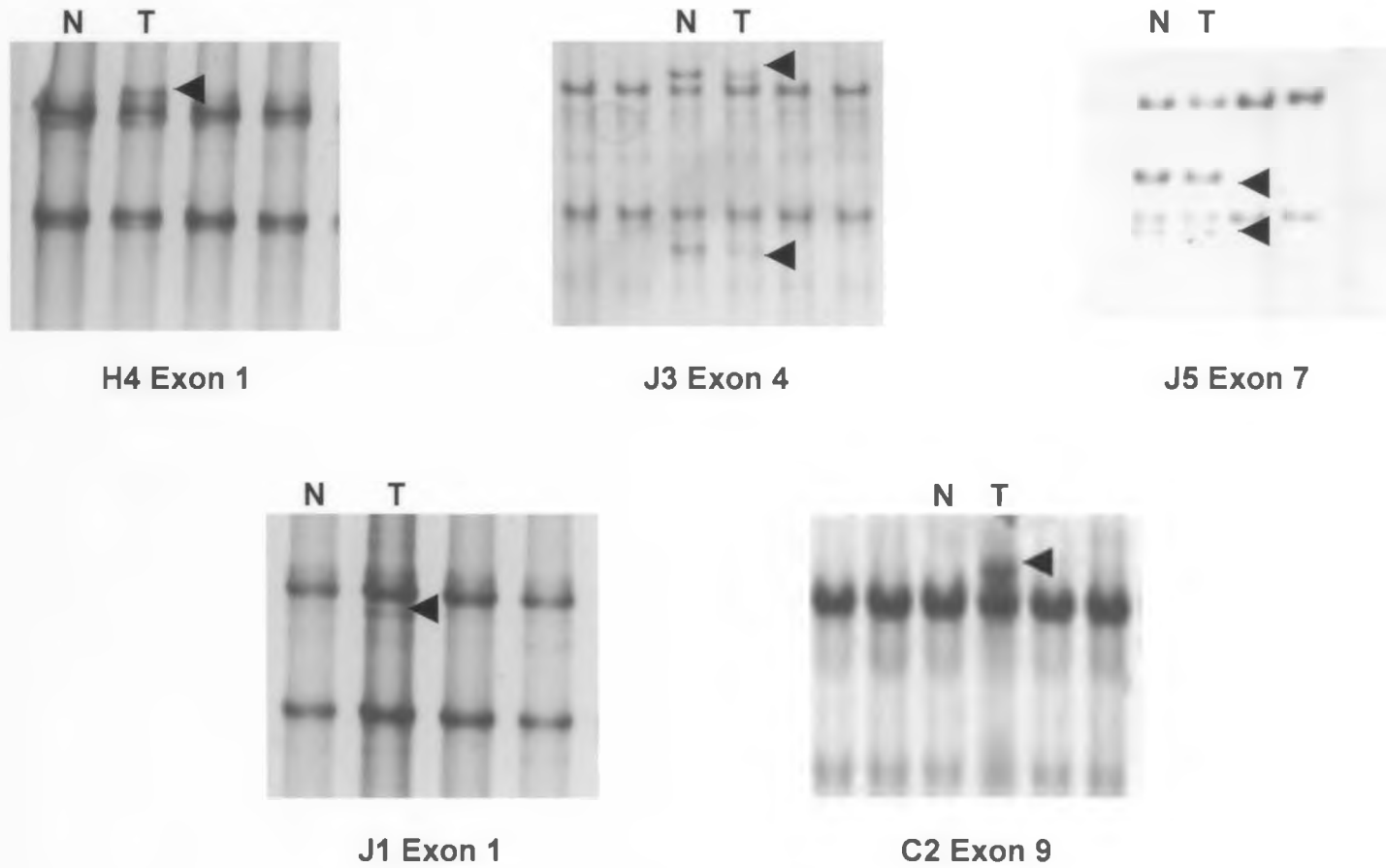


Figure 6.12 Screen captures from Chromas™ displaying nucleotide substitution at 148bp of MLLT3 exon 1 in SCC#H4. Normal sequence is displayed at the top of the figure and the abnormal sequence is displayed on the bottom of the figure. Sequence is displayed in the reverse strand. Sequencing revealed a 148(G→A) transition (black arrows) in MLLT3 exon 1.

Figure 6.12. Sequencing of MLLT3 exon 1 in SCC#H4.

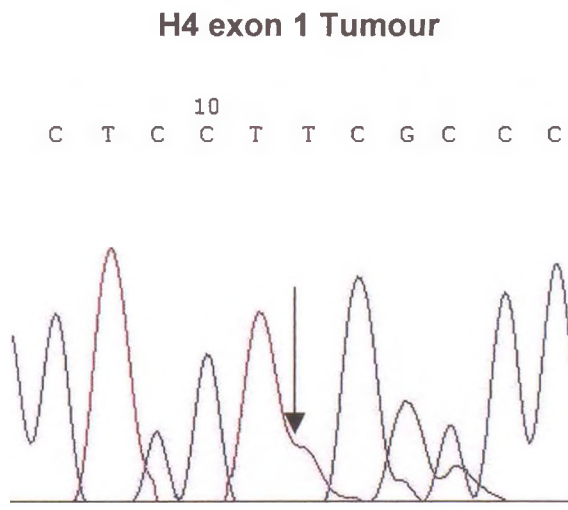
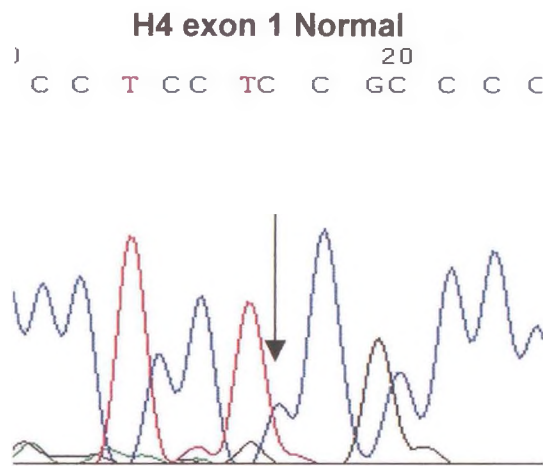
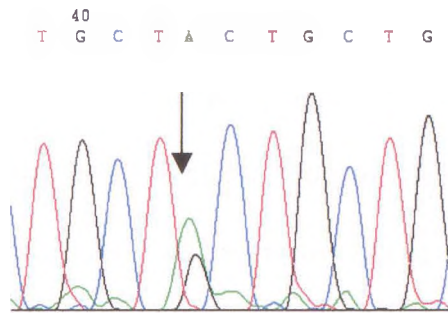
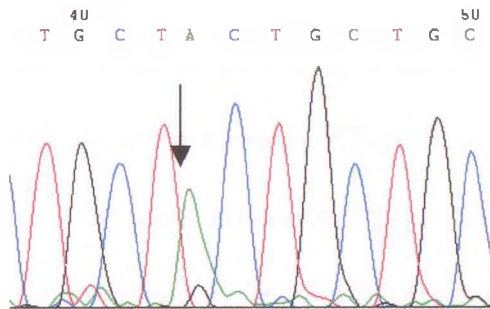


Figure 6.13 Screen captures from Chromas™ displaying nucleotide substitution at 663bp of MLLT3 exon 4 in SCC#J3. Normal and control sequence (from whole blood DNA) is displayed in the top of the figure along with a 663(T→C) transition (black arrows) within a trinucleotide repeat sequence in MLLT3 exon 4 detected in SCC#J3.

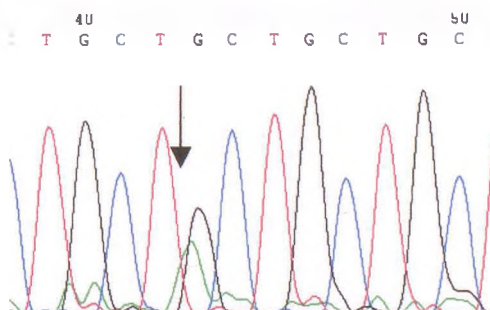
Figure 6.13. Sequencing of MLLT3 exon 4 in SCC#J3.



Control



J3 Exon 4 Normal



J3 Exon 4 Tumour

6.3.7 cDNA analysis of MLLT3

Aberrant expression of the MLLT3 gene was investigated using RT-PCR to amplify exons 4-10 of MLLT3 cDNA extracted from 9 normal human tissues, 3 lung tumours with corresponding normal cDNA, 4 tumourigenic cell lines (3 breast cancer, 1 lymphoblastoid) and 3 foetal cDNA libraries (figure 6.14). Full length cDNA PCR products were detected in all samples except 1 lung tumour (L140, lane 18) and foetal liver cDNA (lane 25) which displayed no PCR product. Further analysis of tumour L140 and foetal liver cDNA using reliable PCR control cDNA's and primers within a different genetic region again displayed no PCR products, indicating a low quality or yield of cDNA probably reflecting inefficient RNA extraction for these samples (data not shown). RT-PCR of MLLT3 cDNA from a further 7 normal and tumour lung pairs again displayed full length cDNA PCR products indicating normal expression of the MLLT3 gene in these lung tumours (figure 6.15). Differences in the intensity of PCR products between normal and tumour cDNA in both investigations were accounted for by varying concentrations of total RNA extracted from the original tissue which affected the RT-PCR.

Figure 6.14. Expression of full length cDNA products from exons 4-10 of the MLLT3 gene in various human and foetal tissues. RNA was extracted from human tissue and cell lines, reverse transcribed then cDNA amplified using MLLT3 cDNA primers from exons 4-10. PCR products were electrophoresed through a 3% agarose gel. Lane order is numbered at the top of each well on both the upper and lower rows. Lane 1: ϕ X174 DNA/Hae III size marker; Lane 2: Lung; Lane 3: Colon; Lane 4: Skeletal Muscle; Lane 5: Spinal chord; Lane 6: Heart; Lane 7: Testis; Lane 8: Breast; Lane 9: Brain; Lane 10: Stomach; Lane 11: L090 Normal lung control; Lane 12: L090 Primary lung tumour; Lane 13: L092 Normal lung control; Lane 14: L092 Primary lung tumour; Lane 15: ϕ X174 DNA/Hae III size marker; Lane 16: ϕ X174 DNA/Hae III size marker; Lane 17: L140 Normal lung control; Lane 18: L140 Primary lung tumour; Lane 19: BT20 Breast cancer cell line; Lane 20: CM51 Breast cancer cell line; Lane 21: 221 Breast cancer cell line; Lane 22: Lymphoblastoid cell line; Lane 23: Normal fibroblasts; Lane 24: Foetal brain; Lane 25: Foetal liver; Lane 26: Foetal lung cDNA library; Lane 27: negative PCR control (sterile water); Lane 28: ϕ X174 DNA/Hae III size marker.

Figure 6.14. Expression of full length cDNA products from exons 4-10 of the MLLT3 gene in various human and foetal tissues.

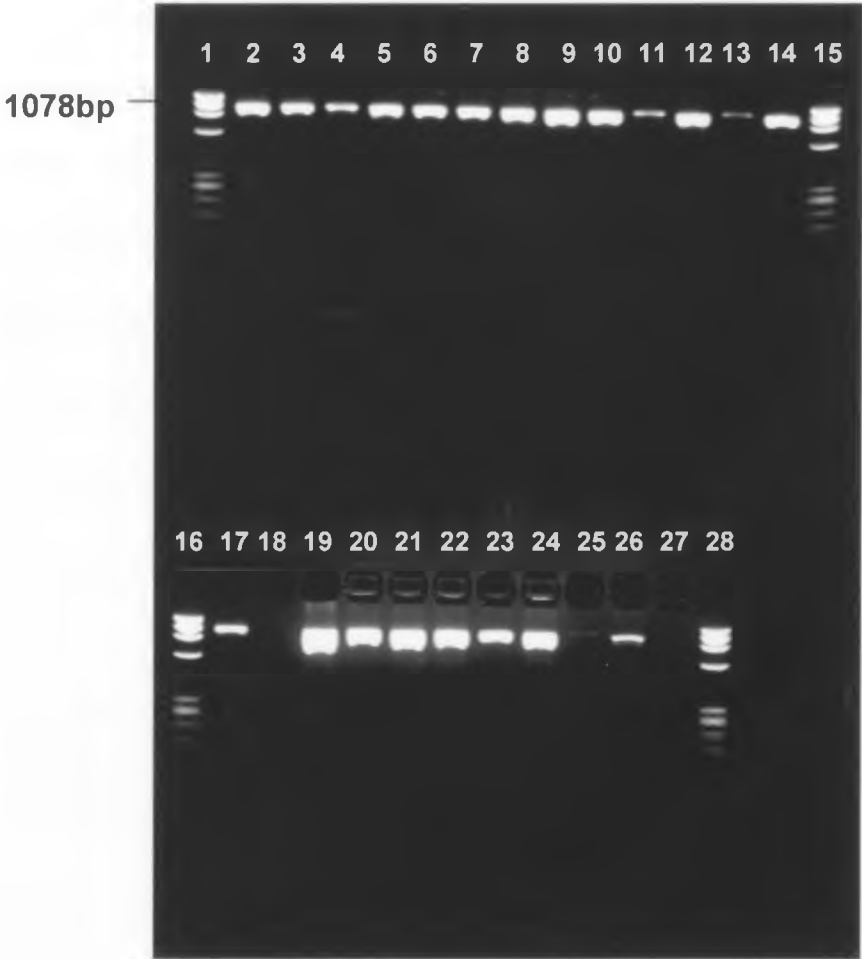
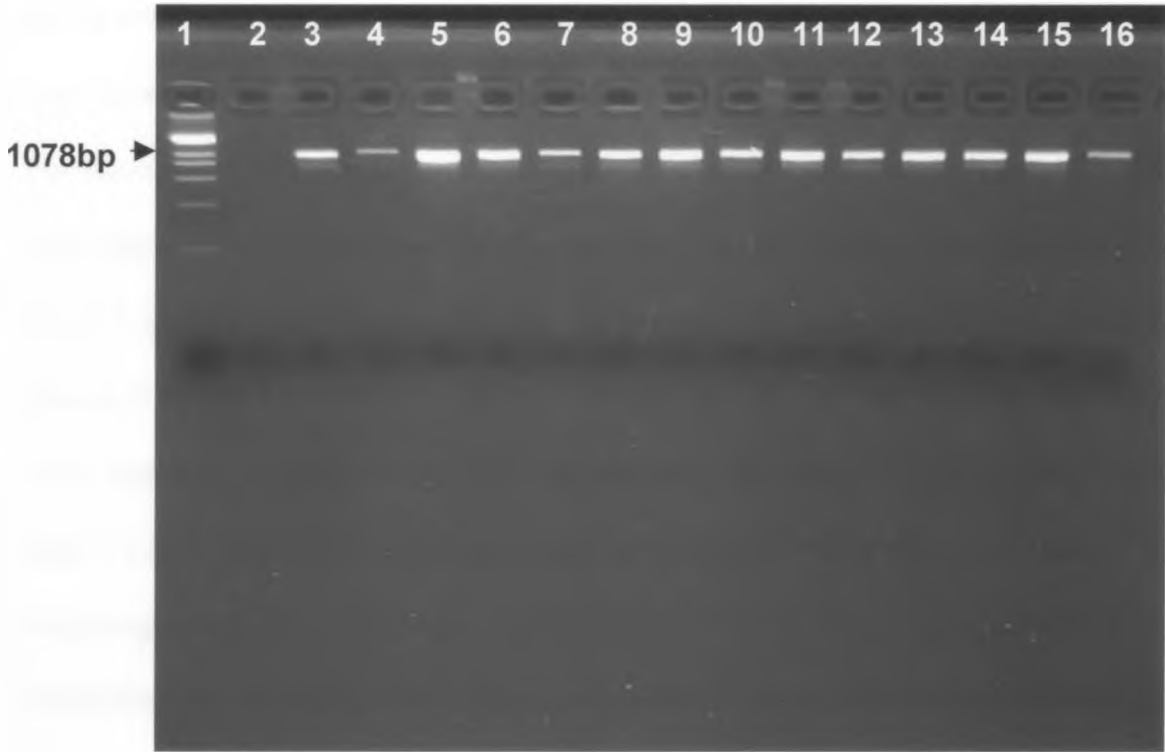


Figure 6.15. Expression of full length cDNA products from exons 4-10 of the MLLT3 gene in lung tumour and normal pairs. RNA was extracted from 7 NSCLC tumours and also paired normal lung tissue, reverse transcribed then cDNA amplified using MLLT3 cDNA primers from exons 4-10. PCR products were electrophoresed through a 3% agarose gel. Lane order is numbered at the top of each well at the top of the gel. Lane 1: ϕ X174 DNA/*Hae* III size marker; Lane 2: negative PCR control sterile water; Lane 3: Lung normal 1; Lane 4: Lung tumour 1; Lane 5: Lung normal 2; Lane 6: Lung tumour 2; Lane 7: Lung normal 3; Lane 8: Lung tumour 3; Lane 9: Lung normal 4; Lane 10: Lung tumour 4; Lane 11: Lung normal 5; Lane 12: Lung tumour 5; Lane 13: Lung normal 6; Lane 14: Lung tumour 6; Lane 15: Lung normal 7; Lane 16: Lung tumour 7

Figure 6.15. Expression of full length cDNA products from exons 4-10 of the MLLT3 gene in lung tumour and normal pairs.



6.3.8 HGMP NIX analysis of deleted region

Full genomic sequences for BACs bA252O2 and bA363E7 and cosmid clones 213a7, 92f5, 34a5 and c48, positioned within the SRD, were obtained from NIH Entrez nucleotide database and analysed using the HGMP resource NIX analysis program. Enquiries using Entrez nucleotide for the genomic sequence of other BACs within this region revealed DNA sequencing was still in progress for multiple unordered pieces. EST search results were only considered as true hits when corresponding sequence results were classed by the NIX database as 'excellent' matches. EST and gene sequences revealed by NIX analysis were confirmed by BLAST database analysis. NIX analysis revealed multiple EST hits on all cosmid (figure 6.16) and both BAC sequences (figure 6.17). Cosmid 35a5 matched with 1 CpG island and 6 EST's, cosmid 92f5 matched with 4 EST's, cosmid c48 matched with 6 EST's and the *Homo sapiens* genomic intron breakpoint sequence of MLL rearrangement, cosmid 213a7 matched with 4 EST's. BAC analysis matched bA252O2 with the sequence of *Homo sapiens* cone sodium-calcium potassium exchanger (NCKX2) mRNA and 6 EST's. BAC bA363E7 was matched with 1 CpG island, rat microtubule-associated proteins 1A/1B light chain 3 (*MAP1A/MAP1B LC3*) mRNA, 2 *Homo sapiens* cone sodium-calcium potassium exchanger splice variants and 16 EST's.

Figure 6.16. Diagrammatic representation of NIX analysis results for cosmid sequences within the SRD. The genomic sequence of cosmids 34a5, 92f5, c48 and 213a7 was analysed using the HGMP NIX database for EST's mapping within the SRD. A diagrammatic representation of the positions of matched EST's and the *MLLT3* exons within the cosmids genomic sequence is shown. Key to figure is as follows : horizontal black lines, cosmid sequence from 1 bp to end of sequence; hatched box, predicted CpG region; red box, *MLLT3* exons (numbered from 1-10); blue triangle, putative GRAIL pol II promoter position; white boxes, EST (ordered E1, E2...etc. according to numerical position within genomic sequence); green box, *Homo sapiens* genomic intron breakpoint sequence of *MLL* rearrangement breakpoint; black arrow, position of microsatellite D9S1778; blue arrow, position of STS stSG43. Identity (including accession number) and position of ESTs within each cosmid is as follows. **Cosmid 34a5:** E1, EM:AA565727 *Homo sapiens* cDNA (10195-11348 bp); E2, EMU:BF510928 *Homo sapiens* cDNA (15093-15274bp); E3, EM:AW936233 *Homo sapiens* cDNA (30855- 313877bp) overlapping with EM:AW936272 *Homo sapiens* cDNA (31103-31498). E4, EM:AV762261 *Homo sapiens* cDNA 5'end, expressed in CD34+ hematopoietic stem/progenitor cells (33434-33756 bp); E5, EM:AA595823 *Homo sapiens* cDNA clone (33766-34114bp). **Cosmid 93f5:** E1, EMU:BF335619 *Homo sapiens* cDNA (16136-16297bp); E2, EM:AW603580 *Homo sapiens* cDNA (23083-23393bp); E3, EM:AW157750 foetal brain *Homo sapiens* cDNA similar to AF-9 protein (23543-23861bp) overlapping with EMU:BF537401 *Mus musculus* cDNA (23604-23861bp). **Cosmid c48:** E1, EMU:BF801725 *Homo sapiens* cDNA (2992-3160bp); E2, EM:AW612910 *Homo sapiens* cDNA similar to human breast basic conserved protein 1 (3304-3661bp); E3, EMU:BF434955 *Homo sapiens* cDNA similar to AF-9

protein (16069-16304bp); E4, EMU:BF858224 Homo sapiens cDNA (8719-8961bp); E5, EM:AI911245 Homo sapiens cDNA clone (11934-12448bp); E6, EM:HSZZ26628 Bone marrow Homo sapiens cDNA (16069-16304bp). **Cosmid 213a7**: E1, EM:BE149610 Homo sapiens cDNA (27774-28031bp); E2, EM:AW451255 Homo sapiens cDNA (34728-35235bp); E3, EM:AI554049 Homo sapiens cDNA clone (38693-39193bp); E4, EM:AW867240 Homo sapiens cDNA (40825-41035bp).

Figure 6.16. Diagrammatic representation of NIX analysis results for cosmid sequences within the SRD.

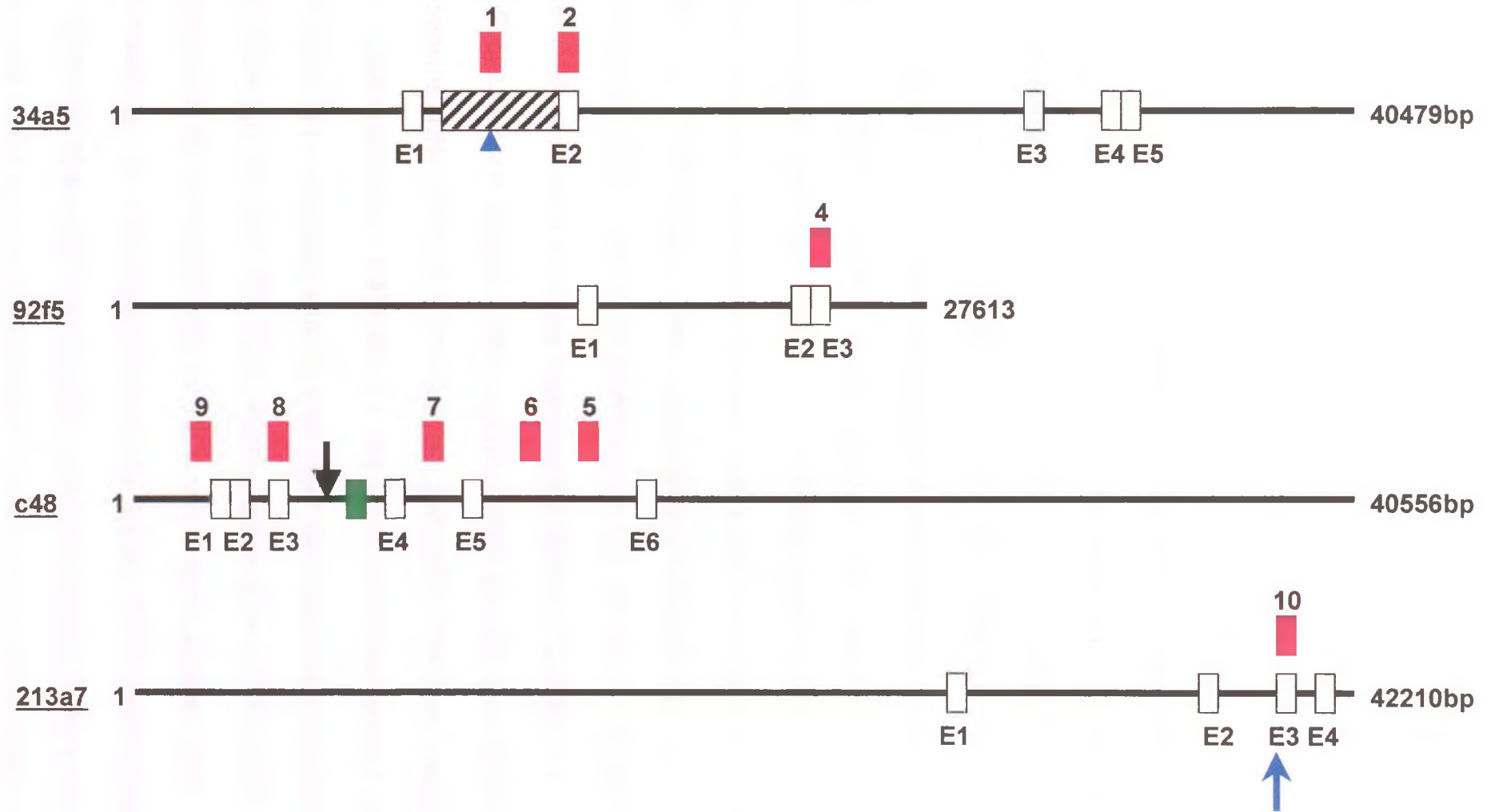


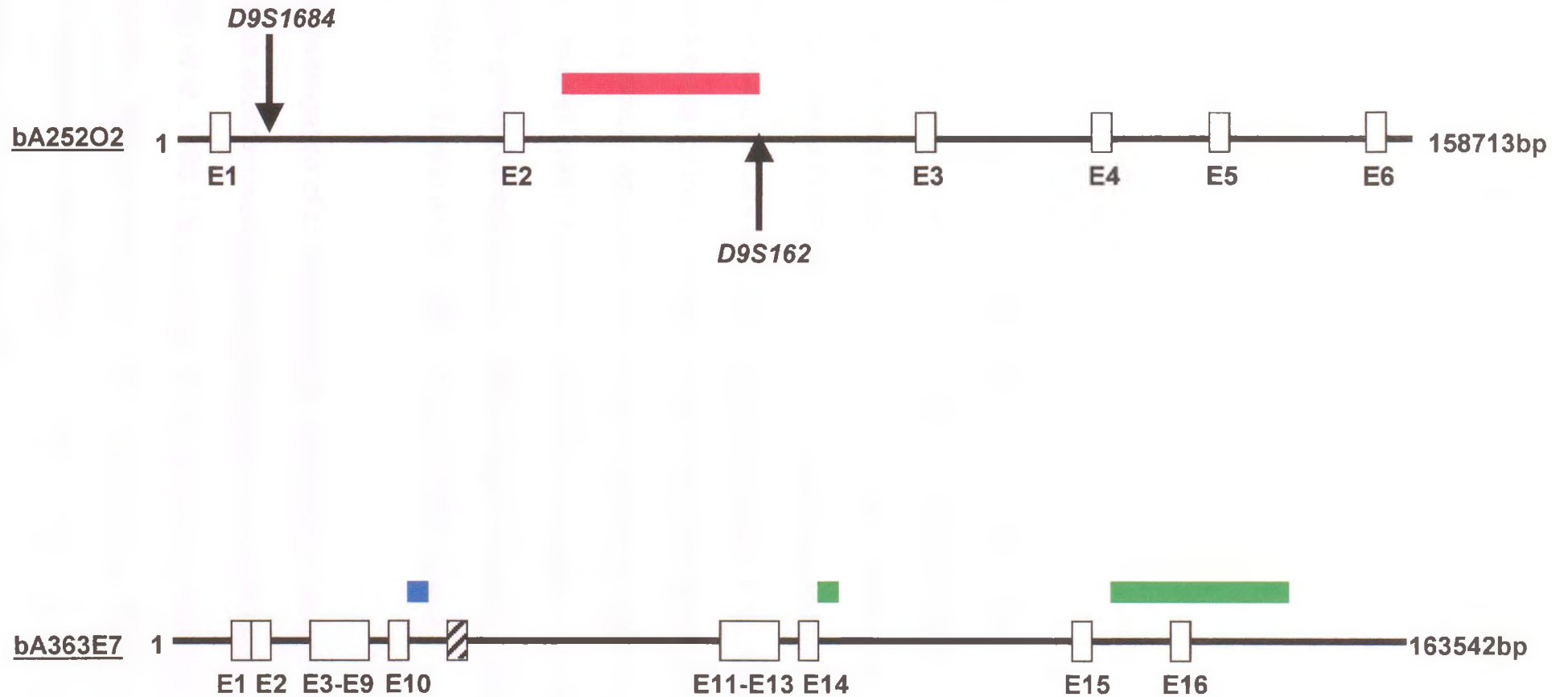
Figure 6.17. Diagrammatic representation of NIX analysis results for BAC sequences within the SRD. The genomic sequences of BACs bA252O2 and bA363E7 were analysed using the NIX database. A diagrammatic representation of the positions of matched EST's and also several genes within the genomic sequence is shown. Key to figure is as follows. Horizontal black lines, genomic DNA sequence from each clone (labelled from 1bp to end of sequence). Red bar, Homo sapiens cone sodium-calcium potassium exchanger (NCKX2) mRNA (49355-74328 bp); blue box, Rat microtubule-associated proteins 1A/1B light chain 3 (MAP1A/MAP1B LC3; 28022-28165 bp); hatched box, predicted CpG region; green box, Homo sapiens cone sodium-calcium potassium exchanger splice variant (NCKX) mRNA (83006-83284 bp); green bar, Homo sapiens cone sodium-calcium potassium exchanger splice variant (117003-143910 bp). Identity (including accession number) and position of ESTs within each BAC is as follows.

bA252O2: E1, EM:HS158280 Homo sapiens cDNA (5918-6096 bp); E2, EMU:BF727132 Human Lens cDNA (40189-40631 bp); E3, EM:AV723006 Homo sapiens cDNA expressed in human hypothalamus (96752-97228 bp); E4, EMU:BF962233 Homo sapiens cDNA (116679-116969 bp); E5, EMU:BF963949 Homo sapiens cDNA (129025-129418 bp); E6, EMU:BF837194 Homo sapiens cDNA (155288-155462 bp).

bA363E7: E1, EM:AW901654 Homo sapiens cDNA (11728-12033 bp) overlapping with E2, EM:AW901644 Homo sapiens cDNA (11888-12134 bp); E3, EMU:BF839613 Homo sapiens cDNA (17752-18192 bp) overlapping with E4, EM:HSZZ14981 Jurkat T-cells V Homo sapiens cDNA (18044-18465 bp); E5, EM:HS930249 foetal liver spleen 1NFLS Homo sapiens cDNA (18440-18769 bp); E6, EM:AI033878 parathyroid tumour Homo sapiens cDNA (18939-19385 bp); E7, EMU:BF816172 Homo sapiens cDNA (19817-20034

bp); E8, EMU:BF992378 Homo sapiens cDNA (20388-20778 bp); E9, EM:AW968664 MAGJ Homo sapiens cDNA (21354-21580 bp) ; E10, EM:AA640952 Homo sapiens cDNA clone similar to light chain 3 subunit of microtubule-associated proteins 1A and 1B (27795-28012 bp); E11, EM:AV728954 Homo sapiens cDNA expressed in human hypothalamus (74336-74669 bp) overlapping with E12, EM:BE766817 Homo sapiens cDNA (74518-75-93 bp); E13, EM:HS880164 infant brain 1NIB Homo sapiens cDNA (75685-75932 bp); E14, EM:AA682997 Homo sapiens brain cDNA clone (80550-80848 bp); E15, EMU:BF365775 Homo sapiens cDNA (114943-115144 bp); E16, EM:AI332676 Homo sapiens cDNA clone (128111-12817 bp).

Figure 6.17. Diagrammatic representation of NIX analysis results for BAC sequences within the SRD.



6.4 Discussion

6.4.1 AI analysis around 9p22-23 revealed multiple regions of AI and further reduced the affected chromosomal region.

Allelic imbalance on chromosome 9p is reported to be a frequent occurrence in multiple tumour types including melanoma, bladder, lung, head and neck and other tumour types (Fountain *et al.*, 1992; Knowles *et al.*, 1994; Merlo *et al.*, 1994a; Lydiatt *et al.*, 1995; Kiaris *et al.*, 1995; Neville *et al.*, 1996; Olopade *et al.*, 1992b). Research has characterised the pattern of allelic imbalance that occurs in NSCLC and studies have identified chromosome 9p as a target in these tumours (Field *et al.*, 1996; Girard *et al.*, 2000; Liloglou *et al.*, 2000; Sato *et al.*, 1994; Tsuchiya *et al.*, 1992). Deletion mapping studies in melanoma identified the critical region of 9p21 (Holland *et al.*, 1994) where putative tumour suppressor genes (TSG) may have existed and subsequent studies located the p16 tumour suppressor gene within this area (Kamb *et al.*, 1994) however evidence exists for the novel site of other TSG(s) on chromosome 9p involved in multiple tumour types (An *et al.*, 1999; Cairns *et al.*, 1995; Keen & Knowles, 1994; Kim *et al.*, 1997; Puig *et al.*, 1995; Simoneau *et al.*, 1996) including the interferon genes (Olopade *et al.*, 1992b; Keen & Knowles, 1994) and a region near D9S171 (Devlin *et al.*, 1994; An *et al.*, 1996) at 9p22.

Closer investigation of chromosome 9p has provided strong evidence that allelic imbalance at chromosome 9p21-p23 occurs at a high frequency in NSCLC (Gazzeri *et al.*, 1998; Liloglou *et al.*, 2000; Mao *et al.*, 1997; Merlo *et al.*, 1994b; Neville *et al.*, 1995; Okami *et al.*, 1997; Olopade *et al.*, 1992a; Olopade *et al.*, 1993; Schmid *et al.*, 1998; Wiest *et al.*, 1997). Merlo *et al.*, (1994b) detected a homozygous deletion within the 9p21 region including the markers IFN-alpha,

D9S126, and D9S171 in SCLC. Mead *et al* (1994) detected deletions centromeric to the IFN locus in proximity to D9S269 and D9S171 and Wiest *et al*, (1997) identified a novel region of chromosomal loss proximal to p16 between the markers D9S265 and D9S269 in NSCLC. Previous research has centred around AI within the p16 region and inactivation of the p16 gene through allele loss or homozygous deletion on chromosome 9p21 (Gazzeri *et al.*, 1998; Gorgoulis *et al.*, 1998; Hamada *et al.*, 1998; Shapiro *et al.*, 1995; Su *et al.*, 1999). More recent studies have identified regions of homozygous deletions proximal to the p16 gene near the methylthioadenosine phosphorylase (MTAP) gene and around the locus D9S171 suggesting the existence of another gene in this region which plays a role in the pathogenesis of NSCLC (Hamada *et al.*, 2000; Schmid *et al.*, 1998). Schmid *et al* (1998) investigated homozygous deletions in both p16 and MTAP genes in 50 NSCLC tumours. They concluded from this study that 38% of NSCLC displayed a MTAP homozygous deletion. In addition, the MTAP deletions were associated with homozygous deletions of p16 in only half of cases, suggesting the existence of another gene telomeric (proximal) to p16 involved in NSCLC. Hamada *et al* (2000) investigated homozygous deletions within the 9p21 region in 145 NSCLC cell lines and detected two clusters of deletions in 26% of the cell lines, one cluster was located within the p16 region and the other was located at the D9S171 region, 3 Mb proximal to the p16 locus. Recently, Girard *et al*, (2000) carried out a genome wide allelotype of lung cancer and detected a minimal region of loss around the locus D9S286 at 9p23. Neville *et al* (1995) analysed 45 NSCLC tumours with a panel of highly informative microsatellites and detected a minimal region of deletion from D9S156-D9S161 on chromosome 9p22-p23.

This investigation has attempted to further characterise the region of loss detected by Neville *et al* on human chromosome 9p22-p23 in a panel of 56 NSCLC tumours. The results presented earlier in this chapter identified several minimal regions of allelic imbalance between the loci D9S269 and D9S1778 on chromosome 9p23. In keeping with previous studies the microsatellites D9S156 and D9S157 displayed the highest frequency of allelic imbalance with rates 71% and 68% respectively (Liloglou *et al.*, 2000; Neville *et al.*, 1995). Fourteen tumours displayed minimal regions of allelic imbalance between the microsatellites D9S254-D9S1778 and the minimal regions could be classed into 4 different categories depending on their breakpoints at D9S254, D9S274, D9S285 and D9S1778. The smallest region of AI (<2Mb in length) was detected between the loci D9S1684 and D9S1778 in three tumours (figure 6.5), hence further narrowing the minimal region of AI within chromosome band 9p22-p23 in NSCLC. In comparison with the previous study by Neville *et al* which detected a minimal region of AI around the locus D9S157 in five NSCLC tumours, a minimal region was detected containing the microsatellite D9S157 in 9/14 tumours in this study as part of a larger region. This implies that AI around D9S157 may play a significant role in NSCLC development as AI at this locus has now been identified in two separate studies using two separate panels of NSCLC tumours. These results supply further evidence for the existence of putative TSG(s) within the 9p22-p23 region.

Statistical analysis revealed no correlation between AI within the 9p22-p23 region and histological/morphological diagnosis. A significant correlation was detected for locus D9S254 with both MLH1 and MSH2 positive immunocytochemistry and D9S285 with MLH1 positive immunocytochemistry.

Association of one locus within the 9p22-p23 region with increased expression of one of the DNA repair genes can be explained as random or chance occurrence, however association between the locus D9S254 and two independent MLH1 and MSH2 expression studies, in addition to another association between locus D9S285 and MLH1 expression does not appear to be a random occurrence. Reduced expression of hMLH1 has been correlated with allelic imbalance at loci D3S1289 and D2S391 (Xinarianos *et al.*, 2000), however there is no documentation describing correlations between AI on chromosome 9p and expression of MLH1 and MLH2 in NSCLC. This suggests the existence of a gene within the 9p22-p23 region which may be involved in the DNA mismatch repair pathway.

6.4.2 Accurate positioning of microsatellite marker order within the 9p22-p23 region and selection of a putative target gene for analysis within the minimal region of AI

The minimal region between D9S1682-D9S1778 was investigated further using mapping and gene searching techniques. The selection of this region for further analysis was based on several criteria.

- 1) The SRD was a relatively small size (<2Mb) when compared to the other minimal regions of deletion and was feasible for mapping strategies.
- 2) The SRD was present in 3 different tumours (H5, C5 and I1).
- 3) Nine out of 14 tumours displayed a breakpoint within the SRD in DM2.
- 4) Databases existed containing YAC and BAC data and a list of candidate genes and EST's were available for this region.

To confirm microsatellite order both the Whitehead Institute and CEPH YAC databases were investigated as previously described in this chapter. Several major problems were encountered during this search which hindered correct positioning of microsatellite loci and accurate mapping of the deleted region. The first problem was the representation of all of the microsatellite loci, which were investigated in this study. Important omissions were the locus D9S1778 from both databases and D9S1684 from the CEPH-Genethon database. Physical maps for the region also contained large gaps between microsatellite loci, which would involve screening YAC libraries for multiple STS's. Screening using PCR to amplify microsatellite DNA from YAC hits within this region produced varied results with several clones not producing microsatellite PCR products. In addition there was a possibility of chimeric YAC clones containing two non-contiguous pieces of DNA from 9p22-p23 not fully representing the region of AI and mapping to other chromosomes. The rate of chimeric clone formation in YACs is estimated to be approximately 50% (usually as a result of recombination) and this limits their usefulness in contig construction (Kouprina *et al.*, 1994; Larionov *et al.*, 1994a; Larionov *et al.*, 1994b). Incorrect YAC physical mapping had placed the order of microsatellites within the 9p22-p23 region as D9S925 – D9S162 – D9S1684 . Initially this order was used for mapping minimal regions of AI during the first part of this study (December 1999). This produced incorrect minimal regions of AI around the locus D9S162 with regions of retention at loci D9S925 and D9S1684 (in tumours C5, H5 and I1), which in fact represented a larger genetic distance. Due to these inconsistencies when attempting to confirm microsatellite order using YAC physical mapping strategies, a BAC mapping approach was investigated. Accurate mapping was achieved using the Sanger Centre AceDB and a BAC

physical map was structured for the region D9S925-D9S162 (February 2000). In addition the correct microsatellite order within the minimal region of AI was confirmed as D9S925 – D9S1684 – D9S162. As this eliminated the original minimal region in the tumours investigated, the ACEDB database was screened for microsatellites distal to D9S162. The locus D9S1778 was found and used to screen the 8 tumours shown in figure 6.5. Further confirmation of loci order was received by PCR amplification of BAC DNA using primers for microsatellite loci within the minimal region. Once the marker order was confirmed, putative target genes within the region were considered for analysis.

6.4.3 Analysis of the MLLT3 gene in NSCLC

Several putative target genes have been assigned to the region between D9S157 and D9S1778 according to the Genemap '99 database (table 6.5) including MTAP and p16 in which previous research has demonstrated these genes to be localised distal to this region (Nobori *et al.*, 1996; Olopade *et al.*, 1992a; Olopade *et al.*, 1993), however the exact localisation of target genes within the region of D9S1684 and D9S1778 is questionable. Physical mapping data and BAC PCR demonstrated that the MLLT3 was located within BACs bA336O12 and bA15P13 along with microsatellite locus D9S1778, and hence was localised within the minimal region of AI. The human MLLT3 (AF9) gene is composed of 10 exons and spans >100 kb distal to the microsatellite locus D9S162. MLLT3 is one of the most common fusion partners with the Myeloid Linkage Leukaemia (MLL) gene at 11q23 and this fusion gene [t(9;11)(p22;q23)] is associated with de-novo acute myelo-genous leukaemia (Corral *et al.*, 1996; Iida *et al.*, 1993; Joh *et al.*, 1996; Takahashi *et al.*, 1996). Previous studies have demonstrated that MLLT3 contains two breakpoint

clusters, one within intron 4 and a second spanning introns 7 and 8 (Super *et al.*, 1997).

The analysis of the MLLT3 in human NSCLC has been described in this chapter. MLLT3 was selected for analysis due to its localisation within the minimal region of AI detected in several tumours. BLAST analysis of the full length human MLLT3/AF9 mRNA revealed homology with several cosmid clones which contained the genomic sequence of the gene. Using a sequence comparison database MLLT3 intron/exon boundaries were obtained and SSCP primers were designed to analyse the gene in NSCLC. SSCP analysis revealed 5 conformation shifts within exons 1, 4, 7 and 9. PCR amplification for SSCP analysis of MLLT3 exon 3 did not produce any PCR products. Due to this primer sequences were re-analysed to discount aberrant synthesis and reamplified using control normal blood DNA. Again no PCR products were detected. SSCP analysis of MLLT3 exon 3 differed from the other exons due to the primer location within the exon as opposed to other MLLT3 primers which were located intronic of the MLLT3 gene. Strissel *et al.*, (2000) investigated the structural properties of the MLLT3 gene and distinguished intron-exon boundaries. This study defined MLLT3 exon 3 as located from nucleotides 389-614 of the primary mRNA sequence and contained in cosmid 232g8 (AC: 002469). BLAST analysis of the genomic cosmid sequence with MLLT3 mRNA revealed no significant homology between the sequences suggesting that exon 3 was not contained within this cosmid. Further evidence was provided by NIX analysis of the cosmid sequence which again showed no homology with MLLT3 mRNA unlike the other cosmids previously described. This may be due to inaccurate sequence submission within the NCBI Entrez nucleotide database. Due to this

MLLT3 intron/exon boundaries for exon 3 could not be established and primers were designed within the exon 3 for analysis of genomic DNA using SSCP. The results of several PCR amplifications using these primers suggest that exon 3 may not be one exon but two or more. Non-amplification of the exon can be accounted for by a large intronic sequence existing within the genomic DNA between manufactured exon 3 primers. Confirmation in the future maybe provided by amplification of cDNA using the current exon 3 primers followed by sequencing.

SSCP PCR products which displayed conformational shifts during PAGE electrophoresis were analysed for mutations using DNA sequencing.

Sequencing revealed single nucleotide substitutions within the MLLT3 gene in 2 tumours. SCC#H4 displayed a 148(G→A) transition in exon 1 and SCC#J3 displayed a 663(T→C) transition within a trinucleotide repeat sequence in exon 4. Normal MLLT3 sequence was obtained for all other tumours which displayed conformation shifts. The single nucleotide substitution detected at 663bp within the trinucleotide repeat sequence (poly serine) of exon 4 in SCC#J3 most likely represents loss of one MLLT3 gene copy in this tumour due to allelic imbalance. Figure 6.13 shows that two alleles can be detected in the MLLT3 sequence from control DNA at different ratios (ACGATTG and ACGACCG) which may represent a polymorphism in the gene sequence. In normal lung DNA from SCC#J3, the intensity of the allele carrying thymine (ACGATTG) has increased with reduction in intensity of the allele carrying cytosine (ACGACCG), representing AI of this allele in the normal tissue. However DNA from the tumour shows AI of the other allele as the intensity of the allele carrying thymine (ACGATTG) is reduced when compared to the allele carrying cytosine

(ACGACG), which has increased. Further evidence for loss of one MLLT3 copy in this tumour is provided by AI analysis of SCC#J3 which showed AI at microsatellite loci from D9S157 – D9S162 proximal to the MLLT3. It is therefore very likely that deletion of the MLLT3 locus in one allele has occurred in the tumour and possibly the normal tissue in this NSCLC patient. Previous studies have demonstrated that SSCP is a fast and sensitive technique for the reliable detection of single base changes within DNA sequences (Fan *et al.*, 1993; Hayashi, 1991; Highsmith *et al.*, 1999).

The low mutation rate detected in the MLLT3 gene using a combination of SSCP analysis and sequencing suggests that the MLLT3 gene is not a target for mutation in NSCLC, however this does not eliminate alterations in RNA processing, gene expression or splicing defects of the MLLT3 gene in lung tumours. In order to investigate whether alterations leading to gene splicing abnormalities (such as alternative splicing) occurred in NSCLC tumours, cDNA from several lung tumours was amplified using primers specific for exons 4-10 of the MLLT3 gene. Amplification of MLLT3 cDNA revealed full length transcripts in several normal tissues, foetal cDNA libraries, tumourigenic cell lines and paired normal and tumour cDNA from NSCLC tumours. This suggested normal expression in these tissues and NSCLC with no visible evidence of alternative splicing, deletion, or the translocations of the MLLT3 gene which are associated with leukaemia in both human and mouse models (Anguita *et al.*, 2000; Corral *et al.*, 1996; Dobson *et al.*, 1999; Iida *et al.*, 1993; Joh *et al.*, 1996; Mitterbauer *et al.*, 1999; Super *et al.*, 1997; Takahashi *et al.*, 1996). This does not exclude alterations in RNA processing or the differential expression of various transcripts in these and other lung tumours and also

alterations within exons 1-3 of the MLLT3 cDNA which were not analysed in this study. An investigation of this kind would be technically demanding and require immuno-expression analysis, direct sequencing and studies of RNA expression using northern blotting, which would be difficult with small clinical samples although feasible with lung tumour cell lines. The results in this chapter suggest that MLLT3 is an infrequent target for abnormalities in NSCLC and that other putative genes may exist in the chromosome 9p22-p23 region and more specifically between the microsatellite loci D9S1684-D9S1778 that could play a role in lung carcinogenesis.

6.4.4 NIX analysis of genomic sequence between D9S1684 and D9S1778 reveals multiple EST's and several candidate genes for analysis.

In order to further define the genomic content of the minimal region of AI and suggest further targets for analysis within this region, NIX analysis of the available genomic sequence in the minimal region was carried out. NIX analysis is a world wide web (WWW) tool for viewing the results of multiple, consecutive DNA sequence analysis programs on a DNA sequence, to aid the identification of interesting regions in nucleic acid sequences (Williams GW, Woollard PM, Hingamp P. (1998) "NIX: A Nucleotide Identification system at the HGMP-RC" URL: <http://www.hgmp.mrc.ac.uk/NIX/>). The NIX database at the HGMP resource centre analyses genomic sequences by first masking repeat regions using Washington University's repeatmasker program. Blast searches are started using the masked sequence against the databases *ecoli*, *est*, *sts*, *embl* (minus the *sts*, *est* and *gss* sections), *trembl* and *swissprot* for DNA and RNA sequence matches. Up to 1000000 alignments can be output. Exon-finding programs (*Grail*, *Genefinder*, *Genemark*, *Fex*, *Hexon*, *Fgene*) are run using the

masked sequence. The tRNAscan program is then run on the sequence. In addition the sequence is analysed for promoter regions, poly(A) tails and CpG regions, all of which can be predictive of gene prediction and regulation. NIX analysis has the advantage of running all the above programs side by side allowing the user to detect a consensus regarding a specific feature in a submitted sequence. However, careful consideration also has to be made with NIX results due to inconsistencies within the individual databases. For instance, EST analysis is commonly undertaken by selecting expressed sequences by virtue of the poly-A regions (poly-A tail) of a mRNA. This method could also pull out genomic poly-A regions and surrounding sequences if the sample was not adequately digested with DNase. In addition, due to the high throughput manner of EST sequencing, the quality is not high and it is probable that many inaccuracies have been made. As a result there are many discrepancies within the EST database, although EST matches can be confirmed using BLAST analysis between the contig and EST sequences.

Genomic sequences were available for cosmid clones 213a7, 92f5, 34a5 and c48 and also for BACs bA252O2 and bA363E7. Other clones within the region are, to date, still being sequenced. NIX analysis revealed homology of the genomic sequence with 42 EST's in total for available sequence within the region between D9S1684 and D9S1778. Several ESTs matched homology with the MLLT3 exons and also the MLL breakpoint sequence especially in the cosmids. As shown in figures 6.16 and 6.17 the cosmid genomic sequences provided the majority of matched sequences within the region. It is interesting to note that cosmid 34a5 was predicted to contain a putative Grail promoter region and CpG island near exon 1 of the MLLT3 gene and hence demonstrating the

effectiveness of the gene prediction programs used by the NIX analysis.

Vertebrate genes have often been associated with CpG regions as methylation of vertebrate DNA (more notably 5-methylcytosine) is related to general repression of transcription and repression of selective gene mechanism. CpG islands (usually short 1-2 kb stretches of DNA in which the frequency of the dinucleotide CpG is approximate to the frequency expected from the bulk GC content of the DNA) are therefore an indicator of actively transcribed genes (Cross & Bird, 1995). A second CpG island was detected in bA363E7, although the only sequence match in close proximity to the CpG island is Rat microtubule-associated proteins 1A/1B light chain 3 (MAP1A/MAP1B LC3) cDNA.

MAP1A/MAP1B microtubule associated proteins are abundant during neuronal microtubule assembly and are thought to interact with cyclin-dependent kinase 5 (CDK5) and play a role in microtubule nucleation and phosphorylation during mitosis (Avila *et al.*, 1994; Dominguez *et al.*, 1994; Fink *et al.*, 1996; Garcia-Perez *et al.*, 1998; Moreno *et al.*, 1999; Pedrotti *et al.*, 1994; Pedrotti & Islam, 1996; Vandre *et al.*, 2000). MAP1A/MAP1B LC3 has previously been localised to chromosome 20cen-p13 (Mann & Hammarback, 1996), suggesting that the homologous sequence detected during NIX analysis may be an isoform of this gene subunit. mRNA from second gene (NCKX2) displayed homologous sequence with both BACs analysed using the NIX database. NCKX2 codes for a protein involved in Na-Ca⁺K exchange and is detected in the rod and cone outer segment plasma membrane and plays an important role in light adaptation (Guerini, 1998; Prinsen *et al.*, 2000). The exchanger is thought to play a housekeeping role in maintaining a low intracellular Ca²⁺ concentration within

numerous cell types (Blaustein & Lederer, 1999). Multiple EST's were mapped to deleted region although most were unknown cDNA library clones from various tissues. Careful positioning and screening of these ESTs and genes using expression and mutation analysis techniques may provide clues to target genes within the region important in the development of NSCLC.

6.5 Conclusion

This work has further refined the minimal region of AI on chromosome 9p22-p23 to a distance of 2-2.5 Mb in human NSCLC. Analysis of a target gene (MLLT3) which is located within this region revealed infrequent conformation abnormalities indicative of mutation, suggesting that MLLT3 was not the target gene involved in this tumour type. The results from NIX analysis of the minimal region of AI revealed the position of multiple EST's and two putative target genes indicating that further screening of this region was required to eliminate these as candidate target genes. Several other minimal regions of deletion were also identified within 9p22-p23, especially near the microsatellite loci D9S157 suggesting the presence of a putative target gene near this loci which plays a role in NSCLC. Refined mapping of this region combined with further AI analysis may reveal target areas for analysis, and this approach combined with EST mapping strategies and sequencing of clones mapping to 9p22-p23 may provide an insight into the development of NSCLC.

CHAPTER 7: ANALYSIS OF CHROMOSOME 5q11-q13 IN HUMAN NON-SMALL CELL LUNG CANCER

7.1 Introduction

Allelotype studies and loss of heterozygosity analysis has identified chromosome 5, and in particular the 5q arm, as a target for AI/LOH in human non-small cell lung cancer (Neville *et al.*, 1996; Virmani *et al.*, 1998; Field *et al.*, 1996; Fong *et al.*, 1995c; Liloglou *et al.*, 2000; Shiseki *et al.*, 1996; Tsuchiya *et al.*, 1992). Deletion mapping studies on chromosome 5q have further narrowed the region of deletion on 5q to 3 chromosomal regions. Miura *et al.* (1992) investigated loss of 5q in 13 tumours and detected deletion of the 5q13-q21 region in 75% of tumours. Hosoe *et al.* (1994 and 1996) detected deletions of 5q in 81% of tumours and found common deletion of the 5q21 region in addition to the deletion of a second region around 5q33-q35. Cooper *et al.* (1996) also detected a high frequency of deletion at 5q21 in 41% of NSCLC tumours. This study investigated mutations in the APC gene, which is located at 5q21, and detected none, indicating another target gene in this region maybe involved in the pathogenesis of NSCLC. Benachenhou *et al.* (1996) investigated LOH around the 5q11-q13 region and detected deletions in 42% of tumours. This study detected a small region of deletion within 1cM, delimited by the microsatellite loci D5S2107 and D5S624. This chapter describes AI analysis of this small region detected by Benachenhou *et al.*, using 5 polymorphic microsatellite loci located in and around 5q11-q13, in a panel of 35 NSCLC which previously demonstrated AI at the polymorphic loci D5S644 on chromosome 5q (Liloglou *et al.*, 2000). BAC and YAC physical mapping data available from web-based databases (BLAST sequence alignment and Whitehead Institute) was also used to construct a physical map of this deleted region and NIX analysis was used to analyse genomic DNA sequences with the

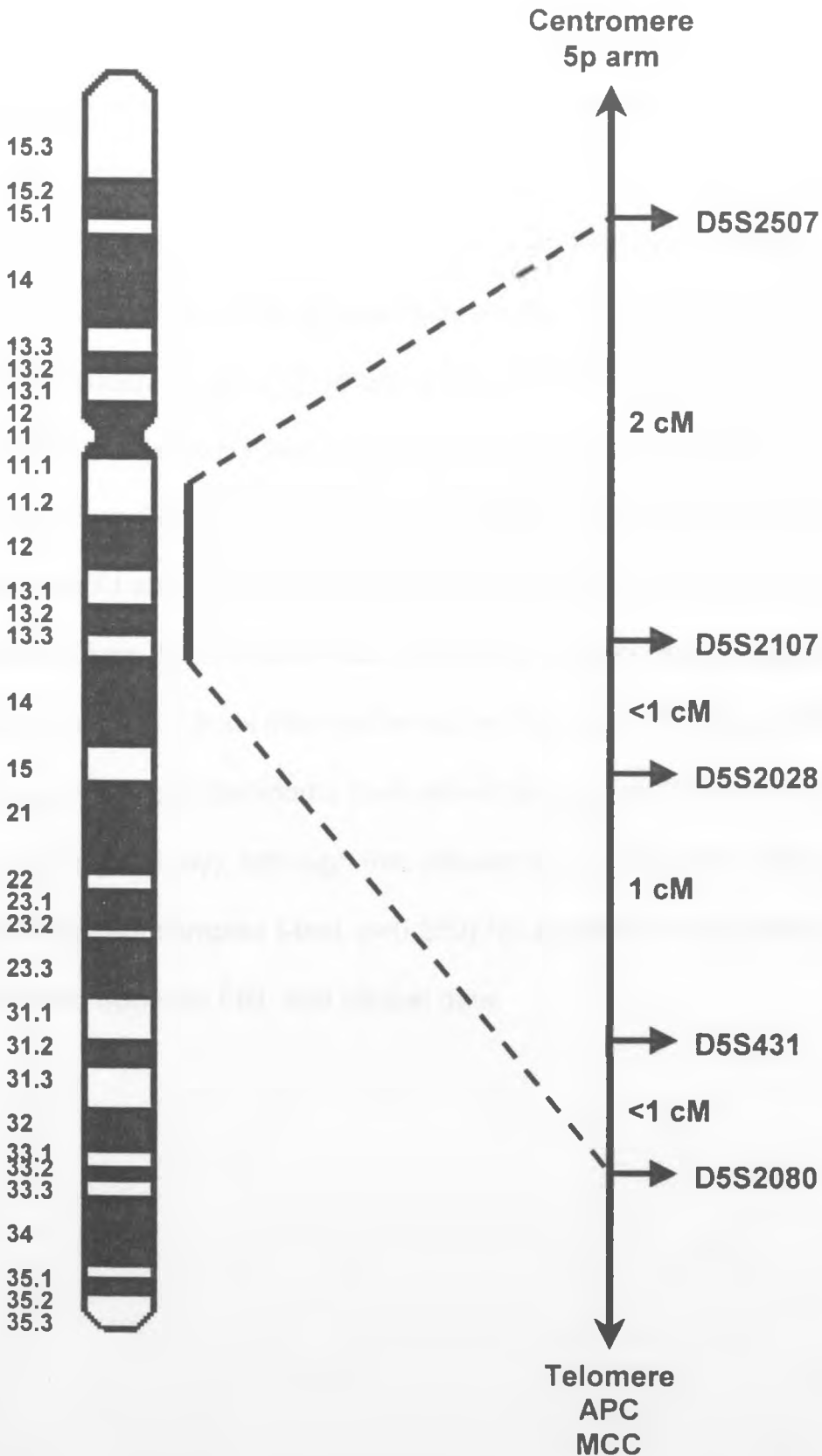
region. A list of candidate genes and EST's within the minimal region were also identified.

7.2 Analysis of chromosome 5q11-q13 in human NSCLC

7.2.1 Methods and sample selection

Tumour and normal DNA was isolated from the resected lung tissue of 35 lung cancer patients (which had previously demonstrated AI at D5S644 on chromosome 5q in Liloglou *et al*, 2000) as described in the materials and methods section. Microsatellite sequences situated in and around the region D5S2107 - D5S624 were investigated using both the Genome Database (GDB) and Whitehead Institute (WI) STS search engines. Suitable microsatellites within and around the region of interest were chosen for analysis based on the availability of the microsatellite sequence and also the average heterozygosity of the microsatellite sequence if available. Once selected for analysis, primers flanking the microsatellite repeat sequences were redesigned using the Oligo™ program as described in Chapter 2 (2.2.1). The panel of 35 tumours were analysed for allelic instability at 5 polymorphic repeat sequence locations spanning the region of interest using PCR followed by PAGE analysis as described in Chapter 2 (2.2.3). Figure 7.1 displays the position of microsatellite loci on chromosome 5q.

Figure 7.1. Diagrammatic representation of the cytogenetic position and genetic distance between microsatellite loci used in 5q11-q13 AI analysis in NSCLC.



7.2.2 AI analysis of chromosome 5q11-q13 in human NSCLC

A panel of 35 NSCLC were analysed using 5 polymorphic microsatellite loci for AI on chromosome 5q11-q13 (table 7.1). Twenty-five out of 35 tumours (71%) displayed AI at one or more microsatellite loci on 5q11-q13. As displayed in figure 7.2, the highest frequency of AI was detected at D5S2107 (16/27, 60%) and the lowest frequency at D5S2028 (7/25, 30%). Five tumours (5/35, 14%) displayed AI at all informative loci indicating complete deletion or amplification of the 5q11-q13 region. Ten tumours (10/35, 28%) displayed no AI at informative loci indicating no microsatellite-related abnormalities of the 5q11-q13 region. AI at D5S2107 was more common in squamous cell carcinoma than adenocarcinoma (12/16 and 4/16 respectively, Pearson Chi-square, $p=0.009$). Several correlations were detected between AI at two loci and these are displayed in table 7.2. No other significant association was detected between AI and clinical data such as node status and morphology. Fractional regional loss (FRL) values were calculated for each tumour as [loci on the arm scored with AI] / [total informative loci on the arm]. The mean FRL was greater in squamous cell carcinoma than adenocarcinomas (0.515 ± 0.074 and 0.372 ± 0.102 respectively), although this difference was not statistically significant (independent samples t-test, $p=0.259$) No significant associations were detected between FRL and clinical data.

Table 7.1. AI analysis of chromosome 5q11-q13 in 35 NSCLC tumours.

Table displaying AI detected in 35 NSCLC at 5q11-q13. Column 1 displays the sample number. Column 2 displays the tumour morphology (SqCCL, squamous cell carcinoma of the lung; AdenoCa, adenocarcinoma; AdenoSq, Adenosquamous carcinoma of the lung), then columns 3 and 4 the TNM classification for each tumour. Note that no tumours were recorded as being positive for metastasis at the time of diagnosis. Columns 5-9 display the results of AI analysis at 5 polymorphic loci on 5q11-q13. Key to table is as follows: no shading and H, heterozygous (no AI); black shading and AI, AI; red shading and MI, microsatellite alteration; grey shading, homozygous. Column 10 displays the fractional regional loss for each tumour.

Table 7.1. AI analysis of chromosome 5q11-q13 in 35 NSCLC tumours.

Sample	Morphology	T	N	D5S2507	D5S2107	D5S2028	D5S431	D5S2080	FRL5q
18	SqCCL	1	0	H	AI		AI	AI	0.750
1	SqCCL	2	0	H	AI	H	AI	H	0.400
2	SqCCL	2	0	AI	AI	AI	H	AI	0.800
7	SqCCL	2	0		H		H	H	0.000
10	SqCCL	2	0			H	H	H	0.000
13	SqCCL	2	0		AI	H	AI	H	0.500
22	SqCCL	2	0	H	AI	H	H	H	0.200
24	SqCCL	2	0	AI			AI	AI	1.000
25	SqCCL	2	0	AI	AI		H	AI	0.750
27	SqCCL	2	0	H	H	H	AI	AI	0.400
29	SqCCL	2	0	MI	AI		H	AI	0.750
32	SqCCL	2	0	AI	AI	AI	AI	AI	1.000
33	SqCCL	2	0		AI		H	H	0.333
38	SqCCL	2	0	H	AI	AI	AI	AI	0.800
34	SqCCL	2	0	AI	AI	H	AI	H	0.600
36	SqCCL	2	1	H			H	H	0.000
28	SqCCL	2	2		AI	H	AI	AI	0.750
37	SqCCL	3	0	H		AI	H	H	0.250
41	SqCCL	3	2	MI		H	AI	H	0.500
6	AdenoSq	2	0	H	H	H	H	H	0.000
15	AdenoSq	3	1	AI	H	H	AI	AI	0.600
30	AdenoCa			AI	AI		H	AI	0.750
4	AdenoCa	2	0				AI	AI	1.000
8	AdenoCa	2	0	H	H	H	H	H	0.000
9	AdenoCa	2	0		AI	H	H	H	0.250
12	AdenoCa	2	0	H	H	H	H	H	0.000
19	AdenoCa	2	0	H	H	H	H	H	0.000
20	AdenoCa	2	0	H	H	H	H	AI	0.200
26	AdenoCa	2	0	AI	H		AI	H	0.500
35	AdenoCa	2	0	H	H	H	H	H	0.000
21	AdenoCa	2	0		AI	AI	AI	AI	1.000
23	AdenoCa	2	0	AI		AI	H		0.667
5	AdenoCa	2	1	H	H	H	H	H	0.000
11	AdenoCa	2	1	H		H	H	H	0.000
31	AdenoCa	3	1	AI	AI	MI	AI	AI	1.000

Figure 7.2. Histogram displaying the frequency of AI at 5q11-q13 in NSCLC. Histogram displaying the frequency of AI (number of samples displaying AI at loci / number of informative samples) at 5 microsatellite loci on 5q11-q13 in 35 NSCLC tumours. The horizontal axis displays the microsatellite loci tested and the vertical axis displays the frequency of AI.

Figure 7.2. Histogram displaying the frequency of AI at 5q11-q13 in NSCLC.

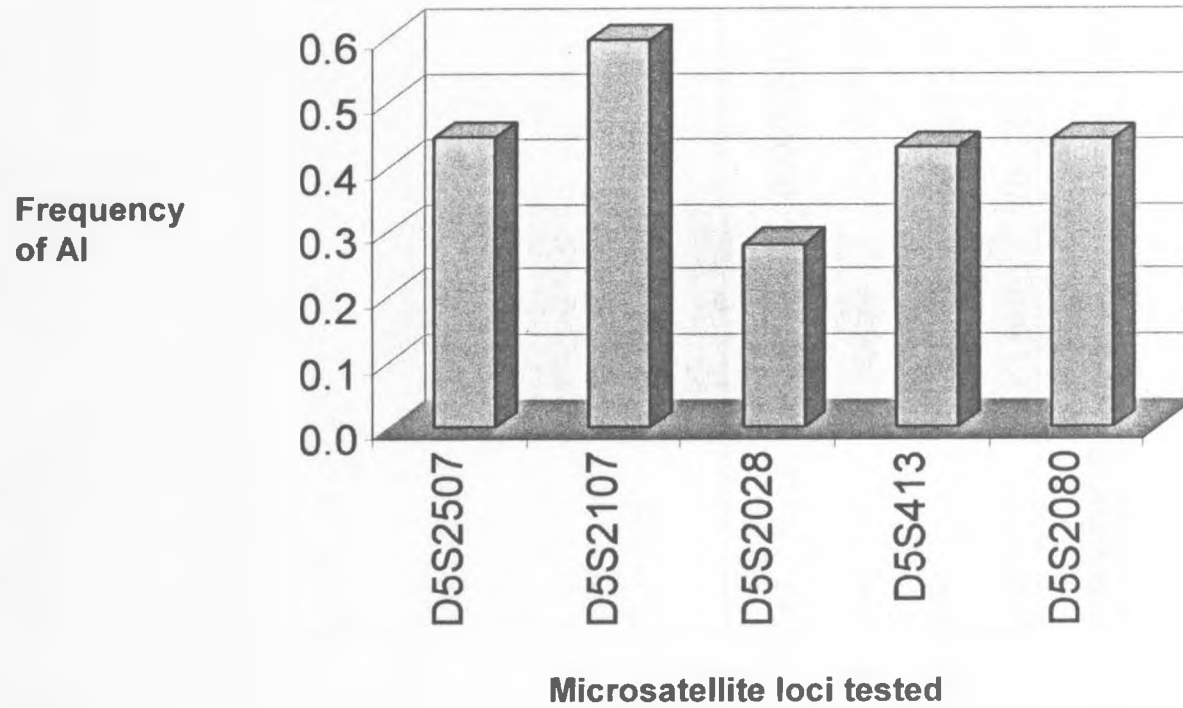


Table 7.2 Table displaying correlations between AI at two loci on chromosome 5q11-q13. Microsatellite loci displaying associated AI are displayed in the first column and the sample number in the second column. The p-value is displayed in the third column. $p \leq 0.05$ was used as a significant correlation between AI at loci.

Markers	Sample #	Significance (two-tailed)
D5S2507 and D5S2107	n=21	p=0.046
D5S2507 and D5S2080	n=26	p=0.019
D5S2028 and D5S2107	n=20	p=0.018
D5S431 and D5S2080	n=34	p=0.018

7.2.3 Multiple minimal regions of AI at 5q11-q13 in NSCLC.

Seven out of thirty-five tumours showed minimal regions of AI on chromosome 5q11-q13 (figures 7.3 and 7.4). The smallest region of AI (SRAI) was detected in four tumours (tumour nos. 1, 13, 34 and 41) around the locus D5S431, and was flanked by the loci D5S2028 and D5S2080 spanning a genetic distance of <2cM. Two tumours (Tumour 22 and the second minimal region in Tumour 1) displayed a minimal region of AI around the marker D5S2107 flanked between the loci D5S2507 and D5S2028. Tumour 26 displayed a minimal region of AI around the microsatellites D5S2028 and D5S431 flanked by the loci D5S2107 and D5S2080. The largest continuous minimal region of AI was detected in tumour 37 which was flanked by the loci D5S2507 and D5S431.

7.2.4 YAC/BAC mapping within the 5q11-q13 region

YAC hits within the 5q11-q13 region were identified by searching the Whitehead Institute (WI) database using microsatellite loci to query the database. The 5q11-q13 region on chromosome 5q was contained in the larger single-linked contig WC5.6 and the region investigated was mapped within the smaller double-linked WC-1412 contig (figure 7.5). YAC hits were obtained for all microsatellite loci except D5S5207 which was not contained in the WI database. The locus D5S2507 (also known as human STS CHLC.GGAT1D10.P18273 clone GGAT1D10) was mapped within the vicinity of D5S2102 and D5S491 using the GDB database and Marshfield chromosome 5 sex-averaged linkage map. BLAST database and Sanger Centre ACEDB analysis using microsatellite sequences revealed only one significant clone match (>20 bp) for each microsatellite loci on chromosome 5. The marker D5S2507 was contained in Homo sapiens chromosome 5 clone CTD-2023N9, D5S2107 in Homo sapiens

chromosome 5 clone CTD-2040H20, D5S2028 in Homo sapiens chromosome 5 clone RP11-479O16, D5S431 in Homo sapiens chromosome 5 clone CTD-2085F10 and D5S2080 in Homo sapiens chromosome 5 clone CTD-2122K7. PCR screening of BAC primary pools (Research Genetics) using microsatellite loci within the SRAI produced four BAC hits with D5S2028. Repeated attempts to amplify BAC colonies with D5S2080 and D5S431 gave no PCR products. As control (genomic) DNA from white blood cells amplified within the BAC PCR and produced PCR products of the correct fragment size, it was determined that D5S431 and D5S2080 may not be represented within the primary BAC library.

Figure 7.3. Diagrammatic representation of the minimal regions of AI detected on 5q11-q13 in seven lung tumours. The current cytogenetic position of microsatellite markers used for AI analysis on chromosome 5q11-q13 are displayed on the left of the figure and are extended into the figure using hatched lines. Deletions in tumours within the region are represented by vertical black bars. Sample codes for each of the seven tumours are displayed at the top of the figure. The minimal region of AI (MRAI) is indicated on the figure in blue.



Figure 7.3 Minimal regions of AI detected on 5q11-q13 in seven NSCLC.

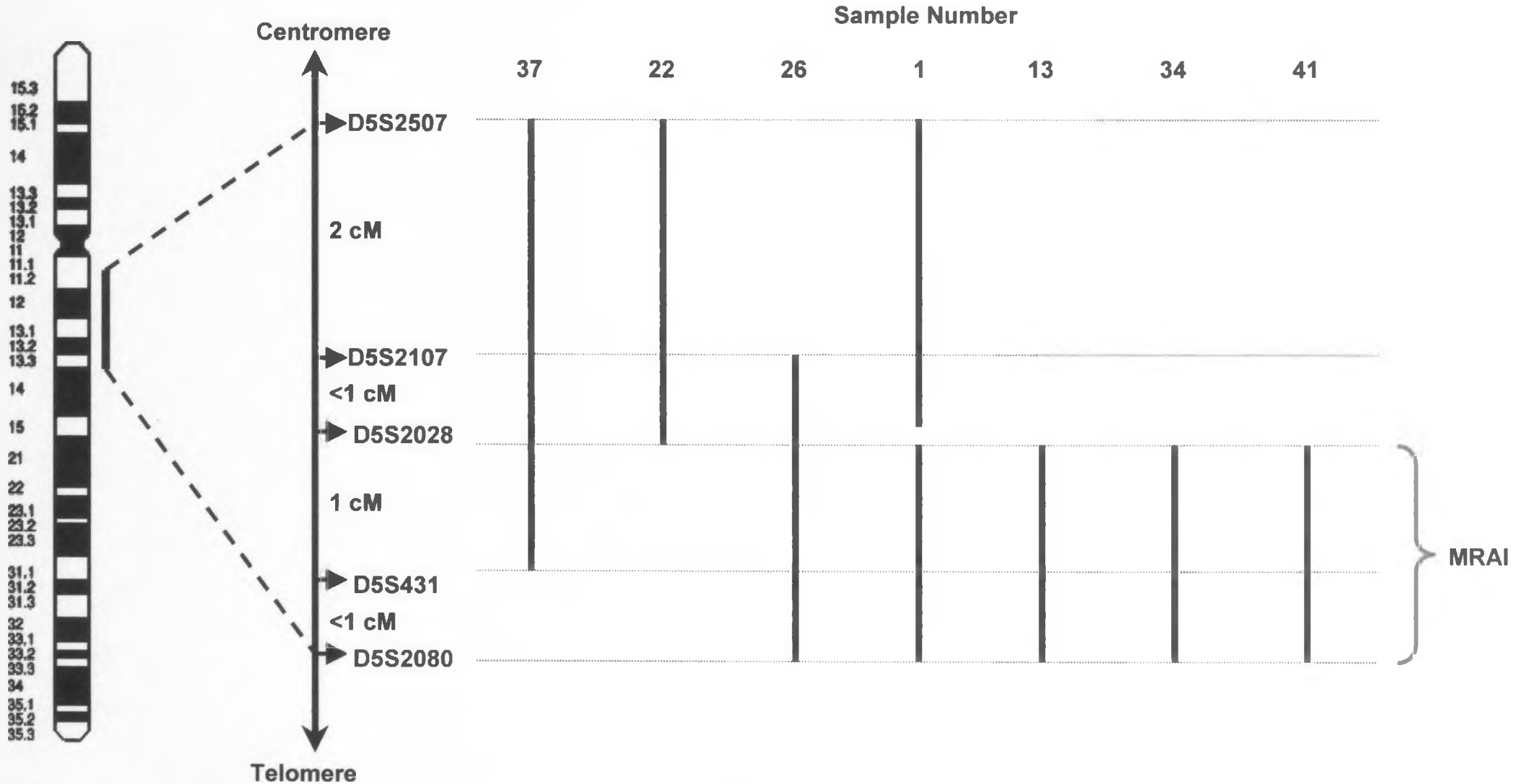


Figure 7.4 Photomicrographs displaying minimal region of AI around D5S431. Photomicrographs displaying minimal region of AI detected in 4 tumours (Tumours 1, 13, 34 and 41) around the microsatellite locus D5S431 at 5q11-q13. Microsatellite loci are displayed above each photomicrograph. Key to figure is as follows: 'N', blood (normal) DNA; 'T', lung tumour DNA. Black arrows indicate loss or amplification of allele.

Figure 7.4 Photomicrographs displaying minimal region of AI around D5S431.

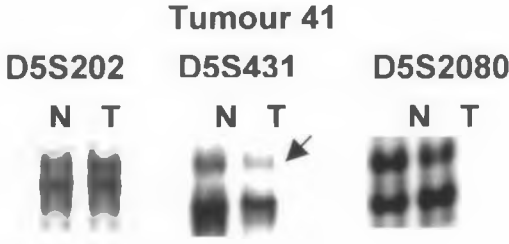
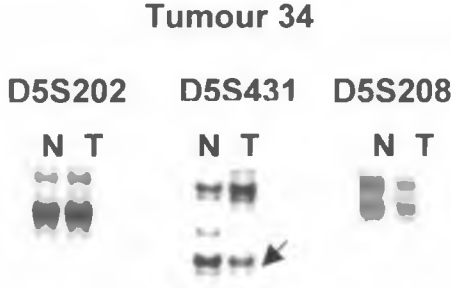
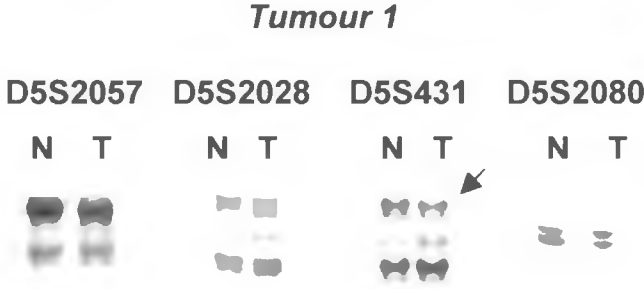
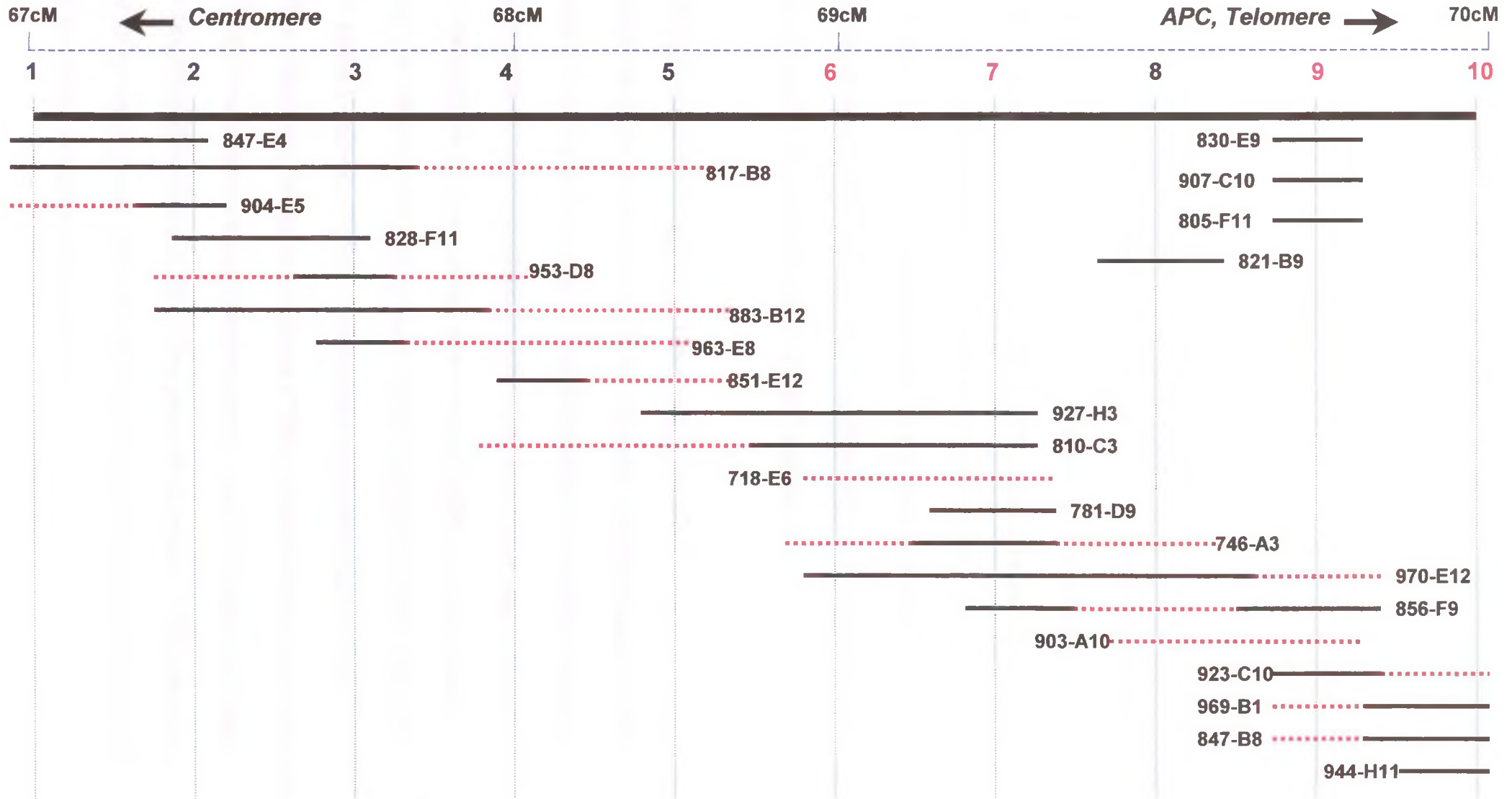


Figure 7.5 Diagrammatic representation of YAC contig WC-1412. Figure displaying the position of YAC clones and microsatellite loci within the 5q11-q13 region according to the YAC contig WC-1412 from WI database. STS markers are numbered 1-10 across the top of the figure and are extended in the figure using blue bars. The genetic distance of the region is displayed as a hatched blue bar at the top of the figure. YAC clones are displayed as black bars and hatched red bars. Black bars, verified or definite YAC/STS hit; red hatched bars, ambiguous YAC/STS hit resolved using CEPH fingerprinting or STS content data. Key to STS markers is as follows: 1, D5S491; 2, AFM183YB8; 3, D5S2102; 4, D5S398; 5, CHLC.GATA23C05; 6, D5S2107; 7, D5S2028; 8, CHLC.GATA67D03; 9, D5S431; 10, D5S2080. STS markers used in this study are marked in red. D5S2507 is not displayed on the map.

Figure 7.5 Diagrammatic representation of YAC contig WC-1412.



7.2.5 NIX database analysis of genomic DNA sequences within the SRAI.

The genomic sequences of clones RP11-479O16, CTD-2085F10 and CTD-2122K7, mapped within the shortest region of AI, were obtained from NCBI entrez nucleotide database and analysed using the HGMP resource NIX analysis database for EST and gene homology within the genomic sequence. EST search results were only considered as true hits when corresponding sequence results were classed by the NIX database as 'excellent' matches. EST and gene sequences revealed by NIX analysis were confirmed by BLAST database analysis. RP11-479O16 matched with 1 EST, Human RAS-related protein (RAB-3B) mRNA sequence and Bovine GTP-binding protein (smg p25C) mRNA. CTD-2085F10 matched with 15 EST's and Human phosphodiesterase mRNA. CTD-2122K7 matched with 18 EST's. The NIX analysis results for the 3 clones are displayed in figure 7.6.

7.2.6 Search for putative candidate genes within the SRAI.

The Genemap'99 database was analysed using the mapped positions of the microsatellite loci D5S2028, D5S431 and D5S2080 as a reference location to search for putative target genes within the SRAI which may play a role in NSCLC development. Seven genes were mapped within the region using Genemap '99. These were Interleukin 6 signal transducer (IL6ST), Integrin CD49B, alpha 2 subunit (ITGA2), Transcription elongation factor B SIII (TCEB1L), Classical Cockayne syndrome (CKN1), Kinesin heavy chain member 2 (KIF2), ADP-ribosylation factor domain protein 1 (ARFD1) and Lymphocyte antigen 64 (mouse) homolog (LY64). The genome database (GDB) was also queried for genes mapped within the 5q11-q13 region. Forty-one genes were located and these are shown in table 7.4.

Figure 7.6 Diagrammatic representation of NIX analysis of genomic sequences within the 5q11-q13 region. Genomic DNA sequences obtained from clones containing the microsatellite loci D5S2028, D5S431 and D5S2080 were analysed using the HGMP NIX analysis database. A diagrammatic representation of the positions of matched EST's and genes within the genomic sequence is displayed. Key to figure is as follows. Horizontal black lines, genomic DNA sequence for each clone (labelled from 1bp to end of sequence). Red bar, M28214 Human Ras-related protein (RAB-3B) mRNA sequence (3509-40362bp); Green bar, M19887 Bovine GTP-binding protein (smg p25C) mRNA (50938-51169bp); Blue bar, L20970 Human phosphodiesterase mRNA. Identity (including accession number) and position of ESTs within each clone is as follows. **RP11-479O16:** E1, EM:AA722515 Homo sapiens cDNA (33198-33438bp). **CTD-2085F10:** E1, EM:AA584372 Homo sapiens cDNA (1357-1581bp); E2, EM:HS1151119 Homo sapiens cDNA (16652-17041bp); E3, EM:AW268093 Homo sapiens cDNA (17144-17469bp); E4, EM:AI650348 Homo sapiens cDNA (30070-30303bp); E5, EM:AW265347 Homo sapiens cDNA (31136-31305bp); E6, EM:AV719908 Homo sapiens cDNA expressed in corresponding non cancerous liver tissue (37644-37869bp); E7, EMU:BF914230 Homo sapiens cDNA (43708-44020bp); E8, EM:HS6641 foetal brain cDNA (50816-51111bp); E9, EMU:BF897577 Homo sapiens cDNA overlapping with EM:AA707118 Homo sapiens cDNA (53881-54509bp); E10, EMU:BF913187 Homo sapiens cDNA (59905-60126bp); E11, EM:HSAA14149 Hela cell Homo sapiens cDNA (71713-71868bp); E12, EM:AA626612 Homo sapiens cDNA from Stratagene lung carcinoma 937218 (74171-74339bp); E13, EM:AV734910 Homo sapiens cDNA expressed in human pheochromocytoma (83431-83822bp); E14, EMU:BG005329 Homo sapiens cDNA (92152-

93883bp); E15, EM:AW263047 Homo sapiens cDNA (114504-115083bp). **CTD-2122K7**: E1, EM:BE042630 Homo sapiens cDNA (10594-10929bp); E2, EM:HSZZ39612 Homo sapiens cDNA (10988-11303bp); E3, EM:BE019686 Homo sapiens cDNA (12225-12505bp); E4, EM:BE019686 Homo sapiens cDNA (12518-12729bp); E5, EM:BE080460 Homo sapiens cDNA (22933-23126bp); E6, EM:HS17086 Homo sapiens foetal liver spleen cDNA (29126-29420bp); E7, EM:HS1228513 Homo sapiens cDNA (32113-32392bp); E8, EM:HS1236613 Homo sapiens cDNA (32515-33023bp); E9, EMU:BF956734 Homo sapiens cDNA (34812-34943bp); E10, EM:HSAA14297 Homo sapiens cDNA (50423-50674bp); E11, EM:AF019350 Homo sapiens brain cDNA (58579-58882bp); E12, EM:AW979271 Homo sapiens cDNA (61036-61483bp); E13, EM:BE161012 Homo sapiens cDNA overlapping with EM:AW800044 Homo sapiens cDNA (60563-60989bp); E14, EM:HS024191 Homo sapiens breast cDNA (71781-72063bp); E15, EM:HS1172892 Homo sapiens cDNA (75238-75608bp); E16, EM:HS1174381 Homo sapiens cDNA (75927-76445bp); E17, EM:AW894399 Homo sapiens cDNA (143907-144269bp); E18, EM:HS49667 Stratagene lung (#937210) Homo sapiens cDNA (165939-166156bp).

Figure 7.6 NIX analysis of genomic sequences within the 5q11-q13 region

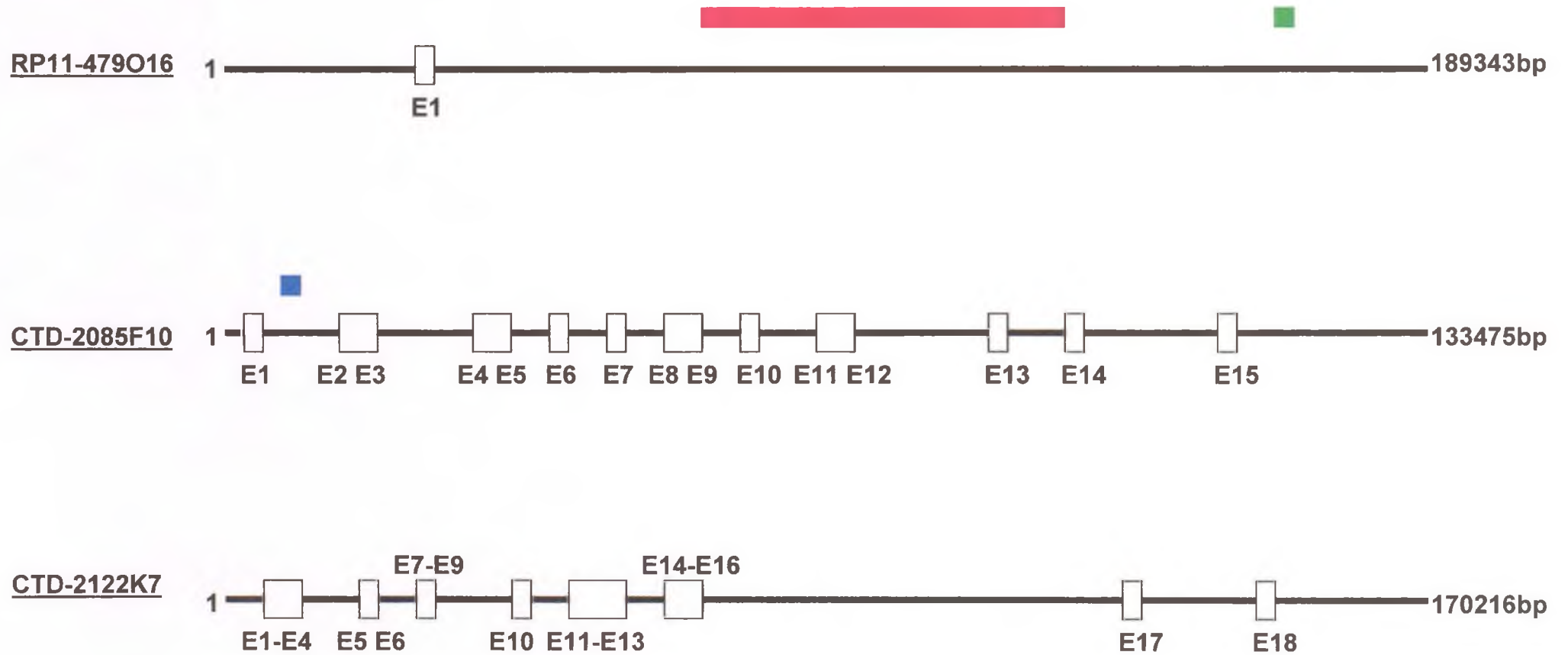


Table 7.4 Table displaying candidate genes mapped within the 5q11-q13 region according to the genome database (GDB). The first column displays the gene name and abbreviation and the second column the currently mapped cytogenetic location (as of date 11.03.01). An asterisk after the cytogenetic location indicates HUGO approval.

Table 7.4 GDB candidate genes within the 5q11-q13 region.

Candidate gene name	Cytogenetic Location
Adaptor-related protein complex 3, beta 1 subunit (AP3B1)	5q12.1-5q13.3
ADP-ribosylation factor domain protein 1 (ARFD1)	5q11.2-5q12.1
Argininosuccinate synthetase pseudogene 9 (ASSP9)	5q11-5q12*
Betaine-homocysteine methyltransferase 2 (BHMT2)	5q13
Cadherin 12 (N-cadherin 2) pseudogene (CDH12P)	5q13
cadherin-related neuronal receptor 2 (CNR2)	5q13
Chondroitin sulfate proteoglycan 2 (CSPG2)	5q12-5q14*
Coagulation factor II (thrombin) receptor (F2R)	5q13*
Creatine kinase, mitochondrial 2 (CKMT2)	5q13.3*
Cyclin B1	5q11-5q12*
Cyclin H	5q13.3-5q14
Dihydrofolate reductase (DHFR)	5q11.2-5q13.2*
Eukaryotic translation elongation factor 1 beta 3 (EEF1B)	5q12-5q14*
Forkhead (Drosophila)-like 8 (FKHL8)	5q12-5q13
General transcription factor IIH, polypeptide 2 (GTF2H2)	5q12.2-5q13.3
Heat shock 27kD protein-like 4 (HSPBL4)	5q11.2
hexosaminidase B (beta polypeptide, HEXB)	5q13*
Kinesin heavy chain member 2 (KIF2)	5q12-5q13
Microtubule-associated protein 1B (MAP1B)	5q13*
MutS (E. coli) homolog 3 (MSH3)	5q11-5q12
Nuclear restricted protein, BTB domain-like (NRPB)	5q12-5q13.3
Orthopedia (Drosophila) homolog (OTP)	5q13.3
Phosphatidylinositol 3-kinase, regulatory subunit, polypeptide 1 (PIK3R1)	5q12-5q13*
Phosphodiesterase 4D, cAMP-specific (dunce (Drosophila)-homolog)	5q12
Phosphodiesterase 8B (PDE8B)	5q12.3-5q13.3
Pro-melanin-concentrating hormone-like 2 (PMCHL20)	5q12-5q13
RAD17 (S. pombe) homolog	5q13
RAS p21 protein activator (GTPase activating protein)	5q13.3
Ras protein-specific guanine nucleotide-releasing factor 2 (RASGRF2)	5q13
Schizophrenia disorder 1 (SCZD1)	5q11.2-5q13.3
Small EDRK-rich factor 1 (SERF1)	5q12.2-5q13.3
Spinal muscular atrophy (Werdnig-Hoffmann disease)	5q12.2-5q13*
Spinal muscular atrophy 3 (SMA3)	5q13
Spinal muscular atrophy 4 (SMA4)	5q13
Spinal muscular atrophy 5 (SMA5)	5q13
Survival of motor neuron 1, telomeric (SMN1)	5q13
Survival of motor neuron 2, centromeric (SMN2)	5q13
Thrombospondin 4 (THBS4)	5q12-5q13
X-ray repair complementing defective repair in Chinese hamster cells 4 (XRCC4)	5q13.1-5q13.3
X-ray repair complementing defective repair in Chinese hamster cells 4-like (XRCC4L)	5q13-5q14
Zinc finger protein-5 (ZNF5)	5q12-5q13

7.3 Discussion

7.3.1 AI analysis of 5q11-q13 further reduced the minimal region of deletion around D5S431 in NSCLC.

Allelic imbalance on chromosome 5 is a frequent occurrence in multiple tumour types including colon, NSCLC, SCLC, oesophageal, gastric, breast and testicular carcinomas (Ashton-Rickardt *et al.*, 1991; Benachenhou *et al.*, 1999; Boynton *et al.*, 1992; Miura *et al.*, 1992; Peng *et al.*, 1999; Sano *et al.*, 1991; Vogelstein *et al.*, 1988; Wieland *et al.*, 1996). Deletion mapping has implicated (and in the case of colorectal cancer demonstrated) several regions on chromosome 5q which may harbour candidate genes involved in the pathogenesis of these tumour types (Ashton-Rickardt *et al.*, 1991; Benachenhou *et al.*, 1998; Boynton *et al.*, 1992; Cooper *et al.*, 1996; Hosoe *et al.*, 1994b; Kinzler *et al.*, 1991; Nakamura *et al.*, 1992; Nakamura *et al.*, 1991; Peng *et al.*, 1999; Ueno *et al.*, 1998; Wieland & Bohm, 1994; Wieland *et al.*, 1996). Studies in NSCLC have identified multiple regions on chromosome 5q which demonstrate allelic loss including 5q21 (Ashton-Rickardt *et al.*, 1991; Fong *et al.*, 1995c) and 5q11-q13 (Benachenhou *et al.*, 1998; Wieland & Bohm, 1994). A recent study by Mendes-da-Silva *et al* (2000) identified allelic loss at 5 regions of chromosome 5q (5q11.2-q12.2, 5q15 (D5S644 locus), 5q22.3-q23.1, 5q31.1, and 5q35.3) in human NSCLC. Nakamura *et al* (1991 and 1992) demonstrated that somatic alterations of the MCC and APC genes on chromosome 5q21 by mutation, insertion or deletion occurred frequently in sporadic colorectal tumours and also in the germline of individuals with hereditary familial polyposis coli (FAP). Cooper *et al* (1996) investigated allelic deletions in NSCLC tumours at 4 polymorphic loci in the APC and MCC genes at 5q21 and detected LOH at two polymorphic loci in 41% of cases, however

SSCP analysis of the mutation cluster region of the APC and MCC genes revealed no mutations indicating another gene within this region involved in NSCLC.

This chapter describes AI analysis of 35 NSCLC tumours, previously demonstrated in Liloglou *et al* (2000) to display AI on chromosome 5q around the microsatellite D5S644 at 5q13, using 5 polymorphic loci in and around the shortest region of deletion detected by Benachenhou in an attempt to narrow the region of deletion on chromosome 5q11-q13. A high frequency of tumours (71%) displayed AI in this chromosomal region, although this was probably due to the tumour population selected and the criteria for selection (AI at D5S644). Statistical analysis of AI data revealed that AI at D5S2107 and the mean FRL was greater in squamous cell carcinomas than adenocarcinomas. Previous studies have demonstrated differences in the frequency of chromosome 5 deletion between tumours with a bronchial origin such as SqCCL and peripheral tumours such as AdenoCa (Cooper *et al.*, 1996; Fong *et al.*, 1995c; Sato *et al.*, 1994). This indicates that 5q11-q13 may be more involved in the pathogenesis of SqCCL than AdenoCa. Minimal regions of AI were detected in 7 tumours within the 5q11-q13 region. In comparison with Benachenhou *et al* (1998) 4 tumours displayed a shortest minimal region of AI around the microsatellite loci D5S431. The SRAI described in this chapter was delimited by the microsatellite loci D5S2028 and D5S2080, which are currently mapped closer to D5S431 than the loci D5S2107 and D5S624 which delimited the shortest region of deletion in Benachenhou *et al*, hence further narrowing the minimal region of AI to 1cM within the 5q11-q13 region.

7.3.2 Physical mapping and NIX analysis within the minimal region of AI.

The Sanger ACeDB and Whitehead Institute databases were queried for YAC and BAC clones mapping to the region D5S2507 – D5S2080 to confirm loci order. Microsatellites loci were mapped to YAC clones within the 5q11-q13 region using the contig WC-1412 in the WI database. D5S2507 was not contained in this contig but was mapped near D5S2102 and D5S491 using a separate genetic map of the region from GDB. Analysis of the Sanger ACeDB gave no BAC colony addresses for the microsatellite loci used in this analysis. This was probably due to under-representation of BAC clones for chromosome 5 within this database. Screening of BAC colony primary pools also proved unsuccessful with matches for 1/5 loci. BLAST analysis using microsatellite sequences matched genomic DNA clones for all 5 loci within the 5q11-q13 region. The genomic DNA sequence for clones RP11-479O16, CTD-2085F10 and CTD-2122K7 were analysed using the HGMP NIX analysis database as they were located within or flanked the minimal region of AI described in this chapter.

RP11-479O16 (containing D5S2028) displayed homology with Human RAS-related protein (RAB3B) mRNA and Bovine GTP-binding protein (smg p25C) mRNA. RAB3B was isolated from a human pheochromocytoma cDNA library and encodes a 23-25 kDa polypeptide which shares approximately 30% homology with the ras superfamily (Zahraoui *et al.*, 1989). RAB3B can bind GTP and exhibit GTPase activities, and the pattern and regulation of RAB3B targeting in epithelial cells implicates the GTPase as a candidate regulator of apical and/or junctional protein traffic in epithelial tissues (Weber *et al.*, 1994).

Another member of the RAB family of genes (RAB5A) previously demonstrated overexpression in lung tumour cell lines (Yu *et al.*, 1999), although no evidence exists for RAB3B abnormalities in NSCLC. Bovine GTP-binding protein was isolated from a bovine brain cDNA library and like RAB3B shares homology with the ras family (Matsui *et al.*, 1988). The Ras p21 oncogene is frequently overexpressed in NSCLC (Spandidos *et al.*, 1990) and GTP-binding proteins are known to play a major role in cell signalling and apoptosis (Leblanc *et al.*, 1999). This evidence suggests that homologous genomic sequences exist within, or nearby the minimal region of AI, and within clone RP11-479O16 that are similar to RAB3B or Bovine GTP-binding protein and may be involved in NSCLC. CTD-2085F10 (containing D5S431) displayed homology with human phosphodiesterase mRNA which is closely related to the *Drosophila dunce* (*dnc*) gene and encodes cyclic AMP-specific phosphodiesterases (Bolger *et al.*, 1993). The *dnc* gene is thought to be involved in learning, memory and female fertility in *Drosophila* (Qiu *et al.*, 1991) which makes human phosphodiesterase mRNA an unlikely candidate for a role in NSCLC. Thirty-four EST's were mapped to the three clones within or flanking the minimal region of AI using NIX analysis. Although almost all EST's which displayed homology to the genomic sequences of the clones were unknown, EST12 which displayed homology with CTD-2085F10 was isolated from a Stratagene lung carcinoma. EST12 therefore makes an interesting candidate for analysis as it is mapped around the locus D5S431 which was located within the minimal region of AI described in this chapter.

7.3.3 Candidate genes within the region 5q11-q13

Using Genemap '99 and the genome database (GDB) web site a large number of candidate genes were mapped to the 5q11-q13 region. The exact positioning of genes within the region is still subject to controversy and it is unknown whether they are mapped near to the minimal region of AI or even between the microsatellite loci described in this chapter due to incomplete mapping data for the region. As described in section 7.2.6 and shown in table 7.4 GDB and Genemap '99 have mapped 48 genes within the 5q11-q13 region. Noticeable candidates for analysis include genes involved in the cell cycle such as Cyclin B1 and Cyclin H and genes involved in DNA repair such as MSH3, XRCC4 and XRCC4L. Benachenhou *et al* (1998) previously investigated the mutation status of the DNA mismatch repair gene MSH3 in NSCLC tumours with hemizygous deletions of the 5q11-13 region and detected two novel polymorphisms. Although no mutations of MSH3 were detected in this study, this does not discount inactivation of the MSH3 gene through other genetic processes such as gene deletion. Cyclin B1 is involved in the G2-M-phase transition during the cell cycle and has shown to be overexpressed in NSCLC and especially SqCCL (Soria *et al.*, 2000) providing further evidence that this gene may be a strong candidate for analysis. There are no documented abnormalities of Cyclin H, XRCC4 or XRCC4L in NSCLC, although studies investigating the involvement of these three genes in lung tumours are infrequent. Until the genetic position of these and other candidate genes within the 5q11-q13 region can be accurately determined through physical mapping techniques, it is unwise to speculate on their role in NSCLC development based on the evidence of minimal regions of deletion detected using AI analysis and deletion mapping techniques. Further studies are therefore needed to correctly map the 5q11-q13 region, especially

around the loci D5S431, which may harbour genes involved in the pathogenesis of NSCLC.

7.4 Conclusion

This chapter has described the analysis of 5q11-q13 in NSCLC using a combination of AI analysis, physical mapping and NIX database analysis. This work further reduced the minimal region of AI/LOH located around the microsatellite locus D5S431 at 5q11-q13 which was detected in previous studies and has indicated the presence of a gene within the region spanned by the microsatellite loci D5S2028 – D5S2080 which may be involved in NSCLC. AI analysis also detected differences in the frequency of AI within the 5q11-q13 region between SqCCL and AdenoCa tumours providing further evidence for different genetic pathways involved in the pathogenesis of these two morphological types. NIX and other database analysis mapped multiple EST's and candidate genes within the region. Three candidates for further analysis included the Ras-like GTP-binding proteins RAB3B and Bovine GTP-binding protein and EST AA626612, mapped near D5S431, isolated from a lung carcinoma cell line.

CHAPTER 8: CONCLUDING REMARKS AND FUTURE CONSIDERATIONS

Lung cancer is the most common human malignancy in the world today, and has been mainly attributed to cigarette smoking. Although advances in diagnosis and chemotherapeutic techniques have recently been made, high mortality and low five-year survival for this disease is a common occurrence due to the advanced nature of lung cancer at the time of presentation (Coleman, 1999). Advances in molecular biology (such as PCR-based microsatellite analysis, gene mapping and expression analysis) to study the genetic changes occurring in human tumours in both the 1980's and 1990's have revealed the importance of proto-oncogenes, DNA repair genes and tumour-suppressor genes in the multistep carcinogenesis pathway during tumour development. This has led to a considerable growth in the knowledge and understanding of critical genes which are involved in different human cancers including lung tumours. The purpose of the work presented in this thesis was to apply a molecular genetic approach to the investigation of genetic alterations associated with the development of human non-small cell lung cancer. The objective of this was to provide information regarding clonality and tumour progression within individuals with lung cancer and also target potential genes involved in the development of lung tumours.

In view of the small biopsy size and limited number of specimens collected from lung samples and the wide range of molecular markers available for genetic analysis of patterns of allelic instability within these tissues, techniques were initially developed for the rapid isolation of DNA from lung tissue and high throughput microsatellite analysis. Chapter three described the selection of a high throughput chromosome 17 microsatellite panel for accurate AI detection in NSCLC based on techniques developed by Liloglou *et al* (2000).

AI was detected in a panel of 64 NSCLC tumours on both 17p and 17q and the frequency of AI detected on both chromosome 17 arms was also significantly higher than previously documented studies in NSCLC. Three minimal regions located at 17p12-p13 were also detected within a genetic distance of 5cM. This work supported evidence for a high frequency of AI on chromosome 17 and demonstrated accurate AI detection using fluorescent markers. Although highly informative with improved accuracy for AI detection, this technique was limited to fluorescent microsatellite markers available within the ABI prism™ high density linkage mapping set. It would have been constructive to study the 5cM region of AI detected in this work using other fluorescent microsatellite markers located at 17p12-p13, narrowing the region of AI with an aim to enable mapping techniques to locate potential genes involved in lung carcinogenesis. However the linkage mapping set was restricted to approximately 5cM distances between markers, a distance which is too large for common mapping/gene detection techniques. Further investigation would therefore involve either analysis using non-fluorescent AI approaches such as silver staining or autoradiography which would reduce sensitivity and efficiency, or the development of 'in-house' fluorescent makers which would require further optimisation and assessment prior to AI analysis.

In view of the high frequency and accurate quantification of AI using fluorescent multiplex analysis within small DNA samples this method was applied to the sequential AI analysis of tissue samples from individuals with lung tumours at multiple chromosomal regions. Previous studies have relied upon classical LOH techniques such as autoradiography and silver staining to detect microsatellite abnormalities in lung tumours and the surrounding tissue (Sozzi *et al.*, 1992;

Sozzi *et al.*, 1995; Tsuchiya *et al.*, 1992; Wistuba *et al.*, 1999a; Wistuba *et al.*, 1999b; Wistuba *et al.*, 1999c; Wistuba *et al.*, 1997a). Although these techniques were generally informative, large studies can be time consuming and require large quantities of DNA of suitable quality, something which is not always available from small tissue samples such as biopsy and paraffin-embedded material. Chapter four described methods for isolating small subpopulations of cells from bronchial epithelial tissue using laser capture microdissection and direct lysis of cells, followed by AI analysis at multiple microsatellite locations using fluorescent multiplex PCR. Initial efforts concentrated on DNA isolation from small amounts of paraffin embedded tissue using microdissection with syringe needles. This method was replaced with Laser Capture Microdissection which provided a fast, reliable method for cell isolation from small tissue samples prior to DNA analysis.

These techniques were employed to undertake AI analysis of sequential regions of bronchial epithelium and tumour in three tumour bearing lungs. Results revealed AI in common with the tumour within histologically normal and metaplastic bronchial epithelium at multiple regions throughout the airway. Individually, each lung displayed different patterns of AI within 1cm intervals and separate cell populations from the same topographical region of the lung displayed different patterns of AI, providing further evidence for field carcinogenesis occurring within NSCLC development. By expanding this analysis and taking closely adjacent slices of each portion of the whole lung in a patient with lung cancer (chapter five) it was demonstrated that multiple patches of AI not only occurred in bronchial tissue near to the site of the tumour, but could occur in every lobe tested at sites distant from the tumour itself. Multiple

clonal cell populations were also detected within each lobe of the lung tested and also within the tumour.

The work presented in chapters four and five raises several questions regarding lung tumours and their development.

- 1) Do other lung cancer patients harbour genetic abnormalities in histologically normal tissue at sites distant from the tumour?
- 2) What is the likelihood of histologically normal but genetically abnormal bronchial cells progressing to form lung tumours?
- 3) Does histologically normal but genetically abnormal lung tissue have clinical implications when dealing with surgical resection margins?

Previous studies have established the existence of genetic abnormalities in histologically normal lung tissue adjacent to the site of the tumour or within bronchial lavage samples (Field *et al.*, 1999; Miozzo *et al.*, 1996; Park *et al.*, 1999; Park *et al.*, 2000; Waber *et al.*, 1996; Walker *et al.*, 1994; Wistuba *et al.*, 1999a; Wistuba *et al.*, 2000b; Wistuba *et al.*, 1999b; Wistuba *et al.*, 1997a). Further analyses similar to that undertaken in chapters 4 and 5 will give an indication whether other lung tumour patients harbour genetic abnormalities in histologically normal tissue. However it is important to note that whole lung specimens should be carefully vetted for this type of study to reduce the likelihood of 'background' or non-tumour specific AI caused by the secondary effects of respiratory tumours such as consolidation, pulmonary infection or other respiratory disease (Liloglou *et al.*, 2001). In addition careful pathological

collaboration/supervision must be recruited to sample lung tissue for a study of this type.

It is unlikely that every genetic lesion within histologically normal bronchial epithelium will result in tumour formation, although this may in part explain the common occurrence of secondary tumours in patients with lung cancer (Hiroshima *et al.*, 1998). Collection of lung tissue from multiple individuals using these multi-sampling methods, together with the development of further fluorescent microsatellite panels incorporating other chromosomal regions would identify common genetic lesions throughout the whole lung in multiple individuals and help distinguish between tumour-specific and non-specific (background) genetic damage. Genetic lesions common to both tumour tissue and histologically normal tissue may then indicate regions where tumour development is likely in these individuals and identify potential target areas for tumour prevention strategies through either surgical resection or chemoprevention measures. For instance if a whole lung has genetic abnormalities but a small pathological lesion, surgical resection around the site of the tumour may be sufficient to remove the tumour itself, however chemoprevention could be developed to target other cells within the lung displaying normal histology but harbouring common genetic abnormalities to prevent secondary tumour development.

The work described in chapter six further refined the minimal region of AI on chromosome 9p22-p23 to a distance of 2-2.5 Mb in human NSCLC. Analysis of a target gene (MLLT3), which is located within a minimal region of AI detected during this study, revealed infrequent conformation abnormalities indicative of

mutation, suggesting that MLLT3 was not the target gene involved in this tumour type. NIX analysis of the minimal region of AI revealed the position of multiple EST's and two other potential target genes within this region positioning, suggesting that further screening of this region was required. Several other minimal regions of deletion were also identified within 9p22-p23, especially near the microsatellite locus D9S157, again suggesting the presence of a putative target gene near this loci which plays a role in NSCLC (Neville *et al.*, 1995). Further investigation of these minimal regions of AI are therefore required possibly incorporating fine microsatellite mapping and sequence analysis methods.

Chapter 7 also described the analysis of 5q11-q13 in NSCLC using a combination of AI analysis, physical mapping and NIX database analysis. This work further reduced the minimal region of AI/LOH located around the microsatellite locus D5S431 at 5q11-q13 which was detected in previous studies (Benachenhou *et al.*, 1998; Mendes-da-Silva *et al.*, 2000; Wieland & Bohm, 1994) and has indicated the presence of a gene within the region spanned by the microsatellite loci D5S2028 – D5S2080 which may be involved in NSCLC. AI analysis detected differences in the frequency of AI within the 5q11-q13 region between SqCCL and AdenoCa tumours providing further evidence for different genetic pathways involved in the pathogenesis of these two morphological types. NIX and other database analysis mapped multiple EST's and candidate genes within the region. Three candidates for further analysis included the Ras-like GTP-binding proteins RAB3B and Bovine GTP-binding protein and EST AA626612, mapped near D5S431, isolated from a lung carcinoma cell line. SSCP followed by mutation/expression analysis,

similar to that described for the MLLT3 gene in chapter 6, would show whether RAB3B, Bovine GTP-binding protein or EST AA626612 were targets in NSCLC. Further fine mapping in NSCLC tumours using microsatellite markers within the region of AI detected in this study would also support current mapping and positional cloning strategies to enable the identification of putative target genes within this region involved in NSCLC.

The recent publication of the draft sequence for the human genome (Lander *et al.*, 2001) provided detailed contig maps and sequence data for these two regions on chromosome 9p22-p23 and 5q11-q13 which will aid and complement future gene searching techniques such as high throughput cDNA array and protein expression analysis. Analysis of DNA sequences using NIX and other databases and comparison with other genomes from bacteria, yeast, *Drosophila*, mouse and *Fugu subripes* will also help to identify putative target genes within these regions which may be involved in NSCLC.

Contemporary advances for targeting lung cancer have shifted from the detection of clinically evident late stage tumours to more asymptomatic preinvasive and early invasive lesions (Hirsch *et al.*, 2001). Traditional approaches to lung cancer detection involve diagnostic tests such as radiography and sputum cytology, however these are now complemented with improved bronchoscopic techniques such as laser-induced fluorescence endoscope (LIFE) bronchoscopy and low dose spiral computed tomography (spiral CT) which can detect smaller tumours than conventional radiography and CT (Ellis & Gleeson, 2001; Shibuya *et al.*, 2001; van Rens *et al.*, 2001). Recognition of molecular biomarkers such as specific patterns of AI or

abnormalities in target genes within lung tumours, preneoplastic tissue and sputum specimens may aid the identification of high risk individuals for further screening using improved diagnostic techniques and/or chemoprevention for lung cancer prior to clinical presentation and before the development of invasive lung carcinoma.

CHAPTER 9: REFERENCES

1. [Faint reference text]

2. [Faint reference text]

3. [Faint reference text]

4. [Faint reference text]

5. [Faint reference text]

6. [Faint reference text]

7. [Faint reference text]

8. [Faint reference text]

9. [Faint reference text]

10. [Faint reference text]

- Aaltonen, L.A. & Peltomaki, P. (1994). Genes involved in hereditary nonpolyposis colorectal carcinoma. *Anticancer Res*, **14**, 1657-60.
- Aaltonen, L.A., Peltomaki, P., Mecklin, J.P., Jarvinen, H., Jass, J.R., Green, J.S., Lynch, H.T., Watson, P., Tallqvist, G., Juhola, M. & et al. (1994a). Replication errors in benign and malignant tumors from hereditary nonpolyposis colorectal cancer patients. *Cancer Res*, **54**, 1645-8.
- Aaltonen, L.A., Sankila, R., Mecklin, J.P., Jarvinen, H., Pukkala, E., Peltomaki, P. & de la Chapelle, A. (1994b). A novel approach to estimate the proportion of hereditary nonpolyposis colorectal cancer of total colorectal cancer burden. *Cancer Detect Prev*, **18**, 57-63.
- Adamson, R., Jones, A.S. & Field, J.K. (1994). Loss of heterozygosity studies on chromosome 17 in head and neck cancer using microsatellite markers. *Oncogene*, **9**, 2077-82.
- Ah-See, K.W., Cooke, T.G., Pickford, I.R., Soutar, D. & Balmain, A. (1994). An allelotype of squamous carcinoma of the head and neck using microsatellite markers. *Cancer Research*, **54**, 1617-1621.
- Amos, C.I., Xu, W. & Spitz, M.R. (1999). Is there a genetic basis for lung cancer susceptibility? *Recent Results Cancer Res*, **151**, 3-12.
- An, H.X., Claas, A., Savelyeva, L., Seitz, S., Schlag, P., Scherneck, S. & Schwab, M. (1999). Two regions of deletion in 9p23-24 in sporadic breast cancer. *Cancer Res*, **59**, 3941-3.
- An, H.X., Niederacher, D., Picard, F., van Roeyen, C., Bender, H.G. & Beckmann, M.W. (1996). Frequent allele loss on 9p21-22 defines a smallest common region in the vicinity of the CDKN2 gene in sporadic breast cancer. *Genes Chromosomes Cancer*, **17**, 14-20.

- Anguita, E., Barrio, C.G., Gonzalez, F.A., Ferro, M.T., del Potro, E., Ropero, P. & Villegas, A. (2000). Association of t(9;11)-MLL AF9 and Trisomy 8 in an AML-M5 Preceded by Pancytopenia. *Cancer Genet Cytogenet*, **120**, 144-147.
- Ashton-Rickardt, P.G., Wyllie, A.H., Bird, C.C., Dunlop, M.G., Steel, C.M., Morris, R.G., Piris, J., Romanowski, P., Wood, R., White, R. & et al. (1991). MCC, a candidate familial polyposis gene in 5q.21, shows frequent allele loss in colorectal and lung cancer. *Oncogene*, **6**, 1881-6.
- Auerbach, O. & Garfinkel, L. (1991). The changing pattern of lung carcinoma. *Cancer*, **68**, 1973-7.
- Augenlicht, L.H., Richards, C., Corner, G. & Pretlow, T.P. (1996). Evidence for genomic instability in human colonic aberrant crypt foci. *Oncogene*, **12**, 1767-72.
- Avila, J., Ulloa, L., Gonzalez, J., Moreno, F. & Diaz-Nido, J. (1994). Phosphorylation of microtubule-associated proteins by protein kinase CK2 in neuritogenesis. *Cell Mol Biol Res*, **40**, 573-9.
- Baker, J.C. & Jacobson, M.K. (1986). Alteration of adenyl dinucleotide metabolism by environmental stress. *Proc Natl Acad Sci U S A*, **83**, 2350-2.
- Banks, R.E., Dunn, M.J., Forbes, M.A., Stanley, A., Pappin, D., Naven, T., Gough, M., Harnden, P. & Selby, P.J. (1999). The potential use of laser capture microdissection to selectively obtain distinct populations of cells for proteomic analysis--preliminary findings. *Electrophoresis*, **20**, 689-700.

- Barnes, L.D., Garrison, P.N., Sibrashvili, Z., Guranowski, A., Robinson, A.K., Ingram, S.W., Croce, C.M., Ohta, M. & Huebner, K. (1996). Fhit, a putative tumor suppressor in humans, is a dinucleoside 5',5'''- P1,P3-triphosphate hydrolase. *Biochemistry*, **35**, 11529-35.
- Barrett, M.T., Reid, B.J. & Joslyn, G. (1995). Genotypic analysis of multiple loci in somatic cells by whole genome amplification. *Nucleic Acids Res*, **23**, 3488-92.
- Baxi, M.D., McLennan, A.G. & Vishwanatha, J.K. (1994). Characterization of the HeLa cell DNA polymerase alpha-associated Ap4A binding protein by photoaffinity labeling. *Biochemistry*, **33**, 14601-7.
- Belinsky, S.A., Nikula, K.J., Palmisano, W.A., Michels, R., Saccomanno, G., Gabrielson, E., Baylin, S.B. & Herman, J.G. (1998). Aberrant methylation of p16(INK4a) is an early event in lung cancer and a potential biomarker for early diagnosis. *Proc Natl Acad Sci U S A*, **95**, 11891-6.
- Benachenhou, N., Guiral, S., Gorska-Flipot, I., Labuda, D. & Sinnett, D. (1998). High resolution deletion mapping reveals frequent allelic losses at the DNA mismatch repair loci hMLH1 and hMSH3 in non-small cell lung cancer. *Int J Cancer*, **77**, 173-80.
- Benachenhou, N., Guiral, S., Gorska-Flipot, I., Labuda, D. & Sinnett, D. (1999). Frequent loss of heterozygosity at the DNA mismatch-repair loci hMLH1 and hMSH3 in sporadic breast cancer. *Br J Cancer*, **79**, 1012-7.
- Bennett, W.P., Alavanja, M.C., Blomeke, B., Vahakangas, K.H., Castren, K., Welsh, J.A., Bowman, E.D., Khan, M.A., Flieder, D.B. & Harris, C.C. (1999a). Environmental tobacco smoke, genetic susceptibility, and risk of lung cancer in never-smoking women. *J Natl Cancer Inst*, **91**, 2009-14.

- Bennett, W.P., Colby, T.V., Travis, W.D., Borkowski, A., Jones, R.T., Lane, D.P., Metcalf, R.A., Samet, J.M., Takeshima, Y., Gu, J.R. & et al. (1993). p53 protein accumulates frequently in early bronchial neoplasia. *Cancer Res*, **53**, 4817-22.
- Bennett, W.P., Hollstein, M.C., Hsu, I.C., Sidransky, D., Lane, D.P., Vogelstein, B. & Harris, C.C. (1992). Mutational spectra and immunohistochemical analyses of p53 in human cancers. *Chest*, **101**, 19S-20S.
- Bennett, W.P., Hussain, S.P., Vahakangas, K.H., Khan, M.A., Shields, P.G. & Harris, C.C. (1999b). Molecular epidemiology of human cancer risk: gene-environment interactions and p53 mutation spectrum in human lung cancer. *J Pathol*, **187**, 8-18.
- Bentley, K.L., Ferguson-Smith, A.C. & Ruddle, F.H. (1988). A review of genomic physical mapping. *Cancer Surv*, **7**, 267-94.
- Betticher, D.C., Heighway, J., Thatcher, N. & Hasleton, P.S. (1997a). Abnormal expression of CCND1 and RB1 in resection margin epithelia of lung cancer patients. *Br J Cancer*, **75**, 1761-8.
- Betticher, D.C., White, G.R., Vonlanthen, S., Liu, X., Kappeler, A., Altermatt, H.J., Thatcher, N. & Heighway, J. (1997b). G1 control gene status is frequently altered in resectable non-small cell lung cancer. *Int J Cancer*, **74**, 556-62.
- Birnboim, H.C. & Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res*, **7**, 1513-23.
- Blaustein, M.P. & Lederer, W.J. (1999). Sodium/calcium exchange: its physiological implications. *Physiol Rev*, **79**, 763-854.

- Boers, J.E., ten Velde, G.P. & Thunnissen, F.B. (1996). P53 in squamous metaplasia: a marker for risk of respiratory tract carcinoma. *Am J Respir Crit Care Med*, **153**, 411-6.
- Bolger, G., Michaeli, T., Martins, T., St John, T., Steiner, B., Rodgers, L., Riggs, M., Wigler, M. & Ferguson, K. (1993). A family of human phosphodiesterases homologous to the dunce learning and memory gene product of *Drosophila melanogaster* are potential targets for antidepressant drugs. *Mol Cell Biol*, **13**, 6558-71.
- Bouck, J.B., Metzker, M.L. & Gibbs, R.A. (2000). Shotgun sample sequence comparisons between mouse and human genomes. *Nat Genet*, **25**, 31-3.
- Boynton, R.F., Blount, P.L., Yin, J., Brown, V.L., Huang, Y., Tong, Y., McDaniel, T., Newkirk, C., Resau, J.H., Raskind, W.H. & et al. (1992). Loss of heterozygosity involving the APC and MCC genetic loci occurs in the majority of human esophageal cancers. *Proc Natl Acad Sci U S A*, **89**, 3385-8.
- Brambilla, E., Gazzeri, S., Moro, D., Lantuejoul, S., Veyrenc, S. & Brambilla, C. (1999a). Alterations of Rb pathway (Rb-p16INK4-cyclin D1) in preinvasive bronchial lesions. *Clin Cancer Res*, **5**, 243-50.
- Brambilla, E., Moro, D., Gazzeri, S. & Brambilla, C. (1999b). Alterations of expression of Rb, p16(INK4A) and cyclin D1 in non-small cell lung carcinoma and their clinical significance [see comments]. *J Pathol*, **188**, 351-60.
- Brash, D.E., Rudolph, J.A., Simon, J.A., Lin, A., McKenna, G.J., Baden, H.P., Halperin, A.J. & Ponten, J. (1991). A role for sunlight in skin cancer: UV-induced p53 mutations in squamous cell carcinoma. *Proc Natl Acad Sci U S A*, **88**, 10124-8.

- Bromen, K., Pohlabeln, H., Jahn, I., Ahrens, W. & Jockel, K.H. (2000). Aggregation of lung cancer in families: results from a population-based case-control study in Germany. *Am J Epidemiol*, **152**, 497-505.
- Brown, L.M., Silverman, D.T., Pottern, L.M., Schoenberg, J.B., Greenberg, R.S., Swanson, G.M., Liff, J.M., Schwartz, A.G., Hayes, R.B., Blot, W.J. & et al. (1994). Adenocarcinoma of the esophagus and esophagogastric junction in white men in the United States: alcohol, tobacco, and socioeconomic factors. *Cancer Causes Control*, **5**, 333-40.
- Burton, M.P., Schneider, B.G., Brown, R., Escamilla-Ponce, N. & Gulley, M.L. (1998). Comparison of histologic stains for use in PCR analysis of microdissected, paraffin-embedded tissues. *Biotechniques*, **24**, 86-92.
- Cairns, P., Tokino, K., Eby, Y. & Sidransky, D. (1995). Localization of tumor suppressor loci on chromosome 9 in primary human renal cell carcinomas. *Cancer Res*, **55**, 224-7.
- Califano, J., van der Riet, P., Westra, W., Nawroz, H., Clayman, G., Piantadosi, S., Corio, R., Lee, D., Greenberg, B., Koch, W. & Sidransky, D. (1996). Genetic progression model for head and neck cancer: implications for field cancerization. *Cancer Res*, **56**, 2488-92.
- Califano, J., Westra, W.H., Koch, W., Meininger, G., Reed, A., Yip, L., Boyle, J.O., Lonardo, F. & Sidransky, D. (1999). Unknown primary head and neck squamous cell carcinoma: molecular identification of the site of origin. *J Natl Cancer Inst*, **91**, 599-604.
- Caputi, M., De Luca, L., Papaccio, G., D'Aponte, A., Cavallotti, I., Scala, P., Scarano, F., Manna, M., Gualdiero, L. & De Luca, B. (1997). Prognostic role of cyclin D1 in non small cell lung cancer: an immunohistochemical analysis. *Eur J Histochem*, **41**, 133-8.

- Cardenas, V.M., Thun, M.J., Austin, H., Lally, C.A., Clark, W.S., Greenberg, R.S. & Heath, C.W. (1997). Environmental tobacco smoke and lung cancer mortality in the American Cancer Society's Cancer Prevention Study. II. *Cancer Causes Control*, **8**, 57-64.
- Carey, T.E. (1996). Field cancerization: are multiple primary cancers monoclonal or polyclonal? [editorial]. *Ann Med*, **28**, 183-8.
- Cavenee, W.K., Dryja, T.P., Phillips, R.A., Benedict, W.F., Godbout, R., Gallie, B.L., Murphree, A.L., Strong, L.C. & White, R.L. (1983). Expression of recessive alleles by chromosomal mechanisms in retinoblastoma. *Nature*, **305**, 779-84.
- Cense, H.A., van Lanschot, J.J., Fockens, P., Obertop, H. & Offerhaus, G.J. (1997). A patient with seven primary tumors of the upper aerodigestive tract: the process of field cancerization versus distant monoclonal expansion [see comments]. *Dis Esophagus*, **10**, 139-42.
- Chen, L.C., Matsumura, K., Deng, G., Kurisu, W., Ljung, B.M., Lerman, M.I., Waldman, F.M. & Smith, H.S. (1994). Deletion of two separate regions on chromosome 3p in breast cancers. *Cancer Research*, **54**, 3021-3024.
- Chen, X.Q., Stroun, M., Magnenat, J.L., Nicod, L.P., Kurt, A.M., Lyautey, J., Lederrey, C. & Anker, P. (1996). Microsatellite alterations in plasma DNA of small cell lung cancer patients [see comments]. *Nat Med*, **2**, 1033-5.
- Cheng, L., Song, S.Y., Pretlow, T.G., Abdul-Karim, F.W., Kung, H.J., Dawson, D.V., Park, W.S., Moon, Y.W., Tsai, M.L., Linehan, W.M., Emmert-Buck, M.R., Liotta, L.A. & Zhuang, Z. (1998). Evidence of independent origin of multiple tumors from patients with prostate cancer. *J Natl Cancer Inst*, **90**, 233-7.

- Cheng, L., Spitz, M.R., Hong, W.K. & Wei, Q. (2000). Reduced expression levels of nucleotide excision repair genes in lung cancer: a case-control analysis. *Carcinogenesis*, **21**, 1527-30.
- Cheung, V.G. & Nelson, S.F. (1996). Whole genome amplification using a degenerate oligonucleotide primer allows hundreds of genotypes to be performed on less than one nanogram of genomic DNA. *Proc Natl Acad Sci U S A*, **93**, 14676-9.
- Chiba, I., Takahashi, T., Nau, M.M., D'Amico, D., Curiel, D.T., Mitsudomi, T., Buchhagen, D.L., Carbone, D., Piantadosi, S., Koga, H. & et al. (1990). Mutations in the p53 gene are frequent in primary, resected non-small cell lung cancer. Lung Cancer Study Group. *Oncogene*, **5**, 1603-10.
- Choi, J.H., Chung, H.C., Yoo, N.C., Lee, H.R., Lee, K.H., Choi, W., Lim, H.Y., Koh, E.H., Kim, J.H., Roh, J.K. & et al. (1994). Changing trends in histologic types of lung cancer during the last decade (1981-1990) in Korea: a hospital-based study. *Lung Cancer*, **10**, 287-96.
- Chumakov, I., Rigault, P., Guillou, S., Ougen, P., Billaut, A., Guasconi, G., Gervy, P., LeGall, I., Soularue, P., Grinas, L. & et al. (1992). Continuum of overlapping clones spanning the entire human chromosome 21q [see comments]. *Nature*, **359**, 380-7.
- Chung, D.C. & Rustgi, A.K. (1995). DNA mismatch repair and cancer. *Gastroenterology*, **109**, 1685-99.
- Chung, G.T., Sundaresan, V., Hasleton, P., Rudd, R., Taylor, R. & Rabbitts, P.H. (1995). Sequential molecular genetic changes in lung cancer development. *Oncogene*, **11**, 2591-8.
- Chung, G.T., Sundaresan, V., Hasleton, P., Rudd, R., Taylor, R. & Rabbitts, P.H. (1996). Clonal evolution of lung tumors. *Cancer Res*, **56**, 1609-14.

- Cohen, A.J., Li, F.P., Berg, S., Marchetto, D.J., Tsai, S., Jacobs, S.C. & Brown, R.S. (1979). Hereditary renal-cell carcinoma associated with a chromosomal translocation. *N Engl J Med*, **301**, 592-5.
- Coleman, M.P. (1999). Cancer Survival Trends In England And Wales 1971-1995. *Office for national statistics, Series 61*, Chapter 22.
- Cooper, C.A., Bubb, V.J., Smithson, N., Carter, R.L., Gledhill, S., Lamb, D., Wyllie, A.H. & Carey, F.A. (1996). Loss of heterozygosity at 5q21 in non-small cell lung cancer: a frequent event but without evidence of apc mutation. *J Pathol*, **180**, 33-7.
- Cooper, J.A. (1990). Oncogenes and anti-oncogenes. *Curr Opin Cell Biol*, **2**, 285-95.
- Corral, J., Lavenir, I., Impey, H., Warren, A.J., Forster, A., Larson, T.A., Bell, S., McKenzie, A.N., King, G. & Rabbitts, T.H. (1996). An Mll-AF9 fusion gene made by homologous recombination causes acute leukemia in chimeric mice: a method to create fusion oncogenes. *Cell*, **85**, 853-61.
- Craig, J.M. & Bickmore, W.A. (1993). Chromosome bands--flavours to savour. *Bioessays*, **15**, 349-54.
- Crnogorac-Jurcevic, T., Brown, J.R., Lehrach, H. & Schalkwyk, L.C. (1997). *Tetraodon fluviatilis*, a new puffer fish model for genome studies. *Genomics*, **41**, 177-84.
- Cross, S.H. & Bird, A.P. (1995). CpG islands and genes. *Curr Opin Genet Dev*, **5**, 309-14.

- Crowell, R.E., Gilliland, F.D., Temes, R.T., Harms, H.J., Neft, R.E., Heaphy, E., Auckley, D.H., Crooks, L.A., Jordan, S.W., Samet, J.M., Lechner, J.F. & Belinsky, S.A. (1996). Detection of trisomy 7 in nonmalignant bronchial epithelium from lung cancer patients and individuals at risk for lung cancer. *Cancer Epidemiol Biomarkers Prev*, **5**, 631-7.
- Curran, S., McKay, J.A., McLeod, H.L. & Murray, G.I. (2000). Laser capture microscopy. *Mol Pathol*, **53**, 64-8.
- D'Amico, D., Carbone, D., Mitsudomi, T., Nau, M., Fedorko, J., Russell, E., Johnson, B., Buchhagen, D., Bodner, S., Phelps, R. & et al. (1992). High frequency of somatically acquired p53 mutations in small-cell lung cancer cell lines and tumors. *Oncogene*, **7**, 339-46.
- de Vos, S., Miller, C.W., Takeuchi, S., Gombart, A.F., Cho, S.K. & Koeffler, H.P. (1995). Alterations of CDKN2 (p16) in non-small cell lung cancer. *Genes Chromosomes Cancer*, **14**, 164-70.
- Devlin, J., Keen, A.J. & Knowles, M.A. (1994). Homozygous deletion mapping at 9p21 in bladder carcinoma defines a critical region within 2cM of IFNA. *Oncogene*, **9**, 2757-60.
- Dietmaier, W., Hartmann, A., Wallinger, S., Heinmoller, E., Kerner, T., Endl, E., Jauch, K.W., Hofstadter, F. & Ruschoff, J. (1999). Multiple mutation analyses in single tumor cells with improved whole genome amplification. *Am J Pathol*, **154**, 83-95.
- Djordjevic, M.V., Hoffmann, D. & Hoffmann, I. (1997). Nicotine regulates smoking patterns. *Prev Med*, **26**, 435-40.

- Dobson, C.L., Warren, A.J., Pannell, R., Forster, A., Lavenir, I., Corral, J., Smith, A.J. & Rabbitts, T.H. (1999). The mll-AF9 gene fusion in mice controls myeloproliferation and specifies acute myeloid leukaemogenesis. *Embo J*, **18**, 3564-74.
- Dolan, K., Garde, J., Gosney, J., Sissons, M., Wright, T., Kingsnorth, A.N., Walker, S.J., Sutton, R., Meltzer, S.J. & Field, J.K. (1998). Allelotype analysis of oesophageal adenocarcinoma: loss of heterozygosity occurs at multiple sites. *Br J Cancer*, **78**, 950-7.
- Doll, R. & Hill, A.B. (1950). Smoking and carcinoma of the lung: a preliminary report. *BMJ*, 739-748.
- Dominguez, J.E., Buendia, B., Lopez-Otin, C., Antony, C., Karsenti, E. & Avila, J. (1994). A protein related to brain microtubule-associated protein MAP1B is a component of the mammalian centrosome. *J Cell Sci*, **107**, 601-11.
- Dosaka-Akita, H., Shindoh, M., Fujino, M., Kinoshita, I., Akie, K., Katoh, M. & Kawakami, Y. (1994). Abnormal p53 expression in human lung cancer is associated with histologic subtypes and patient smoking history. *Am J Clin Pathol*, **102**, 660-4.
- Druck, T., Kastury, K., Hadaczek, P., Podolski, J., Toloczko, A., Sikorski, A., Ohta, M., LaForgia, S., Lasota, J., McCue, P. & et al. (1995). Loss of heterozygosity at the familial RCC t(3;8) locus in most clear cell renal carcinomas. *Cancer Res*, **55**, 5348-53.
- Duhaylongsod, F.G. & Wolfe, W.G. (1991). Barrett's esophagus and adenocarcinoma of the esophagus and gastroesophageal junction. *J Thorac Cardiovasc Surg*, **102**, 36-41; discussion 41-2.

- Dunn, J., Garde, J., Dolan, K., Gosney, J.R., Sutton, R., Meltzer, S.J. & Field, J.K. (1999). Multiple target sites of allelic imbalance on chromosome 17 in Barrett's oesophageal cancer. *Oncogene*, **18**, 987-93.
- Dunn, J.R., Garde, J., Dolan, K., Gosney, J.R., Oates, B.C., Watson, A.J., Fielding, P. & Field, J.K. (2000a). The evolution of loss of heterozygosity on chromosome 17 during the progression to barrett's adenocarcinoma involves a unique combination of target sites in individual specimens. *Clin Cancer Res*, **6**, 4033-42.
- Dunn, J.R., Garde, J., Dolan, K., Gosney, J.R., Oates, B.C., Watson, A.J., Fielding, P. & Field, J.K. (2000b). The evolution of loss of heterozygosity on chromosome 17 during the progression to barrett's adenocarcinoma involves a unique combination of target sites in individual specimens [In Process Citation]. *Clin Cancer Res*, **6**, 4033-42.
- Eccles, D.M., Cranston, G., Steel, C.M., Nakamura, Y. & Leonard, R.C. (1990). Allele losses on chromosome 17 in human epithelial ovarian carcinoma. *Oncogene*, **5**, 1599-601.
- Egeli, U., Karadag, M., Tunca, B. & Ozyardimci, N. (1997). The expression of common fragile sites and genetic predisposition to squamous cell lung cancers. *Cancer Genet Cytogenet*, **95**, 153-8.
- Eguchi, K., Kanai, Y., Kobayashi, K. & Hirohashi, S. (1997). DNA hypermethylation at the D17S5 locus in non-small cell lung cancers: its association with smoking history. *Cancer Res*, **57**, 4913-5.
- Ehlen, T. & Dubeau, L. (1990). Loss of heterozygosity on chromosomal segments 3p, 6q and 11p in human ovarian carcinomas. *Oncogene*, **5**, 219-23.

- el-Deiry, W.S. (1998). Regulation of p53 downstream genes. *Semin Cancer Biol*, **8**, 345-57.
- el-Deiry, W.S., Harper, J.W., O'Connor, P.M., Velculescu, V.E., Canman, C.E., Jackman, J., Pietsenpol, J.A., Burrell, M., Hill, D.E., Wang, Y. & et al. (1994). WAF1/CIP1 is induced in p53-mediated G1 arrest and apoptosis. *Cancer Res*, **54**, 1169-74.
- Ellis, J.R. & Gleeson, F.V. (2001). Lung cancer screening. *Br J Radiol*, **74**, 478-85.
- el-Zein, R., Conforti-Froes, N. & Au, W.W. (1997). Interactions between genetic predisposition and environmental toxicants for development of lung cancer. *Environ Mol Mutagen*, **30**, 196-204.
- Emmert-Buck, M.R., Bonner, R.F., Smith, P.D., Chuaqui, R.F., Zhuang, Z., Goldstein, S.R., Weiss, R.A. & Liotta, L.A. (1996). Laser capture microdissection. *Science*, **274**, 998-1001.
- Engholm, G., Palmgren, F. & Lynge, E. (1996). Lung cancer, smoking, and environment: a cohort study of the Danish population. *BMJ*, **312**, 1259-988.
- Esteller, M., Sanchez-Cespedes, M., Rosell, R., Sidransky, D., Baylin, S.B. & Herman, J.G. (1999). Detection of aberrant promoter hypermethylation of tumor suppressor genes in serum DNA from non-small cell lung cancer patients. *Cancer Res*, **59**, 67-70.
- Fan, E., Levin, D.B., Glickman, B.W. & Logan, D.M. (1993). Limitations in the use of SSCP analysis. *Mutat Res*, **288**, 85-92.

- Farr, S.B., Arnosti, D.N., Chamberlin, M.J. & Ames, B.N. (1989). An apaH mutation causes AppppA to accumulate and affects motility and catabolite repression in Escherichia coli. *Proc Natl Acad Sci U S A*, **86**, 5010-4.
- Fearon, E.R. & Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. *Cell*, **61**, 759-67.
- Ferguson-Smith, M.A. (1997). Genetic analysis by chromosome sorting and painting: phylogenetic and diagnostic applications. *Eur J Hum Genet*, **5**, 253-65.
- Field, J.K., Kiaris, H., Risk, J.M., Tsiriyotis, C., Adamson, R., Zoumpourlis, V., Rowley, H., Taylor, K., Whittaker, J., Howard, P. & et al. (1995). Allelotype of squamous cell carcinoma of the head and neck: fractional allele loss correlates with survival. *Br J Cancer*, **72**, 1180-8.
- Field, J.K., Liloglou, T., Xinarianos, G., Prime, W., Fielding, P., Walshaw, M.J. & Turnbull, L. (1999). Genetic alterations in bronchial lavage as a potential marker for individuals with a high risk of developing lung cancer. *Cancer Res*, **59**, 2690-5.
- Field, J.K., Neville, E.M., Stewart, M.P., Swift, A., Liloglou, T., Risk, J.M., Ross, H., Gosney, J.R. & Donnelly, R.J. (1996). Fractional allele loss data indicate distinct genetic populations in the development of non-small-cell lung cancer. *Br J Cancer*, **74**, 1968-74.
- Fink, J.K., Jones, S.M., Esposito, C. & Wilkowski, J. (1996). Human microtubule-associated protein 1a (MAP1A) gene: genomic organization, cDNA sequence, and developmental- and tissue-specific expression. *Genomics*, **35**, 577-85.

- Fong, K.M., Biesterveld, E.J., Virmani, A., Wistuba, I., Sekido, Y., Bader, S.A., Ahmadian, M., Ong, S.T., Rassool, F.V., Zimmerman, P.V., Giaccone, G., Gazdar, A.F. & Minna, J.D. (1997). FHIT and FRA3B 3p14.2 allele loss are common in lung cancer and preneoplastic bronchial lesions and are associated with cancer-related FHIT cDNA splicing aberrations. *Cancer Res*, **57**, 2256-67.
- Fong, K.M., Kida, Y., Zimmerman, P.V., Ikenaga, M. & Smith, P.J. (1995a). Loss of heterozygosity frequently affects chromosome 17q in non-small cell lung cancer. *Cancer Res*, **55**, 4268-72.
- Fong, K.M., Zimmerman, P.V. & Smith, P.J. (1995b). Microsatellite instability and other molecular abnormalities in non- small cell lung cancer. *Cancer Res*, **55**, 28-30.
- Fong, K.M., Zimmerman, P.V. & Smith, P.J. (1995c). Tumor progression and loss of heterozygosity at 5q and 18q in non-small cell lung cancer. *Cancer Res*, **55**, 220-3.
- Fontanini, G., Bigini, D., Vignati, S., Macchiarini, P., Pepe, S., Angeletti, C.A., Pingitore, R. & Squartini, F. (1993). p53 expression in non small cell lung cancer: clinical and biological correlations. *Anticancer Res*, **13**, 737-42.
- Fountain, J.W., Karayiorgou, M., Ernstoff, M.S., Kirkwood, J.M., Vlock, D.R., Titus-Ernstoff, L., Bouchard, B., Vijayasaradhi, S., Houghton, A.N., Lahti, J. & et al. (1992). Homozygous deletions within human chromosome band 9p21 in melanoma. *Proc Natl Acad Sci U S A*, **89**, 10557-61.
- Frebourg, T., Barbier, N., Yan, Y.X., Garber, J.E., Dreyfus, M., Fraumeni, J., Jr., Li, F.P. & Friend, S.H. (1995). Germ-line p53 mutations in 15 families with Li-Fraumeni syndrome. *Am J Hum Genet*, **56**, 608-15.

- Friedberg, E.C., Bond, J.P., Burns, D.K., Cheo, D.L., Greenblatt, M.S., Meira, L.B., Nahari, D. & Reis, A.M. (2000). Defective nucleotide excision repair in xpc mutant mice and its association with cancer predisposition. *Mutat Res*, **459**, 99-108.
- Froudarakis, M.E., Sourvinos, G., Fournel, P., Bouros, D., Vergnon, J.M., Spandidos, D.A. & Siafakas, N.M. (1998). Microsatellite instability and loss of heterozygosity at chromosomes 9 and 17 in non-small cell lung cancer. *Chest*, **113**, 1091-4.
- Fujimori, M., Tokino, T., Hino, O., Kitagawa, T., Imamura, T., Okamoto, E., Mitsunobu, M., Ishikawa, T., Nakagama, H., Harada, H. & et al. (1991). Allelotype study of primary hepatocellular carcinoma. *Cancer Res*, **51**, 89-93.
- Fullwood, P., Marchini, S., Rader, J.S., Martinez, A., Macartney, D., Broggin, M., Morelli, C., Barbanti-Brodano, G., Maher, E.R. & Latif, F. (1999). Detailed genetic and physical mapping of tumor suppressor loci on chromosome 3p in ovarian cancer. *Cancer Res*, **59**, 4662-7.
- Galipeau, P.C., Prevo, L.J., Sanchez, C.A., Longton, G.M. & Reid, B.J. (1999). Clonal expansion and loss of heterozygosity at chromosomes 9p and 17p in premalignant esophageal (Barrett's) tissue. *J Natl Cancer Inst*, **91**, 2087-95.
- Garcia, S.B., Park, H.S., Novelli, M. & Wright, N.A. (1999). Field cancerization, clonality, and epithelial stem cells: the spread of mutated clones in epithelial sheets. *J Pathol*, **187**, 61-81.

- Garcia-Perez, J., Avila, J. & Diaz-Nido, J. (1998). Implication of cyclin-dependent kinases and glycogen synthase kinase 3 in the phosphorylation of microtubule-associated protein 1B in developing neuronal cells. *J Neurosci Res*, **52**, 445-52.
- Gazdar, A.F. (1994). The molecular and cellular basis of human lung cancer. *Anticancer Res*, **14**, 261-7.
- Gazzeri, S., Gouyer, V., Vour'ch, C., Brambilla, C. & Brambilla, E. (1998). Mechanisms of p16INK4A inactivation in non small-cell lung cancers. *Oncogene*, **16**, 497-504.
- Gilley, J. & Fried, M. (1999). Extensive gene order differences within regions of conserved synteny between the Fugu and human genomes: implications for chromosomal evolution and the cloning of disease genes. *Hum Mol Genet*, **8**, 1313-20.
- GINOT, F., BORDELAIS, I., NGUYEN, S. & GYAPAY, G. (1996). Correction of some genotyping errors in automated fluorescent microsatellite analysis by enzymatic removal of one base overhangs. *Nucleic Acids Res*, **24**, 540-1.
- Girard, L., Zochbauer-Muller, S., Virmani, A.K., Gazdar, A.F. & Minna, J.D. (2000). Genome-wide allelotyping of lung cancer identifies new regions of allelic loss, differences between small cell lung cancer and non-small cell lung cancer, and loci clustering. *Cancer Res*, **60**, 4894-906.
- Glover, T.W., Coyle-Morris, J.F., Li, F.P., Brown, R.S., Berger, C.S., Gemmill, R.M. & Hecht, F. (1988). Translocation t(3;8)(p14.2;q24.1) in renal cell carcinoma affects expression of the common fragile site at 3p14(FRA3B) in lymphocytes. *Am J Hum Genet*, **43**, 265-73.

- Gnarra, J.R., Duan, D.R., Weng, Y., Humphrey, J.S., Chen, D.Y., Lee, S., Pause, A., Dudley, C.F., Latif, F., Kuzmin, I., Schmidt, L., Duh, F.M., Stackhouse, T., Chen, F., Kishida, T., Wei, M.H., Lerman, M.I., Zbar, B., Klausner, R.D. & Linehan, W.M. (1996). Molecular cloning of the von Hippel-Lindau tumor suppressor gene and its role in renal carcinoma. *Biochim Biophys Acta*, **1242**, 201-10.
- Going, J.J. & Lamb, R.F. (1996). Practical histological microdissection for PCR analysis. *J Pathol*, **179**, 121-4.
- Gorgoulis, V.G., Zacharatos, P., Kotsinas, A., Liloglou, T., Kyrroudi, A., Veslemes, M., Rassidakis, A., Halazonetis, T.D., Field, J.K. & Kittas, C. (1998). Alterations of the p16-pRb pathway and the chromosome locus 9p21-22 in non-small-cell lung carcinomas: relationship with p53 and MDM2 protein expression. *Am J Pathol*, **153**, 1749-65.
- Gray, M.R., Donnelly, R.J. & Kingsnorth, A.N. (1993). The role of smoking and alcohol in metaplasia and cancer risk in Barrett's columnar lined oesophagus. *Gut*, **34**, 727-31.
- Green, D.K. (1990). Analysing and sorting human chromosomes. *J Microsc*, **159**, 237-44.
- Greenblatt, M.S., Bennett, W.P., Hollstein, M. & Harris, C.C. (1994). Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. *Cancer Res*, **54**, 4855-78.
- Guerardel, C., Deltour, S., Pinte, S., Monte, D., Begue, A., Godwin, A.K. & Leprince, D. (2000). Identification in the human candidate tumor suppressor gene HIC-1 of a new major alternative TATA-less promoter positively regulated by p53. *J Biol Chem*, **9**, 9.

- Guerini, D. (1998). The Ca²⁺ pumps and the Na⁺/Ca²⁺ exchangers. *Biometals*, **11**, 319-30.
- Haggitt, R.C., Tryzelaar, J., Ellis, F.H. & Colcher, H. (1978). Adenocarcinoma complicating columnar epithelium-lined (Barrett's) esophagus. *Am J Clin Pathol*, **70**, 1-5.
- Hamada, K., Kohno, T., Kawanishi, M., Ohwada, S. & Yokota, J. (1998). Association of CDKN2A(p16)/CDKN2B(p15) alterations and homozygous chromosome arm 9p deletions in human lung carcinoma. *Genes Chromosomes Cancer*, **22**, 232-40.
- Hamada, K., Kohno, T., Takahashi, M., Yamazaki, M., Tashiro, H., Sugawara, C., Ohwada, S., Sekido, Y., Minna, J.D. & Yokota, J. (2000). Two regions of homozygous deletion clusters at chromosome band 9p21 in human lung cancer. *Genes Chromosomes Cancer*, **27**, 308-18.
- Hamilton, S.R. & Smith, R.R. (1987). The relationship between columnar epithelial dysplasia and invasive adenocarcinoma arising in Barrett's esophagus. *Am J Clin Pathol*, **87**, 301-12.
- Harris, C.C. (1993). p53: at the crossroads of molecular carcinogenesis and risk assessment. *Science*, **262**, 1980-1.
- Harris, H., Miller, O.J., Klein, G., Worst, P. & Tachibana, T. (1969). Suppression of malignancy by cell fusion. *Nature*, **223**, 363-8.
- Haugen, A., Ryberg, D., Mollerup, S., Zienolddiny, S., Skaug, V. & Svendsrud, D.H. (2000). Gene-environment interactions in human lung cancer. *Toxicol Lett*, **112-113**, 233-7.
- Haupt, Y., Maya, R., Kazaz, A. & Oren, M. (1997). Mdm2 promotes the rapid degradation of p53. *Nature*, **387**, 296-9.

- Hayashi, K. (1991). PCR-SSCP: a simple and sensitive method for detection of mutations in the genomic DNA. *PCR Methods Appl*, **1**, 34-8.
- Hayashi, N., Sugimoto, Y., Tsuchiya, E., Ogawa, M. & Nakamura, Y. (1994). Somatic mutations of the MTS (multiple tumor suppressor) 1/CDK4I (cyclin-dependent kinase-4 inhibitor) gene in human primary non-small cell lung carcinomas. *Biochem Biophys Res Commun*, **202**, 1426-30.
- Hibi, K., Takahashi, T., Yamakawa, K., Ueda, R., Sekido, Y., Ariyoshi, Y., Suyama, M., Takagi, H. & Nakamura, Y. (1992). Three distinct regions involved in 3p deletion in human lung cancer. *Oncogene*, **7**, 445-9.
- Highsmith, W.E., Nataraj, A.J., Jin, Q., O'Connor, J.M., El-Nabi, S.H., Kusukawa, N. & Garner, M.M. (1999). Use of DNA toolbox for the characterization of mutation scanning methods. II: evaluation of single-strand conformation polymorphism analysis. *Electrophoresis*, **20**, 1195-203.
- Hirano, T., Franzen, B., Kato, H., Ebihara, Y. & Auer, G. (1994). Genesis of squamous cell lung carcinoma. Sequential changes of proliferation, DNA ploidy, and p53 expression. *Am J Pathol*, **144**, 296-302.
- Hiroshima, K., Toyozaki, T., Kohno, H., Ohwada, H. & Fujisawa, T. (1998). Synchronous and metachronous lung carcinomas: molecular evidence for multicentricity. *Pathol Int*, **48**, 869-76.
- Hirsch, F.R., Franklin, W.A., Gazdar, A.F. & Bunn, P.A. (2001). Early detection of lung cancer: clinical perspectives of recent advances in biology and radiology. *Clin Cancer Res*, **7**, 5-22.
- Hittelman, W.N. (1999). Clones and subclones in the lung cancer field (Editorial). *J. Nat. Can. Inst.*, **91**, 1796-1799.

- Hittelman, W.N., Voravud, N., Shin, D.M., Lee, J.S., Ro, J.Y. & Hong, W.K. (1993). Early genetic changes during upper aerodigestive tract tumorigenesis. *J Cell Biochem Suppl*, **17F**, 233-6.
- Hoffmann, D. & Djordjevic, M.V. (1997). Chemical composition and carcinogenicity of smokeless tobacco. *Adv Dent Res*, **11**, 322-9.
- Holland, E.A., Beaton, S.C., Edwards, B.G., Kefford, R.F. & Mann, G.J. (1994). Loss of heterozygosity and homozygous deletions on 9p21-22 in melanoma. *Oncogene*, **9**, 1361-5.
- Hollstein, M., Hergenhahn, M., Yang, Q., Bartsch, H., Wang, Z.Q. & Hainaut, P. (1999). New approaches to understanding p53 gene tumor mutation spectra [see comments]. *Mutat Res*, **431**, 199-209.
- Hollstein, M., Sidransky, D., Vogelstein, B. & Harris, C.C. (1991). p53 mutations in human cancers. *Science*, **253**, 49-53.
- Hollstein, M.C., Wild, C.P., Bleicher, F., Chutimataewin, S., Harris, C.C., Srivatanakul, P. & Montesano, R. (1993). p53 mutations and aflatoxin B1 exposure in hepatocellular carcinoma patients from Thailand. *Int J Cancer*, **53**, 51-5.
- Holmes, J., Jr., Clark, S. & Modrich, P. (1990). Strand-specific mismatch correction in nuclear extracts of human and *Drosophila melanogaster* cell lines. *Proc Natl Acad Sci U S A*, **87**, 5837-41.
- Hosoe, S. (1996). [Search for the tumor-suppressor gene(s) on chromosome 5q, which may play an important role for the progression of lung cancer]. *Nippon Rinsho*, **54**, 482-6.

- Hosoe, S., Shigedo, Y., Ueno, K., Tachibana, I., Osaki, T., Tanio, Y., Kawase, I., Yamakawa, K., Nakamura, Y. & Kishimoto, T. (1994a). Detailed deletion mapping of the short arm of chromosome 3 in small cell and non-small cell carcinoma of the lung. *Lung Cancer*, **10**, 297-305.
- Hosoe, S., Ueno, K., Shigedo, Y., Tachibana, I., Osaki, T., Kumagai, T., Tanio, Y., Kawase, I., Nakamura, Y. & Kishimoto, T. (1994b). A frequent deletion of chromosome 5q21 in advanced small cell and non-small cell carcinoma of the lung. *Cancer Res*, **54**, 1787-90.
- Huang, D.P., Lo, K.W., van Hasselt, C.A., Woo, J.K., Choi, P.H., Leung, S.F., Cheung, S.T., Cairns, P., Sidransky, D. & Lee, J.C. (1994). A region of homozygous deletion on chromosome 9p21-22 in primary nasopharyngeal carcinoma. *Cancer Res*, **54**, 4003-6.
- Hubbert, N.L., Sedman, S.A. & Schiller, J.T. (1992). Human papillomavirus type 16 E6 increases the degradation rate of p53 in human keratinocytes. *J Virol*, **66**, 6237-41.
- Hubert, R., Weber, J.L., Schmitt, K., Zhang, L. & Arnheim, N. (1992). A new source of polymorphic DNA markers for sperm typing: analysis of microsatellite repeats in single cells. *Am J Hum Genet*, **51**, 985-91.
- Huebner, K., Druck, T., Siprashvili, Z., Croce, C.M., Kovatich, A. & McCue, P.A. (1998). The role of deletions at the FRA3B/FHIT locus in carcinogenesis. *Recent Results Cancer Res*, **154**, 200-15.
- Huebner, K., Hadaczek, P., Siprashvili, Z., Druck, T. & Croce, C.M. (1997). The FHIT gene, a multiple tumor suppressor gene encompassing the carcinogen sensitive chromosome fragile site, FRA3B. *Biochim Biophys Acta*, **1332**, M65-70.

- Huncharek, M., Kelsey, K., Muscat, J. & Christiani, D. (1996). Parental cancer and genetic predisposition in malignant pleural mesothelioma: a case-control study. *Cancer Lett*, **102**, 205-8.
- Hung, J., Kishimoto, Y., Sugio, K., Virmani, A., McIntire, D.D., Minna, J.D. & Gazdar, A.F. (1995). Allele-specific chromosome 3p deletions occur at an early stage in the pathogenesis of lung carcinoma [published erratum appears in JAMA 1995 Jun 28;273(24):1908]. *Jama*, **273**, 558-63.
- Hussain, S.P. & Harris, C.C. (1998). Molecular epidemiology of human cancer: contribution of mutation spectra studies of tumor suppressor genes. *Cancer Res*, **58**, 4023-37.
- Hussain, S.P. & Harris, C.C. (1999). p53 mutation spectrum and load: the generation of hypotheses linking the exposure of endogenous or exogenous carcinogens to human cancer. *Mutat Res*, **428**, 23-32.
- Hussain, S.P. & Harris, C.C. (2000). Molecular epidemiology and carcinogenesis: endogenous and exogenous carcinogens. *Mutat Res*, **462**, 311-22.
- Iida, S., Seto, M., Yamamoto, K., Komatsu, H., Tojo, A., Asano, S., Kamada, N., Ariyoshi, Y., Takahashi, T. & Ueda, R. (1993). MLLT3 gene on 9p22 involved in t(9;11) leukemia encodes a serine/proline rich protein homologous to MLLT1 on 19p13. *Oncogene*, **8**, 3085-92.
- Iliopoulos, O., Kibel, A., Gray, S. & Kaelin, W.G., Jr. (1995). Tumour suppression by the human von Hippel-Lindau gene product. *Nat Med*, **1**, 822-6.
- Ish-Horowicz, D. & Burke, J.F. (1981). Rapid and efficient cosmid cloning. *Nucleic Acids Res*, **9**, 2989-98.

- Jacoby, R.F., Marshall, D.J., Kailas, S., Schlack, S., Harms, B. & Love, R. (1995). Genetic instability associated with adenoma to carcinoma progression in hereditary nonpolyposis colon cancer. *Gastroenterology*, **109**, 73-82.
- Jankowski, J.A., Wright, N.A., Meltzer, S.J., Triadafilopoulos, G., Geboes, K., Casson, A.G., Kerr, D. & Young, L.S. (1999). Molecular evolution of the metaplasia-dysplasia-adenocarcinoma sequence in the esophagus. *Am J Pathol*, **154**, 965-73.
- Joh, T., Kagami, Y., Yamamoto, K., Segawa, T., Takizawa, J., Takahashi, T., Ueda, R. & Seto, M. (1996). Identification of MLL and chimeric MLL gene products involved in 11q23 translocation and possible mechanisms of leukemogenesis by MLL truncation. *Oncogene*, **13**, 1945-53.
- Johns, M.M., Westra, W.H., Califano, J.A., Eisele, D., Koch, W.M. & Sidransky, D. (1996). Allelotype of salivary gland tumors. *Cancer Res*, **56**, 1151-4.
- Johnson, L.F., Demeester, T.R. & Haggitt, R.C. (1978). Esophageal epithelial response to gastroesophageal reflux. A quantitative study. *Am J Dig Dis*, **23**, 498-509.
- Kallioniemi, A., Kallioniemi, O.P., Sudar, D., Rutovitz, D., Gray, J.W., Waldman, F. & Pinkel, D. (1992). Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science*, **258**, 818-21.
- Kamb, A., Gruis, N.A., Weaver-Feldhaus, J., Liu, Q., Harshman, K., Tavitgian, S.V., Stockert, E., Day, R.S., Johnson, B.E. & Skolnick, M.H. (1994). A cell cycle regulator potentially involved in genesis of many tumor types. *Science*, **264**, 436-40.

- Kang, G.H., Kim, C.J., Kim, W.H., Kang, Y.K., Kim, H.O. & Kim, Y.I. (1997). Genetic evidence for the multicentric origin of synchronous multiple gastric carcinoma. *Lab Invest*, **76**, 407-17.
- Kanjilal, S., Strom, S.S., Clayman, G.L., Weber, R.S., El-Naggar, A.K., Kapur, V., Cummings, K.K., Hill, L.A., Spitz, M.R., Kripke, M.L. & Ananthaswamy, H.N. (1995). *p53* mutations in nonmelanoma skin cancer of the head and neck: Molecular evidence for field cancerization. *Cancer Res.*, **55**, 3604-3609.
- Kashiwabara, K., Oyama, T., Sano, T., Fukuda, T. & Nakajima, T. (1998). Correlation between methylation status of the *p16/CDKN2* gene and the expression of *p16* and *Rb* proteins in primary non-small cell lung cancers. *Int J Cancer*, **79**, 215-20.
- Kastan, M.B., Onyekwere, O., Sidransky, D., Vogelstein, B. & Craig, R.W. (1991). Participation of *p53* protein in the cellular response to DNA damage. *Cancer Res*, **51**, 6304-11.
- Kastan, M.B., Zhan, Q., el-Deiry, W.S., Carrier, F., Jacks, T., Walsh, W.V., Plunkett, B.S., Vogelstein, B. & Fornace, A.J., Jr. (1992). A mammalian cell cycle checkpoint pathway utilizing *p53* and *GADD45* is defective in ataxia-telangiectasia. *Cell*, **71**, 587-97.
- Kastury, K., Baffa, R., Druck, T., Ohta, M., Cotticelli, M.G., Inoue, H., Negrini, M., Rugge, M., Huang, D., Croce, C.M., Palazzo, J. & Huebner, K. (1996). Potential gastrointestinal tumor suppressor locus at the 3p14.2 *FRA3B* site identified by homozygous deletions in tumor cell lines. *Cancer Res*, **56**, 978-83.

- Kato, M.V., Shimizu, T., Ishizaki, K., Kaneko, A., Yandell, D.W., Toguchida, J. & Sasaki, M.S. (1996). Loss of heterozygosity on chromosome 17 and mutation of the p53 gene in retinoblastoma. *Cancer Lett*, **106**, 75-82.
- Keen, A.J. & Knowles, M.A. (1994). Definition of two regions of deletion on chromosome 9 in carcinoma of the bladder. *Oncogene*, **9**, 2083-8.
- Kemp, C.J., Burns, P.A., Brown, K., Nagase, H. & Balmain, A. (1994). Transgenic approaches to the analysis of *ras* and *p53* function in multistage carcinogenesis. *Cold Spring Harbor Symp. Quant. Biol.*, **59**, 427-434.
- Kersting, M., Friedl, C., Kraus, A., Behn, M., Pankow, W. & Schuermann, M. (2000). Differential frequencies of p16(INK4a) promoter hypermethylation, p53 mutation, and K-ras mutation in exfoliative material mark the development of lung cancer in symptomatic chronic smokers. *J Clin Oncol*, **18**, 3221-9.
- Kessis, T.D., Slebos, R.J., Nelson, W.G., Kastan, M.B., Plunkett, B.S., Han, S.M., Lorincz, A.T., Hedrick, L. & Cho, K.R. (1993). Human papillomavirus 16 E6 expression disrupts the p53-mediated cellular response to DNA damage. *Proc Natl Acad Sci U S A*, **90**, 3988-92.
- Kiaris, H., Spanakis, N., Ergazaki, M., Sourvinos, G. & Spandidos, D.A. (1995). Loss of heterozygosity at 9p and 17q in human laryngeal tumors. *Cancer Lett*, **97**, 129-34.
- Kim, S.K., Ro, J.Y., Kemp, B.L., Lee, J.S., Kwon, T.J., Fong, K.M., Sekido, Y., Minna, J.D., Hong, W.K. & Mao, L. (1997). Identification of three distinct tumor suppressor loci on the short arm of chromosome 9 in small cell lung cancer. *Cancer Res*, **57**, 400-3.

- Kinoshita, I., Dosaka-Akita, H., Mishina, T., Akie, K., Nishi, M., Hiroumi, H., Hommura, F. & Kawakami, Y. (1996). Altered p16INK4 and retinoblastoma protein status in non-small cell lung cancer: potential synergistic effect with altered p53 protein on proliferative activity. *Cancer Res*, **56**, 5557-62.
- Kinzler, K.W., Nilbert, M.C., Vogelstein, B., Bryan, T.M., Levy, D.B., Smith, K.J., Preisinger, A.C., Hamilton, S.R., Hedge, P., Markham, A. & et al. (1991). Identification of a gene located at chromosome 5q21 that is mutated in colorectal cancers [see comments]. *Science*, **251**, 1366-70.
- Kinzler, K.W. & Vogelstein, B. (1996). Lessons from hereditary colorectal cancer. *Cell*, **87**, 159-70.
- Kishimoto, Y., Murakami, Y., Shiraishi, M., Hayashi, K. & Sekiya, T. (1992). Aberrations of the p53 tumor suppressor gene in human non-small cell carcinomas of the lung. *Cancer Res*, **52**, 4799-804.
- Kishimoto, Y., Sugio, K., Hung, J.Y., Virmani, A.K., McIntire, D.D., Minna, J.D. & Gazdar, A.F. (1995a). Allele-specific loss in chromosome 9p loci in preneoplastic lesions accompanying non-small-cell lung cancers. *J Natl Cancer Inst*, **87**, 1224-9.
- Kishimoto, Y., Sugio, K., Hung, J.Y., Virmani, A.K., McIntire, D.D., Minna, J.D. & Gazdar, A.F. (1995b). Allele-specific loss in chromosome 9p loci in preneoplastic lesions accompanying non-small-cell lung cancers [see comments]. *J Natl Cancer Inst*, **87**, 1224-9.
- Kitaguchi, S., Takeshima, Y., Nishisaka, T. & Inai, K. (1998). Proliferative activity, p53 expression and loss of heterozygosity on 3p, 9p and 17p in atypical adenomatous hyperplasia of the lung. *Hiroshima J Med Sci*, **47**, 17-25.

- Kisielewski, A.E., Xiao, G.H., Liu, S.C., Klein-Szanto, A.J., Novara, M., Sina, J., Bleicher, K., Yeung, R.S. & Goodrow, T.L. (1998). Analysis of the FHIT gene and its product in squamous cell carcinomas of the head and neck. *Oncogene*, **17**, 83-91.
- Klein, C.A., Schmidt-Kittler, O., Schardt, J.A., Pantel, K., Speicher, M.R. & Riethmuller, G. (1999). Comparative genomic hybridization, loss of heterozygosity, and DNA sequence analysis of single cells. *Proc Natl Acad Sci U S A*, **96**, 4494-9.
- Knowles, M.A., Elder, P.A., Williamson, M., Cairns, J.P., Shaw, M.E. & Law, M.G. (1994). Allelotype of human bladder cancer. *Cancer Res*, **54**, 531-8.
- Knudson, A.G., Jr. (1971). Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A*, **68**, 820-3.
- Knudson, A.G., Jr., Meadows, A.T., Nichols, W.W. & Hill, R. (1976). Chromosomal deletion and retinoblastoma. *N Engl J Med*, **295**, 1120-3.
- Kohno, H., Hiroshima, K., Toyozaki, T., Fujisawa, T. & Ohwada, H. (1999). p53 mutation and allelic loss of chromosome 3p, 9p of preneoplastic lesions in patients with nonsmall cell lung carcinoma. *Cancer*, **85**, 341-7.
- Konishi, H., Takahashi, T., Kozaki, K., Yatabe, Y., Mitsudomi, T., Fujii, Y., Sugiura, T. & Matsuda, H. (1998). Detailed deletion mapping suggests the involvement of a tumor suppressor gene at 17p13.3, distal to p53, in the pathogenesis of lung cancers. *Oncogene*, **17**, 2095-100.
- Kouprina, N., Eldarov, M., Moyzis, R., Resnick, M. & Larionov, V. (1994). A model system to assess the integrity of mammalian YACs during transformation and propagation in yeast. *Genomics*, **21**, 7-17.

- Kubbutat, M.H., Jones, S.N. & Vousden, K.H. (1997). Regulation of p53 stability by Mdm2. *Nature*, **387**, 299-303.
- Kubbutat, M.H., Ludwig, R.L., Levine, A.J. & Vousden, K.H. (1999). Analysis of the degradation function of Mdm2. *Cell Growth Differ*, **10**, 87-92.
- Kurasono, Y., Ito, T., Kameda, Y., Nakamura, N. & Kitamura, H. (1998). Expression of cyclin D1, retinoblastoma gene protein, and p16 MTS1 protein in atypical adenomatous hyperplasia and adenocarcinoma of the lung. An immunohistochemical analysis. *Virchows Arch*, **432**, 207-15.
- Land, C.E., Shimosato, Y., Saccomanno, G., Tokuoka, S., Auerbach, O., Tateishi, R., Greenberg, S.D., Nambu, S., Carter, D., Akiba, S. & et al. (1993). Radiation-associated lung cancer: a comparison of the histology of lung cancers in uranium miners and survivors of the atomic bombings of Hiroshima and Nagasaki. *Radiat Res*, **134**, 234-43.
- Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., et al. (2001). Initial sequencing and analysis of the human genome. *Nature*, **409**, 860-921.
- Lane, D.P. (1992). Cancer. p53, guardian of the genome [news; comment] [see comments]. *Nature*, **358**, 15-6.
- Lane, D.P. & Crawford, L.V. (1979). T antigen is bound to a host protein in SV40-transformed cells. *Nature*, **278**, 261-3.
- Lane, D.P. & Crawford, L.V. (1980). The complex between simian virus 40 T antigen and a specific host protein. *Proc R Soc Lond B Biol Sci*, **210**, 451-63.
- Larionov, V., Kouprina, N., Eldarov, M., Perkins, E., Porter, G. & Resnick, M.A. (1994a). Transformation-associated recombination between diverged and homologous DNA repeats is induced by strand breaks. *Yeast*, **10**, 93-104.

- Larionov, V., Kouprina, N., Nikolaishvili, N. & Resnick, M.A. (1994b). Recombination during transformation as a source of chimeric mammalian artificial chromosomes in yeast (YACs). *Nucleic Acids Res*, **22**, 4154-62.
- Larson, A.A., Kern, S., Curtiss, S., Gordon, R., Cavenee, W.K. & Hampton, G.M. (1997). High resolution analysis of chromosome 3p alterations in cervical carcinoma. *Cancer Res*, **57**, 4082-90.
- Latif, F., Tory, K., Gnarr, J., Yao, M., Duh, F.M., Orcutt, M.L., Stackhouse, T., Kuzmin, I., Modi, W., Geil, L. & et al. (1993). Identification of the von Hippel-Lindau disease tumor suppressor gene [see comments]. *Science*, **260**, 1317-20.
- Leach, F.S., Nicolaides, N.C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomaki, P., Sistonen, P., Aaltonen, L.A., Nystrom-Lahti, M. & et al. (1993). Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. *Cell*, **75**, 1215-25.
- Leblanc, V., Delumeau, I. & Tocque, B. (1999). Ras-GTPase activating protein inhibition specifically induces apoptosis of tumour cells. *Oncogene*, **18**, 4884-9.
- Lechner, J.F., Neft, R.E., Gilliland, F.D., Crowell, R.E., Auckley, D.H., Temes, R.T. & Belinsky, S.A. (1998). Individuals at high risk for lung cancer have airway epithelial cells with chromosome aberrations frequently found in lung tumor cells. *In Vivo*, **12**, 23-6.
- Lechner, M.S., Mack, D.H., Finicle, A.B., Crook, T., Vousden, K.H. & Laimins, L.A. (1992). Human papillomavirus E6 proteins bind p53 in vivo and abrogate p53-mediated repression of transcription. *Embo J*, **11**, 3045-52.

- Leethanakul, C., Patel, V., Gillespie, J., Pallente, M., Ensley, J.F., Koontongkaew, S., Liotta, L.A., Emmert-Buck, M. & Gutkind, J.S. (2000a). Distinct pattern of expression of differentiation and growth-related genes in squamous cell carcinomas of the head and neck revealed by the use of laser capture microdissection and cDNA arrays. *Oncogene*, **19**, 3220-4.
- Leethanakul, C., Patel, V., Gillespie, J., Shillitoe, E., Kellman, R.M., Ensley, J.F., Limwongse, V., Emmert-Buck, M.R., Krizman, D.B. & Gutkind, J.S. (2000b). Gene expression profiles in squamous cell carcinomas of the oral cavity: use of laser capture microdissection for the construction and analysis of stage-specific cDNA libraries. *Oral Oncol*, **36**, 474-83.
- Levi, F., Ollyo, J.B., La Vecchia, C., Boyle, P., Monnier, P. & Savary, M. (1990). The consumption of tobacco, alcohol and the risk of adenocarcinoma in Barrett's oesophagus. *Int J Cancer*, **45**, 852-4.
- Levin, N.A., Brzoska, P.M., Warnock, M.L., Gray, J.W. & Christman, M.F. (1995). Identification of novel regions of altered DNA copy number in small cell lung tumors. *Genes Chromosomes Cancer*, **13**, 175-85.
- Levine, A.J. (1997). p53, the cellular gatekeeper for growth and division. *Cell*, **88**, 323-31.
- Liloglou, T., Maloney, P., Xinarianos, G., Fear, S. & Field, J.K. (2000). Sensitivity and limitations of high throughput fluorescent microsatellite analysis for the detection of allelic imbalance: application in lung tumors. *Int J Oncol*, **16**, 5-14.

- Liloglou, T., Maloney, P., Xinarianos, G., Hulbert, M., Walshaw, M.J., Gosney, J.R., Turnbull, L. & Field, J.K. (2001). Cancer-specific genomic instability in bronchial lavage: a molecular tool for lung cancer detection. *Cancer Res*, **61**, 1624-8.
- Liloglou, T., Ross, H., Prime, W., Donnelly, R.J., Spandidos, D.A., Gosney, J.R. & Field, J.K. (1997). p53 gene aberrations in non-small-cell lung carcinomas from a smoking population. *Br J Cancer*, **75**, 1119-24.
- Liu, B., Farrington, S.M., Petersen, G.M., Hamilton, S.R., Parsons, R., Papadopoulos, N., Fujiwara, T., Jen, J., Kinzler, K.W., Wyllie, A.H. & et al. (1995). Genetic instability occurs in the majority of young patients with colorectal cancer [see comments]. *Nat Med*, **1**, 348-52.
- Lundberg, C., Skoog, L., Cavenee, W.K. & Nordenskjold, M. (1987). Loss of heterozygosity in human ductal breast tumors indicates a recessive mutation on chromosome 13. *Proc Natl Acad Sci U S A*, **84**, 2372-6.
- Lutchman, M., Pack, S., Kim, A.C., Azim, A., Emmert-Buck, M., van Huffel, C., Zhuang, Z. & Chishti, A.H. (1999). Loss of heterozygosity on 8p in prostate cancer implicates a role for dematin in tumor progression. *Cancer Genet Cytogenet*, **115**, 65-9.
- Lydiatt, W.M., Anderson, P.E., Bazzana, T., Casale, M., Hughes, C.J., Huvos, A.G., Lydiatt, D.D. & Schantz, S.P. (1998). Molecular support for field cancerization in the head and neck. *Cancer*, **82**, 1376-80.
- Lydiatt, W.M., Murty, V.V., Davidson, B.J., Xu, L., Dyomina, K., Sacks, P.G., Schantz, S.P. & Chaganti, R.S. (1995). Homozygous deletions and loss of expression of the CDKN2 gene occur frequently in head and neck squamous cell carcinoma cell lines but infrequently in primary tumors. *Genes Chromosomes Cancer*, **13**, 94-8.

- Lynch, H.T., Smyrk, T.C., Watson, P., Lanspa, S.J., Lynch, J.F., Lynch, P.M., Cavalieri, R.J. & Boland, C.R. (1993). Genetics, natural history, tumor spectrum, and pathology of hereditary nonpolyposis colorectal cancer: an updated review. *Gastroenterology*, **104**, 1535-49.
- Mackay, B., Ordonez, N.G., Bennington, J.L. & Dugan, C.C. (1989). Ultrastructural and morphometric features of poorly differentiated and undifferentiated lung tumors. *Ultrastruct Pathol*, **13**, 561-71.
- Mackay, J., Steel, C.M., Elder, P.A., Forrest, A.P. & Evans, H.J. (1988). Allele loss on short arm of chromosome 17 in breast cancers. *Lancet*, **2**, 1384-5.
- Macoska, J.A., Powell, I.J., Sakr, W. & Lane, M.A. (1992). Loss of the 17p chromosomal region in a metastatic carcinoma of the prostate. *J Urol*, **147**, 1142-6.
- Maehara, Y., Oda, S. & Sugimachi, K. (2001). The instability within: problems in current analyses of microsatellite instability. *Mutat Res*, **461**, 249-63.
- Maestro, R., Gasparotto, D., Vukosavljevic, T., Barzan, L., Sulfaro, S. & Boiocchi, M. (1993). Three discrete regions of deletion at 3p in head and neck cancers. *Cancer Research*, **53**, 5775-5779.
- Malusecka, E., Zborek, A. & Krzyzowska-Gruca, S. (1999). Changes in expression of pRb, p16 and cyclin D1 in non-small cell lung cancer: an immunohistochemical study. *Folia Histochem Cytobiol*, **37**, 19-24.
- Mann, S.S. & Hammarback, J.A. (1996). Gene localization and developmental expression of light chain 3: a common subunit of microtubule-associated protein 1A(MAP1A) and MAP1B. *J Neurosci Res*, **43**, 535-44.

- Mao, L., Fan, Y.H., Lotan, R. & Hong, W.K. (1996). Frequent abnormalities of FHIT, a candidate tumor suppressor gene, in head and neck cancer cell lines. *Cancer Res*, **56**, 5128-31.
- Mao, L., Lee, D.J., Tockman, M.S., Erozan, Y.S., Askin, F. & Sidransky, D. (1994). Microsatellite alterations as clonal markers for the detection of human cancer. *Proc Natl Acad Sci U S A*, **91**, 9871-5.
- Mao, L., Lee, J.S., Kurie, J.M., Fan, Y.H., Lippman, S.M., Lee, J.J., Ro, J.Y., Broxson, A., Yu, R., Morice, R.C., Kemp, B.L., Khuri, F.R., Walsh, G.L., Hittelman, W.N. & Hong, W.K. (1997). Clonal genetic alterations in the lungs of current and former smokers [see comments]. *J Natl Cancer Inst*, **89**, 857-62.
- Marchetti, A., Buttitta, F., Merlo, G., Diella, F., Pellegrini, S., Pepe, S., Macchiarini, P., Chella, A., Angeletti, C.A., Callahan, R. & et al. (1993). p53 alterations in non-small cell lung cancers correlate with metastatic involvement of hilar and mediastinal lymph nodes. *Cancer Res*, **53**, 2846-51.
- Marchetti, A., Buttitta, F., Pellegrini, S., Bertacca, G., Chella, A., Carnicelli, V., Tognoni, V., Filardo, A., Angeletti, C.A. & Bevilacqua, G. (1997). Alterations of P16 (MTS1) in node-positive non-small cell lung carcinomas. *J Pathol*, **181**, 178-82.
- Mate, J.L., Ariza, A., Aracil, C., Lopez, D., Isamat, M., Perez-Piteira, J. & Navas-Palacios, J.J. (1996). Cyclin D1 overexpression in non-small cell lung carcinoma: correlation with Ki67 labelling index and poor cytoplasmic differentiation. *J Pathol*, **180**, 395-9.

- Matsui, Y., Kikuchi, A., Kondo, J., Hishida, T., Teranishi, Y. & Takai, Y. (1988). Nucleotide and deduced amino acid sequences of a GTP-binding protein family with molecular weights of 25,000 from bovine brain. *J Biol Chem*, **263**, 11071-4.
- Mayne, S.T., Buenconsejo, J. & Janerich, D.T. (1999). Familial cancer history and lung cancer risk in United States nonsmoking men and women. *Cancer Epidemiol Biomarkers Prev*, **8**, 1065-9.
- McHale, J.C., McKie, A.B., Tarttelin, E.E. & Inglehearn, C.F. (2000). Expression map of human chromosome region 17p13.3, spanning the RP13 dominant retinitis pigmentosa locus, the Miller-Dieker lissencephaly syndrome (MDLS) region, and a putative tumour suppressor locus. *Cytogenet Cell Genet*, **88**, 225-9.
- Mead, L.J., Gillespie, M.T., Hung, J.Y., Rane, U.S., Rayeroux, K.C., Irving, L.B. & Campbell, L.J. (1997). Frequent loss of heterozygosity in early non-small cell lung cancers at chromosome 9p21 proximal to the CDKN2a gene. *Int J Cancer*, **71**, 213-7.
- Mead, L.J., Gillespie, M.T., Irving, L.B. & Campbell, L.J. (1994). Homozygous and hemizygous deletions of 9p centromeric to the interferon genes in lung cancer. *Cancer Res*, **54**, 2307-9.
- Mendes-da-Silva, P., Moreira, A., Duro-da-Costa, J., Matias, D. & Monteiro, C. (2000). Frequent loss of heterozygosity on chromosome 5 in non-small cell lung carcinoma. *Mol Pathol*, **53**, 184-7.
- Merlo, A., Gabrielson, E., Askin, F. & Sidransky, D. (1994a). Frequent loss of chromosome 9 in human primary non-small cell lung cancer. *Cancer Res*, **54**, 640-2.

- Merlo, A., Gabrielson, E., Mabry, M., Vollmer, R., Baylin, S.B. & Sidransky, D. (1994b). Homozygous deletion on chromosome 9p and loss of heterozygosity on 9q, 6p, and 6q in primary human small cell lung cancer. *Cancer Res*, **54**, 2322-6.
- Mertens, F., Johansson, B., Hoglund, M. & Mitelman, F. (1997). Chromosomal imbalance maps of malignant solid tumors: a cytogenetic survey of 3185 neoplasms. *Cancer Res*, **57**, 2765-80.
- Miller, D.G. (1980). On the nature of susceptibility to cancer. The presidential address. *Cancer*, **46**, 1307-18.
- Miozzo, M., Sozzi, G., Musso, K., Pilotti, S., Incarbone, M., Pastorino, U. & Pierotti, M.A. (1996). Microsatellite alterations in bronchial and sputum specimens of lung cancer patients. *Cancer Res*, **56**, 2285-8.
- Mitelman, F., Mertens, F. & Johansson, B. (1997). A breakpoint map of recurrent chromosomal rearrangements in human neoplasia. *Nat Genet*, **15 Spec No**, 417-74.
- Mitra, A.B., Murty, V.V., Li, R.G., Pratap, M., Luthra, U.K. & Chaganti, R.S. (1994). Allelotype analysis of cervical carcinoma. *Cancer Res*, **54**, 4481-7.
- Mitsudomi, T., Oyama, T., Nishida, K., Ogami, A., Osaki, T., Sugio, K., Yasumoto, K., Sugimachi, K. & Gazdar, A.F. (1996). Loss of heterozygosity at 3p in non-small cell lung cancer and its prognostic implication. *Clin Cancer Res*, **2**, 1185-9.
- Mitterbauer, G., Zimmer, C., Fonatsch, C., Haas, O., Thalhammer-Scherrer, R., Schwarzingler, I., Kalhs, P., Jaeger, U., Lechner, K. & Mannhalter, C. (1999). Monitoring of minimal residual leukemia in patients with MLL-AF9 positive acute myeloid leukemia by RT-PCR. *Leukemia*, **13**, 1519-24.

- Miura, I., Graziano, S.L., Cheng, J.Q., Doyle, L.A. & Testa, J.R. (1992). Chromosome alterations in human small cell lung cancer: frequent involvement of 5q. *Cancer Res*, **52**, 1322-8.
- Momand, J., Zambetti, G.P., Olson, D.C., George, D. & Levine, A.J. (1992). The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation. *Cell*, **69**, 1237-45.
- Moreno, F.J., Diaz-Nido, J., Jimenez, J.S. & Avila, J. (1999). Distribution of CK2, its substrate MAP1B and phosphatases in neuronal cells. *Mol Cell Biochem*, **191**, 201-5.
- Mori, N., Yokota, J., Oshimura, M., Cavenee, W.K., Mizoguchi, H., Noguchi, M., Shimosato, Y., Sugimura, T. & Terada, M. (1989). Concordant deletions of chromosome 3p and loss of heterozygosity for chromosomes 13 and 17 in small cell lung carcinoma. *Cancer Res*, **49**, 5130-5.
- Mori, T., Aoki, T., Matsubara, T., Iida, F., Du, X., Nishihira, T., Mori, S. & Nakamura, Y. (1994). Frequent loss of heterozygosity in the region including BRCA1 on chromosome 17q in squamous cell carcinomas of the esophagus. *Cancer Res*, **54**, 1638-40.
- Morton, N. (1955). *Am. J. Hum. Genet.*, **7**, 277-318.
- Morton, N.E. (1984). Linkage and association. *Prog Clin Biol Res*, **147**, 245-65.
- Moskaluk, C.A. & Kern, S.E. (1997). Microdissection and polymerase chain reaction amplification of genomic DNA from histological tissue sections. *Am J Pathol*, **150**, 1547-52.
- Mueller, J., Werner, M. & Siewert, J.R. (2000). Malignant progression in Barrett's esophagus: pathology and molecular biology. *Recent Results Cancer Res*, **155**, 29-41.

- Murase, T., Inagaki, H. & Eimoto, T. (2000). Influence of histochemical and immunohistochemical stains on polymerase chain reaction. *Mod Pathol*, **13**, 147-51.
- Murthy, S.S. & Testa, J.R. (1999). Asbestos, chromosomal deletions, and tumor suppressor gene alterations in human malignant mesothelioma. *J Cell Physiol*, **180**, 150-7.
- Nakagawa, K., Conrad, N.K., Williams, J.P., Johnson, B.E. & Kelley, M.J. (1995). Mechanism of inactivation of CDKN2 and MTS2 in non-small cell lung cancer and association with advanced stage. *Oncogene*, **11**, 1843-51.
- Nakagawa, S., Yoshikawa, H., Kimura, M., Kawana, K., Matsumoto, K., Onda, T., Kino, N., Yamada, M., Yasugi, T. & Taketani, Y. (1999). A possible involvement of aberrant expression of the FHIT gene in the carcinogenesis of squamous cell carcinoma of the uterine cervix. *Br J Cancer*, **79**, 589-94.
- Nakamura, M., Sugita, K., Inukai, T., Goi, K., Iijima, K., Tezuka, T., Kojika, S., Shiraishi, K., Miyamoto, N., Karakida, N., Kagami, K., T, O.K., Mori, T. & Nakazawa, S. (1999). p16/MTS1/INK4A gene is frequently inactivated by hypermethylation in childhood acute lymphoblastic leukemia with 11q23 translocation. *Leukemia*, **13**, 884-90.
- Nakamura, Y., Nishisho, I., Kinzler, K.W., Vogelstein, B., Miyoshi, Y., Miki, Y., Ando, H. & Horii, A. (1992). Mutations of the APC (adenomatous polyposis coli) gene in FAP (familial polyposis coli) patients and in sporadic colorectal tumors. *Tohoku J Exp Med*, **168**, 141-7.

- Nakamura, Y., Nishisho, I., Kinzler, K.W., Vogelstein, B., Miyoshi, Y., Miki, Y., Ando, H., Horii, A. & Nagase, H. (1991). Mutations of the adenomatous polyposis coli gene in familial polyposis coli patients and sporadic colorectal tumors. *Princess Takamatsu Symp*, **22**, 285-92.
- Nawroz, H., van der Riet, P., Hruban, R.H., Koch, W., Ruppert, J.M. & Sidransky, D. (1994). Allelotype of head and neck squamous cell carcinoma. *Cancer Research*, **54**, 1152-1155.
- Negrini, M., Monaco, C., Vorechovsky, I., Ohta, M., Druck, T., Baffa, R., Huebner, K. & Croce, C.M. (1996). FHIT gene and cancer [news] The FHIT gene at 3p14.2 is abnormal in breast carcinomas. *Environ Health Perspect*, **104**, 818-9.
- Neville, E.M., Stewart, M., Myskow, M., Donnelly, R.J. & Field, J.K. (1995). Loss of heterozygosity at 9p23 defines a novel locus in non-small cell lung cancer. *Oncogene*, **11**, 581-5.
- Neville, E.M., Stewart, M.P., Swift, A., Liloglou, T., Ross, H., Gosney, J.R., Donnelly, R.J. & Field, J.K. (1996). Allelotype of non-small cell lung cancer. *Int. J. Oncology*, **9**, 533-539.
- Nigro, J.M., Baker, S.J., Preisinger, A.C., Jessup, J.M., Hostetter, R., Cleary, K., Bigner, S.H., Davidson, N., Baylin, S., Devilee, P. & et al. (1989). Mutations in the p53 gene occur in diverse human tumour types. *Nature*, **342**, 705-8.
- Nishiyama, H., Kaguraoka, H., Kuroki, M., Takahashi, K., Kitaya, T. & Matsuyama, T. (1989). [Multiple primary cancers in surgically resected lung cancer patients]. *Nippon Kyobu Geka Gakkai Zasshi*, **37**, 56-61.

- Nobori, T., Takabayashi, K., Tran, P., Orvis, L., Batova, A., Yu, A.L. & Carson, D.A. (1996). Genomic cloning of methylthioadenosine phosphorylase: a purine metabolic enzyme deficient in multiple different cancers. *Proc Natl Acad Sci U S A*, **93**, 6203-8.
- Nowell, P.C. (1976). The clonal evolution of tumor cell populations. *Science*, **194**, 23-8.
- Oda, S., Oki, E., Maehara, Y. & Sugimachi, K. (1997). Precise assessment of microsatellite instability using high resolution fluorescent microsatellite analysis. *Nucleic Acids Res*, **25**, 3415-20.
- Oesterreich, S., Allredl, D.C., Mohsin, S.K., Zhang, Q., Wong, H., Lee, A.V., Osborne, C.K. & O'Connell, P. (2001). High rates of loss of heterozygosity on chromosome 19p13 in human breast cancer. *Br J Cancer*, **84**, 493-8.
- Ohno, A. (1996). [p16/CDKN2/MTS1 gene abnormality in lung cancer]. *Nippon Rinsho*, **54**, 503-6.
- Ohta, M., Inoue, H., Cotticelli, M.G., Kastury, K., Baffa, R., Palazzo, J., Siprashvili, Z., Mori, M., McCue, P., Druck, T. & et al. (1996). The FHIT gene, spanning the chromosome 3p14.2 fragile site and renal carcinoma-associated t(3;8) breakpoint, is abnormal in digestive tract cancers. *Cell*, **84**, 587-97.
- Okami, K., Cairns, P., Westra, W.H., Linn, J.F., Ahrendt, S.A., Wu, L., Sidransky, D. & Jen, J. (1997). Detailed deletion mapping at chromosome 9p21 in non-small cell lung cancer by microsatellite analysis and fluorescence in situ hybridization. *Int J Cancer*, **74**, 588-92.

- Olopade, O.I., Bohlander, S.K., Pomykala, H., Maltepe, E., Van Melle, E., Le Beau, M.M. & Diaz, M.O. (1992a). Mapping of the shortest region of overlap of deletions of the short arm of chromosome 9 associated with human neoplasia. *Genomics*, **14**, 437-43.
- Olopade, O.I., Buchhagen, D.L., Malik, K., Sherman, J., Nobori, T., Bader, S., Nau, M.M., Gazdar, A.F., Minna, J.D. & Diaz, M.O. (1993). Homozygous loss of the interferon genes defines the critical region on 9p that is deleted in lung cancers. *Cancer Res*, **53**, 2410-5.
- Olopade, O.I., Jenkins, R.B., Ransom, D.T., Malik, K., Pomykala, H., Nobori, T., Cowan, J.M., Rowley, J.D. & Diaz, M.O. (1992b). Molecular analysis of deletions of the short arm of chromosome 9 in human gliomas. *Cancer Res*, **52**, 2523-9.
- Otterson, G.A., Khleif, S.N., Chen, W., Coxon, A.B. & Kaye, F.J. (1995). CDKN2 gene silencing in lung cancer by DNA hypermethylation and kinetics of p16INK4 protein induction by 5-aza 2'deoxyctidine. *Oncogene*, **11**, 1211-6.
- Panagopoulos, I., Pandis, N., Thelin, S., Petersson, C., Mertens, F., Borg, A., Kristoffersson, U., Mitelman, F. & Aman, P. (1996). The FHIT and PTPRG genes are deleted in benign proliferative breast disease associated with familial breast cancer and cytogenetic rearrangements of chromosome band 3p14. *Cancer Res*, **56**, 4871-5.
- Papadopoulos, N., Nicolaidis, N.C., Wei, Y.F., Ruben, S.M., Carter, K.C., Rosen, C.A., Haseltine, W.A., Fleischmann, R.D., Fraser, C.M., Adams, M.D. & al, e. (1994). Mutation of a mutL homolog in hereditary colon cancer [see comments]. *Science*, **263**, 1625-1629.

- Park, I.W., Wistuba, II, Maitra, A., Milchgrub, S., Virmani, A.K., Minna, J.D. & Gazdar, A.F. (1999). Multiple clonal abnormalities in the bronchial epithelium of patients with lung cancer [In Process Citation]. *J Natl Cancer Inst*, **91**, 1863-8.
- Park, J.Y., Jeon, H., Park, S.H., Park, T.I., Son, J., Kim, C.H., Park, J., Kim, I.S., Jung, T.H. & Jun, S.H. (2000). Microsatellite alteration in histologically normal lung tissue of patients with non-small cell lung cancer [In Process Citation]. *Lung Cancer*, **30**, 83-9.
- Parsons, N.R. & Somerville, L. (2000). Estimation and projection of population lung cancer trends (United Kingdom). *Cancer Causes Control*, **11**, 467-75.
- Passlick, B., Izbicki, J.R., Riethmuller, G. & Pantel, K. (1994). p53 in non-small-cell lung cancer [letter; comment] [see comments]. *J Natl Cancer Inst*, **86**, 801-3.
- Pateromichelakis, S., Lee, G., Langdon, J.D. & Partridge, M. (2000). The FHIT gene in oral squamous cell carcinoma: allelic imbalance is frequent but cDNA aberrations are uncommon. *Oral Oncol*, **36**, 180-188.
- Paulson, T.G., Galipeau, P.C. & Reid, B.J. (1999). Loss of heterozygosity analysis using whole genome amplification, cell sorting, and fluorescence-based PCR. *Genome Res*, **9**, 482-91.
- Pedrotti, B., Colombo, R. & Islam, K. (1994). Microtubule associated protein MAP1A is an actin-binding and crosslinking protein. *Cell Motil Cytoskeleton*, **29**, 110-6.
- Pedrotti, B. & Islam, K. (1996). Dephosphorylated but not phosphorylated microtubule associated protein MAP1B binds to microfilaments. *FEBS Lett*, **388**, 131-3.

- Peng, H.Q., Liu, L., Goss, P.E., Bailey, D. & Hogg, D. (1999). Chromosomal deletions occur in restricted regions of 5q in testicular germ cell cancer. *Oncogene*, **18**, 3277-83.
- Petersen, I., Bujard, M., Petersen, S., Wolf, G., Goeze, A., Schwendel, A., Langreck, H., Gellert, K., Reichel, M., Just, K., du Manoir, S., Cremer, T., Dietel, M. & Ried, T. (1997). Patterns of chromosomal imbalances in adenocarcinoma and squamous cell carcinoma of the lung. *Cancer Res*, **57**, 2331-5.
- Peto, R., Darby, S., Deo, H., Silcocks, P., Whitley, E. & Doll, R. (2000). Smoking, smoking cessation, and lung cancer in the UK since 1950: combination of national statistics with two case-control studies [see comments]. *Bmj*, **321**, 323-9.
- Piccinin, S., Gasparotto, D., Vukosavljevic, T., Barzan, L., Sulfaro, S., Maestro, R. & Boiocchi, M. (1998). Microsatellite instability in squamous cell carcinomas of the head and neck related to field cancerization phenomena. *Br J Cancer*, **78**, 1147-51.
- Prinsen, C.F., Szerencsei, R.T. & Schnetkamp, P.P. (2000). Molecular cloning and functional expression of the potassium-dependent sodium-calcium exchanger from human and chicken retinal cone photoreceptors. *J Neurosci*, **20**, 1424-34.
- Puig, S., Ruiz, A., Lazaro, C., Castel, T., Lynch, M., Palou, J., Vilalta, A., Weissenbach, J., Mascaro, J.M. & Estivill, X. (1995). Chromosome 9p deletions in cutaneous malignant melanoma tumors: the minimal deleted region involves markers outside the p16 (CDKN2) gene. *Am J Hum Genet*, **57**, 395-402.

- Qiu, Y.H., Chen, C.N., Malone, T., Richter, L., Beckendorf, S.K. & Davis, R.L. (1991). Characterization of the memory gene *dunce* of *Drosophila melanogaster*. *J Mol Biol*, **222**, 553-65.
- Quinn, A.G., Sikkink, S. & Rees, J.L. (1994). Basal cell carcinomas and squamous cell carcinomas of human skin show distinct patterns of chromosome loss. *Cancer Res*, **54**, 4756-9.
- Reynolds, J.C., Waronker, M., Pacquing, M.S. & Yassin, R.R. (1999). Barrett's esophagus. Reducing the risk of progression to adenocarcinoma. *Gastroenterol Clin North Am*, **28**, 917-45.
- Risse, E.K., Vooijs, G.P. & van't Hof, M.A. (1988). Diagnostic significance of "severe dysplasia" in sputum cytology. *Acta Cytol*, **32**, 629-34.
- Roest Crollius, H., Jaillon, O., Bernot, A., Dasilva, C., Bouneau, L., Fischer, C., Fizames, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W. & Weissenbach, J. (2000). Estimate of human gene number provided by genome-wide analysis using *Tetraodon nigroviridis* DNA sequence. *Nat Genet*, **25**, 235-8.
- Roth, D. & Sage, H.H. (1969). Defective photochemical repair in epithelium predisposed to field cancerization. *Cancer*, **24**, 511-9.
- Roth, J., Dobbelstein, M., Freedman, D.A., Shenk, T. & Levine, A.J. (1998). Nucleo-cytoplasmic shuttling of the hdm2 oncoprotein regulates the levels of the p53 protein via a pathway used by the human immunodeficiency virus rev protein. *Embo J*, **17**, 554-64.
- Rotter, V., Aloni-Grinstein, R., Schwartz, D., Elkind, N.B., Simons, A., Wolkowicz, R., Lavigne, M., Beserman, P., Kapon, A. & Goldfinger, N. (1994). Does wild-type p53 play a role in normal cell differentiation? *Semin Cancer Biol*, **5**, 229-36.

- Rous, P. (1911). A sarcoma of the fowl transmissible by an agent separable from the tumor. *J. Exp. Med.*, **13**, 397-411.
- Rubin, G.M. (2001). The draft sequences. Comparing species. *Nature*, **409**, 820-1.
- Rusin, M.R., Okamoto, A., Chorazy, M., Czyzewski, K., Harasim, J., Spillare, E.A., Hagiwara, K., Hussain, S.P., Xiong, Y., Demetrick, D.J. & Harris, C.C. (1996). Intragenic mutations of the p16(INK4), p15(INK4B) and p18 genes in primary non-small-cell lung cancers. *Int J Cancer*, **65**, 734-9.
- Russell, S.E., Hickey, G.I., Lowry, W.S., White, P. & Atkinson, R.J. (1990). Allele loss from chromosome 17 in ovarian cancer. *Oncogene*, **5**, 1581-3.
- Sacomanno, G., Archer, V.E., Auerbach, O., Saunders, R.P. & Brennan, L.M. (1974). Development of carcinoma of the lung as reflected in exfoliated cells. *Cancer*, **33**, 256-70.
- Sanchez-Cespedes, M., Ahrendt, S.A., Piantadosi, S., Rosell, R., Monzo, M., Wu, L., Westra, W.H., Yang, S.C., Jen, J. & Sidransky, D. (2001). Chromosomal alterations in lung adenocarcinoma from smokers and nonsmokers. *Cancer Res*, **61**, 1309-13.
- Sanchez-Cespedes, M., Cairns, P., Jen, J. & Sidransky, D. (1998). Degenerate oligonucleotide-primed PCR (DOP-PCR): evaluation of its reliability for screening of genetic alterations in neoplasia. *Biotechniques*, **25**, 1036-8.
- Sanfey, H., Hamilton, S.R., Smith, R.R. & Cameron, J.L. (1985). Carcinoma arising in Barrett's esophagus. *Surg Gynecol Obstet*, **161**, 570-4.
- Sano, T., Tsujino, T., Yoshida, K., Nakayama, H., Haruma, K., Ito, H., Nakamura, Y., Kajiyama, G. & Tahara, E. (1991). Frequent loss of heterozygosity on chromosomes 1q, 5q, and 17p in human gastric carcinomas. *Cancer Res*, **51**, 2926-31.

- Sarr, M.G., Hamilton, S.R., Marrone, G.C. & Cameron, J.L. (1985). Barrett's esophagus: its prevalence and association with adenocarcinoma in patients with symptoms of gastroesophageal reflux. *Am J Surg*, **149**, 187-93.
- Sato, S., Nakamura, Y. & Tsuchiya, E. (1994). Difference of allelotype between squamous cell carcinoma and adenocarcinoma of the lung. *Cancer Res*, **54**, 5652-5.
- Saxena, A., Clark, W.C., Robertson, J.T., Ikejiri, B., Oldfield, E.H. & Ali, I.U. (1992). Evidence for the involvement of a potential second tumor suppressor gene on chromosome 17 distinct from p53 in malignant astrocytomas. *Cancer Res*, **52**, 6716-21.
- Schmid, M., Malicki, D., Nobori, T., Rosenbach, M.D., Campbell, K., Carson, D.A. & Carrera, C.J. (1998). Homozygous deletions of methylthioadenosine phosphorylase (MTAP) are more frequent than p16INK4A (CDKN2) homozygous deletions in primary non-small cell lung cancers (NSCLC). *Oncogene*, **17**, 2669-75.
- Schmidt, H.G., Riddell, R.H., Walther, B., Skinner, D.B. & Riemann, J.F. (1985). Dysplasia in Barrett's esophagus. *J Cancer Res Clin Oncol*, **110**, 145-52.
- Scholes, A.G., Woolgar, J.A., Boyle, M.A., Brown, J.S., Vaughan, E.D., Hart, C.A., Jones, A.S. & Field, J.K. (1998). Synchronous oral carcinomas: independent or common clonal origin? *Cancer Res*, **58**, 2003-6.
- Schultz, D.C., Vanderveer, L., Berman, D.B., Hamilton, T.C., Wong, A.J. & Godwin, A.K. (1996). Identification of two candidate tumor suppressor genes on chromosome 17p13.3. *Cancer Res*, **56**, 1997-2002.
- Schutze, K. & Lahr, G. (1998). Identification of expressed genes by laser-mediated manipulation of single cells. *Nat Biotechnol*, **16**, 737-42.

- Sellers, T.A., Ooi, W.L., Elston, R.C., Chen, V.W., Bailey-Wilson, J.E. & Rothschild, H. (1987). Increased familial risk for non-lung cancer among relatives of lung cancer patients. *Am J Epidemiol*, **126**, 237-46.
- Sermon, K., Lissens, W., Joris, H., Van Steirteghem, A. & Liebaers, I. (1996). Adaptation of the primer extension preamplification (PEP) reaction for preimplantation diagnosis: single blastomere analysis using short PEP protocols. *Mol Hum Reprod*, **2**, 209-12.
- Shapiro, G.I., Park, J.E., Edwards, C.D., Mao, L., Merlo, A., Sidransky, D., Ewen, M.E. & Rollins, B.J. (1995). Multiple mechanisms of p16INK4A inactivation in non-small cell lung cancer cell lines. *Cancer Res*, **55**, 6200-9.
- Shen, C.Y., Yu, J.C., Lo, Y.L., Kuo, C.H., Yue, C.T., Jou, Y.S., Huang, C.S., Lung, J.C. & Wu, C.W. (2000). Genome-wide search for loss of heterozygosity using laser capture microdissected tissue of breast carcinoma: an implication for mutator phenotype and breast cancer pathogenesis. *Cancer Res*, **60**, 3884-92.
- Sheppard, M.N. (1995). *Practical Pulmonary Pathology*. Edward Arnold.
- Shibagaki, I., Shimada, Y., Wagata, T., Ikenaga, M., Imamura, M. & Ishizaki, K. (1994). Allelotype analysis of esophageal squamous cell carcinoma. *Cancer Res*, **54**, 2996-3000.
- Shibata, D. (1994). Extraction of DNA from paraffin-embedded tissue for analysis by polymerase chain reaction: new tricks from an old friend. *Hum Pathol*, **25**, 561-3.

- Shibata, D., Peinado, M.A., Ionov, Y., Malkhosyan, S. & Perucho, M. (1994). Genomic instability in repeated sequences is an early somatic event in colorectal tumorigenesis that persists after transformation. *Nat Genet*, **6**, 273-81.
- Shibuya, K., Fujisawa, T., Hoshino, H., Baba, M., Saitoh, Y., Iizasa, T., Suzuki, M., Otsuji, M., Hiroshima, K. & Ohwada, H. (2001). Fluorescence bronchoscopy in the detection of preinvasive bronchial lesions in patients with sputum cytology suspicious or positive for malignancy. *Lung Cancer*, **32**, 19-25.
- Shields, P.G., Caporaso, N.E., Falk, R.T., Sugimura, H., Trivers, G.E., Trump, B.F., Hoover, R.N., Weston, A. & Harris, C.C. (1993). Lung cancer, race, and a CYP1A1 genetic polymorphism. *Cancer Epidemiol Biomarkers Prev*, **2**, 481-5.
- Shields, P.G. & Harris, C.C. (2000). Cancer risk and low-penetrance susceptibility genes in gene-environment interactions. *J Clin Oncol*, **18**, 2309-15.
- Shih, C., Shilo, B.Z., Goldfarb, M.P., Dannenberg, A. & Weinberg, R.A. (1979). Passage of phenotypes of chemically transformed cells via transfection of DNA and chromatin. *Proc Natl Acad Sci U S A*, **76**, 5714-8.
- Shimizu, E., Shinohara, T., Mori, N., Yokota, J., Tani, K., Izumi, K., Obashi, A. & Ogura, T. (1993). Loss of heterozygosity on chromosome arm 17p in small cell lung carcinomas, but not in neurofibromas, in a patient with von Recklinghausen neurofibromatosis. *Cancer*, **71**, 725-8.

- Shimizu, S., Yatabe, Y., Koshikawa, T., Haruki, N., Hatooka, S., Shinoda, M., Suyama, M., Ogawa, M., Hamajima, N., Ueda, R., Takahashi, T. & Mitsudomi, T. (2000). High frequency of clonally related tumors in cases of multiple synchronous lung cancers as revealed by molecular diagnosis [In Process Citation]. *Clin Cancer Res*, **6**, 3994-9.
- Shimizu, T. & Sekiya, T. (1995). Loss of heterozygosity at 9p21 loci and mutations of the MTS1 and MTS2 genes in human lung cancers. *Int J Cancer*, **63**, 616-20.
- Shiseki, M., Kohno, T., Adachi, J., Okazaki, T., Otsuka, T., Mizoguchi, H., Noguchi, M., Hirohashi, S. & Yokota, J. (1996). Comparative allelotype of early and advanced stage non-small cell lung carcinomas. *Genes Chromosomes Cancer*, **17**, 71-7.
- Shiseki, M., Kohno, T., Nishikawa, R., Sameshima, Y., Mizoguchi, H. & Yokota, J. (1994). Frequent allelic losses on chromosomes 2q, 18q, and 22q in advanced non-small cell lung carcinoma. *Cancer Research*, **54**, 5643-5648.
- Shridhar, R., Shridhar, V., Rivard, S., Siegfried, J.M., Pietraszkiewicz, H., Ensley, J., Pauley, R., Grignon, D., Sakr, W., Miller, O.J. & Smith, D.I. (1996a). Mutations in the arginine-rich protein gene, in lung, breast, and prostate cancers, and in squamous cell carcinoma of the head and neck. *Cancer Res*, **56**, 5576-8.
- Shridhar, V., Rivard, S., Shridhar, R., Mullins, C., Bostick, L., Sakr, W., Grignon, D., Miller, O.J. & Smith, D.I. (1996b). A gene from human chromosomal band 3p21.1 encodes a highly conserved arginine-rich protein and is mutated in renal cell carcinomas. *Oncogene*, **12**, 1931-9.

- Shridhar, V., Siegfried, J., Hunt, J., del Mar Alonso, M. & Smith, D.I. (1994). Genetic instability of microsatellite sequences in many non- small cell lung carcinomas. *Cancer Research*, **54**, 2084-2087.
- Shridhar, V., Wang, L., Rosati, R., Paradee, W., Shridhar, R., Mullins, C., Sakr, W., Grignon, D., Miller, O.J., Sun, Q.C., Petros, J. & Smith, D.I. (1997). Frequent breakpoints in the region surrounding FRA3B in sporadic renal cell carcinomas. *Oncogene*, **14**, 1269-77.
- Sidransky, D. (1995). Importance of chromosome 9p loss in human lung cancer [editorial; comment]. *J Natl Cancer Inst*, **87**, 1201-2.
- Sikkink, S.K., Rehman, I. & Rees, J.L. (1997). Deletion mapping of chromosome 3p and 13q and preliminary analysis of the FHIT gene in human nonmelanoma skin cancer. *J Invest Dermatol*, **109**, 801-5.
- Simoneau, A.R., Spruck, C.H., Gonzalez-Zulueta, M., Gonzalzo, M.L., Chan, M.F., Tsai, Y.C., Dean, M., Steven, K., Horn, T. & Jones, P.A. (1996). Evidence for two tumor suppressor loci associated with proximal chromosome 9p to q and distal chromosome 9q in bladder cancer and the initial screening for GAS1 and PTC mutations. *Cancer Res*, **56**, 5039-43.
- Slaughter, D.P., Southwick, H.W. & Smejkal, W. (1953). "Field cancerization" in oral stratified squamous epithelium: clinical implications for a multicentric origin. *Cancer*, **6**, 693-98.
- Smith, A.L., Hung, J., Walker, L., Rogers, T.E., Vuitch, F., Lee, E. & Gazdar, A.F. (1996). Extensive areas of aneuploidy are present in the respiratory epithelium of lung cancer patients. *Br J Cancer*, **73**, 203-9.
- Smith, R.A., Nigam, B.K. & Thompson, J.M. (1976). Second primary lung carcinoma. *Thorax*, **31**, 507-16.

- Smith, S.A., DiCioccio, R.A., Struewing, J.P., Easton, D.F., Gallion, H.H., Albertsen, H., Mazoyer, S., Johansson, B., Steichen-Gersdorf, E., Stratton, M. & et al. (1994). Localisation of the breast-ovarian cancer susceptibility gene (BRCA1) on 17q12-21 to an interval of ≤ 1 cM. *Genes Chromosomes Cancer*, **10**, 71-6.
- Soria, J.C., Jang, S.J., Khuri, F.R., Hassan, K., Liu, D., Hong, W.K. & Mao, L. (2000). Overexpression of cyclin B1 in early-stage non-small cell lung cancer and its clinical implication. *Cancer Res*, **60**, 4000-4.
- Sozzi, G., Alder, H., Tornielli, S., Corletto, V., Baffa, R., Veronese, M.L., Negrini, M., Pilotti, S., Pierotti, M.A., Huebner, K. & Croce, C.M. (1996a). Aberrant FHIT transcripts in Merkel cell carcinoma. *Cancer Res*, **56**, 2936-9.
- Sozzi, G., Miozzo, M., Donghi, R., Pilotti, S., Cariani, C.T., Pastorino, U., Della Porta, G. & Pierotti, M.A. (1992). Deletions of 17p and p53 mutations in preneoplastic lesions of the lung. *Cancer Res*, **52**, 6079-82.
- Sozzi, G., Miozzo, M., Pastorino, U., Pilotti, S., Donghi, R., Giarola, M., De Gregorio, L., Manenti, G., Radice, P., Minoletti, F. & et al. (1995). Genetic evidence for an independent origin of multiple preneoplastic and neoplastic lung lesions. *Cancer Res*, **55**, 135-40.
- Sozzi, G., Pastorino, U., Moiraghi, L., Tagliabue, E., Pezzella, F., Ghirelli, C., Tornielli, S., Sard, L., Huebner, K., Pierotti, M.A., Croce, C.M. & Pilotti, S. (1998). Loss of FHIT function in lung cancer and preinvasive bronchial lesions. *Cancer Res*, **58**, 5032-7.

- Sozzi, G., Sard, L., De Gregorio, L., Marchetti, A., Musso, K., Buttitta, F.,
Tornielli, S., Pellegrini, S., Veronese, M.L., Manenti, G., Incarbone, M.,
Chella, A., Angeletti, C.A., Pastorino, U., Huebner, K., Bevilaqua, G.,
Pilotti, S., Croce, C.M. & Pierotti, M.A. (1997a). Association between
cigarette smoking and FHIT gene alterations in lung cancer. *Cancer Res*,
57, 2121-3.
- Sozzi, G., Tornielli, S., Tagliabue, E., Sard, L., Pezzella, F., Pastorino, U.,
Minoletti, F., Pilotti, S., Ratcliffe, C., Veronese, M.L., Goldstraw, P.,
Huebner, K., Croce, C.M. & Pierotti, M.A. (1997b). Absence of Fhit
protein in primary lung tumors and cell lines with FHIT gene
abnormalities. *Cancer Res*, **57**, 5207-12.
- Sozzi, G., Veronese, M.L., Negrini, M., Baffa, R., Cotticelli, M.G., Inoue, H.,
Tornielli, S., Pilotti, S., De Gregorio, L., Pastorino, U., Pierotti, M.A.,
Ohta, M., Huebner, K. & Croce, C.M. (1996b). The FHIT gene 3p14.2 is
abnormal in lung cancer. *Cancer Res*, **56**, 2472-4.
- Spandidos, D.A., Zakinthinos, S., Petraki, C., Sotsiou, F., Yiagnisis, M.,
Dimopoulos, A.M., Roussos, C. & Field, J.K. (1990). Expression of ras
p21 and myc p62 oncoproteins in small cell and non small cell carcinoma
of the lung. *Anticancer Res*, **10**, 1105-14.
- Sparkes, R.S., Murphree, A.L., Lingua, R.W., Sparkes, M.C., Field, L.L.,
Funderburk, S.J. & Benedict, W.F. (1983). Gene for hereditary
retinoblastoma assigned to human chromosome 13 by linkage to
esterase D. *Science*, **219**, 971-3.

- Sparkes, R.S. & Sparkes, M.C. (1983). Esterase D studies in human retinoblastoma. *Isozymes Curr Top Biol Med Res*, **11**, 173-82.
- Sparkes, R.S., Sparkes, M.C., Wilson, M.G., Towner, J.W., Benedict, W., Murphree, A.L. & Yunis, J.J. (1980). Regional assignment of genes for human esterase D and retinoblastoma to chromosome band 13q14. *Science*, **208**, 1042-4.
- Spitz, M.R., Wei, Q., Li, G. & Wu, X. (1999). Genetic susceptibility to tobacco carcinogenesis. *Cancer Invest*, **17**, 645-59.
- Stellman, S.D., Muscat, J.E., Hoffmann, D. & Wynder, E.L. (1997a). Impact of filter cigarette smoking on lung cancer histology. *Prev Med*, **26**, 451-6.
- Stellman, S.D., Muscat, J.E., Thompson, S., Hoffmann, D. & Wynder, E.L. (1997b). Risk of squamous cell carcinoma and adenocarcinoma of the lung in relation to lifetime filter cigarette smoking. *Cancer*, **80**, 382-8.
- Strissel, P.L., Strick, R., Tomek, R.J., Roe, B.A., Rowley, J.D. & Zeleznik-Le, N.J. (2000). DNA structural properties of AF9 are similar to MLL and could act as recombination hot spots resulting in MLL/AF9 translocations and leukemogenesis. *Hum Mol Genet*, **9**, 1671-9.
- Su, C., Ye, Y., Cao, X. & Shan, X. (1999). Homozygous deletion of CDKN2/p16 gene in lung cancers [In Process Citation]. *Chung Hua I Hsueh I Chuan Hsueh Tsa Chih*, **16**, 315-7.
- Sun, F., Arnheim, N. & Waterman, M.S. (1995). Whole genome amplification of single cells: mathematical analysis of PEP and tagged PCR. *Nucleic Acids Res*, **23**, 3034-40.

- Sundaresan, V., Ganly, P., Hasleton, P., Rudd, R., Sinha, G., Bleeheh, N.M. & Rabbitts, P. (1992). p53 and chromosome 3 abnormalities, characteristic of malignant lung tumours, are detectable in preinvasive lesions of the bronchus. *Oncogene*, **7**, 1989-97.
- Sundaresan, V., Heppell-Parton, A., Coleman, N., Miozzo, M., Sozzi, G., Ball, R., Cary, N., Hasleton, P., Fowler, W. & Rabbitts, P. (1995). Somatic genetic changes in lung cancer and precancerous lesions. *Ann Oncol*, **6**, 27-31; discussion 31-2.
- Sung, N.S., Zeng, Y. & Raab-Traub, N. (2000). Alterations on chromosome 3 in endemic and nonendemic nasopharyngeal carcinoma. *Int J Cancer*, **86**, 244-50.
- Super, H.G., Strissel, P.L., Sobulo, O.M., Burian, D., Reshmi, S.C., Roe, B., Zeleznik-Le, N.J., Diaz, M.O. & Rowley, J.D. (1997). Identification of complex genomic breakpoint junctions in the t(9;11) MLL-AF9 fusion gene in acute leukemia. *Genes Chromosomes Cancer*, **20**, 185-95.
- Swafford, D.S., Middleton, S.K., Palmisano, W.A., Nikula, K.J., Tesfaigzi, J., Baylin, S.B., Herman, J.G. & Belinsky, S.A. (1997). Frequent aberrant methylation of p16INK4a in primary rat lung tumors. *Mol Cell Biol*, **17**, 1366-74.
- Swift, A., Risk, J.M., Kingsnorth, A.N., Wright, T.A., Myskow, M. & Field, J.K. (1995). Frequent loss of heterozygosity on chromosome 17 at 17q11.2-q12 in Barrett's adenocarcinoma. *Br J Cancer*, **71**, 995-8.

- Takahashi, N., Miura, I., Ohshima, A., Nimura, T., Hashimoto, K., Hatano, Y., Utsumi, S., Kume, M., Saito, K., Kobayashi, Y., Saito, M., Seto, M., Ueda, R. & Miura, A.B. (1996). Translocation (9;11;22)(p22;q23;q11). A new type of complex variant translocation of t(9;11)(p22;q23) with MLL rearrangement. *Cancer Genet Cytogenet*, **88**, 26-9.
- Tammemagi, M.C., McLaughlin, J.R., Mullen, J.B., Bull, S.B., Johnston, M.R., Tsao, M.S. & Casson, A.G. (2000). A study of smoking, p53 tumor suppressor gene alterations and non-small cell lung cancer. *Ann Epidemiol*, **10**, 176-85.
- Tanaka, H., Fujii, Y., Hirabayashi, H., Miyoshi, S., Sakaguchi, M., Yoon, H.E. & Matsuda, H. (1998). Disruption of the RB pathway and cell-proliferative activity in non-small-cell lung cancers. *Int J Cancer*, **79**, 111-5.
- Tanimoto, K., Hayashi, S., Tsuchiya, E., Tokuchi, Y., Kobayashi, Y., Yoshiga, K., Okui, T., Kobayashi, M. & Ichikawa, T. (2000). Abnormalities of the FHIT gene in human oral carcinogenesis. *Br J Cancer*, **82**, 838-43.
- Tatusova, T.A. & Madden, T.L. (1999). BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences [published erratum appears in FEMS Microbiol Lett 1999 Aug 1;177(1):187-8]. *FEMS Microbiol Lett*, **174**, 247-50.
- Telenius, H., Carter, N.P., Bebb, C.E., Nordenskjold, M., Ponder, B.A. & Tunnacliffe, A. (1992a). Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer. *Genomics*, **13**, 718-25.

- Telenius, H., Pelmear, A.H., Tunnacliffe, A., Carter, N.P., Behmel, A., Ferguson-Smith, M.A., Nordenskjold, M., Pfragner, R. & Ponder, B.A. (1992b). Cytogenetic analysis by chromosome painting using DOP-PCR amplified flow-sorted chromosomes. *Genes Chromosomes Cancer*, **4**, 257-63.
- Thiagalingam, S., Lisitsyn, N.A., Hamaguchi, M., Wigler, M.H., Willson, J.K., Markowitz, S.D., Leach, F.S., Kinzler, K.W. & Vogelstein, B. (1996). Evaluation of the FHIT gene in colorectal cancers. *Cancer Res*, **56**, 3173-9.
- Thiberville, L., Payne, P., Vielkinds, J., LeRiche, J., Horsman, D., Nouvet, G., Palcic, B. & Lam, S. (1995). Evidence of cumulative gene losses with progression of premalignant epithelial lesions to carcinoma of the bronchus. *Cancer Res*, **55**, 5133-9.
- Thomas, D.C., Roberts, J.D. & Kunkel, T.A. (1991a). Heteroduplex repair in extracts of human HeLa cells. *J Biol Chem*, **266**, 3744-51.
- Thomas, D.C., Roberts, J.D., Sabatino, R.D., Myers, T.W., Tan, C.K., Downey, K.M., So, A.G., Bambara, R.A. & Kunkel, T.A. (1991b). Fidelity of mammalian DNA replication and replicative DNA polymerases. *Biochemistry*, **30**, 11751-9.
- Thorlacius, S., Tryggvadottir, L., Olafsdottir, G.H., Jonasson, J.G., Ogmundsdottir, H.M., Tulinius, H. & Eyfjord, J.E. (1995). Linkage to BRCA2 region in hereditary male breast cancer. *Lancet*, **346**, 544-5.
- Thun, M.J., Lally, C.A., Flannery, J.T., Calle, E.E., Flanders, W.D. & Heath, C.W. (1997). Cigarette smoking and changes in the histopathology of lung cancer. *J Natl Cancer Inst*, **89**, 1580-6.

- Tian, D., Feng, Z., Hanley, N.M., Setzer, R.W., Mumford, J.L. & DeMarini, D.M. (1998). Multifocal accumulation of p53 protein in esophageal carcinoma: evidence for field cancerization. *Int J Cancer*, **78**, 568-75.
- Todd, R. & Wong, D.T. (1999). Oncogenes. *Anticancer Res*, **19**, 4729-46.
- Tokuchi, Y., Kobayashi, Y., Hayashi, S., Hayashi, M., Tanimoto, K., Hashimoto, T., Nishida, K., Ishikawa, Y., Nakagawa, K., Satoh, Y., Yamamoto, M. & Tsuchiya, E. (1999). Abnormal FHIT transcripts found in both lung cancer and normal lung tissue. *Genes Chromosomes Cancer*, **24**, 105-11.
- Travis, W.D., Colby, T.V., Corrin, B., Shimosato, Y., Brambilla, E. & Sobin, L.H. (1999). World Health Organisation Histological Typing Of Lung And Pleural Tumours (Third Edition). *Springer*.
- Tsao, S.W., Mok, C.H., Oike, K., Muto, M., Goodman, H.M., Sheets, E.E., Berkowitz, R.S., Knapp, R.C. & Lau, C.C. (1991). Involvement of p53 gene in the allelic deletion of chromosome 17p in human ovarian tumors. *Anticancer Res*, **11**, 1975-82.
- Tsuchiya, E., Kitagawa, T., Oh, S., Nakagawa, K., Matsubara, T., Kinoshita, I. & Sugano, H. (1987). Incidence of squamous metaplasia in large bronchi of Japanese lungs: relation to pulmonary carcinomas of various subtypes. *Jpn J Cancer Res*, **78**, 559-64.
- Tsuchiya, E., Nakamura, Y., Weng, S.Y., Nakagawa, K., Tsuchiya, S., Sugano, H. & Kitagawa, T. (1992). Allelotype of non-small cell lung carcinoma-- comparison between loss of heterozygosity in squamous cell carcinoma and adenocarcinoma. *Cancer Res*, **52**, 2478-81.

- Tsuchiya, E., Tanigami, A., Ishikawa, Y., Nishida, K., Hayashi, M., Tokuchi, Y., Hashimoto, T., Okumura, S., Tsuchiya, S. & Nakagawa, K. (2000). Three new regions on chromosome 17p13.3 distal to p53 with possible tumor suppressor gene involvement in lung cancer. *Jpn J Cancer Res*, **91**, 589-96.
- Ueno, K., Kumagai, T., Kijima, T., Kishimoto, T. & Hosoe, S. (1998). Cloning and tissue expression of cDNAs from chromosome 5q21-22 which is frequently deleted in advanced lung cancer. *Hum Genet*, **102**, 63-8.
- Uyama, T., Monden, Y., Tsuyuguchi, M., Harada, K., Kimura, S. & Taniki, T. (1989). Lung cancer in chromate workers: high-risk group for multiple lung cancer. *J Surg Oncol*, **41**, 213-8.
- van Dekken, H., Vissers, C.J., Tilanus, H.W., Tanke, H.J. & Rosenberg, C. (1999). Clonal analysis of a case of multifocal oesophageal (Barrett's) adenocarcinoma by comparative genomic hybridization. *J Pathol*, **188**, 263-266.
- van Oijen, M.G., Leppers Vd Straat, F.G., Tilanus, M.G. & Slootweg, P.J. (2000). The origins of multiple squamous cell carcinomas in the aerodigestive tract. *Cancer*, **88**, 884-93.
- van Rees, B.P., Cleton-Jansen, A.M., Cense, H.A., Polak, M.M., Clement, M.J., Drillenburger, P., van Lanschot, J.J. & Offerhaus, G.J. (2000). Molecular evidence of field cancerization in a patient with 7 tumors of the aerodigestive tract. *Hum Pathol*, **31**, 269-71.
- van Rens, M.T., Schramel, F.M., Elbers, J.R. & Lammers, J.W. (2001). The clinical value of lung imaging fluorescence endoscopy for detecting synchronous lung cancer. *Lung Cancer*, **32**, 13-8.

- Vandre, D.D., Feng, Y. & Ding, M. (2000). Cell cycle-dependent phosphorylation of centrosomes: localization of phosphopeptide specific antibodies to the centrosome. *Microsc Res Tech*, **49**, 458-66.
- Virgilio, L., Shuster, M., Gollin, S.M., Veronese, M.L., Ohta, M., Huebner, K. & Croce, C.M. (1996). FHIT gene alterations in head and neck squamous cell carcinomas. *Proc Natl Acad Sci U S A*, **93**, 9770-5.
- Virmani, A.K., Fong, K.M., Kodagoda, D., McIntire, D., Hung, J., Tonk, V., Minna, J.D. & Gazdar, A.F. (1998). Allelotyping demonstrates common and distinct patterns of chromosomal loss in human lung cancer types. *Genes Chromosomes Cancer*, **21**, 308-19.
- Vogelstein, B., Fearon, E.R., Hamilton, S.R., Kern, S.E., Preisinger, A.C., Leppert, M., Nakamura, Y., White, R., Smits, A.M. & Bos, J.L. (1988). Genetic alterations during colorectal-tumor development. *N Engl J Med*, **319**, 525-32.
- Vogelstein, B. & Kinzler, K.W. (1993). The multistep nature of cancer. [Review]. *Trends in Genetics*, **9**, 138-141.
- von Eggeling, F., Davies, H., Lomas, L., Fiedler, W., Junker, K., Claussen, U. & Ernst, G. (2000). Tissue-specific microdissection coupled with ProteinChip array technologies: applications in cancer research. *Biotechniques*, **29**, 1066-70.
- Waber, P.G., Lee, N.K. & Nisen, P.D. (1996). Frequent allelic loss at chromosome arm 3p is distinct from genetic alterations of the Von-Hippel Lindau tumor suppressor gene in head and neck cancer. *Oncogene*, **12**, 365-9.

- Wakamatsu, K., Nakanishi, Y., Takayama, K., Miyazaki, H., Hayashi, K. & Hara, N. (1999). Frequent expression of p53 protein without mutation in the atypical epithelium of human bronchus. *Am J Respir Cell Mol Biol*, **21**, 209-15.
- Wales, M.M., Biel, M.A., el Deiry, W., Nelkin, B.D., Issa, J.P., Cavenee, W.K., Kuerbitz, S.J. & Baylin, S.B. (1995). p53 activates expression of HIC-1, a new candidate tumour suppressor gene on 17p13.3. *Nat Med*, **1**, 570-7.
- Walker, C., Robertson, L.J., Myskow, M.W., Pendleton, N. & Dixon, G.R. (1994). p53 expression in normal and dysplastic bronchial epithelium and in lung carcinomas. *Br J Cancer*, **70**, 297-303.
- Wang, L., Darling, J., Zhang, J.S., Huang, H., Liu, W. & Smith, D.I. (1999). Allele-specific late replication and fragility of the most active common fragile site, FRA3B. *Hum Mol Genet*, **8**, 431-7.
- Weber, E., Berta, G., Tousson, A., St John, P., Green, M.W., Gopalokrishnan, U., Jilling, T., Sorscher, E.J., Elton, T.S., Abrahamson, D.R. & et al. (1994). Expression and polarized targeting of a rab3 isoform in epithelial cells. *J Cell Biol*, **125**, 583-94.
- Wei, Q. & Spitz, M.R. (1997). The role of DNA repair capacity in susceptibility to lung cancer: a review. *Cancer Metastasis Rev*, **16**, 295-307.
- Wells, D., Sherlock, J.K., Handyside, A.H. & Delhanty, J.D. (1999). Detailed chromosomal and molecular genetic analysis of single cells by whole genome amplification and comparative genomic hybridisation. *Nucleic Acids Res*, **27**, 1214-8.

- Werness, B.A., Parvatiyar, P., Ramus, S.J., Whittemore, A.S., Garlinghouse-Jones, K., Oakley-Girvan, I., DiCioccio, R.A., Wiest, J., Tsukada, Y., Ponder, B.A. & Piver, M.S. (2000). Ovarian carcinoma in situ with germline BRCA1 mutation and loss of heterozygosity at BRCA1 and TP53. *J Natl Cancer Inst*, **92**, 1088-91.
- Weston, A., Willey, J.C., Modali, R., Sugimura, H., McDowell, E.M., Resau, J., Light, B., Haugen, A., Mann, D.L., Trump, B.F. & et al. (1989). Differential DNA sequence deletions from chromosomes 3, 11, 13, and 17 in squamous-cell carcinoma, large-cell carcinoma, and adenocarcinoma of the human lung. *Proc Natl Acad Sci U S A*, **86**, 5099-103.
- Wieland, I. & Bohm, M. (1994). Frequent allelic deletion at a novel locus on chromosome 5 in human lung cancer. *Cancer Res*, **54**, 1772-4.
- Wieland, I., Bohm, M., Arden, K.C., Ammermuller, T., Bogatz, S., Viars, C.S. & Rajewsky, M.F. (1996). Allelic deletion mapping on chromosome 5 in human carcinomas. *Oncogene*, **12**, 97-102.
- Wiest, J.S., Franklin, W.A., Otstot, J.T., Forbey, K., Varella-Garcia, M., Rao, K., Drabkin, H., Gemmill, R., Ahrent, S., Sidransky, D., Saccomanno, G., Fountain, J.W. & Anderson, M.W. (1997). Identification of a novel region of homozygous deletion on chromosome 9p in squamous cell carcinoma of the lung: the location of a putative tumor suppressor gene. *Cancer Res*, **57**, 1-6.
- Williamson, W.A., Ellis, F.H., Gibb, S.P., Shahian, D.M., Aretz, H.T., Heatley, G.J. & Watkins, E. (1991). Barrett's esophagus. Prevalence and incidence of adenocarcinoma. *Arch Intern Med*, **151**, 2212-6.

- Wistuba, II, Behrens, C., Milchgrub, S., Bryant, D., Hung, J., Minna, J.D. & Gazdar, A.F. (1999a). Sequential molecular abnormalities are involved in the multistage development of squamous cell lung carcinoma. *Oncogene*, **18**, 643-50.
- Wistuba, II, Behrens, C., Virmani, A.K., Mele, G., Milchgrub, S., Girard, L., Fondon, J.W., 3rd, Garner, H.R., McKay, B., Latif, F., Lerman, M.I., Lam, S., Gazdar, A.F. & Minna, J.D. (2000a). High resolution chromosome 3p allelotyping of human lung cancer and preneoplastic/preinvasive bronchial epithelium reveals multiple, discontinuous sites of 3p allele loss and three regions of frequent breakpoints. *Cancer Res*, **60**, 1949-60.
- Wistuba, II, Behrens, C., Virmani, A.K., Mele, G., Milchgrub, S., Girard, L., Fondon, J.W., 3rd, Garner, H.R., McKay, B., Latif, F., Lerman, M.I., Lam, S., Gazdar, A.F. & Minna, J.D. (2000b). High resolution chromosome 3p allelotyping of human lung cancer and preneoplastic/preinvasive bronchial epithelium reveals multiple, discontinuous sites of 3p allele loss and three regions of frequent breakpoints [In Process Citation]. *Cancer Res*, **60**, 1949-60.
- Wistuba, II, Behrens, C., Virmani, A.K., Milchgrub, S., Syed, S., Lam, S., Mackay, B., Minna, J.D. & Gazdar, A.F. (1999b). Allelic losses at chromosome 8p21-23 are early and frequent events in the pathogenesis of lung cancer. *Cancer Res*, **59**, 1973-9.
- Wistuba, II, Bryant, D., Behrens, C., Milchgrub, S., Virmani, A.K., Ashfaq, R., Minna, J.D. & Gazdar, A.F. (1999c). Comparison of features of human lung cancer cell lines and their corresponding tumors. *Clin Cancer Res*, **5**, 991-1000.

- Wistuba, II, Lam, S., Behrens, C., Virmani, A.K., Fong, K.M., LeRiche, J., Samet, J.M., Srivastava, S., Minna, J.D. & Gazdar, A.F. (1997a). Molecular damage in the bronchial epithelium of current and former smokers. *J Natl Cancer Inst*, **89**, 1366-73.
- Wistuba, II, Montellano, F.D., Milchgrub, S., Virmani, A.K., Behrens, C., Chen, H., Ahmadian, M., Nowak, J.A., Muller, C., Minna, J.D. & Gazdar, A.F. (1997b). Deletions of chromosome 3p are frequent and early events in the pathogenesis of uterine cervical carcinoma. *Cancer Res*, **57**, 3154-8.
- Wood, M.E., Kelly, K., Mullineaux, L.G. & Bunn, P.A. (2000). The inherited nature of lung cancer: a pilot study [In Process Citation]. *Lung Cancer*, **30**, 135-44.
- Woods, D.B. & Vousden, K.H. (2001). Regulation of p53 function. *Exp Cell Res*, **264**, 56-66.
- Wooster, R., Neuhausen, S.L., Mangion, J., Quirk, Y., Ford, D., Collins, N., Nguyen, K., Seal, S., Tran, T., Averill, D. & et al. (1994). Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. *Science*, **265**, 2088-90.
- Wright, N.A. (2000). Epithelial stem cell repertoire in the gut: clues to the origin of cell lineages, proliferative units and cancer. *Int J Exp Pathol*, **81**, 117-43.
- Wu, C.L., Sloan, P., Read, A.P., Harris, R. & Thakker, N. (1994). Deletion mapping on the short arm of chromosome 3 in squamous cell carcinoma of the oral cavity. *Cancer Research*, **54**, 6484-6488.
- Wu, X., Bayle, J.H., Olson, D. & Levine, A.J. (1993). The p53-mdm-2 autoregulatory feedback loop. *Genes Dev*, **7**, 1126-32.

- Xiao, S., Li, D., Corson, J.M., Vijg, J. & Fletcher, J.A. (1995). Codeletion of p15 and p16 genes in primary non-small cell lung carcinoma. *Cancer Res*, **55**, 2968-71.
- Xinarianos, G., Liloglou, T., Prime, W., Maloney, P., Callaghan, J., Fielding, P., Gosney, J.R. & Field, J.K. (2000). hMLH1 and hMSH2 expression correlates with allelic imbalance on chromosome 3p in non-small cell lung carcinomas. *Cancer Res*, **60**, 4216-21.
- Xu, L., Flynn, B.J., Ungar, S., Pass, H.I., Linnainmaa, K., Mattson, K. & Gerwin, B.I. (1999). Asbestos induction of extended lifespan in normal human mesothelial cells: interindividual susceptibility and SV40 T antigen. *Carcinogenesis*, **20**, 773-83.
- Yamakawa, K., Takahashi, T., Horio, Y., Murata, Y., Takahashi, E., Hibi, K., Yokoyama, S., Ueda, R. & Nakamura, Y. (1993). Frequent homozygous deletions in lung cancer cell lines detected by a DNA marker located at 3p21.3-p22. *Oncogene*, **8**, 327-330.
- Yang, P., Schwartz, A.G., McAllister, A.E., Swanson, G.M. & Aston, C.E. (1999). Lung cancer risk in families of nonsmoking probands: heterogeneity by age at diagnosis. *Genet Epidemiol*, **17**, 253-73.
- Yatabe, Y., Konishi, H., Mitsudomi, T., Nakamura, S. & Takahashi, T. (2000). Topographical distributions of allelic loss in individual non-small-cell lung cancers. *Am J Pathol*, **157**, 985-93.
- Yokota, J. & Sugimura, T. (1993). Multiple steps in carcinogenesis involving alterations of multiple tumor suppressor genes. *Faseb J*, **7**, 920-5.

- Yokota, J., Wada, M., Shimosato, Y., Terada, M. & Sugimura, T. (1987). Loss of heterozygosity on chromosomes 3, 13, and 17 in small-cell carcinoma and on chromosome 3 in adenocarcinoma of the lung. *Proc Natl Acad Sci U S A*, **84**, 9252-6.
- Yoshioka, H., Takeuchi, T., Matsuno, Y., Yamada, T., Shimosato, Y., Hirohashi, S. & Noguchi, M. (1998). Analysis of loss of heterozygosity in small adenocarcinomas of the lung. *Jpn J Clin Oncol*, **28**, 240-4.
- Yu, L., Hui-chen, F., Chen, Y., Zou, R., Yan, S., Chun-xiang, L., Wu-ru, W. & Li, P. (1999). Differential expression of RAB5A in human lung adenocarcinoma cells with different metastasis potential. *Clin Exp Metastasis*, **17**, 213-9.
- Yustein, A.S., Harper, J.C., Petroni, G.R., Cummings, O.W., Moskaluk, C.A. & Powell, S.M. (1999). Allelotype of gastric adenocarcinoma. *Cancer Res*, **59**, 1437-41.
- Zahraoui, A., Touchot, N., Chardin, P. & Tavitian, A. (1989). The human Rab genes encode a family of GTP-binding proteins related to yeast YPT1 and SEC4 products involved in secretion. *J Biol Chem*, **264**, 12394-401.
- Zhang, L., Cui, X., Schmitt, K., Hubert, R., Navidi, W. & Arnheim, N. (1992). Whole genome amplification from a single cell: implications for genetic analysis. *Proc Natl Acad Sci U S A*, **89**, 5847-51.
- Zhang, Z.F., Kurtz, R.C., Sun, M., Karpeh, M., Yu, G.P., Gargon, N., Fein, J.S., Georgopoulos, S.K. & Harlap, S. (1996). Adenocarcinomas of the esophagus and gastric cardia: medical conditions, tobacco, alcohol, and socioeconomic factors. *Cancer Epidemiol Biomarkers Prev*, **5**, 761-8.

- Zhuang, Z., Vortmeyer, A.O., Mark, E.J., Odze, R., Emmert-Buck, M.R., Merino, M.J., Moon, H., Liotta, L.A. & Duray, P.H. (1996). Barrett's esophagus: metaplastic cells with loss of heterozygosity at the APC gene locus are clonal precursors to invasive adenocarcinoma. *Cancer Res*, **56**, 1961-4.
- Zienolddiny, S., Ryberg, D., Gazdar, A.F. & Haugen, A. (1999). DNA mismatch binding in human lung tumor cell lines. *Lung Cancer*, **26**, 15-25.
- Zimonjic, D.B., Druck, T., Ohta, M., Kastury, K., Croce, C.M., Popescu, N.C. & Huebner, K. (1997). Positions of chromosome 3p14.2 fragile sites (FRA3B) within the FHIT gene. *Cancer Res*, **57**, 1166-70.
- Zou, T.T., Lei, J., Shi, Y.Q., Yin, J., Wang, S., Souza, R.F., Kong, D., Shimada, Y., Smolinski, K.N., Greenwald, B.D., Abraham, J.M., Harpaz, N. & Meltzer, S.J. (1997). FHIT gene alterations in esophageal cancer and ulcerative colitis (UC). *Oncogene*, **15**, 101-5.

APPENDIX A : METHODS AND SOLUTIONS

A1 Paraffin embedding of bronchial samples

All Samples were fixed and processed using the Shandon Citadel 2000™ autoprocessor then embedded using the Shandon Histocentre2™. Lung tissue was fixed overnight then sampled and placed in formalin for transfer to the RCICLCR for processing.

1. 10% Neutral buffered formalin x 5 minutes
2. 70% Ethanol x 90 minutes
3. 90% Ethanol x 90 minutes
4. 100% Ethanol x 90 minutes
5. 100% Ethanol x 90 minutes
6. 100% Ethanol x 90 minutes
7. 100% Ethanol x 90 minutes
8. Xylene x 90 minutes
9. Xylene x 90 minutes
10. Xylene x 90 minutes
11. Paraffin Wax x 120 minutes
12. Paraffin Wax x 180 minutes

After impregnation with paraffin wax lung tissue samples were embedded and left overnight at 4°C for wax to harden. Blocks were stored at room temperature.

A2 Hematoxylin and Eosin staining of bronchial samples

All samples were H&E stained using the Shandon Linistain™ autostainer.

Distilled water steps were processed with agitation to ensure complete rinsing of slides.

1. Xylene x 5 minutes
2. Xylene x 5 minutes
3. 100% Ethanol x 1 minute
4. 100% Ethanol x 1 minute
5. 95% Ethanol x 1 minute
6. 70% Ethanol x 1 minute
7. 50% Ethanol x 30 seconds
8. Distilled water x 30 seconds
9. Hematoxylin x 2 minutes
10. Distilled water x 30 seconds
11. 1% Acid alcohol x 30 seconds
12. Scott's tap water x 30 seconds
13. Eosin x 1 minute
14. Distilled water x 30 seconds
15. 70% Ethanol x 30 seconds
16. 95% Ethanol x 1 minute
17. 100% Ethanol x 1 minute
18. 100% Ethanol x 1 minute

19. Xylene x 1 minute

20. Xylene x 1 minute

Slides were stored in Xylene until coverslipped using 100-200 μ l DPX liquid mounting medium then left to dry overnight at room temperature.

A3 Arcturus LCM Hematoxylin and Eosin staining protocol for LCM

Tissue sections prepared for LCM were hematoxylin and eosin stained according to Arcturus recommended protocols (www.arctur.com) with slight variations. Solutions were changed regularly to ensure optimum LCM.

1. Xylene x 5 minutes
2. Xylene x 5 minutes
3. 100% Ethanol x 1 minute
4. 95% Ethanol x 1 minute
5. 70% Ethanol x 30 seconds
6. Distilled water x 30 seconds
7. Hematoxylin x 30 seconds
8. Distilled water x 30 seconds
9. Scott's tap water x 30 seconds
10. 70% Ethanol x 30 seconds
11. 95% Ethanol x 30 seconds
12. Eosin x 30 seconds
13. 70% Ethanol x 30 seconds
14. 95% Ethanol x 1 minute
15. 100% Ethanol x 1 minute
16. 100% Ethanol x 2 minutes
17. Xylene x 5 minutes
18. Xylene x 10 minutes

19. Air dry 20 minutes in a hood or vacuum desiccator

Samples were stored in a desiccator when not in use.

A4 Toluidine blue staining protocol for LCM

1. Xylene x 5 minutes
2. Xylene x 5 minutes
3. 100% Ethanol x 1 minute
4. 100% Ethanol x 1minute
5. 95% Ethanol x 1 minute
6. Distilled water x 30 seconds
- Toluidine blue (0.001% w/v) x 30 seconds
7. Distilled water x 30 seconds
9. 95% Ethanol x 30 seconds
10. 95% Ethanol x 1 minute
11. 100% Ethanol x 1 minute
12. 100% Ethanol x 2 minutes
13. Xylene x 5 minutes
14. Xylene x 10 minutes
15. Air dry 20 minutes in a hood or vacuum desiccator

Samples were stored in a desiccator when not in use.

A5 Solution And Buffer Preparation

1 x P Lysis buffer (LCM lysis buffer)

10 mM Tris-HCl; pH 8.0

1 mM EDTA; pH 8.0

1% (v/v) Tween-20

0.04% (w/v) Proteinase K (Qiagen, UK)

10 x TBE

89 mM Tris-borate

89 mM Boric acid

8 mM EDTA; pH 8.0

To make up 2 litres of 10 x stock dissolve 216g of Tris base, 110g of Boric acid and 80 mls EDTA in a final volume of 2 litres of distilled water. Autoclave for 15 minutes at 10 lb/sq. in. on liquid cycle. Store at room temperature.

SD Media (YAC media)

0.7% Yeast Nitrogen Base without amino acids

2% Glucose

Adenine (55mg/ml)

Tyrosine (55mg/ml)

To make up 800 mls SD media dissolve 5.6g yeast nitrogen base and 16g glucose in 500mls distilled water. Add 44 mg each of adenine and tyrosine then add distilled water to 800 mls. Autoclave for 15 minutes at 10 lb/sq. in. on liquid cycle. Store at room temperature. For solid media add 15 g agar/litre before autoclaving.

LB broth (BAC media)

1% Tryptone (peptone 140)

0.5% Yeast extract

1% NaCl

pH to 7.0 using 10M sodium hydroxide solution

To make up 1 Litre LB media dissolve 10g tryptone, 5g yeast extract and 10g NaCl in 950 mls distilled water. pH to 7.0. Autoclave for 15 minutes at 10 lb/sq. in. on liquid cycle.. Store at room temperature. Add 12.5 µg/ml chloramphenicol (dissolved in ethanol) before use and after autoclaving. For solid media add 15 g agar/litre before autoclaving.

Solution I (BAC DNA extraction)

50 mM glucose

25 mM Tris-HCl; pH 8.0

10 mM EDTA; pH 8.0

Autoclave for 15 minutes at 10 lb/sq. in. on liquid cycle. Store at 4°C.

Solution II (BAC DNA extraction)

0.2 N NaOH (freshly diluted from a 10 N stock)

1% (w/v) SDS

Solution III (BAC DNA extraction)

5 M Potassium acetate 60 ml

Glacial acetic acid 11.5 ml

Distilled water 28.5 ml

The resulting solution is 3 M with respect to potassium and 5 M with respect to acetate.

1 x SSCP Denaturing Buffer

0.1 M NaOH in de-ionised formamide

0.125% (w/v) xylene cyanol FF

6 x PAGE/Agarose Loading Buffer

0.25% bromophenol blue

0.25% xylene cyanol FF

40% glycerol in water

1 mM EDTA; pH 8.0

10% Polyacrylamide Gel Preparation

15 ml 40% acrylamide/*bis*-acrylamide mix (19:1)

3 ml 10 x TBE

300 μ l Ammonium persulphate (25% w/v)

Distilled water to 60 ml

5 ml of acrylamide mix was initially polymerized with 30 μ l TEMED to form a gel plug to eliminate gel leakage. The remaining acrylamide mix was polymerized using 50 μ l TEMED and left to set for 1 hour at room temperature.

10% SSCP Polyacrylamide gel preparation

6 ml 10 X TBE

20 ml 30% acrylamide/*bis*-acrylamide mix (37.5:1)

160 μ l Ammonium persulphate (25% w/v)

Distilled water to 60 ml

5 ml of acrylamide mix was initially polymerized with 30 ml TEMED to form a gel plug to eliminate gel leakage. The remaining acrylamide mix was polymerized using 50 ml TEMED and left to set for 1 hour at room temperature.

**APPENDIX B : CHROMOSOME 9p22-p23 YAC, BAC AND EST
MAPPING DATA.**

Table B1. Whitehead Institute database order and YAC contig location of microsatellites loci within the region D9S269 – D9S1778 on chromosome 9p22-p23. The first column displays the position of each STS (sequence tagged site) within the WI database starting from STS#41 (AFM275YB9) and finishing at STS#73 (D9S1684). Other STS's previous to #41 and after #73 have been edited from the table. The second column displays the STS name within the WI database (note different databases identify STS's under different pseudonym's). The third column displays the map distance between STS's in both Centimorgans (cM) and Centirays (cR). The fourth column displays which contig the STS is mapped to according to the WI database. WC9.1 represents the main WI YAC contig within the chromosome 9p22-p23 region investigated during DM1 and DM2. Microsatellites loci used in DM1 and DM2 are highlighted in bold red text. CHLC.GATA27A11 represents the microsatellite loci D9S925. D9S254 is not represented within the WI database. The WI database was originally queried in December 1999. Subsequent queries (most recently January 2001) revealed no change in marker order.

Table B1. Whitehead Institute database order and YAC contig location of microsatellites loci within the region D9S269 – D9S1778 on chromosome 9p22-p23.

Position #	STS Name	Map Position		Contig	
39	AFM275YB9	-	41 cR	WC9.1	WC-264
40	CHLC.GATA84D07	-	-	WC9.1	WC-264
41	CHLC.GATA21A06	-	46 cR	WC9.1	WC-264
42	D9S168	20 cM	46 cR	WC9.1	WC-264
43	D9S256	23 cM	-	WC9.1	WC-264
44	D9S269	23 cM	-	WC9.1	WC-264
45	AFMB292XG5	-	95 cR	WC9.1	WC-264
46	WI-9748	-	-	WC9.1	-
47	WI-4388	-	51 cR	WC9.1	-
48	D9S267	25 cM	-	WC9.1	WC-194
49	WI-7091	-	55 cR	WC9.1	WC-194
50	D9S268	25 cM	-	WC9.1	WC-194
51	CHLC.GATA50G08	-	-	WC9.1	WC-194
52	WI-2432	-	-	WC9.1	WC-194
53	WI-6883	-	-	WC9.1	WC-194
54	D9S1869	27 cM	-	WC9.1	WC-194
55	D9S274	28 cM	57 cR	WC9.1	WC-194
56	D9S285	28 cM	-	WC9.1	-
57	GCT16G03	-	60 cR	WC9.1	WC-882
58	D9S156	30 cM	60 cR	WC9.1	WC-882
59	D9S1782	30 cM	-	WC9.1	WC-882
60	WI-10873	-	60 cR	WC9.1	WC-882
61	D9S1839	30 cM	-	WC9.1	WC-882
62	WI-6705	-	60 cR	WC9.1	WC-1376
63	CHLC.GAAT2F06	-	60 cR	WC9.1	WC-1376
64	WI-3899	-	64 cR	WC9.1	-
65	WI-8411	-	62 cR	WC9.1	WC-1036
66	WI-10810	-	62 cR	WC9.1	WC-1036
67	IB2409	-	-	WC9.1	WC-1036
68	D9S157	32 cM	62 cR	WC9.1	WC-1036
69	CHLC.GATA27A11	-	62 cR	WC9.1	WC-1036
70	WI-2408	-	-	WC9.1	WC-1036
71	WI-1403	-	62 cR	WC9.1	WC-959
72	D9S162	33 cM	62 cR	WC9.1	WC-959
73	D9S1684	33 cM	-	WC9.1	WC-959

Table B2. Table displaying Whitehead Institute database YAC hits for microsatellite loci used in DM1 and DM2. Microsatellite loci are displayed in the first column. YAC hits are displayed in the second column which is divided into four sub-columns. Unambiguous hits denote a "Definite" or confirmed YAC/STS hit. Ambiguous and disambiguous YAC/STS hits were resolved using CEPH fingerprint data or using STS content data and were not confirmed. YAC hits in the 'Other' column were reported by other labs, primarily CEPH. Unconfirmed YAC addresses are displayed within brackets. YAC hits containing adjacent microsatellites are displayed in bold red font. The loci D9S254 is not represented within the WI database. Data was correct at the time of January 2001.

Table B2. Table displaying WI database YAC hits for microsatellite loci used for DM1 and DM2.

Marker Name	YAC hits			
	Unambiguous	Disambiguous	Ambiguous	Other
D9S269	728-B9, 878-H3 927-B7, 961-E4	738-F8, 794-F11, 804-B2, 810-C7 881-B1, 881-C12 881-F9, 953-E5 967-G7		674-B4, 679-H1 911-D1
D9S254	-	-		-
D9S274	762-D7, 806-F7 912-E9	938-F3, 954-E3 963-F12, 965-B1		924-D12, 924-G8 952-G4 , 956-A5 (765,767)-A(7,9)
D9S285	710-G6	767-F2, 952-G4		
D9S156	830-F3	952-G4		676-A8, 766-A2 954-D10
D9S157	928-E7	877-G12, 960-C7		628-F12, 723-G1 751-C9, 763-F1 884-H10, 922-H6 937-C8, 968-E9
D9S925	748-D1, 751-C9 791-H3, 884-H10 946-E1, 968-E9	804-F5, 822-F4 936-C8		-
D9S1684	711-B5 , 720-H9 776-F2 , 795-B10 936-A3 , 951-G3	755-F7 , 804-B9 877-E3		-
D9S162	711-B5 , 755-F7 776-F2 , 795-B10 804-B9 , 877-E3 936-A3 , 951-G3 965-F8	-		924-E7
D9S1778	-	-	(765-772)_G_3, (781,782)_(A,C)_(9,12) (789-796)_F_10, (818,820)_E_9 (925-932)_C_7	

Table B3. Table displaying CEPH-Genethon database YAC hits for microsatellite loci used for DM1 and DM2. Microsatellite loci are displayed in the first column. YAC hits are displayed in the second column. D9S254, D9S925, D9S1684 and D9S1778 are not represented in the CEPH-Genethon database. YAC hits containing adjacent microsatellites are displayed in bold red font. Data was correct at the time of January 2001.

Table B3. Table displaying CEPH-Genethon database YAC hits for microsatellite loci used for DM1 and DM2.

Marker Name	CEPH-Genethon YAC hits
D9S269	674_b_4, 679_h_1 728_b_9 738_f_8 794_f_11 804_b_2 810_c_7 873_h_3 881_b_1 881_c_12 881_f_9 911_d_1 927_b_7 953_e_5 961_e_4
D9S254	-
D9S274	762_d_7 806_f_7 912_e_9 924_d_12 924_g_8 952_g_4 956_a_5 963_f_1 963_f_12 965_b_1
D9S285	710_g_6 767_f_2 952_g_4
D9S156	676_a_8 766_a_2 830_f_3 952_g_4 954_d_10
D9S157	628_f_12 723_g_1 751_c_9 763_f_1

Marker Name	CEPH-Genethon YAC hits
D9S157 (cont.)	877_g_12
	884_h_10
	922_h_6
	928_e_7
	937_c_8
	960_c_7
	968_e_9
D9S925	-
D9S1684	-
D9S162	755_f_7
	776_f_2
	795_b_10
	804_b_9
	877_e_3
	924_e_7
	936_a_3
	951_g_3
	965_f_8
D9S1778	-

Table B4. Table displaying a list of STS's and EST's within the region of D9S157 – D9S1778 according to Genemap '99 database. The first column displays the category of STS or EST's detected within the region. The second column displays the ID or name of the STS or EST within the Genemap '99 database. The third column displays the generic name of the STS (for example the name of the gene the STS corresponds to). Note that several STS's have been incorrectly positioned within this region.

Table B4. Table displaying a list of STS's within the region of D9S157 – D9S1778 according to Genemap '99 database.

Category	Genemap ID	Other Name
Microsatellite	D9S157	repeat sequence
EST	WI-8411	
EST's	A005O33	
EST's	stSG2674	
EST's	A003M39	
Gene	stSG3237	SH3-domain GRB2-like 2 (SH3GL2)
EST	stSG62636	
EST	stSG29526	
EST's	stSG49361	
EST's	stSG51955	
Gene	stSG54862	CDKN2A cyclin-dependent kinase inhibitor 2A (p16)
EST	stSG49361	
EST	H78384	
Gene	stSG8052	RPS2 ribosomal protein S2
EST	SGC30115	
Gene	WIAF-870	interferon, alpha 21 (IFNA21)
EST	stSG47325	
EST	stSG26364	
Gene	Cda18h01	Human adenovirus protein E3-14.7k interacting protein 1 (FIP-1) mRNA
EST's	H99933	
EST	stSG39878	
Gene	stSG2267	insulin-like growth factor binding protein 6 (IGFBP6)
Gene	SGC35277	interferon, alpha 16 (IFNA16)
Gene	SGC35616	Messenger RNA for human leukocyte (alpha) interferon
EST	H13545	
Gene	M27318	Human interferon (IFN-alpha-M1) mRNA, complete cds
EST	stSG28886	
EST's	WI-13834	
Gene	SGC35314	Human interferon-alpha type I' mRNA, complete cds
Gene	stSG45281	methylthioadenosine phosphorylase (MTAP)
Unknown	stSG26126	Human interferon genes LeIF-L and LeIF-J and pseudogene LeIF-M with intergenic regions
Gene	stSG26994	ribosomal protein S6 (RPS6)

Category	Genemap ID	Other Name
EST'S	WI-12519	
Gene	V00531	interferon, alpha 14 (IFNA14)
EST	sts-AA018603	
Gene	stSG43	myeloid/lymphoid or mixed-lineage leukemia (trithorax (Drosophila) homolog); translocated to, 3 (MLLT3)
EST's	stSG50242	
EST's	stSG54918	
EST	stSG43960	
EST's	stSG26416	
Gene	sts-W55904	adipose differentiation-related protein; adipophilin (ADFP)
EST's	AA251418	
EST's	R72278	
Unknown	sts-J00210	Human leukocyte interferon (IFN-alpha) alpha-d gene
EST	sts-T95022	
Unknown	sts-X02958	
Microsatellite	D9S162	repeat sequence
Gene	SGC35639	interferon, alpha 5 (IFNA5)
Gene	SHGC-35327	interferon, alpha 2 (IFNA2)
Gene	SGC35275	interferon, alpha 14 (IFNA14)
Gene	stSG10352	Homo sapiens phospholipase A2-activating protein mRNA

**APPENDIX C : BLAST ANALYSIS AND SEQUENCE ALIGNMENT
OF MLLT3 GENE.**

C1 MLLT3 BLAST search results

>ref|NM_004529.1| Homo sapiens myeloid/lymphoid or mixed-lineage leukemia (trithorax
(Drosophila) homolog); translocated to, 3 (MLLT3), mRNA
Length = 3376

Score = 2432 bits (1227), Expect = 0.0
Identities = 1297/1332 (97%)
Strand = Plus / Plus

Query: 2045 caaccttcagctttatTTTTCTTTAAAGCCAGTCATCATCTCTTGATAAAGGAGAGGTTA 2104
|||||
Sbjct: 2045 caaccttcagctttatTTTTCTTTAAAGCCAGTCATCATCTCTTGATAAAGGAGAGGTTA 2104

Query: 2105 aagcaaaccagcctcagcggaccactcttctctccaaggaaatccccggaagagtttagc 2164
|||||
Sbjct: 2105 aagcaaaccagcctcagcggaccactcttctctccaaggaaatccccggaagagtttagc 2164

Query: 2165 ctggatagccttgaaaacaaacaaatcaaacacacacaagaaaactcaaagaatgtgta 2224
|||||
Sbjct: 2165 ctggatagccttgaaaacaaacaaatcaaacacacacaagaaaactcaaagaatgtgta 2224

Query: 2225 tggatcatgtatctctctgtggtggttcattccacaggacgaatgcatattcaacacac 2284
|||||
Sbjct: 2225 tggatcatgtatctctctgtggtggttcattccacaggacgaatgcatattcaacacac 2284

Query: 2285 tgccttattacataactgatctatTTTATTATCGCATACAGATATTCTAAGTCGTTGAGGG 2344
|||||
Sbjct: 2285 tgccttattacataactgatctatTTTATTATCGCATACAGATATTCTAAGTCGTTGAGGG 2344

Query: 2345 aatgacaccatcagacattataagtacttggtcccgtggatgctctttcaatgcagcacc 2404

|||||
Sbjct: 2345 aatgacacccatcagacattataagtacttgggcccggtggatgctctttcaatgcagcacc 2404

Query: 2405 cttgccatcccaagcccagtgaccttactcgtataccgtgccactttccaccaacttttt 2464
|||||
Sbjct: 2405 cttgccatcccaagcccagtgaccttactcgtataccgtgccactttccaccaacttttt 2464

Query: 2465 ccaagtcctttaactcgttgcagctctgtatttccaccttttgttttccagttccagga 2524
|||||
Sbjct: 2465 ccaagtcctttaactcgttgcagctctgtatttccaccttttgttttccagttccagga 2524

Query: 2525 cacagattatcaactggggggaccaaatagccaccttgatthttcttctttgtggcttttt 2584
|||||
Sbjct: 2525 cacagattatcaactggggggaccaaatagccaccttgatthttcttctttgtggcttttt 2584

Query: 2585 tcttgaaagtggggcccagtccttggctgtatccatgtaatgatcttgaccatggtag 2644
|||||
Sbjct: 2585 tcttgaaagtggggcccagtccttggctgtatccatgtaatgatcttgaccatggtag 2644

Query: 2645 aaaatgcaccaaataggatcatatgaattgctgtctagccttagtcaataaacttgtagg 2704
|||||
Sbjct: 2645 aaaatgcaccaaataggatcatatgaattgctgtctagccttagtcaataaacttgtagg 2704

Query: 2705 acttttaaacaaaagtgtacctgtaaatgtcctgaatccagcattggtgagctgtcatca 2764
|||||
Sbjct: 2705 acttttaaacaaaagtgtacctgtaaatgtcctgaatccagcattggtgagctgtcatca 2764

Query: 2765 acattcttgtgtctgttttactgttacaatataggtgaatatggaagtaaggcattcc 2824
|||||

Sbjct: 2765 acattcttgtgtctgttttactgttacaatattaggtgaatatggaagtaaaggcattcc 2824

Query: 2825 acaggatcatcatttnnnnnnnngaattctggcctgttttctnnnnnnnnnnctgt 2884

|||||

Sbjct: 2825 acaggatcatcatttaaaaaaaaaagaattctggcctgttttctaaaaaaaaaactgt 2884

Query: 2885 tgtagaaattcttaatttggatctatatttagtcagagtttcagctttcttcagctgcc 2944

|||||

Sbjct: 2885 tgtagaaattcttaatttggatctatatttagtcagagtttcagctttcttcagctgcc 2944

Query: 2945 agtgtgttactcatctttatcctaaaaatctggaatcagagatttttgttgttcacata 3004

|||||

Sbjct: 2945 agtgtgttactcatctttatcctaaaaatctggaatcagagatttttgttgttcacata 3004

Query: 3005 tgattctcttagacacttttatatttgaaaaaattaaaatctttctttggggaaaaattc 3064

|||||

Sbjct: 3005 tgattctcttagacacttttatatttgaaaaaattaaaatctttctttggggaaaaattc 3064

Query: 3065 ttggttattctgccataacagattatgtattaactttagattcagtggttcaataacctg 3124

|||||

Sbjct: 3065 ttggttattctgccataacagattatgtattaactttagattcagtggttcaataacctg 3124

Query: 3125 tttagttgcttgctaataattccagaaggatttcttgtattggtgaaagacggttgggga 3184

|||||

Sbjct: 3125 tttagttgcttgctaataattccagaaggatttcttgtattggtgaaagacggttgggga 3184

Query: 3185 tggggggannnnnngttcttgttgtacccttgttttgaaactagaaatctgtcctgtgg 3244

|||||

Sbjct: 3185 tggggggatttttttgttcttggtgtacccttgttttgaaactagaaatctgtcctgtgg 3244

Query: 3245 catgcaaaagaaagcaaattatTTTTTAAAAGNNNNNNNCCAAAGTACTTTTGGTGTcatt 3304
|||||

Sbjct: 3245 catgcaaaagaaagcaaattatTTTTTAAAAGAAAAAACCAAGTACTTTTGGTGTcatt 3304

Query: 3305 attccatcttctccataagtggagaaatgaaaagtaagaacagctcatcttcaaagtttt 3364
|||||

Sbjct: 3305 attccatcttctccataagtggagaaatgaaaagtaagaacagctcatcttcaaagtttt 3364

Query: 3365 tactagaaattc 3376
|||||

Sbjct: 3365 tactagaaattc 3376

Score = 2290 bits (1155), Expect = 0.0
Identities = 1169/1176 (99%)
Strand = Plus / Plus

Query: 762 tagtaccagtttttcaaagcctcacaattaatgaaggagcacaaggaaaaaccttctaa 821
|||||

Sbjct: 762 tagtaccagtttttcaaagcctcacaattaatgaaggagcacaaggaaaaaccttctaa 821

Query: 822 agactccagagaacataaaaagtgcttcaaagaaccttccagggatcacaacaaatcttc 881
|||||

Sbjct: 822 agactccagagaacataaaaagtgcttcaaagaaccttccagggatcacaacaaatcttc 881

Query: 882 caaagaatcctctaagaaacccaaagaaaataaaccactgaaagaagagaaaatagttcc 941
|||||

Sbjct: 882 caaagaatcctctaagaaacccaaagaaaataaaccactgaaagaagagaaaatagttcc 941

Query: 942 taagatggccttcaaggaacctaaccatgtcaaaagagccaaaaccagatagtaactt 1001

|||||

Sbjct: 942 taagatggccttcaaggaacctaaccatgtcaaaagagccaaaaccagatagtaactt 1001

Query: 1002 actcaccatcaccagtggacaagataagaaggctcctagtaaaaggccgccatttcaga 1061

|||||

Sbjct: 1002 actcaccatcaccagtggacaagataagaaggctcctagtaaaaggccgccatttcaga 1061

Query: 1062 ttctgaagaactctcagccnnnnnnnggaaaaagagtagctcagaggctttattttaaag 1121

|||||

Sbjct: 1062 ttctgaagaactctcagccaaaaaaaggaaaaagagtagctcagaggctttattttaaag 1121

Query: 1122 tttttctagcgcaccaccactgataactcacttgttctgctgacaaaaaacagataaaaga 1181

|||||

Sbjct: 1122 tttttctagcgcaccaccactgataactcacttgttctgctgacaaaaaacagataaaaga 1181

Query: 1182 taaatctcatgtcaagatgggaaaggtaaaaattgaaagtgagacatcagagaagaagaa 1241

|||||

Sbjct: 1182 taaatctcatgtcaagatgggaaaggtaaaaattgaaagtgagacatcagagaagaagaa 1241

Query: 1242 atcaacgttaccgccatttgatgatattgtggatcccaatgattcagatgtggaggagaa 1301

|||||

Sbjct: 1242 atcaacgttaccgccatttgatgatattgtggatcccaatgattcagatgtggaggagaa 1301

Query: 1302 tatatcctctaaatctgattctgaacaaccagtcctgccagctccagctccagctccag 1361

|||||

Sbjct: 1302 tatatcctctaaatctgattctgaacaacccagtcctgccagctccagctccagctccag 1361

Query: 1362 ctccagcttcacaccatcccagaccaggcaacaaggtcctttgaggtctataatgaaaga 1421

|||||

Sbjct: 1362 ctccagcttcacaccatcccagaccaggcaacaaggtcctttgaggtctataatgaaaga 1421

Query: 1422 tctgcattctgatgacaatgaggaggaatcagatgaagtggaggataacgacaatgactc 1481

|||||

Sbjct: 1422 tctgcattctgatgacaatgaggaggaatcagatgaagtggaggataacgacaatgactc 1481

Query: 1482 tgaaatggagaggcctgtaaatagaggaggcagccgaagtcgcagagttagcttaagtga 1541

|||||

Sbjct: 1482 tgaaatggagaggcctgtaaatagaggaggcagccgaagtcgcagagttagcttaagtga 1541

Query: 1542 tggcagcgatagtgaaagcagttctgcttcttcaccctacatcacgaacctccaccacc 1601

|||||

Sbjct: 1542 tggcagcgatagtgaaagcagttctgcttcttcaccctacatcacgaacctccaccacc 1601

Query: 1602 cttactaaaaaccaacaacaaccagattcttgaagtgaaaagtccaataaagcaaagcaa 1661

|||||

Sbjct: 1602 cttactaaaaaccaacaacaaccagattcttgaagtgaaaagtccaataaagcaaagcaa 1661

Query: 1662 atcagataagcaaataaagaatggtgaatgtgacaaggcatacctagatgaactggtaga 1721

|||||

Sbjct: 1662 atcagataagcaaataaagaatggtgaatgtgacaaggcatacctagatgaactggtaga 1721

Query: 1722 gcttcacagaaggttaatgacattgagagaaagacacattctgcagcagatcgtgaacct 1781

|||||

Sbjct: 1722 gcttcacagaaggttaatgacattgagagaaagacacattctgcagcagatcgtgaacct 1781

Query: 1782 tatagaagaaactggacactttcatatcaciaaacacacatttgattttgatctttgctc 1841

|||||

Sbjct: 1782 tatagaagaaactggacactttcatatcaciaaacacacatttgattttgatctttgctc 1841

Query: 1842 gctggacaaaaccacagtccgtaaactacagagttacctggaaacatctggaacatcctg 1901

|||||

Sbjct: 1842 gctggacaaaaccacagtccgtaaactacagagttacctggaaacatctggaacatcctg 1901

Query: 1902 aggatataacaactggatgcatcaagaactattgtg 1937

|||||

Sbjct: 1902 aggatataacaactggatgcatcaagaactattgtg 1937

Score = 955 bits (482), Expect = 0.0
Identities = 482/482 (100%)
Strand = Plus / Plus

Query: 157 taccacacaacaccccagcaaacctccgggccccaggcatggctagctcgtgttccgtg 216

|||||

Sbjct: 157 taccacacaacaccccagcaaacctccgggccccaggcatggctagctcgtgttccgtg 216

Query: 217 caggtgaagctggagctggggcaccgcgccaggtgaggaaaaaccaccgtggagggc 276

|||||

Sbjct: 217 caggtgaagctggagctggggcaccgcgccaggtgaggaaaaaccaccgtggagggc 276

Query: 277 ttcaccacgactggatggtgttcgtacgcggtccggagcacagtaacatacagcacttt 336

|||||

Sbjct: 277 ttcacccacgactggatggtggttcgtacgcgggtccggagcacagtaacatacagcacttt 336

Query: 337 gtggagaaagtcgtcttccacttgcacgaaagctttcctaggccaaaaagagtgtgcaa 396

|||||

Sbjct: 337 gtggagaaagtcgtcttccacttgcacgaaagctttcctaggccaaaaagagtgtgcaa 396

Query: 397 gatccaccttaciaaagtagaagaatctgggtatgctggtttcattttgccaattgaagtt 456

|||||

Sbjct: 397 gatccaccttaciaaagtagaagaatctgggtatgctggtttcattttgccaattgaagtt 456

Query: 457 tattttaaaaaacaaggaagaacctaggaaagtcgctttgattatgacttattcctgcat 516

|||||

Sbjct: 457 tattttaaaaaacaaggaagaacctaggaaagtcgctttgattatgacttattcctgcat 516

Query: 517 cttgaaggccatccaccagtgaatcacctccgctgtgaaaagctaactttcaacaacccc 576

|||||

Sbjct: 517 cttgaaggccatccaccagtgaatcacctccgctgtgaaaagctaactttcaacaacccc 576

Query: 577 acagaggactttaggagaaagttgctgaaggcaggagggaccctaataaggagtattcat 636

|||||

Sbjct: 577 acagaggactttaggagaaagttgctgaaggcaggagggaccctaataaggagtattcat 636

Query: 637 ac 638

||

Sbjct: 637 ac 638

Score = 264 bits (133), Expect = 3e-67

Identities = 133/133 (100%)
Strand = Plus / Plus

Query: 1 tttggggctgagtttaataagcgagcgagcgagcaagcgagcgcggggggaaaaaggcag 60
|||||
Sbjct: 1 tttggggctgagtttaataagcgagcgagcgagcaagcgagcgcggggggaaaaaggcag 60

Query: 61 agaatgtccgccatctaccctccgctcctgggcgcgctctcattcatagcagcctcttca 120
|||||
Sbjct: 61 agaatgtccgccatctaccctccgctcctgggcgcgctctcattcatagcagcctcttca 120

Query: 121 tgaattacagctg 133
|||||
Sbjct: 121 tgaattacagctg 133

Score = 61.9 bits (31), Expect = 2e-06
Identities = 31/31 (100%)
Strand = Plus / Plus

Query: 1966 ggttgtgattttttgttcttggtggtttatat 1996
|||||
Sbjct: 1966 ggttgtgattttttgttcttggtggtttatat 1996

>gb|L13744.1|HUMAF9X Human AF-9 mRNA, complete cds
Length = 3376

Score = 2432 bits (1227), Expect = 0.0
Identities = 1297/1332 (97%)
Strand = Plus / Plus

Query: 2045 caaccttcagctttatTTTTctTTTaaagccagtcacatctcttgataaaggagaggtta 2104
|||||
Sbjct: 2045 caaccttcagctttatTTTTctTTTaaagccagtcacatctcttgataaaggagaggtta 2104

Query: 2105 aagcaaaccagcctcagcggaccactcttctctccaaggaaatccccggaagaggttagc 2164
|||||
Sbjct: 2105 aagcaaaccagcctcagcggaccactcttctctccaaggaaatccccggaagaggttagc 2164

Query: 2165 ctggatagccttgaaaacaaacaaatcaaacacacacaagaaaactcaaagaatgtgta 2224
|||||
Sbjct: 2165 ctggatagccttgaaaacaaacaaatcaaacacacacaagaaaactcaaagaatgtgta 2224

Query: 2225 tggatcatgtatctctctgtggtggttcattccacaggacgaatgcatattcaacacac 2284
|||||
Sbjct: 2225 tggatcatgtatctctctgtggtggttcattccacaggacgaatgcatattcaacacac 2284

Query: 2285 tgccttattacataactgatctatttattatcgcatcacagatattctaagtcggttgaggg 2344
|||||
Sbjct: 2285 tgccttattacataactgatctatttattatcgcatcacagatattctaagtcggttgaggg 2344

Query: 2345 aatgacaccatcagacattataagtaacttgggtcccgtggatgctctttcaatgcagcacc 2404
|||||
Sbjct: 2345 aatgacaccatcagacattataagtaacttgggtcccgtggatgctctttcaatgcagcacc 2404

Query: 2405 ottgccatcccaagcccagtgaccttactcgtataaccgtgccactttccaccaacttttt 2464
|||||
Sbjct: 2405 ottgccatcccaagcccagtgaccttactcgtataaccgtgccactttccaccaacttttt 2464

Query: 2465 ccaagtcctttaactcgttgcagtctgtat tttccacctttgttttccagttccagga 2524
 |||
 Sbjct: 2465 ccaagtcctttaactcgttgcagtctgtat tttccacctttgttttccagttccagga 2524

Query: 2525 cacagattatcaactggggggaccaa atagccaccttgattttcttctttgtggctttt 2584
 |||
 Sbjct: 2525 cacagattatcaactggggggaccaa atagccaccttgattttcttctttgtggctttt 2584

Query: 2585 tcctgaaagttggggcccagtccttggctgtatccatgta atgatcttggaccatggtag 2644
 |||
 Sbjct: 2585 tcctgaaagttggggcccagtccttggctgtatccatgta atgatcttggaccatggtag 2644

Query: 2645 aaaatgcaccaa ataggatcatatgaattgctgtctagccttagtcaataaacttgtagg 2704
 |||
 Sbjct: 2645 aaaatgcaccaa ataggatcatatgaattgctgtctagccttagtcaataaacttgtagg 2704

Query: 2705 acttttaacaaa agtgtacctgtaaatgtcctgaatccagcattgttgagctgtcatca 2764
 |||
 Sbjct: 2705 acttttaacaaa agtgtacctgtaaatgtcctgaatccagcattgttgagctgtcatca 2764

Query: 2765 acattcttgtgtctgttttactg ttacaatattaggtgaat atggaagtaaaggcattcc 2824
 |||
 Sbjct: 2765 acattcttgtgtctgttttactg ttacaatattaggtgaat atggaagtaaaggcattcc 2824

Query: 2825 acaggatcatcatttnnnnnnnnngaattctggtcctg ttttctnnnnnnnnnnnctgt 2884
 ||| ||| |||
 Sbjct: 2825 acaggatcatcatttaaaaaaa agaattctggtcctg ttttctaaaaaaaaaactgt 2884

Query: 2885 tgtagaaattcttaatttggatctatttattagtcagagtttcagctttcttcagctgcc 2944
 |||

Sbjct: 2885 tgtagaaattcttaatttggatctatatttagtcagagtttcagctttcttcagctgcc 2944

Query: 2945 agtgtgttactcatctttatcctaaaaatctggaatcagagatttttgttgttcacata 3004
 |||

Sbjct: 2945 agtgtgttactcatctttatcctaaaaatctggaatcagagatttttgttgttcacata 3004

Query: 3005 tgattctcttagacacttttatatttgaaaaaattaaaatctttctttggggaaaaattc 3064
 |||

Sbjct: 3005 tgattctcttagacacttttatatttgaaaaaattaaaatctttctttggggaaaaattc 3064

Query: 3065 ttggttattctgccataacagattatgtattaacttgtagattcagtggttcaatacctg 3124
 |||

Sbjct: 3065 ttggttattctgccataacagattatgtattaacttgtagattcagtggttcaatacctg 3124

Query: 3125 tttagttgcttgctaataattccagaaggatttcttgattggtgaaagacggttgggga 3184
 |||

Sbjct: 3125 tttagttgcttgctaataattccagaaggatttcttgattggtgaaagacggttgggga 3184

Query: 3185 tggggggannnnnnngttcttggtgtacccttgtttgaaactagaaatctgtcctgtgg 3244
 |||

Sbjct: 3185 tggggggattttttgttcttggtgtacccttgtttgaaactagaaatctgtcctgtgg 3244

Query: 3245 catgcaaaagaaagcaaattatttttaaagnnnnnnccaaagtacttttggtgtcatt 3304
 |||

Sbjct: 3245 catgcaaaagaaagcaaattatttttaaagaaaaaaaccaaagtacttttggtgtcatt 3304

Query: 3305 attccatcttctccataagtggagaaatgaaaagtaagaacagctcatcttcaaagttt 3364
 |||

Sbjct: 3305 attccatcttctccataagtggagaaatgaaaagtaagaacagctcatcttcaaagtttt 3364

Query: 3365 tactagaaattc 3376

|||||

Sbjct: 3365 tactagaaattc 3376

Score = 2290 bits (1155), Expect = 0.0

Identities = 1169/1176 (99%)

Strand = Plus / Plus

Query: 762 tagtaccagtttttcaaagcctcacaattaatgaaggagcacaaggaaaaaccttctaa 821

|||||

Sbjct: 762 tagtaccagtttttcaaagcctcacaattaatgaaggagcacaaggaaaaaccttctaa 821

Query: 822 agactccagagaacataaaaagtgccttcaaagaaccttccagggatcacaacaaatcttc 881

|||||

Sbjct: 822 agactccagagaacataaaaagtgccttcaaagaaccttccagggatcacaacaaatcttc 881

Query: 882 caaagaatcctctaagaaacccaaagaaaataaaccactgaaagaagagaaaatagttcc 941

|||||

Sbjct: 882 caaagaatcctctaagaaacccaaagaaaataaaccactgaaagaagagaaaatagttcc 941

Query: 942 taagatggccttcaaggaacctaaccatgtcaaaagagccaaaaccagatagtaactt 1001

|||||

Sbjct: 942 taagatggccttcaaggaacctaaccatgtcaaaagagccaaaaccagatagtaactt 1001

Query: 1002 actcaccatcaccagtggaacaagataagaaggctcctagtaaaaaggccgccatttcaga 1061

|||||

Sbjct: 1002 actcaccatcaccagtggacaagataagaaggctcctagtaaaaggccgccatttcaga 1061

Query: 1062 ttctgaagaactctcagccnnnnnnnggaaaaagagtagctcagaggctttattttaaag 1121
 |||
 Sbjct: 1062 ttctgaagaactctcagccaaaaaaggaaaaagagtagctcagaggctttattttaaag 1121

Query: 1122 tttttctagcgcaccaccactgatactcacttggtctgctgacaaaaaacagataaaaga 1181
 |||
 Sbjct: 1122 tttttctagcgcaccaccactgatactcacttggtctgctgacaaaaaacagataaaaga 1181

Query: 1182 taaatctcatgtcaagatgggaaaggctaaaattgaaagtgagacatcagagaagaagaa 1241
 |||
 Sbjct: 1182 taaatctcatgtcaagatgggaaaggctaaaattgaaagtgagacatcagagaagaagaa 1241

Query: 1242 atcaacgttaccgccatttgatgatattgtggatcccaatgattcagatgtggaggagaa 1301
 |||
 Sbjct: 1242 atcaacgttaccgccatttgatgatattgtggatcccaatgattcagatgtggaggagaa 1301

Query: 1302 tatatcctctaaatctgattctgaacaaccagtcctgccagctccagctccag 1361
 |||
 Sbjct: 1302 tatatcctctaaatctgattctgaacaaccagtcctgccagctccagctccag 1361

Query: 1362 ctccagcttcacaccatcccagaccaggcaacaaggctccttgaggctataatgaaaga 1421
 |||
 Sbjct: 1362 ctccagcttcacaccatcccagaccaggcaacaaggctccttgaggctataatgaaaga 1421

Query: 1422 tctgcattctgatgacaatgaggaggaatcagatgaagtggaggataacgacaatgactc 1481
 |||

Sbjct: 1422 totgcattctgatgacaatgaggaggaatcagatgaagtggaggataacgacaatgactc 1481

Query: 1482 tgaaatggagaggcctgtaaatagaggaggcagccgaagtcgcagagttagcttaagtga 1541

|||||

Sbjct: 1482 tgaaatggagaggcctgtaaatagaggaggcagccgaagtcgcagagttagcttaagtga 1541

Query: 1542 tggcagcgatagtgaaagcagttctgcttcttcaccctacatcacgaacctccaccacc 1601

|||||

Sbjct: 1542 tggcagcgatagtgaaagcagttctgcttcttcaccctacatcacgaacctccaccacc 1601

Query: 1602 cttactaaaaaccaacaacaaccagattcttgaagtgaaaagtccaataaagcaaagcaa 1661

|||||

Sbjct: 1602 cttactaaaaaccaacaacaaccagattcttgaagtgaaaagtccaataaagcaaagcaa 1661

Query: 1662 atcagataagcaaataaagaatggtgaatgtgacaaggcatacctagatgaactggtaga 1721

|||||

Sbjct: 1662 atcagataagcaaataaagaatggtgaatgtgacaaggcatacctagatgaactggtaga 1721

Query: 1722 gcttcacagaaggttaatgacattgagagaaagacacattctgcagcagatcgtgaacct 1781

|||||

Sbjct: 1722 gcttcacagaaggttaatgacattgagagaaagacacattctgcagcagatcgtgaacct 1781

Query: 1782 tatagaagaaactggacactttcatatcaciaaacacaacatttgattttgatctttgctc 1841

|||||

Sbjct: 1782 tatagaagaaactggacactttcatatcaciaaacacaacatttgattttgatctttgctc 1841

Query: 1842 gctggacaaaaccacagtcggtaaactacagagttacctggaaacatctggaacatcctg 1901

|||||

Sbjct: 1842 gctggacaaaaccacagtcctgtaaactacagagttacctggaacatctggaacatcctg 1901

Query: 1902 aggatataacaactggatgcatcaagaactattgtg 1937

|||||

Sbjct: 1902 aggatataacaactggatgcatcaagaactattgtg 1937

Score = 955 bits (482), Expect = 0.0

Identities = 482/482 (100%)

Strand = Plus / Plus

Query: 157 taccacacaacaccccagcaaacctccgggccccagggcatggctagctcgtggtccgtg 216

|||||

Sbjct: 157 taccacacaacaccccagcaaacctccgggccccagggcatggctagctcgtggtccgtg 216

Query: 217 caggtgaagctggagctggggcaccgcgccaggtgaggaaaaaacccacgtggagggc 276

|||||

Sbjct: 217 caggtgaagctggagctggggcaccgcgccaggtgaggaaaaaacccacgtggagggc 276

Query: 277 ttcaccacgactggatggtggttcgtacgcgggtccggagcacagtaacatacagcacttt 336

|||||

Sbjct: 277 ttcaccacgactggatggtggttcgtacgcgggtccggagcacagtaacatacagcacttt 336

Query: 337 gtggagaaagtctcttccactgacgaaagctttcctaggccaaaaagagtgtgcaa 396

|||||

Sbjct: 337 gtggagaaagtctcttccactgacgaaagctttcctaggccaaaaagagtgtgcaa 396

Query: 397 gatccaccttacaagtagaagaatctgggtatgctggtttcattttgccattgaagtt 456

|||||

Sbjct: 397 gatccaccttacaagtagaagaatctgggtatgctggtttcattttgccaattgaagtt 456

Query: 457 tattttaaaaacaaggaagaacctaggaaagtcgcctttgattatgacttattcctgcat 516
|||||

Sbjct: 457 tattttaaaaacaaggaagaacctaggaaagtcgcctttgattatgacttattcctgcat 516

Query: 517 cttgaaggccatccaccagtgaatcacctccgctgtgaaaagctaactttcaacaacccc 576
|||||

Sbjct: 517 cttgaaggccatccaccagtgaatcacctccgctgtgaaaagctaactttcaacaacccc 576

Query: 577 acagaggactttaggagaaagttgctgaaggcaggaggggaccctaataaggagtattcat 636
|||||

Sbjct: 577 acagaggactttaggagaaagttgctgaaggcaggaggggaccctaataaggagtattcat 636

Query: 637 ac 638

||

Sbjct: 637 ac 638

Score = 264 bits (133), Expect = 3e-67

Identities = 133/133 (100%)

Strand = Plus / Plus

Query: 1 tttggggctgagtttaataagcgagcgagcgagcaagcgagcgcggggggaaaaaggcag 60

|||||

Sbjct: 1 tttggggctgagtttaataagcgagcgagcgagcaagcgagcgcggggggaaaaaggcag 60

Query: 61 agaatgtccgccatctaccctccgctcctgggcgcgctctcattcatagcagcctcttca 120

|||||

Sbjct: 61 agaatgtccgccatctaccctccgctcctgggcgcgctctcattcatagcagcctcttca 120

Query: 121 tgaattacagctg 133
 |
Sbjct: 121 tgaattacagctg 133

Score = 61.9 bits (31), Expect = 2e-06
Identities = 31/31 (100%)
Strand = Plus / Plus

Query: 1966 ggttgtgattttttggttcttggtgtttatat 1996
 |
Sbjct: 1966 ggttgtgattttttggttcttggtgtttatat 1996

>[ref|XM_005654.1|](#) Homo sapiens myeloid/lymphoid or mixed-lineage leukemia (trithorax
(Drosophila) homolog); translocated to, 3 (MLLT3), mRNA
Length = 3375

Score = 2290 bits (1155), Expect = 0.0
Identities = 1169/1176 (99%)
Strand = Plus / Plus

Query: 762 tagtaccagtttttcaaagcctcacaattaatgaaggagcacaaggaaaaaccttctaa 821
 |
Sbjct: 762 tagtaccagtttttcaaagcctcacaattaatgaaggagcacaaggaaaaaccttctaa 821

Query: 822 agactccagagaacataaaaagtgaccttcaaagaaccttccagggatcacaacaaatcttc 881
 |
Sbjct: 822 agactccagagaacataaaaagtgaccttcaaagaaccttccagggatcacaacaaatcttc 881

Query: 882 caaagaatcctctaagaaacccaaagaaaataaaccactgaaagaagagaaaatagttcc 941
 |||
 Sbjct: 882 caaagaatcctctaagaaacccaaagaaaataaaccactgaaagaagagaaaatagttcc 941

Query: 942 taagatggccttcaaggaacctaaacccatgtcaaaagagccaaaaccagatagtaactt 1001
 |||
 Sbjct: 942 taagatggccttcaaggaacctaaacccatgtcaaaagagccaaaaccagatagtaactt 1001

Query: 1002 actcaccatcaccagtggacaagataagaaggctcctagtaaaaggccgccatttcaga 1061
 |||
 Sbjct: 1002 actcaccatcaccagtggacaagataagaaggctcctagtaaaaggccgccatttcaga 1061

Query: 1062 ttctgaagaactctcagccnnnnnnggaaaaagagtagctcagaggctttattttaaag 1121
 |||
 Sbjct: 1062 ttctgaagaactctcagccaaaaaaaggaaaaagagtagctcagaggctttattttaaag 1121

Query: 1122 tttttctagcgcaccaccactgatactcacttgttctgctgacaaaaaacagataaaaga 1181
 |||
 Sbjct: 1122 tttttctagcgcaccaccactgatactcacttgttctgctgacaaaaaacagataaaaga 1181

Query: 1182 taaatctcatgtcaagatgggaaagggtcaaaattgaaagtgagacatcagagaagaagaa 1241
 |||
 Sbjct: 1182 taaatctcatgtcaagatgggaaagggtcaaaattgaaagtgagacatcagagaagaagaa 1241

Query: 1242 atcaacgttaccgccatttgatgatattgtggatcccaatgattcagatgtggaggagaa 1301
 |||
 Sbjct: 1242 atcaacgttaccgccatttgatgatattgtggatcccaatgattcagatgtggaggagaa 1301

Query: 1302 tatatcctctaaatctgattctgaacaaccagtcctgccagctccagctccag 1361

Sbjct: 1302 |tata|tcc|tct|ctaa|tct|gatt|tct|gaaca|acc|cag|tcct|gcc|cag|ctcc|cag|ctcc|cag| 1361

Query: 1362 |ctc|cag|ctt|cac|acc|atc|cc|cag|acc|agg|caaca|agg|tc|ctt|tg|agg|tct|ata|atg|aa|aga| 1421

Sbjct: 1362 |ctc|cag|ctt|cac|acc|atc|cc|cag|acc|agg|caaca|agg|tc|ctt|tg|agg|tct|ata|atg|aa|aga| 1421

Query: 1422 |tct|gca|ttct|gat|gaca|aatg|agg|agga|atc|agat|gaag|tg|gagg|ata|acg|aca|atg|act|c| 1481

Sbjct: 1422 |tct|gca|ttct|gat|gaca|aatg|agg|agga|atc|agat|gaag|tg|gagg|ata|acg|aca|atg|act|c| 1481

Query: 1482 |tgaa|atg|gag|agg|cct|gt|aaat|ag|agg|agg|cag|ccg|aag|tcg|cag|agt|tag|ctt|aag|tga| 1541

Sbjct: 1482 |tgaa|atg|gag|agg|cct|gt|aaat|ag|agg|agg|cag|ccg|aag|tcg|cag|agt|tag|ctt|aag|tga| 1541

Query: 1542 |tgg|cag|cga|tag|tgaa|agc|agt|tct|gct|tct|t|cacc|cct|acat|cac|ga|ac|ctcc|acc|acc| 1601

Sbjct: 1542 |tgg|cag|cga|tag|tgaa|agc|agt|tct|gct|tct|t|cacc|cct|acat|cac|ga|ac|ctcc|acc|acc| 1601

Query: 1602 |ctt|act|aaaa|acca|aca|aca|acc|cag|att|ctt|ga|agt|gaaa|agt|cca|ata|aa|gcaa|ag|caa| 1661

Sbjct: 1602 |ctt|act|aaaa|acca|aca|aca|acc|cag|att|ctt|ga|agt|gaaa|agt|cca|ata|aa|gcaa|ag|caa| 1661

Query: 1662 |atc|aga|taag|caa|ata|aa|aga|atg|gt|gat|gtg|aca|agg|cata|cct|agat|ga|act|gg|taga| 1721

Sbjct: 1662 |atc|aga|taag|caa|ata|aa|aga|atg|gt|gat|gtg|aca|agg|cata|cct|agat|ga|act|gg|taga| 1721

Query: 1722 |gctt|cac|aga|agg|tta|atg|ac|attg|ag|aga|aa|ag|acac|att|ctg|cag|cag|atc|gtg|aac|ct| 1781

Sbjct: 1722 gcttcacagaaggttaatgacattgagagaaagacacattctgcagcagatcgtgaacct 1781

Query: 1782 tatagaagaaactggacactttcatatcacaacacacacatttgatctttgctc 1841

|||||

Sbjct: 1782 tatagaagaaactggacactttcatatcacaacacacacatttgatctttgctc 1841

Query: 1842 gctggacaaaaccacagtcctgtaaacacagagttacctggaaacatctggaacatcctg 1901

|||||

Sbjct: 1842 gctggacaaaaccacagtcctgtaaacacagagttacctggaaacatctggaacatcctg 1901

Query: 1902 aggatataacaactggatgcatcaagaactattgtg 1937

|||||

Sbjct: 1902 aggatataacaactggatgcatcaagaactattgtg 1937

Score = 1580 bits (797), Expect = 0.0
Identities = 815/824 (98%)
Strand = Plus / Plus

Query: 2045 caaccttcagctttatctttctttaaagccagtcacatctcttgataaaggagaggta 2104

|||||

Sbjct: 2045 caaccttcagctttatctttctttaaagccagtcacatctcttgataaaggagaggta 2104

Query: 2105 aagcaaaccagcctcagcggaccactcttctctccaaggaaatccccggaagagttagc 2164

|||||

Sbjct: 2105 aagcaaaccagcctcagcggaccactcttctctccaaggaaatccccggaagagttagc 2164

Query: 2165 ctggatagccttgaaaacaaacaaatcaaacacacacacagaaaactcaaagaatgtgta 2224

|||||

Sbjct: 2165 ctggatagccttgaaaacaaacaaatcaaacacacacaagaaaactcaaagaatgtgta 2224

Query: 2225 tggatcatgtatctctctgtgggtggttcattccacaggacgaatgcatattcaacacac 2284
|||||

Sbjct: 2225 tggatcatgtatctctctgtgggtggttcattccacaggacgaatgcatattcaacacac 2284

Query: 2285 tgccttattacataactgatctatttattatcgcatcacagatattctaagtcggtgaggg 2344
|||||

Sbjct: 2285 tgccttattacataactgatctatttattatcgcatcacagatattctaagtcggtgaggg 2344

Query: 2345 aatgacaccatcagacattataagtacttgggtcccgaggatgctctttcaatgcagcacc 2404
|||||

Sbjct: 2345 aatgacaccatcagacattataagtacttgggtcccgaggatgctctttcaatgcagcacc 2404

Query: 2405 cttgccatcccaagcccagtgacctactcgtataaccgtgccactttccaccaacttttt 2464
|||||

Sbjct: 2405 cttgccatcccaagcccagtgacctactcgtataaccgtgccactttccaccaacttttt 2464

Query: 2465 ccaagtcctttaactcgttgccagtcctgtatctttccaccttttgttttccagttccagga 2524
|||||

Sbjct: 2465 ccaagtcctttaactcgttgccagtcctgtatctttccaccttttgttttccagttccagga 2524

Query: 2525 cacagattatcaactggggggaccaaatagccaccttgatcttctctttgtggtctttt 2584
|||||

Sbjct: 2525 cacagattatcaactggggggaccaaatagccaccttgatcttctctttgtggtctttt 2584

Query: 2585 tcctgaaagtggggcccagtccttggtgtatccatgtaatgatcttgaccatggtag 2644
|||||

Sbjct: 2585 tectgaaagttggggcccagtccttggctgtatccatgtaatgatcttggaccatggtag 2644

Query: 2645 aaaatgcaccaaataaggatcatatgaattgctgtctagccttagtcaataaacttgtagg 2704

|||||

Sbjct: 2645 aaaatgcaccaaataaggatcatatgaattgctgtctagccttagtcaataaacttgtagg 2704

Query: 2705 acttttaaacaaaagtgtacctgtaaagtgcctgaatccagcattggtgagctgtcatca 2764

|||||

Sbjct: 2705 acttttaaacaaaagtgtacctgtaaagtgcctgaatccagcattggtgagctgtcatca 2764

Query: 2765 acattcttgtgtctgttttactgttacaatattaggtgaatatggaagtaaaggcattcc 2824

|||||

Sbjct: 2765 acattcttgtgtctgttttactgttacaatattaggtgaatatggaagtaaaggcattcc 2824

Query: 2825 acaggatcatcatttnnnnnnnngaattctggctctgttttct 2868

|||||

Sbjct: 2825 acaggatcatcatttaaaaaaaagaattctggctctgttttct 2868

Score = 948 bits (478), Expect = 0.0

Identities = 481/482 (99%)

Strand = Plus / Plus

Query: 157 taccacacaacaccccagcaaacctccgggccccaggcatggctagctcgtgttccgtg 216

|||||

Sbjct: 157 taccacacaacaccccagcaaacctccgggccccaggcatggctagctcgtgtgccgtg 216

Query: 217 caggtgaagctggagctggggcaccgcgcccaggtgaggaaaaaacccaccgtggagggc 276

|||||

Sbjct: 217 caggtgaagctggagctggggcaccgcgcccaggtgaggaaaaaacccaccgtggagggc 276

Query: 277 ttcacccacgactggatggtggttcgtacgcgggtccggagcacagtaacatacagcacttt 336

|||||

Sbjct: 277 ttcacccacgactggatggtggttcgtacgcgggtccggagcacagtaacatacagcacttt 336

Query: 337 gtggagaaagtcgtcttccacttgcacgaaagctttcctaggccaaaaagagtgtgcaa 396

|||||

Sbjct: 337 gtggagaaagtcgtcttccacttgcacgaaagctttcctaggccaaaaagagtgtgcaa 396

Query: 397 gatccaccttacaagtagaagaatctgggtatgctggtttcattttgccattgaagtt 456

|||||

Sbjct: 397 gatccaccttacaagtagaagaatctgggtatgctggtttcattttgccattgaagtt 456

Query: 457 tattttaaaaacaaggaagaacctaggaaagtcgctttgattatgacttattcctgcat 516

|||||

Sbjct: 457 tattttaaaaacaaggaagaacctaggaaagtcgctttgattatgacttattcctgcat 516

Query: 517 cttgaaggccatccaccagtgaatcacctccgctgtgaaaagctaactttcaacaacccc 576

|||||

Sbjct: 517 cttgaaggccatccaccagtgaatcacctccgctgtgaaaagctaactttcaacaacccc 576

Query: 577 acagaggactttaggagaaagttgctgaaggcaggagggaccctaataaggagtattcat 636

|||||

Sbjct: 577 acagaggactttaggagaaagttgctgaaggcaggagggaccctaataaggagtattcat 636

Query: 637 ac 638

||

Sbjct: 637 ac 638

Score = 900 bits (454), Expect = 0.0
Identities = 482/496 (97%)
Strand = Plus / Plus

Query: 2881 ctggttagaaaattcctaatttggatctatttatttagtcagagtttcagctttcttcagc 2940
|||||
Sbjct: 2880 ctggttagaaaattcctaatttggatctatttatttagtcagagtttcagctttcttcagc 2939

Query: 2941 tgccagtgtgttactcatctttatcctaaaaatctggaatcagagatttttgttgttca 3000
|||||
Sbjct: 2940 tgccagtgtgttactcatctttatcctaaaaatctggaatcagagatttttgttgttca 2999

Query: 3001 catatgattctcttagacacttttatatttgaaaaattaaaatctttctttggggaaaa 3060
|||||
Sbjct: 3000 catatgattctcttagacacttttatatttgaaaaattaaaatctttctttggggaaaa 3059

Query: 3061 attcttggttattctgccataacagattatgtattaacttgtagattcagtggttcaata 3120
|||||
Sbjct: 3060 attcttggttattctgccataacagattatgtattaacttgtagattcagtggttcaata 3119

Query: 3121 cctgtttagttgcttgctaataattccagaaggatttcttgattggtgaaagacggttg 3180
|||||
Sbjct: 3120 cctgtttagttgcttgctaataattccagaaggatttcttgattggtgaaagacggttg 3179

Query: 3181 gggatggggggannnnnnngttcttggtgtacccttgtttgaaactagaaatctgtcct 3240
|||||
Sbjct: 3180 gggatggggggatttttttggttcttggtgtacccttgtttgaaactagaaatctgtcct 3239

Query: 3241 gtggcatgcaaaagaaagcaaattatTTTTAAAGNNNNNNCCAAAGTACTTTTGGTGT 3300

Sbjct: 3240 gtggcatgcaaaagaaagcaaatatttttaaaagaaaaaaccaaagtacttttgggtgt 3299

Query: 3301 cattattccatcttctccataagtggagaaatgaaaagtaagaacagctcatcttcaaag 3360

Sbjct: 3300 cattattccatcttctccataagtggagaaatgaaaagtaagaacagctcatcttcaaag 3359

Query: 3361 tttttactagaaattc 3376

Sbjct: 3360 tttttactagaaattc 3375

Score = 264 bits (133), Expect = 3e-67
Identities = 133/133 (100%)
Strand = Plus / Plus

Query: 1 tttggggctgagtttaataagcgagcgagcgagcaagcgagcgcggggggaaaaaggcag 60

Sbjct: 1 tttggggctgagtttaataagcgagcgagcgagcaagcgagcgcggggggaaaaaggcag 60

Query: 61 agaatgtccgccatctaccctccgctcctgggcgcgctctcattcatagcagcctcttca 120

Sbjct: 61 agaatgtccgccatctaccctccgctcctgggcgcgctctcattcatagcagcctcttca 120

Query: 121 tgaattacagctg 133

Sbjct: 121 tgaattacagctg 133

Score = 61.9 bits (31), Expect = 2e-06

Identities = 31/31 (100%)
Strand = Plus / Plus

Query: 1966 ggttgtgattttttgttccttgttgtttatat 1996
|||||
Sbjct: 1966 ggttgtgattttttgttccttgttgtttatat 1996

>dbj|D16688.1|HUMMLLT3 Human LTG9/MLLT3 mRNA, C-terminal
Length = 1486

Score = 2070 bits (1044), Expect = 0.0
Identities = 1058/1065 (99%)
Strand = Plus / Plus

Query: 873 caaatcttccaaagaatcctctaagaaacccaaagaaaataaaccactgaaagaagagaa 932
|||||
Sbjct: 1 caaatcttccaaagaatcctctaagaaacccaaagaaaataaaccactgaaagaagagaa 60

Query: 933 aatagttcctaagatggccttcaaggaacctaaccatgtcaaaagagccaaaaccaga 992
|||||
Sbjct: 61 aatagttcctaagatggccttcaaggaacctaaccatgtcaaaagagccaaaaccaga 120

Query: 993 tagtaacttactcaccatcaccagtggacaagataagaaggctcctagtaaaaggccgcc 1052
|||||
Sbjct: 121 tagtaacttactcaccatcaccagtggacaagataagaaggctcctagtaaaaggccgcc 180

Query: 1053 catttcagattctgaagaactctcagccnnnnnnnggaaaaagagtagctcagaggcttt 1112
|||||

Sbjct: 181 catttcagattctgaagaactctcagccaaaaaaggaaaaagagtagctcagaggcttt 240

Query: 1113 atttaaaagtttttctagcgcaccaccactgatactcacttggttctgctgacaaaaaca 1172
 |||

Sbjct: 241 atttaaaagtttttctagcgcaccaccactgatactcacttggttctgctgacaaaaaca 300

Query: 1173 gataaaagataaatctcatgtcaagatgggaaagggtcaaaattgaaagtgagacatcaga 1232
 |||

Sbjct: 301 gataaaagataaatctcatgtcaagatgggaaagggtcaaaattgaaagtgagacatcaga 360

Query: 1233 gaagaagaaatcaacggtaccgccatttgatgatattgtggatcccaatgattcagatgt 1292
 |||

Sbjct: 361 gaagaagaaatcaacggtaccgccatttgatgatattgtggatcccaatgattcagatgt 420

Query: 1293 ggaggagaatatatcctctaaatctgattctgaacaaccagtcctgccagctccagctc 1352
 |||

Sbjct: 421 ggaggagaatatatcctctaaatctgattctgaacaaccagtcctgccagctccagctc 480

Query: 1353 cagctccagctccagcttcacaccatcccagaccaggcaacaagggtcctttgaggtctat 1412
 |||

Sbjct: 481 cagctccagctccagcttcacaccatcccagaccaggcaacaagggtcctttgaggtctat 540

Query: 1413 aatgaaagatctgcattctgatgacaatgaggaggaatcagatgaagtggaggataacga 1472
 |||

Sbjct: 541 aatgaaagatctgcattctgatgacaatgaggaggaatcagatgaagtggaggataacga 600

Query: 1473 caatgactctgaaatggagaggcctgtaaatagaggaggcagccgaagtcgcagagttag 1532
 |||

Sbjct: 601 caatgactctgaaatggagagggcctgtaaatagaggaggcagccgaagtcgcagagttag 660

Query: 1533 cttaagtgatggcagcgatagtgaaagcagttctgcttcttcaccctacatcacgaacc 1592
 |||
 Sbjct: 661 cttaagtgatggcagcgatagtgaaagcagttctgcttcttcaccctacatcacgaacc 720

Query: 1593 tccaccacccttactaaaaaccaacaacaaccagattcttgaagtgaaaagtccaataaa 1652
 |||
 Sbjct: 721 tccaccacccttactaaaaaccaacaacaaccagattcttgaagtgaaaagtccaataaa 780

Query: 1653 gcaaagcaaatacagataagcaaataaagaatggatgaatgtgacaaggcatacctagatga 1712
 |||
 Sbjct: 781 gcaaagcaaatacagataagcaaataaagaatggatgaatgtgacaaggcatacctagatga 840

Query: 1713 actggtagagcttcacagaaggtaatacattgagagaaagacacattctgcagcagat 1772
 |||
 Sbjct: 841 actggtagagcttcacagaaggtaatacattgagagaaagacacattctgcagcagat 900

Query: 1773 cgtgaaccttatagaagaaactggacactttcatatcaciaaacacaacatttgattttga 1832
 |||
 Sbjct: 901 cgtgaaccttatagaagaaactggacactttcatatcaciaaacacaacatttgattttga 960

Query: 1833 tctttgctcgctggacaaaaccacagtcgtaaaactacagagttacctggaacatctgg 1892
 |||
 Sbjct: 961 tctttgctcgctggacaaaaccacagtcgtaaaactacagagttacctggaacatctgg 1020

Query: 1893 aacatcctgaggatataacaactggatgcatcaagaactattgtg 1937
 |||

Sbjct: 1021 aacatcctgaggatataacaactggatgcatcaagaactattgtg 1065

Score = 622 bits (314), Expect = e-175
Identities = 314/314 (100%)
Strand = Plus / Plus

Query: 2045 caaccttcagctttatTTTTCTTTAAAGCCAGTcatcatctcttgataaaggagaggta 2104
|||||
Sbjct: 1173 caaccttcagctttatTTTTCTTTAAAGCCAGTcatcatctcttgataaaggagaggta 1232

Query: 2105 aagcaaaccagcctcagcggaccactcttctctccaaggaaatccccgggaagagttagc 2164
|||||
Sbjct: 1233 aagcaaaccagcctcagcggaccactcttctctccaaggaaatccccgggaagagttagc 1292

Query: 2165 ctggatagccttgaaaacaacaatcaaacacaacacaagaaaactcaaagaatgtgta 2224
|||||
Sbjct: 1293 ctggatagccttgaaaacaacaatcaaacacaacacaagaaaactcaaagaatgtgta 1352

Query: 2225 tggatcatgtatctctctgtggtgggttcattccacaggacgaatgcatattcaacacac 2284
|||||
Sbjct: 1353 tggatcatgtatctctctgtggtgggttcattccacaggacgaatgcatattcaacacac 1412

Query: 2285 tgccttattacataactgatctatttattatcgcatacagatattctaagtcgttgaggg 2344
|||||
Sbjct: 1413 tgccttattacataactgatctatttattatcgcatacagatattctaagtcgttgaggg 1472

Query: 2345 aatgacaccatcag 2358
|||||

Sbjct: 1473 aatgacaccatcag 1486

Score = 61.9 bits (31), Expect = 2e-06
Identities = 31/31 (100%)
Strand = Plus / Plus

Query: 1966 ggttgtgattttttgttcttggtttatat 1996

|||||

Sbjct: 1094 ggttgtgattttttgttcttggtttatat 1124

>gb|AC002050.6|AC002050 Homo sapiens Chromosome 9p22 Cosmid Clone 213a7, complete sequence
Length = 42210

Score = 1580 bits (797), Expect = 0.0
Identities = 815/824 (98%)
Strand = Plus / Minus

Query: 2045 caaccttcagctttatcttttaagccagtcacatctcttgataaaggagaggtta 2104

|||||

Sbjct: 39362 caaccttcagctttatcttttaagccagtcacatctcttgataaaggagaggtta 39303

Query: 2105 aagcaaaccagcctcagcggaccactcttctctccaaggaaatccccgggaagagttagc 2164

|||||

Sbjct: 39302 aagcaaaccagcctcagcggaccactcttctctccaaggaaatccccgggaagagttagc 39243

Query: 2165 ctggatagccttgaaaacaaacaaatcaaacacacacaagaaaactcaagaatgtgta 2224

|||||

Sbjct: 39242 ctggatagccttgaaaacaaacaaatcaaacacacacaagaaaactcaaagaatgtgta 39183

Query: 2225 tggatcatgtatctctctgtggtggttcattccacaggacgaatgcatattcaacacac 2284
|||||

Sbjct: 39182 tggatcatgtatctctctgtggtggttcattccacaggacgaatgcatattcaacacac 39123

Query: 2285 tgccttattacataactgatctatattattatcgcatacagatattctaagtcggtgaggg 2344
|||||

Sbjct: 39122 tgccttattacataactgatctatattattatcgcatacagatattctaagtcggtgaggg 39063

Query: 2345 aatgacaccatcagacattataagtacttgggtcccgaggatgctctttcaatgcagcacc 2404
|||||

Sbjct: 39062 aatgacaccatcagacattataagtacttgggtcccgaggatgctctttcaatgcagcacc 39003

Query: 2405 cttgccatcccaagcccagtgaccttactcgtataaccgtgccactttccaccaacttttt 2464
|||||

Sbjct: 39002 cttgccatcccaagcccagtgaccttactcgtataaccgtgccactttccaccaacttttt 38943

Query: 2465 ccaagtcctttaactcgttgagctctgtatctttccacctttgtttttccagttccagga 2524
|||||

Sbjct: 38942 ccaagtcctttaactcgttgagctctgtatctttccacctttgtttttccagttccagga 38883

Query: 2525 cacagattatcaactggggggaccaaatagccaccttgatctttcttctttgtggtctttt 2584
|||||

Sbjct: 38882 cacagattatcaactggggggaccaaatagccaccttgatctttcttctttgtggtctttt 38823

Query: 2585 tcctgaaagtggggcccagtccttggctgtatccatgtaatgatcttggaccatggtag 2644
|||||

Sbjct: 38822 tcttgaaagttggggcccagtccttggtgtatccatgtaatgatcttgaccatggtag 38763

Query: 2645 aaaatgcaccaaataaggatcatatgaattgctgtctagccttagtcaataaacttgtagg 2704

|||||

Sbjct: 38762 aaaatgcaccaaataaggatcatatgaattgctgtctagccttagtcaataaacttgtagg 38703

Query: 2705 acttttaaacaaaagtgtacctgtaaagtctgaatccagcattggtgagctgtcatca 2764

|||||

Sbjct: 38702 acttttaaacaaaagtgtacctgtaaagtctgaatccagcattggtgagctgtcatca 38643

Query: 2765 acattcttgtgtctgttttactgttacaatattaggtgaatatggaagtaaaggcattcc 2824

|||||

Sbjct: 38642 acattcttgtgtctgttttactgttacaatattaggtgaatatggaagtaaaggcattcc 38583

Query: 2825 acaggatcatcatttannnnnnnngaattctggtcctgttttct 2868

|||||

Sbjct: 38582 acaggatcatcatttaaaaaaaaagaattctggtcctgttttct 38539

Score = 900 bits (454), Expect = 0.0

Identities = 482/496 (97%)

Strand = Plus / Minus

Query: 2881 ctggttagaaaattcttaatttggatctatttattagtcagagtttcagctttcttcagc 2940

|||||

Sbjct: 38527 ctggttagaaaattcttaatttggatctatttattagtcagagtttcagctttcttcagc 38468

Query: 2941 tgccagtgtgttactcatctttatcctaaaaatctggaatcagagattttgttgttca 3000

|||||

Sbjct: 38467 tgccagtgtgttactcatctttatcctaaaaatctggaatcagagattttgttgttca 38408

Query: 3001 catatgattctcttagacacttttatatttgaaaaaattaaaatctttctttggggaaaa 3060
 |||
 Sbjct: 38407 catatgattctcttagacacttttatatttgaaaaaattaaaatctttctttggggaaaa 38348

Query: 3061 attccttggttattctgccataacagattatgtattaacttgtagattcagtggttcaata 3120
 |||
 Sbjct: 38347 attccttggttattctgccataacagattatgtattaacttgtagattcagtggttcaata 38288

Query: 3121 cctgtttagttgcttgctaataattccagaaggatttcttgattggtgaaagacggttg 3180
 |||
 Sbjct: 38287 cctgtttagttgcttgctaataattccagaaggatttcttgattggtgaaagacggttg 38228

Query: 3181 gggatgggggannnnnngttccttggtgtacccttgtttgaaactagaaatctgtcct 3240
 |||
 Sbjct: 38227 gggatgggggatttttttggtccttggtgtacccttgtttgaaactagaaatctgtcct 38168

Query: 3241 gtggcatgcaaaagaaagcaaattatTTTTAAAAGNNNNNNCCAAAGTACTTTGGTGT 3300
 |||
 Sbjct: 38167 gtggcatgcaaaagaaagcaaattatTTTTAAAAGAAAAAACAAAGTACTTTGGTGT 38108

Query: 3301 cattattccatcttctccataagtggagaaatgaaaagtaagaacagctcatcttcaaag 3360
 |||
 Sbjct: 38107 cattattccatcttctccataagtggagaaatgaaaagtaagaacagctcatcttcaaag 38048

Query: 3361 tttttactagaaattc 3376
 |||
 Sbjct: 38047 tttttactagaaattc 38032

Score = 335 bits (169), Expect = 9e-89
Identities = 169/169 (100%)
Strand = Plus / Minus

Query: 1769 agatcgtgaaccttatagaagaaactggacactttcatatcaciaacacaacatttgatt 1828
|||||
Sbjct: 39638 agatcgtgaaccttatagaagaaactggacactttcatatcaciaacacaacatttgatt 39579

Query: 1829 ttgatctttgctcgctggacaaaaccacagtccgtaaactacagagttacctggaaacat 1888
|||||
Sbjct: 39578 ttgatctttgctcgctggacaaaaccacagtccgtaaactacagagttacctggaaacat 39519

Query: 1889 ctggaacatcctgaggatataacaactggatgcatcaagaactattgtg 1937
|||||
Sbjct: 39518 ctggaacatcctgaggatataacaactggatgcatcaagaactattgtg 39470

Score = 61.9 bits (31), Expect = 2e-06
Identities = 31/31 (100%)
Strand = Plus / Minus

Query: 1966 ggttgtgattttttgttcttgttgtttatat 1996
|||||
Sbjct: 39441 ggttgtgattttttgttcttgttgtttatat 39411

>emb|AJ006295.1|RNO6295 Rattus norvegicus mRNA for AF-9
Length = 1591

Score = 1477 bits (745), Expect = 0.0

Identities = 1067/1176 (90%), Gaps = 3/1176 (0%)
Strand = Plus / Plus

Query: 762 tagtaccagtttttcaaagcctcacaaattaatgaaggagcacaaggaaaaaccttctaa 821
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Sbjct: 304 tagtaccagtttttcaaagcctcacaaattaatgaaggagcacaaggaaaaaccttctaa 363

Query: 822 agactccagagAACATAAAAAGTGCOTTCAAAGAACCTTCCAGGGATCACAACAAATCTTC 881
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Sbjct: 364 ggactccagagAACATAAAAAGTGCOTTCAAAGAACCTTCCAGGGATCACAACAAATCTTC 423

Query: 882 caaagaatcctctaagaaacccaaagaaaataaacactgaaagaagagaaaatagttcc 941
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Sbjct: 424 caaagactcctccaagaaacccaaagaaaacaacactgaaagaagagaaaatcgttcc 483

Query: 942 taagatggccttcaaggaacctaAACCCATGTCAAAAGAGCCAAAACCATAGTAACCT 1001
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Sbjct: 484 taagatggccttcaaggagcctaAACCCATGTCAAAAGAGCCAAAACCATAGTAACCT 543

Query: 1002 actcaccatcaccagtgga---caagataagaaggctcctagtaaaaggccgccatttc 1058
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Sbjct: 544 actcaccgtcaccagtgcacagcaggacaagaaggcgcctagtaaaaggccccctgttc 603

Query: 1059 agattctgaagaactctcagccnnnnnnnggaaaaagagtagctcagaggctttatttaa 1118
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Sbjct: 604 agattctgaagaactctcagccaaaaaaaggaaaaagagtagctccgaggctttatttaa 663

Query: 1119 aagtttttctagcgcaccaccactgatactcacttgttctgctgacaaaaaacagataaa 1178
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Sbjct: 1119 aagtttttctagcgcaccaccactgatactcacttgttctgctgacaaaaaacagataaa 1178

Sbjct: 664 aagtttttctagcgcgccccactgatactcacttgttctgctgacaaaaacagataaa 723

Query: 1179 agataaatctcatgtcaagatgggaaagggtcaaaattgaaagtgagacatcagagaagaa 1238
|||||

Sbjct: 724 agataaatctcatgtcaagatgggaaagggtcaaaatcgagagcgagacttcagagaagaa 783

Query: 1239 gaaatcaacgttaccgccatttgatgatattgtggatcccaatgattcagatgtggagga 1298
||| || | | |||

Sbjct: 784 gaagtcgacactgccaccatttgatgatatcgtggatcccaacgactcagacgtggagga 843

Query: 1299 gaatatatcctctaaatctgattctgaacaaccagtcctgccagctccagctccagctc 1358
||| || |||||

Sbjct: 844 gaacatgtcctcaaaatctgactcagagcaaccagtcctgccagttccagttccagttc 903

Query: 1359 cagctccagcttcacaccatcccagaccaggcaacaaggctcctttgaggtctataatgaa 1418
|||||

Sbjct: 904 cagctccagcttcacgccgtctcagaccaggcagcaaggctcctttaaggctgatcatgaa 963

Query: 1419 agatctgcattctgatgacaatgaggaggaatcagatgaagtggaggataacgacaatga 1478
|||||

Sbjct: 964 agatctgcattccgatgacaacgaggaggagtctgatgaggccgaggacaatgacaatga 1023

Query: 1479 ctctgaaatggagaggcctgtaaatagaggaggcagccgaagtgcgagagttagcttaag 1538
|||||

Sbjct: 1024 ctctgagatggaaaggcctgtaaacagagggggaagccgaagtgcgagggtcagcttgag 1083

Query: 1539 tgatggcagcgatagtgaaagcagttctgcttcttcaccctacatcacgaacctccacc 1598
||| |||||

Sbjct: 1084 tgacggcagtgacagtgaaagcagctctgcctcttcccctctccatcacgaacccccacc 1143

Query: 1599 acccttactaaaaaccaacaacaaccagattcttgaagtgaaaagtccaataaagcaaag 1658
|| ||| | ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Sbjct: 1144 gccattattgaaaactaacaacaaccagattcttgaagtgaaaagtccaataaacaag 1203

Query: 1659 caaatcagataagcaaataaagaatggtgaatgtgacaaggcatacctagatgaactggt 1718
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Sbjct: 1204 caaatcagataagcaaataaagaatggtgaatgtgacaaggcttacctagatgagttggt 1263

Query: 1719 agagcttcacagaaggttaatgacattgagagaaagacacattctgcagcagatcgtgaa 1778
||| || ||||||| ||||||| ||||| ||||||| ||||||| ||||||| ||||||| |||||||

Sbjct: 1264 agaactccacagaaggttaatgacactgagggaaagacacattctgcagcagatcgtgaa 1323

Query: 1779 ccttatagaagaaactggacactttcatatcaciaaacacaacatttgattttgatctttg 1838
||||| ||||||| ||||||| ||||||| ||||| ||||||| ||||||| ||||||| |||||||

Sbjct: 1324 ccttatagaggaaactggacactttcatattacaataacaacatttgattttgatctttg 1383

Query: 1839 ctcgctggacaaaaccacagtcgtaactacagagttacctggaaacatctggaacatc 1898
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Sbjct: 1384 ctcgctggacaaaactacagtcgtaattacagagttacctggagacatctggaacatc 1443

Query: 1899 ctgaggatataacaactggatgcatcaagaactatt 1934
||||| || ||||||| ||||||| ||||||| |||||||

Sbjct: 1444 ctgagggtacaacaactggacgcatcaaaaactatt 1479

Score = 272 bits (137), Expect = 1e-69

Identities = 158/165 (95%)
Strand = Plus / Plus

Query: 474 agaacctaggaagtcgctttgattatgacttattcctgcatcttgaaggccatccacc 533
||||||| ||||||||||||||||||||||||||||||||||||||||| |||||
Sbjct: 16 agaacctaggaagtcgctttgattatgacttattcctgcatcttgaaggccaccacc 75

Query: 534 agtgaatcacctccgctgtgaaaagctaactttcaacaaccccacagaggactttaggag 593
||||||||||||||||||| ||||||||||||||| ||||||||||||||||| ||
Sbjct: 76 agtgaatcacctccgctgtgagaagetaactttcaataaccccacagaggactttagaag 135

Query: 594 aaagttgctgaaggcaggaggggaccctaataaggagtattcatac 638
||||||||||||||||||||||||||||||||||| ||||| |||||||||
Sbjct: 136 aaagttgctgaaggcaggaggggaccctaacaggagcattcatac 180

>gb|AC002053.2|AC002053 Cosmid clone 92f5 from human chromosome band 9p22, complete sequence
[Homo sapiens]
Length = 27613

Score = 1067 bits (538), Expect = 0.0
Identities = 552/559 (98%)
Strand = Plus / Minus

Query: 762 tagtaccagtttttcaaagcctcacaattaatgaaggagcacaaggaaaaaccttctaa 821
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 23861 tagtaccagtttttcaaagcctcacaattaatgaaggagcacaaggaaaaaccttctaa 23802

Query: 822 agactccagagaacataaaaagtgcccttcaagaaccttccagggatcacaacaaatcttc 881
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 23801 agactccagagaacataaaaagtgcccttcaagaaccttccagggatcacaacaaatcttc 23742

Query: 882 caaagaatcctctaagaaacccaaagaaaataaaccactgaaagaagagaaaatagttcc 941
 |||
 Sbjct: 23741 caaagaatcctctaagaaacccaaagaaaataaaccactgaaagaagagaaaatagttcc 23682

Query: 942 taagatggccttcaaggaacctaaaccatgtcaaaagagccaaaaccagatagtaactt 1001
 |||
 Sbjct: 23681 taagatggccttcaaggaacctaaacecatgtcaaaagagccaaaaccagatagtaactt 23622

Query: 1002 actcaccatcaccagtggacaagataagaaggctcctagtaaaggccgccatttcaga 1061
 |||
 Sbjct: 23621 actcaccatcaccagtggacaagataagaaggctcctagtaaaggccgccatttcaga 23562

Query: 1062 ttctgaagaactctcagccnnnnnnnggaaaaagagtagctcagaggctttatttaaaag 1121
 |||
 Sbjct: 23561 ttctgaagaactctcagccaaaaaaaggaaaaagagtagctcagaggctttatttaaaag 23502

Query: 1122 tttttctagcgcaccaccactgatactcacttggttctgctgacaaaaaacagataaaaaga 1181
 |||
 Sbjct: 23501 tttttctagcgcaccaccactgatactcacttggttctgctgacaaaaaacagataaaaaga 23442

Query: 1182 taaatctcatgtcaagatgggaaagggtcaaaattgaaagtgagacatcagagaagaagaa 1241
 |||
 Sbjct: 23441 taaatctcatgtcaagatgggaaagggtcaaaattgaaagtgagacatcagagaagaagaa 23382

Query: 1242 atcaacgttaccgccatttgatgatattgtggatcccaatgattcagatgtggaggagaa 1301
 |||
 Sbjct: 23381 atcaacgttaccgccatttgatgatattgtggatcccaatgattcagatgtggaggagaa 23322

Query: 1302 tatatcctctaaatctgat 1320
|||||
Sbjct: 23321 tatatcctctaaatctgat 23303

Score = 48.1 bits (24), Expect = 0.030
Identities = 24/24 (100%)
Strand = Plus / Minus

Query: 615 ggaccctaataaggagtattcatac 638
|||||
Sbjct: 24008 ggaccctaataaggagtattcatac 23985

>[gb|AC002052.8|AC002052](#) Homo sapiens Chromosome 9p22 Cosmid Clone 34a5, complete sequence
Length = 40679

Score = 353 bits (178), Expect = 4e-94
Identities = 181/182 (99%)
Strand = Plus / Plus

Query: 207 gtgttccgtgcaggtgaagctggagctggggcaccgcgccaggtgaggaaaaaacccac 266
|||||
Sbjct: 15093 gtgtgccgtgcaggtgaagctggagctggggcaccgcgccaggtgaggaaaaaacccac 15152

Query: 267 cgtggagggttcacccacgactggatggtgttcgtacgcggtccggagcacagtaacat 326
|||||
Sbjct: 15153 cgtggagggttcacccacgactggatggtgttcgtacgcggtccggagcacagtaacat 15212

Query: 327 acagcactttgtggagaaagtcgtcttccacttgacgaaagctttcctaggccaaaaag 386
|||||

Sbjct: 15213 acagcactttgtggagaaagtcgtcttccacttgcacgaaagctttcctagccaaaaag 15272

Query: 387 ag 388

||

Sbjct: 15273 ag 15274

Score = 264 bits (133), Expect = 3e-67
Identities = 133/133 (100%)
Strand = Plus / Plus

Query: 1 tttggggctgagtttaataagcgagcgagcgagcaagcgagcgcggggggaaaaaggcag 60

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct: 13476 tttggggctgagtttaataagcgagcgagcgagcaagcgagcgcggggggaaaaaggcag 13535

Query: 61 agaatgtccgccatctaccctccgctcctgggcgcgctctcattcatagcagcctttca 120

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct: 13536 agaatgtccgccatctaccctccgctcctgggcgcgctctcattcatagcagcctttca 13595

Query: 121 tgaattacagctg 133

||||||||||||

Sbjct: 13596 tgaattacagctg 13608

Score = 101 bits (51), Expect = 2e-18
Identities = 51/51 (100%)
Strand = Plus / Plus

Query: 157 taccacacaacaccccagcaaactccgggccccagggcatggctagctcg 207

|||||
Sbjct: 13632 taccacacaacaccccagcaaactccgggccccagggcatggctagctcg 13682

>gb|AC000007.8|AC000007 Homo sapiens Chromosome 9 Cosmid Clone c48, complete sequence
Length = 40556

Score = 262 bits (132), Expect = 1e-66
Identities = 132/132 (100%)
Strand = Plus / Minus

Query: 1395 aggtcctttgaggtctataatgaaagatctgcattctgatgacaatgaggaggaatcaga 1454
|||||
Sbjct: 12812 aggtcctttgaggtctataatgaaagatctgcattctgatgacaatgaggaggaatcaga 12753

Query: 1455 tgaagtggaggataacgacaatgactctgaaatggagaggcctgtaaatagaggaggcag 1514
|||||
Sbjct: 12752 tgaagtggaggataacgacaatgactctgaaatggagaggcctgtaaatagaggaggcag 12693

Query: 1515 ccgaagtgcgag 1526
|||||
Sbjct: 12692 ccgaagtgcgag 12681

Score = 204 bits (103), Expect = 2e-49
Identities = 103/103 (100%)
Strand = Plus / Minus

Query: 1524 cagagttagcttaagtgatggcagcgatagtgaaagcagttctgcttcttcaccctaca 1583

|||||
Sbjct: 10049 cagagttagcttaagtgatggcagcgatagtgaaagcagttctgcttcttcaccctaca 9990

Query: 1584 tcacgaacctccaccacccttactaaaaaccaacaacaaccag 1626
|||||
Sbjct: 9989 tcacgaacctccaccacccttactaaaaaccaacaacaaccag 9947

Score = 155 bits (78), Expect = 2e-34
Identities = 78/78 (100%)
Strand = Plus / Minus

Query: 1321 tctgaacaaccagtcctgccagctccagctccagctccagctccagcttcacaccatcc 1380
|||||
Sbjct: 14949 tctgaacaaccagtcctgccagctccagctccagctccagctccagcttcacaccatcc 14890

Query: 1381 cagaccaggcaacaaggt 1398
|||||
Sbjct: 14889 cagaccaggcaacaaggt 14872

Score = 149 bits (75), Expect = 1e-32
Identities = 75/75 (100%)
Strand = Plus / Minus

Query: 1625 agattcttgaagtgaaaagtccaataaagcaaagcaaatcagataagcaaataaagaatg 1684
|||||
Sbjct: 4086 agattcttgaagtgaaaagtccaataaagcaaagcaaatcagataagcaaataaagaatg 4027

Query: 1685 gtgaatgtgacaagg 1699

|||||

Sbjct: 4026 gtgaatgtgacaagg 4012

Score = 147 bits (74), Expect = 4e-32
Identities = 74/74 (100%)
Strand = Plus / Minus

Query: 1697 aggcatacctagatgaactggtagagcttcacagaagggttaatgacattgagagaaagac 1756

|||||

Sbjct: 2803 aggcatacctagatgaactggtagagcttcacagaagggttaatgacattgagagaaagac 2744

Query: 1757 acattctgcagcag 1770

|||||

Sbjct: 2743 acattctgcagcag 2730

>gb|AF312858.1|AF312858 Mus musculus leukemia associated gene protein (Ml1t1) mRNA,
complete cds
Length = 1793

Score = 93.7 bits (47), Expect = 6e-16
Identities = 116/139 (83%)
Strand = Plus / Plus

Query: 217 caggtgaagctggagctggggcaccgcccaggtgaggaaaaaacccaccgtggagggc 276

|||||

Sbjct: 33 caggtgaagttagagctggggcaccgtgccagctacggaagaagcccaccaccgagggc 92

Query: 277 ttcaccacgactggatggtggttcgtacggtccggagcacagtaacatacagcacttt 336

|||||

Sbjct: 93 ttcacccatgaetggatggtcttcgtagcgtggccctgagcagtgtagacatccagcacttt 152

Query: 337 gtggagaaagtcgtcttcc 355

||||||| ||| |||||

Sbjct: 153 gtggagaaggtcatcttcc 171

>[ref|NM_022328.1](#) Mus musculus myeloid/lymphoid or mixed-lineage leukemia (trithorax (Drosophila) homolog); translocated to, 1 (Mllt1), mRNA
Length = 1644

Score = 93.7 bits (47), Expect = 6e-16

Identities = 116/139 (83%)

Strand = Plus / Plus

Query: 217 caggtgaagctggagctggggcaccgcgccaggtgaggaaaaaacccaccgtggagggc 276

||||||| | ||||||||| ||||| | ||| || ||||| |||||

Sbjct: 22 caggtgaagttagagctggggcaccgtgccagctacggaagaagcccaccaccgagggc 81

Query: 277 ttcacccacgactggatggtgttcgtagcgggtccggagcacagtaacatacagcacttt 336

||||||| ||||||||| ||||| || || || ||||| || ||||| |||||

Sbjct: 82 ttcacccatgactggatggtcttcgtagcgtggccctgagcagtgtagacatccagcacttt 141

Query: 337 gtggagaaagtcgtcttcc 355

||||||| ||| |||||

Sbjct: 142 gtggagaaggtcatcttcc 160

>[gb|AF298887.1|AF298887](#) Mus musculus Btk-PH-domain binding protein mRNA, complete cds

Length = 1644

Score = 93.7 bits (47), Expect = 6e-16
Identities = 116/139 (83%)
Strand = Plus / Plus

Query: 217 caggtgaagctggagctggggcaccgcgccaggtgaggaaaaaacccaccgtggagggc 276
||||||||| | ||||||||||||| |||||| | ||| | | |||||| | ||||||

Sbjct: 22 caggtgaagttagagctggggcaccgtgccagctacggaagaagcccaccaccgagggc 81

Query: 277 ttcaccacgactggatggtgttcgtacgcggtccggagcacagtaacatacagcacttt 336
||||||||| ||||||||||| ||||| | | | | ||||| | | |||| | |||||||

Sbjct: 82 ttcaccatgactggatggtcttcgtgcgtggccctgagcagtgtgacatccagcacttt 141

Query: 337 gtggagaaagtcgtttcc 355
||||||||| ||| |||||||

Sbjct: 142 gtggagaaggtcatcttcc 160

>ref|XM 009359.1| Homo sapiens myeloid/lymphoid or mixed-lineage leukemia (trithorax
(Drosophila) homolog); translocated to, 1 (MLLT1), mRNA
Length = 1677

Score = 85.7 bits (43), Expect = 1e-13
Identities = 286/367 (77%)
Strand = Plus / Plus

Query: 212 ccgtgcaggtgaagctggagctggggcaccgcgccaggtgaggaaaaaacccaccgtgg 271
|||| | ||||||| | | ||||||||||| ||||||| | | | | |||||| | |

Sbjct: 14 ccgtccaggtgaggttagagctggggcatcgcgcccaactgcgcaagaagcccaccacgg 73

Query: 272 agggcttcacccacgactggatgggtgttcgtacgcgggtccggagcacagtaacatacagc 331
||||| ||||| ||||||||||||||||| || ||||| || ||||| || ||||| |||||
Sbjct: 74 aggggttcactcacgactggatgggtgttgtccgcggccccgagcaatgtgacatccagc 133

Query: 332 actttgtggagaaagtcgtcttccacttgcacgaaagctttcctaggccaaaaagagtgt 391
||||| ||||||| || ||||||| ||||||| ||||| || | ||| | | |||||
Sbjct: 134 acttcgtggagaaggtggtcttctggctgcacgacagcttcccccaagcccagacgcgtgt 193

Query: 392 gcaaagatccaccttaciaaagtagaagaatctgggtatgctggtttcattttgccaattg 451
||||| || || || ||||||||||||| || || ||||| ||||| ||||| ||||| ||| |||
Sbjct: 194 gcaaggagccccctaciaaagtagaggagtcggggtagcgtggcttcatcatgcccatcg 253

Query: 452 aagtttattttaaaaacaaggaagaacctaggaagtcgcgtttgattatgacttattcc 511
| || | || ||||||||||||| || || ||||| || ||||| || ||||| || |||||
Sbjct: 254 aggtgcacttcaaaaacaaggaggagccgaggaaggtctgcttcacctacgacctgttcc 313

Query: 512 tgcacttgaaggccatccaccagtgaatcacctccgctgtgaaaagctaactttcaaca 571
|| | || ||||||| | || || ||||| ||||| ||||| || ||||| || |||||||
Sbjct: 314 tgaacctggaaggcaaccgcccgtgaaccacctgcgctgcgagaaagctcaccttcaaca 373

Query: 572 accccac 578
|||||||
Sbjct: 374 accccac 380

>ref|NM_005934.1| Homo sapiens myeloid/lymphoid or mixed-lineage leukemia (trithorax
(Drosophila) homolog); translocated to, 1 (MLLT1), mRNA
Length = 1680

Score = 85.7 bits (43), Expect = 1e-13
Identities = 286/367 (77%)

Strand = Plus / Plus

Query: 212 ccgtgcaggtgaagctggagctggggcaccgcgccaggtgaggaaaaaccaccgtgg 271

||||| ||||||| | | ||||||||| ||||||| || | || || || || || ||

Sbjct: 17 ccgtccaggtgaggttagagctggggcatcgcgcccaactgcgcaagaagcccaccacgg 76

Query: 272 agggcttcacccacgactggatggtgttcgtacgcggtccggagcacagtaacatacagc 331

||||| ||||||| ||||||||| ||||||| || ||||||| || ||||||| || ||

Sbjct: 77 aggggttcactcacgactggatggtgtttgtccgcggcccccagcaatgtgacatccagc 136

Query: 332 actttgtggagaaagtcgtcttcacttgcacgaaagctttcctaggccaaaaagagtgt 391

||||| ||||||||| || ||||||| || ||||||| || || || || || || || ||

Sbjct: 137 acttcgtggagaaggtggtcttctggctgcacgacagcttccccaagcccagacgcgtgt 196

Query: 392 gcaaagatccaccttataaaagtagaagaatctgggtatgctggtttcattttgccaattg 451

||||| || || || || || || || || || || || || || || || || || || ||

Sbjct: 197 gcaaggagccccctataaaagtagaggagtcggggtagcgtggcttcatcatgcccatcg 256

Query: 452 aagtttattttaaaaaacaaggaagaacctaggaaagtcgctttgattatgacttattcc 511

| || | || || || || || || || || || || || || || || || || || || ||

Sbjct: 257 aggtgcacttcaaaaaacaaggaggagccgaggaaggtctgcttcacctacgacctgttcc 316

Query: 512 tgcatttgaaggccatccaccagtgaatcacctccgctgtgaaaagctaactttcaaca 571

|| | || || || || || || || || || || || || || || || || || || || ||

Sbjct: 317 tgaacctggaaggcaaccgcgccgtgaaccacctgcgctgcgagaagctcaccttcaaca 376

Query: 572 accccac 578

||||||

Sbjct: 377 accccac 383

>gb|L04285.1|HUMEN1. Human germline ENL mRNA, complete cds
Length = 1680

Score = 85.7 bits (43), Expect = 1e-13
Identities = 286/367 (77%)
Strand = Plus / Plus

Query: 212 ccgtgcaggtgaagctggagctggggcaccgcgccaggtgaggaaaaaacccaccgtgg 271
|||| ||||| | | ||||| ||||| || | || || ||||| ||
Sbjct: 17 ccgtccaggtgaggttagagctggggcatcgcgcccaactgcgcaagaagcccaccacgg 76

Query: 272 agggcttcacccacgactggatggtgttcgtacgcggtccggagcacagtaacatacagc 331
|||| ||||| ||||| ||||| || ||||| || ||||| || ||||| |||||
Sbjct: 77 aggggttcactcacgactggatggtgtttgtccgcggccccgagcaatgtgacatccagc 136

Query: 332 actttgtggagaaagtctcttccacttgacgaaagcttccctaggccaaaaagagtgt 391
|||| ||||| || ||||| ||||| ||||| || | || | | |||||
Sbjct: 137 acttcgtggagaaggtggtcttctggctgacgacagcttccccaagcccagacgcgtgt 196

Query: 392 gcaaagatccaccttacaagtagaagaatctgggtatgctggtttcattttgccaattg 451
|||| || || || ||||| || || ||||| ||||| ||||| ||||| || |
Sbjct: 197 gcaaggagccccctacaagtagaggagtcggggtacgctggcttcatcatgcccatcg 256

Query: 452 aagtttatTTTTAAAAACAAGGAAGAACCTAGGAAAGTCCGCTTTGATTATGACTTATTCC 511
| || | || ||||| || || ||||| ||| |||| | || || | |||||
Sbjct: 257 aggtgcacttcaaaaacaaggaggagccgaggaaggtctgcttcacctacgacctgttcc 316

Query: 512 tgcaccttgaaggccatccaccagtgatcacctccgctgtgaaaagctaactttcaaca 571
|| | ||| ||||| | || | ||||| ||||| ||||| || ||||| || |||||
Sbjct: 317 tgaacctggaaggcaacccgccgctgaaccacctgcgctgcgagaagctcaccttcaaca 376

Query: 572 accccac 578
|||
Sbjct: 377 accccac 383

>dbj|D14539.1|HUMLTG19 Human mRNA for LTG19
Length = 1974

Score = 85.7 bits (43), Expect = 1e-13
Identities = 286/367 (77%)
Strand = Plus / Plus

Query: 212 ccgtgcaggtgaagctggagctggggcaccgcgccaggtgaggaaaaaccaccgtgg 271
|||| ||||| | | ||||| ||||| || || | || || ||||| ||
Sbjct: 181 ccgtccaggtgaggttagagctggggcatcgcgcccaactgcgcaagaagcccaccacgg 240

Query: 272 agggcttcaccacgactggatggtgttcgtacgcggtccggagcacagtaacatacagc 331
|||| ||||| ||||| ||||| || || ||||| || ||||| |||||
Sbjct: 241 aggggttcactcacgactggatggtgtttgtccgcggcccccagcaatgtgacatccagc 300

Query: 332 actttgtggagaaagtcgtcttcacttgcacgaaagctttcctaggccaaaaagagtgt 391
|||| ||||| || ||||| ||||| ||||| || | ||| | | | |||||
Sbjct: 301 acttcgtggagaaggtggtcttctggctgcacgacagcttccccaagcccagacgcgtgt 360

Query: 392 gcaaagatccaccttacaagtagaagaatctgggtatgctggtttcattttgccaattg 451
|||| || || || ||||| || || ||||| ||||| ||||| ||||| |||||

Sbjct: 361 gcaaggagccccctacaaagtagaggagtcggggtacgctggcttcatcatgcccatcg 420

Query: 452 aagtttatttttaaaaacaaggaagaacctaggaaagtcgctttgattatgacttattcc 511

| || | || ||||| || || |||| ||| ||| | || | |||

Sbjct: 421 aggtgcacttcaaaaacaaggaggagccgaggaaggtctgcttcacctacgacctgttc 480

Query: 512 tgcattctgaaggccatccaccagtgaatcacctccgctgtgaaaagctaactttcaaca 571

|| | || ||||| | || || |||| |||| |||| | || |||| | || |||||

Sbjct: 481 tgaacctggaaggcaaccgcgctgaaccacctgcgctgcgagaagctcaccttcaaca 540

Query: 572 accccac 578

||||||

Sbjct: 541 accccac 547

C2 Sequence comparison for MLLT3 versus Cosmid 34a5

Sequence 1 gi 4758719 Homo sapiens myeloid/lymphoid or mixed-lineage leukemia (trithorax (Drosophila) homolog); translocated to, 3 (MLLT3), mRNA

Length 3376 (1..3376)

Sequence 2 gi 7212008 Homo sapiens Chromosome 9p22 Cosmid Clone 34a5

Length 40679 (1..40679)



NOTE: The statistics (bitscore and expect value) is calculated based on the size of nr database

Score = 344 bits (179), Expect = 1e-91
 Identities = 181/182 (99%)
 Strand = Plus / Plus



```
Query: 207  gtgttccgtgcaggtgaagctggagctggggcaccgcgccaggtgaggaaaaaacccac 266
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct: 15093 gtgtgccgtgcaggtgaagctggagctggggcaccgcgccaggtgaggaaaaaacccac 15152
```

Query: 267 cgtggagggcttcacccacgactggatgggtgttcgtacgcgggtccggagcacagtaacat 326
|||||
Sbjct: 15153 cgtggagggcttcacccacgactggatgggtgttcgtacgcgggtccggagcacagtaacat 15212

Query: 327 acagcactttgtggagaaagtcgtcttccacttgcacgaaagctttcctaggccaaaaag 386
|||||
Sbjct: 15213 acagcactttgtggagaaagtcgtcttccacttgcacgaaagctttcctaggccaaaaag 15272

Query: 387 ag 388
||
Sbjct: 15273 ag 15274

Score = 310 bits (161), Expect = 4e-81
Identities = 207/207 (100%)
Strand = Plus / Plus



Query: 1 tttggggctgagtttaataagcgagcgagcgagcaagcgagcgcggggggaaaaaggcag 60
|||||
Sbjct: 13476 tttggggctgagtttaataagcgagcgagcgagcaagcgagcgcggggggaaaaaggcag 13535

Query: 61 agaatgtccgcatctaccctccgctcctgggcgcgctctcattcatagcagcctcttca 120
|||||
Sbjct: 13536 agaatgtccgcatctaccctccgctcctgggcgcgctctcattcatagcagcctcttca 13595

Query: 121 tgaattacagctgnnnnnnnnnnnnnnnnnnnnnntaccacacaacaccccagcaaacc 180

|||||
Sbjct: 13596 tgaattacagctgagggggggcgaggaggggggtaccacacaacaccccagcaaacc 13655

Query: 181 tccgggccccaggcatggctagctcg 207

|||||

Sbjct: 13656 tccgggccccaggcatggctagctcg 13682

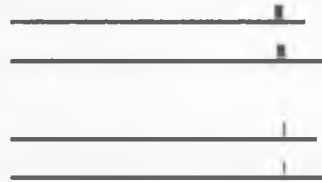
C3 Sequence comparison for MLLT3 versus Cosmid 92f5

Sequence 1 gi 4758719 Homo sapiens myeloid/lymphoid or mixed-lineage leukemia (trithorax (Drosophila) homolog); translocated to, 3 (MLLT3), mRNA

Length 3376 (1..3376)

Sequence 2 gi 4630759 Cosmid clone 92f5 from human chromosome band 9p22

Length 27613 (1..27613)



NOTE: The statistics (bitscore and expect value) is calculated based on the size of nr database

Score = 1048 bits (545), Expect = 0.0
 Identities = 559/559 (100%)
 Strand = Plus / Minus



Query: 762 tagtaccagtttttcaaagcctcacaattaatgaaggagcacaaggaaaaaccttctaa 821
 |||
 Sbjct: 23861 tagtaccagtttttcaaagcctcacaattaatgaaggagcacaaggaaaaaccttctaa 23802

Query: 822 agactccagagaacataaaaagtgccttcaaagaaccttccagggatcacaacaaatcttc 881
|
Sbjct: 23801 agactccagagaacataaaaagtgccttcaaagaaccttccagggatcacaacaaatcttc 23742

Query: 882 caaagaatcctctaagaaacccaaagaaaataaaccactgaaagaagagaaaatagttcc 941
|
Sbjct: 23741 caaagaatcctctaagaaacccaaagaaaataaaccactgaaagaagagaaaatagttcc 23682

Query: 942 taagatggccttcaaggaacctaaacctatgtcaaaagagccaaaaccagatagtaactt 1001
|
Sbjct: 23681 taagatggccttcaaggaacctaaacctatgtcaaaagagccaaaaccagatagtaactt 23622

Query: 1002 actcaccatcaccagtgagacaagataagaaggctcctagttaaaggccgccatttcaga 1061
|
Sbjct: 23621 actcaccatcaccagtgagacaagataagaaggctcctagttaaaggccgccatttcaga 23562

Query: 1062 ttctgaagaactctcagccnnnnnnggaaaaagagtagctcagaggctttattttaaag 1121
| |
Sbjct: 23561 ttctgaagaactctcagccaaaaaaggaaaaagagtagctcagaggctttattttaaag 23502

Query: 1122 ttttctagcgcaccaccactgatactcacttgcttctgctgacaaaaacagataaaaaga 1181
|
Sbjct: 23501 ttttctagcgcaccaccactgatactcacttgcttctgctgacaaaaacagataaaaaga 23442

Query: 1182 taaatctcatgtcaagatgggaaaggtcaaaattgaaagtgagacatcagagaagaagaa 1241
|
Sbjct: 23441 taaatctcatgtcaagatgggaaaggtcaaaattgaaagtgagacatcagagaagaagaa 23382

Query: 1242 atcaacgttaccgccatttgatgatattgtggatcccaatgattcagatgtggaggagaa 1301
|||||
Sbjct: 23381 atcaacgttaccgccatttgatgatattgtggatcccaatgattcagatgtggaggagaa 23322

Query: 1302 tatatcctctaaatctgat 1320
|||||
Sbjct: 23321 tatatcctctaaatctgat 23303

Score = 46.8 bits (24), Expect = 0.071
Identities = 24/24 (100%)
Strand = Plus / Minus

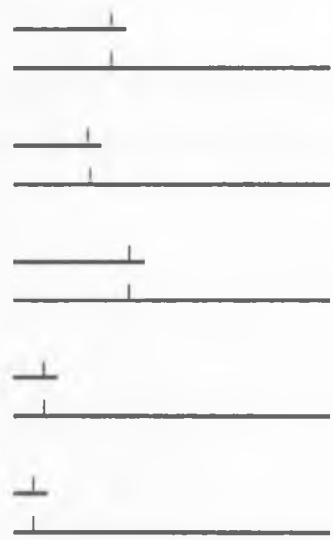


Query: 615 ggaccctaataaggagtattcatac 638
|||||
Sbjct: 24008 ggaccctaataaggagtattcatac 23985

C4 Sequence comparison for MLLT3 versus Cosmid C48

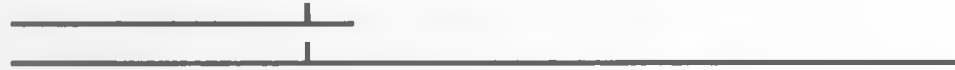
Sequence 1 [gi_4758719](#) Homo sapiens myeloid/lymphoid or mixed-lineage leukemia (trithorax (Drosophila) homolog); translocated to, 3 (MLLT3), mRNA **Length** 3376 (1 .. 3376)

Sequence 2 [gi_4809343](#) Homo sapiens Chromosome 9 Cosmid Clone c48 **Length** 40556 (1 .. 40556)



NOTE: The statistics (bitscore and expect value) is calculated based on the size of nr database

Score = 254 bits (132), Expect = 2e-64
Identities = 132/132 (100%)
Strand = Plus / Minus



Query: 1395 aggtcctttgaggtctataatgaaagatctgcattctgatgacaatgaggaggaatcaga 1454
|||||
Sbjct: 12812 aggtcctttgaggtctataatgaaagatctgcattctgatgacaatgaggaggaatcaga 12753

Query: 1455 tgaagtggaggataacgacaatgactctgaaatggagaggcctgtaaatagaggaggcag 1514
|||||
Sbjct: 12752 tgaagtggaggataacgacaatgactctgaaatggagaggcctgtaaatagaggaggcag 12693

Query: 1515 ccgaagtcgcag 1526
|||||
Sbjct: 12692 ccgaagtcgcag 12681

Score = 198 bits (103), Expect = 1e-47
Identities = 103/103 (100%)
Strand = Plus / Minus



Query: 1524 cagagttagcttaagtgatggcagcgatagtgaaagcagttctgcttcttcaccocctaca 1583
|||||
Sbjct: 10049 eagagttagcttaagtgatggcagcgatagtgaaagcagttctgcttcttcaccocctaca 9990

Query: 1584 tcacgaacctccaccacccttactaaaaaccaacaacaaccag 1626
|||||
Sbjct: 9989 tcacgaacctccaccacccttactaaaaaccaacaacaaccag 9947

Score = 150 bits (78), Expect = 4e-33
Identities = 78/78 (100%)
Strand = Plus / Minus



Query: 1321 tctgaacaaccagtcctgccagctccagctccagctccagctccagcttcacaccatcc 1380
|||||
Sbjct: 14949 tctgaacaaccagtcctgccagctccagctccagctccagctccagcttcacaccatcc 14890

Query: 1381 cagaccaggcaacaaggt 1398
|||||
Sbjct: 14889 cagaccaggcaacaaggt 14872

Score = 144 bits (75), Expect = 2e-31
Identities = 75/75 (100%)
Strand = Plus / Minus



Query: 1625 agattcttgaagtgaaaagtc caataaagcaaagcaaatcagataagcaaataaagaatg 1684
|||||
Sbjct: 4086 agattcttgaagtgaaaagtc caataaagcaaagcaaatcagataagcaaataaagaatg 4027

Query: 1685 gtgaatgtgacaagg 1699

|||||

Sbjct: 4026 gtgaatgtgacaagg 4012

Score = 142 bits (74), Expect = 8e-31

Identities = 74/74 (100%)

Strand = Plus / Minus



Query: 1697 aggcatacctagatgaactggtagagcttcacagaaggttaatgacattgagagaaagac 1756

|||||

Sbjct: 2803 aggcatacctagatgaactggtagagcttcacagaaggttaatgacattgagagaaagac 2744

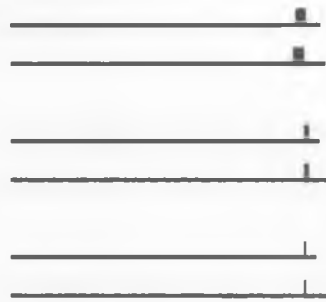
Query: 1757 acattctgcagcag 1770

|||||

Sbjct: 2743 acattctgcagcag 2730

C5 Sequence comparison for MLLT3 versus Cosmid 213a7

Sequence 1	gi <u>4758719</u>	Homo sapiens myeloid/lymphoid or mixed-lineage leukemia (trithorax (Drosophila) homolog); translocated to, 3 (MLLT3), mRNA	Length 3376 (1 .. 3376)
Sequence 2	gi <u>7230795</u>	Homo sapiens Chromosome 9p22 Cosmid Clone 213a7	Length 42210 (1 .. 42210)



NOTE: The statistics (bitscore and expect value) is calculated based on the size of nr database

Score = 2415 bits (1256), Expect = 0.0
Identities = 1331/1332 (99%), Gaps = 1/1332 (0%)
Strand = Plus / Minus



Query: 2045 caaccttcagctttatTTTTtctTTaaagccagtcacatctcttgataaaggagaggtta 2104
 |||
 Sbjct: 39362 caaccttcagctttatTTTTtctTTaaagccagtcacatctcttgataaaggagaggtta 39303

Query: 2105 aagcaaaccagcctcagcggaccactcttctctccaaggaaatccccgggaagagttagc 2164
 |||
 Sbjct: 39302 aagcaaaccagcctcagcggaccactcttctctccaaggaaatccccgggaagagttagc 39243

Query: 2165 ctggatagccttgaaaacaaacaaatcaaacacacacacaagaaaactcaaagaatgtgta 2224
 |||
 Sbjct: 39242 ctggatagccttgaaaacaaacaaatcaaacacacacacaagaaaactcaaagaatgtgta 39183

Query: 2225 tggatcatgtatctctctgtgggtggttcattccacaggacgaatgcatattcaacacac 2284
 |||
 Sbjct: 39182 tggatcatgtatctctctgtgggtggttcattccacaggacgaatgcatattcaacacac 39123

Query: 2285 tgcttattacataactgatctatTTattatcgcatacagatattctaagtcggtgaggg 2344
 |||
 Sbjct: 39122 tgcttattacataactgatctatTTattatcgcatacagatattctaagtcggtgaggg 39063

Query: 2345 aatgacaccatcagacattataagtacttgggtcccgtggatgctctttcaatgcagcacc 2404
 |||
 Sbjct: 39062 aatgacaccatcagacattataagtacttgggtcccgtggatgctctttcaatgcagcacc 39003

Query: 2405 cttgccatcccaagcccagtgaccttactcgtataccgtgccactttccaccaactTTTT 2464
 |||
 Sbjct: 39002 cttgccatcccaagcccagtgaccttactcgtataccgtgccactttccaccaactTTTT 38943

Query: 2465 ccaagtcctttaactcgttgcagctctgtatttccaccttttgttttccagttccagga 2524
 |||
 Sbjct: 38942 ccaagtcctttaactcgttgcagctctgtatttccaccttttgttttccagttccagga 38883

Query: 2525 cacagattatcaactggggggaccaaatagccaccttgattttcttctttgtggctcttt 2584
 |||
 Sbjct: 38882 cacagattatcaactggggggaccaaatagccaccttgattttcttctttgtggctcttt 38823

Query: 2585 tcttgaaagttggggcccagtccttggtgtatccatgtaatgatcttgaccatggtag 2644
 |||
 Sbjct: 38822 tcttgaaagttggggcccagtccttggtgtatccatgtaatgatcttgaccatggtag 38763

Query: 2645 aaaatgcaccaaataggatcatatgaattgctgtctagccttagtcaataaacttgtagg 2704
 |||
 Sbjct: 38762 aaaatgcaccaaataggatcatatgaattgctgtctagccttagtcaataaacttgtagg 38703

Query: 2705 acttttaacaaaagtgtacctgtaaagtctctgaatccagcattggtgagctgtcatca 2764
 |||
 Sbjct: 38702 acttttaacaaaagtgtacctgtaaagtctctgaatccagcattggtgagctgtcatca 38643

Query: 2765 acattcttggtctggttttactggttacaataattaggtgaatatggaagtaaaggcattcc 2824
 |||
 Sbjct: 38642 acattcttggtctggttttactggttacaataattaggtgaatatggaagtaaaggcattcc 38583

Query: 2825 acaggatcatcatttnnnnnnnnngaattctggtcctgttttctnnnnnnnnnnnctgt 2884
 ||| ||| |||
 Sbjct: 38582 acaggatcatcatttaaaaaaaaagaattctggtcctgttttct-aaaaaaaaaactgt 38524

Query: 2885 tgtagaaattcttaatttggatctatatttagtcagagtttcagctttcttcagctgcc 2944
|||||
Sbjct: 38523 tgtagaaattcttaatttggatctatatttagtcagagtttcagctttcttcagctgcc 38464

Query: 2945 agtgtgttactcatctttatcctaaaaatctggaatcagagatTTTTgtttgttcacata 3004
|||||
Sbjct: 38463 agtgtgttactcatctttatcctaaaaatctggaatcagagatTTTTgtttgttcacata 38404

Query: 3005 tgattctcttagacacttttatatttgaaaaaattaaaatctttctttggggaaaaattc 3064
|||||
Sbjct: 38403 tgattctcttagacacttttatatttgaaaaaattaaaatctttctttggggaaaaattc 38344

Query: 3065 ttggttattctgccataacagattatgtattaactttagattcagtggttcaatacctg 3124
|||||
Sbjct: 38343 ttggttattctgccataacagattatgtattaactttagattcagtggttcaatacctg 38284

Query: 3125 tttagttgcttgctaataattccagaaggatttcttgtattggtgaaagacggttgggga 3184
|||||
Sbjct: 38283 tttagttgcttgctaataattccagaaggatttcttgtattggtgaaagacggttgggga 38224

Query: 3185 tggggggannnnnngttcttgttgtaccctgttttgaaactagaaatctgtcctgtgg 3244
|||||
Sbjct: 38223 tggggggatttttggttcttgttgtaccctgttttgaaactagaaatctgtcctgtgg 38164

Query: 3245 catgcaaaagaaagcaaattatTTTTaaaagnnnnnnccaaagtacttttgggtgcatt 3304
|||||
Sbjct: 38163 catgcaaaagaaagcaaattatTTTTaaaagaaaaaaccaaagtacttttgggtgcatt 38104

Query: 3305 attccatcttctccataagtggagaaatgaaaagtaagaacagctcatcttcaaagttt 3364
|||||
Sbjct: 38103 attccatcttctccataagtggagaaatgaaaagtaagaacagctcatcttcaaagttt 38044

Query: 3365 tactagaaattc 3376
|||||
Sbjct: 38043 tactagaaattc 38032

Score = 325 bits (169), Expect = 8e-86
Identities = 169/169 (100%)
Strand = Plus / Minus



Query: 1769 agatcgtgaaccttatagaagaaactggacactttcatatcacaacacacatttgatt 1828
|||||
Sbjct: 39638 agatcgtgaaccttatagaagaaactggacactttcatatcacaacacacatttgatt 39579

Query: 1829 ttgatctttgctcgctggacaaaaccacagtcgtaaaactacagagttacctggaacat 1888
|||||
Sbjct: 39578 ttgatctttgctcgctggacaaaaccacagtcgtaaaactacagagttacctggaacat 39519

Query: 1889 ctggaacatcctgaggatataacaactggatgcatcaagaactattgtg 1937
|||||
Sbjct: 39518 ctggaacatcctgaggatataacaactggatgcatcaagaactattgtg 39470

Score = 60.3 bits (31), Expect = 6e-06
Identities = 31/31 (100%)
Strand = Plus / Minus

Query: 1966 ggttgtgattttttggttcttggtgtttatat 1996
|||||
Sbjct: 39441 ggttgtgattttttggttcttggtgtttatat 39411