

Molecular genetic analysis of paediatric low-grade astrocytoma

Ruth Grace Tatevossian

University College London
and
Cancer Research UK London Research Institute

PhD Supervisor: Professor Denise Sheer

A thesis submitted for the degree of
Doctor of Philosophy
University College London

October 2009

Declaration

I, Ruth Grace Tatevossian, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

The thesis describes the molecular genetic analysis of fifty low-grade paediatric astrocytomas. DNA copy number changes were investigated in paediatric low-grade astrocytomas (WHO grade I and II), using array comparative genomic hybridisation and Affymetrix 250K and 6.0 SNP arrays. A discrete region of DNA copy number gain was identified at chromosome 7q34, primarily although not exclusively in pilocytic astrocytomas of the cerebellum. Further analysis of this region, by PCR and sequencing, demonstrated the presence of gene fusions between *KIAA1549* and *BRAF*. Five *KIAA1549-BRAF* fusion variants were subsequently identified. A further gene fusion between *SRGAP3* and *RAF1* was also found in a single tumour with DNA copy number gain at chromosome 3p25. The fusion genes lacked the auto-inhibitory domains of *BRAF* and *RAF1*. These were replaced in-frame by N-terminal segments of *KIAA1549* and *SRGAP3*, respectively, conferring constitutive kinase activity.

Sequencing confirmed the presence of activating mutations in *KRAS* and *BRAF* in three tumours where gene fusions were not identified. Both gene fusions and activating mutations were shown to cause activation of the ERK/MAP kinase pathway by Western blotting. Further sequencing was performed of *CDKN2A*, *PTEN* and *IDH1/2* to assess the frequency of abnormalities in paediatric low-grade astrocytoma. These genes have previously been found to contain aberrations within adult high-grade astrocytomas. To date, no significant aberrations have been identified in the paediatric astrocytoma samples examined. This confirms previous findings in adult and paediatric astrocytomas, which appear to show distinct molecular changes depending on patient age.

Gene fusions or activating mutations were identified in 100% pilocytic astrocytomas studied and were also found in 28% of grade II astrocytomas. These findings highlight the importance of the ERK/MAPK pathway both in the development of paediatric low-grade astrocytomas and as a possible therapeutic target. Gene fusions may provide a means of molecular classification for pilocytic astrocytomas in the future.

Acknowledgements

I would like to express my warmest thanks and very deepest gratitude to Denise Sheer. She has been a continual source of inspiration, support, guidance and boundless enthusiasm. This work would not have been possible without her dedication to pushing forward the boundaries of knowledge in children's cancer. I am so pleased to have been given the opportunity to work with her.

Many thanks to David Ellison: none of the work presented here would have been possible without him and his team at St Jude. I would like to thank my second supervisor Ian Tomlinson, and my thesis committee member Nic Tapon. I would also like to extend my thanks to Cancer Research UK for supporting me with a clinical research fellowship to enable me to complete my studies.

My most heartfelt thanks to my friends and colleagues; Tim Forshew, Andrew Lawson, Tania Jones, Diego Ottaviani, Chiara Mazzanti, and Paul Mulholland for their endless support and advice over the years. I couldn't have done it without you.

Thank you to all the members of the Human Cytogenetics laboratory; William Ogunkolade, Rowena Carter, Guy Hindley, Jarek Szary, Johan Aarum, Rossitza Christova, Petros Takousis, Radost Vacheva, Alistair Newall and Pei Jun Wu. I would also like to thank Gavin Kelly for his helpful advice and patience, all of the staff at Cancer Research London Research Institute's equipment park for their help and expertise and to Ian Giddings and his staff at the microarray facility at the Institute of Cancer Research in Sutton.

Lastly, I would like to thank my family for their love and encouragement. To my parents, Adolph and Cordelia, my sister and brother-in-law Elsa and Adam, to Victor, Philip, Naomi and Luke, and most particularly of all to James and Sophia: this thesis is dedicated to you.

Table of Contents

Declaration.....	2
Abstract.....	3
Acknowledgements.....	4
Table of Contents.....	5
Table of figures.....	10
List of tables.....	14
List of tables.....	14
Abbreviations.....	16
Abbreviations.....	16
Chapter 1. Introduction.....	18
1.1 Cancer biology – an overview	19
1.2 Genetic and epigenetic changes in cancer – a genomic disease	20
1.2.1 Oncogenes.....	21
1.2.2 Tumour suppressor genes	23
1.3 Epigenetic changes in cancer	24
1.3.1 DNA hypermethylation.....	25
1.3.2 DNA hypomethylation in cancer	26
1.3.3 Post-translational histone modifications	27
1.4 MicroRNAs.....	28
1.5 Chromosomal rearrangements in cancer.....	31
1.6 Cancer – an accumulation of genetic aberrations	35
1.7 Cancer in children – a developmental disease?	37
1.8 Cancer stem cells	38
1.9 An overview of the mitogen-activated protein kinase (MAPK) pathway	40
1.10 Paediatric brain tumours	43
1.10.1 Epidemiology	43
1.10.2 Aetiological factors	44
1.10.3 Early development of the central nervous system.....	47
1.10.4 Histopathological classification of paediatric astrocytomas	47
1.10.5 Clinical presentation and diagnosis.....	49
1.10.6 Treatment	50
1.10.7 Molecular biology of low-grade paediatric astrocytomas.....	51

1.11	Astrocytomas in children and adults.....	53
1.12	Aims of this thesis.....	54
Chapter 2.	Materials and methods	55
2.1	Tumour samples.....	55
2.2	Control DNA samples.....	55
2.2.1	Control RNA.....	57
2.2.2	Control protein.....	57
2.3	Cell lines	57
2.4	General Methods.....	57
2.4.1	Cell culture.....	57
2.4.2	Metaphase spread preparation.....	58
2.4.3	DNA extraction from blood, cell line and tumour samples	59
2.4.4	DNA extraction to create pooled male control DNA.....	59
2.5	RNA extraction from tumour samples.....	59
2.6	Protein extraction from tumour samples.....	60
2.7	DNA and RNA quantification	60
2.8	cDNA synthesis	61
2.9	Genomic DNA whole-genome amplification	61
2.10	Fluorescence in situ hybridisation (FISH)	61
2.11	Interphase fluorescence in-situ hybridisation (i FISH).....	63
2.12	Multiplex fluorescence in situ hybridization (M-FISH).....	63
2.13	Polymerase chain reaction (PCR).....	64
2.13.1	Primer design.....	65
2.13.2	PCR reactions.....	65
2.13.3	Agarose gel electrophoresis of DNA	66
2.14	DNA purification	67
2.14.1	ExoSap	67
2.14.2	Gel purification	67
2.15	Direct sequencing.....	67
Chapter 3.	DNA copy number changes by array CGH	69
3.1	Introduction.....	69
3.2	Aim	72

3.3	Array CGH method.....	72
3.3.1	Array CGH slides.....	72
3.3.2	Random labelling of DNA.....	73
3.3.3	Precipitation of DNA prior to hybridisation.....	75
3.3.4	Preparation of array slide prior to hybridisation.....	76
3.3.5	Pre-hybridisation.....	76
3.3.6	Hybridisation.....	77
3.3.7	Washing.....	78
3.3.8	Drying.....	78
3.3.9	Scanning and data acquisition.....	78
3.3.10	Data analysis.....	82
3.4	Array CGH results.....	85
3.4.1	Male versus female control.....	85
3.5	Medulloblastoma cell lines versus male control.....	86
3.5.1	Characterisation of Daoy by M-FISH and FISH.....	87
3.5.2	Characterisation of Daoy by array CGH.....	87
3.5.3	Characterisation of cell line HTB187 by M-FISH and FISH.....	88
3.5.4	Characterisation of HTB187 by array CGH.....	88
3.5.5	Paediatric low-grade astrocytoma versus pooled male control.....	94
3.5.6	Results of array CGH analysis for paediatric low-grade astrocytomas.....	94
3.5.7	Loss of the terminal end of chromosome 1p.....	95
3.5.8	Gain at 2q11.1.....	95
3.5.9	Loss at 7q22.1.....	96
3.5.10	Comparison of cerebellar samples to non-cerebellar samples.....	96
3.6	Discussion.....	108
Chapter 4.	DNA copy number changes by SNP array analysis.....	118
4.1	Introduction.....	118
4.2	Aim.....	123
4.3	Affymetrix 250K Nsp SNP array.....	123
4.3.1	Materials.....	123
4.3.2	Method.....	124

4.3.3	Analysis.....	125
4.3.4	Results.....	127
4.4	Affymetrix Genome-Wide Human 6.0 SNP array.....	132
4.4.1	Analysis.....	132
4.4.2	6.0 array results.....	134
4.4.3	Comparison of results from SNP arrays and array CGH.....	145
4.5	Discussion.....	149
4.5.1	SNP array investigation of astrocytomas.....	149
4.5.2	Further discrete changes identified by SNP array analysis.....	153
4.6	Summary.....	157
Chapter 5.	Identification and characterisation of fusion genes.....	158
5.1	Introduction.....	158
5.2	Aim.....	159
5.3	Interphase fluorescence in-situ hybridisation (i FISH).....	159
5.3.1	PCR identifies <i>KIAA1549-BRAF</i> fusion.....	160
5.3.2	Sequencing confirms the presence of a <i>KIAA1549-BRAF</i> gene fusion.....	162
5.3.3	Nested PCR identifies three further <i>KIAA1549-BRAF</i> fusions.....	162
5.4	Mapping genomic DNA <i>KIAA1549-BRAF</i> fusion breakpoints.....	167
5.5	Identification of <i>SRGAP3-RAF1</i> fusion in the single pilocytic astrocytoma with 3p25 DNA copy number gain.....	172
5.5.1	Western blotting confirms activation of the MAPK pathway.....	173
5.6	Discussion: RAF fusion genes in paediatric low-grade astrocytoma.....	175
5.6.1	Possible mechanisms for the development of fusions.....	184
5.6.2	MAPK pathway activation in low-grade astrocytomas is confirmed by Western blotting.....	189
Chapter 6.	Candidate gene sequencing in paediatric low-grade astrocytoma.....	190
6.1	Introduction.....	190
6.2	Aim.....	192
6.3	Materials and methods.....	192
6.4	Results.....	192
6.4.1	ERK/MAPK pathway member sequencing.....	192
6.4.2	<i>CDKN2Aα</i> sequencing.....	193

6.4.3	<i>PTEN</i> sequencing.....	196
6.4.4	<i>IDH1</i> and <i>IDH2</i> sequencing.....	198
6.5	Discussion.....	198
6.5.1	MAPK pathway sequencing	198
6.5.2	<i>CDKN2A</i> sequencing	200
6.5.3	<i>PTEN</i> sequencing.....	202
6.5.4	<i>IDH1</i> and <i>IDH2</i> sequencing.....	203
6.6	Conclusion	204
Chapter 7.	Final discussion.....	206
7.1	Review of thesis aims and findings	206
7.1.1	Targeted therapy against the MAPK pathway in low-grade astrocytomas..	211
7.1.2	Lessons from adult glioblastoma studies	212
7.1.3	Fusion genes in cancer	213
Chapter 8.	Appendices.....	215
Appendix A	Reagents	215
	Chemicals.....	215
	Kits.....	216
8.1.1	Loading buffer for gel electrophoresis.....	217
Appendix B	Reagents prepared for array CGH	218
	Hybridisation buffer.....	218
	Array CGH wash 1.....	218
Appendix C.	Primers.....	220
References.	224

Table of figures

Figure 1.1 Mechanisms for tumour suppressor gene inactivation (after Knudson, 1971)...	24
Figure 1.2 MicroRNA generation and function.....	30
Figure 1.3 Mechanisms for the creation of fusion genes in cancer.	32
Figure 1.4 A translocation between chromosomes 9 and 22 at bands 9q34 and 22q11 results in the Philadelphia chromosome.	33
Figure 1.5 Schematic representation of the MAPK pathway, and the possible cellular outcomes from pathway activation.	42
Figure 3.1 G50 spin column preparation.	74
Figure.3.2 Agarose gel image of 12 labelled DNA samples.....	74
Figure 3.3 array CGH scan images.....	80
Figure 3.4 Definition of the background region by GenePix.....	81
Figure 3.5 Composite array CGH whole genome data plot.....	86
Figure 3.6 M-FISH karyotype of HTB186 (Daoy).....	89
Figure 3.7 M-FISH karyotype of cell line HTB187.	90
Figure 3.8 FISH analysis of HTB187 using chromosome 8 paint (red) and <i>MYC</i> probe (green).....	91
Figure 3.9 FISH analysis of HTB187 using chromosome 11 paint (red) and <i>MYC</i> probe (green).....	91
Figure 3.10 Array CGH whole genome plots for medulloblastoma cell lines Daoy and HTB187, obtained using Bioconductor	92
Figure 3.11 Array CGH plots from individual chromosomes from HTB187 (<i>test</i>) compared to male (<i>control</i>), obtained using Bioconductor. Data were smoothed following segmentation.	93
Figure 3.12 Array CGH whole genome plot for tumour sample PA11, showing gains of chromosomes 5 and 6.....	101
Figure 3.13 MA plot using raw array CGH data for sample PA11.	101
Figure 3.14 Array CGH data plots for selected chromosomes from tumour sample PA11.	102

Figure 3.15 Array CGH whole genome plot for tumour sample PMG1, showing gain of chromosomes 6, 8, 11 and 12.....	103
Figure 3.16 MA plot using raw array CGH data for sample PMG1.....	103
Figure 3.17 Array CGH chromosome plots from tumour sample PMG1.....	104
Figure 3.18 Array CGH data plots for discrete regions of gain and loss.....	105
Figure 3.19 Whole genome array CGH plot comparing grouped data from cerebellar tumours with non-cerebellar tumours.	106
Figure 3.20 Array CGH data plots for chromosome 21.....	107
Figure 4.1 Overview of Affymetrix SNP array genotyping.	120
Figure 4.2 Overview of Affymetrix 250K Nsp GeneChip mapping assay.....	124
Figure 4.3 DNA copy number gains and losses in 21 tumours, assessed by 250k Nsp SNP array.	128
Figure 4.4 Copy number views of SNP data showing discrete regions of DNA copy number gain and loss in individual tumour samples, obtained using CNAG.....	129
Figure 4.5 Whole genome DNA copy number view of raw SNP data for sample PA12, showing whole chromosome gains of chromosomes 5 and 6.....	130
Figure 4.6 Chromosome 7 copy number analysis by Affymetrix 6.0 SNP array reveals 7q34 gain in tumour samples PA13, PA5 and PMA1.	137
Figure 4.7 DNA copy number gain at 7q34 in 9/10 representative low-grade astrocytomas, identified by circular binary segmentation.	138
Figure 4.8 DNA copy number gain at 3p25 in PA30.	141
Figure 4.9 DNA copy number loss at 9p21 in samples PXA1 and DA2.....	141
Figure 4.10 DNA copy number changes by SNP 6.0 analysis.	142
Figure 4.11 Array CGH plot for chromosome 7 comparing pooled data from cerebellar samples to non-cerebellar samples.....	147
Figure 4.12 Array CGH clones at the 7q34 region of gain. Figures adapted from Ensembl, release 55.	148
Figure 4.13 Overview of the region of interest within 7q34.....	149
Figure 5.1 Tandem duplication within 7q34 including <i>BRAF</i> , identified by iFISH from paraffin sections.....	160

Figure 5.2. Schematic diagram showing the design of primer placement for the identification of a possible fusion gene.	161
Figure 5.3. 1.5% agarose gel showing results of PCR between <i>KIAA1549</i> and <i>BRAF</i> in sample PA7.	161
Figure 5.4. Sequence analysis of two <i>KIAA1549-BRAF</i> fusions in tumour cDNA.	162
Figure 5.5 1.5% agarose gel showing results of the second stage of the nested PCR.	163
Figure 5.6 Sequence analysis of three further <i>KIAA1549-BRAF</i> fusion variants in tumour cDNA.	164
Figure 5.7 Five <i>KIAA1549-BRAF</i> fusion variants, the thickness of the black lines indicates the frequency of each variant.	166
Figure 5.9 1.5% agarose gel for the identification of <i>KIAA1549-BRAF</i> fusion breakpoints in pooled genomic DNA from 9 tumours.	169
Figure 5.10 Sequence traces for the fusion junctions between <i>KIAA1549</i> intron 16 and <i>BRAF</i> intron 8 in genomic DNA from three tumour samples.	170
Figure 5.11 Sequence traces for the DNA fusion junctions in 9 pilocytic astrocytomas, all containing the <i>KIAA1549-BRAF</i> exon 16-exon 9 fusion.	171
Figure 5.12 Schematic representation of the mapping results for DNA fusion breakpoints within intron 16 of <i>KIAA1549</i> and intron 8 of <i>BRAF</i> in 9 tumours with the <i>KIAA1549-BRAF</i> exon 16-exon 9 fusion.	172
Figure 5.13 Sequence analysis of the <i>SRGAP3-RAF1</i> exon11-exon8 fusion junction in tumour PA30. Exon 11 of <i>SRGAP3</i> has been duplicated, remaining in-frame.	173
Figure 5.14 Schematic representation of the MAPK pathway.	173
Figure 5.15 Western blots showing ERK and phospho-ERK in a representative number of tumour samples and controls.	174
Figure 5.16 Proposed mechanism for the <i>KIAA1549-BRAF</i> fusion to arise by tandem duplication and reinsertion.	176
Figure 5.17 Schematic representations of protein domains within <i>KIAA1549</i> , <i>BRAF</i> , <i>SRGAP3</i> , <i>RAF1</i> and the resulting fusion proteins.	176
Figure 5.18 Non-homologous end-joining (NHEJ) repair of a double strand DNA break.	186

Figure 5.19 Proposed mechanism for FoSTes (fork stalling template switching) to induce genomic rearrangements.	187
Figure 6.1 Activating mutations in <i>KRAS</i> and <i>BRAF</i>	193
Figure 6.2 Homozygous insertion of TT within exon 7 of <i>PTEN</i> in cell line U251.....	196

List of tables

Table 1.1 General classification of oncogenes according to the type of protein produced, with examples for each category (adapted from Croce, 2008)	22
Table 1.2 Examples of cell surface receptors and ligands which may initiate activation of the MAPK pathway.....	42
Table 1.3 Inherited syndromes associated with brain tumours.....	46
Table 2.1 Summary of clinical and pathological details from 50 patients with low-grade astrocytomas. Patient ages are shown in years.	56
Table 3.1 Sex chromosome gains and losses determined by array CGH from tumour samples when compared with pooled male control DNA.	98
Table 3.2 Summary of chromosome gains, losses and point changes, excluding sex chromosomes, found in 20 tumour samples by array CGH.....	99
Table 3.3 Percentage of features meeting inclusion criteria for a ‘good’ feature (>75% spot pixels >2SD above background) for array CGH experiments performed for each sample from Genepix 6.....	112
Table 4.1 Summary of DNA copy number gains and losses identified in 21 low-grade astrocytoma samples by 250K SNP array analysis.....	131
Table 4.2 Segment reporting identified the start and end of the 7q34 region of gain within 23/24 samples containing 7q34 gain by 6.0 SNP analysis.....	139
Table 4.3 Start and end of gained segments at 7q34, identified by circular binary segmentation analysis.	140
Table 4.4 DNA copy number changes in 26 pilocytic astrocytomas by 6.0 SNP array. (- = no abnormality found).....	143
Table 4.5 DNA copy number changes in 11 grade II astrocytomas by 6.0 SNP array. (- = no abnormality found).....	144
Table 4.6 Comparison of whole chromosome gains from 20 tumour samples (11 cerebellar PA, 9 non-cerebellar PA) by array CGH and 250k or 6.0 SNP array analysis.....	146
Table 5.1 <i>KIAA1549-BRAF</i> fusion variants identified in 50 low-grade astrocytomas.	165

Table 5.2 Summary of <i>KIAA1549-BRAF</i> fusion variants found in 50 low-grade paediatric astrocytomas.	166
Table 5.3 Transcript lengths and predicted protein sizes for the five <i>KIAA1549-BRAF</i> fusion variants.	167
Table 5.4 Summary of <i>RAF</i> fusion gene variants.	184
Table.6.1 Summary of results for MAPK pathway member sequencing and fusion screening in 32 pilocytic astrocytomas.	194
Table.6.2 Summary of results for MAPK pathway member sequencing in 18 Grade II astrocytoma samples.	195
Table.6.3 Sequencing results for <i>PTEN</i>	197
Table 8.1 Array CGH clones known to give unexpected ratios	219
Table 8.2 Primers for identification of <i>KIAA1549-BRAF</i> gene fusions in cDNA	220
Table 8.3 Primers for identification of the <i>KIAA1549-BRAF</i> fusion breakpoints in genomic DNA.	221
Table 8.4 Primers for identification of the <i>SRGAP3-RAF1</i> fusion breakpoints in cDNA.	221
Table 8.5 DNA sequencing primers for <i>KRAS</i> , <i>NRAS</i> , <i>HRAS</i> and <i>BRAF</i>	222
Table 8.6 DNA sequencing primers for <i>CDKN2A</i> , <i>PTEN</i> , <i>IDH1</i> and <i>IDH2</i>	223

Abbreviations

ALL	acute lymphoblastic leukaemia
BAC	bacterial artificial chromosome
bp	base pairs
CCD	charged coupled device
cDNA	complementary DNA
cen	centromere
CGH	comparative genomic hybridisation
CCNU	lomustine
DA	diffuse astrocytoma
DAPI	4',6-diamidino-2-phenylindole
DNA	deoxyribonucleic acid
DNase	deoxyribnuclease
EDTA	ethylenediaminetetraacetic acid
FISH	fluorescence <i>in situ</i> hybridisation
FITC	fluorescein isothiocyanate-12-dUTP
GBM	glioblastoma multiforme
HGA/HGG	high-grade astrocytoma/ glioma
Kb	kilobase
kDa	kilodaltons
LOH	loss of heterozygosity
Mb	megabase pairs
M-FISH	multiplex fluorescence <i>in situ</i> hybridisation
mRNA	messenger RNA
NHEJ	non-homologous end joining
PA	pilocytic astrocytoma
PAC	P1 artificial chromosome
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PMA	pilomyxoid astrocytoma

PMG	pilomyxoid glioma
PXA	pleomorphic xanthoastrocytoma
RNA	ribonucleic acid
RNase	ribonuclease
rpm	revolutions per minute
RT	reverse-transcription
SDS	sodium dodecyl sulphate
SNP	single nucleotide polymorphism
SSC	standard saline sodium citrate solution (20%)
TAE	tris acetate EDTA
tel	telomere
TE	tris EDTA
Tris	tris(hydroxymethyl)aminomethane
tRNA	transfer RNA
UCSC	University of California Santa Cruz

Chapter 1. Introduction

Cancer is a genetic disease, which is globally responsible for one in eight deaths (Stratton *et al.*, 2009). Hippocrates (460-370 BC) based the earliest theory on the origins of cancer on the accumulation of an excess of ‘black bile’; imbalances of any of the four humours were thought to be responsible for all illnesses (reviewed by Weinstein) (Weinstein *et al.*, 2008). However, this theory was discarded when anatomic dissection during the Renaissance failed to identify the existence of ‘black bile’. Further credible theories on the origins of cancer did not develop until the nineteenth century, when microscopy enabled Virchow to propose that cancer was a disease of cells. Observations of dividing cancer cells by von Hansemann and Boveri in the late nineteenth and early twentieth centuries linked the presence of chromosomal aberrations to cancer for the first time (Mitelman *et al.*, 2007).

The identification of DNA as the molecular basis of inheritance by Avery *et al.* in 1944, and the elucidation of its molecular structure by Watson and Crick in 1953 suggested that cancer was caused by mutations arising from DNA damage (Watson *et al.*, 1953, Avery *et al.*, 1944). More detailed cytogenetic analysis revealed the presence of specific and recurrent genetic aberrations in cancer, such as the Philadelphia translocation between chromosomes 9 and 22 in chronic myeloid leukaemia (Rowley, 1973). Normal NIH3T3 cells were transformed into cancer cells by the introduction of total genomic DNA from human cancer cells (Krontiris *et al.*, 1981, Shih *et al.*, 1981). These findings led to the discovery of the first activating point mutation to cause human cancer in 1982: the single base substitution from guanine to thymine which leads to a substitution of valine for glycine in the gene *HRAS* (Tabin *et al.*, 1982, Reddy *et al.*, 1982).

The field of cancer genetics has grown rapidly over the last thirty years. The search for aberrations within genes and molecular pathways in the most common form of childhood brain tumour forms the basis for the work presented within this thesis.

1.1 Cancer biology – an overview

Cancer is characterised by the progressive growth of neoplastic cells, which have clonally expanded to create an abnormal tissue (Croce, 2008, Stratton *et al.*, 2009, Vogelstein *et al.*, 2004). All cancers contain aberrations that impair normal physiological processes within cells, causing them to gain new and pathological abilities to develop malignancy. These changes lead to the development of the ‘hallmarks of cancer’ described by Hanahan and Weinberg in 2000: where cancer cells display limitless replicative potential, self-sufficiency in growth signals and insensitivity to signals inhibiting growth, sustained angiogenesis, evasion of apoptosis, chromosomal instability and the ability to invade and metastasise (Hanahan *et al.*, 2000). Each represents the failure of a cellular surveillance system designed to protect against cancer. Several or indeed all of these hallmarks of cancer may be accumulated during the development of cancer, and they may occur in any order. Some tumours may accrue several changes simultaneously. For example, the loss of function of the *TP53* tumour suppressor gene may promote angiogenesis and the development of chromosomal instability in addition to evading apoptosis (Levine, 1997).

All cancers, however diverse they may be, are thought to share a common evolutionary pathway (Cahill *et al.*, 1999, Greaves, 2007). A normal cell is transformed into a cancer cell through the accumulation of genetic and/or epigenetic changes. In general these are somatic changes, but aberrations may occur in germline DNA, leading to hereditary cancer predisposition syndromes such as Li-Fraumeni syndrome, where germline mutations occur in the *TP53* tumour suppressor gene (Malkin *et al.*, 1990). Cancer cells often display genomic instability, allowing the cells to accumulate changes ranging from single point mutations, deletions or amplifications, to complex aberrations including translocations or whole chromosome gains or losses (Lengauer *et al.*, 1998).

1.2 Genetic and epigenetic changes in cancer – a genomic disease

Genetic changes occur often in normal cells, however cellular surveillance systems exist to monitor DNA sequence integrity and facilitate repairs to DNA sequence errors or strand breakages (Hanahan *et al.*, 2000, Loeb, 1991). The tumour suppressor protein p53 plays a vital role here: in response to DNA damage p53 may cause cell cycle arrest to allow repair to occur, or may initiate apoptosis leading to cell death if the damage is severe. Loss of function of the p53 DNA damage repair pathway has been found in the majority of human cancers, highlighting the importance of this pathway in maintaining the integrity of normal cells (Levine, 1997).

Cancers result from modifications of the DNA sequence within oncogenes and tumour suppressor genes, which can occur by many different mechanisms (Croce, 2008, Mitelman *et al.*, 2007). However, cancer may also result from epigenetic modifications affecting levels of gene expression where the DNA sequence is not altered (Esteller *et al.*, 2001). Hence, it may be more correct to term cancer a genomic and epigenomic disease rather than a genetic disease. Epigenetic modifications include DNA methylation and chromatin remodelling (Mathews *et al.*, 2009). Non-coding RNAs including microRNAs have also been identified as important post-transcriptional regulators since they target specific messenger RNAs (mRNAs) for degradation or translational repression, and these may also be regulated epigenetically (Saito *et al.*, 2006).

DNA copy number changes can lead to the activation of oncogenes, inactivation of tumour suppressor genes (TSGs) and alterations in the levels of microRNAs (Croce, 2008). Boveri proposed in 1914 that in every normal cell there was a ‘specific arrangement for inhibiting, which allows the process of division to begin only when the inhibition has been overcome’ and that ‘cells of tumours with unlimited growth would arise if those inhibiting chromosomes were eliminated’ and conversely that cells also contained chromosomes which ‘promote division’, thus stimulating cell growth (Boveri, 1914, as reviewed by

Knudson, 2001(Knudson, 2001, Boveri, 1914)). This theory anticipated the discovery of oncogenes and tumour suppressor genes over 60 years later.

1.2.1 Oncogenes

Oncogenes were initially thought to be transmissible pieces of viral DNA from retroviruses, and were believed to be the underlying cause of all cancer in humans. Rous demonstrated that injection of a cell-free extract from chicken tumour cells induced tumour formation in healthy chickens (Rous, 1911). This extract was found to contain the Rous Sarcoma retrovirus (RSV), which induced the formation of sarcomas in chickens. In fact the transmissible agent was not the virus itself, but a nucleotide sequence encoding the gene *v-src* present in RSV and related to nucleotide sequences found in normal chicken DNA. This led to the discovery of the first oncogene *SRC*, identified by Stehelin *et al.* in 1976 (Stehelin *et al.*, 1976). Stehelin, Bishop and Varmus discovered that oncogenes were normal genes found in many organisms, which are responsible for controlling cell growth and division.

Oncogenes encode proteins that regulate cell proliferation and/or apoptosis, and their products may be broadly classified into six groups, as reviewed by Croce in 2008 (Croce, 2008). These are summarised in Table 1.1. Oncogenes may be activated by genetic aberrations such as translocations creating novel gene fusions (e.g. *EWS-FLII* and *EWS-ERG*) or by the addition of novel enhancer or promoter regions to a gene region (juxtaposition of *MYC* to the human immunoglobulin heavy chain enhancer regions in Burkitt's lymphoma)(Joos *et al.*, 1992, Delattre *et al.*, 1994). Oncogenes may also become activated by gene amplification (*MYCN*) or by activating mutations (*RAF* or *RAS* genes) (Schwab *et al.*, 1983, Downward, 2003). Oncogene activation may confer a growth advantage and/or survival advantage, leading to cancer development.

Category	Examples	Role of protein
Transcription factors	<i>JUN</i>	Forms AP-1 transcription factor as a dimer with Fos transcription protein, increasing proliferation in osteosarcoma, liver and skin cancer (Shaulian <i>et al.</i> , 2002, Shaulian <i>et al.</i> , 2001)
	<i>EWS</i>	Translocations create fusion genes between <i>EWS</i> and multiple partners; the resulting fusion proteins drive Ewing's sarcoma formation (Delattre <i>et al.</i> , 1994)
	<i>MYC</i>	Involved in translocations with <i>IGH</i> in Burkitt lymphoma and in numerous other malignancies. Amplification of <i>MYC</i> is seen in medulloblastoma and neuroblastoma (Joos <i>et al.</i> , 1992, Schwab <i>et al.</i> , 1983)
Chromatin remodelling	<i>MLL</i>	<i>MLL</i> may fuse with multiple partner genes, creating fusion proteins which modulate nucleosomes and histones and deregulate homeobox genes in acute leukaemias (ALL and AML) (Nakamura <i>et al.</i> , 2002)
Growth factors	<i>PDGF</i>	Overexpression of platelet derived growth factor (PDGF) induces unregulated cell growth, for example in glioblastoma (Hermansson <i>et al.</i> , 1988)
Growth factor receptors	EGFR	Deletion of the ligand-binding domain or mutations in the epidermal growth factor receptor (EGFR) leads to constitutive activation of the receptor, as seen in glioblastoma (Huang <i>et al.</i> , 1997)
Signal transduction: 1) Tyrosine kinases	<i>ABL</i>	The <i>ABL-BCR</i> fusion gene in chronic myeloid leukaemia encodes a constitutively activated tyrosine kinase (Lugo <i>et al.</i> , 1990)
	2) Serine/threonine kinases	<i>AKT</i>
Apoptosis regulation	<i>BCL2</i>	<i>BCL2</i> is constitutively expressed after forming a fusion gene with the <i>IGH</i> locus in follicular lymphomas. <i>BCL2</i> encodes a cytoplasmic protein that inhibits apoptosis, increasing cell survival (Tsujimoto <i>et al.</i> , 1984, Tsujimoto <i>et al.</i> , 1985)

Table 1.1 General classification of oncogenes according to the type of protein produced, with examples for each category (adapted from Croce, 2008)

1.2.2 Tumour suppressor genes

Knudson first proposed the concept of tumour suppressor genes from studying the incidence of retinoblastoma (Knudson, 1971). Hereditary retinoblastoma is an autosomal dominant condition, generally leading to bilateral tumour development in the retina of children. Sixty percent of retinoblastoma cases are hereditary, and of these 80% arise bilaterally. Retinoblastoma may also occur *de novo*; these cases are always unilateral. Using incidence data for bilateral and unilateral retinoblastoma, Knudson was able to generate models for the number of mutations required to develop disease. He predicted that bilateral retinoblastomas (which already contained a germline mutation) required one further mutation to arise, and that unilateral retinoblastomas required two somatic mutations to occur – thus proposing the ‘two-hit hypothesis’. These data were supported by the knowledge that children born to a parent affected by retinoblastoma did not always develop retinoblastoma themselves, but could have affected children. Hence the concept arose that both alleles of a tumour suppressor gene must be inactivated before tumorigenesis can occur. This prediction was confirmed by the identification of recessive mutations in the two alleles of the retinoblastoma gene (*RB*), the first tumour suppressor gene to be cloned (Friend *et al.*, 1986).

Inactivation of tumour suppressor genes following the ‘first-hit’ may occur by a variety of different genetic methods, as shown in Figure 1.1. Epigenetic changes such as DNA hypermethylation may also inactivate tumour suppressor genes.

TP53 was initially identified as a tumour suppressor gene in colorectal cancer. Studies using SV40 virus for malignant transformation of cells identified *TP53*, and analysed the coding sequence. Mutations in *TP53* were later found in two colorectal carcinomas: both of which had deletions at 17p13 and both were shown to express *TP53* from the remaining allele (Baker *et al.*, 1989). Germline mutations in *TP53* are found in patients with Li-Fraumeni syndrome, a hereditary syndrome that predisposes to numerous tumours including breast, lung, colorectum and brain. The p53 protein acts as a cellular checkpoint, integrating responses through the cell cycle and DNA repair processes and balancing levels of cell growth and apoptosis in response to internal and external signals (reviewed by Finlay in 1989 and Knudson in 2001) (Knudson, 2001, Finlay *et al.*, 1989). Inactivation of *TP53* and *RB* or inactivation within their genetic pathways has

been implicated in the development of the majority of human cancers. *TP53* is part of a complex cellular surveillance system for monitoring DNA damage, inducing cell cycle arrest to enable DNA repair to be completed. If the damage cannot be repaired apoptosis is triggered. Absence of *TP53*, *BRCA1* and other so-called “caretaker” genes allows DNA damage to accumulate, thus facilitating the multi-step process of tumorigenesis (Deng *et al.*, 2003, Jackson *et al.*, 2009).

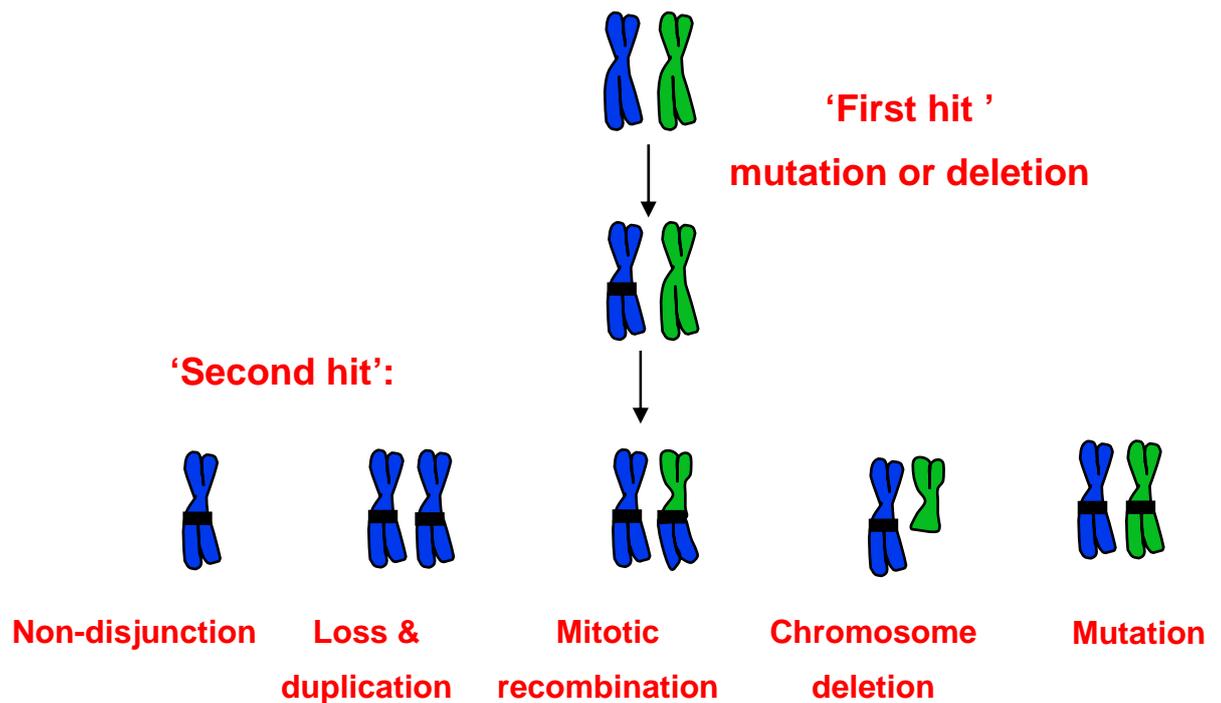


Figure 1.1 Mechanisms for tumour suppressor gene inactivation (after Knudson, 1971).

1.3 Epigenetic changes in cancer

Epigenetic changes are changes affecting gene expression which are not due to alterations in DNA sequence, but are instead related to modifications to the DNA molecule (for example by methylation) or to histone proteins which package the DNA within the nucleus (Pogribny *et al.*, 2009, Mathews *et al.*, 2009).

1.3.1 DNA hypermethylation

DNA hypermethylation is the most fully characterised epigenetic change in cancer. In cancer epigenetics, the relative terms ‘hyper-’ and ‘hypo-methylation’ are used in comparison to normal non-malignant tissue. DNA methylation is known to vary between different tissue types in humans, and also between the same tissues from different mammalian species (Gama-Sosa *et al.*, 1983, Ehrlich *et al.*, 1982). Cultured cells from tumours also display distinct and changed methylation profiles from non-cultured cells (Smiraglia *et al.*, 2001). Therefore, for cancer studies it is most useful to compare methylation profiles between tumour and matched tissue-specific controls. Where this is not possible comparisons should be made between tumour and a range of normal tissues (Ehrlich, 2002).

In DNA methylation, a methyl group is covalently added to the fifth carbon atom in the cytosine pyridine ring by the action of DNA methyltransferases, resulting in the formation of 5-methylcytosine (Chiang *et al.*, 1996). This only occurs where cytosine residues are situated next to a guanine nucleotide; here the dinucleotide is termed CpG where ‘p’ denotes a phosphate separating the two bases in the same sequence (unlike the complementary base pairing of cytosine to guanine bound by hydrogen bonds in the DNA double helix) (Esteller *et al.*, 2002). The majority of CpG dinucleotides are found in unexpressed heterochromatin, however large regions of CpGs, or ‘CpG islands’ are often found within the 5’ regions of genes, containing the promoter, untranslated region and sometimes the first exon (Bird, 1986). Around 75% of gene promoters contain CpG islands, comprising ~ 7% of all CpGs in the genome (Rollins *et al.*, 2006). The majority of CpG islands are not methylated in normal cells.

CpG island methylation is a method of silencing genes at times when their expression is not needed (Esteller *et al.*, 2002). DNA methylation is an essential mechanism used to control normal cellular functions and development. For example, DNA hypermethylation is used to inactivate the genes on one of the two X chromosomes in women (Singer-Sam *et al.*, 1993). In cancer, hypermethylation of CpG islands may function in the same way as a tumour suppressor gene, by switching off gene expression when the gene should be active. Hypermethylation of CpG islands has been found to downregulate genes within the *RB* and *TP53* tumour suppressor gene pathways,

including $p16^{INK4\alpha}$ and $p14^{ARF}$ (Esteller *et al.*, 2000, Robertson *et al.*, 1998, Sakai *et al.*, 1991).

Hypermethylation leading to the inactivation of O-6-methylguanine-DNA methyltransferase (*MGMT*), which encodes a DNA repair protein, has important therapeutic implications in glioblastoma and may be used as a biomarker to direct patient treatment (Nagarajan *et al.*, 2009, Hegi *et al.*, 2005). Reduced expression of *MGMT* due to promoter hypermethylation leads to a defect in DNA repair, thus increasing the sensitivity of cancer cells to DNA damage by alkylating agents such as temozolamide. Newly diagnosed glioblastoma patients with hypermethylation of *MGMT* respond well to treatment with temozolamide and radiotherapy and have been shown to derive a survival benefit, which is not seen in patients with unmethylated *MGMT* (Hegi *et al.*, 2005).

1.3.2 DNA hypomethylation in cancer

Global DNA hypomethylation is a common finding in many cancers, where malignant cells may contain 20-60% less 5-methylcytosine than normal control tissue (Gama-Sosa *et al.*, 1983, Feinberg *et al.*, 1983). The loss of DNA methylation appears to be mainly from the coding regions and introns of genes, and regions of repetitive DNA sequence. The mechanism for de-methylation in cancer is currently unknown, but it is probable that multiple processes are involved. Global hypomethylation has been found in ~ 80% human glioblastomas (Cadieux *et al.*, 2006).

Levels of DNA hypomethylation are known to increase progressively with age. DNA damage, in particular when caused by oxidative stress or exposure to bioflavonoids, alcohol and cigarettes may impair the function of DNA methyltransferases leading to DNA hypomethylation (Weitzman *et al.*, 1994, Valinluck *et al.*, 2004, Lee *et al.*, 2005b, Liu *et al.*, 2007, Seitz *et al.*, 2007).

The degree of DNA hypomethylation appears to correlate with tumour grade, and has been used in some cancers as a prognostic biomarker. Demethylation of DNA is apparent during premalignant stages of urothelial and ovarian cancers, and the degree of hypomethylation increases as the tumour progresses (Widschwendter *et al.*, 2004,

Nakagawa *et al.*, 2005, Watts *et al.*, 2008). Hypomethylation of repetitive satellite sequences and hypermethylation of promoter CpG islands has been used to distinguish between different grades of ovarian cancer (Watts *et al.*, 2008). These data suggest that DNA hypo- and hypermethylation are implicated in the process of tumorigenesis, rather than being a result of damage to the cell by malignant transformation.

Hypomethylation may also be implicated in creating a suitable environment for genomic instability eventually leading to aneuploidy, a common feature of many cancers (Eden *et al.*, 2003). Hypomethylation affects repetitive segments of DNA, found at centromeres, pericentromeric or subtelomeric regions and repetitive elements such as L1 (long interspersed nuclear elements) and Alu repeats (Bestor, 2005, Howard *et al.*, 2008). Hypomethylation of these regions may favour the formation of chromosomal aberrations and has been shown to lead to genomic instability in mice.

Finally, a small number of genes have been identified which show aberrant activation due to promoter hypomethylation, and are therefore highly expressed in cancer (reviewed by Pogribny (Pogribny *et al.*, 2009)). Melanoma antigen family A-1 (*MAGE-1*) gene expression, from the cancer/testis/melanoma antigen family is silenced in normal tissues including the brain; it is usually active only in the testis and placenta (Liu *et al.*, 2004, Cadieux *et al.*, 2006). Increased expression of *MAGE-1* due to specific *MAGE-1* promoter hypomethylation has been found in glioblastomas and glioblastoma cell lines. These increased expression levels also correlated with overall levels of global hypomethylation, but were not associated with increased patient age, indicating that promoter hypermethylation is a regulatory mechanism for gene silencing (Cadieux *et al.*, 2006).

1.3.3 Post-translational histone modifications

DNA is coiled tightly around nucleosome proteins to form chromatin, a DNA-protein complex that enables the entire genome to be packaged within the nucleus (reviewed by Mathews and Esteller (Esteller, 2008, Mathews *et al.*, 2009)). Each nucleosome is composed of eight core histone molecules, with two from each subtype (H2A, H2B, H3 and H4) (Kornberg, 1974). The histone octamer forms a globular structure (around which DNA is coiled), and a flexible histone tail projects from the surface. This histone

tail may be modified by numerous means including methylation, acetylation and phosphorylation, causing conformational changes in the histone structure or charge and hence the nucleosome structure. The conformational changes following histone modifications may allow the chromatin to become highly condensed (heterochromatin) where gene transcription is silenced, or more 'open' and transcriptionally active (euchromatin) (Adkins *et al.*, 2004). Histone modifications have been shown to affect DNA repair and gene transcription, and may be a method of storing information (Jenuwein *et al.*, 2001). The type and site of the modified residue may also influence functional outcome. For example transcriptional activation results from methylation of H3 at lysine position 4, while transcriptional repression results from methylation of lysines at positions 9 or 27 in H3 (Karpf *et al.*, 2005). These and other observations have suggested that a 'histone code' may form an additional regulatory level for the expression of the genetic code (Jenuwein *et al.*, 2001).

Studies of histone modifications in cancers are in their infancy, however initial findings suggest that these may become as significant as genetic alterations in the future. Mutations have been found in numerous genes involved in modifying histones, including histone deacetylases, methyltransferases and demethylases in glioblastomas, by a recent large-scale sequencing study (Parsons *et al.*, 2008). For example, *BM11* is a member of the polycomb complex that regulates histone methylation: deletions and amplifications of *BM11* are frequent findings in adult astrocytomas, and deletions of *BM11* may be associated with poor prognosis in some patients (Hayry *et al.*, 2008, Nagarajan *et al.*, 2009).

1.4 MicroRNAs

MicroRNAs (miRNAs) have recently been recognised as key regulatory factors in cancer development (Croce, 2008). Unlike oncogenes and tumour suppressor genes, which encode proteins, microRNAs encode single strands of RNA between 17-27 nucleotides long (Visone *et al.*, 2009). MiRNAs were first identified in *C. elegans* in the early 1990s, where small complementary transcripts of the gene *lin-4* were found to downregulate the function of *lin-14* mRNA via an antisense RNA interaction (Lee *et al.*, 1993, Wightman *et al.*, 1993). These small transcripts of *Lin-4* were subsequently revealed to be miRNAs.

Mature miRNAs are formed from longer primary single-strand RNA transcripts, by multiple cleavage steps within the nucleus and cytoplasm (Figure 1.2) (Chang *et al.*, 2007). A miRNA strand is taken up by the RNA-induced silencing complex (RISC). If there is perfect sequence complementarity the miRNA-RISC complex may cause cleavage of the target mRNA, if the sequence is partially complementary mRNA translation may be repressed or the mRNA may be degraded. There are now estimated to be between 700-1000 miRNAs in humans. Bioinformatic approaches have predicted that each miRNA may downregulate many hundreds of mRNAs, and estimate that ~30% of all protein coding mRNAs in humans may be regulated by miRNAs (Sethupathy *et al.*, 2006). Aberrant levels of miRNA expression are present in many diseases including cancer (Visone *et al.*, 2009). They may have similar effects to oncogenes or tumour suppressor genes by targeting the messenger RNAs encoding proteins implicated in tumorigenesis.

Over-expression of *miR-155* has been found in B cell lymphomas and breast, lung, colon and thyroid cancers (Eis *et al.*, 2005, Volinia *et al.*, 2006). Transgenic mice overexpressing *miR-155* in B cells display rapid polyclonal pre-B cell expansion in spleen and bone marrow, and go on to develop pre-B cell ALL and high-grade B-cell lymphomas (Costinean *et al.*, 2006). The polyclonal nature of the malignancies developed by these mice indicates that expression of *miR-155* alone may be sufficient for cancer development, hence *miR-155* may be acting as an oncogene. Recently, promoter hypermethylation has been shown to inactivate genes encoding miRNAs, showing that epigenetic modifications and miRNAs may work in tandem to regulate gene function (Lujambio *et al.*, 2007).

Recent studies have shown the importance of microRNAs in tumorigenesis, progression and metastatic processes in cancer, including brain tumours (reviewed by Nicoloso (Nicoloso *et al.*, 2008)). Significant overexpression of *miR-21* has been found in glioblastoma cell lines and tumour samples when compared with control samples (Chan *et al.*, 2005). Knockdown of miR-21 using antisense oligonucleotides and sequence specific miRNA inhibitors increased cell death by apoptosis in glioblastoma cell lines, suggesting that miR-21 upregulation has an anti-apoptotic effect.

It is hoped that microRNAs may also be used therapeutically in the future, by selectively targeting miRNAs or anti-miRNAs to alter levels of gene expression (Visone *et al.*, 2009). Investigation of the roles played by microRNAs will thus form an important part of future cancer studies.

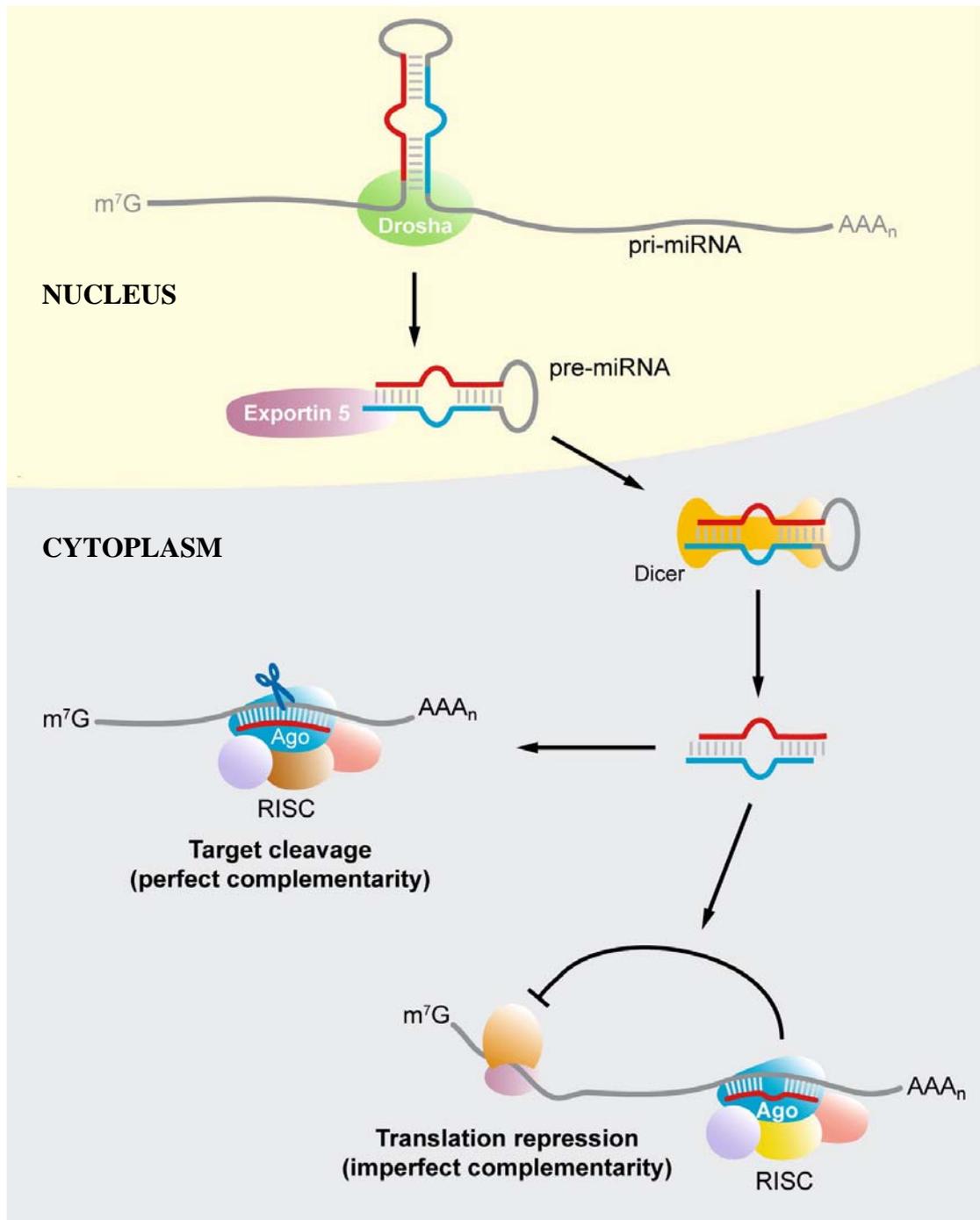


Figure 1.2 MicroRNA generation and function

Figure adapted from Chang *et al.* (Chang *et al.*, 2007)

1.5 Chromosomal rearrangements in cancer

Normal human somatic cells contain 46 chromosomes, and are described as diploid (Heim *et al.*, 2009). Each cell normally contains two homologues of each chromosome and two sex chromosomes, one of maternal and one of paternal origin. The short and long arms of chromosomes are termed 'p' and 'q', respectively. Chromosome rearrangements in cancer may involve changes in the numbers of chromosomes, termed aneuploidy, or aberrations that change the chromosome structure, for example the exchange of genetic material between chromosomes, termed translocations. Translocations may occur within the same chromosome or between different chromosomes. Chromosome segments may be inserted into new positions, inverted (the segment is rotated by 180°), duplicated or deleted. An isochromosome is formed where either the long or short arm of a chromosome is lost, with duplication of the remaining arm. Translocations may be un-balanced with gain or loss of genetic material, or balanced (reciprocal) where no net change occurs. Amplified chromosomal regions may be seen as homogeneously staining regions (HSRs) within a chromosome on imaging, or as small chromosomes known as double minutes.

Chromosomal aberrations have been described in over 55,000 cases of cancer so far, and are catalogued within the Mitelman Database of Chromosome Aberrations in Cancer (<http://cgap.nci.nih.gov/Chromosomes/Mitelman>). Translocations are common in cancer, and are not randomly distributed throughout the genome (Mitelman *et al.*, 2007). Specific chromosomes, regions or chromosomal bands appear to be preferentially involved, and some are highly specific for a particular cancer type (Heim *et al.*, 2009). Translocations may initiate cancer formation by joining together segments of two different genes at the fusion breakpoint (Figure 1.2). This leads to the overexpression of an apparently normal gene at one of the breakpoints, or to the creation of a hybrid fusion genes (e.g. *ABL-BCR* and *BCL2-IGH*, Table 1.1) reviewed by Mitelman (Mitelman *et al.*, 2007).

Fusion genes appear to occur at an early stage in cancer development, as they usually correlate closely with specific tumour types. In animal models, insertion of a fusion construct has been shown to produce the same malignant diseases found in humans with the same fusion, for example insertion of an *IGH-MYC* fusion into mice led to the development of B-cell lymphomas (Adams *et al.*, 1985, Daley *et al.*, 1990). Successful

treatment has been shown to correlate with a decrease or absence of detectable fusion genes, reflecting a reduction in the cancer cells, as shown by monitoring *BCR-ABL* in chronic myeloid leukaemia (Oehler *et al.*, 2003, Deininger *et al.*, 2005). Finally, silencing fusion transcripts, using RNA interference *in vitro*, for example by targeting *BCR-ABL* and *MLL-AF4*, has been shown to reduce tumour proliferation and reverse tumorigenicity (Scherr *et al.*, 2003, Thomas *et al.*, 2005).

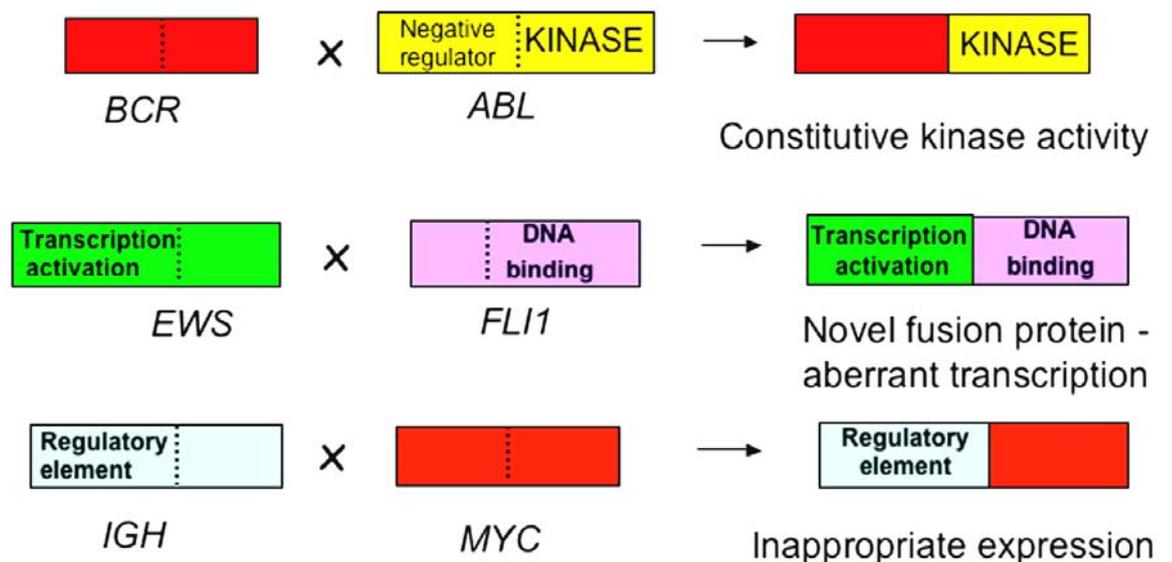


Figure 1.3 Mechanisms for the creation of fusion genes in cancer.

Figure adapted from Sheer and Shipley, 2005 (Sheer *et al.*, 2005).

The first specific and recurrent translocation to be identified in human cancer was a translocation between chromosomes 9 and 22 that was identified by Nowell and Hungerford in 1960, as an unusually small chromosome in chronic myeloid leukaemia (CML) (Nowell *et al.*, 1960). As the discovery was made in Philadelphia, the small chromosome was termed the Philadelphia (Ph) chromosome. It was subsequently shown to be derived from a reciprocal translocation between chromosomes 9 and 22 (Rowley, 1973). This is denoted as t(9;22)(q34;q11), describing the exact position of the translocation breakpoint, at chromosome band q34 on chromosome 9, and at 22q11 (Figure 1.3).

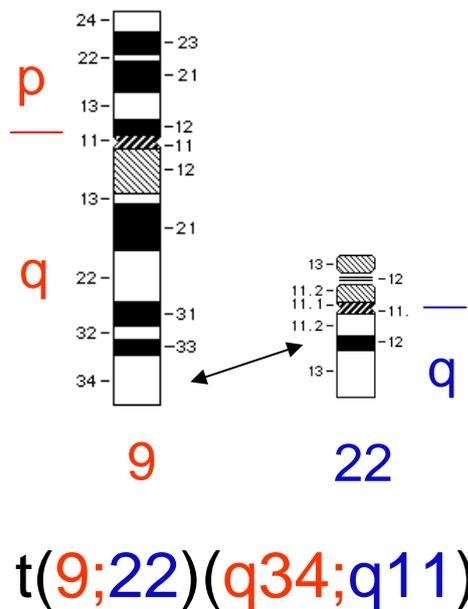


Figure 1.4 A translocation between chromosomes 9 and 22 at bands 9q34 and 22q11 results in the Philadelphia chromosome.

Molecular analysis of the breakpoint in the 1980s revealed a fusion between the genes *ABL* (v-abl Abelson murine leukemia viral oncogene homolog 1) on chromosome 9 and *BCR* (breakpoint cluster region) on chromosome 22 (Groffen *et al.*, 1984). *ABL* encodes a nuclear tyrosine kinase, the activity of the protein is negatively regulated by the SH3 domain. In the *BCR-ABL* fusion, the first exon of *ABL* is replaced with the first twelve or thirteen exons of *BCR*, encoding a chimeric serine tyrosine kinase (Heim *et al.*, 2009). The SH3 regulatory domain from *ABL* is retained within the fusion, but remains in the inactive form; deregulated kinase activity is encoded by a domain within the first exon of *BCR* (Meyn *et al.*, 2006, Lugo *et al.*, 1990). The *BCR-ABL* protein is located in the cytoplasm, and displays deregulated and constitutive kinase activity leading to increased cellular proliferation and leukaemic transformation (Lugo *et al.*, 1990). The *BCR-ABL* fusion alone is sufficient to cause CML, and has been successfully targeted by a specific inhibitor of the *BCR-ABL* tyrosine kinase, imatinib mesylate (Gleevec) (Druker *et al.*, 2001, Sawyers, 2004). This was such an outstanding success that Imatinib is now used as a paradigm to illustrate how other kinase-driven cancers should be treated. Imatinib also targets the kinases KIT and PDGFR in addition to ABL, and is additionally used to treat gastro-intestinal stromal tumours.

Balanced translocations between the genes *EWS* and *FLI1* are found in 90-95% Ewing's sarcomas and produce novel in-frame fusion proteins containing the N-terminal portion from *EWS* and the DNA-binding C-terminal domain from *FLI-1* (Ordonez *et al.*, 2009, Arvand *et al.*, 2001). These novel proteins confer increased transcriptional activation in multiple signalling pathways critical for cellular growth and proliferation and inhibition of apoptosis and senescence, leading to malignant transformation. The presence of *EWS-FLI1* fusions by FISH or RT-PCR may be used as diagnostic tests for Ewing's sarcomas (Delattre *et al.*, 1994). The presence of specific fusion types may also be used to provide prognostic information, for example patients with a fusion between *EWS* exon 7 and *FLI-1* exon 6 are reported to have an improved disease-free survival rate than patients with different fusion types (reviewed by Riley, (Riley *et al.*, 2003)). These fusions are attractive targets for molecularly targeted agents, due to their specificity for Ewing's sarcoma.

The t(8;14)(q24;q32) translocation in Burkitt lymphoma leads to deregulation of the oncogene *MYC* (Kuppers, 2005). The fusion causes the first exon of *MYC* at chromosome band 8q24 to be replaced with the regulatory enhancer region for the immunoglobulin heavy chain gene (*IGH*) at chromosome band 14q32 causing constitutive expression of *MYC*. This mechanism of gene expression driven by immunoglobulin regulatory enhancers is found in over two thirds of mature B-cell lymphomas.

The majority of translocations have been identified in leukaemias and sarcomas. This has been proposed to be mainly because of the technical difficulties in obtaining karyotypes from carcinomas (Heim *et al.*, 2009). However, the balance may shift in the future, as new technologies are developed that allow carcinomas to be investigated at extremely high resolution. Gene fusions have recently been identified in prostate, lung and breast cancers using genome-wide massively parallel sequencing and integrative transcriptome sequencing (Ruan *et al.*, 2007, Tomlins *et al.*, 2007, Campbell *et al.*, 2008). It is possible that all cancers may be found to contain translocations in the future.

Translocations are not always associated with cancers, as they have also been identified in individuals without clinical evidence of cancer (Heim *et al.*, 2009). Patients who have been successfully treated for leukaemia may retain detectable fusion genes for many years without suffering disease relapse. In identical twin studies where one child is

affected with leukaemia the concordance rate is ~10% for the second twin, despite both being born with the same gene fusion in their lymphocytes (Greaves *et al.*, 2003). *TEL-AML1* and *AML1-ETO* fusion genes have been identified in the Guthrie blood spots of babies, however only around 1% subsequently go on to develop leukaemia (Mori *et al.*, 2002). These data indicate that the presence of a fusion gene may not be sufficient in itself for leukaemia to develop. However, *BCR-ABL* alone may drive CML as described above, and expression of *EWS-FLI-1* or the *FUS-CHOP* fusion proteins alone have been shown to transform bone-marrow derived mesenchymal cells into sarcoma-like tumour cells in mice (Riggi *et al.*, 2005, Riggi *et al.*, 2006). It appears, therefore, that some tumour-specific fusions are sufficient in themselves for the initiation of tumorigenesis. An important part of future studies will be to determine which chromosomal aberrations are primary ‘drivers’ of the malignant process, conferring a selective advantage, and which are acquired as secondary changes or ‘passengers’.

The mechanism(s) by which chromosomal translocations arise are currently unknown, however it is likely that DNA double strand breaks occur prior to their formation (Mitelman *et al.*, 2007). Chromosomes have been shown to occupy a distinct location within interphase nuclei (Cremer *et al.*, 2001, Bolzer *et al.*, 2005). Hence, translocations must occur between chromosomes situated near or adjacent to each other. However, the nucleus is a dynamic rather than a static structure. It is probable that spatial proximity of chromosomes or genomic regions may be related to specific periods of cellular proliferation or periods of increased gene expression (Soutoglou *et al.*, 2008). There are inherited cancer-predisposition syndromes where DNA has an increased propensity to break, including ataxia telangectasia and Nijmegen breakage syndrome, where translocations are more likely to occur (Heim *et al.*, 2009).

1.6 Cancer – an accumulation of genetic aberrations

Cancers are thought to arise from the accumulation of genetic aberrations over time, following principles of Darwinian evolution (Cahill *et al.*, 1999, Greaves *et al.*, 2003, Greaves, 2007). Cells are constantly acquiring genetic changes during their lifetime, and natural selection occurs as a result of these genetic changes. Some genetic changes may be detrimental to the cell, and undergo repair or bring about apoptosis. However, some changes will confer a selective advantage to the cell, and this population will then

undergo expansion. Genetic changes can therefore be seen as rate limiting steps within the development of a cancer.

Mathematical models, using incidence data based on patient ages, have proposed that common epithelial carcinomas seen in adults require between 5-7 genetic changes to develop (Armitage *et al.*, 2004, Armitage *et al.*, 1954). Most carcinomas display aneuploidy, where cells contain more or fewer than the normal 46 chromosomes (Heim *et al.*, 2009). Different cells within the same tumour may also contain different chromosome numbers. Some colorectal carcinomas display chromosomal instability (CIN), where whole chromosomes may be gained or lost during cell divisions or microsatellite instability (MSI), which is associated with inactivations in the DNA mismatch repair genes *hMLH1* and *hMSH2* (Lengauer *et al.*, 1997, Lynch *et al.*, 2003). A further subset of colorectal carcinomas contains neither MSI nor CIN (Walther *et al.*, 2008). MSI is seen in around 15% of colorectal cancers, these tumours usually display diploid or near diploid karyotypes. Aneuploidy associated with CIN is found in around 70% of colorectal cancers, and is associated with a poor prognosis for patients when compared to MSI.

Aneuploidy is not always associated with a worse outcome, however. Patient stratification according to ploidy is well established in acute lymphoblastic leukaemia, and here patients with high hyperdiploidy (51-65 chromosomes) respond better to chemotherapy than those with hypodiploidy (<46 chromosomes) (Heerema *et al.*, 1999). Trisomy 8 is the most common chromosomal aberration found in acute myeloid leukaemia (AML), and may be the only change apparent on karyotyping in a small subset of patients (Heim *et al.*, 2009). Trisomy 8 may also be found in some cells of apparently healthy individuals, not all of whom go on to develop AML (Welborn, 2004). The significance of aneuploidy in cancer remains unclear; deleterious effects are not solely due to gene dosage effects. It is uncertain whether aneuploidy represents an essential process for the development of some tumours, or whether these are secondary changes accrued over time as a result of genetic instability.

1.7 Cancer in children – a developmental disease?

Carcinomas comprise the majority of adult cancers, but are rarely seen in children. Over 50% of childhood cancers are leukaemias and central nervous system tumours (Grovas *et al.*, 1997). The accumulation of multiple genetic aberrations over time is a central concept within adult cancer biology, proposed to explain the increased incidence of cancer with increasing age (Armitage *et al.*, 1954). Children's cancers develop over a much shorter time, and some are even present at birth. It seems likely, therefore, that some childhood cancers may require fewer genetic changes to occur, and that some of these changes are present within the germ-line or occur in utero (Scotting *et al.*, 2005, Knudson, 2001).

Solid tumours in children may arise as a consequence of failures within the normal processes of development, and hence cancer may be seen as a 'developmental disease' in this context. For example, the identification of mutations in the patched gene (*PTCH*) in Gorlin's syndrome (an autosomal dominant disorder where patients show increased risk of nevoid basal cell carcinoma and medulloblastoma) led to the discovery that the proliferation of granule cells during normal development of the cerebellum is regulated by the sonic hedgehog (SHH) signalling pathway (Hahn *et al.*, 1996, Johnson *et al.*, 1996, Dahmane *et al.*, 1999). Loss of function mutations in genes within the SHH pathway were subsequently discovered in patients with medulloblastoma, including *PTCH*, suppressor of fused (*SUFU*) and more rarely smoothed (*SMO*) (Taylor *et al.*, 2002, Gilbertson, 2004).

Many solid tumours in children are described as 'embryonal' as they have an undifferentiated histopathological appearance, resembling the corresponding tissue in the developing embryo (reviewed by Scotting (Scotting *et al.*, 2005)). These tumours may also mimic the biological behaviour of the developing tissue from which they arise. In hepatoblastoma, for example, malignant cells secrete α -fetoprotein (AFP), which is usually only produced by the fetal yolk sac and developing liver in early fetal development, reaching a peak at week 14 of gestation and becoming almost undetectable by the age of 1 year (Van Tornout *et al.*, 1997). AFP may be used as a diagnostic marker for hepatoblastoma, and as a highly sensitive method of measuring tumour response during treatment and follow-up.

Many childhood tumours have an age-specific incidence; neuroblastoma, hepatoblastoma and Wilms' tumour mainly affect young children aged less than 4 years, while Hodgkin's lymphoma, osteosarcoma and Ewing's sarcomas are seen more commonly in adolescents (Scotting *et al.*, 2005). The peak incidence for bone tumours occurs at the same time as the adolescent growth spurt, suggesting that there may be a relationship between tumour development and rapid bone growth. Patients with bone tumours have generally undergone puberty earlier than their age-matched peers, however there are conflicting data on whether these patients are taller than age-matched controls (Buckley *et al.*, 1998, Cotterill *et al.*, 2004).

Finally, gene expression studies investigating the profiles from medulloblastoma and Wilms' tumours have shown marked similarities with genes expressed during early fetal development of corresponding normal tissues in mice (Dekel, 2003, Kho *et al.*, 2004). Taken together, these data appear to suggest that cancers in children may arise from defects in the normal processes of growth and development.

1.8 Cancer stem cells

Studies of leukaemias, and more recently of solid tumours suggest that cancer is initiated and maintained by a rare fraction of cells with stem cell properties, termed cancer stem cells (Singh *et al.*, 2004, Taylor *et al.*, 2005, Singh *et al.*, 2003, Bonnet *et al.*, 1997). These cells are phenotypically similar to normal stem cells, but display dysfunctional patterns of self-renewal and differentiation. Cancer stem cells were first described within acute myeloid leukaemia (AML), where a fraction of CD34⁺/CD38⁻ cells were shown to be capable of initiating AML in non-obese diabetic mice with severe combined immunodeficiency (NOD-SCID) (Bonnet *et al.*, 1997). These cells displayed the same cell surface markers as normal haematopoietic stem cells, suggesting that the initiating event in leukaemias occurred in the stem cell compartment.

Similar findings have been made in solid tumours, including paediatric brain tumours. Singh *et al.* established cultures from 14 primary paediatric brain tumours including medulloblastomas, pilocytic astrocytomas, grade II astrocytoma, ependymoma and ganglioglioma (Singh *et al.*, 2003). Neurospheres formed from a small number of cells

(<1%) from all tumour types within 24 to 48 hours. These were found to express the surface glycoprotein CD133, which is normally found on neural stem cells. These CD133⁺ cells displayed properties of self-renewal, proliferation and differentiation and were able to recapitulate the original phenotype of the human brain tumours when cultured *in situ*. Furthermore, when CD133⁺ cells were selectively sorted from primary brain tumours they were able to generate tumours *in vivo* on transplantation into NOD-SCID mice and continued to display stem cell properties of self-renewal, proliferation and differentiation (Singh *et al.*, 2004). CD133 negative cells stopped proliferating, became adherent in culture, and did not form tumours on transplantation into NOD-SCID mice (Singh *et al.*, 2003, Singh *et al.*, 2004). CD133⁺ cells derived from brain tumours have been termed brain tumour initiating cells or brain tumour stem cells (BTSC). BTSC have also been shown to express nestin, Musashi-1, Sox2 and MELK on the cell surface, in addition to CD133 (Sutter *et al.*, 2007).

Brain tumour stem cells isolated from ependymomas have given an insight into the possible cell of origin for these tumours. Ependymomas from different sites display particular genetic aberrations, suggesting they are molecularly distinct diseases; supratentorial ependymomas generally display deletion of *CDKN2A α* while spinal ependymomas often show deletions at chromosome band 22q12 (Taylor *et al.*, 2005). However, gene expression profiles for supratentorial and spinal ependymomas were shown to display similar patterns when compared to those from radial glial cells from corresponding regions of the developing mouse brain (subventricular zone of the lateral ventricles and spinal cord). CD133⁺/Nestin⁺ tumour cells were subsequently isolated from ependymomas. These cells showed similar immunophenotypes to radial glial cells, and were able to form tumours with characteristic histopathological features of ependymoma when transplanted into NOD-SCID mice. These data suggest that ependymomas may arise from radial glia, and hence these may be candidate BTSCs found in distinct regions of the central nervous system.

The distinction between CD133⁺ brain tumours stem cells and CD133⁻ cells may not, however, be as clear as previously thought. Sequential passage of a pure population of CD133⁻ cells, isolated from primary glioblastomas showed that a small subset of CD133⁻ cells were able to proliferate and produce further neurospheres (Beier *et al.*, 2007). These contained differentiated cells, showing markers for different neural lineages, indicating that some CD133⁻ cells from tumours may also contain stem cell

properties. These observations have been confirmed by further studies performing serial passage of glioblastoma xenografts with no detectable CD133 by Western blot analysis (Yi *et al.*, 2008). In addition, Yi *et al.* showed that CD133 is unmethylated in normal colon or brain, but is apparently silenced by promoter hypermethylation in some colorectal and glioblastoma primary tumours and cell lines, suggesting that CD133 activity may also be regulated epigenetically. It is therefore possible that some glioblastoma cells may not express CD133 due to promoter hypermethylation, but that this proportion may be regulated to increase CD133 expression, perhaps during periods of rapid tumour growth. Pallini *et al.* showed that 14/44 glioblastoma tumour samples were able to generate BTSC cultures (Pallini *et al.*, 2008). Patients with tumours able to generate BTSC cultures had a significantly shorter overall survival than those that did not, and higher levels of CD133 and Ki67 expression were also shown to correlate with poor patient survival.

1.9 An overview of the mitogen-activated protein kinase (MAPK) pathway

An overview of the MAPK pathway is presented here, since work presented in this thesis reveals that the RAF protein kinases *BRAF* and *RAF1* within the MAPK pathway are critically important in the development of paediatric low-grade astrocytomas. These findings will be discussed in detail later.

Protein kinases are enzymes, which are capable of covalently attaching phosphate groups to target proteins at particular serine, threonine or tyrosine amino acids. This addition of a phosphate group may affect how the substrate interacts with other proteins, alter its propensity for degradation or regulate its enzymatic activity (Johnson *et al.*, 2002). In the MAPK pathway, many of the substrates are protein kinases, which are ‘switched on’ by phosphorylation; this in turn allows them to phosphorylate subsequent target proteins.

The MAPK pathway is a conserved signalling cascade, which utilizes a series of protein kinases to transduce signals from the cell membrane to the nucleus (Figure 1.4). It plays a crucial role in mediating a range of biological functions, including cell growth, survival and differentiation. The initiating event in MAPK pathway activation occurs

when an extra cellular ligand binds to one of several cell-surface receptors. Examples of the known cell surface receptors and ligands that may initiate activation of the MAPK pathway are shown in Table 1.2. The activated receptor complex then recruits guanine nucleotide-exchange factors (GNEFs), such as members of the SOS family. These GNEFs displace guanine nucleotides from the membrane-bound RAS proteins, thus allowing RAS to bind cytosolic GTP and so adopt its active GTP-bound form (Vetter *et al.*, 2001). This process may be reversed by GTPase-activating proteins, such as neurofibromin 1 (NF1), which catalyse the conversion of the active GTP-bound form of RAS to the inactive GDP-bound form. When RAS is activated, it is capable of interacting with more than 20 substrates, including members of the RAF family and phosphatidylinositol 3-kinase (PI3K) (Schubbert *et al.*, 2007). This interaction allows the three RAF kinases (ARAF, BRAF and RAF1) to phosphorylate MEK1 and MEK2, which subsequently activate ERK1 and ERK2. The activated ERK1/2 proteins translocate to the nucleus, where they may phosphorylate a number of effector molecules such as JUN, ELK1, CREB, TIF-IA and histone H3, which then alter the behaviour of the cell (Schubbert *et al.*, 2007, Sharrocks, 2001).

One possible outcome of ERK activation is cell proliferation (Zhao *et al.*, 2003). PI3K, which is also activated by RAS proteins, catalyses the phosphorylation of phosphatidylinositol-4,5-bisphosphate (PIP₂) to generate phosphatidylinositol-3,4,5-trisphosphate (PIP₃). This process is reversed by the 3'-phosphatase PTEN, which converts PIP₃ back into PIP₂. The accumulation of PIP₃ leads to the recruitment of PDK1 to the membrane where it phosphorylates AKT. Like ERK, AKT can phosphorylate a number of different proteins, many of which are involved in cell survival (Engelman, 2009). The final target of RAS is RASSF1A, which binds RAS in a GTP-dependent manner and primarily mediates apoptosis (Vos *et al.*, 2000).

The MAPK pathway has been found to be constitutively active in a wide range of cancers (Mitra *et al.*, 2009, Russo *et al.*, 2009, Espinosa *et al.*, 2007). Most notably, alterations in the MAPK pathway have been found in 88% of adult glioblastomas (CGARN *et al.*, 2008). Genetic aberrations which contribute to the deregulation of this pathway have been found in both the cell-membrane receptors, such as EGFR, ERBB2, FGFR1 and PDGFRA, and the downstream cytosolic components NF1, BRAF, HRAS, KRAS, NRAS and PTEN (Stratton *et al.*, 2009, Vivanco *et al.*, 2002). This deregulation

of the MAPK pathway either occurs as a result of genes with tumour suppressor functions being inactivated or oncogenes becoming permanently activated

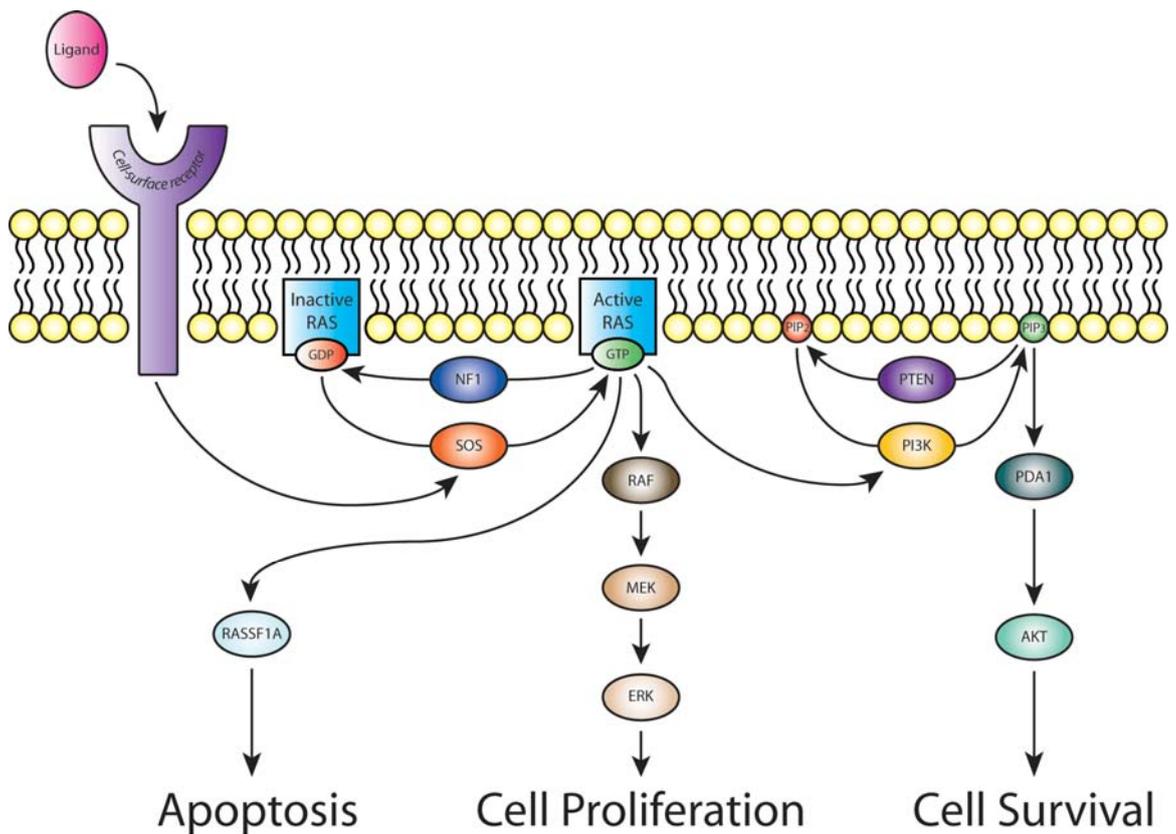


Figure 1.5 Schematic representation of the MAPK pathway, and the possible cellular outcomes from pathway activation.

Receptor	Ligands	References
Estrogen receptor	Estradiol	(Migliaccio <i>et al.</i> , 1996)
	Epidermal growth factor (EGF)	(Bunone <i>et al.</i> , 1996)
Fibroblast growth factor receptor	Fibroblast growth factor	(Muroya <i>et al.</i> , 1992)
Epidermal growth factor receptor (EGFR)	EGF	(Heasley <i>et al.</i> , 1992)
E-cadherins (cell surface molecules)	** Ligand independent stimulation of EGFR	(Pece <i>et al.</i> , 2000)

Table 1.2 Examples of cell surface receptors and ligands which may initiate activation of the MAPK pathway.

1.10 Paediatric brain tumours

1.10.1 Epidemiology

Brain tumours form the second largest group of childhood cancers after leukaemias in children aged 0-14 years. Leukaemias comprise 30% cases, while brain tumours are seen in 20-25% cases (Grovas *et al.*, 1997, Peris-Bonet *et al.*, 2006, Arora *et al.*, 2009). Overall survival rates have improved markedly for childhood leukaemias over the last 20 years: 5-year survival rates are now over 80% (Pui *et al.*, 2003). Childhood brain tumours have not seen a similar improvement. Hence, brain tumours are the leading cause of cancer-related death in children, adolescents and young adults (aged up to 29 years) (Davis *et al.*, 1998, Baldwin *et al.*, 2004, Geraci *et al.*, 2007).

Arora *et al.* conducted a recent study on the epidemiology of primary central nervous system tumours in England, using national cancer registration data from the Office of National Statistics between 1995 and 2003 (Arora *et al.*, 2009). Central nervous system (CNS) tumours formed 25% of all tumours seen in children aged 0-14 years, and 9% for adolescents and young adults aged between 15-24 years. Primary CNS tumours formed 2% of all cancers in adults aged 25-84 years.

In young children infratentorial tumours of the posterior fossa predominate, comprising around 30% of all CNS tumours seen. The proportion of infratentorial tumours decreases with age; the majority of brain tumours in adults and older children arise above the tentorium, in and around the cerebral hemispheres (Pollack, 1994, Arora *et al.*, 2009).

Astrocytomas are the most common type of brain tumour seen in children. Grovas *et al.* studied incidence data from 2100 hospitals across the United States between 1985-1993, estimated to include data from approximately 42% of all childhood cancer patients (Grovas *et al.*, 1997). Grovas found 53% of all childhood brain tumours were astrocytomas, followed by medulloblastoma (24%) and ependymoma (9%). Arora *et al.* studied incidence data over three different age ranges, and found the prevalence of and pathological type of astrocytomas varied with increasing age (Arora *et al.*, 2009). Astrocytomas comprised 40% of all CNS tumours in children between 0-14 years; of these over 50% were pilocytic astrocytomas (PAs), followed by medulloblastomas

(13%) and ependymomas (7%). Astrocytomas formed 30% of CNS tumours in young adults aged 15-24 years, of these 25% were PAs. In this age group tumours of the sellar region were seen in 16% (the majority of these were pituitary adenomas), medulloblastomas and ependymomas together comprised less than 1% of all CNS tumours. In adults aged 25-84 years astrocytomas formed 35% of all CNS tumours, the majority (67%) were glioblastomas. Metastases to the brain, from primary tumours sited elsewhere are very common in adults, seen in 20-40% of cancer patients, however these are rarely seen in children (Pollack, 1994, Patchell, 2003).

Overall age-standardized incidence rates (ASR) for paediatric brain tumours are similar between Europe and North America, ranging from 29.9 to 35.6 per million person-years respectively (Arora *et al.*, 2009, Peris-Bonet *et al.*, 2006). Astrocytoma in children has an ASR of 11.8, followed by medulloblastoma (ASR = 6.5) and ependymoma (ASR = 3.4). The majority of childhood brain tumours occur before the age of 5 years, so prenatal factors must be considered important aetiological influences.

Following this early peak there is a nadir in the incidence of CNS tumours between 15-19 years, before a further peak in incidence between 75-79 years (Arora *et al.*, 2009). These data suggest that genetic factors must play a major role in childhood CNS tumours, and that additional genetic and environmental factors may be implicated in the development of adult brain tumours.

1.10.2 Aetiological factors

Ionising radiation is the only environmental risk factor to have been clearly implicated in childhood brain tumour development. Prenatal exposure to radiation was shown to increase the risk of childhood brain tumours (Harvey *et al.*, 1985, Rodvall *et al.*, 1990, Doll *et al.*, 1997). Medical practice has changed to avoid the use of X-rays in pregnancy, so this aetiology should not be of concern in the future. Children previously treated with radiation for cancer have also shown a significant increase (around 22-fold) in the incidence of CNS malignancies and a 7-fold increase in all types of secondary cancer (Neglia *et al.*, 1991, Meadows *et al.*, 1985). Numerous reviews have been performed, considering a wide variety of environmental factors including electromagnetic fields, use of mobile telephones and dietary factors, but no additional risk factors have been identified to date (Bondy *et al.*, 1994, Baldwin *et al.*, 2004).

Families with higher than expected incidence of brain tumours have provided insights into the genetic basis of disease. Syndromes including Li-Fraumeni syndrome, tuberous sclerosis, Cowden syndrome, Gorlin's syndrome, Turcot syndrome, and neurofibromatosis types I and II are all known to confer an increased risk of childhood brain tumours (Table 1.2) ((Lewis *et al.*, 1983, Blatt *et al.*, 1986, Gorlin, 1987, Li *et al.*, 1988, Martuza *et al.*, 1988, Miyagi *et al.*, 1995, Gustafson *et al.*, 2007, Napolioni *et al.*, 2008)). However, these syndromes are rare, and have been estimated to account for less than 5% of childhood brain tumours overall (Baldwin *et al.*, 2004, Bondy *et al.*, 1994).

Syndrome	Gene	Chromosomal location	Tumour types
Cowden	<i>PTEN</i>	10q23q31 (Steck <i>et al.</i> , 1997)	Gangliocytoma
Hereditary retinoblastoma	<i>Rb</i>	13q14 (Friend <i>et al.</i> , 1986)	Pineoblastoma, glioma, meningioma
Li-Fraumeni	<i>TP53</i>	17p13.3 (Cogen <i>et al.</i> , 1992)	PNET, medulloblastoma, astrocytoma
Neurofibromatosis Type 1	<i>NF-1</i>	17q11.2 (Ledbetter <i>et al.</i> , 1989)	Neurofibroma, optic nerve glioma, astrocytoma
Neurofibromatosis Type 2	<i>NF-2</i>	22q12.2 (Arai <i>et al.</i> , 1994)	Schwannoma, meningioma, spinal ependymoma
Nevoid basal cell carcinoma (Gorlin)	<i>PTCH</i>	9q22.3 (Hahn <i>et al.</i> , 1996)	Medulloblastoma, meningioma
Tuberous sclerosis	<i>TSC1</i> <i>TSC2</i>	9q34 (van Slegtenhorst <i>et al.</i> , 1997) 16p13.3 (Green <i>et al.</i> , 1994)	Subependymal giant cell astrocytoma
Turcot	<i>APC</i> <i>hMLH1</i> <i>hPMS2</i>	5q21q22 (Dolan <i>et al.</i> , 1999) 3p21.3 (Roche <i>et al.</i> , 1996) 7p22 (Risinger <i>et al.</i> , 1995)	Medulloblastoma astrocytoma, ependymoma
Rubenstein-Taybi	<i>CBP</i>	16p13.3 (Petrij <i>et al.</i> , 1995)	Medulloblastoma, oligodendroglioma, meningioma

Table 1.3 Inherited syndromes associated with brain tumours.

Table adapted from Blaney *et al.* (Blaney *et al.*, 2006).

1.10.3 Early development of the central nervous system

The central nervous system begins to form in the 3rd week of embryonal development, when the ectoderm on the dorsal side of the embryo thickens to form the neural plate (Heimer, 1983). The neural plate grows and folds to form the neural tube, which will become the brain and spinal cord. As the neural tube closes, cells migrate from the lateral portions of the walls to become the neural crest. The two main cell types within the brain are derived from the neural crest; neural cells and glial cells. Neural cells differentiate into neurons. Glial cells differentiate into four further types of cells: astrocytes, oligodendrocytes, ependymal cells and microglia.

Astrocytes supply support and nutrition to neurons and are the most abundant glial cell type found in the brain: oligodendrocytes produce the myelin sheath that surrounds neurones. Ependymal cells form the lining within the ventricles and the spinal cord, that contains cerebrospinal fluid. Microglia are involved in tissue repair and neuronal regeneration, and in immune surveillance (Kreutzberg, 1996). They undergo rapid proliferation in response to tissue destruction and migrate to the site of injury where they phagocytose debris or microorganisms. Microglia promote tissue healing by secreting growth factors including IL6 and TGF β .

The cellular origins for brain tumours have not yet been clearly defined. Astrocytomas are thought to be derived from glial cells, specifically astrocytes, which they resemble histologically. Similarly oligodendrogliomas are thought to arise from oligodendrocytes and ependymomas from ependymal cells (Sanai *et al.*, 2005).

1.10.4 Histopathological classification of paediatric astrocytomas

Astrocytomas are classified into grades I to IV by the World Health Organisation (WHO), using defined histopathological and clinical diagnostic criteria reflecting a spectrum of biological behaviour (Kleihues *et al.*, 2000, Louis *et al.*, 2007). The WHO classification for CNS tumours was first published in 2000 based on the consensus recommendations of a working group of experts who met in Lyon in 1999. The fourth edition of the WHO classification of tumours of the central nervous system was

published in 2007, based on the consensus of 25 pathologists and geneticists (and considered additional contributions from a further 70 international experts) (Louis *et al.*, 2007). These are the standard criteria for the definition of brain tumours throughout the world.

The WHO grading is an attempt to predict the behaviour of a tumour; slowly growing tumours are graded I or II. Tumours graded III and IV show increased anaplasia, more mitotic activity and atypical nuclei, and may contain regions of necrosis reflecting more malignant behaviour. All of the work presented in this thesis was conducted on low-grade (WHO grade I and II) paediatric astrocytomas.

Low-grade astrocytomas (LGAs) are a heterogeneous group of tumours, including pilocytic astrocytomas, diffuse astrocytomas (WHO Grade II), pilomyxoid astrocytomas (WHO Grade II) and pleomorphic xanthoastrocytomas (WHO Grade II). Pilocytic astrocytomas (PAs) are well-circumscribed, slow growing, cystic lesions and are the most common type of astrocytoma in children aged 0-19 years (Arora *et al.*, 2009, Kleihues *et al.*, 2000). They can arise throughout the brain and spinal cord, but are most commonly situated in the cerebellum (Ohgaki *et al.*, 2005b). Pilocytic astrocytomas may also arise in the optic pathways (optic pathway gliomas), these tumours are usually associated with the tumour predisposition syndrome neurofibromatosis type I (NF1) (Rubin *et al.*, 2005, Listernick *et al.*, 2007, Brems *et al.*, 2009). NF1 may be inherited as an autosomal dominant disorder or occur following a *de novo* mutation in the tumour suppressor gene *NF1*, located at chromosome band 17q11.2 (von Deimling *et al.*, 1995). *NF1* encodes the protein neurofibromin 1, which acts as a negative regulator of the RAS signal transduction pathway (Schubbert *et al.*, 2007). Optic pathway gliomas are seen in around 15% of patients with NF1, and often behave less aggressively than sporadic pilocytic astrocytomas (Listernick *et al.*, 2007).

Subependymal giant cell astrocytoma (SEGA) is also classified as a grade I astrocytoma by the WHO. This tumour is associated with tuberous sclerosis, a rare multi-system autosomal dominant disease characterised by benign tumours in the brain, kidneys, heart, eyes, lungs and skin (Napolioni *et al.*, 2008). Tuberous sclerosis is caused by mutations in *TSC1* located at chromosome band 9q34 which encodes the protein tuberlin, or *TSC2* located at chromosome band 16p13, encoding hamartin. SEGAs are seen in 6-14% patients with tuberous sclerosis, and usually develop in the first two

decades of life. SEGAs were excluded from the low-grade astrocytomas included within this study.

Diffuse astrocytomas may arise at any site within the central nervous system, but are most commonly found in the cerebral hemispheres, particularly the frontal and temporal lobes (Kleihues *et al.*, 2000). They are characterised by well-differentiated fibrillary astrocytes, showing moderately increased cellularity and occasional nuclear atypia, and diffusely infiltrate adjacent normal brain. Diffuse astrocytomas may progress in young adults, to form secondary glioblastomas (Ohgaki *et al.*, 2007).

Pilomyxoid astrocytomas were originally described as ‘diencephalic pilocytic astrocytoma with clinical onset in infancy’ before being reclassified as a separate entity (Janisch *et al.*, 1985). Pilomyxoid gliomas are characterised by myxoid degeneration and monomorphic piloid cells (Louis *et al.*, 2007). They are primarily located in the hypothalamic/chiasmatic region and appear to have a worse prognosis than pilocytic astrocytomas.

Finally, pleomorphic xanthoastrocytomas are typically found in the cerebral hemispheres, particularly in the temporal lobe. They may also involve the meninges where a cyst may form (Kleihues *et al.*, 2000). These tumours are characterised by large xanthomatous cells containing lipid deposits, and generally have a favourable prognosis.

1.10.5 Clinical presentation and diagnosis

Brain tumours present in a variety of ways, depending on the age of the child and the physical location, type of tumour and the rate of tumour growth (Behrman *et al.*, 1996). The most common presenting features are symptoms and signs of raised intracranial pressure and focal neurological signs. Tumours situated in the posterior fossa, pineal and suprasellar regions are more likely to present with raised intracranial pressure, by obstructing the flow of cerebrospinal fluid and producing hydrocephalus (Pollack, 1994). Tumours situated above the tentorium are more likely to present with focal neurological signs and seizures. Most children present with an alteration to their personality, they may become irritable, lethargic or forgetful and may perform poorly at

school. Many children may also have recurrent episodes of nausea and/or vomiting and headaches.

The diagnosis is usually suggested from the patient's history and by neurological examination. Following this, imaging is performed using computerised tomography (CT) or magnetic resonance imaging (MRI). CT or MRI appearances may suggest a particular tumour type, and will delineate the size and extent of disease. Imaging is also important following surgery, where the presence and estimated volume of residual disease will direct further treatment decisions.

1.10.6 Treatment

Five-year survival rates for childhood brain tumours range from 64% to 72% across Europe and in the United States, respectively (Peris-Bonet *et al.*, 2006). Astrocytomas have the best 5-year survival rates for all childhood brain tumours, around 75% (Grovas *et al.*, 1997). Low-grade astrocytomas (WHO grades I-II) are less aggressive and more responsive to treatment than their high-grade counterparts (WHO grades III-IV). Two-year progression-free survival rates for children with glioblastoma (WHO grade IV) are extremely poor, from 10-30%; 10-year survival rates are >95% for children with pilocytic astrocytoma (Broniscer *et al.*, 2004, Burkhard *et al.*, 2003, Gajjar *et al.*, 1997).

The preferred method of treatment for circumscribed low-grade astrocytomas is radical surgical excision, and if complete resection is achieved no further therapy is necessary (Qaddoumi *et al.*, 2009). However, this may not be possible in cases where the tumour is within the brain stem, diencephalon or optic pathways, or where there is a large amount of infiltration of adjacent normal tissue. The numerical distinction between astrocytomas of grades I-II is not intended to imply that grade II tumours should be treated more aggressively. It is often extremely difficult for pathologists to distinguish between tumours of grade I and II, particularly as many astrocytomas are heterogeneous.

Adjuvant chemotherapy and radiotherapy are used for patients that have unresectable tumours and recurrent or metastatic disease. However, these are not always curative, and further therapies are needed for such clinical situations. The combination of

carboplatin and vincristine is currently the standard approach to treat low-grade astrocytomas in Europe and the United States (Lancaster *et al.*, 2003, Qaddoumi *et al.*, 2009). Carboplatin allergy is a significant problem with the regimen, and may be seen in up to 42% of patients, requiring cessation of treatment with this agent (Lafay-Cousin *et al.*, 2008). Alternative chemotherapy regimens include TPCV, comprised of thioguanine (T), procarbazine (P), CCNU (lomustine) and vincristine (V), and oral temozolamide. There is no apparent difference in the outcome for patients with incompletely resected disease treated with TPCV or carboplatin and vincristine (Ater *et al.*, 2008). Grade IV astrocytomas (glioblastoma multiforme) are highly malignant, and are among the most difficult tumours to treat in children and adults. Changes to treatment strategies have made little impact on survival for children with high-grade gliomas (HGG) in the last three decades, currently 2-year progression-free survival is less than 20% for these patients (Gottardo *et al.*, 2008).

Radiotherapy achieves good tumour response, but is associated with significant neurocognitive and neuroendocrine disabilities, vasculopathy and the risk of second tumours (Packer *et al.*, 1987, Packer *et al.*, 2003, Mulhern *et al.*, 2004). Morbidity following radiotherapy is particularly prevalent in patients with NF1 (Sharif *et al.*, 2006). In general, radiotherapy is delayed until patients are aged > 10 years, or there are signs of tumour progression (Pollack, 1994).

1.10.7 Molecular biology of low-grade paediatric astrocytomas

The majority of low-grade astrocytomas appear to display normal karyotypes on conventional karyotyping. One hundred and thirteen paediatric pilocytic astrocytomas (PAs) and 4 diffuse astrocytomas (DAs) were investigated over eight separate studies (Jenkins *et al.*, 1989, Karnes *et al.*, 1992, Agamanolis *et al.*, 1995, Debiec-Rychter *et al.*, 1995, White *et al.*, 1995, Bhattacharjee *et al.*, 1997, Bigner *et al.*, 1997, Zattara-Cannoni *et al.*, 1998). I combined the data from all of these studies: seventy-six PAs (67%) and all 4 DAs were found to have apparently normal karyotypes. Thirty PAs contained multiple chromosomal aberrations, and 7 contained a single change (usually gain of a single chromosome). The most commonly seen aberrations were whole chromosome gains of chromosome 7 (16 cases), chromosome 8 (11 cases), chromosome 11 (10 cases), chromosomes 5 and 6 (5 cases each) and chromosome 12, 15 and 22 (4

cases each). Small numbers of chromosome losses were also seen in one or two cases for the majority of chromosomes, none were lost with a high prevalence.

Several studies have investigated loss of heterozygosity on chromosome 17 in pilocytic astrocytomas from children and young adults. Von Deimling *et al.* studied 20 PAs, and found loss of heterozygosity of loci including *NF1* on the long arm of chromosome 17 in four patients (one patient was known to have NF1) (von Deimling *et al.*, 1993). Willert *et al.* performed mutational analysis for *TP53* and studied chromosome arm 17p in 8 high-grade astrocytomas and 20 PAs from children (Willert *et al.*, 1995). *TP53* mutations were identified in 1 high-grade astrocytoma and 1 PA. Telomeric losses of 17p were seen in 6/8 high-grade astrocytomas, and in 2/20 PAs. Losses of 17p centromeric to *TP53* were seen in 3 high-grade astrocytomas and 5 PAs. Four of the 7 PA patients with 17p deletions suffered disease recurrence within 1-3 years of completing treatment, compared to 5/13 PAs without 17p loss. These data suggested that genes on 17p might be implicated in disease progression in PAs. Phelan *et al.* conducted mutational analysis for *TP53* in 29 paediatric brain tumours, including 7 PAs and 2 grade II astrocytomas (Phelan *et al.*, 1995). No mutations were detected in *TP53* for any of the tumours investigated, and no LOH was identified in any of the low-grade astrocytomas studied. Finally, Patt *et al.* conducted mutational analysis for *TP53* in 42 astrocytomas, including 7 PAs and 18 low-grade astrocytomas (Patt *et al.*, 1996). A *TP53* mutation was found in 1 PA, but this did not lead to a change in the translated amino acid sequence.

Sanadou investigated 41 paediatric pilocytic astrocytomas and 7 adult pilocytic astrocytomas using comparative genomic hybridisation (Sanoudou *et al.*, 2000). Five of the 41 paediatric PAs (12%) showed chromosomal aberrations, all showed gain of a single chromosome. Two paediatric PAs showed gains of chromosome 7, the remaining three showed gains of chromosomes 5, 6 and 9, respectively. Two of the 7 adult PAs (29%) showed multiple chromosomal aberrations. One showed gains of 1p33-pter, 9q34.1-qter, 17q21.3-ter and whole chromosome gains of 19 and 22. The second showed gains of 2p22-pter, 9q31-qter, 12q13.2-q23 and loss of 16q.

Thus, gain of chromosome 7 was the most common copy number aberration found in paediatric PAs by conventional cytogenetics and comparative genomic hybridisation. These data implied that genes found on chromosome 7 might be involved in the

development of paediatric low-grade astrocytomas. Studies investigating these tumours at higher resolution were required to attempt to identify specific genetic aberrations in more detail.

1.11 Astrocytomas in children and adults

Glioblastomas in adults may arise *de novo* (primary glioblastoma) or evolve from a grade II diffuse astrocytoma or grade III anaplastic astrocytoma (secondary glioblastoma). The molecular pathways leading to the development of adult primary and secondary glioblastomas have been extensively studied and described in detail (Ohgaki *et al.*, 2007, CGARN *et al.*, 2008, Parsons *et al.*, 2008). Paediatric high-grade astrocytomas are less common than their adult counterparts, and few paediatric high-grade astrocytomas appear to arise by progression from lower-grade tumours. Less than 10% paediatric grade II infiltrative astrocytomas undergo malignant transformation to high-grade tumours, compared to 50-90% in adult Grade II tumours (Broniscer *et al.*, 2007).

PTEN mutations and deletions or EGFR amplification, the typical molecular characteristics seen in adult high-grade tumours are present in Grade III and IV paediatric tumours although they appear to be less common, suggesting that childhood tumours may arise via different molecular pathways (Pollack *et al.*, 2006). Similarities and differences between the molecular changes found in high- and low-grade adult and paediatric astrocytomas will be discussed in detail later.

1.12 Aims of this thesis

Despite great technological advances in recent years, little was known about the molecular biology of paediatric low-grade astrocytomas at the start of this research, as described above. The aims of the research presented in this thesis were:

- 1) To identify regions of recurrent DNA copy number gain or loss in low-grade paediatric astrocytomas. A detailed genome-wide assessment of grade I and II paediatric astrocytomas was performed, initially using array comparative genomic hybridisation and later using high resolution single nucleotide polymorphism (SNP) arrays. Comparisons were made between pilocytic astrocytomas arising in the cerebellum and grade II astrocytomas arising in the cerebrum. This work is described in chapter 3 and chapter 4.
- 2) To investigate any discrete or recurrent regions of DNA copy number change identified, and to determine the significance of genes contained within the affected genomic regions. The follow-up of discrete regions of copy number gain at chromosome bands 7q34 and 3p25, and the identification of specific *RAF* gene fusions are described in chapter 5.
- 3) To investigate candidate genes implicated in adult astrocytoma tumorigenesis by mutational analysis. This work is described in chapter 6.
- 4) An overview of the findings and their wider significance is discussed in chapter 7, and also in the paper 'Activation of the ERK/MAPK pathway: a signature genetic defect in posterior fossa pilocytic astrocytomas', which is bound at the end of the thesis (Forsheew *et al.*, 2009).

Chapter 2. Materials and methods

2.1 Tumour samples

Low-grade astrocytomas (WHO grades I and II) from 50 patients aged 1-20 years were obtained from Newcastle General Hospital, United Kingdom and St Jude Children's Research Hospital, Memphis USA in collaboration with Professor David Ellison. Access to the tumours and clinical data were in accordance with the Institutional Review Board and MREC regulations at all the centres involved in the study: Newcastle REC reference number 2002/112; St Jude Children's Research Hospital reference number XPD07-107/IRB and Tissue Resource Request reference number 07-007 and at the Institute of Cell and Molecular Science reference number ICMS/PR/09/77.

Thirty-two patients had cerebellar pilocytic astrocytomas (WHO grade I), 11 had cerebral diffuse (fibrillary) astrocytomas (WHO grade II), and 5 had midline pilomyxoid gliomas. One patient had an infantile hypothalamic pilomyxoid astrocytoma (PMA, WHO grade II) and one had a cerebral pleomorphic xanthoastrocytoma (Table 2.1). All of the patients were negative for Neurofibromatosis type I (NF1) except for the infant with a PMA, who had been diagnosed with NF1 on clinical grounds. Tumours from the optic pathways were not selected for inclusion in the study, so as to exclude optic pathway gliomas occurring secondary to NF1.

All tissue samples were snap frozen in liquid nitrogen at the time of surgical resection, and were obtained from patients before any adjuvant therapy was given. Sections (5-8µm) were prepared from formalin-fixed paraffin-embedded (FFPE) samples of all tumours for detailed histopathological review and selected interphase fluorescence *in situ* hybridization (iFISH).

2.2 Control DNA samples

Control DNA was obtained from blood volunteered from fit and healthy co-workers at Cancer Research UK London Research Institute. Samples were collected from twenty-one males and one female.

Name	Pathology	Site	Age	Sex
PA1	Pilocytic astrocytoma	Cerebellum	20	F
PA2	Pilocytic astrocytoma	Cerebellum	4	M
PA3	Pilocytic astrocytoma	Cerebellum	7	M
PA4	Pilocytic astrocytoma	Cerebellum	2	M
PA5	Pilocytic astrocytoma	Cerebellum	6	F
PA6	Pilocytic astrocytoma	Cerebellum	5	M
PA7	Pilocytic astrocytoma	Cerebellum	13	M
PA8	Pilocytic astrocytoma	Cerebellum	2	M
PA9	Pilocytic astrocytoma	Cerebellum	14	F
PA10	Pilocytic astrocytoma	Cerebellum	19	M
PA11	Pilocytic astrocytoma	Cerebellum	14	M
PA12	Pilocytic astrocytoma	Cerebellum	14	M
PA13	Pilocytic astrocytoma	Cerebellum	10	M
PA14	Pilocytic astrocytoma	Cerebellum	10	M
PA15	Pilocytic astrocytoma	Cerebellum	13	M
PA16	Pilocytic astrocytoma	Cerebellum	10	M
PA17	Pilocytic astrocytoma	Cerebellum	6	M
PA18	Pilocytic astrocytoma	Cerebellum	3	M
PA19	Pilocytic astrocytoma	Cerebellum	3	M
PA20	Pilocytic astrocytoma	Cerebellum	4	M
PA21	Pilocytic astrocytoma	Cerebellum	12	F
PA22	Pilocytic astrocytoma	Cerebellum	6	F
PA23	Pilocytic astrocytoma	Cerebellum	3	F
PA24	Pilocytic astrocytoma	Cerebellum	4	F
PA25	Pilocytic astrocytoma	Cerebellum	4	F
PA26	Pilocytic astrocytoma	Cerebellum	3	F
PA27	Pilocytic astrocytoma	Cerebellum	14	F
PA28	Pilocytic astrocytoma	Cerebellum	7	M
PA29	Pilocytic astrocytoma	Cerebellum	19	M
PA30	Pilocytic astrocytoma	Cerebellum	13	F
PA31	Pilocytic astrocytoma	Cerebellum	1	F
PA32	Pilocytic astrocytoma	Brain stem	9	F
DA1	Diffuse astrocytoma	Cerebral cortex	5	F
DA2	Diffuse astrocytoma	Cerebral cortex	9	M
DA3	Diffuse astrocytoma	Cerebral cortex	17	M
DA4	Diffuse astrocytoma	Cerebral cortex	4	M
DA5	Diffuse astrocytoma	Cerebral cortex	7	F
DA6	Diffuse astrocytoma	Cerebral cortex	5	F
DA7	Diffuse astrocytoma	Cerebral cortex	5	M
DA8	Diffuse astrocytoma	Cerebral cortex	15	F
DA9	Diffuse astrocytoma	Cerebral cortex	9	F
DA10	Diffuse astrocytoma	Cerebral cortex	5	F
DA11	Diffuse astrocytoma	Cerebral cortex	2	M
PMA1	Pilomyxoid astrocytoma	Diencephalon	1	F
PMG1	Pilomyxoid glioma	Diencephalon	17	M
PMG2	Pilomyxoid glioma	Diencephalon	7	F
PMG3	Pilomyxoid glioma	Diencephalon	8	M
PMG4	Pilomyxoid glioma	Spinal cord	12	M
PMG5	Pilomyxoid glioma	Spinal cord	11	F
PXA1	Pleomorphic xanthoastrocytoma	Cerebral cortex	6	F

Table 2.1 Summary of clinical and pathological details from 50 patients with low-grade astrocytomas. Patient ages are shown in years.

2.2.1 Control RNA

Total RNA was obtained from normal control brain samples, available commercially (BioChain, Hayward, CA). The samples comprised four fetal brain samples (two from cerebellum and two from cerebrum) and 2 adult brain samples (one from cerebellum and one from cerebrum)

2.2.2 Control protein

Total protein was obtained from normal control brain samples, available commercially (BioChain, Hayward, CA). They comprised two fetal brain control samples, one from cerebellum and one from cerebrum.

2.3 Cell lines

Two paediatric medulloblastoma cell lines were obtained from the American Type Culture Collection (ATCC). Daoy (HTB186) was originally from a 4 year-old male with a posterior fossa medulloblastoma (Jacobsen *et al.*, 1985). HTB-187 (D341) was originally from a 3 year-old male with a posterior fossa medulloblastoma (Friedman *et al.*, 1988). Cell lines were used as a renewable resource to enable me to optimise molecular techniques before using tumour DNA samples. Paediatric medulloblastoma lines were selected, as no paediatric astrocytoma cell lines were available. Both cell lines were characterised using M-FISH, FISH and array CGH.

2.4 General Methods

2.4.1 Cell culture

Cells were grown on 75 cm² plastic culture flasks, within incubators maintaining a temperature of 37°C and a humidified atmosphere supplemented with 5% CO₂.

Daoy and HTB187 are both cultured as monolayers. They were grown in E4 media supplemented with 10% fetal calf serum (FCS), 1% 2mM L-glutamine, 1% non-essential amino acids and 1% 1mM sodium pyruvate, all v/v. Cells were passaged when

they reached confluence, usually every 3-4 days. Cells were then washed with phosphate-buffered saline (PBS) and detached from the flask by incubating for ~ 3 minutes with 3ml of a solution containing 0.5% v/v porcine trypsin and EDTA (Gibco, Invitrogen) per flask at room temperature. The trypsin digestion was then blocked by the addition of excess culture medium containing FCS. Cells were reseeded in two or three flasks at ~ 30% confluence, or harvested for the preparation of metaphase spreads or extraction of DNA.

2.4.2 Metaphase spread preparation

Prior to the preparation of metaphase spreads from cell lines, 4 drops of colcemid (0.05µg/ml, Sigma) were added to each 75cm² flask, before returning them to the incubator for 1 hour. Colcemid inhibits spindle formation within the cells, so when M-phase is reached cells are unable to progress further within the cell cycle. Hence, cells accumulate in metaphase of mitosis. The duration of exposure to colcemid may be varied to alter the length of chromosomes seen within the metaphase spread. For example, increased exposure will produce shorter chromosomes.

Cells were detached from the flask with trypsin, re-suspended in 10mls of pre-warmed 0.075M KCl and incubated at 37°C for 15 minutes. Following this, the cells were centrifuged and the pellet resuspended in 10mls cold 3:1 methanol/glacial acetic acid fixative (kept at -20 °C). This suspension was centrifuged at 1500rpm for 5 minutes (ALC centrifuge PK120), and the pellet re-suspended in 1-2 mls of cold fixative. Clean glass slides ('Superfrost', 1.0mm thick with a frosted end, VWR international) stored in 100% ethanol were used to make metaphase spreads. Slides were wiped dry of ethanol with lint-free tissue paper and then covered with cold 3:1 methanol/glacial acetic acid fixative. Following this, 2-3 drops of cell suspension were dropped onto the slide using a siliconised glass pipette. Slides were then left to air dry. The quality of the metaphase spreads were assessed under a light microscope. The cell suspension was diluted further with fixative if the slide had too many overlapping nuclei and metaphase spreads. The time of exposure to hypotonic KCl was also adjusted to improve the quality of metaphases seen, for example if multiple scattered single chromosomes were seen the cells had been exposed to KCl for too long.

2.4.3 DNA extraction from blood, cell line and tumour samples

DNA was extracted from blood and frozen tissue samples using the QIAGEN DNA mini kit for blood and cell culture (Qiagen, Crawley, UK), following the manufacturer's instructions. This method involves the use of optimised buffer systems for lysis of nucleated cells in the presence of an anionic detergent to solubilise cellular fragments. Genomic DNA binds to QIAGEN anion-exchange resin within the columns, under appropriate low salt and pH conditions. RNA, proteins and low-molecular weight impurities are removed by a medium-salt wash. Genomic DNA is finally eluted in sterile water, and stored at -20°C .

2.4.4 DNA extraction to create pooled male control DNA

Pooled control DNA was obtained from blood taken from 21 male volunteers working at Cancer Research UK London Research Institute. Genomic DNA was isolated from 10ml whole blood samples using the Chemagic Magnetic Separation Module I (Chemagen AG, Baesweiler, Germany), according to the manufacturer's instructions. Briefly, after lysis of nucleated cells in whole blood, magnetic beads were added that specifically bind DNA. The beads and DNA were bound to metal rods magnetised by an electromagnet to allow removal of cellular debris by washing. Finally, DNA was dissociated into an elution buffer, and the magnetic beads removed using the electromagnet, leaving pure DNA in solution. Each 10ml whole blood sample yielded 1ml of genomic DNA, with concentrations ranging from 171-306 ng/ μl . A stock solution of control pooled male DNA was created with a final concentration of 250ng/ μl , using 400 μl DNA from each of the 21 samples. The pooled male DNA was used as control DNA for PCR primer validation, sequencing and array comparative genomic hybridisation (array CGH).

2.5 RNA extraction from tumour samples

RNA was extracted from the frozen tissue samples using Trizol (Invitrogen, Paisley, UK), initially following the manufacturer's instructions. Once phase separation was complete, the RNA-containing phase was applied to a Qiagen RNeasy spin column

(Qiagen, Crawley, UK). RNA thus bound to the spin column was cleaned up using buffers RW1 and RPE by centrifugation following the instructions within the RNeasy mini kit protocol, and finally eluted into RNase free water. RNA was stored at -80°C , and repeated removal of RNA was avoided where possible, to avoid recurrent episodes of freezing and thawing.

2.6 Protein extraction from tumour samples

Protein was extracted from frozen tissue samples after homogenisation and lysis in 1x RIPA buffer containing 1% (v/v) protease inhibitor cocktail (Roche) and phosphatase inhibitor cocktail (Roche), according to the manufacturer's instructions.

2.7 DNA and RNA quantification

Analysis of DNA quantity and purity was performed using a Nanodrop 2000 spectrophotometer (NanoDrop, ThermoScientific). This instrument measures the light absorbance (A) at specific frequencies, within the diluted DNA sample. DNA absorbs ultraviolet light at a frequency of 260nm, while protein preferentially absorbs light at 280nm. A combination of both the 260nm absorbance reading, and the ratio of the A260nm/A280nm values allow quantification of DNA and an assessment of its purity to be calculated. Pure preparations of DNA will have an absorbance ratio (260nm/280nm) of approximately 1.8. Any contaminants, such as protein, within the DNA sample will cause a reduction in this value. RNA is considered to be of good quality if the absorbance ratio is close to 2.0.

The concentration of nucleic acids may be calculated from the absorbance reading using the following formula:

$$\text{Nucleic acid concentration} = \frac{\text{optical density 260nm}}{\text{Pathlength of laser}} \times \text{standard coefficient} \times \text{sample dilution}$$

In general, a 1cm pathlength standard coefficient of $50\mu\text{g}/\mu\text{l}$ is used for double stranded DNA and $40\mu\text{g}/\mu\text{l}$ for single stranded RNA.

2.8 cDNA synthesis

cDNA was synthesized from total RNA using SuperScript First-Strand cDNA synthesis system and random hexamers (Invitrogen, Carlsbad, CA), according to the manufacturer's instructions.

2.9 Genomic DNA whole-genome amplification

Genomic DNA was whole genome amplified (WGA) using the REPLI-G whole genome amplification kit (Qiagen, Crawley), according to the manufacturer's instructions.

2.10 Fluorescence in situ hybridisation (FISH)

Chromosome paints were used for chromosomes 1, 4, 5, 6, 8, 11, 12, 13, 14, 15, 17, 18 and 22, and a specific probe for *MYC* (all from Cambio). For more recent studies clones were selected from the Sanger Clone Library to probe a specific region at 7q34, and purchased from Geneservice. BAC clone RP11-837G3 was used to target *BRAF* in combination with a control probe to 7p (BAC clones RP11-251I15 + RP11-746C13). Here clones were obtained as stabs, and were plated and cultured on agar plates to obtain single colonies using a standard protocol. DNA was isolated using the Qiagen large construct kit. Clone DNA was labelled by nick-translation with the BioNick kit using biotin-14-dATP (Invitrogen) or digoxigenin-11-dUTP (Roche). Each clone was initially hybridised to normal human metaphase chromosomes to confirm correct localisation.

Prior to hybridisation the chromosomes on each slide were aged with microwave irradiation for 2.5 minutes, using a microwave (Sharp Carousel II) set at medium/high power. To denature the chromosomes, 90µl of denaturation mixture (70% v/v formamide, 2xSSC) was placed on the slide under a 22 x 50 mm coverslip and heated on a hot plate at 73 °C for 3 minutes. The slide was then washed with cool 2xSSC (4 °C)

and dehydrated in an ethanol series (cold 70% then 95% and absolute ethanol, at room temperature, each for 2 minutes).

Labelled DNA (300ng) was precipitated with 5 μ l human Cot-1 DNA (1mg/ml, Invitrogen), 2 μ l salmon sperm DNA (10mg/ml, Invitrogen), 2 μ l 3M NaAc (Fisher Scientific) and 300 μ l absolute ethanol (final volume 309 μ l) on dry ice for 30 minutes (or -20 $^{\circ}$ C overnight). The probe was centrifuged at 15,000 rpm for 15 minutes in an Eppendorf 5415C microcentrifuge at 4 $^{\circ}$ C, dried in air and resuspended in 11 μ l hybridisation mixture (50% formamide, 2x SSC and 10% w/v dextran sulphate). The probe was denatured on a heating block at 85 $^{\circ}$ C for 5 minutes, and then incubated at 37 $^{\circ}$ C for 30 minutes. The slide was then pre-warmed on a heating block at 37 $^{\circ}$ C, and the denatured probe(s) applied to the selected area of the slide. The hybridisation area was sealed under a 22 x 22 mm cover slip with rubber cement, and placed in a moist chamber at 37 $^{\circ}$ C overnight.

Following hybridisation, the slide was washed three times, by gentle shaking for 5 minutes, in 50% v/v formamide and 2x SSC at 42 $^{\circ}$ C, and a further three times in 1x SSC at 60 $^{\circ}$ C. The slides were then incubated in 90 μ l blocking solution containing 4x SSC, 0.1% v/v Tween 20 (SSCT) and 5% w/v Marvel skimmed milk powder and incubated at 37 $^{\circ}$ C for 25 minutes. The blocking solution was drained from the slide, and 90 μ l of the detection antibody (diluted in SSCT) was applied and covered with a coverslip. The slides were then incubated with the appropriate antibodies, anti-avidin-FITC (1:300) or anti-digoxigenin-rhodamine (1:100) at 37 $^{\circ}$ C for 45 minutes, followed by three washes with 4x SSC and 0.1% v/v Tween 20 at 42 $^{\circ}$ C. The chromosomes were then counterstained with DAPI (200 ng/ml) in Citifluor antifade solution and stored at 4 $^{\circ}$ C in the dark.

Slides were examined with a Zeiss Axiophot microscope equipped for epifluorescence using a Zeiss plan-neofluar 100x objective. Separate grey-scale images were captured with a cooled CCD-camera (Hamamatsu). Images were then pseudocoloured and merged. SmartCapture X software (Digital Scientific, Cambridge) was used for image analysis and processing.

2.11 Interphase fluorescence in-situ hybridisation (i FISH)

Dual-colour interphase FISH was performed for selected tumour samples by James Dalton, at St Jude Children's Research Hospital, Memphis. Paraffin tissue sections, 5-8µm thick were used. Probes were derived from BAC clones (Invitrogen, Carlsbad, CA) and labelled with either FITC or rhodamine fluorochromes.

The probe mixtures were diluted 1:50 in DenHyb buffer (Insitus Biotechnologies, Albuquerque, NM) and co-denatured with the target cells on a slide moat at 90°C for 12 minutes. Slides were incubated overnight at 37°C on a slide moat and then washed in 4M urea / 2xSSC at 25°C for 2 minutes. Nuclei were counterstained with DAPI (200ng/ml; Vector Labs) for viewing on a Nikon Eclipse E800 fluorescence microscope equipped with a 100-watt mercury lamp; FITC, Rhodamine, and DAPI filters; 100X PlanApo (1.40) oil objective; and a COHU CCD camera. Images were captured and processed with an exposure time ranging from 0.5-1.5 seconds for each fluorochrome using Cytovision v3.6 software.

2.12 Multiplex fluorescence in situ hybridization (M-FISH)

M-FISH analysis was performed on metaphase spreads prepared from the medulloblastoma cell lines Daoy and HTB187. Slides were pre-treated with 100 µl RNase (100µg/ml) under a 22 x 50 mm coverslip at 37°C for 10 minutes, to remove RNA. Slides were washed twice for 5 minutes in 2 x SSC at room temperature, and then placed in 0.005% v/v pepsin at 37°C for 5 minutes, to remove proteins, cellular debris and aid permeabilization. Following this, the slide was washed twice for 5 minutes in PBS at room temperature, and placed in 1 % formaldehyde fixation solution at room temperature for 2 minutes. The slide was again washed twice in PBS for 5 minutes at room temperature, and dehydrated in an ethanol series (70%, 95% and absolute ethanol) before drying in air. Once dry the slide was placed on a 37°C hotplate to evaporate any remaining ethanol.

The M-FISH Vysis SpectraVysion probe (Abbot Laboratories) was allowed to thaw to room temperature (usually stored at -20 °C), and 10µl of the probe per half-slide was

denatured at 72 °C for 5 minutes. The probe was applied to the slide whilst on the 37 °C hotplate, covered with a 22x22mm coverslip and sealed with rubber cement. The slide was placed in a moist warm chamber, and incubated overnight at 37 °C.

Following hybridisation the slide was washed in 0.4xSSC and 0.3% NP-40 detergent (Calbiochem) at 72 °C for 5 minutes (protecting from light), and then washed in 2x SSC and 0.1% NP-40 at room temperature for 3 minutes before leaving to air dry in the dark. Forty microlitres of DAPI (42 ng/ml) in Citifluor antifade solution was then applied under a fresh 22 x 50 mm coverslip. Images were captured using a Zeiss Axiophot microscope equipped for epifluorescence using a cooled CCD-camera (Hamamatsu). Image analysis was performed using SmartCapture X software (Digital Scientific).

2.13 Polymerase chain reaction (PCR)

PCR is one of the central techniques used in molecular biology, allowing the amplification of DNA from a template using a thermostable DNA polymerase (Mullis *et al.*, 1987). Specifically designed oligonucleotide primers allow the amplification of a target region to take place. This can occur from as little as one molecule of DNA as a starting template. The technique is powerful so it is essential that all reagents are free from any contaminating DNA. To confirm this is the case it is always necessary to run a negative control, which contains all of the reagents for the PCR reaction except the template DNA. Sterile water is used to replace template DNA in the negative control. Hence, any amplification signifies that at least one of the reagents is contaminated, and must be discarded.

The following reagents are required for PCR: template DNA, PCR buffer, deoxyribonucleoside triphosphates (dNTPs), oligonucleotide primers, magnesium chloride (MgCl₂) and *Taq* DNA polymerase, or an equivalent enzyme. The correct concentration of all components is essential for a PCR reaction to be successful. Each PCR reaction must be optimised, by adjusting the PCR cycling conditions, the MgCl₂ concentration or both.

2.13.1 Primer design

Primers were designed using Primer3 software, which may be downloaded from the following website:

http://www.broadinstitute.org/genome_software/other/primer3.html

All oligonucleotide primers were between 18-25bp in length, with an annealing temperature (T_m) between 55-65°C. T_m was estimated using the following formula:

$$T_m = 59.9 + 0.41 \times (\%GC) - 600 / \text{length}$$

Primers were analysed for their dimerization capability and secondary structure formation. When designing primers for sequencing whole genes, primers were designed to amplify each of the coding exons, including the intron-exon boundaries. Fragments were restricted to no larger than 500bp. Any exons larger than 500bp were split into overlapping fragments. A minimal distance of 35bp was left between the primer and the exon/intron boundary.

2.13.2 PCR reactions

All of the PCRs performed throughout the course of the work for this thesis were carried out using QIAGEN Fast-cycling PCR reagents. This uses the HotStarTaq *Plus* DNA Polymerase, a modified form of recombinant DNA polymerase originally isolated from *Thermus Aquaticus* and cloned in *E.Coli*. The enzyme is supplied in an inactive state, with no polymerase activity at ambient temperatures. This prevents extension of nonspecifically annealed primers and primer-dimers formed at low temperatures during the initial stages of setting up the PCR.

Each PCR reaction consisted of 25-50 ng genomic DNA, 0.5µM Forward and Reverse primers, QIAGEN Fast Cycling PCR Master Mix (containing dNTPs and magnesium chloride) and sterile water to a final volume of 20µl.

The PCR reactions were performed on a Peltier thermal cycler PTC-225 (MJ Research). PCR cycling conditions were as follows: initial denaturation for 3 minutes at 95°C followed by 34 cycles of denaturation at 95 °C for 30 seconds, annealing at 57 °C for 30

seconds and extension at 72 °C for 1 minute, followed by final extension at 72 °C for 10 minutes. Once complete, the reaction was kept at 4 °C or on ice. In some reactions the annealing temperature and extension times were varied in order to optimise the production of clean PCR products. If there was difficulty in obtaining a specific PCR product when using certain primers the sterile water was replaced by Q-solution, provided within the QIAGEN Fast Cycling PCR kit. This reagent alters the melting behaviour of DNA, and may be used to optimise annealing of primers that fail to work well under standard conditions.

2.13.3 Agarose gel electrophoresis of DNA

Separation of DNA fragments within a PCR was performed by agarose gel electrophoresis. This technique exploits the negative charge of DNA by causing it to migrate through an agarose gel towards a positive terminal (anode). The rate of progression through the gel decreases as the size of the DNA molecule increases. DNA was visualized by the addition of Ethidium bromide to the gel. Ethidium bromide intercalates between the bases within the DNA molecule and will fluoresce in the presence of ultraviolet light.

One and a half percent agarose gels were prepared, with agarose and 1xTBE. A domestic 600W microwave (Sharp Carousel II) was used to melt the agarose, which was then allowed to cool to around 50°C before ethidium bromide was added, to a final concentration of ~ 0.5µg/ml. The gel was poured into a casting tray and left to set, before being transferred to an electrophoresis tank, and covered with 1xTBE.

DNA samples were mixed with a loading buffer containing OrangeG dye and glycerol (see Appendix A 8.1.3 for recipe). These migrate at the same speed as 300-400bp DNA fragments, allowing their progress to be followed within the gel. Glycerol increases the density of the DNA samples, allowing them to sink into the wells of the gel. Samples were loaded into wells and run for ~30 minutes at 120 volts. All gels were run with a DNA ladder (Hyperladder I or II, Bionline) in the first and last wells, to allow sizing of DNA products. Agarose gels were then visualised on a UV transilluminator (G-Box, Syngene) and photographed using Genesnap (Syngene).

2.14 DNA purification

Prior to DNA sequencing, PCR products of interest were purified by gel purification or by an ExoSap reaction. This is necessary to remove unwanted dNTPs and primers from the PCR product, which would otherwise contaminate the sequencing reaction. ExoSap was used if the PCR product appeared as a clean band on agarose gel, with no additional products or primer dimer seen. Gel purification was used if PCR products additional to the band of interest were seen.

2.14.1 ExoSap

The ExoSap-IT PCR product clean-up kit (USB corporation, USA) was used according to the manufacturer's instructions.

2.14.2 Gel purification

The PCR product of interest was visualised within the gel using a UV-transilluminator, and excised from the gel using a sterile scalpel. DNA was then extracted from the gel slice using the QIAquick Gel extraction kit (Qiagen) in accordance with the manufacturer's instructions.

2.15 Direct sequencing

Direct sequencing allows the detection of a sequence of nucleotides within a known segment of DNA. The most commonly used method for direct sequencing is Sanger dideoxy sequencing (Sanger sequencing) (Sanger *et al.*, 1975). Sanger sequencing chemistry is based on the use of dideoxynucleotides (ddNTPs). These contain a hydrogen on the 3' carbon, instead of the hydroxyl (OH) group found in deoxynucleotides. During the sequencing reaction, a phosphodiester bond cannot form between the dideoxynucleotide and the incoming nucleotide. Hence, DNA chain synthesis is terminated. Dye-terminator sequencing uses fluorescent labelling, specific for each of the four chain terminating dideoxynucleotides (ddATP, ddGTP, ddCTP, or ddTTP) within a single reaction. All sequencing reactions were carried out using

'BigDye Terminator Sequencing' kit version 3 (Applied Biosystems) (Rosenblum *et al.*, 1997).

DNA target sequences were amplified by PCR, purified by ExoSAP, or by gel purification, and then sequenced in both directions using the di-deoxy chain termination method. Each 20µl sequencing reaction contained: 8µl BigDye terminator mix, 4-10ng template DNA, 3.2µl of primer (either forward or reverse) at a final concentration of 3.2pmol, and sterile water to a final volume of 20µl.

The thermal cycling conditions were as follows: initial ramp to 96°C by 2.5°C per second, 96°C for 1 minute followed by 24 cycles of (96 °C for 10 seconds, ramp to optimum annealing temperature for specific PCR by 1°C/second then hold annealing temperature for 5 seconds, ramp to 60°C at 1°C/second followed by 60°C for 4 minutes), ending at 12 °C.

The sequencing reactions were subsequently cleaned up to remove unincorporated dye terminators. If there were fewer than 30 samples DyeEx 2.0 spin columns (Qiagen, Crawley) were used, for larger numbers of samples DyeEx 96-well cleanup plates (Qiagen, Crawley) were used, both according to the manufacturer's instructions.

Staff at the equipment park within Cancer Research UK London Research Institute performed the final sequencing stage, using an ABI 3730 DNA sequencer. Results were screened by eye using the Applied Biosystems software packages Sequence Scanner version 1.0 and Variant Reporter version 1.0. Mutational analysis was performed by eye.

Chapter 3. DNA copy number changes by array CGH

3.1 Introduction

DNA copy number gains and losses are known to contribute to many genetic syndromes and developmental abnormalities. For example, Down's syndrome is due to the addition of a whole or partial copy of chromosome 21 and Cri du chat syndrome is caused by partial deletion of chromosome 5p. Changes in DNA copy number occur in most human cancers, and deletions and gains of DNA copy number may lead to alterations in the expression of tumour suppressor genes and oncogenes, respectively (Hanahan *et al.*, 2000). It is important to characterise the DNA copy number changes in cancers, in order to attempt to understand the pathways and processes involved in tumorigenesis, and thereby to improve tumour diagnosis, classification and treatment.

DNA copy number gains and losses were first identified by karyotyping. This technique uses dividing cells to produce metaphase chromosome spreads. Cells are dropped onto a glass slide, fixed, stained and examined under a microscope. Whole or partial chromosome gains and losses may be determined using this technique, but smaller, subtle changes may not be identified. It is often difficult to obtain dividing cells in many solid cancers, making karyotyping a technical challenge. Furthermore, cancers frequently contain extremely complex chromosomal rearrangements, which are almost impossible to interpret using conventional karyotyping. Hence, new methods were required to investigate DNA copy number alterations in cancer in more detail.

Comparative genomic hybridisation (CGH) was developed in 1992 (Kallioniemi *et al.*, 1992), and entirely revolutionised the fields of genetics and cancer biology. CGH enabled DNA copy number gains and losses to be analysed across an entire genome, in a single hybridisation experiment. Equal amounts of DNA from a reference sample (control) and a test sample (tumour) are differentially labelled with two fluorochromes, and competitively bound to the hybridisation platform. Unlabelled human Cot-1 DNA (placental DNA 50-300bp in size, enriched for repetitive DNA sequences) is used to block non-specific hybridisation across repetitive sequences within heterochromatin, e.g. across centromeric regions (Albertson *et al.*, 2003, Lichter *et al.*, 2000). The relative intensities of the two fluorochromes are measured at a given locus. Differences

in signal intensity represent regions of gain or loss of DNA sequence in the test sample when compared to the reference.

CGH has transformed the study of cancer and inherited conditions, by removing the need to culture cells to produce karyotypes. Previously, the majority of investigations into the molecular basis of cancer had been based on leukaemias, because of the ease of culturing cells and obtaining samples (Mitelman *et al.*, 1997). The few successfully characterised solid tumours contained hugely rearranged karyotypes, making the interpretation of results difficult. Many solid tumours, however, remained inaccessible to detailed study because of an inability to obtain chromosome spreads.

Initially, the CGH technique used human metaphase chromosome spreads as the hybridisation platform, and was termed metaphase CGH (Kallioniemi *et al.*, 1992). The use of metaphase spreads limited the resolution for CGH to between 3~10Mb (Lichter *et al.*, 2000, Beheshti *et al.*, 2003). Subsequently, the technique was replaced by array CGH, using mapped fragments of DNA sequence, arrayed on a glass slide. Array technology allowed much higher resolution mapping of DNA copy number changes, at ~1Mb resolution. Three types of array CGH platform are available: the first utilises genomic DNA clones from bacterial artificial chromosomes (BAC) and P1 artificial chromosomes (PAC) (Pinkel *et al.*, 1998); the second uses complementary DNA (cDNA) clones which are generated by reverse transcription of RNA to DNA before insertion into a cloning vector (Pollack *et al.*, 1999). Recently a third type of platform was developed, using oligonucleotides to perform mapping at higher resolution, for example probes may be spaced at 50 or 140bp intervals for a fine-tiling path array to investigate a particular chromosomal region of interest (Selzer *et al.*, 2005).

The introduction of metaphase CGH, and then array CGH in 1998 allowed investigation of copy number alterations previously beyond the reach of existing technologies to be conducted, in large numbers of solid tumours. This groundbreaking technique paved the way for numerous discoveries in genetic syndromes and cancer. Associations between DNA copy number alterations and prognosis have been identified in many cancers, including prostate cancer, breast cancer and lymphomas (Martinez-Climent *et al.*, 2003, Paris *et al.*, 2004, Callagy *et al.*, 2005, Rubio-Moscardo *et al.*, 2005, Chin *et al.*, 2006).

Knowledge of the copy number variations found in normal individuals is required to inform detailed studies of DNA copy number alterations in disease. Array CGH technology has also enabled large regions of copy number variation or indels (insertion/deletions) to be identified in normal individuals (Iafrate *et al.*, 2004, Sebat *et al.*, 2004). These ranged in size from tens of kilobases to megabases. Such large regions of copy number variation (CNV) were previously unknown in normal individuals. These first two studies identified over 200 large regions of CNV in 75 individuals, and led the way for larger studies to identify regions of CNV within the normal population. The first of these was the HapMap project, which constructed a CNV map of the human genome by investigating 270 individuals from four populations (Redon *et al.*, 2006). These comprised parent and offspring trios from Nigeria and Utah, and unrelated individuals from Japan and China. Investigations were performed using CGH and single-nucleotide polymorphism (SNP) arrays. One thousand four hundred and forty seven regions of CNV were identified, covering 360 megabases (around 12% of the genome) and containing hundreds of genes. Results from these and later studies are freely available to researchers. Common regions of CNV within the genome from normal individuals are catalogued, mapped and regularly updated on web-based genome browsers such as Ensembl and UCSC.

Paediatric astrocytomas had previously been investigated using conventional cytogenetic approaches and metaphase CGH (Raffel, 1996, Schrock *et al.*, 1996, Bhattacharjee *et al.*, 1997, Zattara-Cannoni *et al.*, 1998, Sanoudou *et al.*, 2000, Warr *et al.*, 2001). High-grade paediatric astrocytomas have been shown to contain some of the changes present in adult astrocytomas, including loss of chromosome 10q and the *PTEN* locus, but other changes commonly seen in adult tumours, such as *EGFR* amplification and 17p deletions have not been found at similar frequencies (Rickert *et al.*, 2001, Warr *et al.*, 2001). Most low-grade paediatric astrocytomas display balanced karyotypes by conventional cytogenetics and metaphase CGH (Raffel, 1996, Bhattacharjee *et al.*, 1997, Zattara-Cannoni *et al.*, 1998, Sanoudou *et al.*, 2000). Whole chromosome gains have been identified in paediatric low-grade astrocytomas, most commonly involving chromosomes 5, 7, 19 and 22 (reviewed by Sanoudou) (Sanoudou *et al.*, 2000). The majority of these studies were performed using low-resolution techniques. I decided to use a higher resolution approach to investigate paediatric low-grade astrocytomas using array CGH. Initially the technique was optimised using control normal DNA, and DNA from the paediatric medulloblastoma cell lines Daoy and HTB187. Cell lines were used

as a renewable resource before using tumour DNA samples. I investigated 20 paediatric low-grade astrocytoma samples when I was satisfied with the array CGH optimisation. Initially I was blinded to all clinico-pathological details, other than the tumour site, so began by comparing tumours from the cerebellum with non-cerebellar tumours. I wanted to investigate whether tumours arising from different sites within the brain carried distinct or similar molecular aberrations. Any aberrations identified might provide an insight into the pathways involved in tumour development, within different regions of the brain. When the clinico-pathological details were revealed, it became apparent that the group of non-cerebellar tumours contained different pathologies, which could not be compared directly as a group with the pilocytic astrocytomas of the cerebellum. Hence, the study attempted to identify distinct molecular aberrations within the group of cerebellar pilocytic astrocytomas in comparison to the grade II astrocytomas which arose outside the cerebellum.

3.2 Aim

The aim of the work described within this chapter was to investigate genome-wide DNA copy number changes in paediatric low-grade astrocytomas using array CGH.

3.3 Array CGH method

3.3.1 Array CGH slides

Array CGH was performed using slides produced by the Institute of Cancer Research in Sutton (Hughes *et al.*, 2005). The slides contained 3,122 bacterial artificial chromosome (BAC) and P1 artificial chromosome (PAC) clones. The BAC clones were based on a previously reported clone set, produced using modified degenerate oligonucleotide-primed PCR (DOP-PCR), which allows random amplification of DNA from any source (Fiegler *et al.*, 2003). Clones were selected from the Golden Path published in 2001, with median separation of 1Mb across the genome (range 500kb-4.5Mb). PAC clones were used to cover telomeres, and were selected from a previously described clone set (Knight *et al.*, 2000). Chromosome coverage excluded centromeres, where clone separation was greater than 1Mb, and the short arms of the acrocentric chromosomes (chromosomes 13, 14, 15, 21 and 22). BACs and PACs were spotted in duplicate across

CodeLink activated slides. These slides are specifically designed for microarrays, and are coated with a hydrophilic polymer which covalently binds amine-modified DNA. The slides also had 838 empty spots and 36 spots containing *Drosophila* DNA, as a non-specific hybridisation control.

3.3.2 Random labelling of DNA

DNA from the test sample was labelled with Cy5 and reference control DNA with Cy3 using a Bioprime labelling kit (Invitrogen). Equal amounts of each DNA sample were used, to ensure equality for each sample during the competitive array slide hybridisation step. Sixty microlitres (μl) of Random Primer Solution was added to 450ng of DNA, and made up to 126 μl with distilled water. The solution was then mixed briefly and incubated at 100°C for ten minutes to denature the DNA. The mixture was cooled on ice for 2 minutes. Fifteen μl of 10x dCTP mix; 6 μl of either Cy3 or Cy5 labelled dCTP and 3 μl of Klenow fragment (5 Units/ μl , Applied Biosystems) were then added in sequence to give a total volume of 150 μl . The solution was mixed gently, and incubated at 37 °C overnight. The reaction was stopped by adding 15 μl of EDTA, then put at 100 °C for five minutes before cooling on ice.

Unincorporated labelled nucleotides were removed from the solution using G50 Microspin columns (Amersham) as illustrated below. The columns were prepared by vortexing, mixing the liquid within. Each column base was snapped off, and the lid opened by a quarter turn before placing in a 1.5ml Eppendorf tube. The columns were centrifuged for one minute at 4,000 rpm in an Eppendorf 5415C microcentrifuge, creating a gel filtration column. Excess liquid at the base of each column was blotted off with clean tissue paper.

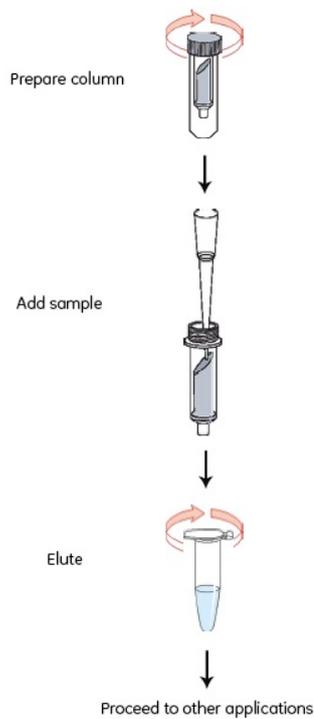


Figure 3.1 G50 spin column preparation.

The DNA is cleaned up using a G50 spin column to ensure removal of unincorporated labelled nucleotides. Diagram reproduced from G50 spin column instruction handbook (Amersham).

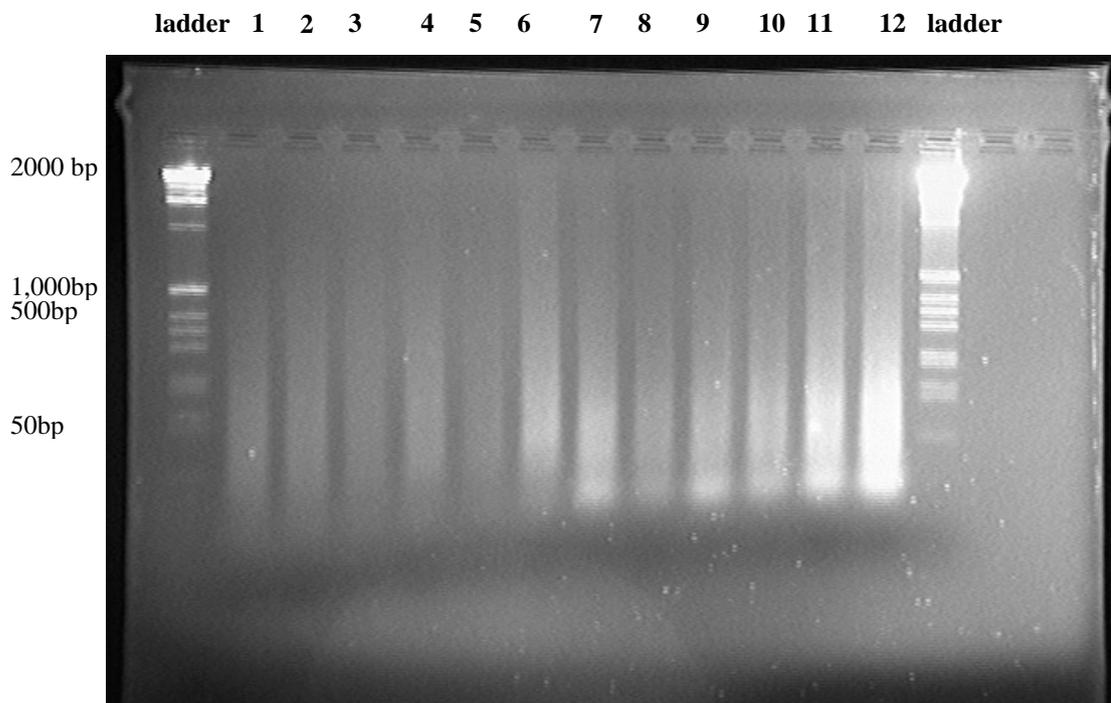


Figure.3.2 Agarose gel image of 12 labelled DNA samples.

The DNA is run on a 1.5% agarose gel to assess the level of fragmentation achieved when compared to the DNA ladder (50-2000bp). The white ‘smear’ for each sample trace indicates that adequate fragmentation has been achieved.

The columns were placed in fresh 1.5ml Eppendorf tubes, and Cy-dye labelled DNA solution was loaded onto the centre of the resin gel column. Three gel columns were used per 150µl sample, as the maximum volume for loading each column was 50µl. The spin column lids were closed tightly, and then loosened by a quarter turn before being centrifuged for two minutes at 4,000 rpm. The eluted DNA was collected into a fresh 1.5ml Eppendorf tube. An aliquot (5µl) of labelled DNA was run on a 1.5% Agarose gel. If adequate fragmentation had occurred, a 'smear' of DNA fragments was visible, with fragments ranging in size from 50 to 500 base pairs when compared to a 1kb ladder (Figure 3.2).

A further 1µl of DNA was analysed using a NanoDrop ND-1000 UV-Vis Spectrophotometer (ThermoScientific, UK). The 'Microarray' setting was used to measure the absorbance of the fluorescent dye, to quantify the concentration of Cy3/Cy5 incorporation (pmol/µl) in the labelled DNA. Five pmol/µl was found to be the minimum concentration of Cy3 or Cy5 necessary for optimum results when scanning the arrays. Samples with fluorescent dye incorporation below this concentration were difficult to visualise during scanning. Random labelling was repeated for DNA samples that did not reach the 5pmol/µl threshold.

3.3.3 Precipitation of DNA prior to hybridisation

Before hybridisation, it is necessary to block repetitive sequences present in the labelled DNA and the array platform clones. Cot-1 DNA was used to block repetitive sequences in the labelled DNA on the surface of the array slide, and a combination of Cot-1 and Herring Sperm DNA was used.

The pre-hybridisation mixture (80µl Herring Sperm DNA, 135µl human Cot-1DNA, 23µl 3M NaAc (pH 5.2) and 400µl cold absolute ethanol) and hybridisation mixture (180µl Cy3 labelled reference DNA, 180µl Cy5 labelled test DNA, 135µl human Cot-1DNA, 55µl 3M NaAc (pH 5.2) and 1000µl cold absolute ethanol) was prepared, and then precipitated at -20 °C overnight or at -80 °C for 1 hour.

3.3.4 Preparation of array slide prior to hybridisation

A frame was prepared from thermo-stable adhesive plastic PCR sheets (ABgene, UK). This frame allowed both the pre-hybridisation and hybridisation mixtures to be retained on the surface of the slide, within the reservoir formed by the frame. Six individual adhesive PCR sheets were stuck together, one on top of the other. Great care was taken to align the sheets as they were stuck together, and air bubbles were pressed out using a plastic slide box to smooth the surface, creating one thick plastic sheet. Any air bubbles or irregularities between the sheets would allow the hybridisation mixture to leak out of the well during the course of the experiment. The exact dimensions of the array slide were drawn on the reverse of the sheet in pencil, including the dimensions of the area containing the array spots. The outer edges of the frame were cut to size using a guillotine. The inner edges were cut out using a scalpel to create the reservoir. The adhesive backing was removed, and the plastic frame placed onto the surface of the array slide. Care was taken both to avoid touching the array spots and to leave a clear border around them with the plastic frame. The barcode at the base of each slide was also left uncovered, to allow reading during the subsequent scanning process. The frame was pressed into place and a pair of forceps used to score down each side of the frame, to ensure adhesion of the reservoir to the slide surface.

3.3.5 Pre-hybridisation

Three hundred microlitres of hybridisation buffer (Appendix B: Reagents prepared for array CGH) was pre-heated to 80°C in a heat block. A humidity chamber was prepared for the hybridisation in a fume hood. Plastic food containers (Lock and Lock) were used, as they were both air- and water-tight due to a rubberised seal between the lid and container. The containers were covered in opaque tape to ensure they were protected from light. Whatman filter paper was used to cover the base of the humidity chamber, and drenched in a mixture of 2xSSC and 20% v/v formamide, to ensure humidity within the chamber.

The precipitated pre-hybridisation and hybridisation mixtures were spun for 15 minutes at 14,000 rpm in an Eppendorf 5415C micro-centrifuge. The supernatant above the pellets was removed, and 500µl 80% ethanol was added to each tube before

centrifuging at maximum speed for 5 minutes. The ethanol supernatant was removed, and the tubes spun for a further minute at maximum speed. Any remaining supernatant was removed with a 10µl pipette tip, and the pellet re-spun until dry.

Sixty microlitres of pre-heated hybridisation buffer and 36ng yeast tRNA (Invitrogen, diluted in distilled water for 60 minutes to a final concentration of 6ng/µl) was added to the hybridisation pellet, which was then resuspended in a heat block at 80°C for 5 minutes. The sample was mixed, and heated for a further 10 minutes at 80°C to allow DNA denaturation. The tube was mixed, and incubated in the dark for 60 minutes.

One hundred and sixty microlitres of pre-heated hybridisation buffer was added to the pre-hybridisation pellet, and allowed to resuspend in a heat block at 80°C for 5 minutes. The sample was heated for 10 minutes at 80°C, mixed and given a pulse spin to remove air bubbles. The pre-hybridisation mixture was applied to the array slide surface, inside the prepared reservoir. Care was taken to avoid touching the array spots with the pipette tip, and to ensure even coverage of the slide with the mixture. Any air bubbles were removed. The slide containing the pre-hybridisation mixture was sealed in the humidity chamber, and placed in a hybridisation oven at 37 °C, on a rocking platform at 5rpm for 1 hour.

3.3.6 Hybridisation

The slide was removed from the humidity chamber. The hybridisation mixture was removed from the dark, mixed thoroughly and given a pulse spin in an Eppendorf 5415C microcentrifuge to remove air bubbles. The hybridisation mixture was then applied to the surface of the slide, into the reservoir containing the pre-hybridisation mixture. The slide was rocked gently to ensure even coverage and any air bubbles were removed. The slide was resealed in the humidity chamber and replaced in the hybridisation oven at 37°C. The slide was rocked gently at 5rpm for 48 hours. The humidity chamber was rotated by 90° after 24 hours, to ensure even coverage of the slide surface by the hybridisation mixture.

3.3.7 Washing

The slide was removed from the humidity chamber, and placed in 2xSSC and 0.03% w/v SDS at room temperature for 1 minute. The slide was then removed, and the plastic frame removed with forceps. The slide was then placed in a Coplin jar containing 2xSSC/0.03% SDS heated to 65 °C, wrapped in foil (to protect from light) and shaken vigorously on a rocking platform for 5 minutes. The slide was transferred to a glass slide holder, open at the base. This was submerged in a large clean container with 1 litre 0.2xSSC at room temperature, wrapped in foil to protect from light, and shaken vigorously for 20 minutes. Three further 20 minute washes were performed in 0.2xSSC, using 1 litre of fresh solution each time. After that a further wash was performed in a Coplin jar containing room temperature PBS/0.05% Tween20 for 10 minutes. Finally, the slide was rinsed in a Coplin jar containing distilled water, and immediately taken for drying.

3.3.8 Drying

The wet slide was placed in a slide box with lint-free tissue paper covering the base, and dried by centrifugation in a plate centrifuge at 700rpm for 3 minutes at room temperature. The slide was stored in a light-proof box until scanning was completed.

3.3.9 Scanning and data acquisition

The slides were scanned using an Axon GenePix 4000B microarray scanner, which uses two lasers to scan the slide simultaneously, by excitation of the fluorophores on the surface. One laser was optimised to identify Cy3 (green) at a wavelength of 532 nm, the other identified Cy5 (red) at 635 nm (Figure 3.3). The intensities of fluorescence at each array spot were collected and quantified using GenePix Pro 6.0 software (Axon Instruments, USA).

The file containing the GenePix array list (.gal) was loaded into the GenePix software. The .gal file describes the size, layout and position of the blocks of spots (known as features), and the names and identifiers of the clones printed within each feature. Initial

background normalisation was performed using the software auto-balance function to even out brightness across the slide. Features were subsequently selected using the automatic grid and feature identification functions within the software. The default feature diameter was 150 μm , although this could be altered on an individual spot-by-spot basis if the feature diameter was smaller than this. The feature indicator is the circle put around each spot, from within which data were collected. The mean, median and standard deviation of pixel intensities within each feature indicator were collected for each feature. The software parameters for acquisition of data from a feature were set as follows. A 'good' feature would contain fluorescent signal from either wavelength covering >75% of the spot diameter at >2 standard deviations above the background signal level. The background signal level was defined as the level of fluorescence in the region surrounding each spot (Figure 3.4). Poorly hybridised arrays were rejected after visual inspection during the scanning process.

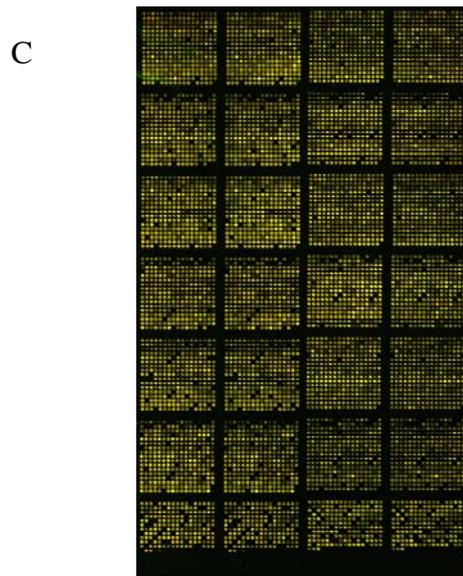
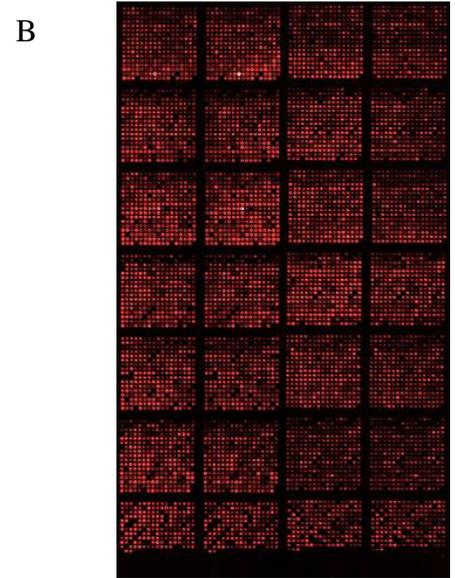
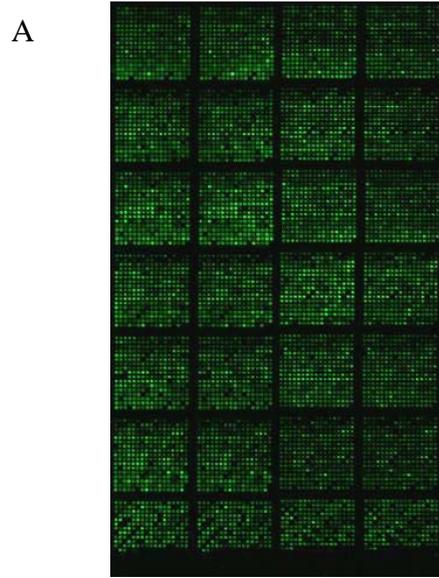


Figure 3.3 array CGH scan images

- (A) Cy 3 image (green)
- (B) Cy 5 image (red)
- (C) Combined scan images of Cy 3 and Cy 5 (yellow).

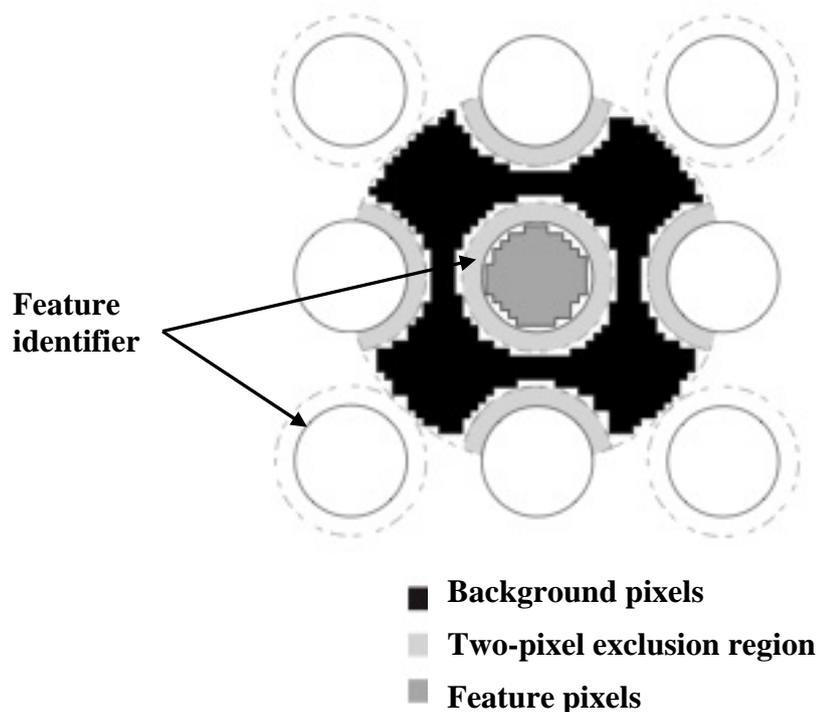


Figure 3.4 Definition of the background region by GenePix.

The black region represents pixels used to calculate the background signal; the dark grey represents the pixels used to calculate the feature signal. The pixels within the light-grey region are excluded. Figure adapted from GenePix Pro 6.0 users handbook.

Features where the signal intensity did not fulfil these inclusion criteria were included in the data collection, but excluded from the subsequent analysis. The array data were acquired as a binary GenePix settings (.gps) file from the feature identifiers surrounding each feature signal.

Each block of features was checked by eye to ensure that features flagged as ‘good’ were truly representative and were not areas of fluorescent debris. Once all features had been checked, the data were saved as GenePix results files (.gpr), which are tab-delimited text files suitable for further analysis.

3.3.10 Data analysis

The .gpr data files were analysed with the assistance of Dr Gavin Kelly at the Bioinformatics and Biostatistics group at the London Research Institute, Cancer Research UK. Analysis was performed in the same way for all of the array experiments using Bioconductor (Gentleman *et al.*, 2004).

An initial background subtraction was performed using GenePix derived values. The test (tumour) samples were labelled in red (Cy 5) and the controls were labelled in green (Cy3). Background corrected values were calculated for each feature as follows:

$$\begin{aligned}\text{Red (R)} &= \text{Red} - \text{local red background signal} \\ \text{Green (G)} &= \text{Green} - \text{local green background signal}\end{aligned}$$

A differential value (M) was then calculated for each feature as follows:

$$M = \log_2[\text{R/G}] \quad \text{i.e. } \log_2 [\text{test signal/control signal}]$$

Any spots with negative values after background correction were deemed to have no signal, and were discarded. A \log_2 scale was used to allow both test and control values to be on the same scale, with a predominantly red signal measured from 0 to ∞ and a predominantly green signal measured from $-\infty$ to 0. Zero change along the x-axis corresponds to no overall gain or loss. Assuming both samples are diploid, gain of 1 copy of DNA relative to the reference sample (ratio 3:2) is represented as +0.58 on the \log_2 scale. Gain of 2 copies of DNA, or gain of the X chromosome in a female versus male control experiment (ratios 4:2 or 2:1) would be represented as +1 by the \log_2 scale. Losses of 1 or 2 DNA copies relative to the reference sample are represented as -1 and $-\infty$ respectively. Hence, +1 represents doubling of DNA copy number, and -1 represents halving of DNA copy number on the \log_2 scale. This also makes the assumption that there is a linear relationship between DNA copy number and signal intensity. However, tumour samples contain varying amounts of non-malignant cells, which affect the ratios for DNA copy number gain and loss. For the purposes of this thesis, I defined a single copy gain to be ≥ 0.3 above the X-axis, and a single copy loss to be ≤ -0.5 below the X-axis.

Features containing *Drosophila* DNA were present on the array to be used as non-specific hybridisation controls. However, many slides were found to contain high background signal or fluorescent debris within the spots defined as empty or containing *Drosophila*. It was apparent that these signals were due to non-specific binding of DNA, so these features were not used in the analysis. Hence, normalisation was performed using data only from the features defined as 'good'. We did not include data from features flagged as 'bad', or from features that were empty or contained *Drosophila* DNA.

We used a median normalisation, which results in each slide having the same median value (zero) as any other slide, where the median is taken across all the 'good' features of that slide. This is reasonable under the assumption that no more than 50% of the clones across the genome are gained in any particular sample, and that no more than 50% are lost. Each BAC clone had two replicates on the array slide. The mean M value was taken across the two replicates on the array. If one replicate had failed, the successful feature value was used. Where multiple array CGH hybridisations were performed for the same tumour sample, the mean M value across all the slide replicates was calculated, to act as a single measurement per-clone per-tumour. We examined the data on both a per-tumour basis and a per-slide basis.

Segmentation (smoothing) of the data was performed using circular binary segmentation. Segmentation allows the data to be divided into regions of equal copy number, and identifies points of DNA copy number change. This process reduces the 'noise' due to the array CGH method. It is unrealistic to expect adjacent clones to have identical aCGH measurements, even when they have identical copy number. Segmentation attempts to remedy this. We followed the approach used by Olshen and Venkatraman et al (Olshen *et al.*, 2004) using the DNA copy number library by the same authors [Venkatraman ES and Olshen A. DNA copy number data analysis. R package version 1.18.0.] within Bioconductor (Gentleman *et al.*, 2004). Briefly, this recursively attempts to identify the two regions within a chromosome that differ most, and then attempts to subdivide these regions, until no further subdivisions are found that are statistically significant. We carried out the standard procedure of merging adjacent regions that had M values within 0.25 of each other.

Tumour samples were further analysed by separating them into two groups; cerebellar and non-cerebellar samples. Data were pooled for each group, and the mean value for each clone was calculated between all patients in the group (using the per-tumour M values). The grouped data then underwent further segmentation. Finally, the segmented grouped data from the non-cerebellar samples were subtracted from the cerebellar sample grouped data (delta cerebellar group), to identify any significant differences present between the two groups.

The average intensity (A) for each feature was calculated as follows;

$$A = 1/2 \times [\log_2 R + \log_2 G]$$

MA plots were calculated for each slide and for each tumour sample using a single measurement per-clone per-tumour. The MA plot gives an overview of the raw data, under the assumption that the majority of features would contain equal amounts of red and green signal. Hence, the majority of data points would be at zero. The MA plot also gives an indication of any bias due to excessive amounts of green or red fluorescence (possibly due to failure of one fluorochrome), which would skew the plot from the zero axis.

All of the data plots shown in this thesis were obtained from Bioconductor, although visualisation tools improved over time. Plots obtained later, from the tumour versus control studies are displayed in a different way from earlier experiments. However all experiments were analysed in the same way as described above. For the early experiments involving male versus female and cell line characterisation, large numbers of data points were missing from the 23 clone duplicates representing the Y chromosome on some of the arrays. This caused problems when using the algorithms; hence the Y chromosome was excluded from the earliest analysis. This may have been a batch problem, as loss of data from the Y chromosome was not encountered in later experiments involving tumour DNA. Information on both the sex chromosomes were included in these analyses.

The following weblink provides access to the raw data plots for all tumour samples:

<https://bioinformatics.cancerresearchuk.org/~kelly02/index.php?id=Ruth>

Data are available for whole karyotype aCGH, MA plots and individual chromosomes. Data may be viewed by individual array, by tumour sample and by group. Grouped data are available for cerebellar samples, non-cerebellar samples and the delta cerebellar group (identifies differences between the two groups). The weblink also contains access to the Bioconductor analysis script that generated the visualisations, provided by Gavin Kelly.

3.4 Array CGH results

3.4.1 Male versus female control

Control aCGH experiments were performed using male versus female and male versus male (same/same) DNA from two individuals in the lab. Dye-swaps were performed for the male versus female experiments: each DNA was labelled with either Cy5 or Cy3 for two replicates to remove any possible competitive advantage due to dye incorporation.

Female versus male aCGH analysis showed relative gain of an additional copy of the X chromosome (as expected), but also revealed a number of unexpected small regions of gain and loss across all of the chromosomes on the whole genome plot (Figure 3.5). The Y chromosome was not included in the earliest analysis using Bioconductor, as insufficient data were present.

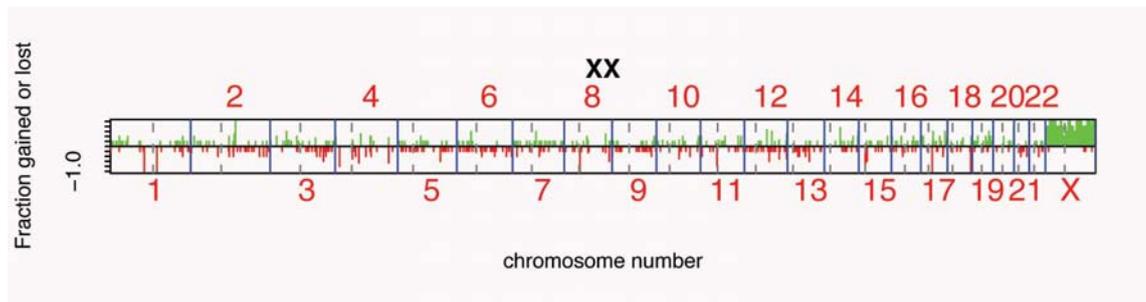


Figure 3.5 Composite array CGH whole genome data plot.

A \log_2 scale is used for each chromosome (1 – 22 and X). The baseline (zero) indicates an equal amount of DNA at each point in the test and control samples. Green lines above the baseline indicate a gain of DNA relative to the control sample. Red lines below the baseline indicate a loss of DNA relative to the control. This plot compares female (test) DNA versus male (control) DNA, and clearly demonstrates that an additional copy of the X chromosome is present. The plot also shows multiple regions of gain and loss across the genome, likely to be due to genetic polymorphic differences between the two individuals.

This suggested that the additional regions of gains and losses were probably due to genetic polymorphisms between the two normal individuals. The affected clones were not found to be mis-assigned after checking on Ensembl, or among the few clones known to give unexpected ratios from the experience of other researchers using these arrays (Appendix B 8.1 Array CGH clones known to give unexpected ratios). No significant differences were noted between the dye-swap replicates, dye swaps were therefore not performed for later experiments.

Pooled male DNA from 21 individuals was used as the control DNA for subsequent array CGH experiments involving tumour samples, to reduce the possibility of identifying copy number changes due to polymorphisms between two individuals.

3.5 Medulloblastoma cell lines versus male control

Further experiments were performed using cell lines to assess whether the technique could correctly identify known regions of DNA copy number gain and loss. Since no cell lines derived from paediatric low-grade astrocytomas were available, array CGH was performed using two paediatric medulloblastoma cell lines, Daoy (HTB186) and HTB187 (D341). Both contained chromosomal aberrations including whole chromosome gains, amplifications and deletions.

I had previously characterised these cell lines, using M-FISH and FISH as described in chapter 2, working with Tania Jones and Radost Vatcheva in the Human Cytogenetics laboratory at the London Research Institute. These results are summarised below, with a full description of the karyotype for each cell line and representative M-FISH and FISH images (Figures 3.6-3.11).

3.5.1 Characterisation of Daoy by M-FISH and FISH

Daoy is a cell line derived from a medulloblastoma in a four year old male. It contains a hypertetraploid modal range of 93-99 chromosomes, two normal X chromosomes, and no Y chromosome as is commonly seen in immortalised malignant cell lines. In addition this cell line contained multiple complex chromosomal rearrangements.

Whole chromosome changes demonstrated by M-FISH included gain of chromosome 1 (up to 5 copies, with 3-5 additional marker chromosomes containing chromosome 1 material), loss of chromosomes 4 (2 copies), 13 (3 copies), and numerous other rearrangements (Figure 3.6). These changes were confirmed using FISH.

3.5.2 Characterisation of Daoy by array CGH

DNA from Daoy was compared to single male control DNA by array CGH in two replicates. Array CGH analysis identified the following regions of chromosomal gain and loss:

Copy number gain: 1p32-36, 6p12-25, 7p and 7q, 9q22-34, 15q11.1-14, 17p11.1-13, chr 19, chr 20, chr 21q11.1-22 and chr 22q12.1-13

Copy number loss: 1q12, 3p12-14, 4q11-32, 8p11.2-23, 8q11.1-21.1, 10 cen, 13p and 13q11.1-31, 16q11.1-12.2, Xq13-22, and complete loss of Y.

The composite whole genome plot from the array CGH analysis is shown in Figure 3.10 (A). All of the changes found by FISH and M-FISH were identified by array CGH.

3.5.3 Characterisation of cell line HTB187 by M-FISH and FISH

HTB187 is a hyperdiploid line derived from a medulloblastoma in a three year old male. It has a modal range of 44-50 chromosomes and one copy of X and Y.

M-FISH analysis showed a male karyotype with the following aberrations: trisomy 6, trisomy 8 and trisomy 18, amplification of *MYC* to form a homogeneously staining region (HSR) of chromosome 8 material within a derivative chromosome 11, isochromosome 17q, and translocation of 1q onto 22p, producing a derivative chromosome 22, der(22)t(1;22)(q12;p12) (Figure 3.7).

These changes were confirmed by FISH: the HSR on chromosome 11 was shown to be derived from amplification of *MYC* using a specific *MYC* probe, in conjunction with chromosome paints for chromosomes 8 and 11. One copy of chromosome 8 did not contain the *MYC* signal; this was translocated to chromosome 11 and was highly amplified. Two representative FISH images are shown to demonstrate these findings (Figures 3.8-3.9)

3.5.4 Characterisation of HTB187 by array CGH

DNA from HTB187 was compared to DNA from a single male control by array CGH in two replicates.

Array CGH analysis identified the following regions of chromosomal gain:

Copy number gain: 1q, isochromosome 17q and chromosomes 6, 8 and 18

Copy number loss: 17p

The composite whole genome plot from the array CGH analysis is shown in Figure 3.10 (B). Selected smoothed chromosome plots, demonstrating regions of copy number gain and loss on individual chromosomes are shown in Figure 3.11.

All of the copy number changes found by FISH and M-FISH were identified by array CGH.

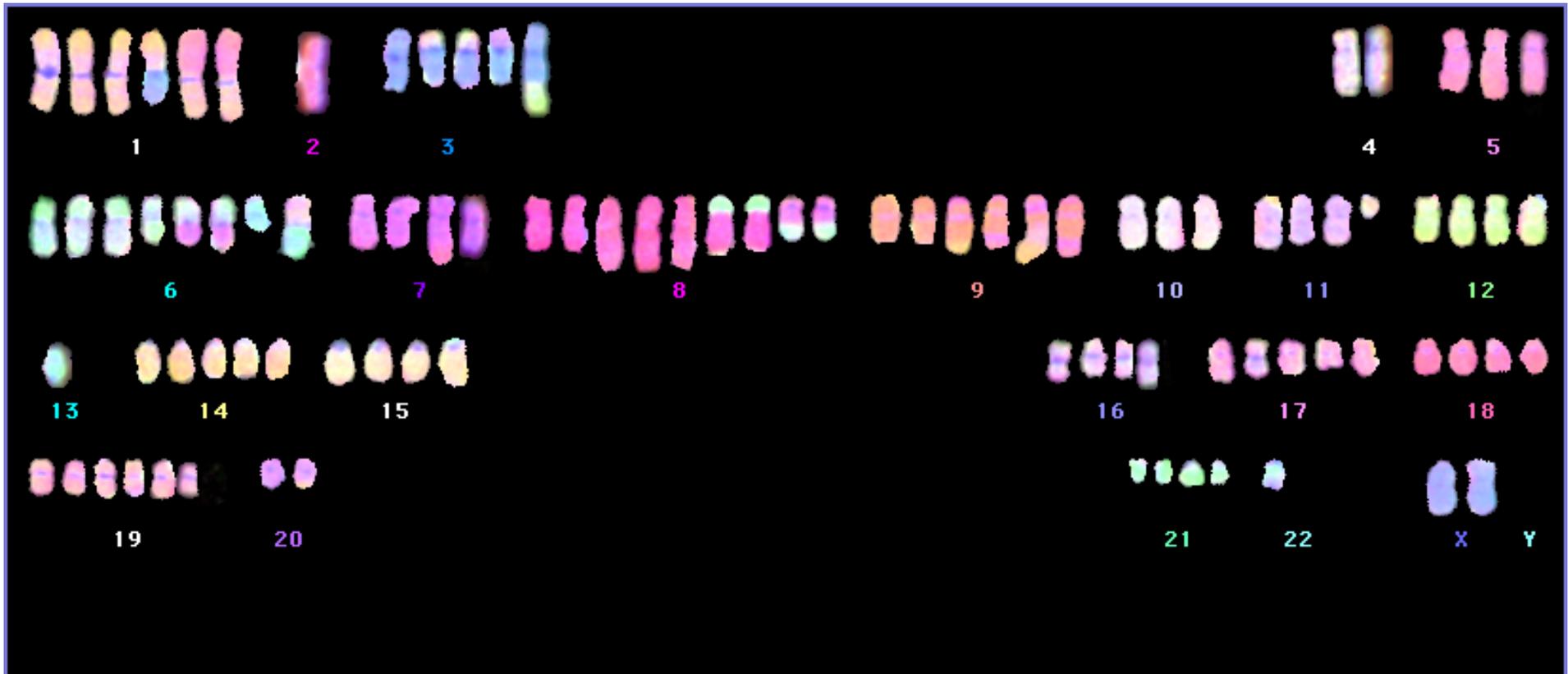


Figure 3.6 M-FISH karyotype of HTB186 (Daoy).

The cell is hypertetraploid, having a modal range of 93-99 chromosomes with two normal X no Y and multiple rearrangements.

M-FISH revealed the following karyotype: 93-99 X, -(X)(q?), -Y, Y, +(1)(p?), +(1)(q?), t(1;13), +(2)(p?), +(2)(q?), del(3)(p?), +(3)(q?), -4, -4, +der(?)t(4:?) (p:?)x2, +(5)(p?), +(5)(q?), +(6)(p?), +(6)(q?), +7, del(8)(pter>q21.1), del(8)(q24.3), del(9)(p?), del(9)(q?), del(10)(p-q), -13, -13, +14, del(14)(p11.1), del(14)(p13), +(15)(q12-13), del(15)(p11.2-13), +(16)(p?), del(16)(q?), +(17)(p?), +(17)(q?), del(18)(p?), +19, +20, +(21)(q), +(22)(q?). 4n

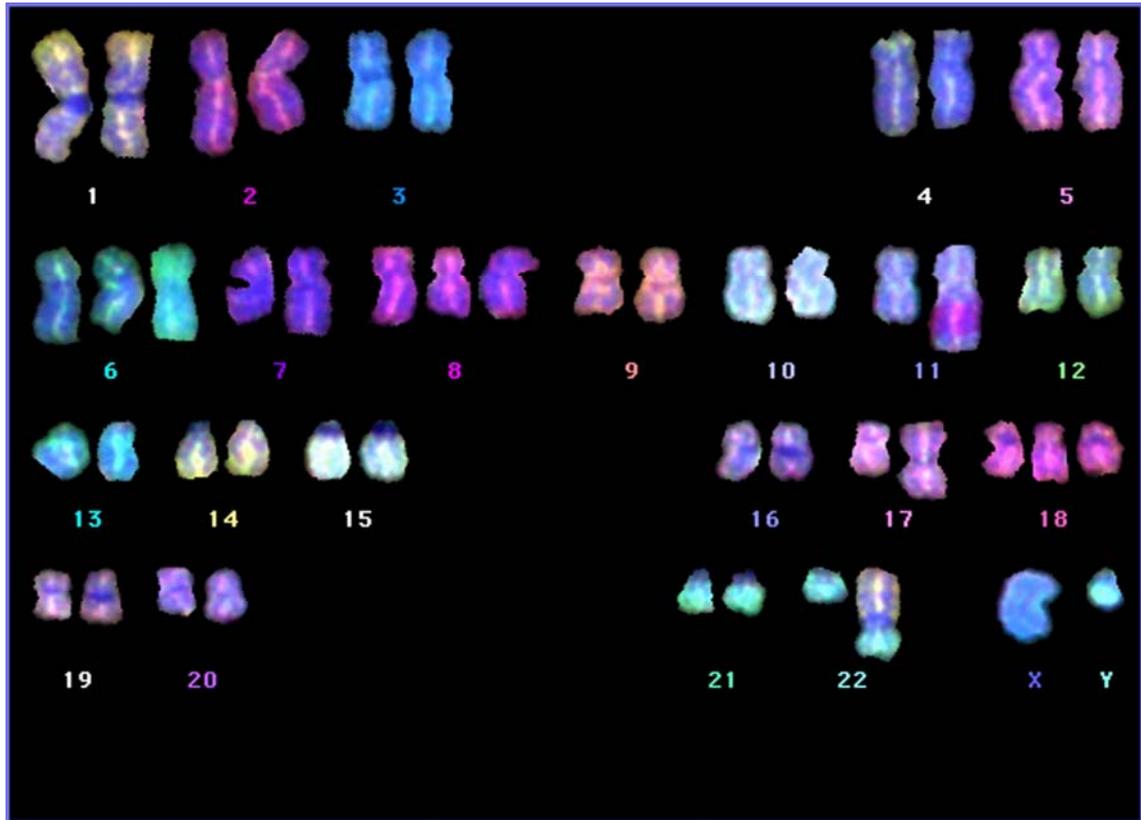


Figure 3.7 M-FISH karyotype of cell line HTB187.

The cell is hyperdiploid, having a modal range of 44-50 chromosomes with 1 copy of X and Y. M-FISH analysis showed a male karyotype with the following aberrations:

49, XY, +6, +8, -11, 11(HSR), -17, i(17q), +18, -22, der(22)t(1;22)(q12;p12).

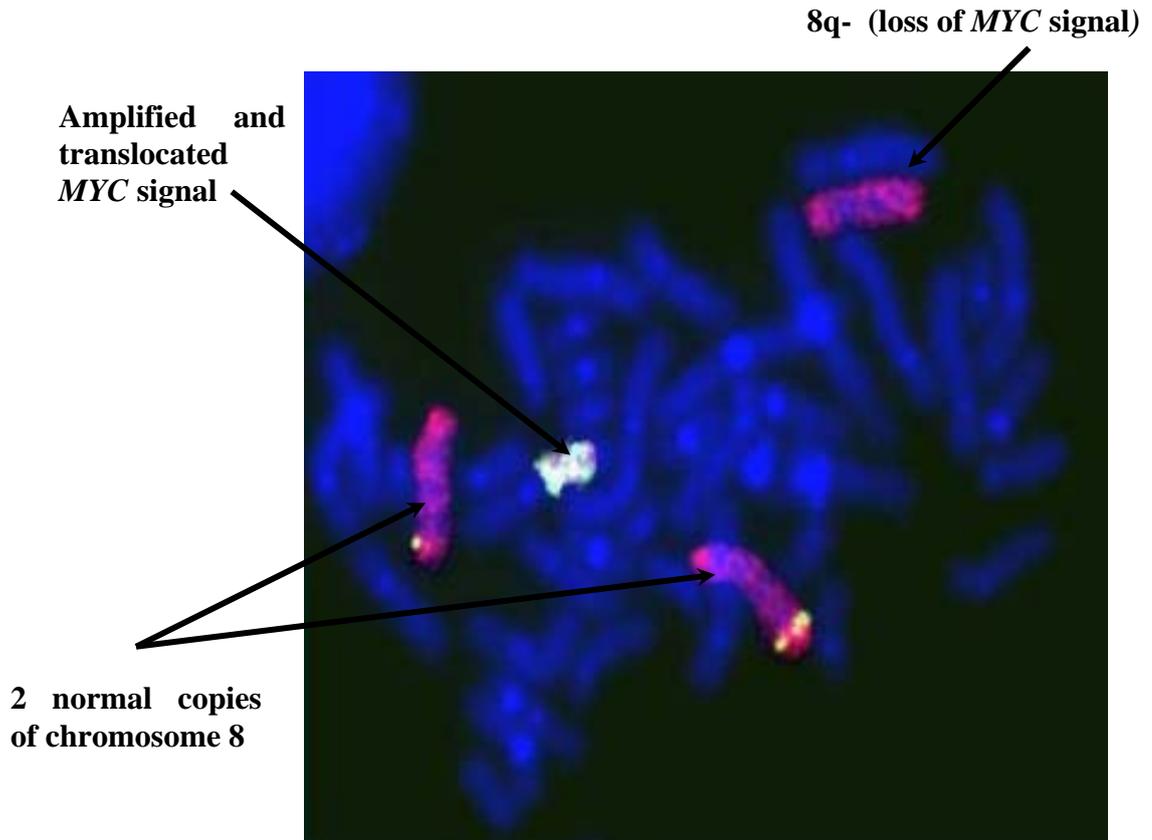


Figure 3.8 FISH analysis of HTB187 using chromosome 8 paint (red) and *MYC* probe (green).

Two normal copies of chromosome 8 and one copy of 8q- (with loss of the *MYC* signal) are visualised. Amplified *MYC* is translocated to chromosome 11 (see figure below).

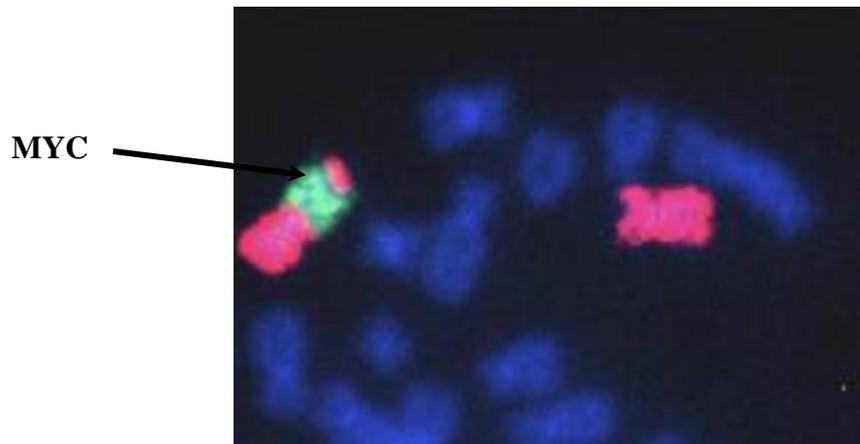


Figure 3.9 FISH analysis of HTB187 using chromosome 11 paint (red) and *MYC* probe (green).

An HSR containing amplified *MYC* (arrowed) is present on chromosome 11.

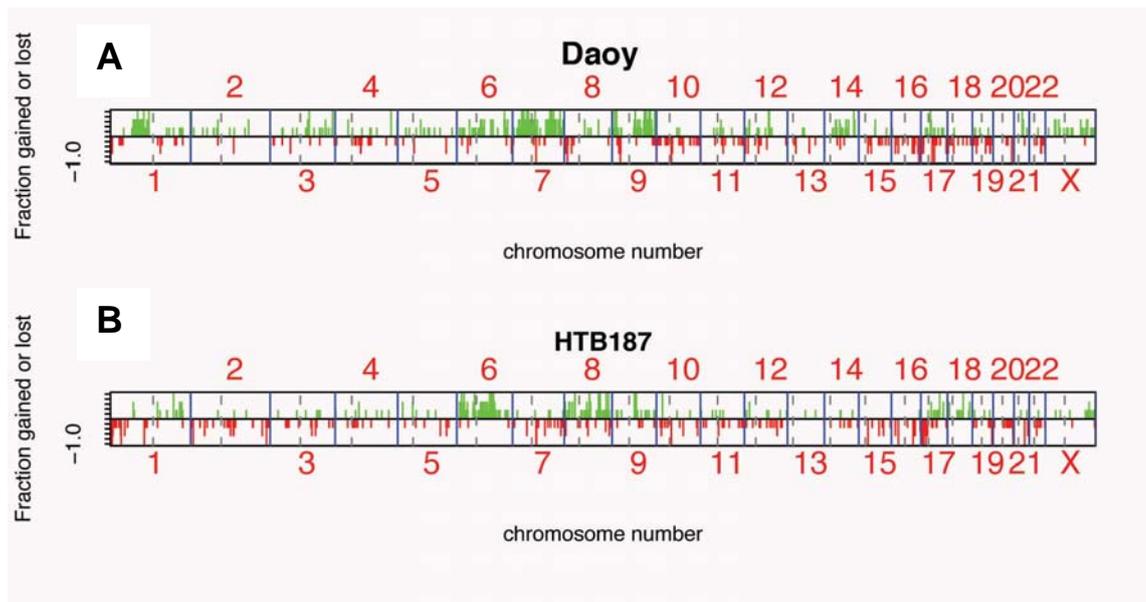


Figure 3.10 Array CGH whole genome plots for medulloblastoma cell lines Daoy and HTB187, obtained using Bioconductor

A) Daoy (*test*) versus single male (*control*). Changes identified in Daoy included gained copies of 1p32-36, 9q22-34, whole gain of 7 and 14. Chromosomal losses including 8p11-23 and loss of chromosome 16 were identified

B) HTB187 (*test*) versus single male (*control*), showing additional copies of chromosomes 6, 8 and 18, gain of 1p and 17q and loss of 17p.

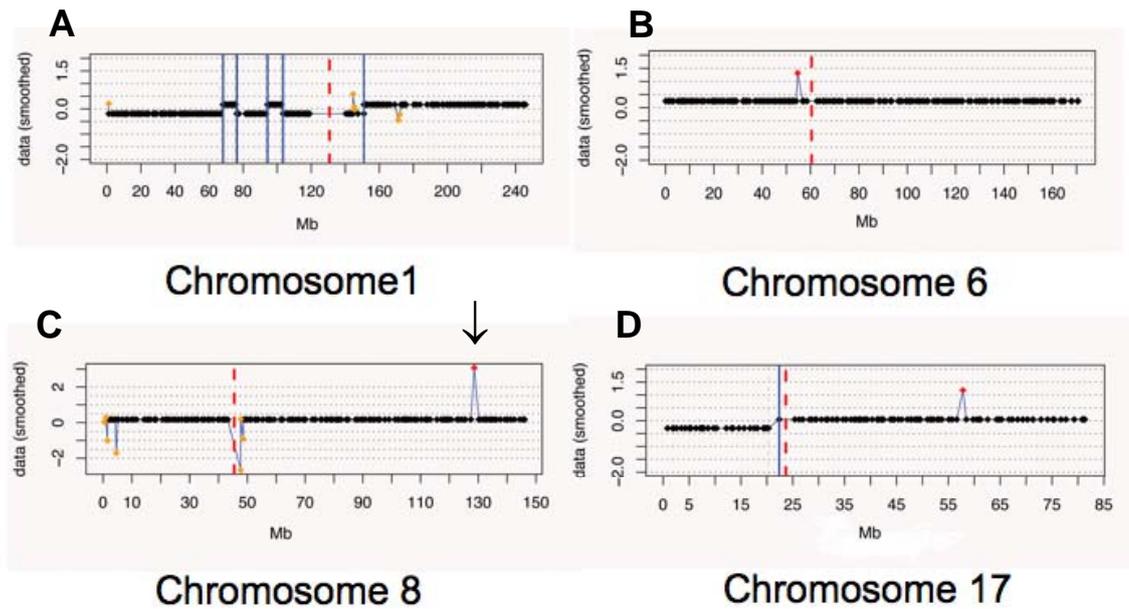


Figure 3.11 Array CGH plots from individual chromosomes from HTB187 (*test*) compared to male (*control*), obtained using Bioconductor. Data were smoothed following segmentation.

- A)** Gain of distal chromosome 1q from 150-245Mb, with additional regions of gain on 1p at ~70 and ~100Mb
- B)** Trisomy of chromosome 6. The baseline level is consistently above 0, indicating gain of a whole copy of the chromosome
- C)** Amplification of *MYC* on chromosome 8 (arrowed)
- D)** Loss of chromosome arm 17p and gain of 17q.

3.5.5 Paediatric low-grade astrocytoma versus pooled male control

Array CGH was performed using DNA from 20 paediatric low-grade astrocytoma samples. These were provided by Professor David Ellison, as described in chapter 2. Information on tumour site alone was available at the time of sample selection. Eleven cerebellar and 9 non-cerebellar samples were selected for array CGH according to the quality and amount of DNA available. I was blind to all other clinico-pathological details, including age, sex and diagnosis at the time the experiments were performed. The cohort contained 10 pilocytic astrocytomas (PA) from the cerebellum and 1 PA from the brainstem, 5 diffuse astrocytomas (DA), 3 pilomyxoid gliomas (PMG) and 1 pilomyxoid astrocytoma (PMA), all from non-cerebellar sites, respectively.

Each tumour sample was co-hybridised with pooled normal male control DNA, obtained from 21 co-workers as described in chapter 2. Examples of array CGH plots obtained for individual samples and for pooled sample groups (cerebellar and non-cerebellar) are shown in figures 3.12-3.20.

3.5.6 Results of array CGH analysis for paediatric low-grade astrocytomas

Sex chromosome gains and losses from tumour samples compared with male control DNA are shown in table 3.1. Male patients would be expected to display no net gain of X or Y chromosome material when compared to male control DNA. Female patients should display gain of X and loss of Y material when compared to male control DNA. All 20 patients were correctly identified as male or female, with complete agreement for the sex chromosome changes identified by array CGH.

Gains and losses, both of discrete genomic regions and whole chromosomes were identified in the tumour samples. These results, and details of the clones flanking the regions of gain or loss are displayed in Table 3.2.

Three patients had apparently normal karyotypes (PA3, PA15, PA17). A further 6 patients had apparently normal karyotypes with 1-2 discrete regions of chromosome

gain or loss (PA13, PA18, PA23, PMA1, DA7, DA9). Whole chromosome gains were identified in 6/20 tumours (range 1-4 whole chromosomes gained). Whole chromosome gains were seen for chromosome 6 (gained in 4 samples, respectively), chromosome 5 (gained in 2 samples) and chromosomes 7, 8, 9, 11, 12, 13 and 21 (gained in 1 sample each) (Figures 3.12-3.17). The mean age of patients with whole chromosome gains was 14.6 years, compared to a mean age of 6.5 years for patients without chromosome gains.

Whole chromosome losses were identified in 5/20 patients (range 1-4 whole chromosomes lost). Whole chromosome losses were seen for chromosomes 17 and 22 (lost in 3 samples each), and chromosomes 19 and 20 (each lost in a single sample).

Smaller regions, involving copy number changes of chromosomal arms, and discrete regions of chromosome gains or losses were found in 11 patients. The majority of these changes were identified in single samples. However, three regions of chromosomal gain or loss were identified involving 2 or more samples, and are described in detail below.

3.5.7 Loss of the terminal end of chromosome 1p

Loss of the terminal end of chromosome 1p was identified in four samples (PA32, PMG2, DA2, DA4 (Figure 3.18 (A))). In two samples, PA32 and DA2, the region of 1p loss ended at clone RP5-1043G4, at 55Mb, and included over 62 BAC/PAC clones. This region includes the genes *C1orf175*, *TTC4*, *PARS2* and *TTC22*. The region of 1p loss differed in the remaining two samples: DA4 showed 1p loss over 53 clones, ending at clone RP11-8J9 (45Mb), which contains the genes *ATPAF1*, *KIAA0494* and a region of known copy number variation (CNV). PMG2 showed 1p loss ending at clone RP11-117D22 (52Mb), and including over 60 clones containing the genes *ZYG11B*, *ECHD2* and a region of known CNV.

3.5.8 Gain at 2q11.1

Two samples (DA7, DA9) revealed identical regions of copy number gain of 2q11.1 at 102Mb (Figure 3.18(B)). The gained region covered 1.8Mb, starting at clone RP11-451C2, ending at clone RP11-451G1, and flanked by RP11-299H21 and RP11-350B7.

No genes are contained within the start and end clones, which both contain regions of known CNV, however the region gained contains the genes *SLC9A2* and *TMEM182*.

3.5.9 Loss at 7q22.1

Two samples (PA13, DA4) revealed loss of a discrete region at 99Mb on chromosome 7q22.1 (Figure 3.18 (C)), with a small region of loss common to both. Sample PA13 showed loss covering 1.5Mb at 7q22.1, starting at clone RP11-506M12 (which contains, among others, the genes *MCM7* and *TAF6*) and ending at clone RP5-1059M17, containing *EMID2*. Both of these clones are known to contain regions of CNV. Sample DA4 showed a smaller region of loss of 800Kb at 7q22.1. Here, the region of loss began at clone RP4-550A13, containing the gene *SMURF1*, and ended at clone RP11-506M12 (the starting point for the loss in PA13).

3.5.10 Comparison of cerebellar samples to non-cerebellar samples

Further analysis was performed by separating the tumour samples into two groups, of cerebellar and non-cerebellar samples. The cerebellar group was entirely composed of pilocytic astrocytomas. The non-cerebellar group contained 5 diffuse astrocytomas (DA), 3 pilomyxoid gliomas (PMG), 1 pilomyxoid astrocytoma (PMA), and the single pilocytic astrocytoma from the brainstem.

The mean intensity value for each clone was calculated between all patients in the group (using the per-tumour M values). The grouped data underwent further segmentation before subtracting the non-cerebellar group data from the cerebellar group data to create a delta-cerebellar group. This was to identify any significant DNA copy number differences present in the cerebellar samples, but not in the non-cerebellar group.

No regions of copy number change were identified in the cerebellar group, compared to the non-cerebellar group (Figure 3.19). A single region of gain was identified in the non-cerebellar samples at 21q22.2 (~42Mb) when the two groups were compared. This group was, however, formed of a number of different tumour types. Individual data plots for chromosome 21 are shown (Figure 3.20 (A), (B) and (C)), from cerebellar and non-cerebellar groups individually and when compared with each other. The region of

21q22.2 gain involved 3 clones and spanned a region of ~2.5Mb. The three clones in the gained region were; RP1-128M19 at 40.7Mb, containing *BRWDA*, *HMGNI* and *WRB*, RP1-171F15 at 41.7Mb, containing *DSCAM* and RP1-265B9 at 42.6-42.7Mb, containing *BACE2*, *FAM3B* and *MX2*. Further investigation of the raw data showed the regions of gain occurred most frequently in diffuse astrocytomas. RP1-128MG was gained in 4 samples (DA3, DA7 and DA9), RP1-171F15 was gained in 4 samples (PMA1, DA2, DA4) and RP1-265B9 was gained in 2 samples (DA3 and DA9) respectively. However, none of these regions were identified as significant regions of gain when examining each sample individually, or within the non-cerebellar group alone.

Sample	Pathology, site	Age	Sex	X chromosome	Y chromosome	Agrees with patient sex
PA1	Cerebellum	20	F	Gain	Loss	Yes
PA3	Cerebellum	7	M	No gain	No gain	Yes
PA11	Cerebellum	14	M	No gain	No gain	Yes
PA13	Cerebellum	10	M	No gain	No gain	Yes
PA14	Cerebellum	10	M	No gain	No gain	Yes
PA15	Cerebellum	13	M	No gain	No gain	Yes
PA17	Cerebellum	6	M	No gain	No gain	Yes
PA18	Cerebellum	3	M	No gain	No gain	Yes
PA20	Cerebellum	4	M	No gain	No gain	Yes
PA23	Cerebellum	3	F	Gain	Loss	Yes
PA32	Brain stem	9	F	Gain	Loss	Yes
PMA1	Diencephalon	1	F	Gain	Loss	Yes
DA2	Cerebrum	9	M	No gain	No gain	Yes
DA3	Cerebrum	17	M	No gain	No gain	Yes
DA4	Cerebrum	4	M	No gain	No gain	Yes
DA7	Cerebrum	5	M	No gain	No gain	Yes
DA9	Cerebrum	9	F	Gain	Loss	Yes
PMG1	Diencephalon	17	M	No gain	No gain	Yes
PMG2	Diencephalon	7	F	Gain	Loss	Yes
PMG5	Spinal cord	11	F	Gain	Loss	Yes

Table 3.1 Sex chromosome gains and losses determined by array CGH from tumour samples when compared with pooled male control DNA.

Male patients should display no net gain of X or Y, and female patients should display gain of X and loss of Y. There was 100% concordance between patient sex and the gains and losses in the sex chromosomes identified by array CGH. PA (pilocytic astrocytoma), DA (diffuse astrocytoma), PMG (pilomyxoid glioma), PMA (pilomyxoid astrocytoma).

Sample	Site	Age	Sex	Chromosome gains	Chromosome losses	Discrete changes	Position (Mb)	Clones at start and end of region	Size of change	Clones flanking region
PA1	Cerebellum	20	F	Gain 5, 6	None					
PA3	Cerebellum	7	M	None	None					
PA11	Cerebellum	14	M	Gain 5, 6	None					
PA13	Cerebellum	10	M	None	None	7q22.1 loss	99Mb	RP11-506M12*, RP5-1059M17*	1.5Mb	RP4-604G5, RP11-333G13
PA14	Cerebellum	10	M	None	Loss 17					
PA15	Cerebellum	13	M	None	None					
PA17	Cerebellum	6	M	None	None					
PA18	Cerebellum	3	M	None	None	2p22.1 loss	39Mb	RP11-173C1, RP11-395B14*	500Kb	RP11-278G12, RP11-457F14
PA20	Cerebellum	4	M	None	Loss 22					
PA23	Cerebellum	3	F	None	None	1q42.12 loss	220Mb	RP11-100E13*, RP11-375H24	2.4Mb	RP11-105I12, RP11-276J4
						4p15.31 gain	20Mb	RP11-362J17, RP11-10G12*	1.8Mb	PAC963K6, RP11-406E22
PA32	Brainstem	9	F	Gain 13	Loss 22	Loss terminal end 1p	tel - 55Mb	PAC14E10*, RP5-1043G4	55Mb (62 clones)	RP5-1070D5
PMA1	Diencephalon	1	F	None	None	5q33.3 gain 8q11.23 gain	157Mb 53Mb	RP11-391B7*, CTC-279E3 RP11-197I11, RP11-182E14*	1.8Mb 860Kb	RP11-102A1, RP11-505G12 RP11-11C20, RP11-53M11

Table 3.2 Summary of chromosome gains, losses and point changes, excluding sex chromosomes, found in 20 tumour samples by array CGH.

Asterisks denote clones containing known copy number variations or copy number polymorphisms from UCSC Genome Browser (March 2006 assembly).

Sample	Site	Age	Sex	Chromosome gains	Chromosome losses	Discrete changes	Position (Mb)	Clones at start and end of region	Size of change	Clones flanking region
DA2	Cerebrum	9	M	None	Loss 17	Loss terminal end 1p	tel-55Mb	PAC14E10*, RP5-1043G4	55Mb (62 clones)	RP5-1070D5
DA3	Cerebrum	17	M	Gain 6, 7, 21	None	5q32 gain	149Mb	CTB-1013H5, CTB-137O22*	500Kb	RP5-1070D5
DA4	Cerebrum	4	M	None	Loss 17, 19, 20, 22	Loss terminal end 1p 7 cen-7q11.23 loss 7q22.1 loss 7q36.1-tel loss	tel-46Mb cen-74Mb 99Mb 148Mb-tel	PAC14E10*, RP11-8J9* RP4-725G10*, RP4-771P4 RP4-550A13, RP11-506M12* RP5-1136G13*, PAC3K23	45Mb (53 clones) 17Mb (17 clones) 800Kb 10Mb (17 clones)	RP11-330M19 RP11-339F13, RP4-799O8 RP11-380G21, RP11-44M6 RP4-811H12
DA7	Cerebrum	5	M	None	None	2q11.2 gain 17p12 gain	102Mb 13Mb	RP11-451C2*, RP11-451G1* RP11-388F14*, RP1-27J12	1.8Mb 1.1Mb	RP11-299H21, RP11-350B7 RP11-488L1, RP11-385D13
DA9	Cerebrum	9	F	None	None	2q11.2 gain	102Mb	RP11-451C2*, RP11-451G1*	1.8Mb	RP11-299H21, RP11-350B7
PMG1	Diencephalon	17	M	Gain 6, 8, 11, 12	None					
PMG2	Diencephalon	7	F	None	None	Loss terminal end 1p	tel- 53Mb	PAC14E10*, RP11-117D22*	52Mb (60 clones)	RP11-243A18
PMG5	Spinal cord	11	F	Gain 9	None					

Table 3.2 Continued.

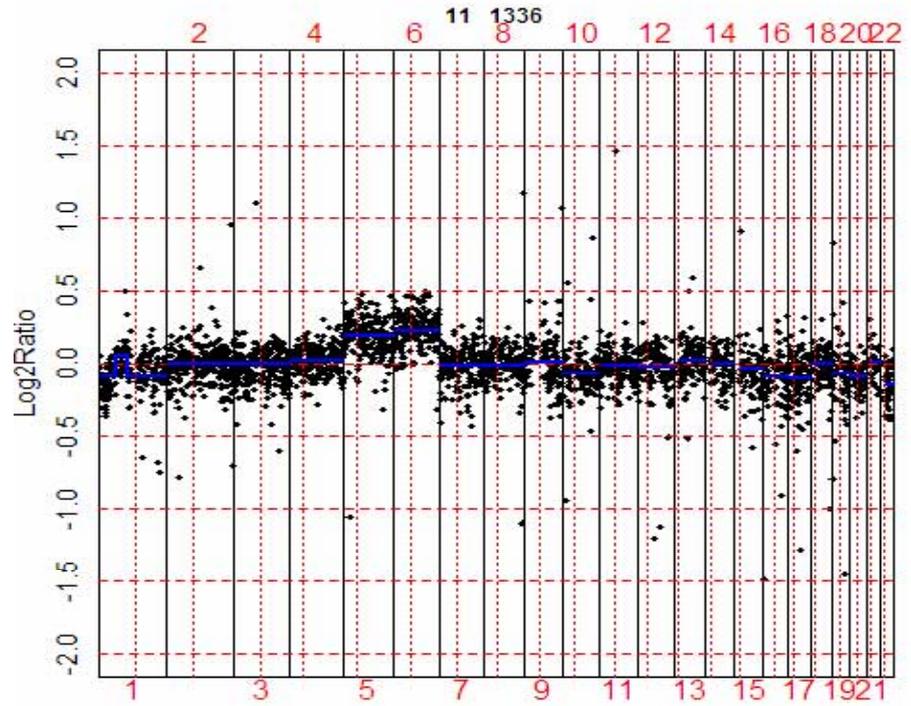


Figure 3.12 Array CGH whole genome plot for tumour sample PA11, showing gains of chromosomes 5 and 6.

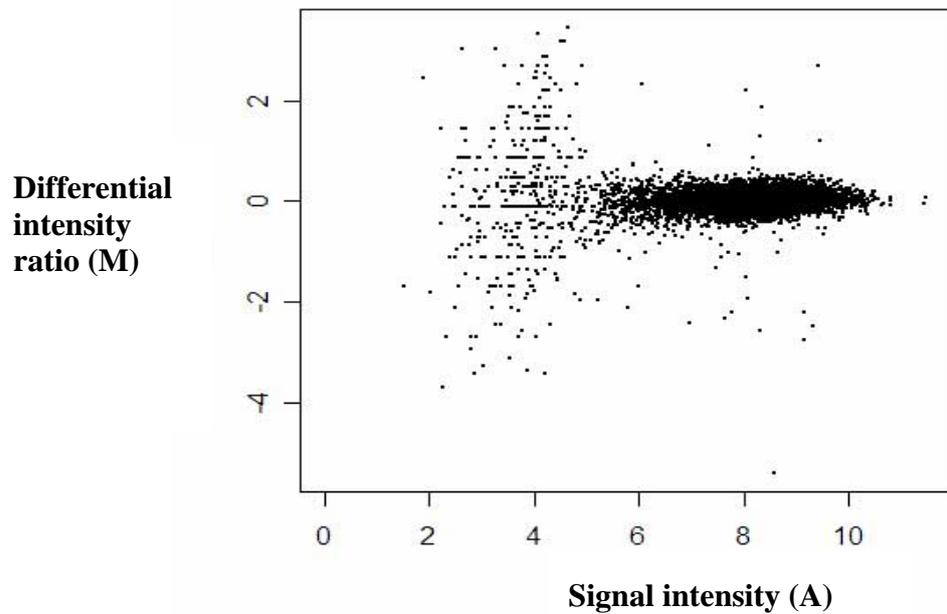


Figure 3.13 MA plot using raw array CGH data for sample PA11.

Equal hybridisation between fluorochromes is shown for the majority of data points, along the zero axis. At the lowest signal intensity values the feature values become scattered, as they become similar to the background signal intensity.

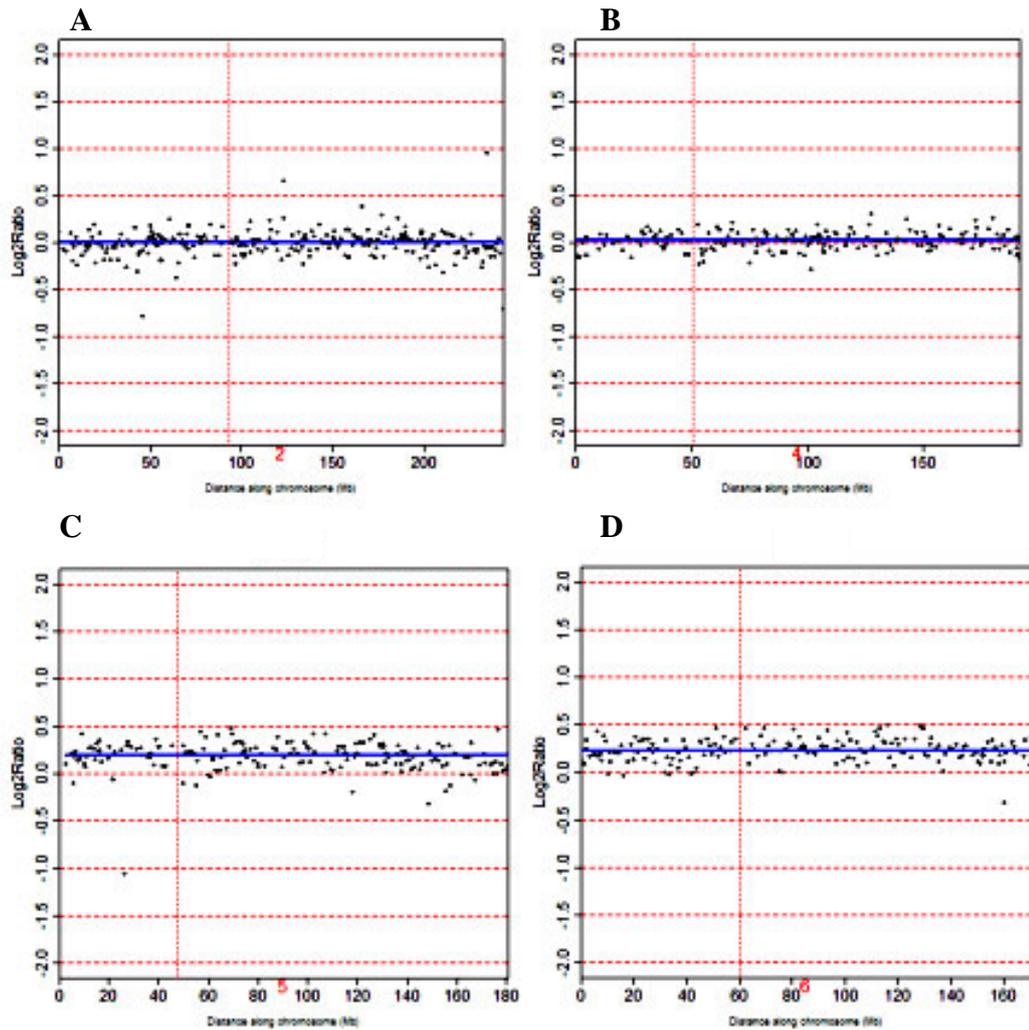


Figure 3.14 Array CGH data plots for selected chromosomes from tumour sample PA11.

- A) Chromosome 2, no net gain or loss of DNA**
- B) Chromosome 4, no net gain or loss of DNA**
- C) Chromosome 5, whole chromosome copy number gain**
- D) Chromosome 6, whole chromosome copy number gain**

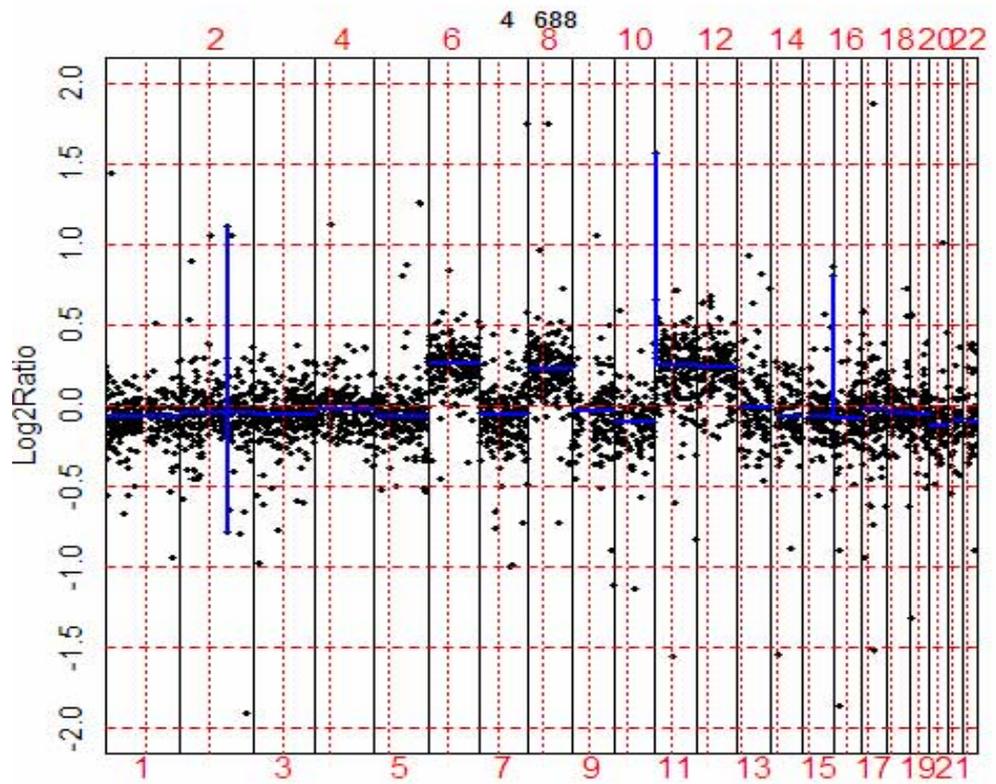


Figure 3.15 Array CGH whole genome plot for tumour sample PMG1, showing gain of chromosomes 6, 8, 11 and 12.

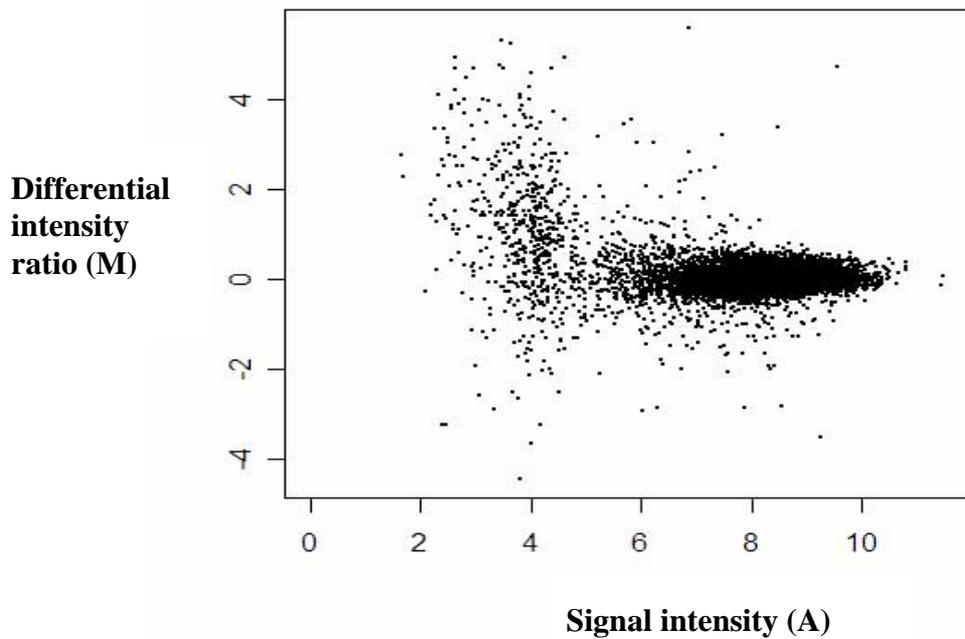


Figure 3.16 MA plot using raw array CGH data for sample PMG1.

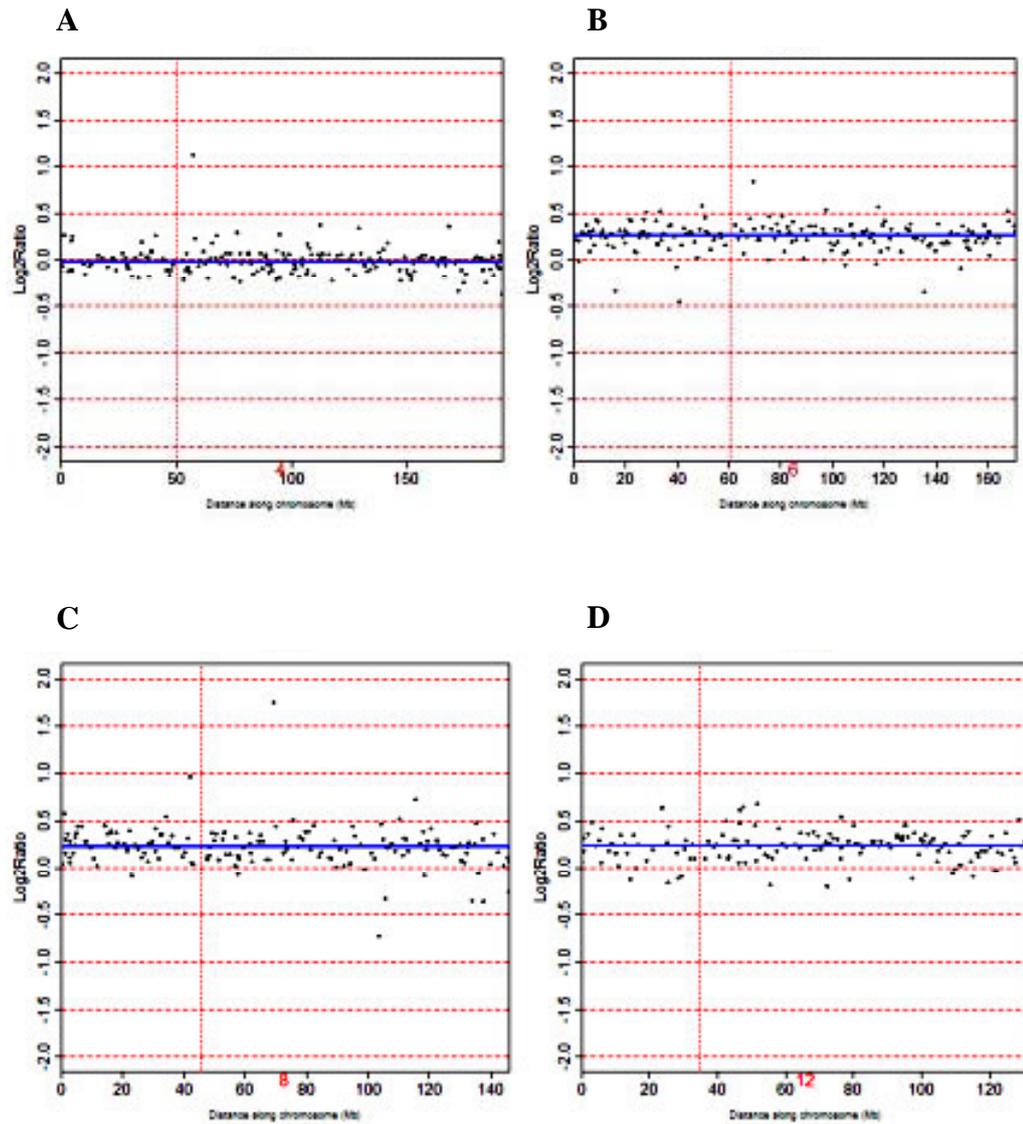


Figure 3.17 Array CGH chromosome plots from tumour sample PMG1.

- A) Chromosome 4, no net gain or loss of DNA**
- B) Chromosome 6, whole chromosome copy number gain**
- C) Chromosome 8, whole chromosome copy number gain**
- D) Chromosome 12, whole chromosome copy number gain**

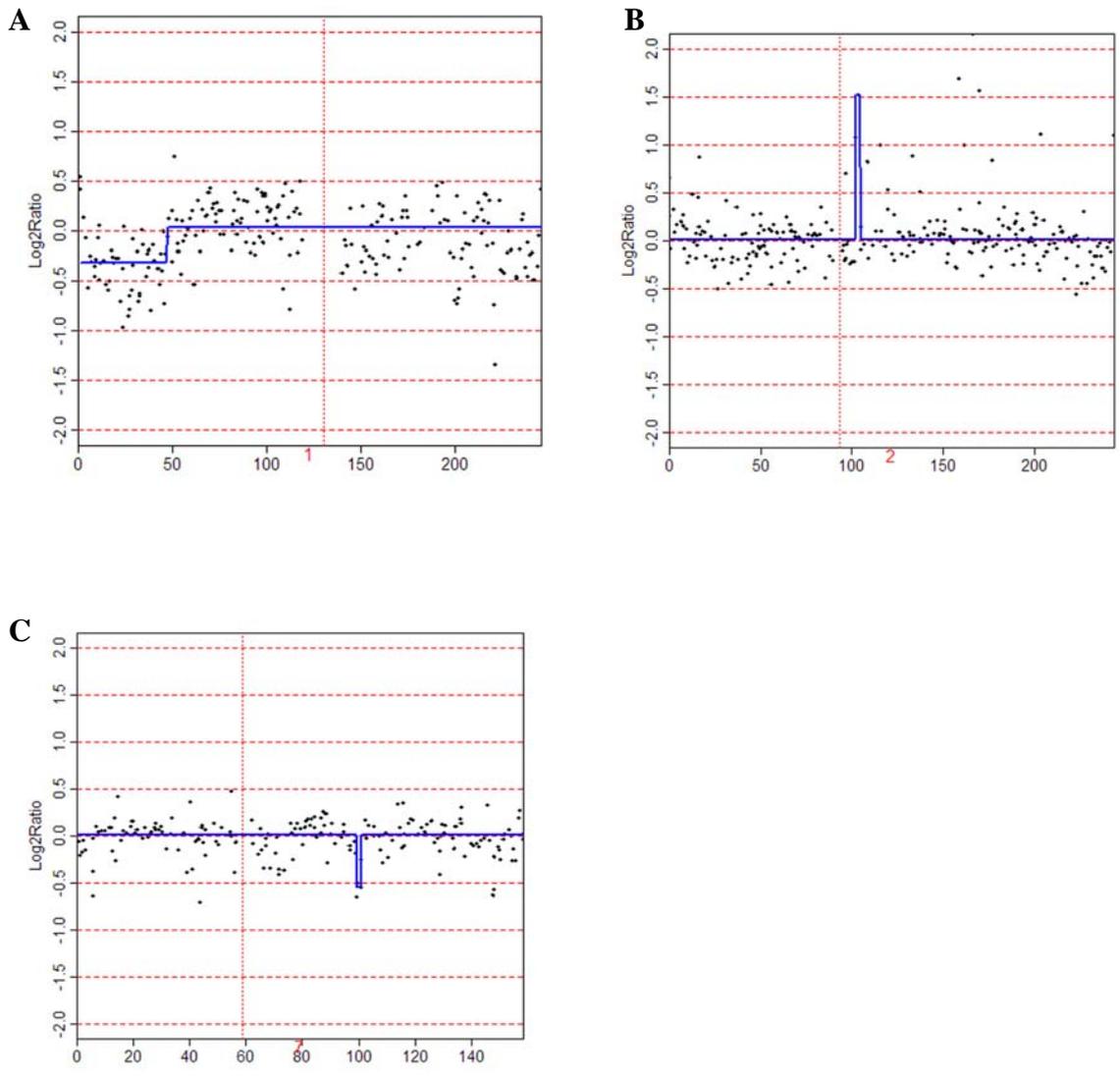


Figure 3.18 Array CGH data plots for discrete regions of gain and loss.

- A) Loss of chromosome 1p in sample DA4.**
- B) Gain in chromosome band 2q11.2 (102Mb) in sample DA7.**
- C) Loss in chromosome band 7q22.1 (99Mb) in sample PA13.**

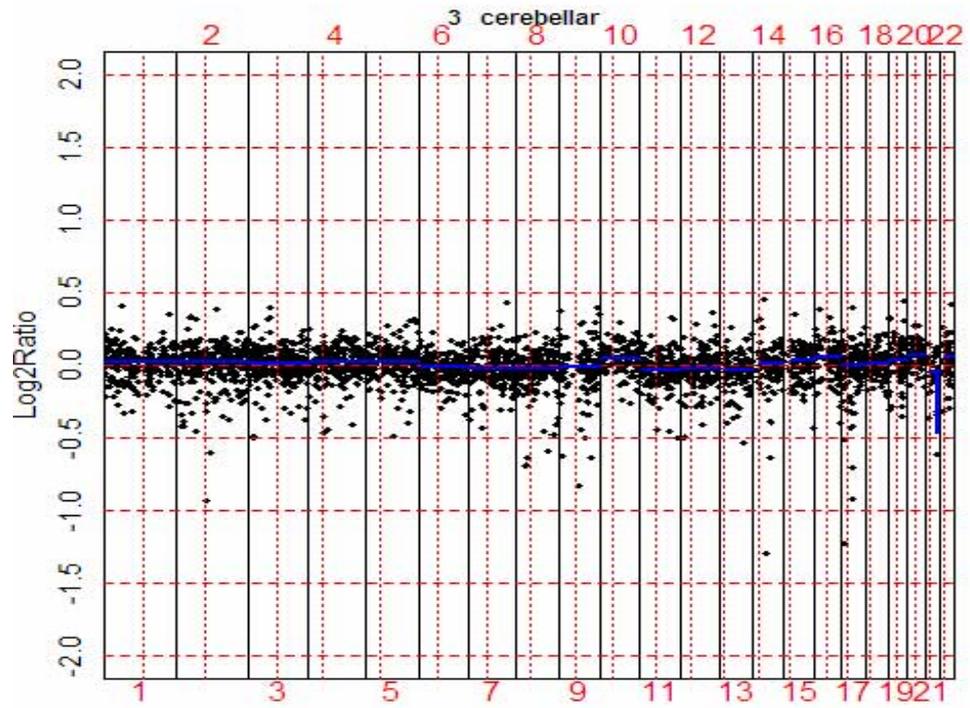


Figure 3.19 Whole genome array CGH plot comparing grouped data from cerebellar tumours with non-cerebellar tumours.

No significant regions of gain or loss were identified within the cerebellar tumour group.

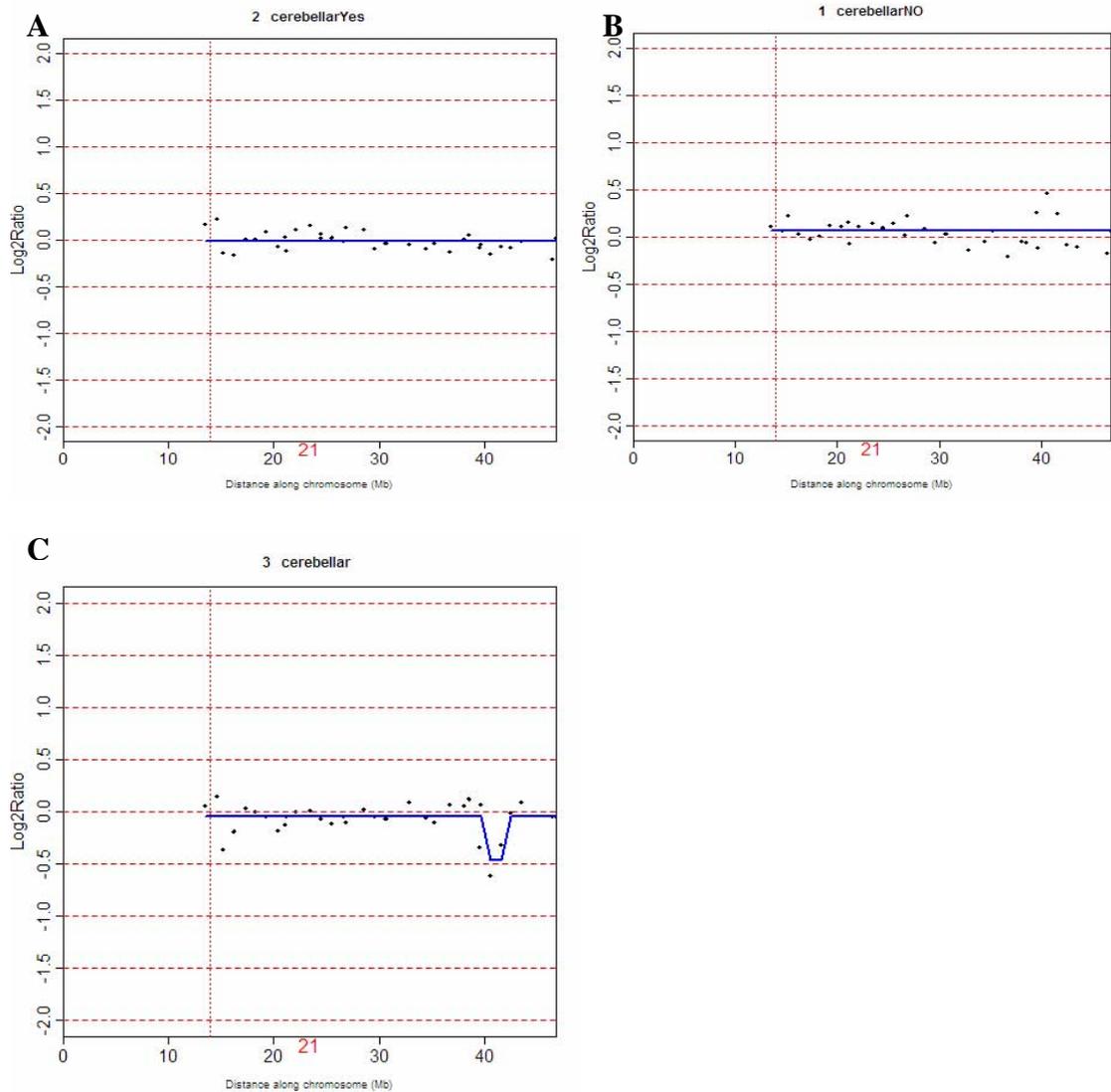


Figure 3.20 Array CGH data plots for chromosome 21.

A) Array CGH data plot for chromosome 21 using pooled data from all cerebellar samples, showing no net gain or loss of DNA.

B) Array CGH plot for chromosome 21 using pooled data from all non-cerebellar samples. This indicates a region of gain at ~42Mb which does not achieve statistical significance.

C) Array CGH plot for chromosome 21, subtracting non-cerebellar data from cerebellar data to identify differences between the two groups. A significant region of copy number loss is identified at ~40Mb, indicating a region of gain in the non-cerebellar group, when compared to the cerebellar group.

3.6 Discussion

In this chapter, gains and losses of whole chromosomes and discrete genomic regions were identified by array CGH, in paediatric low-grade astrocytoma samples and in the medulloblastoma cell lines Daoy and HTB187.

Array CGH has enabled detailed analysis of copy number changes to be performed in cancer. There are, however, limitations to the technique. Gains or losses of genetic material are identified relative to the whole genome. Balanced translocations are not identified by the technique, as there is no net change of genetic dosage (Gaasenbeek *et al.*, 2006, Lichter *et al.*, 2000). Aneuploidy (e.g. monosomy or trisomy) is identified by array CGH, this is common in epithelial cancers (Storchova *et al.*, 2004). However the data represent an average of genetic imbalances, which is entirely dependent on the input DNA. Data plots for array CGH from cell lines contain a pure or almost pure population of cells containing the same genetic aberration. DNA extracted from tumour samples may contain a heterogeneous population of cells, a proportion of which may contain germline karyotypes, while a further proportion have undergone malignant transformation. Using a \log_2 scale, a single DNA copy gain in a tumour sample when compared with a control would be expected to be + 0.58 above the x-axis, and a gain of 2 copies +1. In general, array CGH plots from tumour samples rarely give such clear ratios due to contamination by non-malignant cells. This underlies one of the main difficulties with array CGH – where to set the thresholds for gain and loss. There has not been general consensus of what constitutes gain or loss within the literature. For the purposes of this thesis, I designated a single copy gain to be ≥ 0.3 above the x-axis, and a single copy loss to be ≤ -0.5 below the Y-axis, to reflect the probable heterogeneity of the tumour tissue studied.

Three types of array CGH platform are available. Arrays comprised of BAC/PAC genomic DNA clones or cDNA clones may be used for genome-wide analysis, synthetic oligonucleotide arrays may be used to study particular regions of interest at higher resolution. More recently genome-wide oligonucleotide arrays have become available, but these had not been developed at the time this study began (Ylstra *et al.*, 2006).

The array platform chosen for a series of experiments should be made after considering the merits and the disadvantages of each system. Genomic BAC or PAC clones may exhibit significant homology between paralogues or related gene family members at different chromosomal locations, which affects binding specificity for individual clones. Ideally, all genomic clones selected for use on array slides should undergo mapping by fluorescence in situ hybridisation (FISH) to confirm their genomic location (Fiegler *et al.*, 2003). Unfortunately this is time-consuming and costly, so is usually not possible for each and every clone used to construct an array. Arrays produced using cDNA clones provide gene-specific data, and are not subject to the clone mapping errors that may occur in genomic arrays. However, many genes exhibit alternative splicing, which will produce differences in sequence not covered by individual cDNA clones. Hence, genomic BAC DNA arrays may be preferred to probe gene-poor regions that do not have representative cDNA clones.

The production of the numerous BAC clones required for an array is hugely expensive and time-consuming. Individual BAC clones are produced in large-scale bacterial cultures, often giving low yields of DNA per clone (Albertson *et al.*, 2003, Fiegler *et al.*, 2003). BAC DNA is generally viscous in solution as large clone inserts have a high molecular weight, which makes printing array slides technically difficult (Albertson *et al.*, 2003). Early array CGH production methods occasionally used sonication of DNA to reduce fragment sizes to overcome this (Cai *et al.*, 2002). Currently, clones can be produced by PCR amplification from smaller-scale single-colony DNA clone preparations using a variety of PCR amplification methods (Snijders *et al.*, 2001, Veltman *et al.*, 2002, Fiegler *et al.*, 2003). Amplification by modified degenerate-oligonucleotide primed PCR (DOP-PCR), using three different primers selected to amplify human DNA sequences in preference to DNA from *E.Coli*, gave the most reproducible and complete results (Fiegler *et al.*, 2003). This was due to reduced contamination of clones by *E.Coli* DNA, and to improved primer efficiency for human DNA amplification.

For this study, I selected genomic DNA array slides, as they provided more comprehensive information on copy number changes across the whole genome, including gene-poor regions.

The resolution of array CGH data is determined not only by the quality of the input DNA, but also by the size and density of the BAC and PAC clones on the array. The methods of printing and drying the slides are very important, to ensure features hybridise evenly with the input DNA. Features should be uniformly round to allow identification by spot-finding software packages such as Genepix or Spotfinder. Deficiencies in printing or drying the slides can compromise array slide quality, and hence the sensitivity and specificity of the results obtained.

Sensitivity (i.e. the number of true positive spots identified) was low in the array slides used to prepare data for this thesis. I found that there were high numbers of features flagged as 'bad' when scanning, due to deficiencies with printing. Significant numbers of features were rejected by the Genepix software, due to their appearances on scanning. Many spots appeared doughnut or rugby-ball shaped, giving uneven hybridisation of fluorochrome, and hence not meeting the inclusion criteria for a 'good' feature. The number of features meeting inclusion criteria for array CGH experiments and replicates performed for each sample are shown as percentages in Table 3.3. The range of features included in the analyses varied from 23-77% between samples. Twelve out of 20 samples had 23-49% 'good' features available for analysis. However, samples where <40% features were used for analysis were still able to assign patient sex correctly by gain, loss or no change in the sex chromosomes. For example, DA3 only included 22 features from the Y chromosome, but showed that there was no net gain or loss of DNA when compared to the male control DNA, correctly indicating that the patient was male. These data indicate that whole chromosome changes were identified correctly even where a low proportion of features were included. However, identification of discrete regions of gain or loss would obviously be compromised by the lack of data points across the chromosomes caused by the low numbers of 'good' features.

High background signal was also a problem when scanning the slides, despite numerous changes to the washing protocol during optimisation of the technique. Hence specificity (proportion of true negative spots identified) was low. This appeared to be due to non-specific hybridisation of test and control DNA to the slide surface, which again appeared to be related to problems with slide production. The MA plots demonstrated equal hybridisation between fluorochromes for the majority of the data. However, at lower signal intensities, the feature values became scattered away from the zero axis, as they became similar to the background signal intensity. Overall, it appeared that the

array slides produced by the ICR did not achieve the quality of arrays previously available to the laboratory from another source, which had been used in our laboratory to investigate adult glioblastomas (Mulholland *et al.*, 2006).

Array CGH has been used to identify regions of DNA copy number gain and loss in adult astrocytomas, particularly in glioblastomas (WHO grade IV). Amplifications of the oncogenes *EGFR* (36%), *PDGFR* (8%), *CDK4* (13%), *MDM2* (9%) and *MDM4* (8%) occur frequently in glioblastomas (Roerig *et al.*, 2005, Ruano *et al.*, 2006, Rao *et al.*, 2009, Korshunov *et al.*, 2006, Misra *et al.*, 2005). These data indicate that the genes involved in gliomagenesis are part of well-known oncogenic pathways, including growth factor signalling via receptor tyrosine kinases and inhibition of cell cycle progression and apoptosis via the *p53* and Rb pathways. Array CGH data have also been used to define further sub-classifications within glioblastomas. Primary adult glioblastomas (GBMs) arise *de novo* with a short clinical history, generally less than 3 months, and comprise over 90% of adult GBMs. Secondary GBMs arise from a pre-existing low-grade astrocytoma over a longer period, often 5-10 years. Secondary GBMs are much less common, constituting 5% of GBMs, and tend to affect younger patients (Maher *et al.*, 2006, Ohgaki *et al.*, 2007). Amplification of *EGFR* is the most common abnormality seen in primary GBMs, being found in 36% compared to ~ 8% in secondary GBMs. Mutations in *PTEN* are also seen more commonly in primary GBMs. Loss of function mutations affecting *p53* are most commonly seen in secondary GBMs. These appear to occur early in their malignant progression, as mutations are most commonly seen in WHO grade II and III astrocytomas (Raffel *et al.*, 1999, Ohgaki *et al.*, 2007). However, both primary and secondary GBMs contain similar levels of some aberrations, including homozygous loss of *CDKN2A*, and amplification of *CDK4* and *MDM2*.

Relatively few studies have investigated adult low-grade astrocytomas using metaphase or array CGH (Schrock *et al.*, 1996, Hirose *et al.*, 2003). Large copy number gains were the most frequent aberrations reported, most commonly involving complete or partial gains of chromosomes 5 and 7, and losses of chromosome 1p. The characteristic changes present in adult GBMs have not been identified in low-grade adult astrocytoma samples.

Sample	Slide 1	Slide 1	Slide 2	Slide 2	Slide 3	Slide 3
	Good features	Plot medians on 45° line	Good features	Plot medians on 45° line	Good features	Plot medians on 45° line
PA1	48%	No	65%	Yes		
PA3	29%	No	16%	Yes		
PA11	77%	Yes				
PA13	55%	Yes				
PA14	18%	Yes	45%	Yes	67%	Yes
PA15	18%	Yes	30%	Yes	72%	Yes
PA17	27%	Yes	41%	No	10%	No
PA18	34%	Yes	37%	No	48%	Yes
PA20	32%	No	20%	No		
PA23	24%	No	77%	Yes		
PA32	19%	Yes	37%	No	48%	Yes
PMA1	48%	Yes				
DA2	56%	Yes				
DA3	15%	No	23%	Yes		
DA4	30%	Yes	31%	Yes		
DA7	38%	Yes				
DA9	18%	Yes	37%	Yes		
PMG1	65%	Yes				
PMG2	25%	No	36%	Yes		
PMG5	23%	No	26%	Yes	17%	No

Table 3.3 Percentage of features meeting inclusion criteria for a ‘good’ feature (>75% spot pixels >2SD above background) for array CGH experiments performed for each sample from Genepix 6.

Paediatric astrocytomas have also been investigated using metaphase and array CGH. Studies investigating paediatric high-grade astrocytomas have identified losses of chromosome 10q including the *PTEN* locus, similar to those seen in adults, but amplification of *EGFR* (7p12) and *CDK4* (12q), or chromosome 10p loss have not been frequent findings. Warr *et al.* studied 13 paediatric high-astrocytomas (grades III and IV), and found loss of the *PTEN* locus concurrent with complete loss of chromosome 10 in 1 case, but did not find amplification of *EGFR*, or loss of *CDKN2A/B* or *RB* in any samples (Warr *et al.*, 2001). Similarly, Rickert *et al.* found a much lower frequency of chromosome 7 gains in paediatric high-grade astrocytomas (8%), when compared to those seen in adults (32%) (Rickert *et al.*, 2001). No paediatric cases were shown to contain loss of 10p, compared with 51% in adult GBMs. In this study: 23% paediatric GBMs showed gains of chromosome 1p versus 9% in adult GBMs and 23% paediatric GBMs showed 2q gain versus 3% adult GBM. Similar findings were found for chromosome 21q gain, 6q loss, 11q loss and 16q loss – which were all more common in the paediatric astrocytomas. These results suggest that high-grade paediatric astrocytomas appear to contain some similarities, but also contain different genetic aberrations to those seen in adult astrocytomas.

Few studies have been performed in paediatric low-grade astrocytomas using CGH. Again, the frequent genetic aberrations present in adult GBMs were not found. The majority of low-grade astrocytomas from children and adults display balanced karyotypes. The most commonly seen aberrations in children are whole chromosome gains, involving chromosomes 5, 6, 7, 9, 11 and 19 or losses of 2 or 22, usually occurring as a single aberration (Schrock *et al.*, 1996, Sanoudou *et al.*, 2000, Jones *et al.*, 2006). The most common aberrations in adult low-grade astrocytomas were partial or whole gains of chromosome 7 (Schrock *et al.*, 1996), gains of chromosomes 19 or 22, or discrete losses at 1p33-pter, 9q34.1-ter, 17q21.3-ter, 2p22-pter, 9q31-qter and 12q13.2-q23 (Sanoudou *et al.*, 2000). Chromosome gains have been noted to be more common in older patients, aged >10 years with pilocytic astrocytoma (Jones *et al.*, 2006). Jones *et al.* showed a median of 4-5 chromosome gains per case in patients aged >10 years, whereas younger patients showed one or fewer chromosome gains. These results concur with the whole copy number gains identified in 6/20 tumours here, which were all identified in patients between the ages of 9 and 20 years (mean 14.6 years, standard deviation 4.13 years). Patients with no whole chromosome gains had a mean age of 6.5 years, standard deviation 3.4 years.

Array CGH analysis of the paediatric tumour samples here revealed that 40% of the samples had either a balanced karyotype, or a balanced karyotype with a single small region of chromosome gain or loss, within the limits of resolution of the technique. Whole chromosome gains were identified in 30% of the tumour samples, most commonly in chromosomes 6 and 5, and whole chromosome losses were most commonly seen for chromosomes 17 and 22, in order of prevalence. The results for whole chromosome gains from this study agree with previous findings for chromosome gains in low-grade astrocytoma. This study also revealed small numbers of chromosome losses and discrete regions of gain and loss, which had not been seen in earlier studies.

Partial losses of chromosome arm 1p were identified in 4 samples (PA32, PMG2, DA2 and DA4). The regions of loss varied in size from 45 to 55Mb. Two samples showed identical regions of terminal 1p loss, ending at the same BAC clone at 55Mb. Currently the functions of genes within this region (*C1orf175*, *TTC4*, *PARS2* and *TTC2*) are little understood; none are known to play any role in cancer. Two further samples also showed partial 1p losses, of different sizes. Both ended within regions of known copy number variation. The loss of chromosome 1p has previously been identified in adult astrocytomas. Complete loss of 1p, in tandem with 19q loss is a characteristic finding in oligodendrogliomas (Idbaih *et al.*, 2005). Oligodendrogliomas have distinct morphological appearances from other grade II astrocytomas. Patients with oligodendrogliomas, with complete 1p and 19q loss have been shown to have longer overall survival and progression-free survival when compared with grade II-III astrocytomas and oligoastrocytomas. Partial chromosome 1p losses, without concurrent 19q losses have also been identified in adult astrocytomas. These are most commonly found in GBMs, and appear to correlate with a poorer prognosis. None of the samples with 1p loss in this study had concurrent 19q loss – this was as expected, as no oligodendrogliomas were included in this study. However, the size of the region of loss made further investigation difficult, as so many genes are contained within the region. Further studies on additional samples may help to define the region of loss more closely. Additionally, the use of short-term tumour cultures, if available, would allow M-FISH and FISH to be performed, to map the boundaries of the region of loss more closely.

Two samples (DA7, DA9) displayed identical regions of gain covering 1.8Mb, at 102Mb on chromosome 2q11.1. The clones at both the start and end of this region of gain contained regions of known CNV, so the region of gain identified could represent normal variation between individuals. Two genes were within the gained region: solute carrier family 9 (sodium/hydrogen exchanger), member 2 (*SLC9A2*), and transmembrane protein 182 (*TMEM182*). Neither of these genes is known to have a role in brain tumours or other malignancies.

Two samples (PA13, DA4) revealed losses of a discrete region on chromosome 7q22.1, with a minimal common region of loss covering the clone RP11-506M12. This region contains the genes *MCM7* and *TAF6*. The protein encoded by mini-chromosome maintenance complex component 7 (*MCM7*) is one of a group of highly conserved MCM proteins, including MCM2, 4, 6 and 7. These proteins form a complex, which is essential for the initiation of DNA replication, and may act as a helicase to unwind the DNA (Tachibana *et al.*, 2005). CDK4 is known to associate with this protein complex, and may regulate its binding to the tumour suppressor protein Rb. *TAF6* RNA polymerase II (*TAF6*) encodes one of the subunits within transcription factor IID, one of the many proteins involved in the initiation of RNA transcription. *TAF6* has 4 isoforms, and one of these, *TAF6* delta has been shown to be part of an apoptosis signalling pathway, including activity in tumour cell lines lacking *p53* (Wilhelm *et al.*, 2008).

Comparison of cerebellar pilocytic astrocytoma samples with non-cerebellar samples did not reveal any significant regions of copy number change in the cerebellar group. A small region of gain at chromosome band 21q22.2 was identified in the pooled data from the group of varied non-cerebellar tumour samples, represented by 3 BAC clones. However, this region was only identified when the data were subtracted from the pooled cerebellar samples, and was not identified as a significant region of gain (ratio ≥ 0.3 in raw data, or by CBS) in individual patient samples. The entire length of chromosome 21 was covered by 26 BAC clones which were sparsely dispersed along the chromosome. Further examination of the data plots across chromosome 21 for individual tumours and individual arrays revealed a large amount of signal intensity variability. Hence, this apparent region of gain might purely represent experimental noise. One of the three gained clones was found in each of the 5 diffuse astrocytomas in the non-cerebellar group. It was possible that combining the sample data in this way had allowed recurrent aberrations present in multiple samples to be revealed, which were not considered

significant when considering analyses from single samples. Genes of interest in the 21q22.2 region include *BRWD1*, *HMGNI* and *DSCAM*. *BRWD1* (bromodomain and WD repeat domain containing 1) is a WD repeat protein family member, known to be involved in cell cycle progression, signal transduction, apoptosis and gene regulation. The encoded protein function is currently unknown. The protein encoded by *HMGNI* (high mobility group nucleosome binding domain 1) may alter the chromatin conformation of transcribed genes, by affecting DNA interaction with histone proteins (Cherukuri *et al.*, 2008). Finally, *DSCAM* (Down syndrome cell adhesion molecule) is known to play a role in CNS development as an axonal guidance molecule (Liu *et al.*, 2009). *DSCAM* expression is involved in the development of the choroid plexus, floor and roof of the 4th ventricle, pons and medulla (Barlow *et al.*, 2002), which makes it an attractive candidate to investigate further in non-cerebellar astrocytomas. I felt that these changes were most probably due to experimental noise – but was interested to follow up gain of this region, particularly in diffuse astrocytomas using a different technique, which will be described in chapter 4.

Additional regions of discrete gain and loss were also identified in single tumour samples. Discrete changes in single samples were less convincing as regions containing possible candidate genes, as these could be attributed to experimental noise. The array CGH slides used for this series of experiments were sufficiently powerful to identify gains or losses involving whole chromosomes, or large chromosomal regions. The overall resolution and quality of the array slides used was poor. Affymetrix 250k SNP arrays became commercially available (and affordable) towards the end of 2006. I decided to move to a higher resolution array platform to conduct further analysis of the low-grade astrocytoma samples, as I did not feel that the 1Mb array CGH platform had sufficient power to identify discrete regions of DNA copy number change. Hence, the focus of my DNA copy number investigation moved away from 1Mb array CGH to higher resolution SNP array platforms. This decision was subsequently justified by the identification of a recurrent discrete region of copy number gain in cerebellar pilocytic astrocytomas using Affymetrix 250K and 6.0 SNP arrays, which will be described in further detail in chapter 4.

In summary, array CGH analysis correctly identified regions of chromosome gain and loss within two medulloblastoma cell lines previously characterised by M-FISH and FISH. Examination of 20 cerebellar and non-cerebellar paediatric low-grade

astrocytomas showed 100% sensitivity in identification of patient sex, by sex chromosome copy number change. Balanced karyotypes, some containing a single discrete region of gain or loss were demonstrated in 40% of the tumour samples, and small numbers of whole chromosome gains and losses were identified, consistent with previously described findings in the literature.

Whole chromosome gains were found in patients aged 9-20 years. No whole chromosome gains were found in patients younger than 9 years, supporting data from a previous study (Jones *et al.*, 2006). There were no recurrent aberrations identified in the array CGH profiles of tumours from the cerebellum. Three recurrent aberrations involving discrete copy number changes were found, and these were the primary candidate regions of interest for the higher resolution SNP array study.

Chapter 4. DNA copy number changes by SNP array analysis

4.1 Introduction

All human genomes differ, except those from identical twins. Around 99.9% of the 3.2 billion base pairs forming the genome are common between individuals. However, the tiny 0.01% difference involves millions of base pairs, which form the basis of genetic variation (Kruglyak *et al.*, 2001). These genetic differences are inherited, and include traits for disease and disease susceptibility. Alleles are alternative forms of a particular gene. Humans contain two copies of each chromosome, and two alleles (one from each chromosome) form an individual's genotype for a particular gene. Alternate alleles are often termed A and B. An individual is homozygous where both alleles for a gene are the same (AA, BB), where the alleles are different the individual is heterozygous for a particular genetic trait (AB).

A single nucleotide polymorphism (SNPs) is a variation within the DNA sequence caused by the change of a single nucleotide. There are estimated to be ~ 10 million SNPs in the human genome and around 7 million common SNPs with minor allele frequency (MAF) greater than 5% (Kruglyak *et al.*, 2001). The MAF is the frequency of a less frequent allele for a SNP within a population. It is estimated that there are a further 4 million SNPs with a MAF from 1-5% (Hinds *et al.*, 2005). Previously SNPs were defined as polymorphisms occurring at a frequency of >1% within the population (Wang *et al.*, 1998), however the definition of a SNP has recently been refined, and a minimum frequency is no longer required. Hence, many SNP databases now include SNPs with a lowest allele frequency of <1%.

SNPs are the most frequent type of variation found in the human genome, and may vary within the same individual (where SNPs differ between alleles) or between individuals or species (LaFramboise, 2009). These changes may be used as genetic markers to study disease (Bignell *et al.*, 2004, Huang *et al.*, 2004). Large international consortia including the International HapMap Consortium and the 1000 Genomes Project are

attempting to identify, genotype and catalogue SNPs from large numbers of individuals from different ethnic backgrounds (Siva, 2008, HapMap, 2003). These data are freely available to researchers, and may be used to link SNP variation with specific disease traits.

Array technology was developed for use in SNP identification in 1998 (Wang *et al.*, 1998). Wang *et al.* designed 25-oligonucleotide probes centred on an individual SNP nucleotide, which was substituted at position 13 by adenine (A), cytosine (C), guanine (G) or thymine (T) in four otherwise identical probes. Biotinylated DNA was then hybridised to these four probes. DNA homozygous for the SNP sequence (AA) would hybridise strongly to the complementary probe, however heterozygotes (AB) and different homozygotes (BB) would show different hybridisation patterns at the central nucleotide (Figure 4.1). The SNPs within a DNA sample were amplified using multiplex PCR, using multiple primer pairs, however the formation of primer dimers limited the numbers of primers per reaction. Sample preparation was very labour intensive using this technique. A significant proportion (around 12%) of SNPs identified using this method were found to be false positives when compared to sequence data (Wang *et al.*, 1998). Non-specific signal hybridisation became more of a problem as the resolution of SNP arrays increased. Larger numbers of possible target sequences led to increases in cross-hybridisation of genomic DNA across probes, with a higher proportion of non-specific signals, as reviewed by Kennedy *et al.* in 2003 (Kennedy *et al.*, 2003).

Kennedy and co-researchers at Affymetrix developed the whole-genome sampling assay (WGSA) method currently used for SNP arrays in 2003, which will be described in more detail in the Methods section below (Kennedy *et al.*, 2003). The WGSA method increased hybridisation specificity by fragmenting the genomic DNA with restriction enzymes, followed by ligation to a specific adaptor sequence. PCR amplification, using a single primer set for the adaptor sequence, reduced both the complexity and fragment sizes of the DNA applied to the array, thus improving specificity. SNPs are subsequently genotyped by hybridisation to perfect match or mis-matched 25-mer probes for each SNP on the array. Mis-matched probes contain a single base mis-match at the centre of each probe (Figure 4.1). Probes are arranged in quartets, which comprise perfect match probes for alleles A and B, and mismatch probes for alleles A and B. This allows three possible genotypes (AA, AB, BB) to be determined for each SNP, using

the intensity measurements for each probe within the quartet. The signal intensity depends on the affinity between target and probes, and also on the amount of target DNA in the sample.

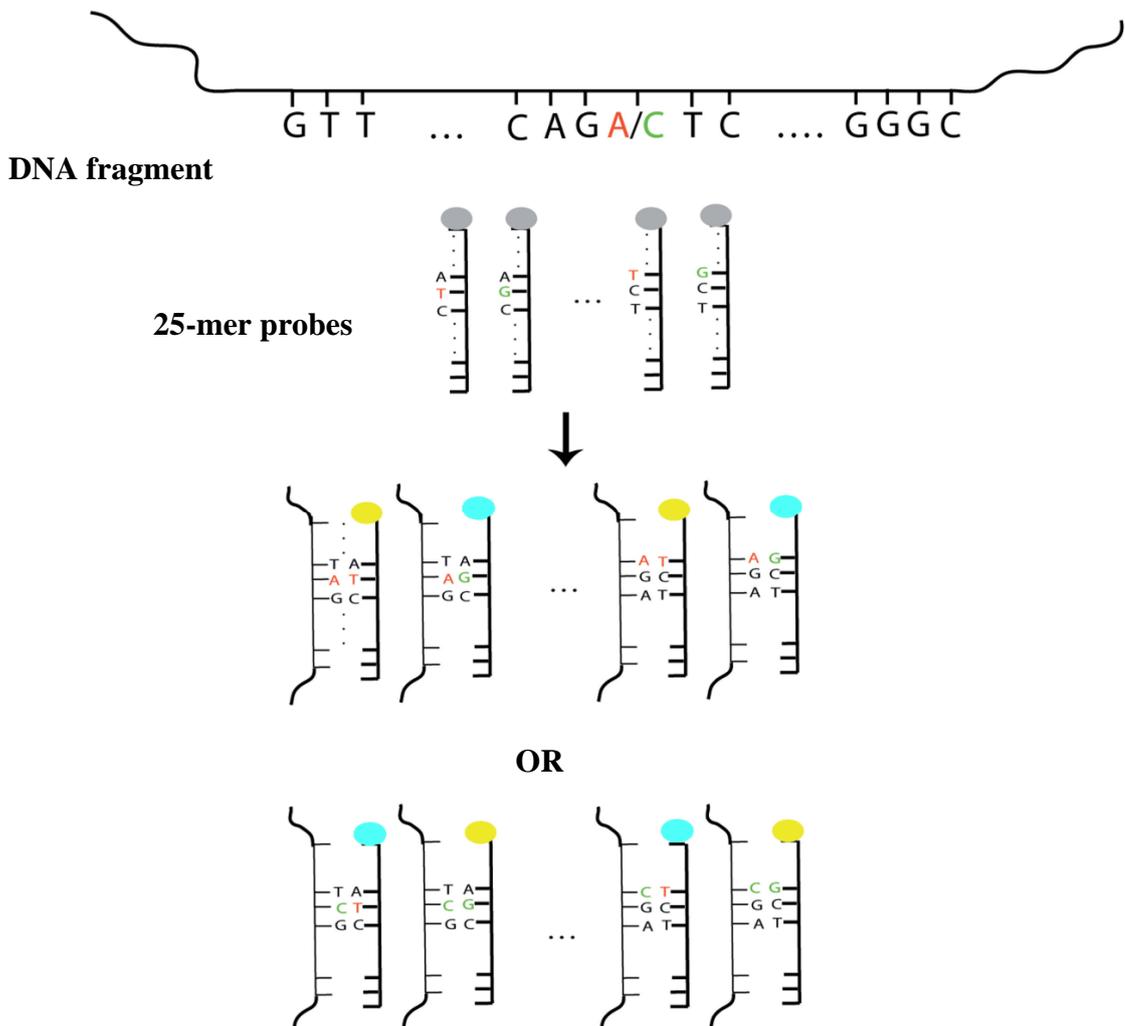


Figure 4.1 Overview of Affymetrix SNP array genotyping.

The DNA fragment containing a SNP with alternate alleles adenine/cytosine (A/C) is investigated with the 25-mer probes shown below. The DNA will bind to all of the probes, but binding is most efficient when all of the 25-mer probe bases are complementary (yellow), rather than the single base mismatch at the SNP site (blue). The perfectly matched and bound probes produce a strong fluorescent signal on scanning, the mismatched probes produce dimmer signals. Figure adapted from LaFramboise, Nucleic Acids Research 2009 (LaFramboise, 2009).

SNP arrays were originally designed for genotyping to investigate genetic diversity and disease associations and to undertake population studies. They have been used extensively for linkage disequilibrium studies to investigate genetic associations with disease, and have more recently been used for genome wide association studies (GWAS). GWAS investigates large numbers of individuals with and without a disease, and aims to find genetic variants that are more prevalent in those with the disease. The genetic variants associated with a disease may directly cause disease or increase susceptibility to a condition, or the regions of variation may be linked to nearby alleles. Linkage may be autosomal, where alleles are physically located close to each other on the same chromosome, or may be via linkage disequilibrium. Here, alleles at different loci, sometimes even on different chromosomes occur within the same haplotype more frequently than expected by chance (Wall *et al.*, 2003). Such large studies have huge statistical power to find genes implicated in disease; the first GWAS looking at genetic risk factors in adult astrocytoma was recently published (Shete *et al.*, 2009). In this study, 550 000 single nucleotide polymorphisms (SNPs) from 1,878 glioma cases and 3,670 controls were genotyped. Five genomic regions with strong disease associations were identified, including novel candidate genes. These included the telomerase gene *TERT* at chromosome 5p15.33 (odds ratio: 1.27, $p=1.5 \times 10^{-17}$), the *CCDC26* gene at chromosome 8q24.21 (odds ratio: 1.36, $p=2.34 \times 10^{-18}$) known to modulate cell differentiation and cell death as well as to down-regulate telomerase, the cyclin dependent kinase inhibitor genes *CDKN2A* and *CDKN2B* on chromosome 9p12.3 (odds ratio: 1.24, $p=7.24 \times 10^{-15}$) which play a pivotal role in cell cycle control, the gene *PHLDB1* on chromosome 11q23.3 (odds ratio: 1.18, $p=1.07 \times 10^{-8}$) the function of which is little understood, as well as the RAD3-like helicase encoding gene *RTEL1* at chromosome 20q13.33 (odds ratio: 1.28, $p=2.52 \times 10^{-12}$) involved in telomere maintenance. Interestingly, telomere maintenance appears to be a recurring theme among the genes identified, and this will need to be investigated further in adult astrocytoma (Shete *et al.*, 2009).

In addition to the detection of genes implicated in disease by association studies, SNP array technology has become an important tool in the detection and characterisation of copy number variation and loss of heterozygosity (LOH) in cancer. The identification of LOH, which may occur at chromosomal regions containing tumour suppressor genes, was previously performed by genotyping using microsatellite markers (short tandem repeat polymorphisms containing 1-4 bp repeats) (Weber *et al.*, 1989). However, this

process was extremely time consuming and labour intensive. The use of SNP arrays to identify LOH revolutionised the field (Lindblad-Toh *et al.*, 2000), as the technique is simple, quick and allows large numbers of samples to be analysed in a high-throughput fashion, which would previously have been unthinkable.

SNP arrays have also been used to investigate DNA copy number in the cancer genome. The fluorescent intensity of SNP array probe features vary with DNA copy number changes, in exactly the same way as previously described in array CGH (Bignell *et al.*, 2004, Huang *et al.*, 2004). Hence, SNP arrays are now used to investigate both DNA copy number and loss of heterozygosity in tumour samples, providing a greater depth of information from a single DNA sample that is not possible from array CGH alone. This is extremely important as some regions within tumours that display LOH do not show DNA copy number changes, and would not be identified by CGH. SNP arrays therefore provide huge advantages over CGH arrays, in terms of the data produced per sample. SNP arrays are manufactured under strict quality control conditions, providing accurate and highly reproducible genotyping. However they do not provide a comparative hybridisation between test and control DNA in a single experiment. Therefore, samples from normal individuals must be run using the same platform, to allow comparisons to be drawn between tumour samples and normal individuals. Data must be compared between 'normal' and experimental samples, using statistical methods to determine relative gains and losses of DNA across a sample set to improve statistical power. Ethnicity must also be considered when selecting control samples, as polymorphisms in SNPs are known to vary between ethnic groups.

The first Affymetrix GeneChip Mapping 10k array, developed in 2003, contained 10,044 SNP markers with a mean inter-marker distance of 105kb. Since then, Affymetrix have produced several generations of SNP arrays, each at higher resolution. The GeneChip Human Mapping 500K array set enables 500,000 SNPs to be genotyped using two separate arrays (each 250K). The Affymetrix Genome-Wide Human SNP Array 6.0 contains 1.8 million markers for genetic variation on a single array; >906,600 SNPs and >946,000 probes to detect regions of copy number variation. This SNP array provided the highest genotyping coverage to date, with a median marker spacing of 680 bases (Affymetrix 6.0 datasheet). However, Affymetrix have recently announced the development of a new SNP array platform containing 2.7 million probes, which will be available later in 2009. Illumina has developed an alternative platform using 50-bp

oligonucleotides attached to indexed beads, randomly arrayed on glass slides. Their current highest-resolution array contains 650,000 oligonucleotides, with a median SNP spacing of 2kb.

4.2 Aim

The aim of the work described within this chapter was to investigate genome-wide DNA copy number changes in paediatric low-grade astrocytomas using Affymetrix 250K and 6.0 SNP arrays.

4.3 Affymetrix 250K Nsp SNP array

4.3.1 Materials

An initial study of 21 tumour samples was run with Tim Forsheew, using the Affymetrix GeneChip Human Mapping 250K Nsp Array.

The samples comprised 14 pilocytic astrocytomas, 4 diffuse astrocytomas, one infantile hypothalamic pilomyxoid astrocytoma, one pilomyxoid glioma and one pleomorphic xanthoastrocytoma. Data from the tumour samples were compared to those from two healthy male DNA controls (extracted from co-workers at the Cancer Research UK London Research Institute), which were run on the same platform.

Briefly, the GeneChip Human Mapping 500K array set comprises two separate arrays (each covering 250,000 SNPs or 250K) and two separate assays. The assays each utilise one of two restriction enzymes, *Nsp I* or *Sty I* to digest the genomic DNA into fragments (Figure 4.2). Oligonucleotide adaptor sequences are then ligated onto the ends of the DNA fragments. The adaptors bind specifically to 4bp overhangs at the *Nsp I* or *Sty I* restriction sites. PCR is then performed, with conditions optimised to amplify fragments between 200-1100 base pairs long. The amplified DNA is then fragmented, denatured and labelled prior to hybridisation onto the array. The process is summarised in Figure 4.1.

For these experiments, the Nsp 250K array was selected which contains probes for 262,264 SNPs, each 25 base pairs long. Each SNP is represented by 6 or 10 probe quartets. Each quartet contains perfect match and mis-match probes for each of the two alleles, A and B. Therefore 24 or 40 different 25-bp oligonucleotides are present on the array to probe each SNP.

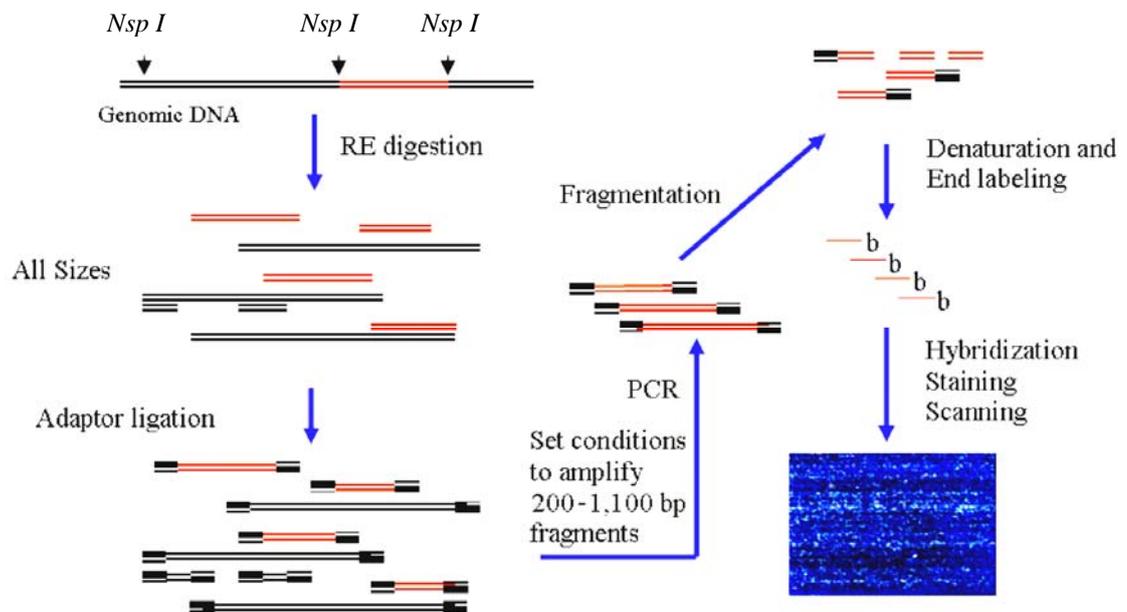


Figure 4.2 Overview of Affymetrix 250K Nsp GeneChip mapping assay.

Genomic DNA was digested with the restriction enzyme *Nsp I*. The resulting fragments were ligated with adaptors specific for the restriction sites. A single PCR primer was used, which recognised the adaptor sequence used to amplify the ligated fragments. PCR conditions were optimised to amplify fragments between 200 and 1,100 base pairs. Amplified DNA was fragmented, denatured and labelled before hybridisation to the 250K Nsp array. The arrays were then washed and scanned. Figure adapted from the Affymetrix 500K protocol.

4.3.2 Method

The 250K experiments were performed following the protocol provided by the manufacturer (Affymetrix Inc., Santa Clara, CA). Two hundred and fifty nanograms of genomic DNA at 50ng/ μ l were used per array. DNA was digested using the *Nsp I* restriction enzyme master mix, and ligated with the *Nsp I* adaptor. Following PCR, the PCR products were purified using the Ultrafree-MC filtration column (Millipore,

Billerica, MA); this was the only modification from the manufacturer's protocol. After DNase-I digestion, the PCR products were labelled with a biotinylated nucleotide analogue using terminal deoxynucleotidyl transferase, and hybridised to the array.

Array hybridisation, washing and scanning was performed by Tracy Chaplin at the John Vane Science Centre, Charterhouse Square, London. Arrays were scanned using a GeneChip 3000 7G scanner (Affymetrix, Santa Clara, CA). Data were exported as cell intensity (CEL) and genotyping analysis (CHP) files.

4.3.3 Analysis

Analysis of the Affymetrix 250K Nsp SNP array data was performed with Tim Forsheew.

Genotype calling was performed using GTYPE (Affymetrix GeneChip Genotyping Analysis Software), which converts the 6 or 10 probe quartet intensity measures into a genotype inference (AA, AB or BB). GTYPE uses an automated genotype-calling algorithm, which calculates the sum of signals from the perfect match probes from the array, and subtracts the sum of signals from the mismatched probes for each SNP. A confidence score is then assigned for each genotype. This can be used as a quality control, where SNPs with low confidence scores and SNPs where no call is assigned are not used for further copy number analysis. Samples with a call rate of >95% were used for further analysis in this study.

Copy number analysis was performed using Copy Number Analyzer for Affymetrix GeneChip Mapping (CNAG) version 2.0 ((Nannya *et al.*, 2005). CNAG is an algorithm for SNP array copy number analysis, which uses hidden Markov modelling and local mean analysis to reduce the variation in the raw signal ratios, thus improving the signal to noise ratio. Accurate genotyping information derived from CNAG enables inference of LOH in addition to allele-based copy number analysis. CNAG was developed to allow analysis of samples from different experiments, by compensating for variation between different experimental conditions. CNAG also allows comparison of tumour samples with multiple normal reference samples in the absence of constitutional DNA

as a paired control, which was the case for all of the tumour samples in this study. Data from the tumour samples were compared to data from two healthy male DNA controls.

Signal intensity, genotype information and gender calls are extracted from the .CEL and .CHP files. CNAG normalises the data from the sum of signals over the 6 or 10 perfect match probes for the A and B alleles to perform genotyping. Linear regression is then performed, to compensate for the length of PCR fragments and GC content of the PCR products, which reduces variation between experiments. The \log_2 ratios of these values are then used to compare between different arrays. Signal intensity, genotype information and gender calls are used to create .CFH files, which are used for copy number calculations. For these, CNAG summarises the data from the sum of signals of the 6 or 10 perfect match probes for the A and B alleles to create a single 'raw' copy number measure for each SNP. These 'raw' copy number measures are rough measurements of the true underlying copy number state. CNAG performs pair-wise analysis between each sample and the available normal reference samples, and calculates the best combination of references to minimise the standard deviation for diploid autosomal (non-sex chromosome) regions. Gender information is used to infer copy number accurately in the X chromosome. Where patient sex is unknown, gender can be inferred by a significant number of heterozygous SNP calls from the X chromosome. However, LOH on the X chromosome could incorrectly indicate male gender in tumour specimens with deletion of one X chromosome. The sex was known for all of the cases included in this study.

CNAG uses two methods of analysis to infer DNA copy number from SNPs. The overall aim is to infer segments of locally constant copy number change in the tumour samples from the noisy raw copy number SNP data. In local mean analysis, the mean \log_2 ratio is calculated between 3 or 10 SNPs depending on the standard deviation of the initial \log_2 values after linear regression. This further eliminates random noise, and helps to visualise the array copy number data. An alternate method employed by CNAG is hidden Markov modelling, which infers the copy number in a 'hidden state' (the true integer copy number) by comparison with the 'observed' state (the raw copy number, produced by the summarisation step). Hidden Markov modelling uses the assumption that a copy number change is due to genetic recombination between adjacent SNP loci, and calculates the probability of a given copy number change by comparison with neighbouring SNPs.

4.3.4 Results

The data obtained from the Affymetrix 250K array analysis are available in the GEO database under Accession number GSE14960.

Copy number analysis of the 21 tumour samples revealed a novel finding – a discrete copy number gain of ~1.9 Mb situated at chromosome band 7q34, in 12/14 cerebellar pilocytic astrocytomas and in the single hypothalamic pilomyxoid astrocytoma (Figures 4.3, 4.4). Furthermore, two samples showed a discrete region of copy number loss at 9p21, containing the *CDKN2A* locus. These were both grade II astrocytomas: 1 diffuse astrocytoma from the cerebral cortex (DA2) and 1 pleomorphic xanthoastrocytoma (PXA1) (Figure 4.4). Additional small regions of gain and loss were identified, however the majority of these were situated in known regions of copy number variation.

Whole chromosome gains of chromosomes 5 and 6 were identified in 3 samples, respectively. These were all pilocytic astrocytomas from the cerebellum (PA12, PA14, PA27), patients were aged 14, 10 and 14 years, respectively (Figure 4.5). Single whole chromosome gains for chromosomes 7, 11, 13, 14, 15 and 20 were also identified, again in samples PA14 and PA27. The data are summarised in Table 4.1.

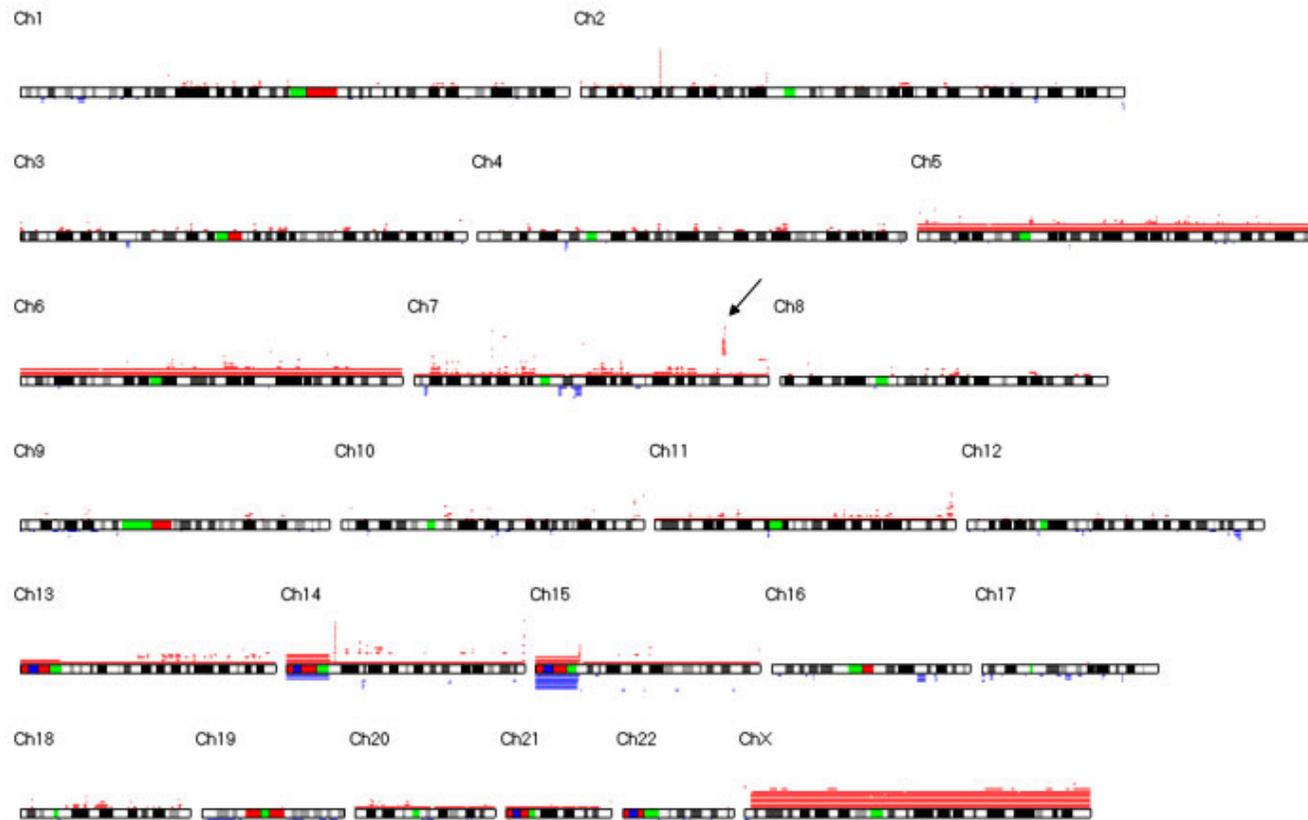


Figure 4.3 DNA copy number gains and losses in 21 tumours, assessed by 250k Nsp SNP array.

Red lines above the chromosome indicate gains, blue lines below indicate losses. A discrete copy number gain of ~1.9Mb in chromosome band 7q34 was identified (indicated by arrow). Whole chromosome gains of chromosomes 5 and 6 were identified in 3 samples, respectively. Data obtained from CNAG.

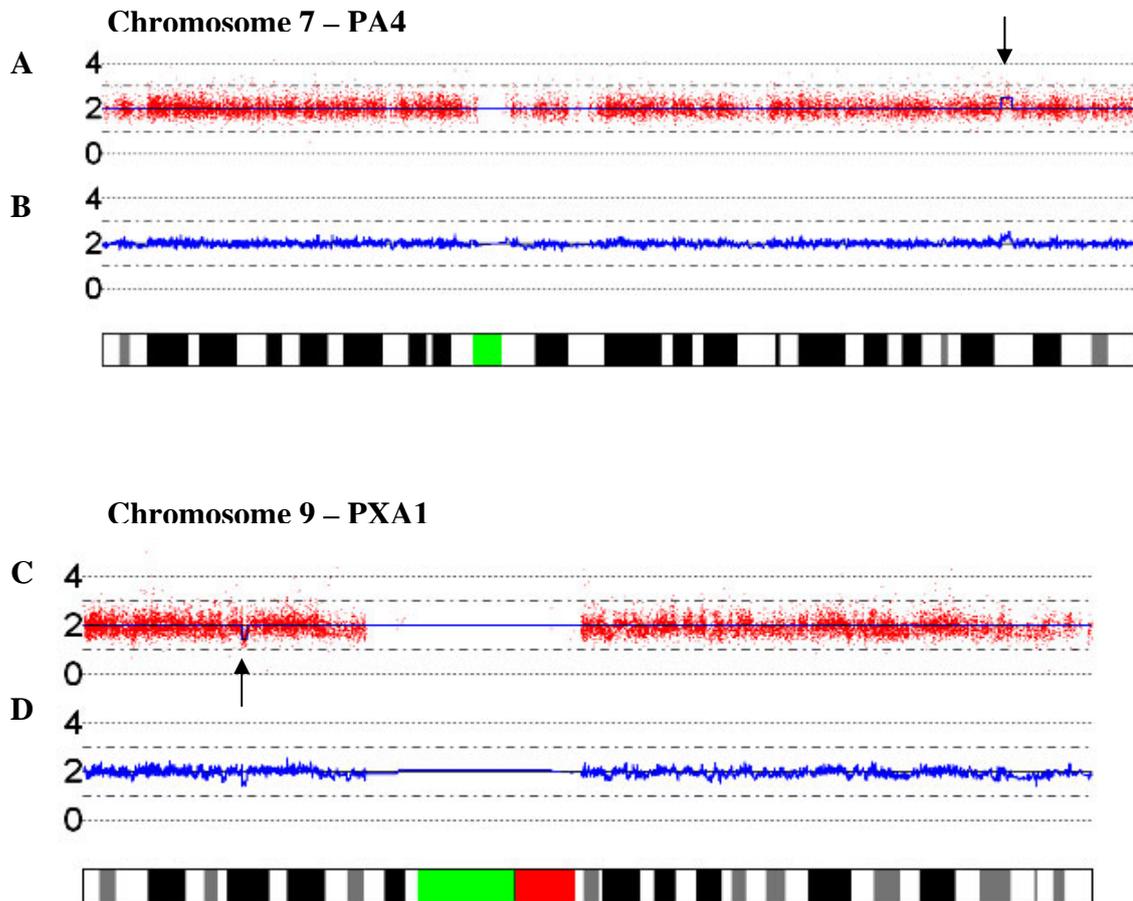


Figure 4.4 Copy number views of SNP data showing discrete regions of DNA copy number gain and loss in individual tumour samples, obtained using CNAG.

The X-axis displays genomic position, as shown on the chromosome ideogram below. The Y-axis displays DNA copy number. The upper traces in red (A, C) show raw SNP data. Here, the majority of SNPs are present in 2 copies, consistent with a diploid karyotype. There is considerable variation between individual SNPs in the raw trace, and the blue line within this upper trace shows a smoothed copy number trace, created using hidden Markov modelling. The lower traces in blue (B, D) are smoothed SNP data traces, created using local mean analysis.

- (A) DNA copy number gain at chromosome band 7q34 in sample PA4
- (B) Smoothed SNP copy number trace showing copy number gain at 7q34.
- (C) DNA copy number loss at chromosome band 9p21 in sample PXA1.
- (D) Smoothed SNP data trace showing copy number loss at 9p21

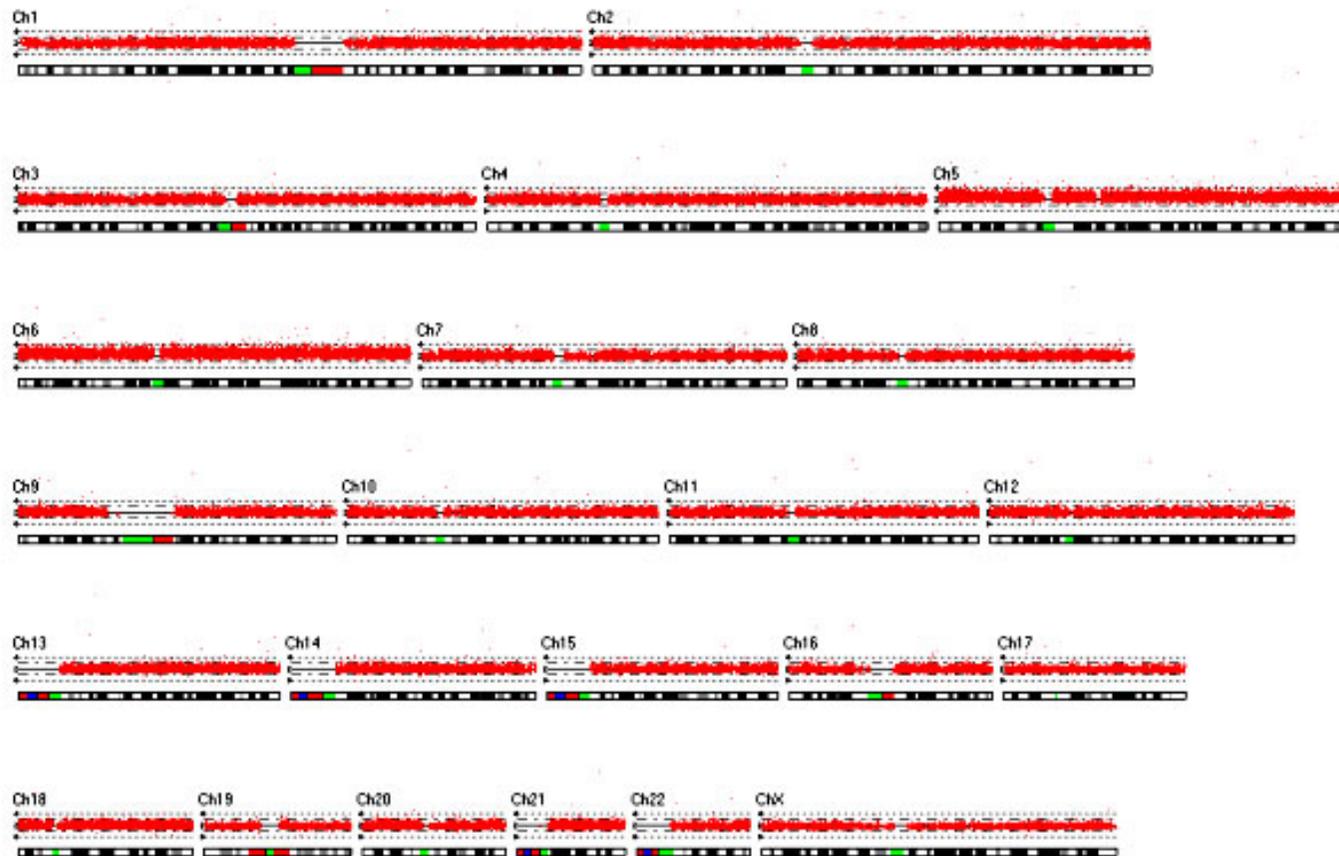


Figure 4.5 Whole genome DNA copy number view of raw SNP data for sample PA12, showing whole chromosome gains of chromosomes 5 and 6.

Name	Site	Age (years)	Sex	7q34 gain	9p21 loss	Whole chromosome gains
PXA1	Cerebral cortex	6	F	-	+	
DA2	Cerebral cortex	9	M	-	+	
DA7	Cerebral cortex	5	M	-	-	
DA8	Cerebral cortex	15	F	-	-	
DA9	Cerebral cortex	9	F	-	-	
PMA1	Diencephalon	1	F	+	-	
PMG2	Diencephalon	7	F	-	-	
PA3	Cerebellum	7	M	+	-	
PA4	Cerebellum	2	M	+	-	
PA7	Cerebellum	13	M	+	-	
PA10	Cerebellum	19	M	+	-	5, 6, 7, 14
PA11	Cerebellum	14	M	+	-	5, 6
PA13	Cerebellum	10	M	+	-	
PA14	Cerebellum	10	M	+	-	
PA16	Cerebellum	10	M	-	-	
PA17	Cerebellum	6	M	-	-	
PA18	Cerebellum	3	M	+	-	
PA20	Cerebellum	4	M	+	-	
PA22	Cerebellum	6	F	+	-	
PA26	Cerebellum	3	F	+	-	
PA29	Cerebellum	14	F	+	-	5, 6, 11, 13, 15, 20

Table 4.1 Summary of DNA copy number gains and losses identified in 21 low-grade astrocytoma samples by 250K SNP array analysis.

(PXA = pleomorphic xanthoastrocytoma, DA = diffuse astrocytoma, PMA = pilomyxoid astrocytoma, PMG = pilomyxoid glioma, PA = pilocytic astrocytoma)

4.4 Affymetrix Genome-Wide Human 6.0 SNP array

The study was extended, to investigate the copy number changes identified above at higher resolution, using the Affymetrix 6.0 array. Thirty-seven tumour samples were included in this study. These comprised 26 pilocytic astrocytomas, 6 diffuse astrocytomas, 3 pilomyxoid astrocytoma, one infantile hypothalamic pilomyxoid glioma and one pleomorphic xanthoastrocytoma. Seventeen of these tumour samples had previously been investigated using the Affymetrix 250K array.

Affymetrix Genome-Wide Human 6.0 SNP array experiments were carried out at the Hartwell Centre for Bioinformatics and Biotechnology, St Jude Children's Research Hospital, Memphis, USA, in collaboration with Professor David Ellison. The experiments were performed following the manufacturer's instructions. Briefly, 500 nanograms of genomic DNA were digested with both NspI and StyI restriction enzymes, and ligated to adaptors specific for the 4bp overhangs as described above. PCR was performed using a single generic primer to amplify the ligated fragments, with conditions set to preferentially amplify fragments between 200-1,100 bp. PCR products from both restriction enzyme digests were combined and purified using polystyrene beads. The amplified DNA was then fragmented, labelled and hybridised to the SNP array.

4.4.1 Analysis

Affymetrix Genotyping Console 3.0.2 software was used for most of the analysis for the 6.0 SNP arrays. Additional analysis tools were used, which gave slightly different results, particularly on mapping the boundaries for regions of copy number change.

Control .CEL data from Affymetrix Genome-Wide Human 6.0 SNP arrays were obtained from the International HapMap project for 48 unrelated Caucasian controls (27 males and 21 females). All .CEL data, from tumour samples and the 48 controls were uploaded into Affymetrix Genotyping Console 3.0.2. Initial genotyping was performed using the Birdseed (version 2) algorithm (Korn *et al.*, 2008). Birdseed estimates the signal intensity for each allele of each SNP, and assigns AA, AB and BB genotype calls

to each sample containing 2 SNP alleles. SNPs are then assigned to AA, AB or BB clusters. These clusters are initially defined by the expected location for each SNP, derived from 270 HapMap sample genotypes. SNP positions are then refined according to each tumour sample SNP assignment.

A reference copy number file (.REF) was created, within the Affymetrix Genotyping Console 3.0.2, using the genotype data from the 48 control samples. This was then used as the baseline for copy number comparisons at each SNP for the tumour samples. Tumour sample data were normalised using quantile normalisation. This aims to make the distribution of probe intensities the same for each array within a wider set of arrays. Probes are sorted by intensity across each array, and the mean intensities are then calculated by rank across the wider array set. Individual probe intensities at each rank are then replaced by the mean intensity from the wider group. However, each tumour sample will retain an individual genotype ‘fingerprint’, as the order of probes is specific for each when sorted by intensity. \log_2 ratios between each tumour and the reference were calculated for each SNP. Smoothed signals were also constructed, using a weighted mean of the \log_2 ratios of surrounding probes. The weights are proportional to the genomic distance from the location of each test probe. Allele differences were calculated, by subtracting the signals for allele B from allele A, each standardised with respect to their median values in the reference copy number file. These data may be used to infer regions of LOH.

Copy number states were calculated using hidden Markov modelling, as previously described. The overall aim is for the copy number analysis to be able to infer segments of constant copy number change in the tumour samples when compared to the control samples, from the raw SNP copy number data. Further analysis was performed using the segment reporting tool within Affymetrix Genotyping Console 3.0.2. This aims to identify regions or segments of copy number change in tumour samples, which are not within known regions of copy number variation (CNV). The Affymetrix CNV map files, denoting regions of CNV, were obtained from the Toronto database of genomic variants (<http://projects.tcag.ca/variation/>) (Iafrate *et al.*, 2004). The segment reporting tool uses the copy number states derived using hidden Markov modelling to define regions of copy number change. For this analysis the parameters were defined to identify segments larger than 200Kb with <50% overlap with known CNVs. Individual segments were also reviewed by eye.

Regions of copy number gain or loss were also analysed independently by Dr Jing Ma at the Hartwell Centre for Bioinformatics and Biotechnology, St Jude Children's Research Hospital, Memphis, USA. Here, the tumour data were compared to 90 reference samples obtained from children undergoing long-term follow-up, >5 years from diagnosis of leukaemia. Affymetrix SNP 6.0 array data were processed from .CEL files to extract raw signal intensity values using dChip PM-only model-based expression analysis (Zhao *et al.*, 2004). Data were then normalized using a reference normalization algorithm (Pounds *et al.*, 2009), which uses only markers from chromosomes known or predicted to be disomic to guide array normalization. A \log_2 ratio of tumour versus the median signal of 90 reference samples was calculated for each marker in each array using the reference-normalized signal intensity values. The circular binary segmentation algorithm (Olshen *et al.*, 2004) implemented in the DNA copy package from Bioconductor (Gentleman *et al.*, 2004) was applied to the above \log_2 ratio data, to identify change points denoting regions of copy number alteration for each tumour sample.

4.4.2 6.0 array results

The results in this section were obtained working with Tim Forsheew, Andrew Lawson, Jing Ma and David Ellison.

Copy number analysis of the 37 tumour samples revealed a discrete region of copy number gain of ~1.9 Mb, situated at chromosome band 7q34 in 24/37 low-grade astrocytomas. These included 23 pilocytic astrocytomas and the hypothalamic pilomyxoid astrocytoma (Figure 4.6). Further analysis, using the segmentation tool from Affymetrix Genotyping Console 3.0.2, revealed that the gained region at 7q34 shared similar start and end points in the majority of tumours. Segmentation analysis identified 23/24 tumours with 7q34 gains detected by SNP copy number analysis, and provided detailed information on the exact genomic locations for the start and end of the gained regions. Using this method of analysis there appeared to be some variation between the start and end of the gained region between samples. Ten of the 23 samples with 7q34 gain by segmental analysis showed the start of the gained segment within

KIAA1549 and 14/23 showed termination of the gain within *BRAF*. The exact positions for these segments, in megabases, are shown in Table 4.2.

Start and end points within *KIAA1549* and *BRAF* were also revealed by separate analysis using circular binary segmentation (CBS), performed by Jing Ma at St Jude. Some of the 6.0 SNP data were excluded from CBS analysis; these included four of the PA samples (PAs 12, 15, 26, 29). For these cases the raw SNP data appeared to be more 'noisy' than for other samples, despite all of the samples passing initial DNA quality control checks. The gained region identified by CBS, started within the middle of the gene *KIAA1549*, and ended within the middle of the gene *BRAF* in the majority of the tumours containing 7q34 gain (Figure 4.7). Considerably less variation was apparent between the start and end of the gained regions in samples by this method of analysis. CBS found 19/19 samples with the start of the gained region within *KIAA1549*, and 17/19 with termination of the gained region within *BRAF* (Table 4.3).

One pilocytic astrocytoma sample (PA30) contained a discrete ~3Mb region of copy number gain at chromosome band 3p25. The region of gain began within the middle of the gene *SRGAP3*, and ended within the middle of the gene *RAF1* (Figure 4. 8). Segmental analysis identified the region of gain between 9,067,444-12040620Mb, also placing the start within *SRGAP3*. Here, however, the region of gain ended within *SYN2*, ~ 500Kb upstream from *RAF1*. Analysis by CBS was not performed for this sample, as the raw SNP data appeared too 'noisy'.

Two grade II astrocytoma samples, one diffuse astrocytoma and one pleomorphic xanthoastrocytoma (DA2 and PXA1), showed discrete copy number loss at chromosome band 9p21, containing the gene *CDKN2A* (Figure 4.9). Segmental analysis of PXA1 revealed heterozygous copy number loss of ~600kb, between 21,791,530 - 22,413,279Mb. This region contained *CDKN2A* and also a known region of CNV. The second sample DA2 showed a larger ~4Mb region of homozygous loss, within the context of heterozygous loss for the majority of chromosome 9p. This region of loss was between 21,571,712 – 25,614,770, and contained the genes *MTAP*, *CDKN2A* and *ELAVL2*. CBS was not carried out for DA2. However, CBS analysis was performed for sample PXA1, and agreed exactly with the 21,791,530 - 22,413,279Mb region identified by segmental analysis.

One diffuse astrocytoma, DA3, contained a discrete region of loss at 10q24.32, between 104,264,147-104,846,119Mb (Figure 4.10(A)). This region contained the genes *SUFU* (suppressor of fused homolog Drosophila), where loss of function is known to be associated with the development of medulloblastomas, through overactivity of the Sonic Hedgehog signalling pathway, and *TRIM8* (tripartite motif-containing 8), which may also be a tumour suppressor gene (Toniato *et al.*, 2002, Taylor *et al.*, 2002). Analysis using CBS identified a very similar region of loss, between 104,017,691-105,023,374Mb, also containing *SUFU* and *TRIM8*.

This same sample, DA3, also contained a discrete region of gain at chromosome 11p13, of ~300Kb (Figure 4.10(B)). The region of gain was between 33,196,607-33,514,710MB by the segment reporting tool. A single gene, *HIPK3* (homeodomain interacting protein kinase 3), is found within this region. *HIPK3* is a serine/threonine kinase/homeodomain-interacting protein kinase, which modulates the Fas signalling pathway (Curtin *et al.*, 2004). CBS identified a similar region of gain, between 33,099,702-33,519,549Mb, which also contained *HIPK3*.

One pilocytic astrocytoma sample (PA2) contained a discrete ~2Mb region of gain at 11q23.1, in addition to gain at 7q34 (Figure 4.10(C)). This was between 111,018,800 – 113,008,870Mb by segmental analysis and CBS, and contained, among others, the gene *NCAM1* (neural cell adhesion molecule 1).

Whole chromosome gains were identified in 9/37 samples. These aberrations are summarised in Tables 4.4-4.5.

Loss of heterozygosity (LOH) analysis was performed, using the LOH analysis tool within Affymetrix Genotyping Console 3.0.2. LOH analysis did not identify any large regions of homozygosity. Particular attention was paid to the LOH analysis for chromosome 17 from patient PMA1, who had been diagnosed with Neurofibromatosis type I on clinical grounds. No LOH was demonstrated at the *NF1* locus on chromosome band 17q11.2.

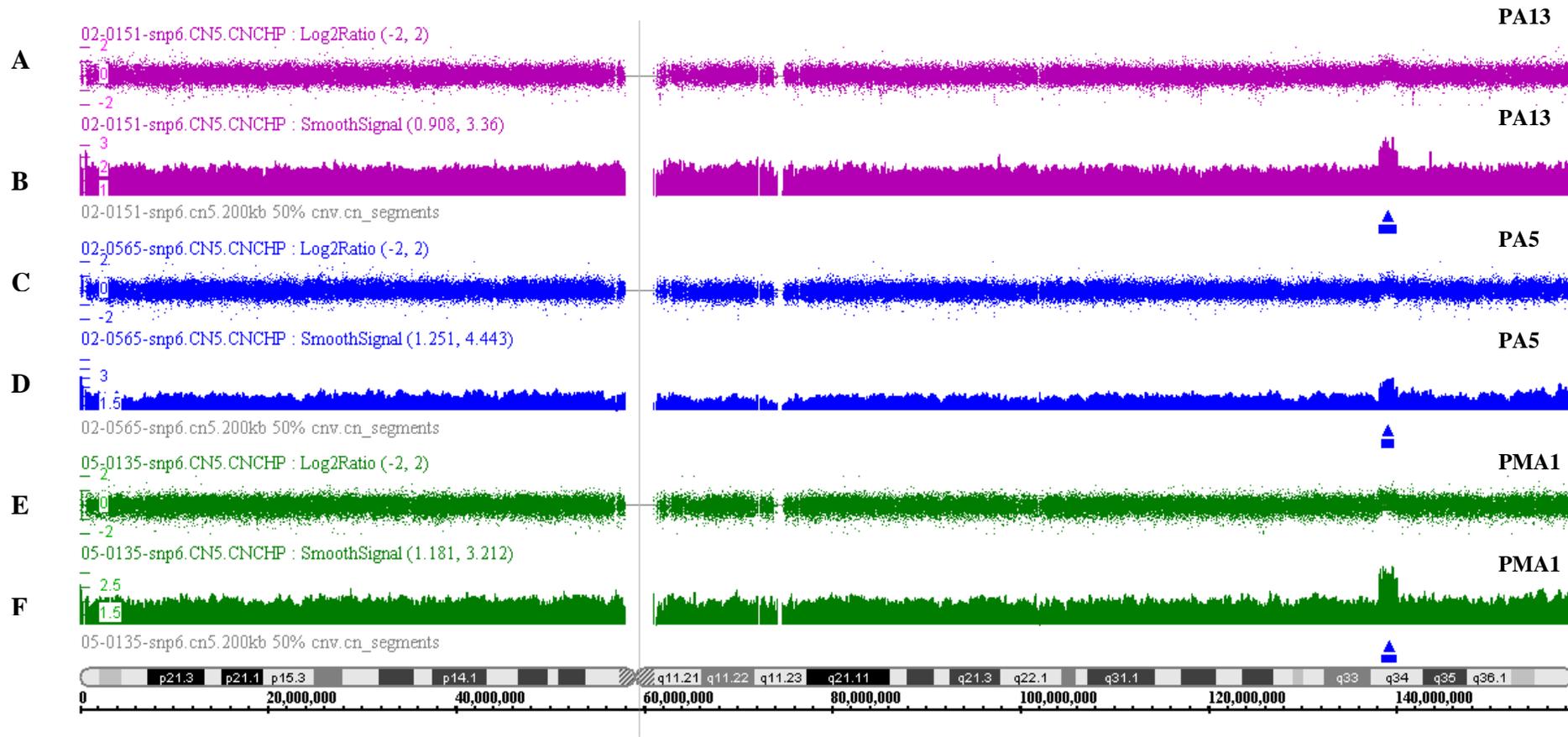


Figure 4.6 Chromosome 7 copy number analysis by Affymetrix 6.0 SNP array reveals 7q34 gain in tumour samples PA13, PA5 and PMA1. The upper traces (A, C, E) show raw log₂ ratios and the lower traces (B, D, F) show smoothed signals. Regions of copy number gain identified by the segmentation tool are shown as a blue bar overlaid with an arrow.

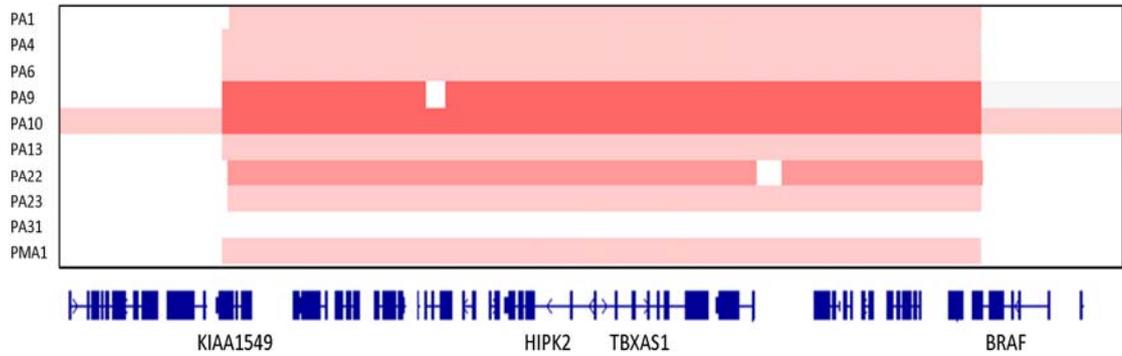


Figure 4.7 DNA copy number gain at 7q34 in 9/10 representative low-grade astrocytomas, identified by circular binary segmentation.

Red shading represents regions of DNA copy number gain, where increased colour intensity indicates the degree of amplification. The boundaries of the gained region fall within the genes *KIAA1549* and *BRAF*, as shown below. Sample PA31 did not contain gain at 7q34.

Sample	Start of segment, Mb	End of segment, Mb
PA1	138,207,608	140,096,654
PA2	138,207,608	139,656,577
PA3	138,413,825	139,719,145
PA4	138,182,168	140,134,713
PA5	138,414,600	139,754,451
PA6	138,172,165	140,144,933
PA9	138,182,168	140,134,713
PA10	137,807,619	141,409,151
PA11	138,814,009	139,754,451
PA12	139,520,846	140,090,492
PA13	138,182,168	140,142,094
PA14	138,182,168	140,142,094
PA15	138,154,050	140,090,492
PA17	138,216,399	139,754,451
PA18	138,547,517	139,754,451
PA19	138,182,168	140,096,654
PA22	138,467,471	140,082,280
PA23	138,401,224	140,134,713
PA25	138,814,009	139,689,768
PA26	138,839,105	139,590,942
PA27	-	-
PA29	138,592,363	140,090,492
PA32	138,188,628	140,141,917
PMA1	138,467,471	140,134,713

Table 4.2 Segment reporting identified the start and end of the 7q34 region of gain within 23/24 samples containing 7q34 gain by 6.0 SNP analysis.

Start segments highlighted in grey are contained within the gene *KIAA1549*, end segments highlighted in grey are within *BRAF*. The genomic regions for *KIAA1549* are 138,166,667 - 138,255,110Mb and *BRAF* 140,080,282 - 140,271,033Mb (University of California Santa Cruz [UCSC] genome browser, March 2006 assembly). Sample PA27 did not appear to contain segmental gain ≥ 200 Kb, however on visual inspection of the raw data PA27 did contain copy number gain at 7q34. The height of the gain was < 2.5 copies; this sample may have contained a proportion of normal tissue, thus reducing the size of the copy number gain.

Sample	Start of segment, Mb	End of segment, Mb
PA1	138207878	140141903
PA2	138216385	140134561
PA3	138191752	140134713
PA4	138188628	140134713
PA5	138253456	140129896
PA6	138188768	140141917
PA9	138189780	140134713
PA10	138189780	140141917
PA11	138207608	140121296
PA12	Not analysed	Not analysed
PA13	138189780	140142320
PA14	138184937	140240542
PA15	Not analysed	Not analysed
PA17	138192396	139696552
PA18	138188768	139486243
PA19	138184937	140129287
PA22	138202445	140144933
PA23	138202445	140141903
PA25	138192396	140134462
PA26	Not analysed	Not analysed
PA27	-	-
PA29	Not analysed	Not analysed
PA32	138189780	140141903
PMA1	138191752	140134713

Table 4.3 Start and end of gained segments at 7q34, identified by circular binary segmentation analysis.

Start segments highlighted in grey are contained within the gene *KIAA1549*, end segments highlighted in grey are within *BRAF*. The genomic regions for *KIAA1549* are 138,166,667 - 138,255,110Mb and *BRAF* 140,080,282 - 140,271,033Mb (UCSC, genome browser, March 2006 assembly). Sample PA27 did not appear to contain segmental gain ≥ 200 Kb, however on visual inspection of the raw data copy number gain was seen at 7q34. The height of the gain was < 2.5 copies; this sample may have contained a proportion of normal tissue, thus reducing the size of the copy number gain.

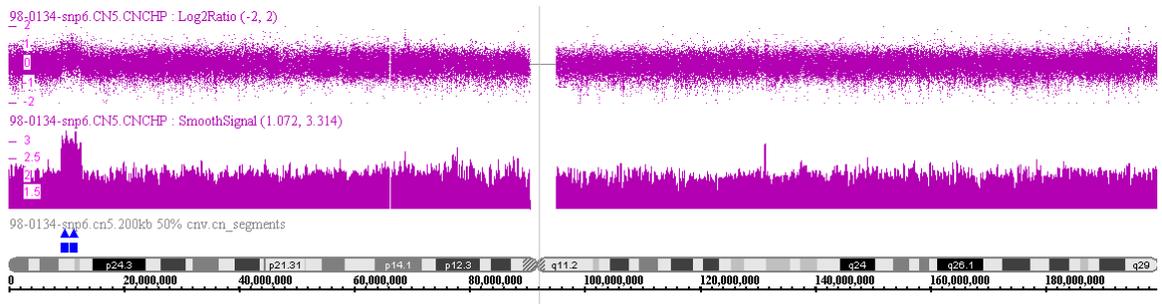


Figure 4.8 DNA copy number gain at 3p25 in PA30.

The upper trace shows raw \log_2 ratios, the lower trace shows smoothed copy number data. The gained region identified by the segmentation tool, is denoted by a blue bar and arrow.

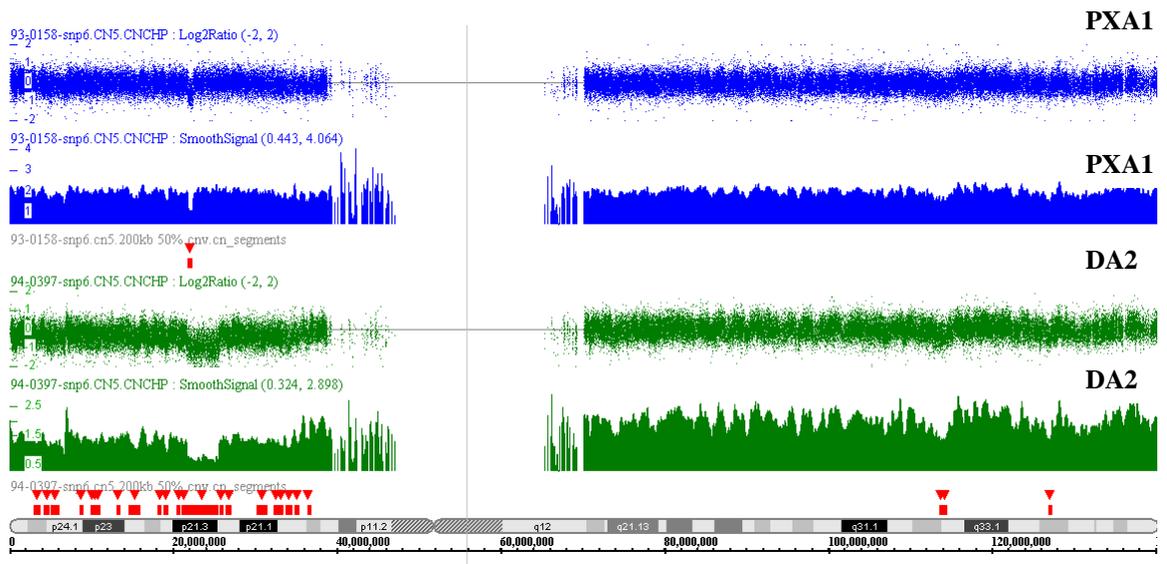


Figure 4.9 DNA copy number loss at 9p21 in samples PXA1 and DA2.

The upper traces show raw \log_2 ratios, the lower traces smoothed copy number data and regions of copy number loss identified by segmentation (red bars and arrows).

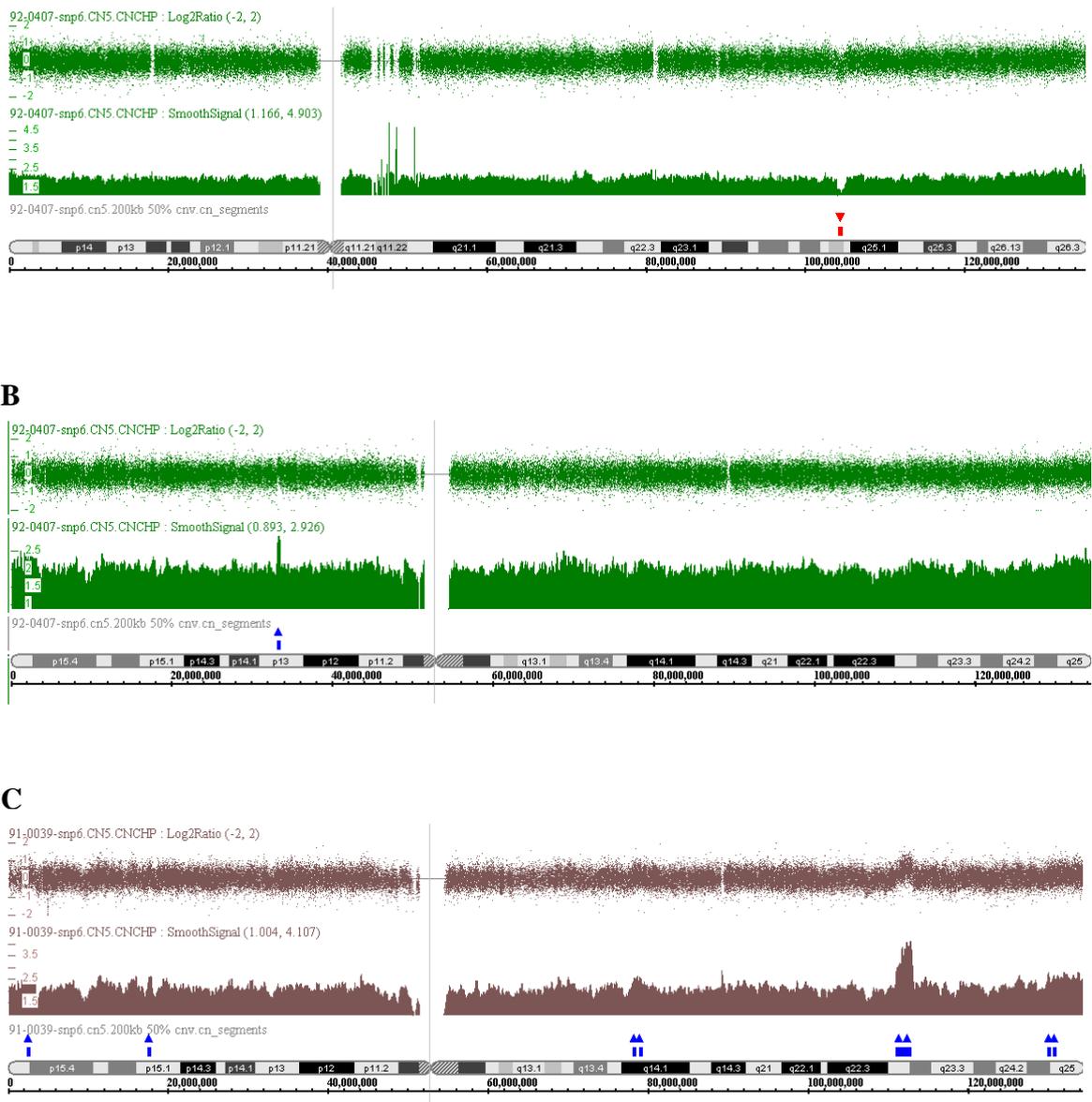


Figure 4.10 DNA copy number changes by SNP 6.0 analysis.

Raw log₂ ratios and smoothed copy number data are shown. Regions of copy number gain identified by segmentation are shown as blue bars and arrows and regions of loss as red bars and arrows.

(A) DNA copy number loss at 10q24.32 in sample DA3.

(B) DNA copy number gain at 11p13 in sample DA3.

(C) DNA copy number gain at 11q23.1 in sample PA2.

Name	Site	Age (yrs)	Sex	3p25 gain	7q34 gain	9p21 loss	11q32 gain	Whole chromosome gains
PA1	Cerebellum	20	F	-	+	-	-	5,6
PA2	Cerebellum	4	M	-	+	-	+	-
PA3	Cerebellum	7	M	-	+	-	-	-
PA4	Cerebellum	2	M	-	+	-	-	-
PA5	Cerebellum	6	F	-	+	-	-	-
PA6	Cerebellum	5	M	-	+	-	-	-
PA9	Cerebellum	14	F	-	+	-	-	5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 17, 18, 20
PA10	Cerebellum	19	M	-	+	-	-	5, 6, 7, 14
PA11	Cerebellum	14	M	-	+	-	-	5, 6
PA12	Cerebellum	14	M	-	+	-	-	-
PA13	Cerebellum	10	M	-	+	-	-	-
PA14	Cerebellum	10	M	-	+	-	-	-
PA15	Cerebellum	13	M	-	+	-	-	-
PA16	Cerebellum	10	M	-	-	-	-	-
PA17	Cerebellum	6	M	-	+	-	-	-
PA18	Cerebellum	3	M	-	+	-	-	-
PA19	Cerebellum	3	M	-	+	-	-	-
PA22	Cerebellum	6	F	-	+	-	-	-
PA23	Cerebellum	3	F	-	+	-	-	-
PA25	Cerebellum	4	F	-	+	-	-	-
PA26	Cerebellum	3	F	-	+	-	-	-
PA27	Cerebellum	14	F	-	+	-	-	-
PA29	Cerebellum	19	M	-	+	-	-	5, 6, 11, 13, 15, 20
PA30	Cerebellum	13	F	+	-	-	-	-
PA31	Cerebellum	1	F	-	-	-	-	-
PA32	Brain stem	9	F	-	+	-	-	-

Table 4.4 DNA copy number changes in 26 pilocytic astrocytomas by 6.0 SNP array. (- = no abnormality found).

Name	Pathology	Site	Age (years)	Sex	3p25 gain	7q34 gain	9p21 loss	10q24.32 loss	11p13 gain	Whole chromosome gains
PXA1	Pleomorphic xanthoastrocytoma	Cerebral cortex	6	F	-	-	+	-	-	7
DA2	Diffuse astrocytoma	Cerebral cortex	9	M	-	-	+	-	-	-
DA3	Diffuse astrocytoma	Cerebral cortex	17	M	-	-	-	+	+	5, 6, 7, 15, 20, 21
DA4	Diffuse astrocytoma	Cerebral cortex	4	M	-	-	-	-	-	-
DA5	Diffuse astrocytoma	Cerebral cortex	7	F	-	-	-	-	-	-
DA6	Diffuse astrocytoma	Cerebral cortex	5	F	-	-	-	-	-	-
DA7	Diffuse astrocytoma	Cerebral cortex	5	M	-	-	-	-	-	-
PMA1	Pilomyxoid astrocytoma	Diencephalon	1	F	-	+	-	-	-	-
PMG1	Pilomyxoid glioma	Diencephalon	17	M	-	-	-	-	-	6, 8, 11, 12
PMG2	Pilomyxoid glioma	Diencephalon	7	F	-	-	-	-	-	-
PMG5	Pilomyxoid glioma	Spinal cord	11	F	-	-	-	-	-	6, 7, 8, 9, 11, 12

Table 4.5 DNA copy number changes in 11 grade II astrocytomas by 6.0 SNP array. (- = no abnormality found).

4.4.3 Comparison of results from SNP arrays and array CGH

The results obtained by Affymetrix 250k or 6.0 SNP arrays were compared to data obtained by array CGH (Table 4.6). There was some concordance between array CGH and SNP array data for 3 samples where no whole chromosome gains or losses were identified, and between 4 tumour samples with whole chromosome gains. Whole copy number gains and losses were identified by array CGH in a further 6 cases, but these were not subsequently found by SNP array analysis. In general, the SNP array data did not identify any whole chromosome losses.

The region of gain at 7q34 was not identified in cerebellar pilocytic astrocytoma samples by array CGH analysis (Figure 4.11). I investigated the density of clones on the array slides covering the 7q34 region of interest, to see where probes in this region were situated.

Three clones were present on the array CGH slides in the region of interest at 7q34, spanning a region of ~ 2.6Mb. The first clone in the region was RP11-269N18 (101980 bps), situated at position 138341354-138443333, around 70Kb downstream towards the centromere from the 3' end of the gene *KIAA1549* (Figure 4.12 (A)). The second clone present in the region was RP5-886O8 (143569 bps), situated at 139539734-139683302, within the gene *TBXAS1* (Figure 4.12 (B)). The final probe in the region of interest was RP5-1173P7 (129223 bps). This was situated ~100Kb upstream from the start of *BRAF*, the next gene of interest, at position 140,080,751-140,271,033 (Figure 4.12(C)). The two clones RP11-269N18 and RP5-1173P7 were both outside the region of gain, however RP5-886O8 was situated within the centre of the gained region (Figure 4.13). This probe would have been expected to identify the copy number gain present in the cerebellar samples. None of the cerebellar pilocytic astrocytomas showed any significant region of gain at clone RP5-886O8 (at position 138947011 on chromosome 7 in raw and smoothed wig data files).

These data are available on the Cancer Research UK London Research Institute's bioinformatics and biostatistics website at the following link:

<https://bioinformatics.cancerresearchuk.org/~kelly02/index.php?id=Ruth>

Sample	Site	Age	Sex	aCGH copy number gains / losses	SNP array copy number gains/ losses
PA1	Cerebellum	20	F	Gain 5, 6	Gain 5, 6, 7q34 gain
PA3	Cerebellum	7	M	-	7q34 gain
PA11	Cerebellum	14	M	Gain 5, 6	Gain 5, 6, 7q34 gain
PA13	Cerebellum	10	M	7q22.1 loss	7q34 gain
PA14	Cerebellum	10	M	Loss 17	7q34 gain
PA15	Cerebellum	13	M	-	7q34 gain
PA17	Cerebellum	6	M	-	7q34 gain
PA18	Cerebellum	3	M	2p22.1 loss	7q34 gain
PA20	Cerebellum	4	M	Loss 22	7q34 gain
PA23	Cerebellum	3	F	1q42.12 loss Gain 4p15.31	7q34 gain
PA32	Brain stem	9	F	Gain 13, loss terminal end 1p	7q34 gain
PMA1	Diencephalon	1	F	Gain 5q33.3, 8q11.23	7q34 gain
DA2	Cerebrum	9	M	Loss 17 Loss terminal end 1p	-
DA3	Cerebrum	17	M	Gain 6, 7, 21 Gain 5q32	Gain 5, 6, 7, 15, 20, 21
DA4	Cerebrum	4	M	Loss 17, 19, 20, 22 Loss terminal end 1p, 7 cen-7q11.23, 7q22.1 and 7q36.1-tel	-
DA7	Cerebrum	5	M	Gain 2q11.2, 17p12	-
DA9	Cerebrum	9	F	Gain 2q11.2	-
PMG1	Diencephalon	17	M	Gain 6, 8, 11, 12	Gain 6, 8, 11, 12
PMG2	Diencephalon	7	F	Loss terminal end 1p	-
PMG5	Spinal cord	11	F	Gain chromosome 9	Gain 6, 7, 8, 9, 11, 12

Table 4.6 Comparison of whole chromosome gains from 20 tumour samples (11 cerebellar PA, 9 non-cerebellar PA) by array CGH and 250k or 6.0 SNP array analysis.

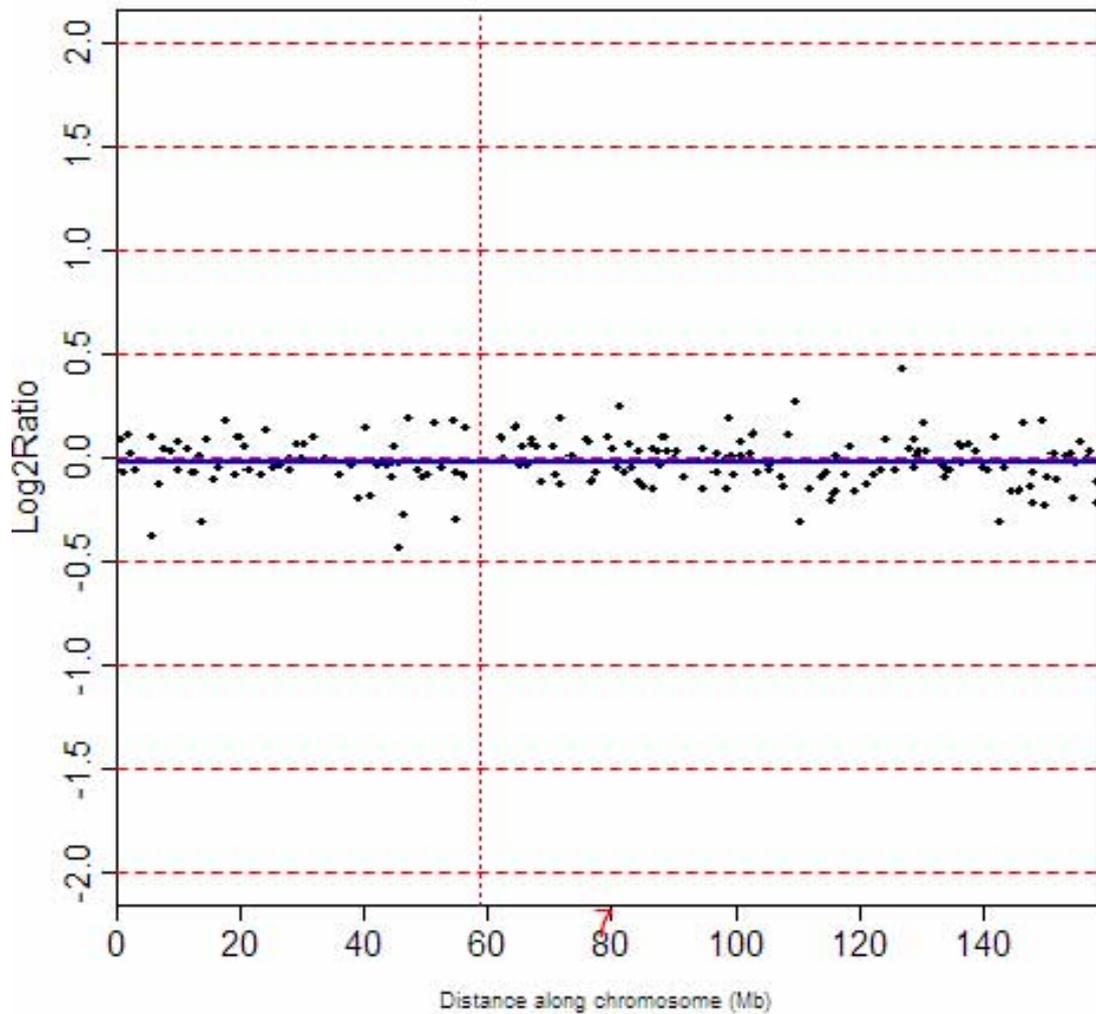


Figure 4.11 Array CGH plot for chromosome 7 comparing pooled data from cerebellar samples to non-cerebellar samples.

No copy number differences were identified between the two groups on chromosome 7. Clone RP5-886O8 (within the region of gain at 7q34) was situated at ~ 139.5Mb, and was not gained in any of the pilocytic astrocytoma samples.

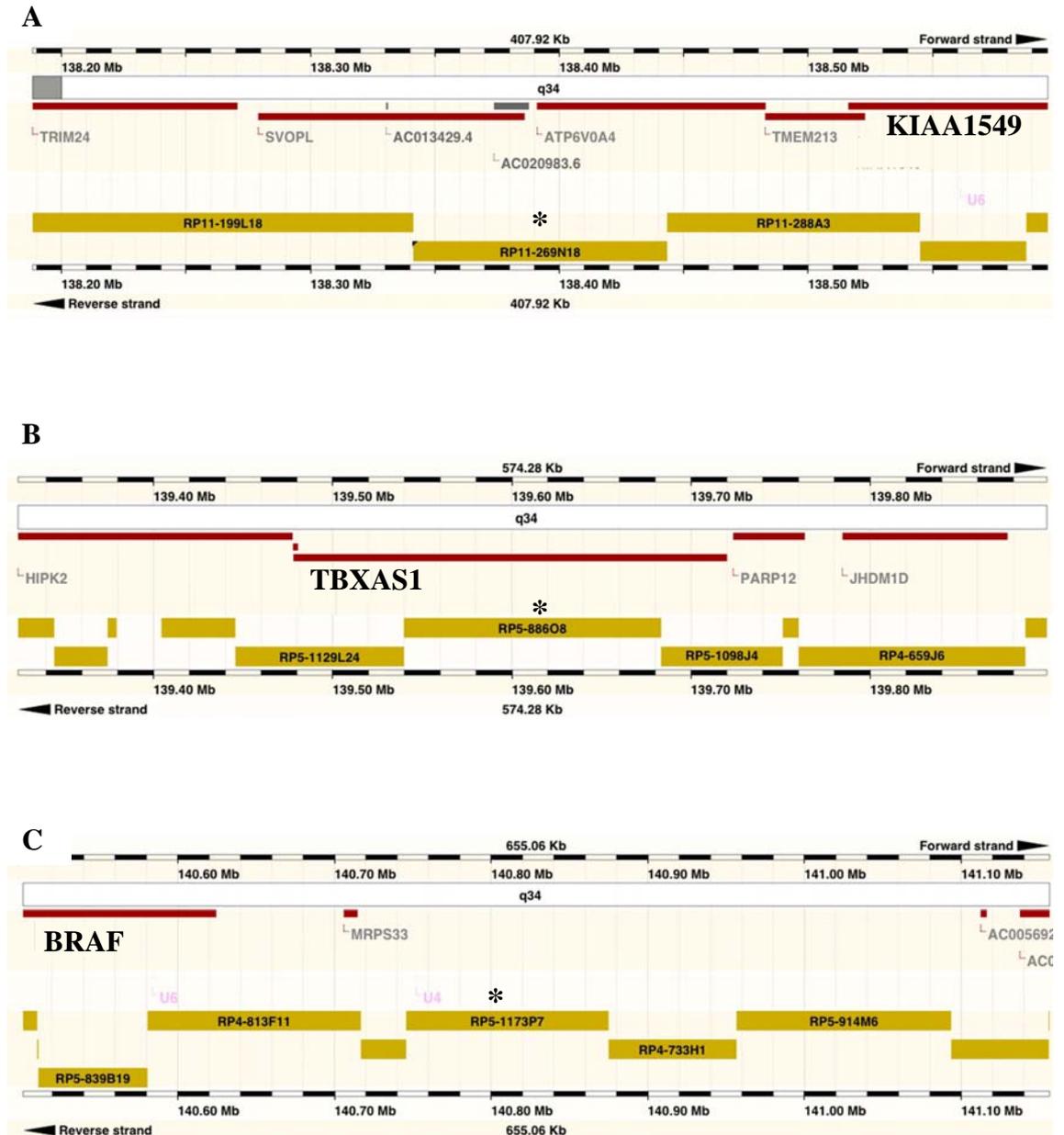


Figure 4.12 Array CGH clones at the 7q34 region of gain. Figures adapted from Ensembl, release 55.

- (A) Clone RP11-269N18 (highlighted with an asterisk) is situated ~ 70Kb downstream from *KIAA1549* towards the centromere.
- (B) Clone RP5-886O8, situated in the centre of the gained region, within *TBXAS1*.
- (C) Clone RP5-1173P7, situated ~ 100Kb upstream from *BRAF*.

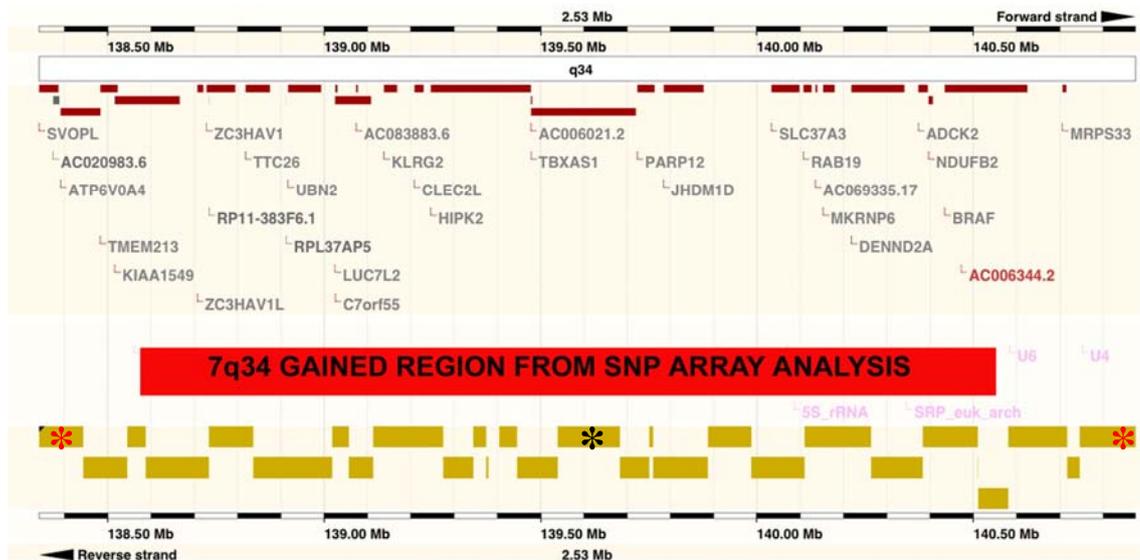


Figure 4.13 Overview of the region of interest within 7q34.

The gained region, identified by SNP array analysis is overlaid by a red bar. Array CGH clones within the region are marked by asterisks. The clones marked in red lie outside the region of gain (RP11-269N18, RP5-1173P7), the clone marked in black (RP5-88608) is the single clone present within the gained region. Figure adapted from Ensembl, release 55.

4.5 Discussion

4.5.1 SNP array investigation of astrocytomas

SNP array studies have begun to yield a wealth of detailed information on the biology of cancer. Two comprehensive studies have recently been published which yield new insights into the genetic aberrations involved in the development of adult glioblastoma (GBM) (CGARN *et al.*, 2008, Parsons *et al.*, 2008). These studies utilised SNP data to investigate frequent copy number changes seen in GBM, in combination with mutation and methylation analysis. Both of these studies revealed that deregulation of the RB, TP53 and RAS/PI(3)K pathways are obligatory events in the majority of adult GBM, affecting much larger proportions of both patients and genes within each pathway than had previously been suspected. New genes and signalling pathways not previously known to be implicated in the development of GBMs were also identified, for example mutations in the isocitrate dehydrogenase genes *IDH1* and *IDH2* were found to occur

early in the development of adult grade II-III astrocytomas which progress to become secondary GBMs (Parsons *et al.*, 2008).

Few studies to date have used SNP arrays to investigate paediatric astrocytomas. Wong *et al.* published the first study in 2006, using Affymetrix 10K SNP arrays (Wong *et al.*, 2006). Fourteen low-grade astrocytomas (6 pilocytic astrocytomas, 5 gangliogliomas, 2 subependymal giant cell astrocytomas (SEGA) and one unspecified low-grade astrocytoma) and 14 high-grade astrocytomas (13 GBMs, 1 anaplastic astrocytoma) were studied. No regions of LOH were detected in the pilocytic astrocytomas, and the majority of low-grade astrocytomas showed either no detectable LOH, or LOH affecting a single chromosome arm. Two gangliogliomas showed LOH of a single chromosome arm (3q and 9p), the unspecified astrocytoma had LOH at 6q and the two SEGAs showed LOH of 16p and 21q, respectively. In contrast the high-grade astrocytomas studied showed highly complex patterns of LOH involving multiple different chromosomes. Amplification of *EGFR* and *PDGFR* were found to be relatively uncommon in paediatric high-grade astrocytoma; each was seen in 2/13 GBMs, and amplification was confirmed by real-time PCR. This study was, however, performed using one of the earliest SNP array platforms, which was still at a fairly low resolution.

A later study by Deshmukh *et al.* compared array CGH and SNP array analysis for 10 paediatric patients with pilocytic astrocytomas (PAs), using matched blood and tumour samples (Deshmukh *et al.*, 2008). Here the high-resolution 385K NimbleGen oligonucleotide array and the Affymetrix Mapping 500K SNP array sets were used. One of the 10 tumour samples showed whole chromosome gains of 6 and 21, identified by both array platforms, the remaining 9 PAs showed no whole chromosome gains or losses. Multiple focal copy number changes were seen in individual samples, but no recurrent regions of gain or loss were identified by visual inspection of the data. However, segmental analysis revealed a single recurrent region of copy number gain on chromosome 7q34, between 138,151,200-139,456,000 (NCBI build 35). This was seen in 8/10 PAs by SNP array and in 6/10 PAs by array CGH. The common minimal region of gain between the two data sets was localised in the upstream region of homeodomain-interacting protein kinase 2 (*HIPK2*). Gain of *HIPK2* was validated by quantitative PCR using matched tumour and peripheral blood DNA in 6/10 PAs. In a separate cohort of 61 PAs 42% (26/61) showed amplification of *HIPK2*. The gained

region identified here also included partial gain of *BRAF* in three PAs, and partial gain of *KIAA1549* in two PAs.

The 7q34 region of gain in pilocytic astrocytomas was also identified by array CGH in two studies published last year (Jones *et al.*, 2008, Pfister *et al.*, 2008). Pfister *et al.* found a 0.97Mb region of gain at 7q34 involving 2 BAC clones (RP5-886O8 and RP4-726N20) in 30/66 low-grade paediatric astrocytomas (53 PAs, 13 diffuse astrocytomas). Similarly, Jones *et al.* found a recurrent 2Mb region of gain by array CGH at 7q34 and followed this up using a chromosome 7 tiling path array, where 7q34 gains between clones RP11-355D18 (within *KIAA1549*) and RP11-543P6 (within *BRAF*) were seen in 29/44 pilocytic astrocytomas. These, and further studies which identified 7q34 gain in pilocytic astrocytomas will be discussed in more detail in chapter 5.

These data confirm the most striking change found by SNP array analysis in this study: a discrete ~1.9Mb region of gain situated at chromosome band 7q34 in 24/37 samples. Twenty-three of 26 cerebellar pilocytic astrocytomas and the single hypothalamic pilomyxoid astrocytoma contained 7q34 gains. *HIPK2* was contained within this region of gain, as found by Deshmukh *et al.* (Deshmukh *et al.*, 2008). However, further analysis by segmentation revealed that this gained region had highly recurrent start and end points, partially contained within the genes *BRAF* and *KIAA1549*. *BRAF* is a partner in two different fusion genes, with *AKAP9* in radiation induced thyroid cancer and with *FCHSD1* in large melanocytic naevi (Dessars *et al.*, 2007, Ciampi *et al.*, 2005). In both cases only the 3' portion of *BRAF*, which encodes the active kinase domain, were included in the gene fusion. These findings led to speculation that *BRAF* might be involved in a gene fusion with *KIAA1549* in pilocytic astrocytomas. I will discuss this further in chapter 5.

Array CGH, as described in chapter 3, did not identify this region of gain at 7q34 in the pilocytic astrocytoma samples. This was, perhaps, understandable as there was only a single BAC clone situated in the region of gain on the array slides used. Furthermore, none of the potential regions of interest identified by array CGH (1p loss, 2q11.1 gain, 7q22.1 loss and 21q22.2 gain) were subsequently identified by either 250K or 6.0 array analyses. Overall the quality and resolution of the array CGH data appeared noisy, and few significant regions of copy number change were identified.

Whole chromosome gains and losses were, however, identified by array CGH, and some of these were subsequently confirmed by SNP array analysis. The array CGH and SNP array results concurred in 8 cases (Table 4). Three samples (PA3, PA15, PA17) contained no whole chromosome gains or losses (although all contained 7q34 gains by SNP analysis). Five cases contained whole chromosome gains; the majority of the results for whole chromosome gains were the same between the two platforms. The 250K and 6.0 SNP arrays showed whole chromosome gains in 9/37 patients.

The most commonly affected chromosomes were: 6 (8 cases), 5 (6 cases), 7 (5 cases), 11 (4 cases); 8, 12, 15 and 20 (3 cases each); 9 and 14 (2 cases each). The mean age of patients with whole chromosome gains was 15.2 years, compared to a mean age of 7 years for patients with no chromosome gains. A single young patient aged 6 years, showed trisomy 7. The cases aged >10 years had an average of 5 chromosomes gained per patient. These data confirm findings from the previous chapter, and from the literature that whole chromosome gains are more prevalent in older patients with pilocytic astrocytoma, aged > 10 years (Jones *et al.*, 2006).

In general, the low-grade astrocytomas appeared to display stable genomes, with few discrete regions of copy number gain and loss. This observation is in line with previous analyses by conventional cytogenetics and array CGH, and also fits with the natural history of low-grade astrocytomas. Pilocytic astrocytomas are slow-growing tumours, with extremely good 5- and 10-year survival data following surgical resection (Burkhard *et al.*, 2003, Ohgaki *et al.*, 2005a). Paediatric glioblastomas, on the other hand, are highly aggressive tumours, and display hugely rearranged karyotypes, in keeping with the general observation that high-grade malignancies appear to accumulate large numbers of genomic aberrations (reviewed by Stratton, 2009) (Stratton *et al.*, 2009).

4.5.2 Further discrete changes identified by SNP array analysis

SNP array analysis enabled the detection of regions of copy number change at much higher resolution than array CGH, with considerably less experimental noise. Four further regions of copy number change were identified in low-grade astrocytomas, in addition to the gains at 7q34, although these were found at significantly lower frequencies.

4.5.2.1 Gain of 3p25 in sample PA30

The pilocytic astrocytoma PA30, from a 13 year old female, was found to contain a discrete ~3Mb region of copy number gain at chromosome band 3p25, within parts of the genes *SRGAP3* and *RAF1* (Figure 4. 8). No other copy number changes were found, and there was no gain at 7q34. Further segmental analysis identified the start of the gained region within *SRGAP3*, however, the region of gain appeared to end within *SYN2*, ~ 500Kb upstream from *RAF1*. Analysis by CBS was not possible for this sample.

RAF1 was an attractive candidate gene, particularly in light of the similar partial gains within *BRAF* seen in the majority of pilocytic astrocytomas. Both *BRAF* and *RAF1* are serine-threonine-specific protein kinases within the mitogen activated protein kinase pathway (MAPK pathway), activated downstream from RAS. Deregulation of the MAPK pathway has recently been identified as highly significant in low-grade astrocytomas, and this will be discussed in detail in chapter 5 (Sharma *et al.*, 2005, Pfister *et al.*, 2008, Jones *et al.*, 2008, Forsheew *et al.*, 2009, Jones *et al.*, 2009, Sievert *et al.*, 2009).

4.5.2.2 Loss of 9p21 in DA2 and PXA1

Two grade II astrocytoma samples DA2 and PXA1, showed discrete copy number loss at chromosome band 9p21, containing the gene *CDKN2A* (cyclin-dependent kinase inhibitor 2A) (Figure 4.9). DA2 showed homozygous loss of *CDKN2A*, with general

heterozygous loss of the whole short arm of chromosome 9. PXA1 revealed heterozygous copy number loss of *CDKN2A* at 9p21 and single copy gain encompassing the whole of chromosome 7. For PXA1, analysis using the segmentation tool within Affymetrix Genotyping Console 3.0.2 and further analysis by segmentation and CBS showed exact agreement on the boundaries of the region of loss.

CDKN2A is a known tumour suppressor gene, and encodes several transcript variants, which have different first exons. *CDKN2A* is deleted or mutated in a wide range of cancers, including leukaemias and solid tumours (Kohno *et al.*, 2006). Homozygous loss of *CDKN2A* is a common finding in adult glioblastomas: two recent studies reported *CDKN2A* loss in ~ 50% adult GBMs, with overall alterations in the RB pathway in 77% (CGARN *et al.*, 2008, Parsons *et al.*, 2008). Only one of the samples identified by SNP array analysis here had homozygous loss of *CDKN2A*. Hence, sequencing was undertaken in 50 low-grade astrocytomas to search for additional inactivating mutations in *CDKN2A*. These results will be described in further detail in chapter 6.

4.5.2.3 Loss of 10q24.32 in sample DA3

One diffuse astrocytoma in a 17 year-old male, DA3, contained a discrete region of loss at 10q24.32 containing, among others, the genes *SUFU* (suppressor of fused homolog Drosophila) and *TRIM8* (tripartite motif-containing 8)(Figure 4.10(A)). *SUFU* is a negative regulator of the Sonic Hedgehog (SHH) signalling pathway: deletions or inactivating mutations in *SUFU* are known to be associated with the development of desmoplastic medulloblastomas, through overactivity of the SHH signalling pathway (Rubin *et al.*, 2002, Taylor *et al.*, 2002, Pasca di Magliano *et al.*, 2003, Gilbertson *et al.*, 2008). *TRIM8* is a putative tumour suppressor gene; the protein it encodes has a similar structure to other tumour suppressor proteins, and the 10q24.32 genomic region is known to be a common site for deletions and rearrangements in glioblastoma, suggesting that the region may be involved in tumorigenesis (Toniato *et al.*, 2002, Carinci *et al.*, 2007). SNP copy number analysis indicated heterozygous loss at 10q24.32. Hence, further follow-up of the genes in this region will be required, to search for a possible inactivating mutation in the remaining allele of *SUFU* or *TRIM8*.

The tumour from this patient contained a further discrete copy number change discussed below, and numerous additional whole chromosome gains of 5, 6, 7, 15, 20 and 21.

4.5.2.4 Gain of 11p13 in sample DA3

Sample DA3 also contained a ~300Kb region of gain at chromosome 11p13, which contains the gene *HIPK3* (homeodomain interacting protein kinase 3) (Figure 4.10(B)). This region of gain was identified by segmental analysis and CBS. *HIPK3* is a serine/threonine kinase/homeodomain-interacting protein kinase, which modulates the Fas signalling pathway. Increased activation of the mitogen-activated protein kinase JNK, leading to increased levels of *HIPK3* expression has been shown to promote resistance to apoptosis in prostate cancer cell lines (Curtin *et al.*, 2004). Sensitivity to apoptosis in the prostate cancer cells was restored in this study using RNA interference directed against *HIPK3* and a specific JNK inhibitor (Curtin *et al.*, 2004). *HIPK3* does not currently have any known association with brain tumours, but is widely expressed in mammalian tissues including the brain and spinal cord (Rochat-Steiner *et al.*, 2000). As increased levels of *HIPK3* expression were found to promote resistance to apoptosis, and sample DA3 shows genomic gain at this region it is possible that this might play a role in promoting tumour growth. As previously mentioned, this sample contained numerous whole chromosome gains in addition to the two discrete copy number changes described.

4.5.2.5 Gain of 11q23.1 in sample PA2

Finally, one pilocytic astrocytoma sample in a 4-year old male, PA2, contained a discrete ~2Mb region of gain at 11q23.1, in addition to gain at 7q34 (Figure 4.10(C)). This contained, among others, the gene *NCAM1* (neural cell adhesion molecule 1). *NCAM1* is expressed on the surface of the majority of neural cells, and has been shown to play a major role in the development of the nervous system (reviewed by Ditlevsen, 2008) (Ditlevsen *et al.*, 2008). In the adult brain *NCAM1* is associated with neuronal plasticity, which is part of the process for learning and consolidation of memory. *NCAM1* binds neuronal cells together, but also has a significant function in coordinating intracellular signalling, resulting from stimuli on or in the vicinity of the cell surface. Stimulation of *NCAM1* has also been shown to lead to activation of the

MAPK pathway via ERK1/2 phosphorylation, which may ultimately lead to neurite outgrowth, although the method of MAPK pathway activation is currently unclear. It may be that stimulation of NCAM1 leads to activation of *RAS* via the cytoplasmic protein tyrosine kinases Fyn and FAK, but this remains to be determined. Initiation of neurite outgrowth is known to be coordinated via a highly complex web of signalling pathways, so activation of a single pathway may not necessarily induce neurite growth. NCAM1 and a related neural cell adhesion molecule L1 have also been shown to promote neuronal survival, so it is possible that a survival advantage may be conferred by increased expression of *NCAM1* (Ditlevsen *et al.*, 2008). Further investigations using quantitative PCR for *NCAM1* in sample PA2 will be required to confirm whether increased levels of expression are found.

4.5.2.6 Astrocytomas with no copy number changes identified by SNP array analysis

One pilocytic astrocytoma (PA31) and seven grade II astrocytomas (DA4, DA5, DA6, DA7, DA8, DA9 and PMG2) had no apparent copy number changes found by analysis of the SNP data. Other mechanisms may underlie tumorigenesis in this group, and further investigations will be required to investigate whether activating mutations, epigenetic changes or possible aberrations in microRNA expression are present in these samples

4.6 Summary

SNP array analysis identified a striking copy number change in pilocytic astrocytomas, at a frequency that was both unprecedented and completely unexpected, in tumours that were previously thought to contain few or no genetic rearrangements. A discrete ~1.9Mb region of gain was seen at 7q34 in 23/26 cerebellar pilocytic astrocytomas and in the single hypothalamic pilomyxoid astrocytoma. Additional discrete copy number changes were also identified in further samples, although these were present at much lower frequencies.

Two separate statistical analyses were performed concurrently with the Affymetrix 6.0 SNP data: each was entirely independent of the other. Analysis using the Affymetrix Genotyping console segmentation tool, and further segmental analysis were performed at our laboratory in London: CBS was performed by Dr Jing Ma at St Jude Children's Research Hospital in Memphis. Both analysis methods identified the region of gain at 7q34 and the additional discrete regions of gain and loss described. The 3p25 gain in PA30 was the only discrete change found using a single method, as CBS was not possible for this sample. All of the methods for analysing SNP data gave essentially similar results. Although there were subtle differences in the exact boundaries for segments of loss or gain, the majority of genes contained within these regions were the same.

However, these data highlight the importance of verifying copy number changes by different techniques. Additional methods of ascertaining DNA copy number changes include fluorescence *in situ* hybridisation (FISH), quantitative real-time PCR and multiplex ligation-dependent probe amplification (MLPA). FISH was used to investigate the region of 7q34 gain in paraffin tissue sections from pilocytic astrocytomas, and these results will be discussed further in chapter 5. The additional regions of copy number change identified by SNP array analysis require further validation, and I will continue with this work following the completion of my PhD.

Chapter 5. Identification and characterisation of fusion genes

5.1 Introduction

A discrete region of DNA copy number gain was identified by Affymetrix 250K and 6.0 SNP arrays, at chromosome band 7q34, in 23/26 pilocytic astrocytomas and the single pilomyxoid astrocytoma. The gained region showed recurrent start and end points, partially contained within the genes *BRAF* and *KIAA1549*. This suggested that a duplication event had occurred across this genomic region, including parts of *BRAF* and *KIAA1549*. *BRAF* has been shown to be involved in chromosomal rearrangements leading to the creation of two different fusion genes in thyroid carcinoma and large melanocytic naevi (Ciampi *et al.*, 2005, Dessars *et al.*, 2007). In both of these cases only part of *BRAF* is included in the fusion. Both result in constitutive activation of *BRAF*, through loss of the N-terminal auto-inhibitory domain of *BRAF*, however this is achieved via different genomic rearrangements.

The *AKAP9-BRAF* fusion was identified in radiation-induced papillary thyroid carcinoma (Ciampi *et al.*, 2005). FISH demonstrated duplication of the *BRAF* locus, in the presence of two copies of chromosome 7, and subsequent analysis revealed a gene fusion between exons 1-8 of A-kinase anchor protein 9 (*AKAP9*) and exons 9-18 of *BRAF*. Both genes are situated on chromosome 7: *AKAP9* at 7q21-q22 and *BRAF* at 7q34. The gene fusion results from paracentric inversion of the q arm of chromosome 7, creating an in-frame *AKAP9-BRAF* fusion. The 5' region encoding the auto-inhibitory N-terminal Ras-binding domain of *BRAF* is lost. The C-terminal protein kinase domain of *BRAF* is retained in the *AKAP9-BRAF* fusion, which was shown to display elevated kinase activity and induced transformation of NIH3T3 cells (Ciampi *et al.*, 2005).

A further fusion between *BRAF* and *FCHSD1* has been identified in large melanocytic naevi (Dessars *et al.*, 2007). This arises by a different mechanism, from a translocation between *FCHSD1* at chromosome 5q31 and *BRAF* at 7q34. Here the fusion junction lies between exon 13 of *FCHSD1* and exon 9 of *BRAF*. Again, the auto-inhibitory N-

terminal Ras-binding domain of *BRAF* was lost from the *FCHSD1-BRAF* fusion transcript, and the protein kinase domain was retained.

The SNP copy number analyses of low-grade astrocytomas described in chapter 4 of this thesis suggested that the region of 7q34 copy number gain did not include the 5' end of *BRAF* or the 3' end of *KIAA1549*. This led to the hypothesis that *KIAA1549* and *BRAF* were involved in a gene fusion in the pilocytic astrocytomas.

5.2 Aim

The aim of the work described within this chapter was to investigate whether fusion genes were present in paediatric low-grade astrocytomas, using the following techniques:

- 1) PCR from tumour cDNA and DNA
- 2) DNA sequencing

5.3 Interphase fluorescence in-situ hybridisation (i FISH)

Dual-colour interphase FISH was performed by James Dalton, at St Jude Children's Research Hospital, Memphis, to confirm whether the 7q34 region of gain included a duplicated region. The iFISH method is described in detail in section 2.11. Tumours containing 7q34 gain were selected for interphase FISH analysis, using 5-8µm paraffin tissue sections. Probes were derived from BAC clones (Invitrogen, Carlsbad, CA) and labelled with either FITC or rhodamine fluorochromes. BAC clone RP11-837G3 was selected to target exons 11-18 of *BRAF*, and labelled with rhodamine. BAC clones RP11-251I15 + RP11-746C13 were selected as control probes for chromosome arm 7p, and labelled with fluorescein isothiocyanate (FITC).

Interphase FISH indicated the presence of a tandem duplication at 7q34, in the presence of two homologues of chromosome 7 (Figure 5.1). This duplicated region included exons 11-18 of the *BRAF* locus, encoding the C-terminal protein kinase domain. These results confirmed the 7q34 copy number gain identified by SNP analysis, and led to further experiments designed to identify a possible fusion gene between *KIAA1549* and *BRAF*.



Figure 5.1 Tandem duplication within 7q34 including *BRAF*, identified by iFISH from paraffin sections.

BAC clone RP11-837G3 was used to target exons 11-18 of *BRAF* (red) in combination with control probes to 7p (BAC clones RP11-251I15 + RP11-746C13, green). Images were selected from nuclei in tumour PA1 and showed signals from both homologues of chromosome 7 (green) and 3 signals from *BRAF* (red), indicating duplication of the *BRAF* region on one homologue. Many nuclei showed an incomplete set of FISH signals as these images were from FFPE sections. All iFISH was performed at St Jude Children's Research Hospital, Memphis.

5.3.1 PCR identifies *KIAA1549-BRAF* fusion

Oligonucleotide primers were designed to identify any possible gene fusion between *BRAF* and *KIAA1549*. Twelve forward primers were designed stepping along *KIAA1549* at approximately 500bp intervals (Appendix C, Table 8.2). A single reverse primer was designed within the kinase domain of *BRAF* in exon 12, based on the presumption that the kinase domain would be retained within any possible fusion gene (Figure 5.2).

PCRs were performed to identify possible fusion gene products in cDNA from two tumours with 7q34 copy number gain (PA7, PA29) and cDNA prepared from control whole brain total RNA. PCR products were visualised using 1.5% agarose gels. Both tumours revealed PCR products, indicating the presence of a *KIAA1549-BRAF* gene fusion. PCR products were seen in sample PA7 between *BRAF* exon 12 and *KIAA1549* primers 9 (exon 12), 10 (exons 14/15) and 11 (exon 16). These three bands stepped down in size, consistent with the primer positions moving along *KIAA1549* in 500bp increments (Figure 5.3). Sample PA29 showed PCR products between *BRAF* exon 12 and *KIAA1549* primers 10 (exons 14/15) and 11 (exon 16). No PCR products were identified in normal brain cDNA.

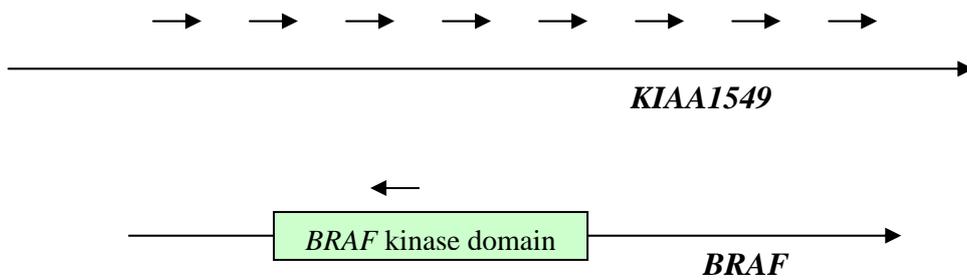


Figure 5.2. Schematic diagram showing the design of primer placement for the identification of a possible fusion gene.

Forward primers are situated at ~ 500 bp intervals along *KIAA1549*. The single reverse primer was in exon 12 of *BRAF* within the kinase domain. Both genes are transcribed in the same direction. The exact locations and sequences of the primers are shown in Appendix C Table 8.2.

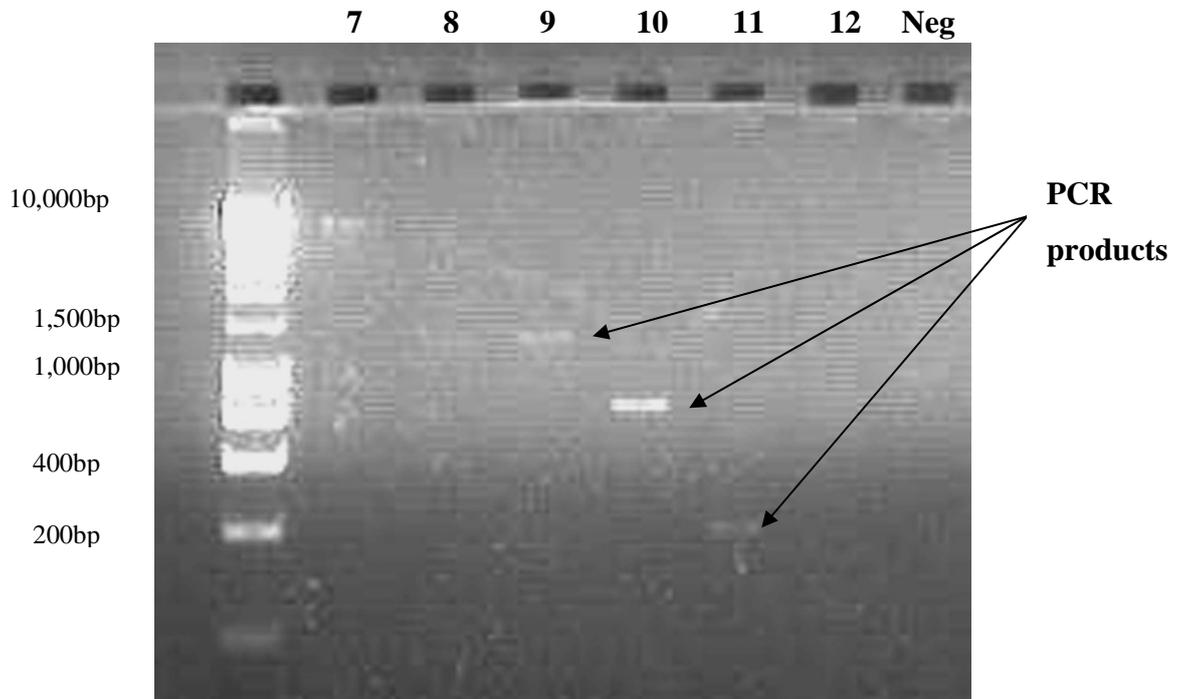


Figure 5.3. 1.5% agarose gel showing results of PCR between *KIAA1549* and *BRAF* in sample PA7.

The PCR products indicate the presence of a *KIAA1549- BRAF* gene fusion. From left to right the wells contain: DNA ladder (Hyperladder I; 200 base pairs (bp) - 10kb, Bioline), *KIAA1549* forward primers 7-12 and negative control (water). All PCRs used the same reverse primer within *BRAF* exon 12. A non-specific product was seen for *KIAA1549* forward primer 7. Product sizes are 1300, 800 and 300 base pairs, in descending order.

5.3.2 Sequencing confirms the presence of a *KIAA1549-BRAF* gene fusion

Direct sequencing was performed for the PCR products identified from samples PA7 and PA29. These confirmed the presence of a fusion gene product between *KIAA1549* exon 16 and *BRAF* exon 9 in sample PA7 (Figure 5.4.A). A different fusion product was identified in sample PA29, between *KIAA1549* exon 19 and *BRAF* exon 9 (Figure 5.4.B). These data were reproducible, and were repeated on cDNA from separate reverse transcription reactions.



Figure 5.4. Sequence analysis of two *KIAA1549-BRAF* fusions in tumour cDNA.

(A) Fusion junction between *KIAA1549* exon 16 and *BRAF* exon 9 in PA7.

(B) Fusion junction between *KIAA1549* exon 19 and *BRAF* exon 9 in PA29.

5.3.3 Nested PCR identifies three further *KIAA1549-BRAF* fusions

Primers were designed for a nested PCR, to increase the specificity of the reaction to detect fusions. Previously a few tumour samples had produced weak PCR products with *KIAA1549* and *BRAF* primer combinations. These products did not reveal gene fusions on direct sequencing, suggesting that some non-specific binding of PCR primers was present. Nested PCR uses two sets of PCR primers, in two successive runs. The second primer set is specifically designed to amplify a product contained within the first PCR

product (Garson *et al.*, 1990). This increases the overall sensitivity of the PCR reaction, as it is very unlikely that any non-specific PCR products will contain binding sites for the second primer set.

The primers for the first stage of the nested PCR were *KIAA1549* primer 9 (forward) and *BRAF* exon 13 (reverse) (Appendix C, Table 8.2). The primers for the second stage of the nested PCR were *KIAA1549* primer 10 (forward) and *BRAF* exon 12 (reverse). Primer products were visualised using 1.5% agarose gels.

Three further *KIAA1549-BRAF* fusion variants were identified using the nested PCR primer combination: between *KIAA1549* exon 15 and *BRAF* exon 9, *KIAA1549* exon 16 and *BRAF* exon 11 (Figure 5.5) and *KIAA1549* exon 18 and *BRAF* exon 10. Validation of the additional fusion variants was performed by direct sequencing (Figure 5.6). During these experiments, both the *KIAA154-BRAF* exon 16-exon 9 and the *KIAA154-BRAF* exon 19-exon 9 fusion variants were used as positive controls. Both were identified successfully using the nested PCR primers. Hence, the nested PCR primer combination was used to detect all five *KIAA1549-BRAF* fusion variants described within this thesis. Again, these data were reproducible, and were repeated on cDNA from separate reverse transcription reactions.

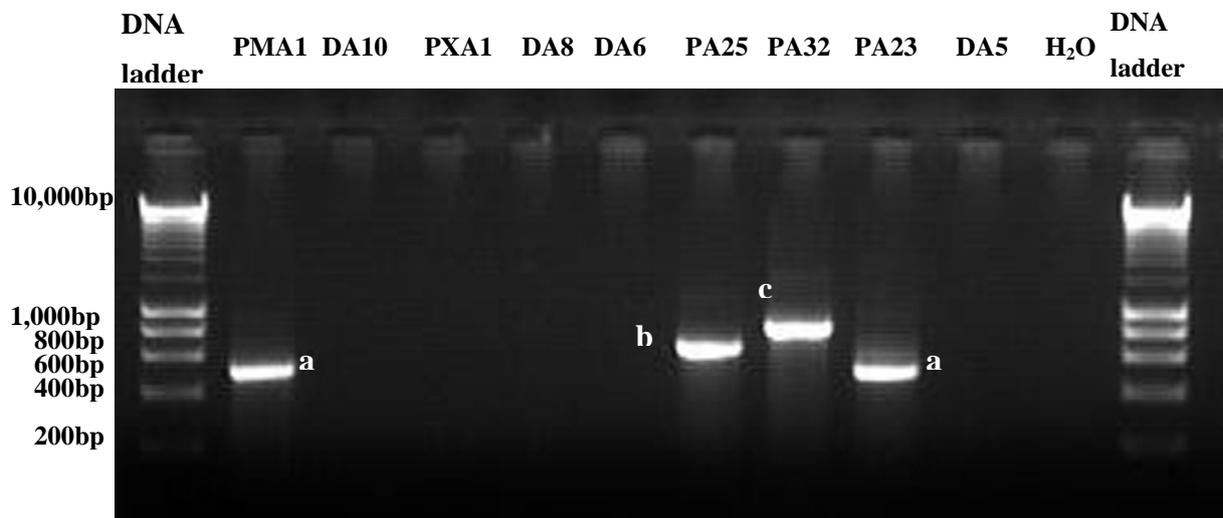


Figure 5.5 1.5% agarose gel showing results of the second stage of the nested PCR. PCR was performed for 9 tumour samples and a negative control (water). The PCR products indicate the presence of 3 fusion variants. Products indicated by (a) contain a *KIAA154-BRAF* exon 15-exon 9 variant of 505bp, (b) a *KIAA154-BRAF* exon 16-exon 11 variant of 649bp and (c) a *KIAA154-BRAF* exon 16-exon 9 variant of 823bp. The DNA ladder is Hyperladder I (200 bp-10kb, Bionline).

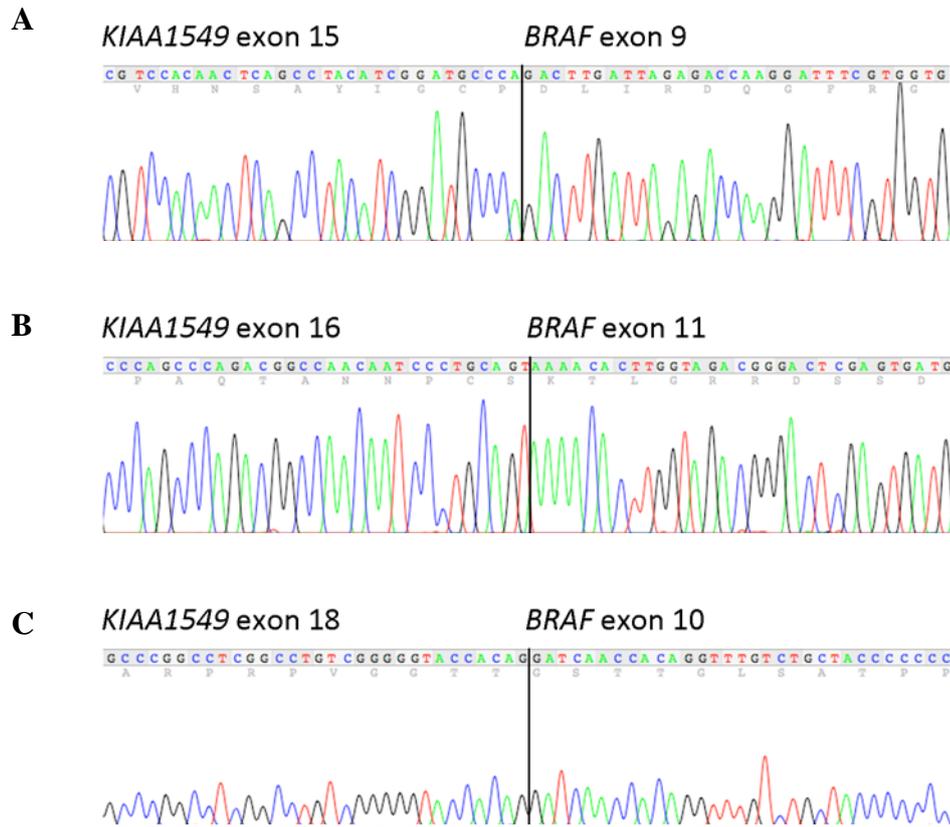


Figure 5.6 Sequence analysis of three further *KIAA1549-BRAF* fusion variants in tumour cDNA.

- (A) Fusion junction between *KIAA1549* exon 15 and *BRAF* exon 9 in PMG4.**
- (B) Fusion junction between *KIAA1549* exon 16 and *BRAF* exon 11 in PA26.**
- (C) Fusion junction between *KIAA1549* exon 18 and *BRAF* exon 10 in PA28.**

Due to the success of the nested PCR in the identification of *KIAA1549-BRAF* fusion variants, the study was extended to include 9 additional low-grade astrocytoma samples, which had not been analysed by 250k or 6.0 SNP array. In total, 50 low-grade astrocytoma samples were investigated using the nested PCR. *KIAA1549-BRAF* fusion variants were identified in 33/50 low-grade astrocytomas, including 30/32 pilocytic astrocytomas (Table 5.1). The majority were the *KIAA1549-BRAF* exon 16-exon 9 variant, which was found in 22/32 pilocytic astrocytomas. The frequencies of the different fusion variants are summarised in Table 5.2. Fusions were found in all 26 samples that contained 7q34 copy number gain, and in 6/9 samples with no copy number data. A *KIAA1549-BRAF* exon 16-exon 9 fusion was found in one sample, PA16, which did not contain 7q34 gain by 250k or 6.0 SNP array analyses.

Name	Site	Age	Sex	<i>KIAA1549-BRAF</i> fusion	7q34 gain by 250K/6.0 SNP analysis
PA1	Cerebellum	20	F	Exon 16 - Exon 9	+
PA2	Cerebellum	4	M	Exon 16 - Exon 9	+
PA3	Cerebellum	7	M	Exon 16 - Exon 9	+
PA4	Cerebellum	2	M	Exon 16 - Exon 9	+
PA5	Cerebellum	6	F	Exon 16 - Exon 9	+
PA6	Cerebellum	5	M	Exon 16 - Exon 9	+
PA7	Cerebellum	13	M	Exon 16 - Exon 9	+
PA8	Cerebellum	2	M	Exon 16 - Exon 9	Not tested
PA9	Cerebellum	14	F	Exon 16 - Exon 9	+
PA10	Cerebellum	19	M	Exon 16 - Exon 9	+
PA11	Cerebellum	14	M	Exon 16 - Exon 9	+
PA12	Cerebellum	14	M	Exon 16 - Exon 9	+
PA13	Cerebellum	10	M	Exon 16 - Exon 9	+
PA14	Cerebellum	10	M	Exon 16 - Exon 9	+
PA15	Cerebellum	13	M	Exon 16 - Exon 9	+
PA16	Cerebellum	10	M	Exon 16 - Exon 9	-
PA17	Cerebellum	6	M	Exon 16 - Exon 9	+
PA18	Cerebellum	3	M	Exon 16 - Exon 9	+
PA19	Cerebellum	3	M	Exon 16 - Exon 9	+
PA20	Cerebellum	4	M	Exon 16 - Exon 9	+
PA21	Cerebellum	12	F	Exon 16 - Exon 9	Not tested
PA22	Cerebellum	6	F	Exon 15 - Exon 9	+
PA23	Cerebellum	3	F	Exon 15 - Exon 9	+
PA24	Cerebellum	4	F	Exon 15 - Exon 9	Not tested
PA25	Cerebellum	4	F	Exon 16 - Exon 11	+
PA26	Cerebellum	3	F	Exon 16 - Exon 11	+
PA27	Cerebellum	14	F	Exon 16 - Exon 11	+
PA28	Cerebellum	7	M	Exon 18 - Exon 10	Not tested
PA29	Cerebellum	19	M	Exon 19 - Exon 9	+
PA30	Cerebellum	13	F	-	-
PA31	Cerebellum	1	F	-	-
PA32	Brain stem	9	F	Exon 16 - Exon 9	+
DA1	Cerebral cortex	5	F	Exon 15 - Exon 9	Not tested
DA2	Cerebral cortex	9	M	-	-
DA3	Cerebral cortex	17	M	-	-
DA4	Cerebral cortex	4	M	-	-
DA5	Cerebral cortex	7	F	-	-
DA6	Cerebral cortex	5	F	-	-
DA7	Cerebral cortex	5	M	-	-
DA8	Cerebral cortex	15	F	-	-
DA9	Cerebral cortex	9	F	-	-
DA10	Cerebral cortex	5	F	-	Not tested
DA11	Cerebral cortex	2	M	-	Not tested
PMA1	Diencephalon	1	F	Exon 15 - Exon 9	+
PMG1	Diencephalon	17	M	-	-
PMG2	Diencephalon	7	F	-	-
PMG3	Diencephalon	8	M	-	Not tested
PMG4	Spinal cord	12	M	Exon 15 - Exon 9	Not tested
PMG5	Spinal cord	11	F	-	-
PXA1	Cerebral cortex	6	F	-	-

Table 5.1 *KIAA1549-BRAF* fusion variants identified in 50 low-grade astrocytomas.

PA (pilocytic astrocytoma), DA (diffuse astrocytoma), PMA (pilomyxoid astrocytoma), PMG (pilomyxoid glioma), PXA (pleomorphic xanthoastrocytoma), - (no abnormality found).

Fusion variant	Frequency	
	Pilocytic astrocytoma	Non-pilocytic astrocytoma
<i>KIAA1549-BRAF</i> exon 15-exon 9	3/32	-
<i>KIAA1549-BRAF</i> exon 16-exon 9	22/32	3/18 (1 DA, 1 PMA, 1 PMG)
<i>KIAA1549-BRAF</i> exon 16-exon 11	3/32	-
<i>KIAA1549-BRAF</i> exon 18-exon 10	1/32	-
<i>KIAA1549-BRAF</i> exon 19-exon 9	1/32	-

Table 5.2 Summary of *KIAA1549-BRAF* fusion variants found in 50 low-grade paediatric astrocytomas.

All of the fusion variants would theoretically produce in-frame gene fusions. Each variant was found to occur between exons in the same start and end phases for transcription, ensuring the production of an in-frame protein (Figure 5.7). The most frequent fusion combination was between *KIAA1549-BRAF* exon 16-exon 9. Several other fusion combinations producing in-frame proteins are theoretically possible, but have not, as yet, been identified. Based on the fusion pairings discovered to date, the two most likely to occur are *KIAA1549-BRAF* exon 19-exon 11 and *KIAA1549-BRAF* exon 17-exon 8. The fusion transcript and predicted fusion protein sizes are shown in Table 5.3.

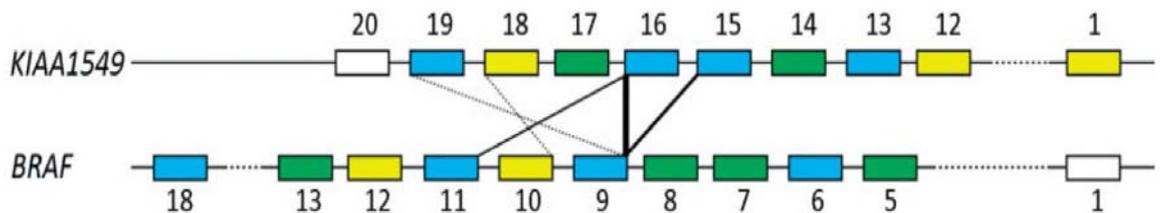


Figure 5.7 Five *KIAA1549-BRAF* fusion variants, the thickness of the black lines indicates the frequency of each variant.

The start phase of the exons in *BRAF* and the end phases of the exons in *KIAA1549* are denoted by the following colours: blue = 0, yellow = 1, green = 2, white = untranslated. Only a fusion between exons in the same colour (i.e. start and end phase) will produce an in-frame protein.

<i>KIAA1549-BRAF</i> fusion	Transcript length	Predicted protein size
Exon 15-Exon 9 (short transcript)	2626bp	90.77kDa
Exon 15-Exon 9 (long transcript)	6257bp	221.4kDa
Exon 16-Exon 9 (short transcript)	2944bp	101.8kDa
Exon 16-Exon 9 (long transcript)	6575bp	232.5kDa
Exon 16-Exon 11 (short transcript)	2770bp	95.7kDa
Exon 16-Exon 11 (long transcript)	6401bp	226.4kDa
Exon 18-Exon 10 (short transcript)	3112bp	107.7kDa
Exon 18-Exon 10 (long transcript)	6743bp	238.3kDa
Exon 19-Exon 9 (short transcript)	3195bp	114.1kDa
Exon 19-Exon 9 (long transcript)	6926bp	244.8kDa

Table 5.3 Transcript lengths and predicted protein sizes for the five *KIAA1549-BRAF* fusion variants.

Transcript lengths and predicted protein sizes are shown for the short form where transcription may have been initiated from the promotor within intron 8, as in the short form of the *KIAA1549-BRAF* fusion described by Jones *et al.* (Jones *et al.*, 2008). The long transcript is shown where protein transcription may have commenced from the promotor within exon 1 of *KIAA1549*.

5.4 Mapping genomic DNA *KIAA1549-BRAF* fusion breakpoints

The *KIAA1549-BRAF* fusion breakpoints were confirmed by sequencing whole-genome amplified (WGA) genomic DNA from 9 tumours (PAs 1, 2, 3, 4, 5, 6, 7, 8 and 32). All contained the most prevalent *KIAA1549-BRAF* exon 16-exon 9 fusion variant. Forward primers were designed, spaced approximately 600bp apart throughout *KIAA1549* intron 16, reverse primers were designed similarly across *BRAF* intron 8. A 96-well format was used, where each column corresponded to one of 12 *KIAA1549* DNA primers, and each row corresponded to one of 8 *BRAF* DNA primers (Appendix C, Table 8.3). PCRs were performed using pooled WGA genomic DNA from the 9 tumour samples, and pooled male control DNA. PCR products were visualised using 1.5% agarose gels (Figure 5.8). All primer combinations producing a tumour-specific product were then tested in individual tumour DNA samples, and validated by direct sequencing (Figure 5.9).

The genomic DNA fusion breakpoints were identified within intron 16 of *KIAA1549* and intron 8 of *BRAF* for all 9 tumours tested. These are summarised in Figure 5.10. Two tumours (PA4 and PA7) showed a sequence change from *KIAA1549* to *BRAF* across a single base within the intronic sequence. Six tumours (PAs 1, 2, 3, 5, 6 and 8) showed small regions of microhomology, ranging from 2-4 bases, between the two intronic sequences across the fusion junction. One of these samples (PA1) also contained three additional bases (TAA) inserted at the fusion junction, a single base after the region of microhomology. Finally, one sample (PA32) contained a single cytosine base insertion at the fusion junction.

All of the tumours showed breakpoints at different sites within these introns, suggesting that the fusion breakpoints were not associated with a specific sequence motif. No known fragile sites were identified within this region, or within the wider genomic region involving *KIAA1549* and *BRAF*. The positions of Alu elements, which are implicated in tandem duplications in acute myeloid leukaemia, were mapped across the region of DNA fusion breakpoints. The fusion breakpoints were not related to the situation of Alu elements (Figure 5.11). Further scrutiny of the region was performed to identify regions of segmental duplication or low-copy repeats. These duplicated segments of DNA contain nearly identical sequence, which may act as foci for non-allelic homologous recombination, leading to translocations or other genomic rearrangements (Cheung *et al.*, 2003, Bailey *et al.*, 2006). Human segmental duplications are defined as continuous portions of DNA >1kb in size containing >90% sequence identity, which map to more than one genomic location (Bailey *et al.*, 2006). No segmental duplications were identified in the genomic sequences of *KIAA1549* or *BRAF*, after searching within the Human Genome Segmental Duplication Database (<http://projects.tcag.ca/humandup/>) or the Human Segmental Duplication Database (<http://humanparalogy.gs.washington.edu/>).

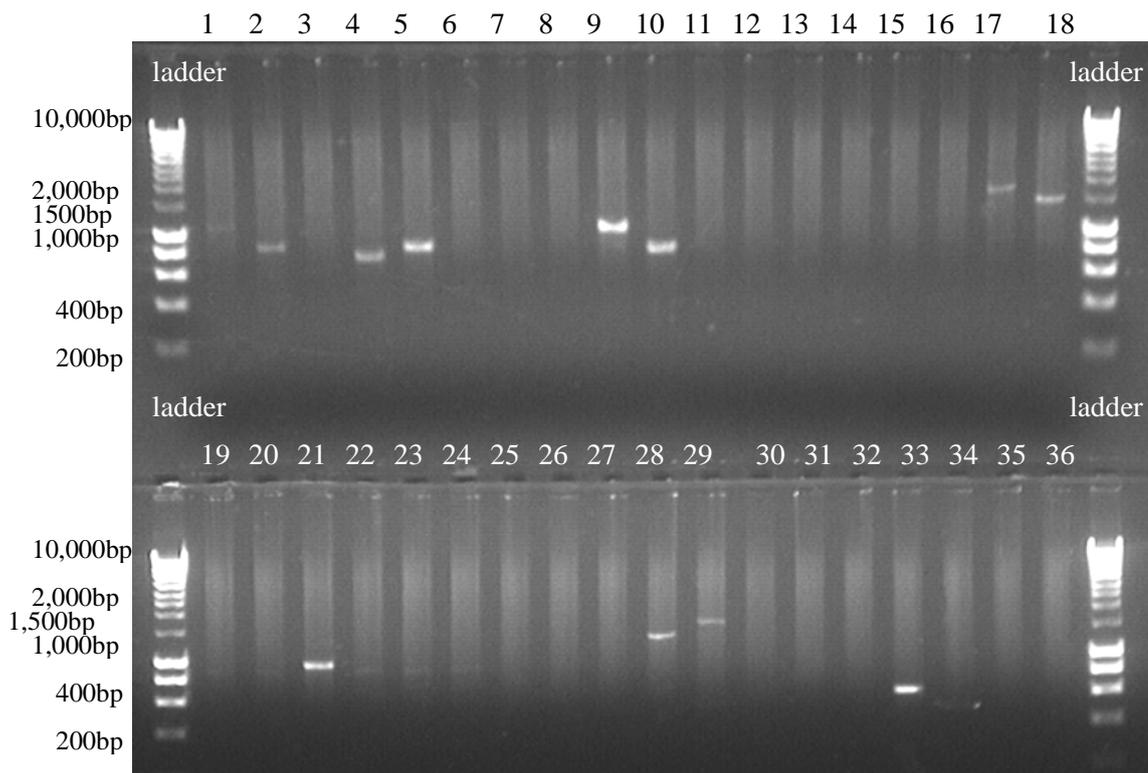


Figure 5.9 1.5% agarose gel for the identification of *KIAA1549-BRAF* fusion breakpoints in pooled genomic DNA from 9 tumours.

Each well contains the PCR from one primer combination in the 96-well plate format, and each PCR product indicates a gene fusion identified by a specific primer combination. Thirty-six separate reaction results are shown here. Positive PCR products were then validated by PCR in separate tumour DNA samples followed by direct sequencing.

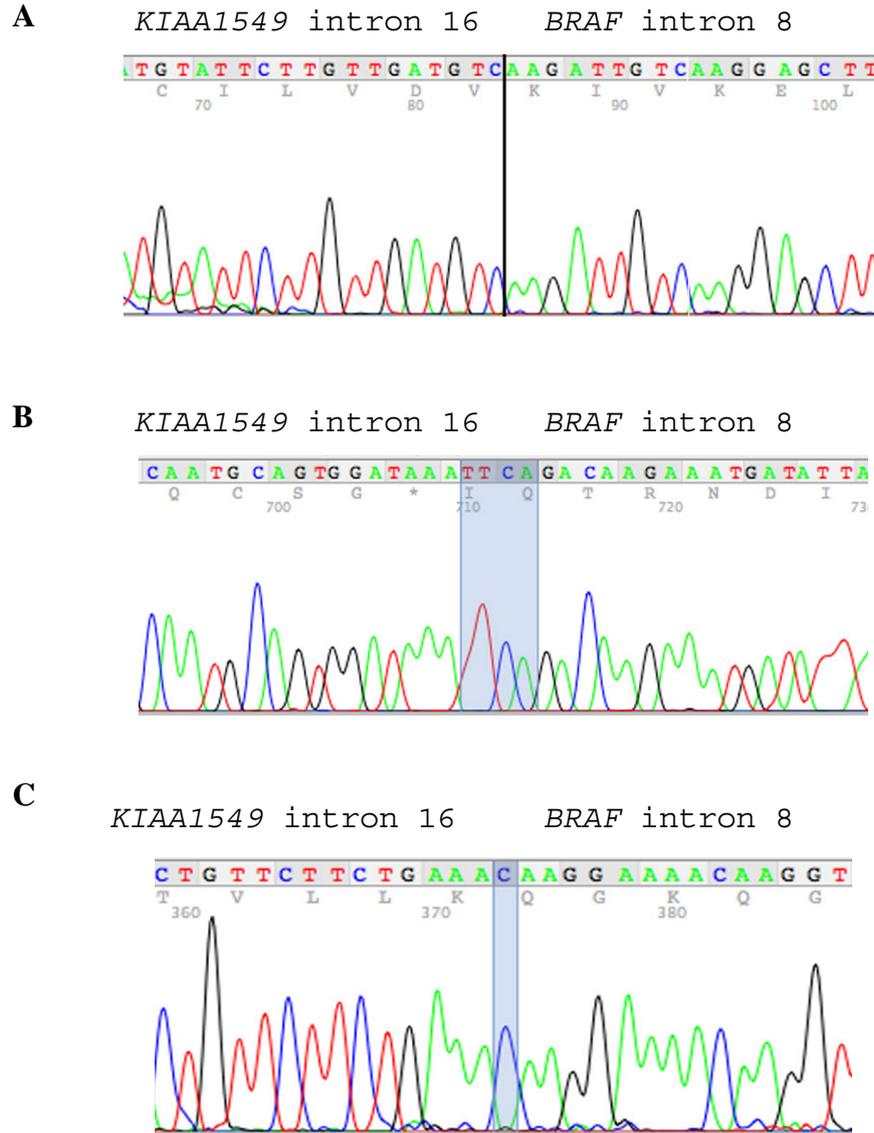


Figure 5.10 Sequence traces for the fusion junctions between *KIAA1549* intron 16 and *BRAF* intron 8 in genomic DNA from three tumour samples.

(A) Fusion junction in PA4, showing the sequence change at a distinct site between *KIAA1549* and *BRAF*.

(B) The fusion junction in PA3 demonstrates microhomology across 4 bases (highlighted in blue) between *KIAA1549* and *BRAF*.

(C) The fusion junction in PA32 demonstrates the insertion of a single cytosine base (highlighted in blue) between *KIAA1549* and *BRAF*.

KIAA1549 TTGCATTTGATAAG AATT A---AAACTCGGGTGATGGGTC
 PA1 TTGCATTTGATAAG AATT A T A T A T A T T T T T A A A T T A T A C
BRAF ACCAGTGTACTTTT AATT T---TATATTTTAAATTATAC

KIAA1549 TGCTATGGATTGAACTGT GTC CTCCCCAAAATCACATGTT
 PA2 TGCTATGGATTGAACTGT GTC ATTTGTTAAGGAAATGAAA
BRAF CATAAGGGGGTTAGGAAA GTC ATTTGTTAAGGAAATGAAA

KIAA1549 TGCAATGCAGTGGATAAA TTCA AGAGGAAGAATCAGCGAG
 PA3 TGCAATGCAGTGGATAAA TTCA GACAAGAAATGATATTTA
BRAF ATGTGGTTTTTACAGTAT TTCA GACAAGAAATGATATTTA

KIAA1549 AATGTATTCTTGTGATGTC TCCAGAAAAGTTTCAGTTTA
 PA4 AATGTATTCTTGTGATGTC AAGATTGTCAAGGAGCTTTA
BRAF GTAACCTAGAATGAGTAGAA AAGATTGTCAAGGAGCTTTA

KIAA1549 AGAAGGAAAGATCTAGAAA TG AGTTGTGTGTTTTCTAAAG
 PA5 AGAAGGAAAGATCTAGAAA TG CACCAGTATAATTTTTTTT
BRAF ATTTTATCTCTCTTAATG TG CACCAGTATAATTTTTTTT

KIAA1549 GTATGCATATGTGTTTGT ATGG AGAGAGTGAGGGGGAGAG
 PA6 GTATGCATATGTGTTTGT ATGG CCTTTGCCCTCAAGGAAG
BRAF TTATTTATGAATTAGCAC ATGG CCTTTGCCCTCAAGGAAG

KIAA1549 GCAGGAGGACTTTTGCTTTG TGGAGGAACGGAGGGGGATT
 PA7 GCAGGAGGACTTTTGCTTTG AAGATGCCTGTGGGACTTTC
BRAF TGGGGGCATATATTGACTTTA AAGATGCCTGTGGGACTTTC

KIAA1549 TGAGCATTCTGGGGATG GGCTT AGGAGCCACATCGTGCTG
 PA8 TGAGCATTCTGGGGATG TGCTT GCTTGGGAATAAATATACA
BRAF GCTACTCTCTGATTATAT TGCTT GCTTGGGAATAAATATACA

KIAA1549 TAGTTACTGTTCTTCTGAAA-TATCTAATTCCAAAAGGAA
 PA32 TAGTTACTGTTCTTCTGAAAC AAGGAAAACAAGGTTTGCA
BRAF GCACAGTTTGTTAAATACCA-AAGGAAAACAAGGTTTGCA

Figure 5.11 Sequence traces for the DNA fusion junctions in 9 pilocytic astrocytomas, all containing the *KIAA1549-BRAF* exon 16-exon 9 fusion.

The tumour sequence traces are compared with *KIAA1549* (identical sequence is highlighted in green) and *BRAF* (identical sequence is highlighted in yellow). Regions of microhomology between all three sequences are highlighted in red. Two tumour samples (PA1, PA32) contain base insertions (highlighted in blue), both occurring at the fusion junction. Two samples (PA4, PA7) do not show any region of microhomology, as the sequence changes abruptly in-frame from *KIAA1549* to *BRAF*.

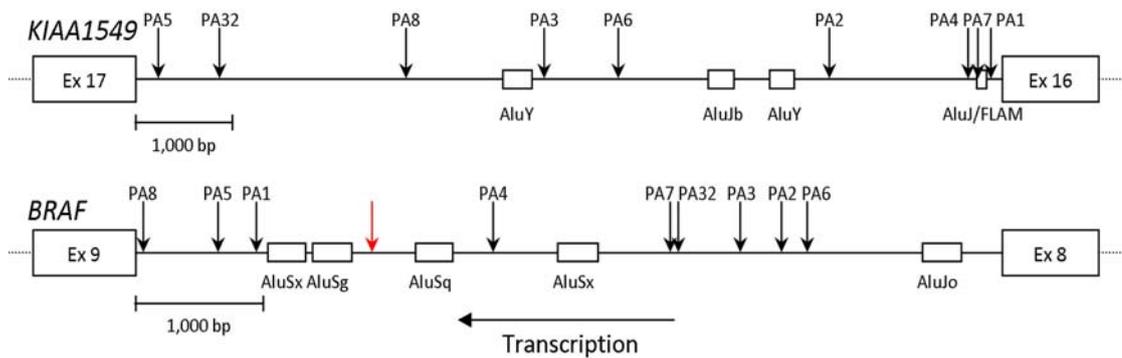


Figure 5.12 Schematic representation of the mapping results for DNA fusion breakpoints within intron 16 of *KIAA1549* and intron 8 of *BRAF* in 9 tumours with the *KIAA1549-BRAF* exon 16-exon 9 fusion.

Both genes are transcribed in the same direction. The fusion breakpoints are different in all of the 9 tumours, and are not related to the positions of Alu elements. The red arrow indicates the position of the *AKAP9-BRAF* fusion in thyroid carcinoma.

5.5 Identification of *SRGAP3-RAF1* fusion in the single pilocytic astrocytoma with 3p25 DNA copy number gain

Two of the 32 pilocytic astrocytomas studied did not contain a *KIAA1549-BRAF* fusion. One of these tumours, PA30, contained a discrete region of DNA copy number gain at chromosome 3p25, starting within *SRGAP3* and ending within *RAF1*. *RAF1* is also a Raf kinase, from the same family as *BRAF*. Following the successful identification of the *KIAA1549-BRAF* gene fusion, a similar approach was used to search for a fusion gene between *SRGAP3* and *RAF1*.

Nine primers were designed throughout *SRGAP3*, and 3 primers were designed within the kinase domain of *RAF1* (Appendix C, Table 8.4). PCR indicated the presence of a gene fusion, from a number of different primer combinations. The PCR product from *SRGAP3* forward primer 5 (exon 10) and *RAF1* reverse primer 2 (exon 11) was sequenced (as other products were too large). This identified a fusion gene between *SRGAP3* exon 11 and *RAF1* exon 8 (figure 5.12). Sequence analysis also identified an in-frame duplication of *SRGAP3* exon 11, before the fusion junction with *RAF1*.

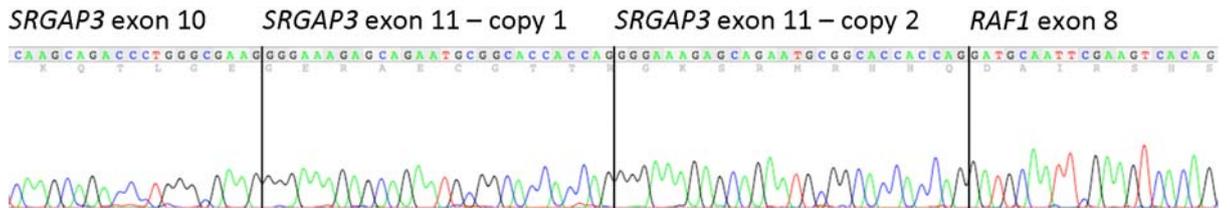


Figure 5.13 Sequence analysis of the *SRGAP3-RAF1* exon11-exon8 fusion junction in tumour PA30. Exon 11 of *SRGAP3* has been duplicated, remaining in-frame.

5.5.1 Western blotting confirms activation of the MAPK pathway

BRAF and RAF1 are serine-threonine-specific protein kinases, activated downstream of RAS. They activate the extracellular signal-regulated kinases MEK 1/2 by phosphorylation, which in turn activate the kinases ERK1/2 by phosphorylation (Figure 5.13). Once activated, ERK1 and ERK2 translocate into the nucleus where numerous physiological responses are elicited, including cell proliferation, differentiation and apoptosis (Roberts *et al.*, 2007).

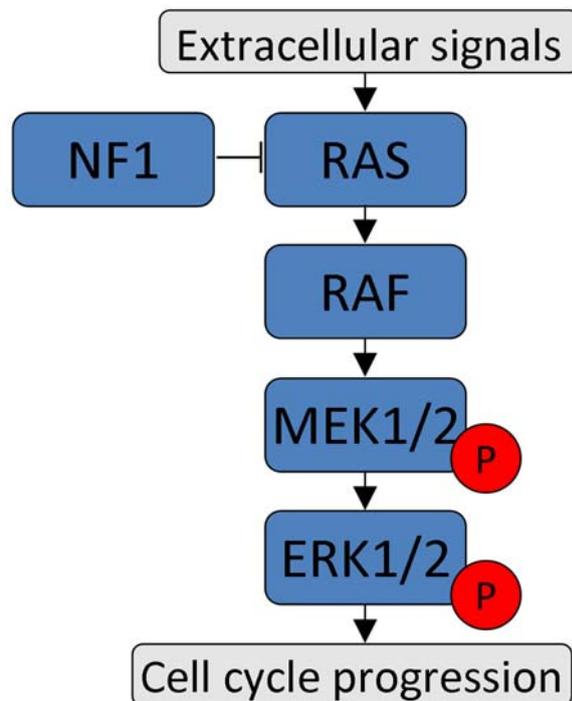


Figure 5.14 Schematic representation of the MAPK pathway.

Both MEK and ERK are activated by phosphorylation. Once activated, phospho-ERKs translocate to the nucleus, where a variety of responses are elicited. These include the promotion of cell proliferation, differentiation and apoptosis.

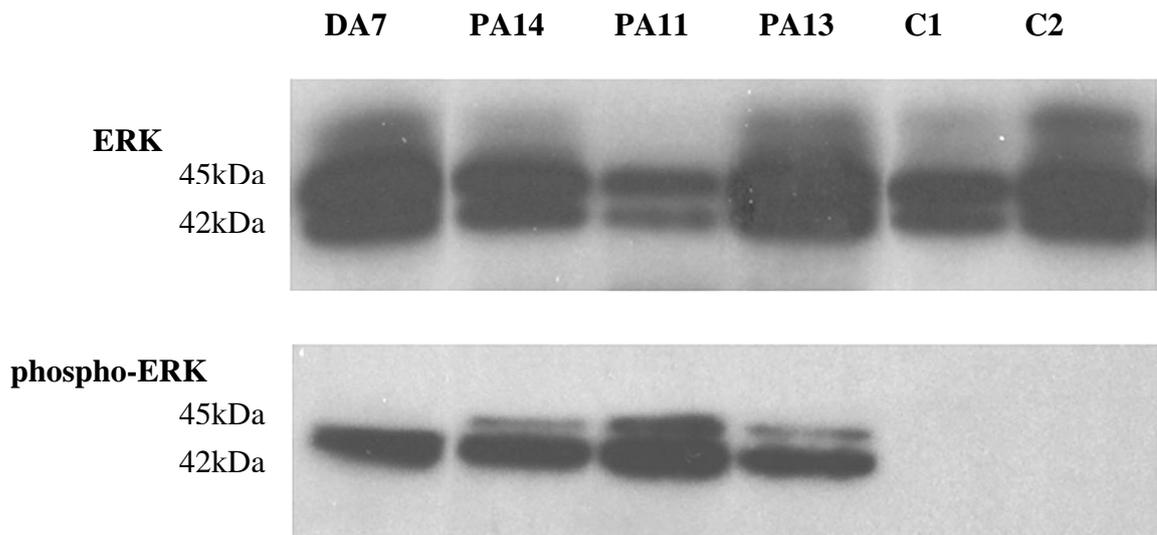


Figure 5.15 Western blots showing ERK and phospho-ERK in a representative number of tumour samples and controls.

Protein sizes are indicated in kilodaltons. C1 (fetal cerebellum control), C2 (fetal frontal lobe control).

Western blotting was performed by Dr William Ogunkolade, at the Blizard Institute of Cell and Molecular Science, London, to confirm whether the MAPK pathway was activated in samples with fusion genes. Eleven tumours were examined: 7 contained *KIAA1549-BRAF* fusions (PAs 2, 11, 13, 14, 15, 25 and 29), one contained the *SRGAP3-RAF1* fusion (PA30) and three did not contain gene fusions (DAs 3, 4 and 7). Total protein was obtained from two fetal brain controls, one from cerebellum and one from the cerebrum (frontal lobe), (BioChain, Hayward, CA).

Twenty micrograms of protein lysates and control brain protein were separated on Nupage®Novex® 4-12% Bis-Tris Gel 4 (Invitrogen) and transferred to polyvinylidene difluoride membranes (GE Healthcare, Buckinghamshire, UK). Membranes were then probed with antibodies specific to MEK1/2, phospho-MEK1/2, ERK1/2 and phospho-ERK1/2 (p44/42) (all from Cell Signaling Technology, Hertfordshire, UK). After probing the PVDF membranes with appropriate peroxidase-labelled secondary antibodies (GE Healthcare, Buckinghamshire, UK), signals were developed using ECL Plus Western blotting detection reagents (GE Healthcare, Buckinghamshire, UK) according to the manufacturer's recommendations, and visualized on a Kodak film (BioMax MR; Kodak, Rochester, NY).

Western blot analysis showed no differences in the levels of MEK and ERK between tumours and controls. However, high levels of phospho-MEK (data not shown) and phospho-ERK were found in all of the tumour samples when compared to controls (figure 5.14). These results confirmed that activation of the MAPK pathway was present in tumours containing *RAF* gene fusions, implying that activation of the pathway may have occurred due to the presence of gene fusions. However, these data do not show a causal role for activation of the MAPK pathway by gene fusions, as MAPK pathway activation was also seen in three tumours without gene fusions. These data suggest that activation of the MAPK pathway is, however, a critical event leading to the development of paediatric low-grade astrocytomas.

5.6 Discussion: RAF fusion genes in paediatric low-grade astrocytoma

Fusions involving two *RAF* genes, *RAF1* and *BRAF* were identified in 34/50 low-grade astrocytomas, including 31/32 pilocytic astrocytomas and 3/18 Grade II astrocytomas. Both were identified by the presence of discrete regions of DNA copy number gain at 3p25 and 7q34, respectively. Each *RAF* gene appeared to show duplication only at the 3' end, which contains the coding sequence for the kinase domain. Duplication at 7q34 was confirmed using iFISH in paraffin sections from tumours known to contain 7q34 gain by SNP analysis. *BRAF* has been shown to fuse with *AKAP9* and *FCHSD* in other malignancies, retaining the kinase domain in each case. These data suggested a possible mechanism for the *RAF* gene fusions to arise, via tandem duplications (Figure 5.15).

The *SRGAP3-RAF1* fusion was found in a single pilocytic astrocytoma. One of the five *KIAA1549-BRAF* fusion variants was found in 30/32 pilocytic astrocytomas. In both *RAF* fusion genes the autoinhibitory domains within BRAF and RAF1 are replaced in-frame with the N-terminal segments of *KIAA1549* and *SRGAP3*, respectively (Figure 5.16). Hence, the fusion proteins retain the kinase domains of BRAF and RAF1.

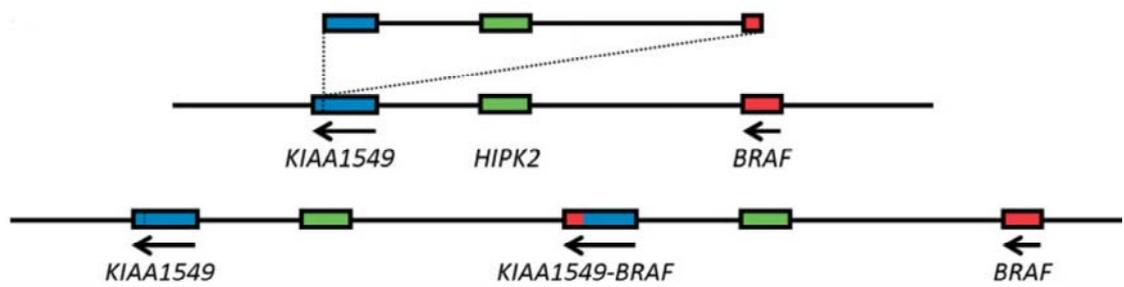


Figure 5.16 Proposed mechanism for the *KIAA1549-BRAF* fusion to arise by tandem duplication and reinsertion.

The fusion would also arise if the duplicated segment inserted into *BRAF* rather than *KIAA1549*. The *SRGAP3-RAF1* fusion is proposed to arise by the same mechanism.

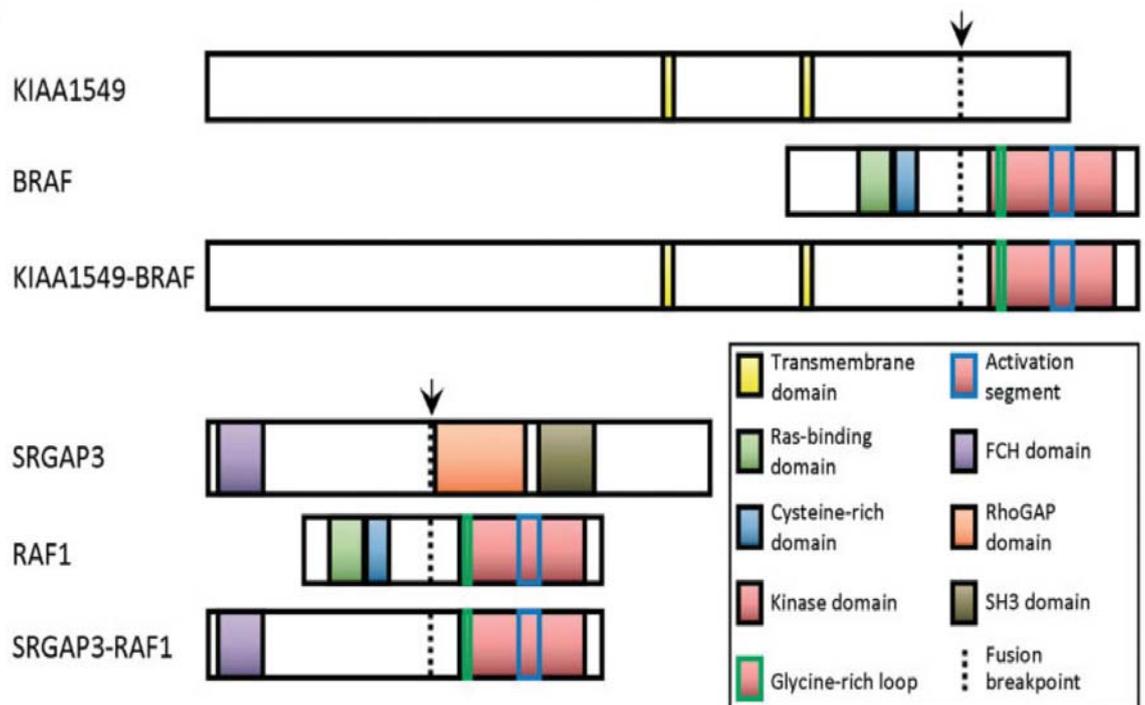


Figure 5.17 Schematic representations of protein domains within *KIAA1549*, *BRAF*, *SRGAP3*, *RAF1* and the resulting fusion proteins.

The dotted lines indicate the positions of the fusion breakpoints. The auto-inhibitory domains within *BRAF* and *RAF1* are found within the Ras-binding and cysteine-rich domains.

BRAF and RAF1 are well-characterised protein kinases within the MAPK pathway, and both retain their kinase domains in the gene fusions described here. However, their fusion partner genes, *KIAA1549* and *SRGAP3*, may also have an important part to play in the process of tumorigenesis.

KIAA1549 is expressed in many tissues, including the brain, however, nothing is known about the normal cellular role of the protein it encodes (Nagase *et al.*, 2000). An additional exon was recently identified at the start of *KIAA1549* (Jones *et al.*, 2008), which was confirmed by sequencing here. Hence, all exon numbering here for *KIAA1549* contains the additional first exon, giving a total of 20 exons. *KIAA1549* is thought to contain two transmembrane domains: it is possible that these may be used to attach the fusion gene product to the cell membrane. *KIAA1549* has also recently been shown to have two alternative transcripts, one long and one short (Jones *et al.*, 2008). The shorter transcript originates from a promoter within intron 8, and may contain only a single transmembrane domain. The relative abundance of the two transcripts, and whether one or both contribute to the fusion with *BRAF* is currently unknown.

The second fusion partner *SRGAP3* (Slit-Robo GTPase activating protein) has been characterised much more fully, and has been shown to play roles in neuronal migration, axonal branching and in the regulation of neuronal development in a complex with the scaffolding protein WAVE1 (Wiskott-Aldrich syndrome protein family member 1) (Wong *et al.*, 2001, Soderling *et al.*, 2007). *SRGAP3* expression has also been shown at sites of neural proliferation in the developing murine brain, suggesting that *SRGAP3* has a role in the control of neurogenesis (Bacon *et al.*, 2009). *SRGAP3* is involved in a balanced constitutive translocation t(X;3)(p11.2;p25) in a patient with severe mental retardation, however the breakpoint is within intron 3, unlike the intron 11 breakpoint described here (Endris *et al.*, 2002). *SRGAP3* contains three functional motifs: the N-terminal FCH (Fes/CIP4 homology) domain is retained in the fusion, but the RhoGAP and C-terminal SH3 domains are lost. FCH domains are present in other proteins, where they have a putative function in regulating the microtubule cytoskeleton (Yang *et al.*, 2006). The FCH domain is retained in the *FCHSD1-BRAF* fusion in large congenital melanocytic naevi, which suggests that a selective advantage may be conferred to cells containing these fusion proteins.

The identification of *KIAA1549-BRAF* and *SRGAP3-RAF1* fusion genes, particularly at such a high frequency in pilocytic astrocytoma, has been a huge breakthrough in the understanding the importance of MAPK pathway activation in the origins of paediatric low-grade astrocytoma. These findings were published in the Journal of Pathology in March 2009, entitled ‘Activation of the ERK/MAPK pathway: a signature genetic defect in pilocytic astrocytomas’, where I am joint first author (Forsheew *et al.*, 2009).

The MAPK pathway was first implicated in the development of astrocytomas from studies of patients with the tumour-predisposition syndrome Neurofibromatosis type 1 (NF-1). NF-1 is a common autosomal dominant disorder, which affects around 1 in 3,000 individuals, and is caused by a germ-line inactivating mutation in the *NF-1* tumour suppressor gene located at chromosome band 17q11.2 (von Deimling *et al.*, 1995, Rubin *et al.*, 2005, Brems *et al.*, 2009). Neurofibromin, the protein encoded by *NF-1* acts as a negative regulator of the RAS signal transduction pathway, as shown in figure 5.13. Neurofibromin interacts with RAS via a GTPase activating domain, converting active RAS-GTP to its inactive form. Loss of a single allele of NF-1 is insufficient for tumorigenesis – the second allele must also be lost, as outlined in Knudson’s two-hit hypothesis (Knudson, 1985). Loss of heterozygosity (LOH), by partial chromosome deletion, whole chromosome loss or mitotic recombination causes the second hit in NF-1 associated tumours. Penetrance of NF-1 is 100% by 5 years of age (Huson *et al.*, 1989). Nearly all individuals with NF-1 will develop a benign or malignant nervous system tumour during their lifetime – these include optic pathway gliomas (pilocytic astrocytomas of the optic nerve and chiasm), neurofibromas and hamartomas of the iris (Lisch nodules) (Rubin *et al.*, 2005, Brems *et al.*, 2009). Pilocytic astrocytomas of the optic pathway are seen in 15-20% children with NF-1 (Listernick *et al.*, 2007). These may remain indolent for years, however, between 30-50% will cause symptoms including sudden and unpredictable loss of vision. Their histological appearance is usually indistinguishable from that of sporadic pilocytic astrocytoma (PA). However, very few sporadic PAs exhibit NF-1 inactivation, either by LOH or chromosomal deletions at 17q11.2, or DNA methylation, and expression of neurofibromin has also been demonstrated in these tumours (Kluwe *et al.*, 2001, Wimmer *et al.*, 2002, Ebinger *et al.*, 2005). These indications that sporadic PAs arise by different mechanisms to those for NF-1 associated PAs led to a number of studies investigating alternative methods of MAPK pathway activation in low-grade astrocytomas.

The ERK/MAPK pathway is upregulated in about a third of all cancers, and in many cases this is due to activating point mutations in the *RAF* and *RAS* genes (Downward, 2003, Dhillon *et al.*, 2007). Activating mutations in *RAS* are most commonly found at codons 12, 13 and 61: all prevent the hydrolysis of GTP. Thus, *RAS* remains in an active GTP-bound form, activating its effectors including *RAF*. The most common *RAF* mutation is *BRAF* V600E, seen in over 90% cases, which causes a conformational change in the kinase activation loop, inducing constitutive catalytic activation. In contrast, *ARAF* and *RAF1* (*CRAF*) mutations are extremely rare (Lee *et al.*, 2005a, Dhillon *et al.*, 2007). The reason for this may be the fact that serine 445 is constitutively activated in *BRAF*, while *ARAF* and *RAF1* (*CRAF*) require further phosphorylation to become activated (Wellbrock *et al.*, 2004). Hence, *BRAF* appears to be poised for activation, which may explain the propensity for mutations in cancer to occur in *BRAF*. *RAS* and *RAF* mutations appear to be mutually exclusive (Petti *et al.*, 2006, Dhillon *et al.*, 2007). Melanomas containing both *NRAS* and *BRAF* mutations have been described, but these do not appear to occur within the same cell: co-expression of *NRAS* Q61R and *BRAF* V600E was found to induce senescence within cultured melanoma cells (Petti *et al.*, 2006).

Activation of the MAPK pathway was found in 21/21 sporadic pilocytic astrocytomas using immunohistochemistry for phospho-AKT, a downstream effector from *RAS* (Sharma *et al.*, 2005). However, this high level of pathway activation was not explained by the presence of *RAS* and *RAF* activating mutations. *KRAS* activating mutations affecting codons 12, 13 and 61 have been identified in pilocytic astrocytomas, however this is a rare event affecting between 4-7% of tumours (Maltzman *et al.*, 1997, Sharma *et al.*, 2005, Janzarik *et al.*, 2007). Similarly low incidence rates (~ 6%) have also been found for *BRAF* V600E mutation, which has been identified in 5 pilocytic astrocytomas and 2 diffuse astrocytomas to date (Jones *et al.*, 2008, Pfister *et al.*, 2008, Sievert *et al.*, 2009).

Gene transcription and expression may be inactivated epigenetically, by methylation of CpG islands in gene promoters. *RASSF1* is a tumour suppressor gene located on 3p21.3, which acts downstream in the *RAS* signal transduction pathway. *RASSF1A* is inactivated by promoter hypermethylation in many cancers, including lung, breast, ovarian and thyroid cancer (Agathangelou *et al.*, 2001). *RASSF1A* promoter hypermethylation has been identified in astrocytomas of all grades, including pilocytic

astrocytomas (Yu *et al.*, 2004, Lorente *et al.*, 2009). Thus, a range of mechanisms for deregulating the ERK/MAPK pathway have been identified in low-grade astrocytomas, including *NFI* deletion and mutations, *RASSF1A* methylation, *BRAF* V600E and *KRAS* activating mutations. Taken together, these data indicated that other factors were implicated in the pathogenesis of low-grade astrocytomas, as none appeared to account for the high numbers of samples with MAPK pathway activation identified by Sharma *et al.* (Sharma *et al.*, 2005).

DNA copy number analysis by array CGH of 66 paediatric low-grade astrocytomas, by Pfister *et al.* identified a discrete 0.97Mb region of gain at chromosome 7q34 in 13 pilocytic astrocytomas, encompassing the majority of *BRAF* (Pfister *et al.*, 2008). Dual colour FISH, using a chromosome 7 centromeric probe and RP4-726N20 to probe the 3' end of *BRAF* was conducted for 28 adult low-grade astrocytomas (3 pilocytic astrocytomas and 25 diffuse astrocytomas) in a tissue microarray. Results confirmed duplication of the *BRAF* locus in the presence of two homologues of chromosome 7 in 6 diffuse astrocytomas. Furthermore, ten tumours (3 PAs and 7 DAs) displayed gains of both *BRAF* and the chromosome 7 centromeric probes, indicating large genomic gains of chromosome 7. Western blot analysis of tumour samples containing the *BRAF* duplication showed high levels of expression of phosphorylated ERK when compared to non-neoplastic brain, suggesting constitutive activation of the MAPK pathway in pilocytic astrocytomas.

A further study by Deshmukh *et al.*, published the same month, identified a 1Mb region of copy number gain at 7q34 in 10 paediatric patients with cerebellar pilocytic astrocytomas by array CGH (Deshmukh *et al.*, 2008). Here the gained region encompassed the 3' terminal portion of *BRAF* in three samples, however the majority contained a common region of copy number gain spanning the genes *HIPK2* (homeodomain-interacting kinase 2) and *TBXAS1* (thromboxane A synthase 1). Similarly, Bar *et al.* identified the 7q34 region of gain in 17/20 pilocytic astrocytomas, commenting that the gained region only encompassed exons coding for the *BRAF* kinase domain (Bar *et al.*, 2008). Quantitative RT-PCR confirmed the presence of 3 or more copies of the 7q34 locus, using sequences within *HIPK2*, *RAB19B* and *BRAF*. Western blotting also revealed high levels of phospho-ERK in 3/6 pilocytic astrocytomas with 7q34 gain, when compared to one tumour without 7q34 gain and two normal brain specimens.

Two studies published within a week of each other in late 2008 revealed gene fusions between *KIAA1549* and *BRAF* at regions of 7q34 duplication in low-grade astrocytomas (Jones *et al.*, 2008, Sievert *et al.*, 2009). The first study by Sievert *et al.* used Illumina HumanHap550K SNP arrays, and found 7q34 gain in 17/22 paediatric PAs and 3/6 fibrillary astrocytomas. A novel *KIAA1549-BRAF* fusion product was identified by reverse transcription PCR, and sequence analysis revealed an in-frame fusion between *KIAA1549* exon 15 and *BRAF* exon 9. The predicted fusion product contained the protein kinase domain of *BRAF* and lacked the N-terminal regulatory domains, conferring constitutive kinase activity. However, an additional first exon was subsequently found in *KIAA1549*, approximately 60kb upstream of the previously documented first exon using 5' RACE analysis (Jones *et al.*, 2008). Hence Sievert *et al.* identified the *KIAA1549-BRAF* exon 16-exon 9 variant, once the exon numbering was altered due to the additional first exon in *KIAA1549*.

Jones *et al.* found the 7q34 region of gain in 44 pilocytic astrocytomas using 1Mb array CGH, and a chromosome 7 tiling path array to investigate the region at higher resolution (Jones *et al.*, 2008). Three fusion variants were identified: between *KIAA1549-BRAF* exon 16-exon 9 (20 cases), *KIAA1549-BRAF* exon 16-exon 11 (2 cases) and *KIAA1549-BRAF* exon 15-exon 9 (7 cases). An additional first exon was identified in *KIAA1549* by 5' RACE analysis. Two possible *KIAA1549* transcripts, one long and one short (beginning at a promoter within intron 8) were confirmed by PCR, and an alternative splice site was found in intron 19, giving rise to two possible variants of *KIAA1549* exon 20, one 48bp longer than the other. The long and short isoforms of the *KIAA1549-BRAF* exon 16-exon 9 were cloned, and used alongside *BRAF* V600E and wild-type *BRAF* to transfect Cos-7 cells. Kinase activity was tested for all using a chemiluminescence assay. Both fusion isoforms showed constitutive kinase activity, the highest levels were seen for the long isoform, the short form displayed similar levels to those seen for *BRAF* V600E when compared to wild-type *BRAF*. NIH3T3 cells were stably transduced with an empty vector, *HRAS* V12, *BRAF* V600E and the short *KIAA1549-BRAF* exon 16-exon 9 isoform. Anchorage independent growth was seen in cells containing the short *KIAA1549-BRAF* isoform, *BRAF* V600E and *HRAS* V12, but not in cells containing wild-type *BRAF* or the empty vector. This study concluded that the 7q34 gains seen in 26/29 PAs were highly specific for pilocytic astrocytomas. Six grade II astrocytomas were also found to have 7q34 gains, however, these were felt to

represent false positives or misdiagnosed pilocytic astrocytomas as patient survival was >12 years for all 6 cases. The 7q34 gain was not seen in 60 anaplastic astrocytomas or 184 glioblastomas, suggesting that the gain was highly specific for pilocytic astrocytomas. These latter cases were not described in detail, however, and it is likely that the majority of these tumours were from adults.

Five *KIAA1549-BRAF* fusion variants were identified in 30/32 PAs and 3/18 Grade II astrocytomas, from the work described in this chapter. The most common variant was between *KIAA1549-BRAF* exon 16-exon 9, which was found in 22/32 PAs. Further fusion variants were present at much lower frequencies, as summarised in Table 5.2. Some fusion gene variants were only revealed by sequencing, as not all of the 50 tumour samples had undergone DNA copy number analysis. Where present, copy number data confirmed 7q34 gain in all of the samples with fusions except for one, PA16. This tumour was found to contain a *KIAA1549-BRAF* exon 16-exon 9 fusion, but did not show DNA copy number gain at 7q34 by either 250k or 6.0 SNP array analyses. It is possible that this sample contained a higher amount of contaminating normal tissue, which obscured the presence of DNA copy number gain. Interphase FISH has not been possible on this sample to date, but would be an important piece of evidence to show whether DNA duplication at 7q34 is present in this tumour. If not, it is possible that the *KIAA1549-BRAF* fusion has occurred by another mechanism, not involving tandem duplication. Sievert et al described a similar case, where a single pilocytic astrocytoma was found to contain the *KIAA1549-BRAF* exon 16-exon 9 fusion by sequencing, but did not show 7q34 gain by SNP analysis (Sievert *et al.*, 2009). However, this sample was shown to contain three copies of *LUC7L2* by FISH, which lies between *KIAA1549* and *BRAF* within the region of 7q34 gain.

All of the studies have found *KIAA1549-BRAF* fusions in a large proportion of pilocytic astrocytomas, but their role in grade II astrocytomas is less clear. Three *KIAA1549-BRAF* exon 15-exon9 fusions were found in Grade II astrocytomas (DA1, PMA1, PMG4) in this study. PMA1 had previously been shown to contain 7q34 gain; the two further samples had not undergone copy number analysis. Copy number gains at 7q34, indicative of the *KIAA1549-BRAF* fusion, have been identified by CGH or SNP copy number analysis in 7/52 Grade II astrocytomas and 5/96 mixed astrocytic/oligodendroglial tumours studied to date (Jones *et al.*, 2008, Pfister *et al.*, 2008, Forsheew *et al.*, 2009, Sievert *et al.*, 2009). Pfister et al also identified duplication

of the *BRAF* locus without concurrent gain of chromosome 7 or 7q by FISH in 6 adult diffuse astrocytomas (Pfister *et al.*, 2008). Further studies are necessary to elucidate whether these grade II astrocytomas with *RAF* fusions are PAs with unusual histopathological features, a benign sub-group of non-pilocytic astrocytomas, which arose *de novo* or tumours which have progressed from PAs by accumulating additional, as yet unidentified, changes.

An *SRGAP3-RAF1* fusion variant was described by Jones *et al.* from further analysis of a discrete region of gain at chromosome 3p25 in a single pilocytic astrocytoma (Jones *et al.*, 2009). RT-PCR identified an in-frame fusion gene between *SRGAP3* exon 12 and *RAF1* exon 10. This *SRGAP3-RAF1* fusion variant was cloned, and conferred an anchorage independent growth phenotype and increased ability to phosphorylate endogenous MEK in NIH3T3 cells when compared to cells transduced with wild-type *RAF1*. The breakpoints described by Jones *et al.* are different to those found for the *SRGAP3-RAF1* fusion in this thesis. Here, the breakpoints are between *SRGAP3* exon 11 and *RAF1* exon 8. Intriguingly, an additional duplication of *SRGAP3* exon 11 is seen within this variant, creating an in-frame gene fusion with partial tandem duplication of *SRGAP3* exon 11. Both fusions retain the *RAF1* kinase domain. These are currently the only two *SRGAP3-RAF1* fusion variants to have been described in the literature.

In summary, this thesis has described the identification of an *SRGAP3-RAF1* and five *KIAA1549-BRAF* fusion variants: with the former and two of the latter previously undescribed in the literature. Combining these results with those from other studies, in-frame *KIAA1549-BRAF* fusion variants have been found in 76/95 (80%) sporadic pilocytic astrocytomas. The majority (59/76) are the *KIAA1549-BRAF* exon 16-exon 9 fusion variant. These results are summarised in Table 5.4.

Fusion gene variant	Relative frequency in PAs
<i>KIAA1549-BRAF</i> exon 16-exon 9	59/76
<i>KIAA1549-BRAF</i> exon 15-exon 9	10/76
<i>KIAA1549-BRAF</i> exon 16-exon 11	5/76
<i>KIAA1549-BRAF</i> exon 18-exon 10	1/76
<i>KIAA1549-BRAF</i> exon 19-exon 9	1/76
<i>SRGAP3-RAF1</i> exon11-exon 8	1/2
<i>SRGAP3-RAF1</i> exon 12-exon 10	1/2

Table 5.4 Summary of *RAF* fusion gene variants.

Fusion variants were identified in 78/95 pilocytic astrocytomas by three separate studies (Jones *et al.* 2008, Forshew *et al.* 2009 and Sievert *et al.* 2009). The study by Sievert did not include the additional exon for *KIAA1549*. Hence, the *KIAA1549-BRAF* exon 15-exon 9 fusions described have been recorded as *KIAA1549-BRAF* exon 16-exon 9 fusions.

5.6.1 Possible mechanisms for the development of fusions

The *KIAA1549-BRAF* and *SRGAP3-RAF1* fusions appear to have arisen from tandem duplication events. Two well-characterised mechanisms are used to repair DNA double strand breaks (DSB) in eukaryotes: homologous recombination (HR) and non-homologous end joining (NHEJ). Both processes have been described in the genesis of tandem duplications in cancer and other genetic disorders, although in this context HR becomes non-allelic homologous recombination (Batzer *et al.*, 2002, Gu *et al.*, 2008, Lieber *et al.*, 2008).

Non-allelic homologous recombination (NAHR) occurs between non-allelic repetitive sequences that may occur frequently throughout the genome, such as Alu elements (short repetitive sequences <500bp long), or low-copy repeats, which may only occur a few times and may range from one to hundreds of kilobases in size (Batzer *et al.*, 2002, Gu *et al.*, 2008, Hastings *et al.*, 2009). Homologous recombination is primarily used to repair DSBs during G2 or late S phase during the cell cycle, as the sister chromatid is used as a template. Sequences which lie between the regions that recombine may be duplicated or deleted, changing the DNA copy number. NAHR may give rise to deletions, duplications or inversions, depending on the orientation of the repetitive

sequences, and whether the rearrangement is intrachromatid, interchromatid or interchromosomal (Gu *et al.*, 2008). NAHR has been implicated in the partial tandem duplication of *MLL* in acute myeloid leukaemia (Strout *et al.*, 1998, Basecke *et al.*, 2006).

Genomic DNA sequencing was carried out for 9 PAs with the same *KIAA1549-BRAF* exon 16-exon9 fusion variant, with the expectation of finding a common site for the DNA breakpoints. However, sequencing revealed that the DNA breakpoints were distributed widely throughout intron 16 of *KIAA1549* and intron 8 of *BRAF*, with no apparent relationship to Alu elements or to any other repetitive sequences (Figure 5.11). Repetitive sequences are essential to create the origins for NAHR. Hence, these data suggest that the *KIAA1549-BRAF* fusion did not occur by Alu-Alu mediated recombination or NAHR.

Non-homologous end joining (NHEJ) has also been implicated in the creation of tandem duplications. NHEJ is used to repair DSBs, which may be physiological, for example in VDJ recombination to generate the T-cell receptor or pathological, for example in response to ionizing radiation (reviewed by Lieber) (Lieber, 2008). Unlike NAHR, NHEJ may repair DSBs during any stage of the cell cycle, as the process does not use a homologous chromosome as a template. When the DSB is detected in NHEJ, the protein heterodimer Ku binds to the broken DNA ends. Ku interacts with the nuclease Artemis-DNA-dependent protein kinase catalytic subunit (Artemis-DNA-Pkcs), polymerases μ and λ and a DNA ligase (XLF-XRCC4-DNA ligase IV), which appear to be capable of binding in any order. Binding of these elements can take place without Ku, but efficiency is improved by its presence. In essence, Ku functions as a 'work bench', where repair of the DSB can take place. Artemis-DNA-Pkcs resects any overhangs at the broken DSB ends, making them compatible in order to allow the DNA to re-anneal (Figure 5.17). This editing process may require deletion and/or addition of bases by the polymerases μ and λ , which creates an 'information scar', the hallmark of NHEJ (Gu *et al.*, 2008, Lieber, 2008).

Genomic rearrangements may arise through NHEJ by ligating DNA ends together from different DSBs: repetitive sequences are not required for the process to occur. The NHEJ mechanism is an attractive candidate for the tandem duplications identified in the *KIAA1549-BRAF* fusion. Two of the nine PA samples sequenced show insertions of 1

and 3 base pairs, which would be consistent with the ‘information scar’ created by the process. However, none of the samples show sequence deletions, and it is difficult to explain how the in-frame duplication of exon 11 in the *SRGAP3-RAF1* fusion could have occurred by this method.

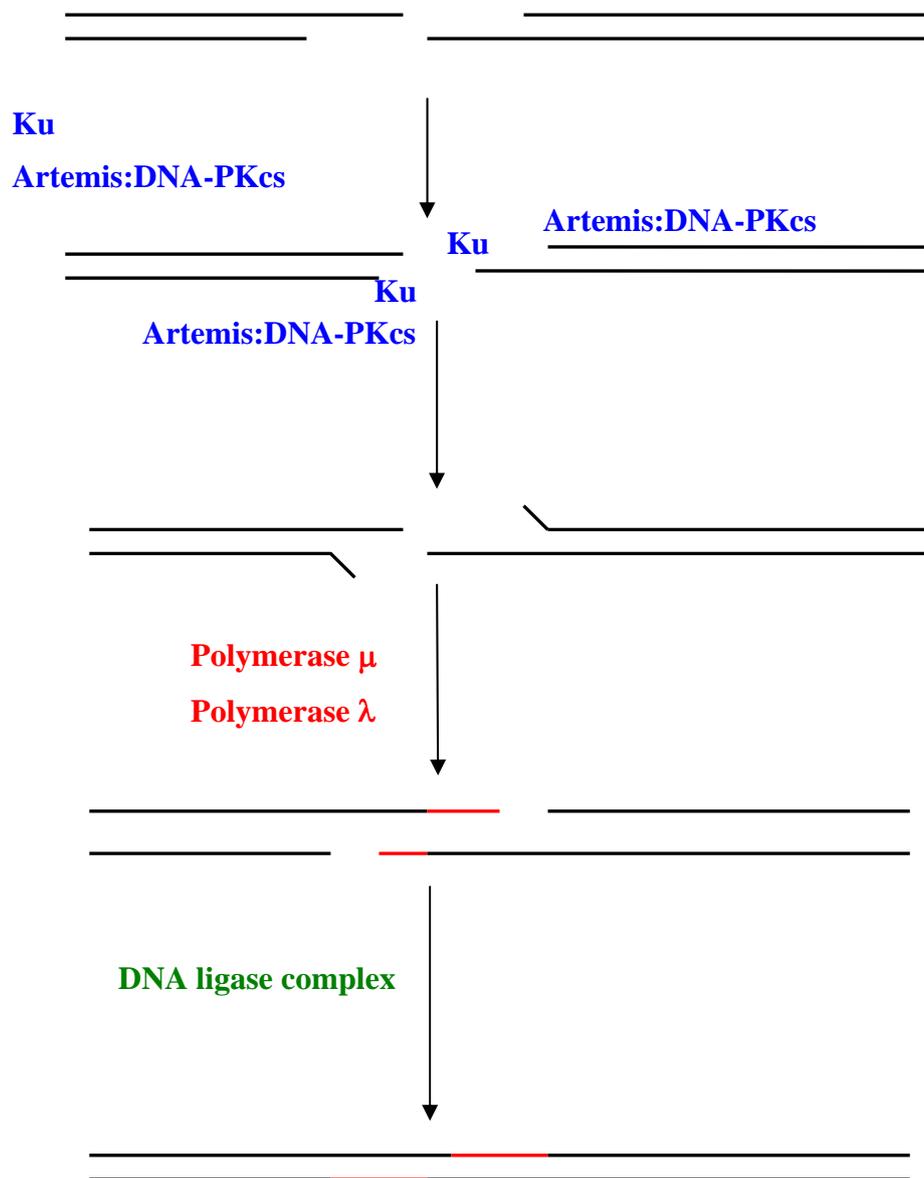


Figure 5.18 Non-homologous end-joining (NHEJ) repair of a double strand DNA break.

Ku binds to the broken DNA ends, and allows the Artemis complex to resect the overhanging ends. The polymerases μ and λ add nucleotides (shown in red) to complete the sequence and the DNA ligase complex then repairs the double stranded DNA. Figure adapted from Lieber, 2008 (Lieber, 2008).

A number of alternative mechanisms for the creation of genetic rearrangements, relating to errors in DNA replication have recently been proposed, however these are all currently hypothetical models. These include fork stalling and template switching (FoSTeS) and microhomology-mediated break-induced replication (MMBIR). FoSTeS suggests that stalling of the DNA replication fork may allow the ‘lagging’ DNA strand to detach from the original template (Figure 5.18) (Lee *et al.*, 2007). This ‘lagging’ DNA strand may then attach to another replication fork nearby through regions of microhomology, allowing DNA replication to continue. Switching to a replication fork downstream of the original would result in deletion, and switching to a replication fork upstream would create a region of duplication. The FoSTeS model proposes that the lagging strand may switch DNA templates multiple times before resuming DNA replication on the original template strand, however the reasons for this to occur are not entirely clear. A single FoSTeS event could explain the tandem duplication leading to the *KIAA1549-BRAF* fusion, and two FoSTeS events occurring during one round of replication could explain the duplicated exon 11 in the *SRGAP3-RAF1* fusion.

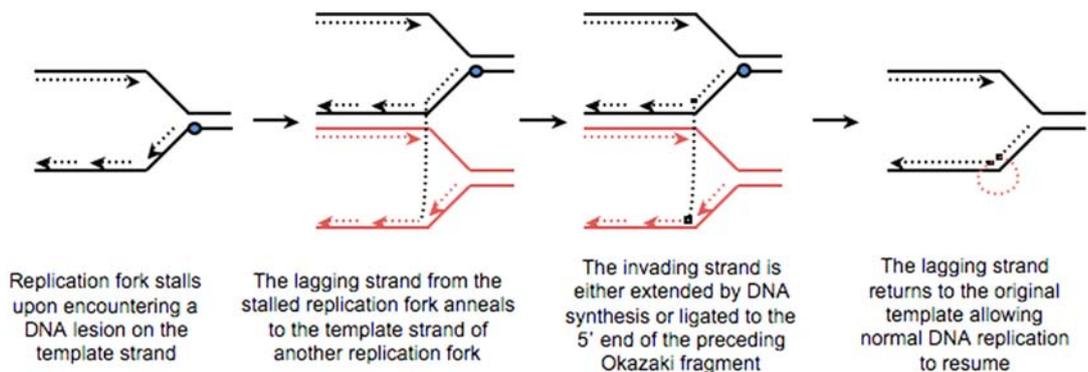


Figure 5.19 Proposed mechanism for FoSTeS (fork stalling template switching) to induce genomic rearrangements.

Figure adapted from Lee, 2007 (Lee *et al.*, 2007).

FoSTeS was originally described in an attempt to explain the complex genetic rearrangements in Pelisaeus-Merzbacher disease, a recessive X-linked genetic disorder affecting the central nervous system. This condition arises through complex nonrecurrent chromosome changes involving the *PLP1* gene. Microarray copy number analysis revealed sections of duplicated or tripled regions of DNA, interspersed with regions of deletion or normal copy number (Lee *et al.*, 2007). Regions of microhomology between 2 to 5 nucleotides long were found by sequencing the duplication junctions. Detailed analysis of the regions of microhomology within the

sequence data suggested DNA duplication had occurred in different orientations within the same DNA strand, leading to the development of the FoSTes model. This model is attractive to explain the genesis of the *RAF* fusion genes, however the situation in Pelisaeus-Merzbacher disease is not directly comparable. The genomic region surrounding *PLP1* contains numerous LCRs, which may act as foci for FoSTes to occur, which is not the case in the region surrounding *KIAA1549* and *BRAF*. Furthermore two of the samples, PA4 and PA7, do not show any region of microhomology at the fusion junction. In addition, the region of duplication involving the 7q34 region is generally the only copy number change present in an otherwise balanced genome, and there are no further regions of duplication or deletion nearby, which might be expected to be present if FoSTes had occurred.

In budding yeast, Payen et al found that segmental duplications arise from accidents in the replication process, via two alternative replication-based mechanisms (Payen *et al.*, 2008). The first involves break-induced DNA replication, by strand invasion of the sister chromatid, however only half of the segmental duplications seen in yeast contained homologous repeats at their junctions, suggesting another mechanism could also lead to segmental duplications. This was termed microhomology/microsatellite-induced replication (MMIR). In MMIR replication forks collapse and break down rather than stalling, before rejoining via regions of microhomology or at microsatellites (simple repeating sequences between 1-4 bases long (Weber *et al.*, 1989)). The MMIR mechanism has yet to be identified in humans.

A similar mechanism, micro-homology mediated break-induced replication (MMBIR), has also been suggested for the development of regions of copy number variation (Hastings *et al.*, 2009). Again this model was proposed to explain genetic rearrangements in Pelisaeus-Merzbacher disease, by the same group who proposed the FoSTes model. Here, collapse of a replication fork occurs rather than stalling, possibly due to hypoxic stress, creating a single strand DNA break. The free DNA end dissociates, and may attach and form a further replication fork on a nearby DNA strand through regions of microhomology, allowing DNA replication to continue. The proposed MMBIR mechanism is theoretical, and suggests that microhomology is required for the process to occur. There is, as yet, no experimental evidence for this mechanism in humans. At the moment it does not appear that MMBIR can be used to

explain the genesis of the *KIAA1549-BRAF* fusion in the two PAs with no regions of microhomology at the fusion junctions.

Both FoSTes and MMBIR are highly speculative. Currently it appears that none of these proposed mechanisms are able to explain how all of the fusion junctions identified by sequencing could have arisen. Hence, further investigations will be required to elucidate how the *RAF* gene fusions in PAs are formed.

5.6.2 MAPK pathway activation in low-grade astrocytomas is confirmed by Western blotting

Western blot analysis revealed high levels of phospho-MEK and phospho-ERK in eleven tumour samples when compared to two fetal brain controls, consistent with results from previous studies. No differences in the levels of MEK and ERK were found between the two groups. These results confirmed that activation of the MAPK pathway was present in tumours. Seven of the tumours contained *KIAA1549-BRAF* fusions and the tumour containing the *SRGAP3-RAF1* fusion were studied. However, three of the tumours examined did not, to our knowledge, contain any *RAF* gene fusions. Hence, the MAPK pathway was also activated in three tumours without gene fusions. These data led me to search for different mechanisms for MAPK activation in these samples, by sequencing for activating point mutations in *BRAF* and the three *RAS* genes. These data are discussed in chapter 6.

So far no groups, including ours, have demonstrated either the *KIAA1549-BRAF* or *SRGAP3-RAF1* fusion proteins by Western blot analysis. The predicted size for the most common *KIAA1549-BRAF* fusion would be 232 kDa if protein transcription commenced from the promoter within exon 1 of *KIAA154*. If transcription commenced from the promoter within intron 8, as in the short form described by Jones *et al.* the shorter predicted transcript would be 101 kDa (Jones *et al.*, 2008). The *SRGAP3-RAF1* fusion protein is predicted to be 99 kDa. These experiments are ongoing, but initial attempts have not, as yet, identified a protein band of the expected size. Efforts are currently hampered by the lack of a suitable antibody with which to probe for *KIAA1549*.

Chapter 6. Candidate gene sequencing in paediatric low-grade astrocytoma

6.1 Introduction

Gene fusions between *KIAA1549* and *BRAF* were identified in 30/32 pilocytic astrocytomas and 3/18 non-pilocytic astrocytomas (one cerebral diffuse astrocytoma, one hypothalamic pilomyxoid astrocytoma and one spinal pilomyxoid glioma). An additional gene fusion between *SRGAP3* and *RAF1* was identified in a single pilocytic astrocytoma. These gene fusions were shown to be associated with activation of the ERK/MAPK pathway, using Western blotting for MEK1/2, phospho-MEK1/2, ERK1/2 and phospho-ERK1/2. However, MAPK pathway activation was also identified in three diffuse astrocytomas that did not contain gene fusions, suggesting that the pathway could be activated by another mechanism in these cases. The single pilocytic astrocytoma, which did not contain the *KIAA1549-BRAF* or *SRGAP3-RAF1* fusions, was of particular interest in this respect, as were the 15/18 non-pilocytic astrocytomas without *KIAA1549-BRAF* fusions. The possibility of MAPK pathway activation through specific gene mutations was thus considered.

BRAF and *RAF1* are critical serine-threonine-specific protein kinases, activated downstream of *RAS* within the MAPK pathway. Malignant transformation is known to occur in many cancers, due to activating point mutations in *BRAF* and the *RAS* gene family (Downward, 2003). Hence, screening for possible activating mutations at known mutation hotspots in *KRAS*, *NRAS*, *HRAS* (exons 2 and 3) and *BRAF* (exons 11 and 15) was undertaken in 50 paediatric low-grade astrocytomas. Mutation screening was also performed for *CDKN2A*, *PTEN*, *IDH1* and *IDH2*, as aberrations in these genes have been implicated in the development of adult astrocytomas (Ichimura *et al.*, 1996, Rasheed *et al.*, 1997, Wang *et al.*, 1997, Raffel *et al.*, 1999, Furnari *et al.*, 2007, Parsons *et al.*, 2008, De Carli *et al.*, 2009, Watanabe *et al.*, 2009). *CDKN2A* was also selected for mutation screening as deletions were identified within the *CDKN2A* locus by SNP copy number analysis in two grade II astrocytomas. Neither of these tumours contained the *KIAA1549-BRAF* or *SRGAP3-RAF1* fusions.

The prevalence of mutations in members of the receptor tyrosine kinase (RTK) /PI3K pathway were recently elucidated by two large studies in adult glioblastoma, one from the Cancer Genome Atlas Research Network, and one from Williams Parsons et al (CGARN *et al.*, 2008, Parsons *et al.*, 2008). The Cancer Genome Atlas study investigated 91 glioblastomas (72 treated and 19 untreated), with matched normal tissue or peripheral blood for mutations in 601 selected genes. This study found that 86% of adult glioblastomas contained at least one genetic aberration in the RTK/PI3K pathway. The most common abnormalities within this pathway were mutations and amplifications in *EGFR* (45%), followed by mutations and homozygous deletions of *PTEN* (36%) and *NF1* (18%). Mutations in *RAS* family members were found in 2%. Aberrations in the p53 pathway were found in 87% of glioblastomas, and mutations or homozygous deletions of *CDKN2A* and *TP53* were found in 49% and 35% of these cases, respectively. Finally, aberrations in RB were seen in 78% cases, and mutations or homozygous deletions of *CDKN2A* and *CDKN2B* were found in 52% and 47% of these cases, respectively. Deregulation of all three signalling pathways was found in 74% glioblastomas studied here, suggesting that aberrations in all three pathways are required for glioblastoma development.

The study by Williams Parsons et al looked at 22 glioblastoma tumour samples, and sequenced 20,661 protein-coding genes (Parsons *et al.*, 2008). This study found the most commonly mutated genes in glioblastomas to be *TP53* (35%), *PTEN* (26%), *NF1* (15%) and *EGFR* (14%). *CDKN2A* showed homozygous deletions in 50% of glioblastomas studied. This study also identified recurrent mutations within the active site of the novel gene isocitrate dehydrogenase I (*IDH1*), as discussed above.

Both of these large studies were conducted in adult glioblastomas, which are highly rearranged, and appear to contain different molecular signatures when compared to those seen in paediatric low-grade astrocytomas. *EGFR* and *MDM2* amplification, mutations in *PTEN* and LOH of 10p are typical findings in primary glioblastomas, which tend to develop rapidly after a short clinical history with no evidence of a precursor lesion. Mutations in *TP53*, and LOH at 19q and 22q are more typical of secondary glioblastomas, which have arisen more slowly from a less malignant low-grade astrocytoma (Watanabe *et al.*, 2009).

6.2 Aim

The aim of the work described within this chapter was to investigate the presence of gene mutations in paediatric low-grade astrocytomas using DNA sequencing.

6.3 Materials and methods

Mutational analysis of *KRAS*, *NRAS*, *HRAS* (exons 2 and 3) and *BRAF* (exons 11 and 15) was undertaken by direct sequencing in 50 paediatric low-grade astrocytomas and one healthy male control, as described in sections 2.13.2-2.15 in chapter 2. Water was included as a negative control. The primers used are shown in Appendix C Table 8.5.

Mutational analysis of *CDKN2A*, *PTEN*, *IDH1* and *IDH2* was undertaken by direct sequencing in 41 paediatric low-grade astrocytomas, one healthy male and human glioma cell line U251. Water was included as a negative control. The primers used are shown in Appendix C, Table 8.6.

Target exons were amplified by PCR from whole-genome amplified (WGA) tumour DNA, followed by direct sequencing in both directions using the di-deoxy chain termination method on an ABI 3730 DNA sequencer. Results were screened by eye using the Applied Biosystems software packages Sequence Scanner version 1.0 and Variant Reporter version 1.0. Further mutational analysis was performed by eye.

6.4 Results

6.4.1 ERK/MAPK pathway member sequencing

50 paediatric low-grade astrocytomas were screened for activating mutations at known hotspots in *KRAS*, *NRAS*, *HRAS* and *BRAF*. The single cerebellar pilocytic astrocytoma (PA31), which did not contain any fusions in *KIAA1549-BRAF* or *SRGAP-RAF1*, was found to contain the G12A activating mutation in *KRAS* (Figure 6.1.A). This causes substitution of glycine to alanine in codon 12. Two grade II astrocytomas (1 diffuse astrocytoma, DA2, and the pleomorphic xanthoastrocytoma, PXA1) were found to contain the V600E activating mutation in *BRAF* (Figure 6.1.B). This leads to the substitution of valine by glutamic acid at position 600.

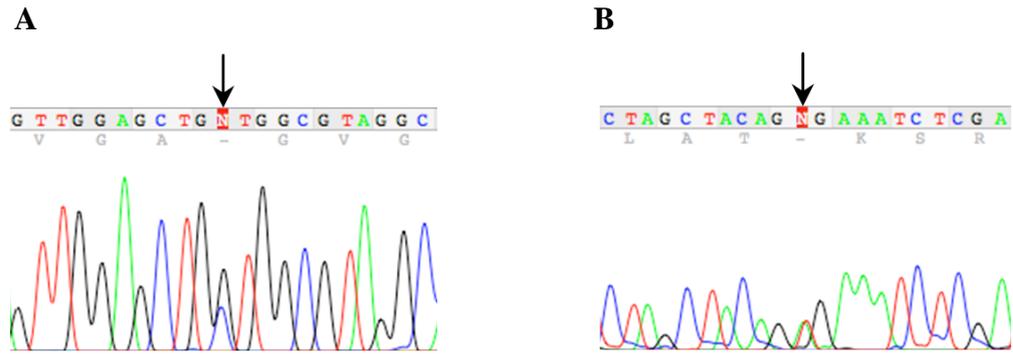


Figure 6.1 Activating mutations in *KRAS* and *BRAF*.

(A) *KRAS* G12A activating mutation (G to C) in tumour PA31.

(B) *BRAF* V600E activating mutation (T to A) in tumour PXA1.

The sequencing results for all the MAPK pathway members are summarised below for 32 pilocytic astrocytomas (Table 6.1) and 18 non-pilocytic astrocytomas (Table 6.2).

6.4.2 *CDKN2A* α sequencing

CDKN2A α exons 1, 2 and 3 were sequenced in 43 tumour samples, normal male DNA as a positive control and the human glioma cell line U251, as a negative control. U251 is known to contain homozygous deletion of *CDKN2A* (Fuxe *et al.*, 2000). Water was also included as a negative control. Sequencing was successful in the normal male DNA and failed, as expected, in U251, due to homozygous loss of *CDKN2A*.

Exon 1 sequencing was successful in all of the 41 tumour samples. Exon 2 sequencing was successful in 32 tumour samples, and failed in the following 9 samples: PA4, PA32, PA18, DA4, PA8, PMA1, PMG3, PA7 and PMG4. Exon 3 sequencing was successful in 36 tumour samples, and failed in the following 5 samples: PMG5, PA17, PA15, DA8 and PA13. No *CDKN2A* α mutations or variants were identified in any of the tumour samples or in the male DNA.

Name	Site	Age	Sex	KIAA1549- <i>BRAF</i> fusion	SRGAP3- <i>RAF1</i> fusion	MAPK pathway sequencing	Discrete copy Number changes
PA1	Cerebellum	20	F	Ex16-Ex9	-	-	7q34 gain
PA2	Cerebellum	4	M	Ex16-Ex9	-	-	7q34 gain
PA3	Cerebellum	7	M	Ex16-Ex9	-	-	7q34 gain
PA4	Cerebellum	2	M	Ex16-Ex9	-	-	7q34 gain
PA5	Cerebellum	6	F	Ex16-Ex9	-	-	7q34 gain
PA6	Cerebellum	5	M	Ex16-Ex9	-	-	7q34 gain
PA7	Cerebellum	13	M	Ex16-Ex9	-	-	Not run
PA8	Cerebellum	2	M	Ex16-Ex9	-	-	Not run
PA9	Cerebellum	14	F	Ex16-Ex9	-	-	7q34 gain
PA10	Cerebellum	19	M	Ex16-Ex9	-	-	7q34 gain
PA11	Cerebellum	14	M	Ex16-Ex9	-	-	7q34 gain
PA12	Cerebellum	14	M	Ex16-Ex9	-	-	7q34 gain
PA13	Cerebellum	10	M	Ex16-Ex9	-	-	7q34 gain
PA14	Cerebellum	10	M	Ex16-Ex9	-	-	7q34 gain
PA15	Cerebellum	13	M	Ex16-Ex9	-	-	7q34 gain
PA16	Cerebellum	10	M	Ex16-Ex9	-	-	-
PA17	Cerebellum	6	M	Ex16-Ex9	-	-	7q34 gain
PA18	Cerebellum	3	M	Ex16-Ex9	-	-	7q34 gain
PA19	Cerebellum	3	M	Ex16-Ex9	-	-	7q34 gain
PA20	Cerebellum	4	M	Ex16-Ex9	-	-	Not run
PA21	Cerebellum	12	F	Ex16-Ex9	-	-	Not run
PA22	Cerebellum	6	F	Ex15-Ex9	-	-	7q34 gain
PA23	Cerebellum	3	F	Ex15-Ex9	-	-	7q34 gain
PA24	Cerebellum	4	F	Ex15-Ex9	-	-	Not run
PA25	Cerebellum	4	F	Ex16-Ex11	-	-	7q34 gain
PA26	Cerebellum	3	F	Ex16-Ex11	-	-	7q34 gain
PA27	Cerebellum	14	F	Ex16-Ex11	-	-	7q34 gain
PA28	Cerebellum	7	M	Ex18-Ex10	-	-	Not run
PA29	Cerebellum	19	M	Ex19-Ex9	-	-	7q34 gain
PA30	Cerebellum	13	F	-	Ex11-Ex8	-	3p25 gain
PA31	Cerebellum	1	F	-	-	<i>KRAS</i> p.G12A	-
PA32	Brain stem	9	F	Ex16-Ex9	-	-	7q34 gain

Table.6.1 Summary of results for MAPK pathway member sequencing and fusion screening in 32 pilocytic astrocytomas.

Name	Pathology	Site	Age	Sex	KIAA1549- <i>BRAF</i> fusion	SRGAP3- <i>RAF1</i> fusion	MAPK pathway sequencing	Discrete copy number changes
DA1	Diffuse astrocytoma	cerebral cortex	5	F	Ex15-Ex9	-	-	Not run
DA2	Diffuse astrocytoma	cerebral cortex	9	M	-	-	<i>BRAF</i> p.V600E	9p21 loss
DA3	Diffuse astrocytoma	cerebral cortex	17	M	-	-	-	-
DA4	Diffuse astrocytoma	cerebral cortex	4	M	-	-	-	-
DA5	Diffuse astrocytoma	cerebral cortex	7	F	-	-	-	-
DA6	Diffuse astrocytoma	cerebral cortex	5	F	-	-	-	-
DA7	Diffuse astrocytoma	cerebral cortex	5	M	-	-	-	-
DA8	Diffuse astrocytoma	cerebral cortex	15	F	-	-	-	Not run
DA9	Diffuse astrocytoma	cerebral cortex	9	F	-	-	-	Not run
DA10	Diffuse astrocytoma	cerebral cortex	5	F	-	-	-	Not run
DA11	Diffuse astrocytoma	cerebral cortex	2	M	-	-	-	Not run
PMA1	Pilomyxoid astrocytoma	diencephalon	1	F	Ex15-Ex9	-	-	7q34 gain
PMG1	Pilomyxoid glioma	diencephalon	17	M	-	-	-	-
PMG2	Pilomyxoid glioma	diencephalon	7	F	-	-	-	-
PMG3	Pilomyxoid glioma	diencephalon	8	M	-	-	-	Not run
PMG4	Pilomyxoid glioma	spinal cord	12	M	Ex15-Ex9	-	-	Not run
PMG5	Pilomyxoid glioma	spinal cord	11	F	-	-	-	-
PXA1	Pleomorphic xanthoastrocytoma	cerebral cortex	6	F	-	-	<i>BRAF</i> p.V600E	9p21 loss

Table.6.2 Summary of results for MAPK pathway member sequencing in 18 Grade II astrocytoma samples.

Results for *KRAS*, *NRAS*, *HRAS* and *BRAF* are shown, with the presence of fusion genes and specific DNA copy number changes.

6.4.3 *PTEN* sequencing

Exons 1 to 9 of *PTEN* were sequenced in 41 tumour samples, normal male DNA as a positive control and the human glioma cell line U251. U251 is known to contain a homozygous insertion of TT within exon 7 of *PTEN*. This causes a frameshift within the translated protein sequence at glutamic acid 242, which is thought to cause an oncogenic mutation (Ikediobi *et al.*, 2006). Water was also included as a negative control.

The homozygous c.723_724insTT insertion mutation was identified successfully in U251 (Figure 6.2).

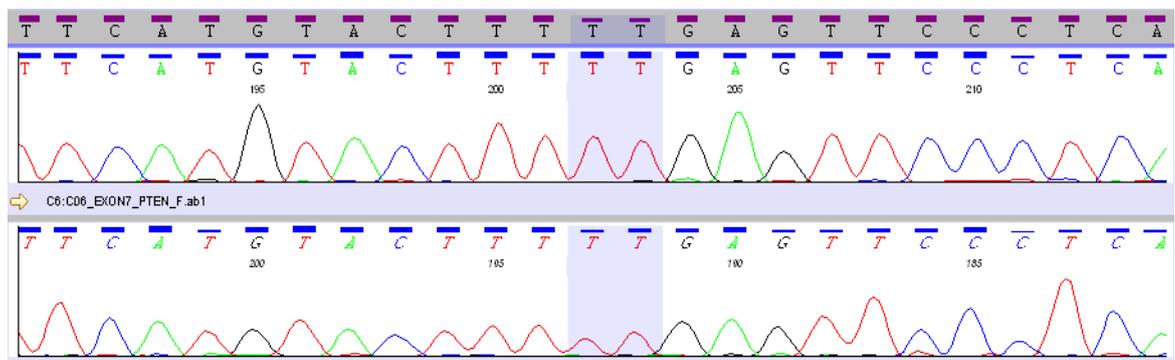


Figure 6.2 Homozygous insertion of TT within exon 7 of *PTEN* in cell line U251. The sequence traces for forward and reverse strands are shown.

Sequencing was successfully completed for all 41 samples in exons 2, 5, 7, 8 and 9 of *PTEN*. Sequencing failed in both directions for 1 sample in exon 1, 9 samples in exon 3, 3 samples in exon 4 and 1 sample in exon 6. A single polymorphism, where a single cytosine was substituted by thiamine, was identified within exon 1 in sample PA6. However, this change was within the intron of exon 1, eleven base pairs from the exon 1/intron boundary, and would not have altered the transcribed protein sequence. No other abnormal variants were identified. Results are shown in Table 6.3.

	Exon 1	Exon 2	Exon 3	Exon 4	Exon 5	Exon 6	Exon 7	Exon 8	Exon 9
PA1	-	-	X	-	-	-	-	-	-
PA2	-	-	X	-	-	-	-	-	-
PA3	-	-	-	-	-	-	-	-	-
PA4	-	-	X	-	-	-	-	-	-
PA5	-	-	-	-	-	-	-	-	-
PA6	C > T	-	-	-	-	-	-	-	-
PA7	-	-	-	-	-	-	-	-	-
PA8	-	-	X	-	-	-	-	-	-
PA9	-	-	X	X	-	-	-	-	-
PA10	-	-	-	-	-	X	-	-	-
PA11	-	-	-	-	-	-	-	-	-
PA13	-	-	-	-	-	-	-	-	-
PA14	-	-	-	-	-	-	-	-	-
PA15	-	-	X	-	-	-	-	-	-
PA16	-	-	-	-	-	-	-	-	-
PA17	-	-	-	-	-	-	-	-	-
PA18	-	-	-	-	-	-	-	-	-
PA19	-	-	-	-	-	-	-	-	-
PA20	-	-	-	-	-	-	-	-	-
PA22	-	-	-	-	-	-	-	-	-
PA23	-	-	-	-	-	-	-	-	-
PA25	-	-	-	-	-	-	-	-	-
PA26	-	-	-	-	-	-	-	-	-
PA27	X	-	-	-	-	-	-	-	-
PA29	-	-	-	-	-	-	-	-	-
PA31	-	-	-	-	-	-	-	-	-
PA32	-	-	-	-	-	-	-	-	-
PMA1	-	-	-	-	-	-	-	-	-
DA1	-	-	-	-	-	-	-	-	-
DA2	-	-	-	-	-	-	-	-	-
DA3	-	-	-	-	-	-	-	-	-
DA4	-	-	-	-	-	-	-	-	-
DA7	-	-	X	X	-	-	-	-	-
DA8	-	-	X	-	-	-	-	-	-
DA9	-	-	-	-	-	-	-	-	-
PMG1	-	-	-	-	-	-	-	-	-
PMG2	-	-	-	-	-	-	-	-	-
PMG3	-	-	-	-	-	-	-	-	-
PMG4	-	-	X	-	-	-	-	-	-
PMG5	-	-	-	-	-	-	-	-	-
PXA1	-	-	-	X	-	-	-	-	-
U251	-	-	-	-	-	-	TT ins	-	-
Control	-	-	-	-	-	-	-	-	-
H ₂ O	X	X	X	X	X	X	X	X	X

Table.6.3 Sequencing results for *PTEN*

Exons 1-9 were sequenced in 41 tumour samples, U251, male control and water.

X = sequencing failed in both directions, - = no abnormality detected.

The c.723_724insTT mutation was detected in U251. The C>T substitution within PA6 is within the intronic sequence of exon 1, so would not have altered the protein sequence.

6.4.4 *IDH1* and *IDH2* sequencing

IDH1 exon 4 and *IDH2* exons 4 and 5 were sequenced in 41 tumour samples, normal male DNA as a positive control and the human glioma cell line U251. *IDH1* exon 4 was sequenced successfully in all samples except PXA1. A variant sequence within the coding region of exon 4 was found in 3 samples (DA4, DA8 and PA17). This corresponds to a known SNP, rs11554137, which does not alter the amino acid sequence. No other abnormal variants were detected. *IDH2* exons 4 and 5 were successfully sequenced in all of the 41 tumour samples, and no abnormal variants were identified.

6.5 Discussion

6.5.1 MAPK pathway sequencing

Direct sequencing identified activating mutations within *KRAS* and *BRAF* in three tumour samples. Hence, alternate mechanisms were identified for activation of the MAPK pathway, as none of these three tumours contained the *KIAA1549-BRAF* or *SRGAP3-RAF1* gene fusions. This was particularly important in the case of the single pilocytic astrocytoma, PA31, which was found to contain the *KRAS* G12A activating mutation. This result confirmed that MAPK pathway activation was present in 100% of the 32 pilocytic astrocytomas examined for this thesis. The discovery of genetic changes leading to activation of a molecular pathway in an entire cohort of pilocytic astrocytomas was unprecedented.

Tumours containing *BRAF*- and *KRAS*-activating mutations were separate from those containing *BRAF* or *RAF1* gene fusions. Earlier work presented in this thesis has shown that gene fusions are associated with MAPK pathway activation in pilocytic astrocytomas. Activating mutations in *KRAS* and *BRAF* are uncommon in low-grade astrocytomas. However, their presence highlights the fact that activation of the MAPK pathway, however this may be achieved, is of critical importance in the development of low-grade astrocytomas.

RAS proteins have an essential role in regulating normal rates of cellular proliferation, and were some of the first proteins to be identified as constitutively activated in tumours due to point mutations in their coding sequences (Bos, 1989). It is estimated that ~20% of all tumours contain an activating mutation in one of the three *RAS* genes. The majority of these mutations (85%) are within *KRAS*, a further 15% are within *NRAS*, and ~1% in *HRAS* (Downward, 2003). *RAS* mutations are most commonly found in pancreatic, lung, thyroid and colorectal cancers. Activating mutations in *KRAS* at codon 12 (G to C transversion, resulting in amino acid change from glycine to alanine), codon 13 (G to C transversion, resulting in amino acid change from glycine to arginine) and codon 61 (A to T transversion, resulting in amino acid change from glutamic acid to glutamine) have previously been identified in paediatric pilocytic astrocytomas. However, these are rare events, seen in around 4-7% tumours (Maltzman *et al.*, 1997, Sharma *et al.*, 2005, Janzarik *et al.*, 2007). This incidence is slightly higher than the 3% incidence found in the 32 pilocytic astrocytomas examined here, where a single *KRAS* G12A mutation was found.

Mutations in *BRAF* have also been associated with many cancers, most commonly malignant melanoma, colorectal cancer, and lung cancer. The majority of mutations within *BRAF* involve the substitution of valine for glutamate at residue 600 (V600E) within the kinase domain. The resulting mutant BRAF protein has kinase activity elevated tenfold from wild-type BRAF, conferring a huge advantage in the transformation of NIH3T3 cells (Davies *et al.*, 2002). A low incidence rate, between 4-6%, for *BRAF* V600E activating mutations has also been found in paediatric grade I-II astrocytomas (Jones *et al.*, 2008, Pfister *et al.*, 2008, Sievert *et al.*, 2009). These mutations were shown to be somatic in 5 patients (3 PA, 2 DA), where germline blood did not show the V600E mutations (Pfister *et al.*, 2008, Sievert *et al.*, 2009). A novel *BRAF* activating mutation was recently identified in a single pilocytic astrocytoma, where insertion of 3 nucleotides in the DNA sequence leads to the addition of a threonine residue close to the mutational hotspot at position 600 (Jones *et al.*, 2009). This mutation conferred constitutive kinase activity and transformed NIH3T3 fibroblasts, however it was not identified in the set of patient samples examined for this thesis.

From the three samples with activating mutations identified by this study, it appears that low-grade astrocytomas do not contain concurrent *BRAF* and *KRAS* activating

mutations. Concurrent mutations in *BRAF* and *NRAS* have been found in studies on malignant melanoma and naevi, although these are rare (Pollock *et al.*, 2003, Goel *et al.*, 2006). The majority of benign melanocytic naevi show mutations in *BRAF*, but most do not progress to malignant melanoma. The prognostic significance of mutations in *BRAF* within benign naevi is unclear, as most studies do not appear to show any significant link with malignant transformation or patient outcome. In melanoma, it appears that further changes, in addition to mutations in *BRAF*, are required for malignant transformation (Goel *et al.*, 2006). It remains to be seen if this is also the case in paediatric low-grade astrocytoma.

6.5.2 *CDKN2A* sequencing

Two proteins are produced from the *CDKN2A* locus, using alternate first exons. Exons 2 and 3 are common to both transcripts. p16^{INK4A} is encoded from *CDKN2A α* , and p14^{ARF} by transcript *CDKN2A β* (Stott *et al.*, 1998). Both transcripts of *CDKN2A* are negative regulators of the cell cycle, affecting different pathways. p16^{INK4A} inhibits the retinoblastoma (Rb) pathway by inhibition of CDK4, p14^{ARF} inhibits the p53 pathway via MDM2. Deletion of *CDKN2A* was identified by SNP array analysis in two grade II tumour samples, as described in chapter 4 of this thesis. Hence, *CDKN2A* was selected as a candidate for sequencing.

Sequencing was performed successfully for *CDKN2A α* exons 1-3. Primers were also designed for sequencing the alternate first exon in *CDKN2A β* . Unfortunately attempts to sequence this exon were unsuccessful, despite multiple experiments, and will be repeated in the future. Previous studies have shown, however, that mutations within *CDKN2A* are most commonly situated within exon 2, which is shared by both *CDKN2A α* and *β* (Ichimura *et al.*, 2000). *CDKN2A* exon 2 was sequenced successfully in 32/41 tumour samples.

Homozygous deletion of *CDKN2A α* (p16^{INK4A}) has been identified in many studies of high-grade astrocytomas in adults and children (Ichimura *et al.*, 1996, Raffel *et al.*, 1999, Furnari *et al.*, 2007, CGARN *et al.*, 2008, Parsons *et al.*, 2008). However, many studies of adult astrocytomas have shown that inactivation of the RB pathway through loss of *CDKN2A α* alone is not sufficient to initiate malignant transformation in

astrocytomas (Furnari *et al.*, 2007). A recent microarray study, comparing tumour tissue samples from 16 adult patients with low-grade gliomas who subsequently developed high-grade gliomas found that loss of the *CDKN2A α* locus at 9p21 was the most frequent change seen following malignant transformation (Idbaih *et al.*, 2008). These data suggest that *CDKN2A α* loss is a major factor involved in secondary glioblastoma formation in adult low-grade glioma patients, but it does not appear to be an initiating factor in low-grade astrocytoma formation.

No mutations or variants were identified in *CDKN2A α* by direct sequencing in the tumour samples studied for this thesis. From these results, and those from previous studies, it appears that *CDKN2A α* does not play a significant role in the development of low-grade paediatric astrocytomas.

Both *CDKN2A* deletion and *BRAF* V600E activating mutations were identified in two grade II astrocytomas (DA2 and PXA1). Both patients were relatively young, aged 9 and 6 years, respectively. It is possible that these grade II tumours display more genetic rearrangements, and that these two samples had acquired a second aberration. One sample, DA2, contained homozygous loss of *CDKN2A*, the second, PXA1 had lost a single copy of *CDKN2A* as identified by SNP copy number analysis. Both samples were sequenced successfully, and no *CDKN2A α* mutations were identified in either sample. The fact that sequencing was successful for *CDKN2A α* in both samples with copy number loss, particularly in the sample with homozygous loss of *CDKN2A* by SNP analysis suggests that some normal stromal tissue was present in the tumour samples prior to DNA extraction. Hence, these sequencing results may in fact represent germline DNA from normal tissue within the sample rather than purely tumour DNA. Alternatively the lack of an identified *CDKN2A α* mutation in the remaining copy within sample PXA1, which contains single copy loss by SNP analysis could also suggest that this gene is silenced by other mechanisms, for example by DNA hypermethylation, which is known to occur in lung and breast cancer (Baylin *et al.*, 2006). The effects of haploinsufficiency may also be important here, where reduction in gene dosage caused by loss of one allele may cause a phenotypic change leading to tumorigenesis (Santarosa *et al.*, 2004). *CDKN2A* is a cyclin dependent kinase, which acts as a cell cycle checkpoint. Hence, haploinsufficiency here may facilitate cell cycle progression. Initial studies of homozygous deletions involving *CDKN2A* in urothelial cell

carcinomas indicate that haploinsufficiency may affect tumour invasion and disease recurrence (Chapman *et al.*, 2005).

Further research is necessary to determine what additional genetic changes are present in histological subtypes of low-grade astrocytoma in larger patient sets. The two patients identified here with both *BRAF* V600E activating mutations and deletion of *CDKN2A* had tumours with very different histopathological features. As yet unidentified genetic changes may help to uncover the molecular biology underlying different histological subtypes of grade II astrocytoma.

6.5.3 *PTEN* sequencing

The tumour-suppressor gene phosphatase with tensin homology (*PTEN*) is mutated or deleted in many tumours, including glioblastomas, breast cancer and melanoma. Mutations in *PTEN* are common in high-grade astrocytomas, but have not previously been identified in low-grade astrocytomas in adults or children (Rasheed *et al.*, 1997, Wang *et al.*, 1997, Yan *et al.*, 2009). *PTEN* encodes a protein containing a phosphatase domain, with both lipid and protein phosphatase activity. The phosphatase domain is encoded by exon 5, where ~30% germline and somatic mutations are found (Bonneau *et al.*, 2000). *PTEN* is a negative regulator of the phosphatidylinositol 3-kinase (PI3K) cell-signalling pathway. The protein inactivates phosphatidylinositol (3,4,5)-trisphosphate (PIP₃) by removing the phosphate at position D3, regenerating phosphatidylinositol (3,4)-biphosphate (PIP₂), which functions as an “on-off” switch for the pathway (Sansal *et al.*, 2004). PIP₃ recruits kinases from the protein kinase B (Akt) family. Many downstream targets are activated by Akt, including Forkhead, the Bcl-2-associated death promoter (BAD) and mTOR, which are involved in critical functions including cell proliferation, apoptosis, and DNA damage repair (Wu *et al.*, 2003, Sansal *et al.*, 2004). When *PTEN* is inactivated by mutation, deletion or promoter hypermethylation, PI3K pathway activation occurs. *PTEN* may also downregulate MAPK pathway signalling. Dephosphorylation of the adaptor proteins FAK and Shc by *PTEN* leads to decreased MEK and MAPK activity. Studies in malignant melanoma have indicated that *PTEN* loss and oncogenic *RAS* activation may occur in a reciprocal manner (Wu *et al.*, 2003). Hence, the possibility of loss of function mutations in *PTEN* leading to tumorigenesis was a further mechanism to explore within the paediatric astrocytoma tumour samples, particularly in the grade II tumours without gene fusions.

No loss-of-function mutations were detected in any of the tumour samples studied. Hence, it appears that loss of function of *PTEN* is not implicated in the development of grade II astrocytomas in children, in agreement with the findings from previous studies.

6.5.4 *IDH1* and *IDH2* sequencing

Recurrent mutations within the active site of the gene isocitrate dehydrogenase I (*IDH1*), on chromosome 2q33 were recently discovered in 12% of glioblastoma patients (Parsons *et al.*, 2008). The *IDH1* mutations were more prevalent in young patients, with a mean age of 33 years compared to patients with wild-type *IDH1*, whose mean age was 53 years. This finding was highly novel and was particularly exciting, as the mutations appeared to be more prevalent in young patients. Isocitrate dehydrogenase catalyses the production of α -ketoglutarate from isocitrate, producing nicotinamide adenine dinucleotide phosphate (NADPH). Mutations occurred at codon 132 within exon 4 of *IDH1*, resulting in an amino acid sequence change from arginine to histidine. An equivalent residue is also present in *IDH2*, at codon 172.

Hence, mutation analysis by direct sequencing of the relevant exons in *IDH1* and *IDH2* was performed in the set of paediatric low-grade astrocytomas investigated in this thesis. No mutations were found, suggesting that mutations in *IDH1* and *IDH2* are not implicated in the development of paediatric low-grade astrocytomas. This result was subsequently confirmed by more recent studies, which have found *IDH1* mutations within >70% WHO grade II and III astrocytomas, oligodendrogliomas and secondary glioblastomas which had evolved from lower-grade tumours in adults (Watanabe *et al.*, 2009, Yan *et al.*, 2009). Yan *et al.* studied 445 central nervous system tumours. The majority were from adults, but 52 paediatric astrocytoma samples were included. These comprised 21 pilocytic astrocytomas, 2 sub-ependymal giant cell astrocytomas (SEGA – WHO grade I), 14 WHO grade II astrocytomas and 15 paediatric glioblastomas (grade IV) (Yan *et al.*, 2009). *IDH1/2* mutations were identified in a small number of paediatric astrocytomas: *IDH1* mutations were found in 4/14 WHO grade II astrocytomas and 1/14 grade II astrocytomas contained an *IDH2* mutation. No *IDH1/2* mutations were found in 21 pilocytic astrocytomas, or in 15 paediatric glioblastomas.

The paediatric cases with mutations in *IDH1/2* were older than those without mutations; median ages were 17 years and 5 years, respectively ($p=0.002$). These findings were complemented by a further study by De Carli et al., who found *IDH1* mutations in 4/73 (5%) children with non-pilocytic astrocytomas (De Carli *et al.*, 2009). This group did not identify any *IDH2* mutations in children. Paediatric patients with *IDH1* mutations were again found to be older than children without mutations in *IDH1*. In this study median ages were 16 years for patients with mutations, and 7 years for patients without mutations, respectively ($p=0.002$). These data suggest that low-grade astrocytomas from adolescents contain similar biological markers to those seen in adult astrocytomas. Paediatric low-grade astrocytomas remain a distinct biological group from those seen in adults and adolescents.

6.6 Conclusion

In conclusion, this chapter has identified activating mutations within the MAPK pathway members *BRAF* and *KRAS* in a small but highly significant number of patients with low-grade astrocytoma. The discovery of an activating mutation in *KRAS* in the only pilocytic astrocytoma without gene fusions confirmed that MAPK pathway associated defects were present in 100% of the 32 pilocytic astrocytomas examined in this thesis.

The high frequency of MAPK pathway activation by genetic aberrations in all of the pilocytic astrocytomas studied suggests that this pathway is of critical importance in the development of these tumours. The extent of MAPK pathway activation in low-grade astrocytomas was previously unknown. This exciting finding provides the opportunity to exploit an entirely novel therapeutic target in this patient group. Fortuitously, activation of the MAPK pathway is a common abnormality in many adult cancers, including lung cancer and malignant melanoma. Hence, drug development aimed at producing new agents to target the MAPK pathway is already in progress. Specific inhibitors of the MAPK pathway are currently being tested in clinical trials for use in adult lung cancer and melanoma. This will be discussed in further detail in the final chapter of this thesis.

Additional sequencing of *CDKN2A*, *PTEN* and *IDH1/2* did not identify any mutations within the tumour samples. Preliminary studies to investigate *TP53* by sequencing were also undertaken for this thesis. A large number of exons were completed. However, sequencing was unsuccessful for exons 4 and 7, which contain the two main mutation hotspots. These experiments will be repeated in the future, and may provide further information on genetic aberrations present in grade II astrocytomas.

Chapter 7. Final discussion

7.1 Review of thesis aims and findings

The initial aim for this thesis was to study DNA copy number changes in paediatric low-grade astrocytomas using array CGH, the best available technique at the time the study began. I tested the hypothesis that differences in DNA copy number were present between pilocytic astrocytomas arising in the cerebellum and grade II astrocytomas from the cerebrum. There were a variety of pathologies within the group of grade II astrocytomas, so conclusions could only be drawn regarding the presence or absence of changes in pilocytic astrocytomas from the cerebellum when compared to the other group. No data were found to support the hypothesis of a difference in DNA copy number between the two groups by array CGH. Forty percent of the samples tested showed an apparently balanced karyotype. A few recurrent regions of discrete copy number change were, however, identified in a small number of samples.

When SNP array technology became available and began to supersede array CGH, the study was extended at higher resolution using Affymetrix 250K and 6.0 SNP arrays. None of the regions of interest suggested by array CGH analysis were confirmed by the SNP arrays. However, for the very first time a discrete region of copy number gain was identified at chromosome 7q34, primarily although not exclusively in pilocytic astrocytomas of the cerebellum. This region was initially identified using the 250K SNP arrays. The 7q34 gain and further regions of discrete copy number change were identified in additional samples by 6.0 SNP analysis.

Duplication of *BRAF* was confirmed within the gained region at 7q34 by interphase FISH in paraffin tissue samples. Further analysis of this region, by PCR and sequencing, demonstrated the presence of gene fusions between *KIAA1549* and *BRAF*. Five *KIAA1549-BRAF* fusion variants were subsequently identified. A further gene fusion between *SRGAP3* and *RAF1* was also found in a single tumour with DNA copy number gain at chromosome band 3p25. Both fusion genes lacked the auto-inhibitory domains of *BRAF* and *RAF1*, which were replaced in-frame by N-terminal segments of *KIAA1549* and *SRGAP3*, respectively, conferring constitutive kinase activity.

Sequencing of the DNA breakpoints for 9 tumours with the same *KIAA1549-BRAF* fusion variant, suggested that the *RAF* fusions may arise through the formation of tandem duplications and non-homologous end-joining (NHEJ) within the same chromosome. Further investigations are necessary to consider the important questions of how, where and when these fusions arise within the developing brain. Currently these questions are beyond the scope of this thesis, but the data presented here may begin to provide some insights into how to address these issues.

Sequencing also confirmed the presence of activating mutations in *KRAS* and *BRAF* in some tumours where gene fusions had not been found. Both the gene fusions and the activating mutations were shown to be associated with activation of the ERK/MAP kinase pathway by Western blotting.

Surprisingly, gene fusions or activating mutations were identified in 100% pilocytic astrocytomas studied. This was an unprecedented finding considering that pilocytic astrocytomas had previously been thought to contain very few genetic changes, with the majority of samples analysed previously showing apparently normal karyotypes. In addition, gene fusions and activating mutations were found in 28% of grade II astrocytomas, highlighting the importance of the ERK/MAPK pathway in the development of other paediatric low-grade gliomas. These findings have been complemented by other studies in the literature, although few studies have looked at significant numbers of grade II astrocytomas so far (Jones *et al.*, 2008, Pfister *et al.*, 2008, Forsheew *et al.*, 2009, Sievert *et al.*, 2009). It is currently unclear whether the grade II astrocytomas with *RAF* fusions are in fact pilocytic astrocytomas with unusual histopathological features, a benign sub-group of *de novo* grade II astrocytomas or a group of tumours that have progressed from pilocytic astrocytomas.

Very few published studies have investigated all known methods of activation for the MAPK pathway in low-grade astrocytomas. Furthermore, additional novel mechanisms for MAPK pathway activation, such as *ARAF* fusions may be as yet unidentified in these tumours. This raises a further question, whether different mechanisms of MAPK activation can be used as diagnostic or prognostic markers for children with low-grade astrocytomas. Pilocytic astrocytomas associated with NF-1 have been shown to be more indolent than sporadic PAs (Rubin *et al.*, 2005), and are often treated with a “watch and wait” strategy, as up to 50% may not progress further (Hargrave, 2009). Mutations

within the *RAS* genes may affect branches of the MAPK pathway that are unaffected by alterations in *BRAF*. Are there differences in clinical outcomes for these patients, depending on whether the MAPK pathway is activated by a gene fusion or an activating mutation? Do the different variants of the *KIAA1549-BRAF* fusion result in differences in kinase activity? Collaborative studies investigating large cohorts of well-characterised tumour samples with corresponding clinical information are needed to address these questions, which may lead to the development of molecular diagnostic and prognostic criteria for low-grade astrocytomas.

Sequencing was performed for the genes *CDKN2A*, *PTEN* and *IDH1/2* to assess the frequency of abnormalities in paediatric low-grade astrocytoma. These genes have all previously been found to contain aberrations within adult high-grade astrocytomas. To date, no significant aberrations have been identified in the paediatric astrocytoma samples examined. This confirms previous findings in adult and paediatric astrocytomas, where distinct molecular changes appear to vary depending on patient age.

The additional regions of copy number change identified by SNP array analysis are awaiting further validation and investigation. The sample numbers in the study cohort will be increased to include more grade II astrocytomas, and I look forward to continuing with this work following the completion of my PhD.

There were 13 grade II astrocytomas that did not contain *RAF* fusions or activating mutations. Other mechanisms may underlie tumorigenesis in this group, and further studies are required to investigate the role of epigenetic changes or aberrations in microRNA profiles in this sample group.

The relative specificity of the *KIAA1549-BRAF* fusion for pilocytic astrocytomas has led to speculation on the possibility of FISH being used as a diagnostic test to identify the fusion. A recent paper by Korshunov et al. described a combined diagnostic approach, using FISH for *BRAF* and *KIAA1549* to detect *KIAA1549-BRAF* fusions and sequencing for *IDH1* to distinguish between pilocytic astrocytomas and diffuse astrocytomas (Korshunov *et al.*, 2009). This study used FITC-labelled RP11-355D18 probe (green), situated between introns 8-9 and 16-17 on *KIAA1549*, and digoxigenin-labelled RP4-726N20 probe (red), just downstream from the 3' end of *BRAF*, for

interphase FISH analysis. A tissue microarray was prepared, using 5µm paraffin sections from 70 pilocytic astrocytomas and 50 diffuse astrocytomas. Signals were scored by two investigators in >100 non-overlapping nuclei. Samples with 7q34 gain were defined as having >10% nuclei containing 3 or more signals for *KIAA1549* and *BRAF*. Samples with the fusion were defined as those showing co-localisation of a red and green probe resulting in a yellow signal, however it was unclear from the method if this was the presence of a single yellow signal within a sample or co-localisation in >10% nuclei, as for 7q34 gain. The Korshunov study found 7q34 gain in 52/70 pilocytic astrocytomas and 31/50 diffuse astrocytomas. Co-localisation of signals, indicating the presence of *KIAA1549-BRAF* fusions were seen in 49/70 pilocytic astrocytomas and in no diffuse astrocytomas. *IDH1* mutations were only identified in diffuse astrocytomas (41/50), leading the authors to suggest that a combination of FISH and mutation analysis for *IDH1* could be used to distinguish between pilocytic and diffuse astrocytomas. The inclusion criteria for both 7q34 gain or the presence of a co-localised fusion signal appear to be very low in this study, considering that the technique visualises interphase nuclei in 2D where partial nuclei may be present above or below a selected nucleus within the section. In similar circumstances I would have used a proportion of >50% nuclei with 3 signals to imply 7q34 gain or the presence of a fusion, or to have considered the use of 3D FISH to be sure that there were no possible artifacts from partially overlapping nuclei.

An attempt was made to develop a technique to identify the *KIAA1549-BRAF* fusion by interphase FISH, in collaboration between our laboratory and David Ellison and his team at St Jude in Memphis. BAC probes were selected within the 5' proximal and 3' distal portions of *BRAF* (RP11-837G3 and RP11-948019), and within the 5' proximal portion of *KIAA1549* (RPL11-148L5). It was hoped that these probes would enable us to identify separation between the proximal and distal portions of *BRAF*, and subsequently to identify co-localisation of probes between *KIAA1549* and *BRAF* to indicate the presence of a fusion gene. *KIAA1549* and *BRAF* are separated by ~ 1.72Mb, and the probes used were separated by 1.65Mb. Normally a 1.65MB distance between probes would be expected to give good signal separation in interphase FISH. Unfortunately this was not the case in our experience, and neither technique proved to be sufficiently sensitive to enable it be used as a diagnostic test for the *KIAA1549-BRAF* fusion in paraffin tissue sections (data not shown). Many overlapping nuclei were seen, which gave confusing and highly variable signal patterns. Very occasionally co-

localised signals were seen, indicating the presence of the fusion, but these were rare events and did not occur with sufficient frequency in samples known to contain a fusion to be used with confidence as a diagnostic test. Korshunov et al. suggest that FISH would be particularly suitable to detect the presence of the fusions in small samples, obtained by stereotactic biopsy. In our experience this would not be possible using interphase FISH alone.

It is possible that there might be a specific chromatin configuration, such as a loop within the region, which may complicate the probe separation for these genes when studied by FISH. It is also possible that if such a configuration were present this might be implicated in the genesis of the fusion. The two genes may be brought together spatially during a particular phase of growth or development, perhaps during specific periods of rapid brain growth in the developing embryo. If DNA double strand breaks should occur during this time of spatial proximity, repair by NHEJ could conceivably bring about the creation of the fusion. This process could be associated with increased levels of gene expression or DNA replication, for example both genes may be brought together within the same transcription machinery or replication factory (Ottaviani *et al.*, 2008). There is currently no information on chromatin folding within this genomic region, however, and this will be an interesting avenue for further study in an attempt to understand how the fusions might arise.

Currently I feel that nested PCR for the *KIAA1549-BRAF* fusion using cDNA from fresh or frozen tumour samples provides the most reliable diagnostic test for identification of the fusions. This technique is also potentially possible for cDNA generated from paraffin samples, which would allow retrospective studies including larger sample numbers to be conducted. Successful studies have been conducted on RNA extracted from formalin-fixed paraffin embedded archival tissue (Specht *et al.*, 2001, Macabeo-Ong *et al.*, 2002). It would also be interesting to consider whether evidence of the fusion could be detected from the peripheral blood of patients with fusion positive tumours.

7.1.1 Targeted therapy against the MAPK pathway in low-grade astrocytomas

One of the most exciting outcomes to emerge from the findings in this thesis, and from other similar studies has been the identification of the MAPK pathway as a possible therapeutic target in low-grade astrocytomas. Surgery is currently the best available treatment for low-grade astrocytomas in childhood, but for some patients this may not be possible, for example where tumours are sited in the brainstem, optic pathways or hypothalamus. Overall survival rates for children with low-grade gliomas are excellent, but around 50% of patients will show evidence of tumour progression at some stage during their treatment (Hargrave, 2009). Adjuvant chemotherapy and/or radiotherapy are used for patients with unresectable, recurrent or metastatic disease. These strategies are not always curative, and leave a significant number of patients with severe long-term side effects from treatment including neuroendocrine abnormalities, growth failure, learning disabilities, vasculopathy and secondary neoplasms (Packer *et al.*, 1987, Packer *et al.*, 2003, Mulhern *et al.*, 2004).

MAPK pathway inhibitors are already in development for use in adult cancers, and these agents may also be suitable for use in clinical trials for paediatric patients with low-grade astrocytomas. Sorafenib is a small molecule inhibitor of tyrosine protein kinases including RAF and RET, and has additional activity against vascular endothelial growth factor receptors (VEGFR) and platelet-derived growth factor receptor β (PDGFR β), among others (Escudier *et al.*, 2007). Sorafenib has shown efficacy in the treatment of hepatocellular carcinoma and renal cell carcinoma in adults (Escudier *et al.*, 2007, Llovet *et al.*, 2008). AZD6244 is a MEK inhibitor, currently undergoing early clinical trials in adult patients with a variety of solid tumours (Friday *et al.*, 2008). Pre- and post-treatment biopsies showed an 83% reduction in nuclear phospho-ERK immunostaining following AZD6244 treatment, showing evidence of significant drug activity within tumour tissue. Clinical trials using targeted therapy against the MAPK pathway in low-grade astrocytoma, for patients with recurrent or metastatic disease, or where surgery is ineffective are likely to start in the near future.

However, a note of caution must be introduced here, remembering that these treatments are aimed at paediatric patients. Altering the patterns of activity in the MAPK pathway

within the developing brain may have devastating, and entirely unexpected consequences, potentially as serious as those left by current treatment modalities. A future goal for treatment of low-grade astrocytomas could be to use RNA inhibition to specifically target the *RAF* gene fusions, leaving cells without the fusion unaffected.

7.1.2 Lessons from adult glioblastoma studies

The comprehensive genome-wide analyses of copy number and mutations, and the recently published study using GWAS to identify novel candidate genes in adult glioblastoma show the power of combining research efforts in large collaborative studies (CGARN *et al.*, 2008, Parsons *et al.*, 2008, Shete *et al.*, 2009). Paediatric brain tumours are rare compared to brain tumours seen in adults, therefore future research strategies should include more national and international collaborations to increase the power of studies to reveal new insights into these diseases.

A recent paper has indicated that *NF1* inactivation may play a more significant role in adult glioblastomas than previously thought, and that differences in the method of inactivation may affect responses to treatment (McGillicuddy *et al.*, 2009). Two distinct mechanisms for *NF1* inactivation were identified in GBMs. Firstly, hyperactivation of protein kinase C (PKC) triggers proteasomal degradation of neurofibromin, the protein encoded by *NF1*, as a result these tumours show sensitivity to PKC inhibitors. More rarely, GBMs may have complete homozygous loss of *NF1* with concurrent inactivation of p53. These tumours are not sensitive to PKC inhibitors, but homozygous loss of *NF1* confers high sensitivity to mTOR inhibitors. The same authors have proposed a mechanism for the development of pilocytic astrocytomas in neurofibromatosis type 1 (NF1) (McGillicuddy *et al.*, 2009). Immunohistochemical analysis of two NF1-associated pilocytic astrocytomas revealed high levels of expression of p15^{INK4 β} , p16^{INK4 α} and p53; all three have been described as markers of oncogene-induced senescence (Braig *et al.*, 2005, Collado *et al.*, 2005). McGillicuddy *et al.* propose that in patients with NF1 the second genetic ‘hit (as outlined in Knudson’s two-hit hypothesis (Knudson, 1971)) results in complete inactivation of *NF1*, or may drive an initial period of rapid cellular proliferation, resulting in the development of a pilocytic astrocytoma. The p53 pathway is unaffected in these patients, and is subsequently activated to induce

senescence, preventing further progression. What triggers tumour progression to occur at a later date in these patients is currently unknown. Malignant astrocytoma formation occurs with 100% penetrance in mouse models with homozygous loss of function mutations in *NFI* and concurrent inactivation of the p53 pathway (Zhu *et al.*, 2005). Mice with homozygous *NFI* loss of function with intact p53 rarely develop tumours, implying that the intact p53 pathway prevents the formation of tumours by activation of the *RAS* pathway alone.

7.1.3 Fusion genes in cancer

Numerous karyotypic abnormalities have been identified in cancers, however these do not appear to be evenly distributed throughout the genome (Heim *et al.*, 2009). Some genomic regions are preferentially involved in rearrangements found in cancer, and some of these genetic aberrations appear to be highly specific for particular cancer types.

Translocations and their corresponding gene fusions have been implicated in the development of cancer for many years (reviewed by Mitelman in 2007) (Mitelman *et al.*, 2007). To date, the majority of fusion genes have been found in leukaemias and sarcomas, however new techniques including array CGH and SNP arrays have enabled fusions to be identified in other solid tumours. Recent studies have identified novel gene fusions in solid tumours using entirely new technologies and approaches including genome-wide massively parallel paired-end sequencing and integrative transcriptome sequencing (Ruan *et al.*, 2007, Campbell *et al.*, 2008, Maher *et al.*, 2009). These studies have identified gene fusions in prostate, lung and breast cancers, all of which had breakpoints situated so closely together that they could not have been identified by previously available cytogenetic techniques (Heim *et al.*, 2009). It is likely that there will be an exponential growth in our knowledge of genetic aberrations and gene fusions in solid tumours, as these new technologies are exploited further. However, a large number of chromosomal abnormalities may not, in themselves, be causative of cancer, and may merely reflect an accumulation of aberrations in a cell already displaying evidence of neoplastic transformation as reviewed by Stratton (Stratton *et al.*, 2009). Distinguishing which aberrations constitute ‘drivers’ of malignant transformation, and which are ‘passengers’, conferring no particular advantage to the malignant process,

which may have occurred at any time before or after the ‘driver’ mutation arose will be particularly important for future studies.

Chromosomal aberrations in themselves may not be sufficient for tumorigenesis to occur. Trisomy 8 has been found in both benign and malignant diseases, but has also been found in tissue from apparently healthy normal individuals (Heim *et al.*, 2009). Fusion genes associated with leukaemias, including *BCR-ABL* t(9;22)(q34;q11) and *ETV6/RUNX1* t(12;21)(p13;q22) have been found in the blood of healthy infants, at ~ 100 fold higher than the rates of childhood leukaemias (Mori *et al.*, 2002, Greaves, 2006). Only a small proportion of these children will go on to develop leukaemia, which suggests that further stimuli in addition to the presence of a fusion gene are necessary to drive leukaemic transformation. Numerous epidemiological data, and a recent GWAS study which has implicated genes involved in transcriptional regulation and differentiation of B-cell progenitors suggest that infections may drive the genesis of childhood leukaemias in patients with an underlying genetic predisposition (Greaves, 2006, Papaemmanuil *et al.*, 2009). Additional stimuli may also be required to drive the process of tumorigenesis in solid tumours: their identification must await future studies.

Chapter 8. Appendices

Appendix A Reagents

Chemicals

Agarose	(Invitrogen, Carlsbad, CA)
Anti-avidin-FITC	(Roche Diagnostics)
Anti-digoxigenin-rhodamine	(Roche Diagnostics)
Biotin-14-dATP	(Invitrogen, Carlsbad, CA)
1mM Cy3 dCTP	(NEN Life Science)
1mM Cy5 dCTP	(NEN Life Science)
Citifluor (Glycerol/PBS)	(Citifluor, Ltd)
Chloroform	(Sigma)
Colcemid	(Sigma)
Cot-1 DNA	(Invitrogen, Carlsbad, CA)
DAPI III	(Invitrogen, Carlsbad, CA)
Dextran sulphate	(Fisher Scientific)
Digoxigenin-11dUTP	(Roche Diagnostics)
E4 media	(Invitrogen, Carlsbad, CA)
EDTA	(Sigma)
Ethanol	(Fisher Scientific)
Ethidium bromide	(Sigma)
Fetal calf serum	(Gibco)
Fluorescein isothiocyanate-12-dUTP (FITC)	(NEN Life Science)
Formamide	(Applied Biosystems)
Formaldehyde	(Fisher Scientific)
Glacial acetic acid	(Fisher Scientific)
L-glutamine	(Fisher Scientific)
Glycerol	(Fisher Scientific)
Herring sperm DNA	(Sigma, D7290)
Human Cot-1 DNA	(Invitrogen, Carlsbad, CA)
Hyperladder I-II, DNA ladders	(Bioline)
Klenow fragment	(Applied Biosystems)
2-mercaptoethanol	(Sigma)
Methanol	(Fisher Scientific)
NP-40	(Calbiochem)
Non-essential amino acids	(Sigma)

Orange-G	(Sigma)
Pepsin	(Vysis)
Phosphate buffered saline (PBS)	(Fisher Scientific)
Potassium Chloride (KCl)	(Fisher Scientific)
RNase	(Sigma)
Salmon sperm DNA	(Invitrogen, Carlsbad, CA)
Sodium dodecyl sulphate (SDS)	(Fisher Scientific)
20x SSC (0.3M trisodium citrate/3.0M sodium chloride)	(Fisher Scientific)
Sodium acetate (NaAc)	(Fisher Scientific)
Sodium pyruvate	(Fisher Scientific)
Texas Red-5-dUTP (TR)	(NEN Life Science)
Tris(hydroxymethyl)aminomethane (Tris)	(Fisher Scientific)
Trizol	(Invitrogen, Carlsbad, CA)
Tween 20	(Pierce 28320)
Vysis SpectraVysion probe	(Abbot Laboratories)
Water (distilled RNase/Dnase free)	(Sigma)
Yeast tRNA	(Invitrogen, Carlsbad, CA)

Kits

BigDye Terminator Cycle Sequencing kit version 3.0	(Applied Biosystems)
Bioprime labelling kit	(Invitrogen, Carlsbad, CA)
BioNick kit	(Invitrogen, Carlsbad, CA)
DyeEx 2.0 spin columns/96-well plates	(Qiagen, Crawley)
Exosap-IT	(USB, Affymetrix)
QIAGEN DNeasy Blood and Tissue Mini Kit	(Qiagen, Crawley)
QIAGEN RNeasy Cell and Tissue Mini Kit	(Qiagen, Crawley)
QIAGEN® Fast-cycling PCR kit	(Qiagen, Crawley)
QIAquick™ Gel extraction kit	(Qiagen, Crawley)
QIAGEN large-construct kit	(Qiagen, Crawley)
Micro-spin G50 columns	(Pharmacia Amersham)
Repli-G mini kit	(Qiagen, Crawley)
SuperScript First-Strand cDNA synthesis kit	(Invitrogen, Carlsbad, CA)

8.1.1 Loading buffer for gel electrophoresis

Orange G - 1-Phenylazo-2-naphthol-6,8-disulfonic acid disodium salt (Sigma)

2% orange G is made from 0.2g in 10 ml of water.

Glycerol	5 ml
0.5 M EDTA	1 ml
2% orange G	1 ml
10% SDS	0.1 ml
Distilled water	2.9 ml

Appendix B Reagents prepared for array CGH

Hybridisation buffer

Dextran sulphate	2 grams
De-ionised formamide	10ml
20 x SSC	2ml
1M Tris	200µl
10% Tween 20	200µl
Distilled water	make up final volume to 20mls

Dissolve overnight on shaking platform at room temperature.

Once made, divide into 800µl aliquots and store at -20°C .

Array CGH wash 1

20 x SSC	100ml
20% SDS	1.5ml

Make up to 1 litre with distilled water

Array CGH wash 2

20 x SSC	10ml
----------	------

Make up to 1 litre with distilled water

Array CGH wash 3

10% Tween 20	10ml
--------------	------

Make up to 1 litre with distilled water

Clone	Chr	Position (bp)
bA326G21	1	143533568
bA5K23	1	159201779.5
dJ1108M17	1	104815713
dJ97P20	1	167462445
bA32C20	2	128082133
bA400O18	2	184873490
963K6	4	191632378
bA94E2	5	18186914.5
dJ159G19	6	80412976
dJ93N13	6	32596730.5
bA350F16	8	46413667
bB445N5	10	38414815
bA13E1	10	48233085
221K18	12	131110572
bA25J23	13	78143302
bA279F15	13	55656472
820M16	14	104124908
bA2F9	15	18507931
bA161M6	16	1085464
dJ843B9	17	43741286
bA220N20	17	44390565.5
bA416K7	17	45289524.5
bA294I20	19	86263
bA50L23	22	19845427.5

Table 8.1 Array CGH clones known to give unexpected ratios

The following autosomal clones used on the array slides have been shown to display unexpected linear ratios in normal/normal hybridisations by a number of independent groups.

Appendix C. Primers

Gene	cDNA Primers
<i>KIAA1549</i> primer 1 F (exon 2)	5'AGCACAGCAGTCCTCTCCAT3'
<i>KIAA1549</i> primer 2 F (exon 2)	5'CACCTTTCGGTCAGCTTTTC3'
<i>KIAA1549</i> primer 3 F (exon 2)	5'CCTAGCGACTCCAGCAAAAC3'
<i>KIAA1549</i> primer 4 F (exon 2)	5'CACTGCATTTTTCTCGGTCA3'
<i>KIAA1549</i> primer 5 F (exon 2)	5'ACGGTTTCACTGACGGATTC3'
<i>KIAA1549</i> primer 6 F (exon 2)	5'GCCTTCACTCTCGAAGCAAC3'
<i>KIAA1549</i> primer 7 F (exon 6)	5'TTGCGGTTAAAAGCACACAG3'
<i>KIAA1549</i> primer 8 F (exon 9)	5'AAATGGACCTCCAGAGAGCA3'
<i>KIAA1549</i> primer 9 F (exon 12)	5'GAGGGACGCAGGAGATAAGA3'
<i>KIAA1549</i> primer 10 F (exon 14/15)	5'CCAGGAAGAGCTCACGGATA3'
<i>KIAA1549</i> primer 11 F (exon 16)	5'CAGTGGGGGTCCTTCTACAG3'
<i>KIAA1549</i> primer 12 F (exon 20)	5'CTCCACGGAAGACCTCCAG3'
<i>BRAF</i> exon 12 R	5'TGCTGAGGTGTAGGTGCTGT3'
<i>KIAA1549</i> nested primer 1 F	5'GAGGGACGCAGGAGATAAGA3'
<i>KIAA1549</i> nested primer 2 F	5'CCAGGAAGAGCTCACGGATA3'
<i>BRAF</i> nested primer 1 R	5'AAGTAATCCATGCCCTGTGC3'
<i>BRAF</i> nested primer 2 R	5'TGCTGAGGTGTAGGTGCTGT3'

Table 8.2 Primers for identification of *KIAA1549-BRAF* gene fusions in cDNA
Forward primers (F) and reverse primers (R) are shown.

Gene	Primers
<i>KIAA1549</i> DNA 1 F	5'CAGTGGGGGTCCTTCTACAG3'

<i>KIAA1549</i> DNA 2 F	5'GGGACCAGTAAACGGTCAGA3'
<i>KIAA1549</i> DNA 3 F	5'GCCAGTCTGTGTGACTCCAA3'
<i>KIAA1549</i> DNA 4 F	5'ACTCTGTCGGCCTGTGTTTT3'
<i>KIAA1549</i> DNA 5 F	5'AGGTGAGAGGGAGCAGTTGA3'
<i>KIAA1549</i> DNA 6 F	5'TGGTGAAACCCCGTCTCTAC3'
<i>KIAA1549</i> DNA 7 F	5'AACCAGGTGCTTCAAGGATG3'
<i>KIAA1549</i> DNA 8 F	5'CCAAAGAAAACAGCAGCACAA3'
<i>KIAA1549</i> DNA 9 F	5'GAGTGCAGTGGCACAATGTC3'
<i>KIAA1549</i> DNA 10 F	5'GCTGGGTTTACCTTCCTTCC3'
<i>KIAA1549</i> DNA 11 F	5'ACATGAAATTGGGTGGCAAT3'
<i>KIAA1549</i> DNA 12 F	5'GCAGGAGGACTTTTGCTTTG3'
<i>BRAF</i> DNA 1 R	5'CACCACGAAATCCTTGGTCT3'
<i>BRAF</i> DNA 2 R	5'ATGTCCATTTCAGGCCATTC3'
<i>BRAF</i> DNA 3 R	5'TGAATGGCCCCAAATATCTC3'
<i>BRAF</i> DNA 4 R	5'GCACCCAGCCGATAAGATT3'
<i>BRAF</i> DNA 5 R	5'TCCTTCACAGCCAAGCTTCT3'
<i>BRAF</i> DNA 6 R	5'TAGTCTCCCTCCCCAAAACA3'
<i>BRAF</i> DNA 7 R	5'CCCAACATACCAGGGAACAG3'
<i>BRAF</i> DNA 8 R	5'CATGCATGGCTAATTTGTTTT3'

Table 8.3 Primers for identification of the *KIAA1549-BRAF* fusion breakpoints in genomic DNA

Forward primers (F) and reverse primers (R) are shown.

Gene	Primers
<i>SRGAP3</i> 1 F (exon 1)	5'ATCTGGGGCAGGTTACTGTG3'
<i>SRGAP3</i> 2 F (exon 2)	5'GCTGGTGGAGCAGTTCAAAT3'
<i>SRGAP3</i> 3 F (exon 5)	5'GCTGAGAAGCAGGAGGAGAA3'
<i>SRGAP3</i> 4 F (exon 7)	5'AGTCTTCTGCCCTCCACTCA3'
<i>SRGAP3</i> 5 F (exon 10)	5'TGGCAGTAACCTCATCACCA3'
<i>SRGAP3</i> 6 F (exon 14/15)	5'AGAGGTGAAGACCCCCTTGT3'
<i>SRGAP3</i> 7 F (exon 18)	5'TTGACCATGACAATGGCACT3'
<i>SRGAP3</i> 8 F (exon 21)	5'GCGGTTACGATCTGATGGAG3'
<i>SRGAP3</i> 9 F (exon 22)	5'CGAGATGATGACCACCTTCA3'
<i>RAF1</i> 1 R (exon 10)	5'AAAAGAGCCTGACCCAATCC3'
<i>RAF1</i> 2 R (exon 11)	5'GGTTGGGTCGACAACCTTTA3'
<i>RAF1</i> 3 R (exon 12)	5'GCCAGGTTGTCCTTTGTCAT3'

Table 8.4 Primers for identification of the *SRGAP3-RAF1* fusion breakpoints in cDNA

Forward primers (F) and reverse primers (R) are shown.

Exon	Primers
<i>KRAS</i> exon 2 F	5'CCTTTATCTGTATCAAAGAATGGTC3'
<i>KRAS</i> exon 2 R	5'TTTGTATTAAGGTTACTGGTGGAG3'
<i>KRAS</i> exon 3 F	5'TGCATGGCATTAGCAAAGAC3'
<i>KRAS</i> exon 3 R	5'TCTTTGGAGCAGGAACAATG3'

<i>NRAS</i> exon 2 F	5'GAACCAAATGGAAGGTCACA3'
<i>NRAS</i> exon 2 R	5'TGGGTAAAGATGATCCGACA3'
<i>NRAS</i> exon 3 F	5'TGCCCCCTTACCCTCCACA3'
<i>NRAS</i> exon 3 R	5'CCTCATTTCCCCATAAAGATTCAGA3'
<i>HRAS</i> exon 2 F	5'CAGGAGACCCTGTAGGAGGA3'
<i>HRAS</i> exon 2 R	5'CCTATCCTGGCTGTGTCCTG3'
<i>HRAS</i> exon 3 F	5'AGAGGCTGGCTGTGTGAACT3'
<i>HRAS</i> exon 3 R	5'TCACGGGGTTCACCTGTA3'
<i>BRAF</i> exon 11 F	5'TTTTCTGTTTGGCTTGA3'
<i>BRAF</i> exon 11 R	5'TGTCACAATGTCACCACATTACA3'
<i>BRAF</i> exon 15 F	5'TCATAATGCTTGCTCTGATAGGA3'
<i>BRAF</i> exon 15 R	5'GGCCAAAATTTAATCAGTGGA3'

Table 8.5 DNA sequencing primers for *KRAS*, *NRAS*, *HRAS* and *BRAF*
Forward primers (F) and reverse primers (R) are shown.

Exon	Primers
<i>CDKN2A</i> exon 1 F	5'GAAAGAGGAGGGGCTGGCTGGTC3'
<i>CDKN2A</i> exon 1 R	5'GCGCTACCTGATTCCAATTCCCCTGC3'
<i>CDKN2A</i> exon 2 F	5'GGCTCTACACAAGCTTCCTT3'
<i>CDKN2A</i> exon 2 R	5'TGAGCTTTGGAAGCTCTCAG3'
<i>CDKN2A</i> exon 3 F	5'CAACAGTGTTCAGAAACGATGC3'
<i>CDKN2A</i> exon 3 R	5'TGTGCCACACATCTTTGACC3'
<i>PTEN</i> exon 1 F	5'TTCCATCCTGCAGAAGAAGC3'
<i>PTEN</i> exon 1 R	5'ATCCGTCTACTCCCACGTTC3'
<i>PTEN</i> exon 2 F	5'TCCAGCTATAGTGGGGAAAAC3'
<i>PTEN</i> exon 2 R	5'AAGTCCATTAGGTACGGTAAGCC3'
<i>PTEN</i> exon 3 F	5'AAACCCATAGAAGGGGTATTTG3'
<i>PTEN</i> exon 3 R	5'AACAATGCTCTTGGACTTCTTG3'
<i>PTEN</i> exon 4 F	5'AAAGATTCAGGCAATGTTTGTAG3'
<i>PTEN</i> exon 4 R	5'TGTATCTCACTCGATAATCTGGATG3'
<i>PTEN</i> exon 5 F	5'GGAATCCAGTGTTCCTTTTAAATACC3'
<i>PTEN</i> exon 5 R	5'TTTTCCAATAAATTCCTCAGATCCAG3'
<i>PTEN</i> exon 6 F	5'ATGGCTACGACCCAGTTACC3'
<i>PTEN</i> exon 6 R	5'CAAATGCTTCAGAAATATAGTCTCC3'
<i>PTEN</i> exon 7 F	5'TGCTTGAGATCAAGATTGCAG3'
<i>PTEN</i> exon 7 R	5'CTCACCAATGCCAGAGTAAGC3'
<i>PTEN</i> exon 8 F	5'CTCAGATTGCCTTATAATAGTCTTTG3'
<i>PTEN</i> exon 8 R	5'TCAAGCAAGTTCTTCATCAGC3'
<i>PTEN</i> exon 9 F	5'AGCTTGGCAACAGAGCAAG3'
<i>PTEN</i> exon 9 R	5'TAAAACGGGAAAGTGCCATC3'
<i>IDH1</i> exon 4 F	5'GAGCTCTATATGCCATCACTGC3'
<i>IDH1</i> exon 4 R	5'CAAGTTGGAATTTCTGGGC3'
<i>IDH2</i> exon 4-5 F	5'ATTCTGGTTGAAAGATGGCG3'
<i>IDH2</i> exon 4-5 R	5'CAGAAGAAAGGAAAGCCACG3'

Table 8.6 DNA sequencing primers for *CDKN2A*, *PTEN*, *IDH1* and *IDH2*. Forward primers (F) and reverse primers (R) are shown.

References

- Adams, J. M., Harris, A. W., Pinkert, C. A., Corcoran, L. M., Alexander, W. S., Cory, S., Palmiter, R. D. & Brinster, R. L. (1985) The c-myc oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice. *Nature*, 318, 533-8.
- Adkins, N. L., Watts, M. & Georgel, P. T. (2004) To the 30-nm chromatin fiber and beyond. *Biochim Biophys Acta*, 1677, 12-23.
- Agamanolis, D. P. & Malone, J. M. (1995) Chromosomal abnormalities in 47 pediatric brain tumors. *Cancer Genet Cytogenet*, 81, 125-34.
- Agathangelou, A., Honorio, S., Macartney, D. P., Martinez, A., Dallol, A., Rader, J., Fullwood, P., Chauhan, A., Walker, R., Shaw, J. A., Hosoe, S., Lerman, M. I., Minna, J. D., Maher, E. R. & Latif, F. (2001) Methylation associated inactivation of RASSF1A from region 3p21.3 in lung, breast and ovarian tumours. *Oncogene*, 20, 1509-18.
- Albertson, D. G. & Pinkel, D. (2003) Genomic microarrays in human genetic disease and cancer. *Hum Mol Genet*, 12 Spec No 2, R145-52.
- Arai, E., Ikeuchi, T. & Nakamura, Y. (1994) Characterization of the translocation breakpoint on chromosome 22q12.2 in a patient with neurofibromatosis type 2 (NF2). *Hum Mol Genet*, 3, 937-9.
- Armitage, P. & Doll, R. (1954) The age distribution of cancer and a multi-stage theory of carcinogenesis. *Br J Cancer*, 8, 1-12.
- Armitage, P. & Doll, R. (2004) The age distribution of cancer and a multi-stage theory of carcinogenesis. 1954. *Int J Epidemiol*, 33, 1174-9.
- Arora, R. S., Alston, R. D., Eden, T. O., Estlin, E. J., Moran, A. & Birch, J. M. (2009) Age-incidence patterns of primary CNS tumors in children, adolescents, and adults in England. *Neuro Oncol*, 11, 403-13.
- Arvand, A. & Denny, C. T. (2001) Biology of EWS/ETS fusions in Ewing's family tumors. *Oncogene*, 20, 5747-54.
- Ater, J., Holmes, E., Zhou, T., Mazewski, C., Roberts, W. & Vezina, G. E. A. (2008) Results of COG protocol A9952: A randomized phase 3 study of two chemotherapy regimens for incompletely resected low-grade glioma in young children. *Neuro-Oncology*, 10:, 451-452.

Avery, O. T., Macleod, C. M. & McCarty, M. (1944) Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types : Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii. *J Exp Med*, 79, 137-158.

Bacon, C., Endris, V. & Rappold, G. (2009) Dynamic expression of the Slit-Robo GTPase activating protein genes during development of the murine nervous system. *J Comp Neurol*, 513, 224-36.

Bailey, J. A. & Eichler, E. E. (2006) Primate segmental duplications: crucibles of evolution, diversity and disease. *Nat Rev Genet*, 7, 552-64.

Baker, S. J., Fearon, E. R., Nigro, J. M., Hamilton, S. R., Preisinger, A. C., Jessup, J. M., Vantuinen, P., Ledbetter, D. H., Barker, D. F., Nakamura, Y., White, R. & Vogelstein, B. (1989) Chromosome 17 deletions and p53 gene mutations in colorectal carcinomas. *Science*, 244, 217-21.

Baldwin, R. T. & Preston-Martin, S. (2004) Epidemiology of brain tumors in childhood--a review. *Toxicol Appl Pharmacol*, 199, 118-31.

Bar, E. E., Lin, A., Tihan, T., Burger, P. C. & Eberhart, C. G. (2008) Frequent gains at chromosome 7q34 involving BRAF in pilocytic astrocytoma. *J Neuropathol Exp Neurol*, 67, 878-87.

Barlow, G. M., Lyons, G. E., Richardson, J. A., Sarnat, H. B. & Korenberg, J. R. (2002) DSCAM: an endogenous promoter drives expression in the developing CNS and neural crest. *Biochem Biophys Res Commun*, 299, 1-6.

Basecke, J., Whelan, J. T., Griesinger, F. & Bertrand, F. E. (2006) The MLL partial tandem duplication in acute myeloid leukaemia. *Br J Haematol*, 135, 438-49.

Batzler, M. A. & Deininger, P. L. (2002) Alu repeats and human genomic diversity. *Nat Rev Genet*, 3, 370-9.

Baylin, S. B. & Ohm, J. E. (2006) Epigenetic gene silencing in cancer - a mechanism for early oncogenic pathway addiction? *Nat Rev Cancer*, 6, 107-16.

Beheshti, B., Braude, I., Marrano, P., Thorner, P., Zielenska, M. & Squire, J. A. (2003) Chromosomal localization of DNA amplifications in neuroblastoma tumors using cDNA microarray comparative genomic hybridization. *Neoplasia*, 5, 53-62.

Behrman, R. E., Kliegman, R. M. & Arvin, A. M. (1996) Chapter 555. Brain tumors in children. IN Nelson, W. E. (Ed.) *Nelson Textbook of Pediatrics*. 15th Edition ed. Philadelphia, W.B. Saunders.

- Beier, D., Hau, P., Proescholdt, M., Lohmeier, A., Wischhusen, J., Oefner, P. J., Aigner, L., Brawanski, A., Bogdahn, U. & Beier, C. P. (2007) CD133(+) and CD133(-) glioblastoma-derived cancer stem cells show differential growth characteristics and molecular profiles. *Cancer Res*, 67, 4010-5.
- Bestor, T. H. (2005) Transposons reanimated in mice. *Cell*, 122, 322-5.
- Bhattacharjee, M. B., Armstrong, D. D., Vogel, H. & Cooley, L. D. (1997) Cytogenetic analysis of 120 primary pediatric brain tumors and literature review. *Cancer Genet Cytogenet*, 97, 39-53.
- Bignell, G. R., Huang, J., Greshock, J., Watt, S., Butler, A., West, S., Grigorova, M., Jones, K. W., Wei, W., Stratton, M. R., Futreal, P. A., Weber, B., Shaperro, M. H. & Wooster, R. (2004) High-resolution analysis of DNA copy number using oligonucleotide microarrays. *Genome Res*, 14, 287-95.
- Bigner, S. H., McLendon, R. E., Fuchs, H., Mckeever, P. E. & Friedman, H. S. (1997) Chromosomal characteristics of childhood brain tumors. *Cancer Genet Cytogenet*, 97, 125-34.
- Bird, A. P. (1986) CpG-rich islands and the function of DNA methylation. *Nature*, 321, 209-13.
- Blaney, S. M., Kun, L. E., Hunter, J., Rorke-Adams, L. B., Ching, L., Strother, D. & Pollack, I. F. (2006) Tumours of the Central Nervous System. IN Pizzo, P. A. A. P., D.G. (Ed.) *Principles and Practice of Paediatric Oncology*. Fifth edition ed. Philadelphia, Lippincott Williams and Wilkins.
- Blatt, J., Jaffe, R., Deutsch, M. & Adkins, J. C. (1986) Neurofibromatosis and childhood tumors. *Cancer*, 57, 1225-9.
- Bolzer, A., Kreth, G., Solovei, I., Koehler, D., Saracoglu, K., Fauth, C., Muller, S., Eils, R., Cremer, C., Speicher, M. R. & Cremer, T. (2005) Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLoS Biol*, 3, e157.
- Bondy, M., Wiencke, J., Wrensch, M. & Kyritsis, A. P. (1994) Genetics of primary brain tumors: a review. *J Neurooncol*, 18, 69-81.
- Bonneau, D. & Longy, M. (2000) Mutations of the human PTEN gene. *Hum Mutat*, 16, 109-22.
- Bonnet, D. & Dick, J. E. (1997) Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med*, 3, 730-7.
- Bos, J. L. (1989) ras oncogenes in human cancer: a review. *Cancer Res*, 49, 4682-9.

- Boveri, T. (1914) *Zur Frage der Entstehung Maligner Tumoren. (Translation: The origin of Malignant Tumours)*. Williams and Wilkins.
- Braig, M., Lee, S., Loddenkemper, C., Rudolph, C., Peters, A. H., Schlegelberger, B., Stein, H., Dorken, B., Jenuwein, T. & Schmitt, C. A. (2005) Oncogene-induced senescence as an initial barrier in lymphoma development. *Nature*, 436, 660-5.
- Brems, H., Beert, E., De Ravel, T. & Legius, E. (2009) Mechanisms in the pathogenesis of malignant tumours in neurofibromatosis type 1. *Lancet Oncol*, 10, 508-15.
- Broniscer, A., Baker, S. J., West, A. N., Fraser, M. M., Proko, E., Kocak, M., Dalton, J., Zambetti, G. P., Ellison, D. W., Kun, L. E., Gajjar, A., Gilbertson, R. J. & Fuller, C. E. (2007) Clinical and molecular characteristics of malignant transformation of low-grade glioma in children. *J Clin Oncol*, 25, 682-9.
- Broniscer, A. & Gajjar, A. (2004) Supratentorial high-grade astrocytoma and diffuse brainstem glioma: two challenges for the pediatric oncologist. *Oncologist*, 9, 197-206.
- Brunet, A., Bonni, A., Zigmond, M. J., Lin, M. Z., Juo, P., Hu, L. S., Anderson, M. J., Arden, K. C., Blenis, J. & Greenberg, M. E. (1999) Akt promotes cell survival by phosphorylating and inhibiting a Forkhead transcription factor. *Cell*, 96, 857-68.
- Buckley, J. D., Pendergrass, T. W., Buckley, C. M., Pritchard, D. J., Nesbit, M. E., Provisor, A. J. & Robison, L. L. (1998) Epidemiology of osteosarcoma and Ewing's sarcoma in childhood: a study of 305 cases by the Children's Cancer Group. *Cancer*, 83, 1440-8.
- Bunone, G., Briand, P. A., Miksicek, R. J. & Picard, D. (1996) Activation of the unliganded estrogen receptor by EGF involves the MAP kinase pathway and direct phosphorylation. *Embo J*, 15, 2174-83.
- Burkhard, C., Di Patre, P. L., Schuler, D., Schuler, G., Yasargil, M. G., Yonekawa, Y., Lutolf, U. M., Kleihues, P. & Ohgaki, H. (2003) A population-based study of the incidence and survival rates in patients with pilocytic astrocytoma. *J Neurosurg*, 98, 1170-4.
- Cadioux, B., Ching, T. T., Vandenberg, S. R. & Costello, J. F. (2006) Genome-wide hypomethylation in human glioblastomas associated with specific copy number alteration, methylenetetrahydrofolate reductase allele status, and increased proliferation. *Cancer Res*, 66, 8469-76.
- Cahill, D. P., Kinzler, K. W., Vogelstein, B. & Lengauer, C. (1999) Genetic instability and darwinian selection in tumours. *Trends Cell Biol*, 9, M57-60.

- Cai, W. W., Mao, J. H., Chow, C. W., Damani, S., Balmain, A. & Bradley, A. (2002) Genome-wide detection of chromosomal imbalances in tumors using BAC microarrays. *Nat Biotechnol*, 20, 393-6.
- Callagy, G., Pharoah, P., Chin, S. F., Sangan, T., Daigo, Y., Jackson, L. & Caldas, C. (2005) Identification and validation of prognostic markers in breast cancer with the complementary use of array-CGH and tissue microarrays. *J Pathol*, 205, 388-96.
- Campbell, P. J., Stephens, P. J., Pleasance, E. D., O'meara, S., Li, H., Santarius, T., Stebbings, L. A., Leroy, C., Edkins, S., Hardy, C., Teague, J. W., Menzies, A., Goodhead, I., Turner, D. J., Clee, C. M., Quail, M. A., Cox, A., Brown, C., Durbin, R., Hurler, M. E., Edwards, P. A., Bignell, G. R., Stratton, M. R. & Futreal, P. A. (2008) Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. *Nat Genet*, 40, 722-9.
- Carinci, F., Arcelli, D., Lo Muzio, L., Francioso, F., Valentini, D., Evangelisti, R., Volinia, S., D'angelo, A., Meroni, G., Zollo, M., Pastore, A., Ionna, F., Mastrangelo, F., Conti, P. & Tete, S. (2007) Molecular classification of nodal metastasis in primary larynx squamous cell carcinoma. *Transl Res*, 150, 233-45.
- Cgarn & Network, C. G. A. R. (2008) Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature*, 455, 1061-8.
- Chan, J. A., Krichevsky, A. M. & Kosik, K. S. (2005) MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. *Cancer Res*, 65, 6029-33.
- Chang, T. C. & Mendell, J. T. (2007) microRNAs in vertebrate physiology and human disease. *Annu Rev Genomics Hum Genet*, 8, 215-39.
- Chapman, E. J., Harnden, P., Chambers, P., Johnston, C. & Knowles, M. A. (2005) Comprehensive analysis of CDKN2A status in microdissected urothelial cell carcinoma reveals potential haploinsufficiency, a high frequency of homozygous co-deletion and associations with clinical phenotype. *Clin Cancer Res*, 11, 5740-7.
- Cherukuri, S., Hock, R., Ueda, T., Catez, F., Rochman, M. & Bustin, M. (2008) Cell cycle-dependent binding of HMGN proteins to chromatin. *Mol Biol Cell*, 19, 1816-24.
- Cheung, J., Estivill, X., Khaja, R., Macdonald, J. R., Lau, K., Tsui, L. C. & Scherer, S. W. (2003) Genome-wide detection of segmental duplications and potential assembly errors in the human genome sequence. *Genome Biol*, 4, R25.
- Chiang, P. K., Gordon, R. K., Tal, J., Zeng, G. C., Doctor, B. P., Pardhasaradhi, K. & Mccann, P. P. (1996) S-Adenosylmethionine and methylation. *Faseb J*, 10, 471-80.

Chin, K., Devries, S., Fridlyand, J., Spellman, P. T., Roydasgupta, R., Kuo, W. L., Lapuk, A., Neve, R. M., Qian, Z., Ryder, T., Chen, F., Feiler, H., Tokuyasu, T., Kingsley, C., Dairkee, S., Meng, Z., Chew, K., Pinkel, D., Jain, A., Ljung, B. M., Esserman, L., Albertson, D. G., Waldman, F. M. & Gray, J. W. (2006) Genomic and transcriptional aberrations linked to breast cancer pathophysiology. *Cancer Cell*, 10, 529-41.

Ciampi, R., Knauf, J. A., Kerler, R., Gandhi, M., Zhu, Z., Nikiforova, M. N., Rabes, H. M., Fagin, J. A. & Nikiforov, Y. E. (2005) Oncogenic AKAP9-BRAF fusion is a novel mechanism of MAPK pathway activation in thyroid cancer. *J Clin Invest*, 115, 94-101.

Cogen, P. H., Daneshvar, L., Metzger, A. K., Duyk, G., Edwards, M. S. & Sheffield, V. C. (1992) Involvement of multiple chromosome 17p loci in medulloblastoma tumorigenesis. *Am J Hum Genet*, 50, 584-9.

Collado, M., Gil, J., Efeyan, A., Guerra, C., Schuhmacher, A. J., Barradas, M., Benguria, A., Zaballos, A., Flores, J. M., Barbacid, M., Beach, D. & Serrano, M. (2005) Tumour biology: senescence in premalignant tumours. *Nature*, 436, 642.

Costinean, S., Zanesi, N., Pekarsky, Y., Tili, E., Volinia, S., Heerema, N. & Croce, C. M. (2006) Pre-B cell proliferation and lymphoblastic leukemia/high-grade lymphoma in E(mu)-miR155 transgenic mice. *Proc Natl Acad Sci U S A*, 103, 7024-9.

Cotterill, S. J., Wright, C. M., Pearce, M. S. & Craft, A. W. (2004) Stature of young people with malignant bone tumors. *Pediatr Blood Cancer*, 42, 59-63.

Cremer, T. & Cremer, C. (2001) Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet*, 2, 292-301.

Croce, C. M. (2008) Oncogenes and cancer. *N Engl J Med*, 358, 502-11.

Curtin, J. F. & Cotter, T. G. (2004) JNK regulates HIPK3 expression and promotes resistance to Fas-mediated apoptosis in DU 145 prostate carcinoma cells. *J Biol Chem*, 279, 17090-100.

Dahmane, N. & Ruiz I Altaba, A. (1999) Sonic hedgehog regulates the growth and patterning of the cerebellum. *Development*, 126, 3089-100.

Daley, G. Q., Van Etten, R. A. & Baltimore, D. (1990) Induction of chronic myelogenous leukemia in mice by the P210bcr/abl gene of the Philadelphia chromosome. *Science*, 247, 824-30.

Davies, H., Bignell, G. R., Cox, C., Stephens, P., Edkins, S., Clegg, S., Teague, J., Woffendin, H., Garnett, M. J., Bottomley, W., Davis, N., Dicks, E., Ewing, R., Floyd,

Y., Gray, K., Hall, S., Hawes, R., Hughes, J., Kosmidou, V., Menzies, A., Mould, C., Parker, A., Stevens, C., Watt, S., Hooper, S., Wilson, R., Jayatilake, H., Gusterson, B. A., Cooper, C., Shipley, J., Hargrave, D., Pritchard-Jones, K., Maitland, N., Chenevix-Trench, G., Riggins, G. J., Bigner, D. D., Palmieri, G., Cossu, A., Flanagan, A., Nicholson, A., Ho, J. W., Leung, S. Y., Yuen, S. T., Weber, B. L., Seigler, H. F., Darrow, T. L., Paterson, H., Marais, R., Marshall, C. J., Wooster, R., Stratton, M. R. & Futreal, P. A. (2002) Mutations of the BRAF gene in human cancer. *Nature*, 417, 949-54.

Davis, F. G., Freels, S., Grutsch, J., Barlas, S. & Brem, S. (1998) Survival rates in patients with primary malignant brain tumors stratified by patient age and tumor histological type: an analysis based on Surveillance, Epidemiology, and End Results (SEER) data, 1973-1991. *J Neurosurg*, 88, 1-10.

De Carli, E., Wang, X. & Puget, S. (2009) IDH1 and IDH2 mutations in gliomas. *N Engl J Med*, 360, 2248; author reply 2249.

Debiec-Rychter, M., Alwasiak, J., Liberski, P. P., Nedoszytko, B., Babinska, M., Mrozek, K., Imielinski, B., Borowska-Lehman, J. & Limon, J. (1995) Accumulation of chromosomal changes in human glioma progression. A cytogenetic study of 50 cases. *Cancer Genet Cytogenet*, 85, 61-7.

Deininger, M., Buchdunger, E. & Druker, B. J. (2005) The development of imatinib as a therapeutic agent for chronic myeloid leukemia. *Blood*, 105, 2640-53.

Dekel, B. (2003) Profiling gene expression in kidney development. *Nephron Exp Nephrol*, 95, e1-6.

Delattre, O., Zucman, J., Melot, T., Garau, X. S., Zucker, J. M., Lenoir, G. M., Ambros, P. F., Sheer, D., Turc-Carel, C., Triche, T. J. & Et Al. (1994) The Ewing family of tumors--a subgroup of small-round-cell tumors defined by specific chimeric transcripts. *N Engl J Med*, 331, 294-9.

Deng, C. X. & Wang, R. H. (2003) Roles of BRCA1 in DNA damage repair: a link between development and cancer. *Hum Mol Genet*, 12 Spec No 1, R113-23.

Deshmukh, H., Yeh, T. H., Yu, J., Sharma, M. K., Perry, A., Leonard, J. R., Watson, M. A., Gutmann, D. H. & Nagarajan, R. (2008) High-resolution, dual-platform aCGH analysis reveals frequent HIPK2 amplification and increased expression in pilocytic astrocytomas. *Oncogene*, 27, 4745-51.

Dessars, B., De Raeve, L. E., El Housni, H., Debouck, C. J., Sidon, P. J., Morandini, R., Roseeuw, D., Ghanem, G. E., Vassart, G. & Heimann, P. (2007) Chromosomal

translocations as a mechanism of BRAF activation in two cases of large congenital melanocytic nevi. *J Invest Dermatol*, 127, 1468-70.

Dhillon, A. S., Hagan, S., Rath, O. & Kolch, W. (2007) MAP kinase signalling pathways in cancer. *Oncogene*, 26, 3279-90.

Ditlevsen, D. K., Povlsen, G. K., Berezin, V. & Bock, E. (2008) NCAM-induced intracellular signaling revisited. *J Neurosci Res*, 86, 727-43.

Dolan, K., Garde, J., Walker, S. J., Sutton, R., Gosney, J. & Field, J. K. (1999) LOH at the sites of the DCC, APC, and TP53 tumor suppressor genes occurs in Barrett's metaplasia and dysplasia adjacent to adenocarcinoma of the esophagus. *Hum Pathol*, 30, 1508-14.

Doll, R. & Wakeford, R. (1997) Risk of childhood cancer from fetal irradiation. *Br J Radiol*, 70, 130-9.

Downward, J. (2003) Targeting RAS signalling pathways in cancer therapy. *Nat Rev Cancer*, 3, 11-22.

Druker, B. J., Talpaz, M., Resta, D. J., Peng, B., Buchdunger, E., Ford, J. M., Lydon, N. B., Kantarjian, H., Capdeville, R., Ohno-Jones, S. & Sawyers, C. L. (2001) Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *N Engl J Med*, 344, 1031-7.

Ebinger, M., Senf, L., Wachowski, O. & Scheurlen, W. (2005) No aberrant methylation of neurofibromatosis 1 gene (NF1) promoter in pilocytic astrocytoma in childhood. *Pediatr Hematol Oncol*, 22, 83-7.

Eden, A., Gaudet, F., Waghmare, A. & Jaenisch, R. (2003) Chromosomal instability and tumors promoted by DNA hypomethylation. *Science*, 300, 455.

Ehrlich, M. (2002) DNA methylation in cancer: too much, but also too little. *Oncogene*, 21, 5400-13.

Ehrlich, M., Gama-Sosa, M. A., Huang, L. H., Midgett, R. M., Kuo, K. C., Mccune, R. A. & Gehrke, C. (1982) Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res*, 10, 2709-21.

Eis, P. S., Tam, W., Sun, L., Chadburn, A., Li, Z., Gomez, M. F., Lund, E. & Dahlberg, J. E. (2005) Accumulation of miR-155 and BIC RNA in human B cell lymphomas. *Proc Natl Acad Sci U S A*, 102, 3627-32.

Endris, V., Wogatzky, B., Leimer, U., Bartsch, D., Zatyka, M., Latif, F., Maher, E. R., Tariverdian, G., Kirsch, S., Karch, D. & Rappold, G. A. (2002) The novel Rho-

GTPase activating gene MEGAP/ srGAP3 has a putative role in severe mental retardation. *Proc Natl Acad Sci U S A*, 99, 11754-9.

Engelman, J. A. (2009) Targeting PI3K signalling in cancer: opportunities, challenges and limitations. *Nat Rev Cancer*, 9, 550-62.

Escudier, B., Eisen, T., Stadler, W. M., Szczylik, C., Oudard, S., Siebels, M., Negrier, S., Chevreau, C., Solska, E., Desai, A. A., Rolland, F., Demkow, T., Hutson, T. E., Gore, M., Freeman, S., Schwartz, B., Shan, M., Simantov, R. & Bukowski, R. M. (2007) Sorafenib in advanced clear-cell renal-cell carcinoma. *N Engl J Med*, 356, 125-34.

Espinosa, A. V., Porchia, L. & Ringel, M. D. (2007) Targeting BRAF in thyroid cancer. *Br J Cancer*, 96, 16-20.

Esteller, M. (2008) Epigenetics in cancer. *N Engl J Med*, 358, 1148-59.

Esteller, M., Corn, P. G., Baylin, S. B. & Herman, J. G. (2001) A gene hypermethylation profile of human cancer. *Cancer Res*, 61, 3225-9.

Esteller, M. & Herman, J. G. (2002) Cancer as an epigenetic disease: DNA methylation and chromatin alterations in human tumours. *J Pathol*, 196, 1-7.

Esteller, M., Tortola, S., Toyota, M., Capella, G., Peinado, M. A., Baylin, S. B. & Herman, J. G. (2000) Hypermethylation-associated inactivation of p14(ARF) is independent of p16(INK4a) methylation and p53 mutational status. *Cancer Res*, 60, 129-33.

Feinberg, A. P. & Vogelstein, B. (1983) Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature*, 301, 89-92.

Fiegler, H., Carr, P., Douglas, E. J., Burford, D. C., Hunt, S., Scott, C. E., Smith, J., Vetrie, D., Gorman, P., Tomlinson, I. P. & Carter, N. P. (2003) DNA microarrays for comparative genomic hybridization based on DOP-PCR amplification of BAC and PAC clones. *Genes Chromosomes Cancer*, 36, 361-74.

Finlay, C. A., Hinds, P. W. & Levine, A. J. (1989) The p53 proto-oncogene can act as a suppressor of transformation. *Cell*, 57, 1083-93.

Forshe, T., Tatevossian, R. G., Lawson, A. R., Ma, J., Neale, G., Ogunkolade, B. W., Jones, T. A., Aarum, J., Dalton, J., Bailey, S., Chaplin, T., Carter, R. L., Gajjar, A., Broniscer, A., Young, B. D., Ellison, D. W. & Sheer, D. (2009) Activation of the ERK/MAPK pathway: a signature genetic defect in posterior fossa pilocytic astrocytomas. *J Pathol*, 218, 172-81.

Friday, B. B. & Adjei, A. A. (2008) Advances in targeting the Ras/Raf/MEK/Erk mitogen-activated protein kinase cascade with MEK inhibitors for cancer therapy. *Clin Cancer Res*, 14, 342-6.

Friedman, H. S., Burger, P. C., Bigner, S. H., Trojanowski, J. Q., Brodeur, G. M., He, X. M., Wikstrand, C. J., Kurtzberg, J., Berens, M. E., Halperin, E. C. & Et Al. (1988) Phenotypic and genotypic analysis of a human medulloblastoma cell line and transplantable xenograft (D341 Med) demonstrating amplification of c-myc. *Am J Pathol*, 130, 472-84.

Friend, S. H., Bernards, R., Rogelj, S., Weinberg, R. A., Rapaport, J. M., Albert, D. M. & Dryja, T. P. (1986) A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature*, 323, 643-6.

Furnari, F. B., Fenton, T., Bachoo, R. M., Mukasa, A., Stommel, J. M., Stegh, A., Hahn, W. C., Ligon, K. L., Louis, D. N., Brennan, C., Chin, L., Depinho, R. A. & Cavenee, W. K. (2007) Malignant astrocytic glioma: genetics, biology, and paths to treatment. *Genes Dev*, 21, 2683-710.

Fuxe, J., Akusjarvi, G., Goike, H. M., Roos, G., Collins, V. P. & Pettersson, R. F. (2000) Adenovirus-mediated overexpression of p15INK4B inhibits human glioma cell growth, induces replicative senescence, and inhibits telomerase activity similarly to p16INK4A. *Cell Growth Differ*, 11, 373-84.

Gaasenbeek, M., Howarth, K., Rowan, A. J., Gorman, P. A., Jones, A., Chaplin, T., Liu, Y., Bicknell, D., Davison, E. J., Fiegler, H., Carter, N. P., Roylance, R. R. & Tomlinson, I. P. (2006) Combined array-comparative genomic hybridization and single-nucleotide polymorphism-loss of heterozygosity analysis reveals complex changes and multiple forms of chromosomal instability in colorectal cancers. *Cancer Res*, 66, 3471-9.

Gajjar, A., Sanford, R. A., Heideman, R., Jenkins, J. J., Walter, A., Li, Y., Langston, J. W., Muhlbauer, M., Boyett, J. M. & Kun, L. E. (1997) Low-grade astrocytoma: a decade of experience at St. Jude Children's Research Hospital. *J Clin Oncol*, 15, 2792-9.

Gama-Sosa, M. A., Midgett, R. M., Slagel, V. A., Githens, S., Kuo, K. C., Gehrke, C. W. & Ehrlich, M. (1983) Tissue-specific differences in DNA methylation in various mammals. *Biochim Biophys Acta*, 740, 212-9.

Garson, J. A., Tedder, R. S., Briggs, M., Tuke, P., Glazebrook, J. A., Trute, A., Parker, D., Barbara, J. A., Contreras, M. & Aloysius, S. (1990) Detection of hepatitis

C viral sequences in blood donations by "nested" polymerase chain reaction and prediction of infectivity. *Lancet*, 335, 1419-22.

Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A. J., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J. Y. & Zhang, J. (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol*, 5, R80.

Geraci, M., Birch, J. M., Alston, R. D., Moran, A. & Eden, T. O. (2007) Cancer mortality in 13 to 29-year-olds in England and Wales, 1981-2005. *Br J Cancer*, 97, 1588-94.

Gilbertson, R. J. (2004) Medulloblastoma: signalling a change in treatment. *Lancet Oncol*, 5, 209-18.

Gilbertson, R. J. & Ellison, D. W. (2008) The origins of medulloblastoma subtypes. *Annu Rev Pathol*, 3, 341-65.

Goel, V. K., Lazar, A. J., Warneke, C. L., Redston, M. S. & Haluska, F. G. (2006) Examination of mutations in BRAF, NRAS, and PTEN in primary cutaneous melanoma. *J Invest Dermatol*, 126, 154-60.

Gorlin, R. J. (1987) Nevoid basal-cell carcinoma syndrome. *Medicine (Baltimore)*, 66, 98-113.

Gottardo, N. G. & Gajjar, A. (2008) Chemotherapy for malignant brain tumors of childhood. *J Child Neurol*, 23, 1149-59.

Greaves, M. (2006) Infection, immune responses and the aetiology of childhood leukaemia. *Nat Rev Cancer*, 6, 193-203.

Greaves, M. (2007) Darwinian medicine: a case for cancer. *Nat Rev Cancer*, 7, 213-21.

Greaves, M. F. & Wiemels, J. (2003) Origins of chromosome translocations in childhood leukaemia. *Nat Rev Cancer*, 3, 639-49.

Green, A. J., Johnson, P. H. & Yates, J. R. (1994) The tuberous sclerosis gene on chromosome 9q34 acts as a growth suppressor. *Hum Mol Genet*, 3, 1833-4.

Groffen, J., Stephenson, J. R., Heisterkamp, N., De Klein, A., Bartram, C. R. & Grosveld, G. (1984) Philadelphia chromosomal breakpoints are clustered within a limited region, bcr, on chromosome 22. *Cell*, 36, 93-9.

- Grovas, A., Fremgen, A., Rauck, A., Ruymann, F. B., Hutchinson, C. L., Winchester, D. P. & Menck, H. R. (1997) The National Cancer Data Base report on patterns of childhood cancers in the United States. *Cancer*, 80, 2321-32.
- Gu, W., Zhang, F. & Lupski, J. R. (2008) Mechanisms for human genomic rearrangements. *Pathogenetics*, 1, 4.
- Gustafson, S., Zbuk, K. M., Scacheri, C. & Eng, C. (2007) Cowden syndrome. *Semin Oncol*, 34, 428-34.
- Hahn, H., Wicking, C., Zaphiropoulous, P. G., Gailani, M. R., Shanley, S., Chidambaram, A., Vorechovsky, I., Holmberg, E., Unden, A. B., Gillies, S., Negus, K., Smyth, I., Pressman, C., Leffell, D. J., Gerrard, B., Goldstein, A. M., Dean, M., Toftgard, R., Chenevix-Trench, G., Wainwright, B. & Bale, A. E. (1996) Mutations of the human homolog of Drosophila patched in the nevoid basal cell carcinoma syndrome. *Cell*, 85, 841-51.
- Hanahan, D. & Weinberg, R. A. (2000) The hallmarks of cancer. *Cell*, 100, 57-70.
- Hapmap, T. I. (2003) The International HapMap Project. *Nature*, 426, 789-96.
- Hargrave, D. (2009) Paediatric high and low grade glioma: the impact of tumour biology on current and future therapy. *Br J Neurosurg*, 23, 351-63.
- Harvey, E. B., Boice, J. D., Jr., Honeyman, M. & Flannery, J. T. (1985) Prenatal x-ray exposure and childhood cancer in twins. *N Engl J Med*, 312, 541-5.
- Hastings, P. J., Ira, G. & Lupski, J. R. (2009) A microhomology-mediated break-induced replication model for the origin of human copy number variation. *PLoS Genet*, 5, e1000327.
- Hayry, V., Tanner, M., Blom, T., Tynninen, O., Roselli, A., Ollikainen, M., Sariola, H., Wartiovaara, K. & Nupponen, N. N. (2008) Copy number alterations of the polycomb gene BMI1 in gliomas. *Acta Neuropathol*, 116, 97-102.
- Heasley, L. E. & Johnson, G. L. (1992) The beta-PDGF receptor induces neuronal differentiation of PC12 cells. *Mol Biol Cell*, 3, 545-53.
- Heerema, N. A., Nachman, J. B., Sather, H. N., Sensel, M. G., Lee, M. K., Hutchinson, R., Lange, B. J., Steinherz, P. G., Bostrom, B., Gaynon, P. S. & Uckun, F. (1999) Hypodiploidy with less than 45 chromosomes confers adverse risk in childhood acute lymphoblastic leukemia: a report from the children's cancer group. *Blood*, 94, 4036-45.
- Hegi, M. E., Diserens, A. C., Gorlia, T., Hamou, M. F., De Tribolet, N., Weller, M., Kros, J. M., Hainfellner, J. A., Mason, W., Mariani, L., Bromberg, J. E., Hau, P.,

- Mirimanoff, R. O., Cairncross, J. G., Janzer, R. C. & Stupp, R. (2005) MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med*, 352, 997-1003.
- Heim, S. & Mitelman, F. (2009) *Cancer Cytogenetics*, Wiley-Blackwell.
- Heimer, L. (1983) *The human brain and spinal cord*, New York, Springer-Verlag.
- Hermansson, M., Nister, M., Betsholtz, C., Heldin, C. H., Westermark, B. & Funai, K. (1988) Endothelial cell hyperplasia in human glioblastoma: coexpression of mRNA for platelet-derived growth factor (PDGF) B chain and PDGF receptor suggests autocrine growth stimulation. *Proc Natl Acad Sci U S A*, 85, 7748-52.
- Hinds, D. A., Stuve, L. L., Nilsen, G. B., Halperin, E., Eskin, E., Ballinger, D. G., Frazer, K. A. & Cox, D. R. (2005) Whole-genome patterns of common DNA variation in three human populations. *Science*, 307, 1072-9.
- Hirose, Y., Aldape, K. D., Chang, S., Lamborn, K., Berger, M. S. & Feuerstein, B. G. (2003) Grade II astrocytomas are subgrouped by chromosome aberrations. *Cancer Genet Cytogenet*, 142, 1-7.
- Howard, G., Eiges, R., Gaudet, F., Jaenisch, R. & Eden, A. (2008) Activation and transposition of endogenous retroviral elements in hypomethylation induced tumors in mice. *Oncogene*, 27, 404-8.
- Huang, H. S., Nagane, M., Klingbeil, C. K., Lin, H., Nishikawa, R., Ji, X. D., Huang, C. M., Gill, G. N., Wiley, H. S. & Cavenee, W. K. (1997) The enhanced tumorigenic activity of a mutant epidermal growth factor receptor common in human cancers is mediated by threshold levels of constitutive tyrosine phosphorylation and unattenuated signaling. *J Biol Chem*, 272, 2927-35.
- Huang, J., Wei, W., Zhang, J., Liu, G., Bignell, G. R., Stratton, M. R., Futreal, P. A., Wooster, R., Jones, K. W. & Shaperro, M. H. (2004) Whole genome DNA copy number changes identified by high density oligonucleotide arrays. *Hum Genomics*, 1, 287-99.
- Hughes, S., Damato, B. E., Giddings, I., Hiscott, P. S., Humphreys, J. & Houlston, R. S. (2005) Microarray comparative genomic hybridisation analysis of intraocular uveal melanomas identifies distinctive imbalances associated with loss of chromosome 3. *Br J Cancer*, 93, 1191-6.
- Huson, S. M., Compston, D. A., Clark, P. & Harper, P. S. (1989) A genetic study of von Recklinghausen neurofibromatosis in south east Wales. I. Prevalence, fitness,

mutation rate, and effect of parental transmission on severity. *J Med Genet*, 26, 704-11.

Iafrate, A. J., Feuk, L., Rivera, M. N., Listewnik, M. L., Donahoe, P. K., Qi, Y., Scherer, S. W. & Lee, C. (2004) Detection of large-scale variation in the human genome. *Nat Genet*, 36, 949-51.

Ichimura, K., Bolin, M. B., Goike, H. M., Schmidt, E. E., Moshref, A. & Collins, V. P. (2000) Deregulation of the p14ARF/MDM2/p53 pathway is a prerequisite for human astrocytic gliomas with G1-S transition control gene abnormalities. *Cancer Res*, 60, 417-24.

Ichimura, K., Schmidt, E. E., Goike, H. M. & Collins, V. P. (1996) Human glioblastomas with no alterations of the CDKN2A (p16INK4A, MTS1) and CDK4 genes have frequent mutations of the retinoblastoma gene. *Oncogene*, 13, 1065-72.

Idbaih, A., Carvalho Silva, R., Criniere, E., Marie, Y., Carpentier, C., Boisselier, B., Taillibert, S., Rousseau, A., Mokhtari, K., Ducray, F., Thillet, J., Sanson, M., Hoang-Xuan, K. & Delattre, J. Y. (2008) Genomic changes in progression of low-grade gliomas. *J Neurooncol*, 90, 133-40.

Idbaih, A., Marie, Y., Pierron, G., Brennetot, C., Hoang-Xuan, K., Kujas, M., Mokhtari, K., Sanson, M., Lejeune, J., Aurias, A., Delattre, O. & Delattre, J. Y. (2005) Two types of chromosome 1p losses with opposite significance in gliomas. *Ann Neurol*, 58, 483-7.

Ikediobi, O. N., Davies, H., Bignell, G., Edkins, S., Stevens, C., O'meara, S., Santarius, T., Avis, T., Barthorpe, S., Brackenbury, L., Buck, G., Butler, A., Clements, J., Cole, J., Dicks, E., Forbes, S., Gray, K., Halliday, K., Harrison, R., Hills, K., Hinton, J., Hunter, C., Jenkinson, A., Jones, D., Kosmidou, V., Lugg, R., Menzies, A., Mironenko, T., Parker, A., Perry, J., Raine, K., Richardson, D., Shepherd, R., Small, A., Smith, R., Solomon, H., Stephens, P., Teague, J., Tofts, C., Varian, J., Webb, T., West, S., Widaa, S., Yates, A., Reinhold, W., Weinstein, J. N., Stratton, M. R., Futreal, P. A. & Wooster, R. (2006) Mutation analysis of 24 known cancer genes in the NCI-60 cell line set. *Mol Cancer Ther*, 5, 2606-12.

Jackson, S. P. & Bartek, J. (2009) The DNA-damage response in human biology and disease. *Nature*, 461, 1071-8.

Jacobsen, P. F., Jenkyn, D. J. & Papadimitriou, J. M. (1985) Establishment of a human medulloblastoma cell line and its heterotransplantation into nude mice. *J Neuropathol Exp Neurol*, 44, 472-85.

Janisch, W., Schreiber, D., Martin, H. & Gerlach, H. (1985) [Diencephalic pilocytic astrocytoma with clinical onset in infancy. Biological behavior and pathomorphological findings in 11 children]. *Zentralbl Allg Pathol*, 130, 31-43.

Janzarik, W. G., Kratz, C. P., Loges, N. T., Olbrich, H., Klein, C., Schafer, T., Scheurlen, W., Roggendorf, W., Weiller, C., Niemeyer, C., Korinthenberg, R., Pfister, S. & Omran, H. (2007) Further evidence for a somatic KRAS mutation in a pilocytic astrocytoma. *Neuropediatrics*, 38, 61-3.

Jenkins, R. B., Kimmel, D. W., Moertel, C. A., Schultz, C. G., Scheithauer, B. W., Kelly, P. J. & Dewald, G. W. (1989) A cytogenetic study of 53 human gliomas. *Cancer Genet Cytogenet*, 39, 253-79.

Jenuwein, T. & Allis, C. D. (2001) Translating the histone code. *Science*, 293, 1074-80.

Johnson, G. L. & Lapadat, R. (2002) Mitogen-activated protein kinase pathways mediated by ERK, JNK, and p38 protein kinases. *Science*, 298, 1911-2.

Johnson, R. L., Rothman, A. L., Xie, J., Goodrich, L. V., Bare, J. W., Bonifas, J. M., Quinn, A. G., Myers, R. M., Cox, D. R., Epstein, E. H., Jr. & Scott, M. P. (1996) Human homolog of patched, a candidate gene for the basal cell nevus syndrome. *Science*, 272, 1668-71.

Jones, D. T., Ichimura, K., Liu, L., Pearson, D. M., Plant, K. & Collins, V. P. (2006) Genomic analysis of pilocytic astrocytomas at 0.97 Mb resolution shows an increasing tendency toward chromosomal copy number change with age. *J Neuropathol Exp Neurol*, 65, 1049-58.

Jones, D. T., Kocialkowski, S., Liu, L., Pearson, D. M., Backlund, L. M., Ichimura, K. & Collins, V. P. (2008) Tandem duplication producing a novel oncogenic BRAF fusion gene defines the majority of pilocytic astrocytomas. *Cancer Res*, 68, 8673-7.

Jones, D. T., Kocialkowski, S., Liu, L., Pearson, D. M., Ichimura, K. & Collins, V. P. (2009) Oncogenic RAF1 rearrangement and a novel BRAF mutation as alternatives to KIAA1549:BRAF fusion in activating the MAPK pathway in pilocytic astrocytoma. *Oncogene*, 28, 2119-23.

Joos, S., Falk, M. H., Lichter, P., Haluska, F. G., Henglein, B., Lenoir, G. M. & Bornkamm, G. W. (1992) Variable breakpoints in Burkitt lymphoma cells with chromosomal t(8;14) translocation separate c-myc and the IgH locus up to several hundred kb. *Hum Mol Genet*, 1, 625-32.

- 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.
- Karnes, P. S., Tran, T. N., Cui, M. Y., Raffel, C., Gilles, F. H., Barranger, J. A. & Ying, K. L. (1992) Cytogenetic analysis of 39 pediatric central nervous system tumors. *Cancer Genet Cytogenet*, 59, 12-9.
- Karpf, A. R. & Matsui, S. (2005) Genetic disruption of cytosine DNA methyltransferase enzymes induces chromosomal instability in human cancer cells. *Cancer Res*, 65, 8635-9.
- Kennedy, G. C., Matsuzaki, H., Dong, S., Liu, W. M., Huang, J., Liu, G., Su, X., Cao, M., Chen, W., Zhang, J., Liu, W., Yang, G., Di, X., Ryder, T., He, Z., Surti, U., Phillips, M. S., Boyce-Jacino, M. T., Fodor, S. P. & Jones, K. W. (2003) Large-scale genotyping of complex DNA. *Nat Biotechnol*, 21, 1233-7.
- Kho, A. T., Zhao, Q., Cai, Z., Butte, A. J., Kim, J. Y., Pomeroy, S. L., Rowitch, D. H. & Kohane, I. S. (2004) Conserved mechanisms across development and tumorigenesis revealed by a mouse development perspective of human cancers. *Genes Dev*, 18, 629-40.
- Kleihues, P. & Cavenee, W. K. (2000) *Pathology and genetics of tumours of the nervous system.*, Lyon, International Agency for Research on Cancer (IARC Press).
- Kluwe, L., Hagel, C., Tatagiba, M., Thomas, S., Stavrou, D., Ostertag, H., Von Deimling, A. & Mautner, V. F. (2001) Loss of NF1 alleles distinguish sporadic from NF1-associated pilocytic astrocytomas. *J Neuropathol Exp Neurol*, 60, 917-20.
- Knight, S. J., Lese, C. M., Precht, K. S., Kuc, J., Ning, Y., Lucas, S., Regan, R., Brenan, M., Nicod, A., Lawrie, N. M., Cardy, D. L., Nguyen, H., Hudson, T. J., Riethman, H. C., Ledbetter, D. H. & Flint, J. (2000) An optimized set of human telomere clones for studying telomere integrity and architecture. *Am J Hum Genet*, 67, 320-32.
- Knudson, A. G. (2001) Two genetic hits (more or less) to cancer. *Nat Rev Cancer*, 1, 157-62.
- 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. (1985) Hereditary cancer, oncogenes, and antioncogenes. *Cancer Res*, 45, 1437-43.

Kohno, T. & Yokota, J. (2006) Molecular processes of chromosome 9p21 deletions causing inactivation of the p16 tumor suppressor gene in human cancer: deduction from structural analysis of breakpoints for deletions. *DNA Repair (Amst)*, 5, 1273-81.

Korn, J. M., Kuruvilla, F. G., Mccarroll, S. A., Wysoker, A., Nemesh, J., Cawley, S., Hubbell, E., Veitch, J., Collins, P. J., Darvishi, K., Lee, C., Nizzari, M. M., Gabriel, S. B., Purcell, S., Daly, M. J. & Altshuler, D. (2008) Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nat Genet*, 40, 1253-60.

Kornberg, R. D. (1974) Chromatin structure: a repeating unit of histones and DNA. *Science*, 184, 868-71.

Korshunov, A., Meyer, J., Capper, D., Christians, A., Remke, M., Witt, H., Pfister, S., Von Deimling, A. & Hartmann, C. (2009) Combined molecular analysis of BRAF and IDH1 distinguishes pilocytic astrocytoma from diffuse astrocytoma. *Acta Neuropathol*, 118, 401-5.

Korshunov, A., Sycheva, R. & Golanov, A. (2006) Genetically distinct and clinically relevant subtypes of glioblastoma defined by array-based comparative genomic hybridization (array-CGH). *Acta Neuropathol*, 111, 465-74.

Kreutzberg, G. W. (1996) Microglia: a sensor for pathological events in the CNS. *Trends Neurosci*, 19, 312-8.

Krontiris, T. G. & Cooper, G. M. (1981) Transforming activity of human tumor DNAs. *Proc Natl Acad Sci U S A*, 78, 1181-4.

Kruglyak, L. & Nickerson, D. A. (2001) Variation is the spice of life. *Nat Genet*, 27, 234-6.

Kuppers, R. (2005) Mechanisms of B-cell lymphoma pathogenesis. *Nat Rev Cancer*, 5, 251-62.

Lafay-Cousin, L., Sung, L., Carret, A. S., Hukin, J., Wilson, B., Johnston, D. L., Zelcer, S., Silva, M., Odame, I., Mpofo, C., Strother, D. & Bouffet, E. (2008) Carboplatin hypersensitivity reaction in pediatric patients with low-grade glioma: a Canadian Pediatric Brain Tumor Consortium experience. *Cancer*, 112, 892-9.

Laframboise, T. (2009) Single nucleotide polymorphism arrays: a decade of biological, computational and technological advances. *Nucleic Acids Res*, 37, 4181-93.

- Lancaster, D. L., Hoddes, J. A. & Michalski, A. (2003) Tolerance of nitrosurea-based multiagent chemotherapy regime for low-grade pediatric gliomas. *J Neurooncol*, 63, 289-94.
- Ledbetter, D. H., Rich, D. C., O'connell, P., Leppert, M. & Carey, J. C. (1989) Precise localization of NF1 to 17q11.2 by balanced translocation. *Am J Hum Genet*, 44, 20-4.
- Lee, J. A., Carvalho, C. M. & Lupski, J. R. (2007) A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *Cell*, 131, 1235-47.
- Lee, J. W., Soung, Y. H., Kim, S. Y., Park, W. S., Nam, S. W., Min, W. S., Kim, S. H., Lee, J. Y., Yoo, N. J. & Lee, S. H. (2005a) Mutational analysis of the ARAF gene in human cancers. *Apmis*, 113, 54-7.
- Lee, R. C., Feinbaum, R. L. & Ambros, V. (1993) The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*, 75, 843-54.
- Lee, W. J., Shim, J. Y. & Zhu, B. T. (2005b) Mechanisms for the inhibition of DNA methyltransferases by tea catechins and bioflavonoids. *Mol Pharmacol*, 68, 1018-30.
- Lengauer, C., Kinzler, K. W. & Vogelstein, B. (1997) Genetic instability in colorectal cancers. *Nature*, 386, 623-7.
- Lengauer, C., Kinzler, K. W. & Vogelstein, B. (1998) Genetic instabilities in human cancers. *Nature*, 396, 643-9.
- Levine, A. J. (1997) p53, the cellular gatekeeper for growth and division. *Cell*, 88, 323-31.
- Lewis, J. H., Ginsberg, A. L. & Toomey, K. E. (1983) Turcot's syndrome. Evidence for autosomal dominant inheritance. *Cancer*, 51, 524-8.
- Li, F. P., Fraumeni, J. F., Jr., Mulvihill, J. J., Blattner, W. A., Dreyfus, M. G., Tucker, M. A. & Miller, R. W. (1988) A cancer family syndrome in twenty-four kindreds. *Cancer Res*, 48, 5358-62.
- Lichter, P., Joos, S., Bentz, M. & Lampel, S. (2000) Comparative genomic hybridization: uses and limitations. *Semin Hematol*, 37, 348-57.
- Lieber, M. R. (2008) The mechanism of human nonhomologous DNA end joining. *J Biol Chem*, 283, 1-5.
- Lieber, M. R., Lu, H., Gu, J. & Schwarz, K. (2008) Flexibility in the order of action and in the enzymology of the nuclease, polymerases, and ligase of vertebrate non-homologous DNA end joining: relevance to cancer, aging, and the immune system. *Cell Res*, 18, 125-33.

- Lindblad-Toh, K., Tanenbaum, D. M., Daly, M. J., Winchester, E., Lui, W. O., Villapakkam, A., Stanton, S. E., Larsson, C., Hudson, T. J., Johnson, B. E., Lander, E. S. & Meyerson, M. (2000) Loss-of-heterozygosity analysis of small-cell lung carcinomas using single-nucleotide polymorphism arrays. *Nat Biotechnol*, 18, 1001-5.
- Listernick, R., Ferner, R. E., Liu, G. T. & Gutmann, D. H. (2007) Optic pathway gliomas in neurofibromatosis-1: controversies and recommendations. *Ann Neurol*, 61, 189-98.
- Liu, G., Li, W., Wang, L., Kar, A., Guan, K. L., Rao, Y. & Wu, J. Y. (2009) DSCAM functions as a netrin receptor in commissural axon pathfinding. *Proc Natl Acad Sci U S A*, 106, 2951-6.
- Liu, G., Ying, H., Zeng, G., Wheeler, C. J., Black, K. L. & Yu, J. S. (2004) HER-2, gp100, and MAGE-1 are expressed in human glioblastoma and recognized by cytotoxic T cells. *Cancer Res*, 64, 4980-6.
- Liu, H., Zhou, Y., Boggs, S. E., Belinsky, S. A. & Liu, J. (2007) Cigarette smoke induces demethylation of prometastatic oncogene synuclein-gamma in lung cancer cells by downregulation of DNMT3B. *Oncogene*, 26, 5900-10.
- Llovet, J. M., Ricci, S., Mazzaferro, V., Hilgard, P., Gane, E., Blanc, J. F., De Oliveira, A. C., Santoro, A., Raoul, J. L., Forner, A., Schwartz, M., Porta, C., Zeuzem, S., Bolondi, L., Greten, T. F., Galle, P. R., Seitz, J. F., Borbath, I., Haussinger, D., Giannaris, T., Shan, M., Moscovici, M., Voliotis, D. & Bruix, J. (2008) Sorafenib in advanced hepatocellular carcinoma. *N Engl J Med*, 359, 378-90.
- Loeb, L. A. (1991) Mutator phenotype may be required for multistage carcinogenesis. *Cancer Res*, 51, 3075-9.
- Lorente, A., Mueller, W., Urdangarin, E., Lazcoz, P., Lass, U., Von Deimling, A. & Castresana, J. S. (2009) RASSF1A, BLU, NORE1A, PTEN and MGMT expression and promoter methylation in gliomas and glioma cell lines and evidence of deregulated expression of de novo DNMTs. *Brain Pathol*, 19, 279-92.
- Louis, D. N., Ohgaki, H., Wiestler, O. D., Cavenee, W. K., Burger, P. C., Jouvett, A., Scheithauer, B. W. & Kleihues, P. (2007) The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol*, 114, 97-109.
- Lugo, T. G., Pendergast, A. M., Muller, A. J. & Witte, O. N. (1990) Tyrosine kinase activity and transformation potency of bcr-abl oncogene products. *Science*, 247, 1079-82.

Lujambio, A., Ropero, S., Ballestar, E., Fraga, M. F., Cerrato, C., Setien, F., Casado, S., Suarez-Gauthier, A., Sanchez-Cespedes, M., Git, A., Spiteri, I., Das, P. P., Caldas, C., Miska, E. & Esteller, M. (2007) Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. *Cancer Res*, 67, 1424-9.

Lynch, H. T. & De La Chapelle, A. (2003) Hereditary colorectal cancer. *N Engl J Med*, 348, 919-32.

Macabeo-Ong, M., Ginzinger, D. G., Dekker, N., Mcmillan, A., Regezi, J. A., Wong, D. T. & Jordan, R. C. (2002) Effect of duration of fixation on quantitative reverse transcription polymerase chain reaction analyses. *Mod Pathol*, 15, 979-87.

Maher, C. A., Kumar-Sinha, C., Cao, X., Kalyana-Sundaram, S., Han, B., Jing, X., Sam, L., Barrette, T., Palanisamy, N. & Chinnaiyan, A. M. (2009) Transcriptome sequencing to detect gene fusions in cancer. *Nature*, 458, 97-101.

Maher, E. A., Brennan, C., Wen, P. Y., Durso, L., Ligon, K. L., Richardson, A., Khatry, D., Feng, B., Sinha, R., Louis, D. N., Quackenbush, J., Black, P. M., Chin, L. & Depinho, R. A. (2006) Marked genomic differences characterize primary and secondary glioblastoma subtypes and identify two distinct molecular and clinical secondary glioblastoma entities. *Cancer Res*, 66, 11502-13.

Malkin, D., Li, F. P., Strong, L. C., Fraumeni, J. F., Jr., Nelson, C. E., Kim, D. H., Kassel, J., Gryka, M. A., Bischoff, F. Z., Tainsky, M. A. & Et Al. (1990) Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science*, 250, 1233-8.

Maltzman, T. H., Mueller, B. A., Schroeder, J., Rutledge, J. C., Patterson, K., Preston-Martin, S. & Faustman, E. M. (1997) Ras oncogene mutations in childhood brain tumors. *Cancer Epidemiol Biomarkers Prev*, 6, 239-43.

Martinez-Climent, J. A., Alizadeh, A. A., Seagraves, R., Blesa, D., Rubio-Moscardo, F., Albertson, D. G., Garcia-Conde, J., Dyer, M. J., Levy, R., Pinkel, D. & Lossos, I. S. (2003) Transformation of follicular lymphoma to diffuse large cell lymphoma is associated with a heterogeneous set of DNA copy number and gene expression alterations. *Blood*, 101, 3109-17.

Martuza, R. L. & Eldridge, R. (1988) Neurofibromatosis 2 (bilateral acoustic neurofibromatosis). *N Engl J Med*, 318, 684-8.

Mathews, L. A., Crea, F. & Farrar, W. L. (2009) Epigenetic gene regulation in stem cells and correlation to cancer. *Differentiation*, 78, 1-17.

Mcgillicuddy, L. T., Fromm, J. A., Hollstein, P. E., Kubek, S., Beroukhir, R., De Raedt, T., Johnson, B. W., Williams, S. M., Nghiemphu, P., Liao, L. M., Cloughesy, T. F., Mischel, P. S., Parret, A., Seiler, J., Moldenhauer, G., Scheffzek, K., Stemmer-Rachamimov, A. O., Sawyers, C. L., Brennan, C., Messiaen, L., Mellingshoff, I. K. & Cichowski, K. (2009) Proteasomal and genetic inactivation of the NF1 tumor suppressor in gliomagenesis. *Cancer Cell*, 16, 44-54.

Meadows, A. T., Baum, E., Fossati-Bellani, F., Green, D., Jenkin, R. D., Marsden, B., Nesbit, M., Newton, W., Oberlin, O., Sallan, S. G. & Et Al. (1985) Second malignant neoplasms in children: an update from the Late Effects Study Group. *J Clin Oncol*, 3, 532-8.

Meyn, M. A., 3rd, Wilson, M. B., Abdi, F. A., Fahey, N., Schiavone, A. P., Wu, J., Hochrein, J. M., Engen, J. R. & Smithgall, T. E. (2006) Src family kinases phosphorylate the Bcr-Abl SH3-SH2 region and modulate Bcr-Abl transforming activity. *J Biol Chem*, 281, 30907-16.

Migliaccio, A., Di Domenico, M., Castoria, G., De Falco, A., Bontempo, P., Nola, E. & Auricchio, F. (1996) Tyrosine kinase/p21ras/MAP-kinase pathway activation by estradiol-receptor complex in MCF-7 cells. *Embo J*, 15, 1292-300.

Misra, A., Pellarin, M., Nigro, J., Smirnov, I., Moore, D., Lamborn, K. R., Pinkel, D., Albertson, D. G. & Feuerstein, B. G. (2005) Array comparative genomic hybridization identifies genetic subgroups in grade 4 human astrocytoma. *Clin Cancer Res*, 11, 2907-18.

Mitelman, F., Johansson, B. & Mertens, F. (2007) The impact of translocations and gene fusions on cancer causation. *Nat Rev Cancer*, 7, 233-45.

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. P. & Cote, R. J. (2009) Molecular pathogenesis and diagnostics of bladder cancer. *Annu Rev Pathol*, 4, 251-85.

Miyagi, K., Mukawa, J., Kinjo, N., Horikawa, K., Mekar, S., Nakasone, S., Koga, H., Higa, Y. & Naito, M. (1995) Astrocytoma linked to familial ataxia-telangiectasia. *Acta Neurochir (Wien)*, 135, 87-92.

Mori, H., Colman, S. M., Xiao, Z., Ford, A. M., Healy, L. E., Donaldson, C., Hows, J. M., Navarrete, C. & Greaves, M. (2002) Chromosome translocations and covert leukemic clones are generated during normal fetal development. *Proc Natl Acad Sci U S A*, 99, 8242-7.

- Mulhern, R. K., Merchant, T. E., Gajjar, A., Reddick, W. E. & Kun, L. E. (2004) Late neurocognitive sequelae in survivors of brain tumours in childhood. *Lancet Oncol*, 5, 399-408.
- Mulholland, P. J., Fiegler, H., Mazzanti, C., Gorman, P., Sasieni, P., Adams, J., Jones, T. A., Babbage, J. W., Vatcheva, R., Ichimura, K., East, P., Poullikas, C., Collins, V. P., Carter, N. P., Tomlinson, I. P. & Sheer, D. (2006) Genomic profiling identifies discrete deletions associated with translocations in glioblastoma multiforme. *Cell Cycle*, 5, 783-91.
- Mullis, K. B. & Faloona, F. A. (1987) Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol*, 155, 335-50.
- Muroya, K., Hattori, S. & Nakamura, S. (1992) Nerve growth factor induces rapid accumulation of the GTP-bound form of p21ras in rat pheochromocytoma PC12 cells. *Oncogene*, 7, 277-81.
- Nagarajan, R. P. & Costello, J. F. (2009) Epigenetic mechanisms in glioblastoma multiforme. *Semin Cancer Biol*, 19, 188-97.
- Nagase, T., Kikuno, R., Nakayama, M., Hirose, M. & Ohara, O. (2000) Prediction of the coding sequences of unidentified human genes. XVIII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro. *DNA Res*, 7, 273-81.
- Nakagawa, T., Kanai, Y., Ushijima, S., Kitamura, T., Kakizoe, T. & Hirohashi, S. (2005) DNA hypomethylation on pericentromeric satellite regions significantly correlates with loss of heterozygosity on chromosome 9 in urothelial carcinomas. *J Urol*, 173, 243-6.
- Nakamura, T., Mori, T., Tada, S., Krajewski, W., Rozovskaia, T., Wassell, R., Dubois, G., Mazo, A., Croce, C. M. & Canaani, E. (2002) ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. *Mol Cell*, 10, 1119-28.
- Nannya, Y., Sanada, M., Nakazaki, K., Hosoya, N., Wang, L., Hangaishi, A., Kurokawa, M., Chiba, S., Bailey, D. K., Kennedy, G. C. & Ogawa, S. (2005) A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. *Cancer Res*, 65, 6071-9.
- Napolioni, V. & Curatolo, P. (2008) Genetics and molecular biology of tuberous sclerosis complex. *Curr Genomics*, 9, 475-87.

- Neglia, J. P., Meadows, A. T., Robison, L. L., Kim, T. H., Newton, W. A., Ruymann, F. B., Sather, H. N. & Hammond, G. D. (1991) Second neoplasms after acute lymphoblastic leukemia in childhood. *N Engl J Med*, 325, 1330-6.
- Nicoloso, M. S. & Calin, G. A. (2008) MicroRNA involvement in brain tumors: from bench to bedside. *Brain Pathol*, 18, 122-9.
- Nowell, P. C. & Hungerford, D. A. (1960) A minute chromosome in human granulocytic leukaemia. *Science*, 1497.
- Oehler, V. G. & Radich, J. P. (2003) Monitoring bcr-abl by polymerase chain reaction in the treatment of chronic myeloid leukemia. *Curr Oncol Rep*, 5, 426-35.
- Ohgaki, H. & Kleihues, P. (2005a) Epidemiology and etiology of gliomas. *Acta Neuropathol*, 109, 93-108.
- Ohgaki, H. & Kleihues, P. (2005b) Population-based studies on incidence, survival rates, and genetic alterations in astrocytic and oligodendroglial gliomas. *J Neuropathol Exp Neurol*, 64, 479-89.
- Ohgaki, H. & Kleihues, P. (2007) Genetic pathways to primary and secondary glioblastoma. *Am J Pathol*, 170, 1445-53.
- Olshen, A. B., Venkatraman, E. S., Lucito, R. & Wigler, M. (2004) Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics*, 5, 557-72.
- Ordóñez, J. L., Osuna, D., Herrero, D., De Alava, E. & Madoz-Gurpide, J. (2009) Advances in Ewing's sarcoma research: where are we now and what lies ahead? *Cancer Res*, 69, 7140-50.
- Ottaviani, D., Lever, E., Takousis, P. & Sheer, D. (2008) Anchoring the genome. *Genome Biol*, 9, 201.
- Packer, R. J., Gurney, J. G., Punyko, J. A., Donaldson, S. S., Inskip, P. D., Stovall, M., Yasui, Y., Mertens, A. C., Sklar, C. A., Nicholson, H. S., Zeltzer, L. K., Neglia, J. P. & Robison, L. L. (2003) Long-term neurologic and neurosensory sequelae in adult survivors of a childhood brain tumor: childhood cancer survivor study. *J Clin Oncol*, 21, 3255-61.
- Packer, R. J., Meadows, A. T., Rorke, L. B., Goldwein, J. L. & D'Angio, G. (1987) Long-term sequelae of cancer treatment on the central nervous system in childhood. *Med Pediatr Oncol*, 15, 241-53.
- Pallini, R., Ricci-Vitiani, L., Banna, G. L., Signore, M., Lombardi, D., Todaro, M., Stassi, G., Martini, M., Maira, G., Larocca, L. M. & De Maria, R. (2008) Cancer stem

cell analysis and clinical outcome in patients with glioblastoma multiforme. *Clin Cancer Res*, 14, 8205-12.

Papaemmanuil, E., Hosking, F. J., Vijayakrishnan, J., Price, A., Olver, B., Sheridan, E., Kinsey, S. E., Lightfoot, T., Roman, E., Irving, J. A., Allan, J. M., Tomlinson, I. P., Taylor, M., Greaves, M. & Houlston, R. S. (2009) Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. *Nat Genet*, 41, 1006-10.

Paris, P. L., Andaya, A., Fridlyand, J., Jain, A. N., Weinberg, V., Kowbel, D., Brebner, J. H., Simko, J., Watson, J. E., Volik, S., Albertson, D. G., Pinkel, D., Alers, J. C., Van Der Kwast, T. H., Vissers, K. J., Schroder, F. H., Wildhagen, M. F., Febbo, P. G., Chinnaiyan, A. M., Pienta, K. J., Carroll, P. R., Rubin, M. A., Collins, C. & Van Dekken, H. (2004) Whole genome scanning identifies genotypes associated with recurrence and metastasis in prostate tumors. *Hum Mol Genet*, 13, 1303-13.

Parsons, D. W., Jones, S., Zhang, X., Lin, J. C., Leary, R. J., Angenendt, P., Mankoo, P., Carter, H., Siu, I. M., Gallia, G. L., Olivi, A., McLendon, R., Rasheed, B. A., Keir, S., Nikolskaya, T., Nikolsky, Y., Busam, D. A., Tekleab, H., Diaz, L. A., Jr., Hartigan, J., Smith, D. R., Strausberg, R. L., Marie, S. K., Shinjo, S. M., Yan, H., Riggins, G. J., Bigner, D. D., Karchin, R., Papadopoulos, N., Parmigiani, G., Vogelstein, B., Velculescu, V. E. & Kinzler, K. W. (2008) An integrated genomic analysis of human glioblastoma multiforme. *Science*, 321, 1807-12.

Pasca Di Magliano, M. & Hebrok, M. (2003) Hedgehog signalling in cancer formation and maintenance. *Nat Rev Cancer*, 3, 903-11.

Patchell, R. A. (2003) The management of brain metastases. *Cancer Treat Rev*, 29, 533-40.

Patt, S., Gries, H., Giraldo, M., Cervos-Navarro, J., Martin, H., Janisch, W. & Brockmoller, J. (1996) p53 gene mutations in human astrocytic brain tumors including pilocytic astrocytomas. *Hum Pathol*, 27, 586-9.

Payen, C., Koszul, R., Dujon, B. & Fischer, G. (2008) Segmental duplications arise from Pol32-dependent repair of broken forks through two alternative replication-based mechanisms. *PLoS Genet*, 4, e1000175.

Pece, S. & Gutkind, J. S. (2000) Signaling from E-cadherins to the MAPK pathway by the recruitment and activation of epidermal growth factor receptors upon cell-cell contact formation. *J Biol Chem*, 275, 41227-33.

Peris-Bonet, R., Martinez-Garcia, C., Lacour, B., Petrovich, S., Giner-Ripoll, B., Navajas, A. & Steliarova-Foucher, E. (2006) Childhood central nervous system tumours--incidence and survival in Europe (1978-1997): report from Automated Childhood Cancer Information System project. *Eur J Cancer*, 42, 2064-80.

Petrij, F., Giles, R. H., Dauwerse, H. G., Saris, J. J., Hennekam, R. C., Masuno, M., Tommerup, N., Van Ommen, G. J., Goodman, R. H., Peters, D. J. & Et Al. (1995) Rubinstein-Taybi syndrome caused by mutations in the transcriptional co-activator CBP. *Nature*, 376, 348-51.

Petti, C., Molla, A., Vegetti, C., Ferrone, S., Anichini, A. & Sensi, M. (2006) Coexpression of NRASQ61R and BRAFV600E in human melanoma cells activates senescence and increases susceptibility to cell-mediated cytotoxicity. *Cancer Res*, 66, 6503-11.

Pfister, S., Janzarik, W. G., Remke, M., Ernst, A., Werft, W., Becker, N., Toedt, G., Wittmann, A., Kratz, C., Olbrich, H., Ahmadi, R., Thieme, B., Joos, S., Radlwimmer, B., Kulozik, A., Pietsch, T., Herold-Mende, C., Gnekow, A., Reifenberger, G., Korshunov, A., Scheurlen, W., Omran, H. & Lichter, P. (2008) BRAF gene duplication constitutes a mechanism of MAPK pathway activation in low-grade astrocytomas. *J Clin Invest*, 118, 1739-49.

Phelan, C. M., Liu, L., Ruttledge, M. H., Muntzning, K., Ridderheim, P. A. & Collins, V. P. (1995) Chromosome 17 abnormalities and lack of TP53 mutations in paediatric central nervous system tumours. *Hum Genet*, 96, 684-90.

Pinkel, D., Seagraves, R., Sudar, D., Clark, S., Poole, I., Kowbel, D., Collins, C., Kuo, W. L., Chen, C., Zhai, Y., Dairkee, S. H., Ljung, B. M., Gray, J. W. & Albertson, D. G. (1998) High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet*, 20, 207-11.

Pogribny, I. P. & Beland, F. A. (2009) DNA hypomethylation in the origin and pathogenesis of human diseases. *Cell Mol Life Sci*, 66, 2249-61.

Pollack, I. F. (1994) Brain tumors in children. *N Engl J Med*, 331, 1500-7.

Pollack, I. F., Hamilton, R. L., James, C. D., Finkelstein, S. D., Burnham, J., Yates, A. J., Holmes, E. J., Zhou, T. & Finlay, J. L. (2006) Rarity of PTEN deletions and EGFR amplification in malignant gliomas of childhood: results from the Children's Cancer Group 945 cohort. *J Neurosurg*, 105, 418-24.

Pollack, J. R., Perou, C. M., Alizadeh, A. A., Eisen, M. B., Pergamenschikov, A., Williams, C. F., Jeffrey, S. S., Botstein, D. & Brown, P. O. (1999) Genome-wide

analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet*, 23, 41-6.

Pollock, P. M., Harper, U. L., Hansen, K. S., Yudt, L. M., Stark, M., Robbins, C. M., Moses, T. Y., Hostetter, G., Wagner, U., Kakareka, J., Salem, G., Pohida, T., Heenan, P., Duray, P., Kallioniemi, O., Hayward, N. K., Trent, J. M. & Meltzer, P. S. (2003) High frequency of BRAF mutations in nevi. *Nat Genet*, 33, 19-20.

Pounds, S., Cheng, C., Mullighan, C., Raimondi, S. C., Shurtleff, S. & Downing, J. R. (2009) Reference alignment of SNP microarray signals for copy number analysis of tumors. *Bioinformatics*, 25, 315-21.

Pui, C. H., Cheng, C., Leung, W., Rai, S. N., Rivera, G. K., Sandlund, J. T., Ribeiro, R. C., Relling, M. V., Kun, L. E., Evans, W. E. & Hudson, M. M. (2003) Extended follow-up of long-term survivors of childhood acute lymphoblastic leukemia. *N Engl J Med*, 349, 640-9.

Qaddoumi, I., Sultan, I. & Broniscer, A. (2009) Pediatric low-grade gliomas and the need for new options for therapy: Why and how? *Cancer Biol Ther*, 8, 4-10.

Raffel, C. (1996) Molecular biology of pediatric gliomas. *J Neurooncol*, 28, 121-8.

Raffel, C., Frederick, L., O'fallon, J. R., Atherton-Skaff, P., Perry, A., Jenkins, R. B. & James, C. D. (1999) Analysis of oncogene and tumor suppressor gene alterations in pediatric malignant astrocytomas reveals reduced survival for patients with PTEN mutations. *Clin Cancer Res*, 5, 4085-90.

Rao, S. K., Edwards, J., Joshi, A. D., Siu, I. M. & Riggins, G. J. (2009) A survey of glioblastoma genomic amplifications and deletions. *J Neurooncol*.

Rasheed, B. K., Stenzel, T. T., Mclendon, R. E., Parsons, R., Friedman, A. H., Friedman, H. S., Bigner, D. D. & Bigner, S. H. (1997) PTEN gene mutations are seen in high-grade but not in low-grade gliomas. *Cancer Res*, 57, 4187-90.

Reddy, E. P., Reynolds, R. K., Santos, E. & Barbacid, M. (1982) A point mutation is responsible for the acquisition of transforming properties by the T24 human bladder carcinoma oncogene. *Nature*, 300, 149-52.

Redon, R., Ishikawa, S., Fitch, K. R., Feuk, L., Perry, G. H., Andrews, T. D., Fiegler, H., Shapero, M. H., Carson, A. R., Chen, W., Cho, E. K., Dallaire, S., Freeman, J. L., Gonzalez, J. R., Gratacos, M., Huang, J., Kalaitzopoulos, D., Komura, D., Macdonald, J. R., Marshall, C. R., Mei, R., Montgomery, L., Nishimura, K., Okamura, K., Shen, F., Somerville, M. J., Tchinda, J., Valsesia, A., Woodwark, C., Yang, F., Zhang, J., Zerjal, T., Zhang, J., Armengol, L., Conrad, D. F., Estivill, X., Tyler-Smith, C.,

Carter, N. P., Aburatani, H., Lee, C., Jones, K. W., Scherer, S. W. & Hurles, M. E. (2006) Global variation in copy number in the human genome. *Nature*, 444, 444-54.

Rickert, C. H., Strater, R., Kaatsch, P., Wassmann, H., Jurgens, H., Dockhorn-Dworniczak, B. & Paulus, W. (2001) Pediatric high-grade astrocytomas show chromosomal imbalances distinct from adult cases. *Am J Pathol*, 158, 1525-32.

Riggi, N., Cironi, L., Provero, P., Suva, M. L., Kaloulis, K., Garcia-Echeverria, C., Hoffmann, F., Trumpp, A. & Stamenkovic, I. (2005) Development of Ewing's sarcoma from primary bone marrow-derived mesenchymal progenitor cells. *Cancer Res*, 65, 11459-68.

Riggi, N., Cironi, L., Provero, P., Suva, M. L., Stehle, J. C., Baumer, K., Guillou, L. & Stamenkovic, I. (2006) Expression of the FUS-CHOP fusion protein in primary mesenchymal progenitor cells gives rise to a model of myxoid liposarcoma. *Cancer Res*, 66, 7016-23.

Riley, R. D., Burchill, S. A., Abrams, K. R., Heney, D., Sutton, A. J., Jones, D. R., Lambert, P. C., Young, B., Wailoo, A. J. & Lewis, I. J. (2003) A systematic review of molecular and biological markers in tumours of the Ewing's sarcoma family. *Eur J Cancer*, 39, 19-30.

Risinger, J. I., Umar, A., Barrett, J. C. & Kunkel, T. A. (1995) A hPMS2 mutant cell line is defective in strand-specific mismatch repair. *J Biol Chem*, 270, 18183-6.

Roberts, P. J. & Der, C. J. (2007) Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. *Oncogene*, 26, 3291-310.

Robertson, K. D. & Jones, P. A. (1998) The human ARF cell cycle regulatory gene promoter is a CpG island which can be silenced by DNA methylation and down-regulated by wild-type p53. *Mol Cell Biol*, 18, 6457-73.

Rochat-Steiner, V., Becker, K., Micheau, O., Schneider, P., Burns, K. & Tschopp, J. (2000) FIST/HIPK3: a Fas/FADD-interacting serine/threonine kinase that induces FADD phosphorylation and inhibits fas-mediated Jun NH(2)-terminal kinase activation. *J Exp Med*, 192, 1165-74.

Roche, J., Boldog, F., Robinson, M., Robinson, L., Varella-Garcia, M., Swanton, M., Waggoner, B., Fishel, R., Franklin, W., Gemmill, R. & Drabkin, H. (1996) Distinct 3p21.3 deletions in lung cancer and identification of a new human semaphorin. *Oncogene*, 12, 1289-97.

- Rodvall, Y., Pershagen, G., Hrubec, Z., Ahlbom, A., Pedersen, N. L. & Boice, J. D. (1990) Prenatal X-ray exposure and childhood cancer in Swedish twins. *Int J Cancer*, 46, 362-5.
- Roerig, P., Nessling, M., Radlwimmer, B., Joos, S., Wrobel, G., Schwaenen, C., Reifenberger, G. & Lichter, P. (2005) Molecular classification of human gliomas using matrix-based comparative genomic hybridization. *Int J Cancer*, 117, 95-103.
- Rollins, R. A., Haghghi, F., Edwards, J. R., Das, R., Zhang, M. Q., Ju, J. & Bestor, T. H. (2006) Large-scale structure of genomic methylation patterns. *Genome Res*, 16, 157-63.
- Rosenblum, B. B., Lee, L. G., Spurgeon, S. L., Khan, S. H., Menchen, S. M., Heiner, C. R. & Chen, S. M. (1997) New dye-labeled terminators for improved DNA sequencing patterns. *Nucleic Acids Res*, 25, 4500-4.
- Rous, P. (1911) Transmission of a malignant new growth by means of a cell-free filtrate. *Jama*, 56, 198.
- Rowley, J. D. (1973) Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature*, 243, 290-3.
- Ruan, Y., Ooi, H. S., Choo, S. W., Chiu, K. P., Zhao, X. D., Srinivasan, K. G., Yao, F., Choo, C. Y., Liu, J., Ariyaratne, P., Bin, W. G., Kuznetsov, V. A., Shahab, A., Sung, W. K., Bourque, G., Palanisamy, N. & Wei, C. L. (2007) Fusion transcripts and transcribed retrotransposed loci discovered through comprehensive transcriptome analysis using Paired-End diTags (PETs). *Genome Res*, 17, 828-38.
- Ruano, Y., Mollejo, M., Ribalta, T., Fiano, C., Camacho, F. I., Gomez, E., De Lope, A. R., Hernandez-Moneo, J. L., Martinez, P. & Melendez, B. (2006) Identification of novel candidate target genes in amplicons of Glioblastoma multiforme tumors detected by expression and CGH microarray profiling. *Mol Cancer*, 5, 39.
- Rubin, J. B. & Gutmann, D. H. (2005) Neurofibromatosis type 1 - a model for nervous system tumour formation? *Nat Rev Cancer*, 5, 557-64.
- Rubin, J. B. & Rowitch, D. H. (2002) Medulloblastoma: a problem of developmental biology. *Cancer Cell*, 2, 7-8.
- Rubio-Moscardo, F., Climent, J., Siebert, R., Piris, M. A., Martin-Subero, J. I., Nielander, I., Garcia-Conde, J., Dyer, M. J., Terol, M. J., Pinkel, D. & Martinez-Climent, J. A. (2005) Mantle-cell lymphoma genotypes identified with CGH to BAC

microarrays define a leukemic subgroup of disease and predict patient outcome. *Blood*, 105, 4445-54.

Russo, A. E., Torrisi, E., Bevelacqua, Y., Perrotta, R., Libra, M., Mccubrey, J. A., Spandidos, D. A., Stivala, F. & Malaponte, G. (2009) Melanoma: molecular pathogenesis and emerging target therapies (Review). *Int J Oncol*, 34, 1481-9.

Saito, Y., Liang, G., Egger, G., Friedman, J. M., Chuang, J. C., Coetzee, G. A. & Jones, P. A. (2006) Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. *Cancer Cell*, 9, 435-43.

Sakai, T., Toguchida, J., Ohtani, N., Yandell, D. W., Rapaport, J. M. & Dryja, T. P. (1991) Allele-specific hypermethylation of the retinoblastoma tumor-suppressor gene. *Am J Hum Genet*, 48, 880-8.

Sanai, N., Alvarez-Buylla, A. & Berger, M. S. (2005) Neural stem cells and the origin of gliomas. *N Engl J Med*, 353, 811-22.

Sanger, F. & Coulson, A. R. (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol*, 94, 441-8.

Sanoudou, D., Tingby, O., Ferguson-Smith, M. A., Collins, V. P. & Coleman, N. (2000) Analysis of pilocytic astrocytoma by comparative genomic hybridization. *Br J Cancer*, 82, 1218-22.

Sansal, I. & Sellers, W. R. (2004) The biology and clinical relevance of the PTEN tumor suppressor pathway. *J Clin Oncol*, 22, 2954-63.

Santarosa, M. & Ashworth, A. (2004) Haploinsufficiency for tumour suppressor genes: when you don't need to go all the way. *Biochim Biophys Acta*, 1654, 105-22.

Sawyers, C. (2004) Targeted cancer therapy. *Nature*, 432, 294-7.

Scherr, M., Battmer, K., Winkler, T., Heidenreich, O., Ganser, A. & Eder, M. (2003) Specific inhibition of bcr-abl gene expression by small interfering RNA. *Blood*, 101, 1566-9.

Schrock, E., Blume, C., Meffert, M. C., Du Manoir, S., Bersch, W., Kiessling, M., Lozanowa, T., Thiel, G., Witkowski, R., Ried, T. & Cremer, T. (1996) Recurrent gain of chromosome arm 7q in low-grade astrocytic tumors studied by comparative genomic hybridization. *Genes Chromosomes Cancer*, 15, 199-205.

Schubbert, S., Shannon, K. & Bollag, G. (2007) Hyperactive Ras in developmental disorders and cancer. *Nat Rev Cancer*, 7, 295-308.

Schwab, M., Alitalo, K., Klempnauer, K. H., Varmus, H. E., Bishop, J. M., Gilbert, F., Brodeur, G., Goldstein, M. & Trent, J. (1983) Amplified DNA with limited homology to myc cellular oncogene is shared by human neuroblastoma cell lines and a neuroblastoma tumour. *Nature*, 305, 245-8.

Scotting, P. J., Walker, D. A. & Perilongo, G. (2005) Childhood solid tumours: a developmental disorder. *Nat Rev Cancer*, 5, 481-8.

Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P., Maner, S., Massa, H., Walker, M., Chi, M., Navin, N., Lucito, R., Healy, J., Hicks, J., Ye, K., Reiner, A., Gilliam, T. C., Trask, B., Patterson, N., Zetterberg, A. & Wigler, M. (2004) Large-scale copy number polymorphism in the human genome. *Science*, 305, 525-8.

Seitz, H. K. & Stickel, F. (2007) Molecular mechanisms of alcohol-mediated carcinogenesis. *Nat Rev Cancer*, 7, 599-612.

Selzer, R. R., Richmond, T. A., Pofahl, N. J., Green, R. D., Eis, P. S., Nair, P., Brothman, A. R. & Stallings, R. L. (2005) Analysis of chromosome breakpoints in neuroblastoma at sub-kilobase resolution using fine-tiling oligonucleotide array CGH. *Genes Chromosomes Cancer*, 44, 305-19.

Sethupathy, P., Megraw, M. & Hatzigeorgiou, A. G. (2006) A guide through present computational approaches for the identification of mammalian microRNA targets. *Nat Methods*, 3, 881-6.

Sharif, S., Ferner, R., Birch, J. M., Gillespie, J. E., Gattamaneni, H. R., Baser, M. E. & Evans, D. G. (2006) Second primary tumors in neurofibromatosis 1 patients treated for optic glioma: substantial risks after radiotherapy. *J Clin Oncol*, 24, 2570-5.

Sharma, M. K., Zehnauer, B. A., Watson, M. A. & Gutmann, D. H. (2005) RAS pathway activation and an oncogenic RAS mutation in sporadic pilocytic astrocytoma. *Neurology*, 65, 1335-6.

Sharrocks, A. D. (2001) The ETS-domain transcription factor family. *Nat Rev Mol Cell Biol*, 2, 827-37.

Shaulian, E. & Karin, M. (2001) AP-1 in cell proliferation and survival. *Oncogene*, 20, 2390-400.

Shaulian, E. & Karin, M. (2002) AP-1 as a regulator of cell life and death. *Nat Cell Biol*, 4, E131-6.

Sheer, D. & Shipley, J. (2005) *Chapter 6. Molecular cytogenetics of cancer.*, Oxford, Oxford University Press.

Shete, S., Hosking, F. J., Robertson, L. B., Dobbins, S. E., Sanson, M., Malmer, B., Simon, M., Marie, Y., Boisselier, B., Delattre, J. Y., Hoang-Xuan, K., El Hallani, S., Idbaih, A., Zelenika, D., Andersson, U., Henriksson, R., Bergenheim, A. T., Feychting, M., Lonn, S., Ahlbom, A., Schramm, J., Linnebank, M., Hemminki, K., Kumar, R., Hepworth, S. J., Price, A., Armstrong, G., Liu, Y., Gu, X., Yu, R., Lau, C., Schoemaker, M., Muir, K., Swerdlow, A., Lathrop, M., Bondy, M. & Houlston, R. S. (2009) Genome-wide association study identifies five susceptibility loci for glioma. *Nat Genet*, 41, 899-904.

Shih, C., Padhy, L. C., Murray, M. & Weinberg, R. A. (1981) Transforming genes of carcinomas and neuroblastomas introduced into mouse fibroblasts. *Nature*, 290, 261-4.

Sievert, A. J., Jackson, E. M., Gai, X., Hakonarson, H., Judkins, A. R., Resnick, A. C., Sutton, L. N., Storm, P. B., Shaikh, T. H. & Biegel, J. A. (2009) Duplication of 7q34 in pediatric low-grade astrocytomas detected by high-density single-nucleotide polymorphism-based genotype arrays results in a novel BRAF fusion gene. *Brain Pathol*, 19, 449-58.

Singer-Sam, J. & Riggs, A. D. (1993) X chromosome inactivation and DNA methylation. *Exs*, 64, 358-84.

Singh, S. K., Clarke, I. D., Terasaki, M., Bonn, V. E., Hawkins, C., Squire, J. & Dirks, P. B. (2003) Identification of a cancer stem cell in human brain tumors. *Cancer Res*, 63, 5821-8.

Singh, S. K., Hawkins, C., Clarke, I. D., Squire, J. A., Bayani, J., Hide, T., Henkelman, R. M., Cusimano, M. D. & Dirks, P. B. (2004) Identification of human brain tumour initiating cells. *Nature*, 432, 396-401.

Siva, N. (2008) 1000 Genomes project. *Nat Biotechnol*, 26, 256.

Smiraglia, D. J., Rush, L. J., Fruhwald, M. C., Dai, Z., Held, W. A., Costello, J. F., Lang, J. C., Eng, C., Li, B., Wright, F. A., Caligiuri, M. A. & Plass, C. (2001) Excessive CpG island hypermethylation in cancer cell lines versus primary human malignancies. *Hum Mol Genet*, 10, 1413-9.

Snijders, A. M., Nowak, N., Segreaves, R., Blackwood, S., Brown, N., Conroy, J., Hamilton, G., Hindle, A. K., Huey, B., Kimura, K., Law, S., Myambo, K., Palmer, J., Ylstra, B., Yue, J. P., Gray, J. W., Jain, A. N., Pinkel, D. & Albertson, D. G. (2001) Assembly of microarrays for genome-wide measurement of DNA copy number. *Nat Genet*, 29, 263-4.

Soderling, S. H., Guire, E. S., Kaech, S., White, J., Zhang, F., Schutz, K., Langeberg, L. K., Banker, G., Raber, J. & Scott, J. D. (2007) A WAVE-1 and WRP signaling complex regulates spine density, synaptic plasticity, and memory. *J Neurosci*, 27, 355-65.

Soutoglou, E. & Misteli, T. (2008) On the contribution of spatial genome organization to cancerous chromosome translocations. *J Natl Cancer Inst Monogr*, 16-9.

Specht, K., Richter, T., Muller, U., Walch, A., Werner, M. & Hofler, H. (2001) Quantitative gene expression analysis in microdissected archival formalin-fixed and paraffin-embedded tumor tissue. *Am J Pathol*, 158, 419-29.

Steck, P. A., Pershouse, M. A., Jasser, S. A., Yung, W. K., Lin, H., Ligon, A. H., Langford, L. A., Baumgard, M. L., Hattier, T., Davis, T., Frye, C., Hu, R., Swedlund, B., Teng, D. H. & Tavtigian, S. V. (1997) Identification of a candidate tumour suppressor gene, MMAC1, at chromosome 10q23.3 that is mutated in multiple advanced cancers. *Nat Genet*, 15, 356-62.

Stehelin, D., Varmus, H. E., Bishop, J. M. & Vogt, P. K. (1976) DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. *Nature*, 260, 170-3.

Storchova, Z. & Pellman, D. (2004) From polyploidy to aneuploidy, genome instability and cancer. *Nat Rev Mol Cell Biol*, 5, 45-54.

Stott, F. J., Bates, S., James, M. C., McConnell, B. B., Starborg, M., Brookes, S., Palmero, I., Ryan, K., Hara, E., Vousden, K. H. & Peters, G. (1998) The alternative product from the human CDKN2A locus, p14(ARF), participates in a regulatory feedback loop with p53 and MDM2. *Embo J*, 17, 5001-14.

Stratton, M. R., Campbell, P. J. & Futreal, P. A. (2009) The cancer genome. *Nature*, 458, 719-24.

Strout, M. P., Marcucci, G., Bloomfield, C. D. & Caligiuri, M. A. (1998) The partial tandem duplication of ALL1 (MLL) is consistently generated by Alu-mediated homologous recombination in acute myeloid leukemia. *Proc Natl Acad Sci U S A*, 95, 2390-5.

Sutter, R., Yadirgi, G. & Marino, S. (2007) Neural stem cells, tumour stem cells and brain tumours: dangerous relationships? *Biochim Biophys Acta*, 1776, 125-37.

Tabin, C. J., Bradley, S. M., Bargmann, C. I., Weinberg, R. A., Papageorge, A. G., Scolnick, E. M., Dhar, R., Lowy, D. R. & Chang, E. H. (1982) Mechanism of activation of a human oncogene. *Nature*, 300, 143-9.

- Tachibana, K. E., Gonzalez, M. A. & Coleman, N. (2005) Cell-cycle-dependent regulation of DNA replication and its relevance to cancer pathology. *J Pathol*, 205, 123-9.
- Taylor, M. D., Liu, L., Raffel, C., Hui, C. C., Mainprize, T. G., Zhang, X., Agatep, R., Chiappa, S., Gao, L., Lowrance, A., Hao, A., Goldstein, A. M., Stavrou, T., Scherer, S. W., Dura, W. T., Wainwright, B., Squire, J. A., Rutka, J. T. & Hogg, D. (2002) Mutations in SUFU predispose to medulloblastoma. *Nat Genet*, 31, 306-10.
- Taylor, M. D., Poppleton, H., Fuller, C., Su, X., Liu, Y., Jensen, P., Magdaleno, S., Dalton, J., Calabrese, C., Board, J., Macdonald, T., Rutka, J., Guha, A., Gajjar, A., Curran, T. & Gilbertson, R. J. (2005) Radial glia cells are candidate stem cells of ependymoma. *Cancer Cell*, 8, 323-35.
- Thomas, M., Gessner, A., Vornlocher, H. P., Hadwiger, P., Greil, J. & Heidenreich, O. (2005) Targeting MLL-AF4 with short interfering RNAs inhibits clonogenicity and engraftment of t(4;11)-positive human leukemic cells. *Blood*, 106, 3559-66.
- Tomlins, S. A., Laxman, B., Dhanasekaran, S. M., Helgeson, B. E., Cao, X., Morris, D. S., Menon, A., Jing, X., Cao, Q., Han, B., Yu, J., Wang, L., Montie, J. E., Rubin, M. A., Pienta, K. J., Roulston, D., Shah, R. B., Varambally, S., Mehra, R. & Chinnaiyan, A. M. (2007) Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. *Nature*, 448, 595-9.
- Toniato, E., Chen, X. P., Losman, J., Flati, V., Donahue, L. & Rothman, P. (2002) TRIM8/GERP RING finger protein interacts with SOCS-1. *J Biol Chem*, 277, 37315-22.
- Tsujimoto, Y., Finger, L. R., Yunis, J., Nowell, P. C. & Croce, C. M. (1984) Cloning of the chromosome breakpoint of neoplastic B cells with the t(14;18) chromosome translocation. *Science*, 226, 1097-9.
- Tsujimoto, Y., Gorham, J., Cossman, J., Jaffe, E. & Croce, C. M. (1985) The t(14;18) chromosome translocations involved in B-cell neoplasms result from mistakes in VDJ joining. *Science*, 229, 1390-3.
- Valinluck, V., Tsai, H. H., Rogstad, D. K., Burdzy, A., Bird, A. & Sowers, L. C. (2004) Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). *Nucleic Acids Res*, 32, 4100-8.
- Van Slegtenhorst, M., De Hoogt, R., Hermans, C., Nellist, M., Janssen, B., Verhoef, S., Lindhout, D., Van Den Ouweland, A., Halley, D., Young, J., Burley, M., Jeremiah,

S., Woodward, K., Nahmias, J., Fox, M., Ekong, R., Osborne, J., Wolfe, J., Povey, S., Snell, R. G., Cheadle, J. P., Jones, A. C., Tachataki, M., Ravine, D., Sampson, J. R., Reeve, M. P., Richardson, P., Wilmer, F., Munro, C., Hawkins, T. L., Sepp, T., Ali, J. B., Ward, S., Green, A. J., Yates, J. R., Kwiatkowska, J., Henske, E. P., Short, M. P., Haines, J. H., Jozwiak, S. & Kwiatkowski, D. J. (1997) Identification of the tuberous sclerosis gene TSC1 on chromosome 9q34. *Science*, 277, 805-8.

Van Tornout, J. M., Buckley, J. D., Quinn, J. J., Feusner, J. H., Krailo, M. D., King, D. R., Hammond, G. D. & Ortega, J. A. (1997) Timing and magnitude of decline in alpha-fetoprotein levels in treated children with unresectable or metastatic hepatoblastoma are predictors of outcome: a report from the Children's Cancer Group. *J Clin Oncol*, 15, 1190-7.

Veltman, J. A., Schoenmakers, E. F., Eussen, B. H., Janssen, I., Merckx, G., Van Cleef, B., Van Ravenswaaij, C. M., Brunner, H. G., Smeets, D. & Van Kessel, A. G. (2002) High-throughput analysis of subtelomeric chromosome rearrangements by use of array-based comparative genomic hybridization. *Am J Hum Genet*, 70, 1269-76.

Vetter, I. R. & Wittinghofer, A. (2001) The guanine nucleotide-binding switch in three dimensions. *Science*, 294, 1299-304.

Visone, R. & Croce, C. M. (2009) MiRNAs and cancer. *Am J Pathol*, 174, 1131-8.

Vivanco, I. & Sawyers, C. L. (2002) The phosphatidylinositol 3-Kinase AKT pathway in human cancer. *Nat Rev Cancer*, 2, 489-501.

Vogelstein, B. & Kinzler, K. W. (2004) Cancer genes and the pathways they control. *Nat Med*, 10, 789-99.

Volinia, S., Calin, G. A., Liu, C. G., Ambs, S., Cimmino, A., Petrocca, F., Visone, R., Iorio, M., Roldo, C., Ferracin, M., Prueitt, R. L., Yanaihara, N., Lanza, G., Scarpa, A., Vecchione, A., Negrini, M., Harris, C. C. & Croce, C. M. (2006) A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad Sci U S A*, 103, 2257-61.

Von Deimling, A., Krone, W. & Menon, A. G. (1995) Neurofibromatosis type 1: pathology, clinical features and molecular genetics. *Brain Pathol*, 5, 153-62.

Von Deimling, A., Louis, D. N., Menon, A. G., Von Ammon, K., Petersen, I., Ellison, D., Wiestler, O. D. & Seizinger, B. R. (1993) Deletions on the long arm of chromosome 17 in pilocytic astrocytoma. *Acta Neuropathol*, 86, 81-5.

- Vos, M. D., Ellis, C. A., Bell, A., Birrer, M. J. & Clark, G. J. (2000) Ras uses the novel tumor suppressor RASSF1 as an effector to mediate apoptosis. *J Biol Chem*, 275, 35669-72.
- Wall, J. D. & Pritchard, J. K. (2003) Haplotype blocks and linkage disequilibrium in the human genome. *Nat Rev Genet*, 4, 587-97.
- Walther, A., Houlston, R. & Tomlinson, I. (2008) Association between chromosomal instability and prognosis in colorectal cancer: a meta-analysis. *Gut*, 57, 941-50.
- Wang, D. G., Fan, J. B., Siao, C. J., Berno, A., Young, P., Sapolsky, R., Ghandour, G., Perkins, N., Winchester, E., Spencer, J., Kruglyak, L., Stein, L., Hsie, L., Topaloglou, T., Hubbell, E., Robinson, E., Mittmann, M., Morris, M. S., Shen, N., Kilburn, D., Rioux, J., Nusbaum, C., Rozen, S., Hudson, T. J., Lipshutz, R., Chee, M. & Lander, E. S. (1998) Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science*, 280, 1077-82.
- Wang, S. I., Puc, J., Li, J., Bruce, J. N., Cairns, P., Sidransky, D. & Parsons, R. (1997) Somatic mutations of PTEN in glioblastoma multiforme. *Cancer Res*, 57, 4183-6.
- Warr, T., Ward, S., Burrows, J., Harding, B., Wilkins, P., Harkness, W., Hayward, R., Darling, J. & Thomas, D. (2001) Identification of extensive genomic loss and gain by comparative genomic hybridisation in malignant astrocytoma in children and young adults. *Genes Chromosomes Cancer*, 31, 15-22.
- Watanabe, T., Nobusawa, S., Kleihues, P. & Ohgaki, H. (2009) IDH1 mutations are early events in the development of astrocytomas and oligodendrogliomas. *Am J Pathol*, 174, 1149-53.
- Watson, J. D. & Crick, F. H. (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, 171, 737-8.
- Watts, G. S., Futscher, B. W., Holtan, N., Degeest, K., Domann, F. E. & Rose, S. L. (2008) DNA methylation changes in ovarian cancer are cumulative with disease progression and identify tumor stage. *BMC Med Genomics*, 1, 47.
- Weber, J. L. & May, P. E. (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet*, 44, 388-96.
- Weinstein, I. B. & Case, K. (2008) The history of Cancer Research: introducing an AACR Centennial series. *Cancer Res*, 68, 6861-2.

- Weitzman, S. A., Turk, P. W., Milkowski, D. H. & Kozlowski, K. (1994) Free radical adducts induce alterations in DNA cytosine methylation. *Proc Natl Acad Sci U S A*, 91, 1261-4.
- Welborn, J. (2004) Constitutional chromosome aberrations as pathogenetic events in hematologic malignancies. *Cancer Genet Cytogenet*, 149, 137-53.
- Wellbrock, C., Karasarides, M. & Marais, R. (2004) The RAF proteins take centre stage. *Nat Rev Mol Cell Biol*, 5, 875-85.
- White, F. V., Anthony, D. C., Yunis, E. J., Tarbell, N. J., Scott, R. M. & Schofield, D. E. (1995) Nonrandom chromosomal gains in pilocytic astrocytomas of childhood. *Hum Pathol*, 26, 979-86.
- Widschwendter, M., Jiang, G., Woods, C., Muller, H. M., Fiegl, H., Goebel, G., Marth, C., Muller-Holzner, E., Zeimet, A. G., Laird, P. W. & Ehrlich, M. (2004) DNA hypomethylation and ovarian cancer biology. *Cancer Res*, 64, 4472-80.
- Wightman, B., Ha, I. & Ruvkun, G. (1993) Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell*, 75, 855-62.
- Wilhelm, E., Pellay, F. X., Benecke, A. & Bell, B. (2008) TAF6delta controls apoptosis and gene expression in the absence of p53. *PLoS One*, 3, e2721.
- Willert, J. R., Daneshvar, L., Sheffield, V. C. & Cogen, P. H. (1995) Deletion of chromosome arm 17p DNA sequences in pediatric high-grade and juvenile pilocytic astrocytomas. *Genes Chromosomes Cancer*, 12, 165-72.
- Wimmer, K., Eckart, M., Meyer-Puttlitz, B., Fonatsch, C. & Pietsch, T. (2002) Mutational and expression analysis of the NF1 gene argues against a role as tumor suppressor in sporadic pilocytic astrocytomas. *J Neuropathol Exp Neurol*, 61, 896-902.
- Wong, K., Ren, X. R., Huang, Y. Z., Xie, Y., Liu, G., Saito, H., Tang, H., Wen, L., Brady-Kalnay, S. M., Mei, L., Wu, J. Y., Xiong, W. C. & Rao, Y. (2001) Signal transduction in neuronal migration: roles of GTPase activating proteins and the small GTPase Cdc42 in the Slit-Robo pathway. *Cell*, 107, 209-21.
- Wong, K. K., Tsang, Y. T., Chang, Y. M., Su, J., Di Francesco, A. M., Meco, D., Riccardi, R., Perlaky, L., Dauser, R. C., Adesina, A., Bhattacharjee, M., Chintagumpala, M. & Lau, C. C. (2006) Genome-wide allelic imbalance analysis of pediatric gliomas by single nucleotide polymorphic allele array. *Cancer Res*, 66, 11172-8.

- Wu, H., Goel, V. & Haluska, F. G. (2003) PTEN signaling pathways in melanoma. *Oncogene*, 22, 3113-22.
- Yan, H., Parsons, D. W., Jin, G., Mclendon, R., Rasheed, B. A., Yuan, W., Kos, I., Batinic-Haberle, I., Jones, S., Riggins, G. J., Friedman, H., Friedman, A., Reardon, D., Herndon, J., Kinzler, K. W., Velculescu, V. E., Vogelstein, B. & Bigner, D. D. (2009) IDH1 and IDH2 mutations in gliomas. *N Engl J Med*, 360, 765-73.
- Yang, Y., Marcello, M., Endris, V., Saffrich, R., Fischer, R., Trendelenburg, M. F., Sprengel, R. & Rappold, G. (2006) MEGAP impedes cell migration via regulating actin and microtubule dynamics and focal complex formation. *Exp Cell Res*, 312, 2379-93.
- Yi, J. M., Tsai, H. C., Glockner, S. C., Lin, S., Ohm, J. E., Easwaran, H., James, C. D., Costello, J. F., Riggins, G., Eberhart, C. G., Laterra, J., Vescovi, A. L., Ahuja, N., Herman, J. G., Schuebel, K. E. & Baylin, S. B. (2008) Abnormal DNA methylation of CD133 in colorectal and glioblastoma tumors. *Cancer Res*, 68, 8094-103.
- Ylstra, B., Van Den Ijssel, P., Carvalho, B., Brakenhoff, R. H. & Meijer, G. A. (2006) BAC to the future! or oligonucleotides: a perspective for micro array comparative genomic hybridization (array CGH). *Nucleic Acids Res*, 34, 445-50.
- Yu, J., Zhang, H., Gu, J., Lin, S., Li, J., Lu, W., Wang, Y. & Zhu, J. (2004) Methylation profiles of thirty four promoter-CpG islands and concordant methylation behaviours of sixteen genes that may contribute to carcinogenesis of astrocytoma. *BMC Cancer*, 4, 65.
- Zattara-Cannoni, H., Gambarelli, D., Lena, G., Dufour, H., Choux, M., Grisoli, F. & Vagner-Capodano, A. M. (1998) Are juvenile pilocytic astrocytomas benign tumors? A cytogenetic study in 24 cases. *Cancer Genet Cytogenet*, 104, 157-60.
- Zhao, J., Yuan, X., Frodin, M. & Grummt, I. (2003) ERK-dependent phosphorylation of the transcription initiation factor TIF-IA is required for RNA polymerase I transcription and cell growth. *Mol Cell*, 11, 405-13.
- Zhao, X., Li, C., Paez, J. G., Chin, K., Janne, P. A., Chen, T. H., Girard, L., Minna, J., Christiani, D., Leo, C., Gray, J. W., Sellers, W. R. & Meyerson, M. (2004) An integrated view of copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays. *Cancer Res*, 64, 3060-71.
- Zhu, Y., Guignard, F., Zhao, D., Liu, L., Burns, D. K., Mason, R. P., Messing, A. & Parada, L. F. (2005) Early inactivation of p53 tumor suppressor gene cooperating with NF1 loss induces malignant astrocytoma. *Cancer Cell*, 8, 119-30.

