UNIVERSITY OF SOUTHAMPTON

FACULTY OF MEDICINE HEALTH AND LIFE SCIENCES

School of Medicine

A Microarray Approach to Detect Chromosome 18 Deletions in Colorectal Cancer

by

Jonathan P Trickett

MB BS, FRCS (Gen)

Thesis for the degree of DM

September 2007

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF MEDICINE HEALTH AND LIFE SCIENCES

SCHOOL OF MEDICINE

DM

A MICROARRAY APPROACH TO DETECT CHROMOSOME 18 DELETIONS IN COLORECTAL CANCER

by Jonathan P Trickett

Conventional techniques have identified Chromosome 18 as having the highest frequency of deletion in colorectal cancer (CRC); it is the likely site of a causative tumour suppressor gene inactivated by deletion in CRC carcinogenesis. These techniques have not however accurately identified the deleted region that is likely to contain the causative tumour suppressor gene.

A high-resolution technique, microarray comparative genomic hybridisation (CGH) was utilised, directly comparing normal DNA against tumour DNA for differences. The microarray was designed and constructed to target chromosome 18 at the highest resolution with a complete "tiling path" of 860 overlapping BAC clones to represent this region.

The accuracy of the constructed micro array and comparative genomic hybridisation technique was confirmed by analysis of normal against normal DNA and a cell line with a known large deletion and amplification.

Micro array CGH was performed on 47 cell lines and 69 primary cancers, identifying deletions in 81% of the cell lines and 55% of the primary cancers. Three minimal regions of deletion were identified, spanning only 371kb to 3.5 Mb. Two of the minimal regions of deletion were common to 92% of all deletions in the cell lines and 79% of primary cancers. The other minimal region of deletion was common to 92% of all deletions in the cell lines and 63% in the primary cancers. Fluorescent in situ hybridisation (FISH) confirmed the presence and location of these minimal deletions.

The 3 minimal regions of deletion contained only 8 genes; including 3 strong candidate tumour suppressor genes, SMAD 7, CADHERIN 7 and CADHERIN 19. SMAD 7 is part of the tumour growth factor (TGF) cascade, a potent inhibitor of cell growth and inducer of apoptosis, important in colorectal carcinogenesis. The CADHERINS are involved in intercellular adhesion, disturbance of which is a prerequisite for invasion and metastasis of tumour cells.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

ACKNOWLEDGEMENTS

This research was funded by the Royal College of Surgeons of England and HOPE charity.

I would like to thank my supervisors Professor David Wilson and Mr Nick Beck for all their support and guidance.

Professor David Wilson's knowledge, dedication and accessibility extended to me as part of the Human Genetics Division was vital in the completion of a thesis of high quality. I would like to thank all of the Human Genetics Division for their support especially Kelly Wilkinson and Mirella Spalluto.

I am also very grateful to the major collaborative group lead by Professor Ian Tomlinson for their help and allowing a proportion of my work to be undertaken at the Molecular and Population genetics Laboratory, Cancer Research UK, London.

CHAPTER 1: INTRODUCTION

1.1 TUMOURIGENESIS: THE GENETIC BASIS OF CANCER

All cells normally exist under strict regulation of growth, differentiation, interactions and apoptosis. Escape from this tight regulation can lead to tumourigenesis.

The normal function of tumour suppressor genes is to inhibit or control cell division. They ensure genomic stability through low mutation rates, accurate replication, controlled cell cycle progression and apoptosis, inhibiting progression to a malignant phenotype. In contrast oncogenes control cellular proliferation which when overactive transforms a normal cell to a tumour cell.

Alterations in the function of these key genes results in conversion to a malignant phenotype. Loss or gain of gene function is influenced by allelic copy number change, deletion of tumour suppressor genes and amplification of oncogenes, in conjunction with mutations or epigenetic influences on gene transcription [1].

A mutation that provides a cell with a survival advantage will be selected for and result in clonal expansion of the cell in which it arose. Conversion to a malignant phenotype may requires successive somatic mutations but DNA repair mechanisms can be overwhelmed by key mutations, enhancing cellular proliferation and chromosomal instability, so in turn greatly increasing mutation rates.

Knudson's "two-hit hypothesis", describes allelic loss and mutation as the "two hits" on a pair of alleles that render complete loss of tumour suppressor gene function [2]. Identification of allelic copy number loss has been used successfully to locate tumour suppressor genes.

1.2 INHERETED COLORECTAL CANCER

Knowledge of the constitutional changes responsible for inherited colorectal cancer (CRC) provided the initial understanding of CRC tumourigenesis.

1.2.1 Familial adenomatous polyposis

Familial Adenomatous Polyposis (FAP) is an autosomal dominant disease that accounts for approximately 1% of all CRCs [3]. It is characterised by hundreds of colorectal adenomatous polyps evident from the second or third decade of life, if untreated individuals with FAP will develop CRC in their 30's [4]. The phenotype also includes duodenal adenomatous polyps[5] and extra intestinal manifestations, including hypertrophy of the retinal pigmented epithelium[6] and desmoid tumours [7-9] [4, 10]. The disease is due to mutations in the adenomatous polyposis coli (*APC*) tumour suppressor gene located on chromosome 5 (see section 1.4.1). The site of mutation in the *APC* gene can influence the FAP phenotype [11, 12]. Mutations at codons beyond 1250 [12] and particularly 1309 and 1328 in exon 15G are uniformly associated with a severe polyposis phenotype [11] including a higher risk of rectal cancer. This information is potentially useful to determine, in patients with FAP, if a prophylactic colonic resection should be extended to include a rectal resection (panproctocolectomy) which has an associated higher surgical morbidity.

1.2.2 Hereditary non polyposis colorectal cancer

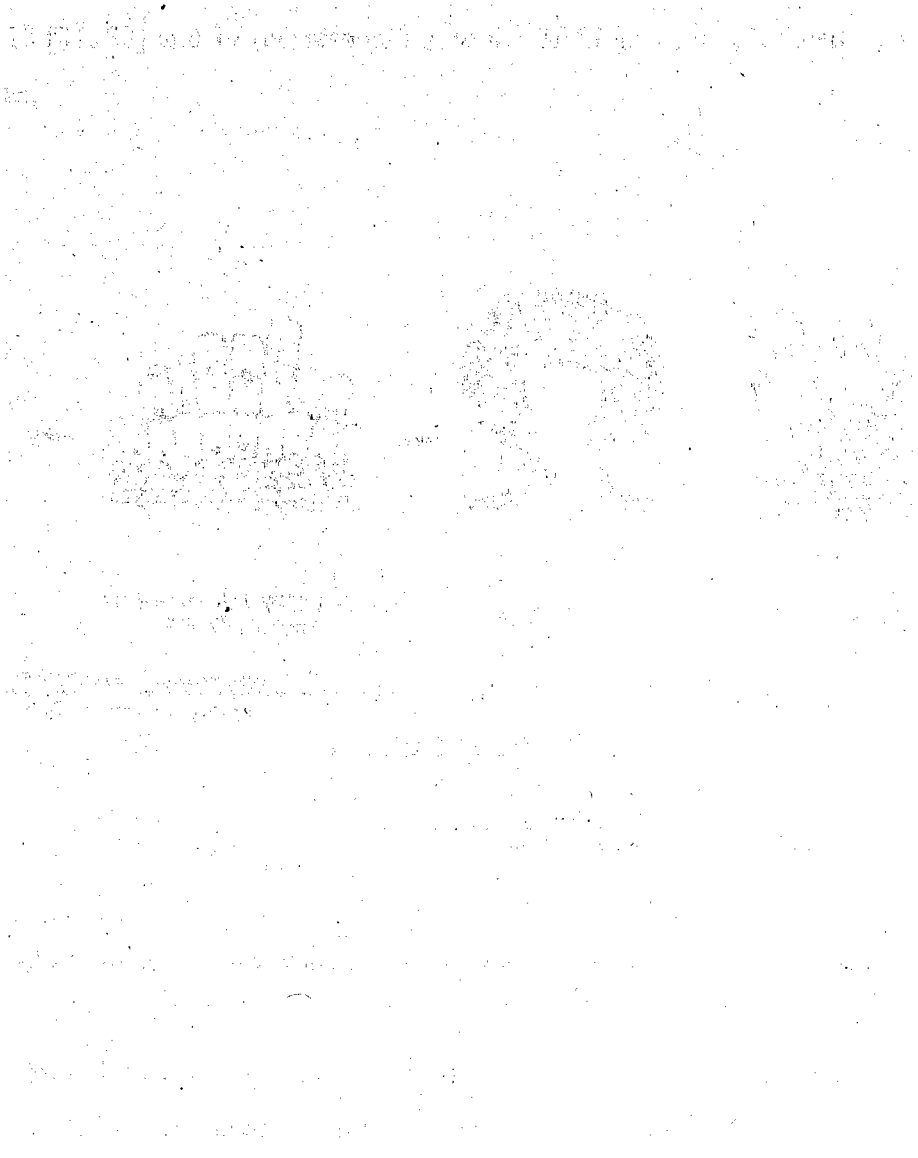
Henry Lynch described a second form of familial colorectal cancer[2,3]. Initially, this was believed to encompass two separate syndromes, colon cancer with or without associated extra-colonic malignancies. Lynch introduced the term 'hereditary non-polyposis colorectal cancer' (HNPCC) to encompass both Lynch syndromes.[4]

Lynch viewed HNPCC as a syndrome characterised by an autosomal dominant pattern of inheritance, early onset of malignancy with a predilection for the proximal colon, multiple CRCs, the absence of premonitory lesions (e.g. adenomas), and the occurrence of cancer in certain extracolonic sites, notably endometrium and ovary[13]

Confusion has arisen over the last decade in the in accurate use of the term HNPCC, the term has been applied to heterogeneous groups of families meeting limited clinical criteria that may suggest Lynch syndrome and in clinical scenarios in which CRCs with DNA microsatellite instability are diagnosed but in which there is no vertical transmission of an altered DNA mismatch repair (MMR) gene. A term that has multiple, mutually incompatible meanings is highly problematic, particularly when it may influence the management of an individual family. The Lynch syndrome is best understood as a hereditary predisposition to malignancy that is explained by a germline mutation in a DNA MMR gene. The diagnosis does not depend in an absolute sense on any particular family pedigree structure or age of onset of malignancy. Families with a strong family history of colorectal cancer that do not have Lynch syndrome can been grouped as 'Familial Colorectal Cancer Type-X'. The first step in characterizing these cancer families is to distinguish them from Lynch syndrome. The term HNPCC no longer serves any useful purpose and should be phased out [14]

Lynch syndrome is associated with germ-line mutations in one of two mismatch repair genes *MLH1* and *MLH2* (see section 1.4.2) [15, 16]. Inactivation of the mismatch repair mechanism responsible for the correction of base pair replication errors results in the accumulation of mutations including short repeat DNA sequences, known as microsatellite instability (MSI) [17]. Fifteen percent of sporadic cancers are MSI+ and have mutations or epigenetic changes [18, 19] in one of a larger group of mismatch repair genes *MLH1, MSH2, MSH6, PMS1, PMS2* (see section 1.5.1). These MSI+ sporadic

cancers retain a near diploid karyotype with a balanced "chromosomal complement". The remaining sporadic cancers demonstrate chromosomal instability (CIN+), with a higher frequency of loss or addition to their chromosomal set (aneuploidy) [20, 21].



1.3 COLORECTAL CANCER TUMOURIGENESIS

1.3.1 Genetic pathway of colorectal cancer tumourigenesis

The histological progression from healthy colonic epithelium, through increasingly dysplastic adenomas to carcinoma is well established [22]. A parallel genetic pathway driving this transition was suggested by Vogelstein [23] Bodmer[I] and others [24, 25] following evidence of allelic loss on chromosome 5, (corresponding to the *APC* gene) [26-28] and *KRAS2* mutations [29, 30] in colorectal adenomas and loss of specific regions of chromosome 18 [31, 32] and 17 (corresponding to the *TP53* gene) in advanced adenomas and carcinomas.

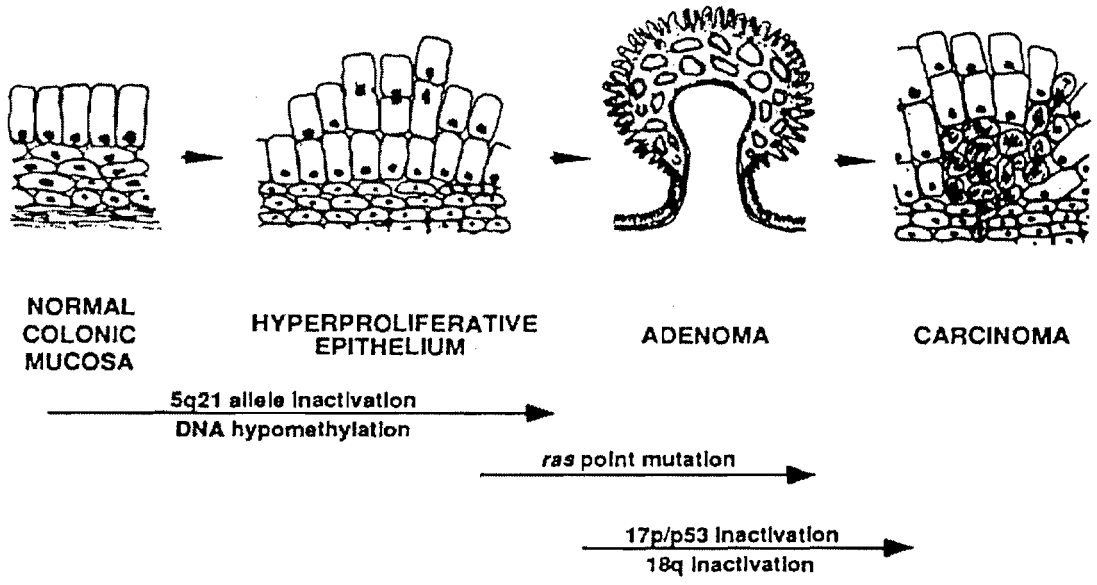
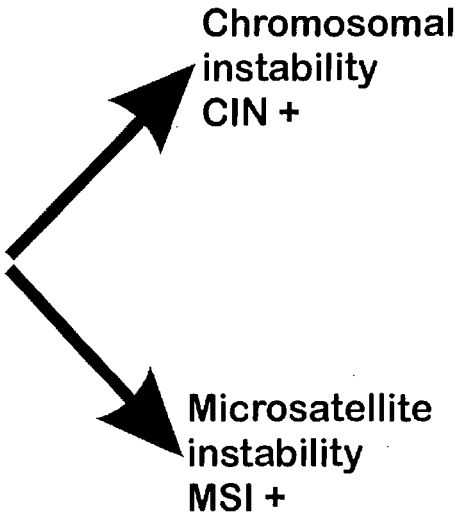


Fig. 1.1 Histological adenoma carcinoma sequence and suggested parallel genetic pathway

This widely accepted model of genetic events illustrates one possible multistep process. Less than 7% of CRCs have mutations in all three identified genes, *APC, K-RAS2* and *TP-53* [33]. These mutations and copy number changes rather than lying on one single

common pathway, probably lie on different pathways, with no one single pattern of development applicable to all CRCs.



A simple broad molecular classification for CRCs [1] has been based on the identification of 2 distinct groups. Microsatellite unstable (MSI+) sporadic CRCs with a stable near diploid karyotype and microsatellite stable (MSI-) CRCs demonstrating chromosomal instability (CIN+) with an aneuploid/ polyploidy karyotype.

Fig 1.2 Sporadic colorectal cancers classified by genomic stability, MSI /CIN divergence.

The classification as either MSI+ or CIN+ is appropriate for most CRCs, there is however evidence that some CRCs are not exclusive to either group [34, 35].

DNA ploidy patterns of sporadic CRCs have been determined by DNA flow cytometry reporting a diploid karyotype in 27% and aneuploid in 73% [36].

As about 15% of sporadic CRCs are MSI+, there are a small group of diploid CRCs that are MSI-ve, phenotypically they appear to conform to the MSI- group with an associated shorter survival [37].

1.4 TUMOUR SUPPRESSOR GENES AND ONCOGENES IN COLORECTAL CANCER

1.4.1 Adenosis Polyposis Coli gene

The Adenosis Polyposis Coli (*APC*)(OMIM 175100: APC) gene is a tumour suppressor gene located on chromosome 5q21 [38]; its product is involved in the Wnt-signalling pathway. The APC protein is cytoplasmic, acting in a multiprotein complex (with GSK3b and axin) that degrades the oncogene B-catenin by promoting the phosphorylation of its N terminus which thereby targets it for degradation by the ubiquitin proteasome pathway [39] (Fig. 1.3a).

If the multiprotein complex with the APC protein does not occur, B-catenin escapes degradation, stabilising, translocating to the nucleus and forming a complex with T cell factor [40], a member of the Tcf transcription factor family, within the nuclei of WNT responsive cells [41, 42] . The T cell factor guides the complex to promoters of target genes, B-catenin then promotes their transcription. These WNT target genes include *C-MYC*, and *CYCLIN D1* [43] (Fig. 1.3b).

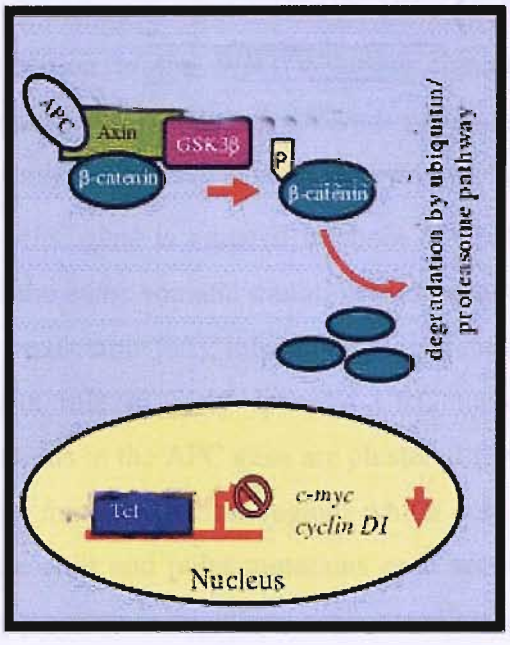


Fig. 1.3a: Controlled cell growth

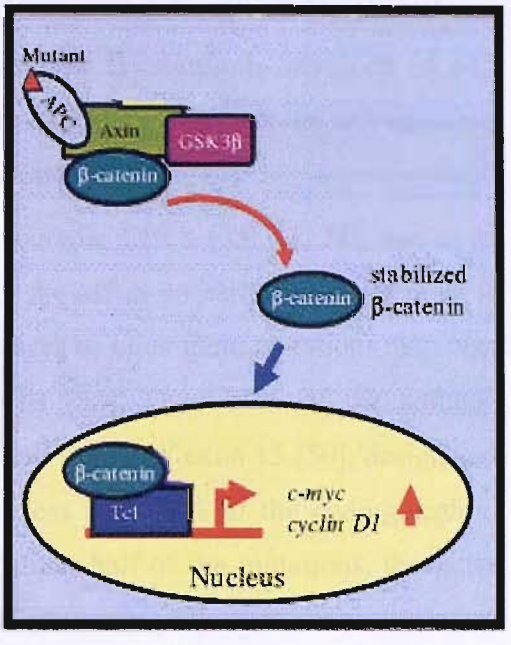


Fig. 1.3b: Uncontrolled cell growth

Fig. 1.3a/b: A model for the Wnt-signalling pathway.

Fig. 1.3a: The normal APC multiprotein complex leads to the degradation of the oncogene B-catenin by the ubiquitin/proteasome pathway in normal colonic epithelial cells.

Fig. 1.3b: If there is an APC mutation, B-catenin escapes degradation, translocates to the nucleus, promoting the transcription of target genes.

CRCs with *APC* or *B-CATENIN* mutations therefore exhibit increased *B-CATENIN*/Tcf-mediated transcription[44], over expression of the oncogene *C-MYC* [45] has been observed in both early and late CRC causing neoplastic growth [46]. *CYCLIN D1* is a major regulator of the progression of cells into the proliferative stage of the cell cycle [47], its increased expression has been observed in CRCs [48, 49] .

In addition to the WNT/Wingless signalling pathways B-catenin is involved in cell adhesion; within cell-cell adherens junctions, B-catenin functions to link the cell adhesion molecule E-cadherin to the actin cytoskeleton via A-catenin [50-52]

The *APC* gene is mutated in about 60% of early sporadic CRCs [33, 53, 54], and even with the same somatic mutation that occurs in FAP the affect on early tumourigenesis is less predictable [55]; inferring that epigenetic influences or other gene mutations may also have a role in early sporadic CRC tumourigenesis [56]. Over 60% of the somatic mutations in the APC gene are clustered within a small region of exon 15 [50], designated as the mutation cluster region, which accounts for less than 10% of the coding region. Frame shift and point mutations each account for about half of the mutations, the latter mostly nonsense mutations but some missense mutations occur [50].

1.4.2 Missmatch repair genes

MLH1 (OMIM 120436: MLH1) gene is mapped to chromosome 3p21.3 [57].
Mismatch repair is required for the cell to accurately copy its genome during cellular proliferation. Missmatch repair increases the fidelity of DNA replication by identifying and excising single-base mismatches and insertion-deletion loops that may arise during DNA replication. Thus, the mismatch repair system serves a DNA damage surveillance function by preventing incorrect base pairing or avoiding insertion-deletion loops by slippage of DNA polymerase [58]. Deficiencies of this system result in mutation rates 100-fold greater than those observed in normal cells [59, 60]. These mutations are particularly evident in microsatellite sequences, consisting of repeats of 1–4 base pairs, this microsatellite instability (MSI+) is thereby a hallmark of mismatch repair gene-deficient cancers. MSI has been observed in approximately 15% of sporadic CRCs [18, 19] and in virtually all CRC arising in patients with hereditary nonpolyposis colorectal cancer (HNPCC) [61, 62].
Mismatch repair gene mutations have been identified in about 26% of sporadic MSI+ cancers [63-65] , in the remaining 74% there may be mutations that have not as yet been identified or there is an epigenetic mode of gene inactivation. Hypermethylation of the promoter region is responsible for inactivation of several tumour suppressor genes [66-68]

and often occurs in *MLH1* in sporadic primary colorectal cancers with MSI. This methylation is also present in MSI+ tumours with known mutations of a mismatch repair gene and indeed it is present in a small subset of MSI-ve tumours [18]. It may therefore not be causative in all cases.

1.4.3 *KRAS2* oncogene

The *KRAS2* oncogene (OMIM 190070: V-KI-RAS2 KIRSTEN RAT SARCOMA 2 VIRAL ONCOGENE HOMOLOG; KRAS2) is mapped to 12p12.1-p11.1 [69]

The RAS oncogene product is a membrane-localised protein of 21 kd that functions as a molecular switch linking receptor and non receptor tyrosine kinase activation from the cell membrane to downstream cytoplasmic or nuclear events, mainly controlling the differentiation or proliferation of cells. The *RAS* family includes several distinct members *RAS* (Harvey, Kirsten, M, Neuroblastoma, and R), Rap (1 and 2) and Ral; they share at least 50% sequence identity. Activated *RAS* operates via a number of distinct signalling cascades, either leading to activation of genes, such as transforming growth factor- and vascular endothelial growth factor (VEGF), or changes in the actin cytoskeleton.

Oncogenic mutations in the gene are present in approximately 30% of all human cancers and are frequently limited to only one of the *RAS* genes in each cancer type. *KRAS* mutations occur frequently in non-small-cell lung, colorectal, and pancreatic carcinomas; *HRAS* mutations are common in bladder, kidney, and thyroid carcinomas; *NRAS* mutations are found in melanoma, hepatocellular carcinoma, and haematological malignancies. Thyroid carcinomas are unique in having mutations in *K,N* and *HRAS* genes [70]. Mutations most frequently occur in *KRAS* and least often in *HRAS*.

The *RAS*-signalling pathway has attracted considerable attention as a target for anticancer therapy because of its important role in carcinogenesis [70].

Cancer	Mutation frequency, %
Non-small-cell lung cancer	33
Colorectal	44
Pancreas	90
Thyroid	
Follicular	53
Undifferentiated papillary	60
Papillary	0
Seminoma	43
Melanoma	13
Bladder	10
Liver	30
Kidney	10
Myelodysplastic syndrome	40
Acute myelogenous leukaemia	30

Table 1.1: RAS mutations in human tumours

The *KRAS2* oncogene on chromosome 12, is mutated in 90 % of pancreatic carcinomas [70, 71] and 37% to 58% of CRCs [33, 72]. The mutations in CRC are on codons 12, 13, and 61, with different patterns in different populations [71] .

1.4.4 *TP53* gene

The *TP53* tumour suppressor gene (OMIM 191170: TUMOR PROTEIN p53; TP53) is located on chromosome17p13, the transcription factor TP53 regulates an essential growth checkpoint that both protects against genomic rearrangement and the accumulation of mutations. It controls both cell cycle progression and apoptosis.

DNA damage leads to rapid induction of TP53, which activates transcription of the Cdk inhibitor, this blocks cell cycle progression, both by acting as a general inhibitor of Cdk/cyclin complexes and by inhibiting DNA replication by binding to proliferating cell nuclear antigen [73]. The resulting cell cycle arrest allows time for damaged DNA to be repaired before it is replicated. Loss of TP53 prevents this damage-induced cell cycle arrest, leading to increased mutation frequencies and a general instability of the cell genome [74, 75]. TP53 is also required for apoptosis induced by DNA damage [73], a response that is presumably advantageous because it eliminates cells carrying potentially deleterious mutations. Cells lacking TP53 fail to undergo apoptosis in response to agents that damage DNA, including radiation and chemotherapeutic agents, such tumour cells can therefore be resistant to treatment.

Germ line mutations in *TP53* were initially reported in Li-Fraumeni syndrome [76-78], a rare autosomal-dominant disease with a predisposition to a wide range of malignancies including breast carcinoma, sarcomas and brain tumours [77]. The inheritance of a mutant TP53 allele results in a 25-fold increase in the chance of developing cancer by the age of 50, compared with the general population and almost 50% of affected members develop more than one malignancy.

TP53 mutations are very common in the development of somatic tumours [79], occurring late in tumourigenesis.[78]. Sporadic TP53 mutations occur in 61-70% of CRCs [33, 80], commonly missense mutations [81] and in more than 50% of pancreatic adenocarcinomas.

1.5 COLORECTAL CANCER GENETIC MARKERS OF SURVIVAL

Approximately 15,000 of the 29,000 individuals diagnosed with colorectal cancer (CRC) in England and Wales each year will die from the disease [82]. The mortality rate for CRC declined in the 1980's and early 1990s, but the rate has remained virtually static since the mid 1990's [83]. The American Society of Clinical Oncology (ASCO) guidelines recommend adjuvant chemotherapy for all stage III [84] and selected stage II CRCs [85, 86] but not all patients with stage III CRC will benefit and accurate selection of patients with stage II CRC who will benefit is difficult.

It is evident from the literature on CRC tumourigenesis that molecular genetic changes in the development of all CRCs is not the same [87]. Mutations in key genes in the progression of CRC lie on different, not one synergistic evolutionary pathway [33]. The MSI+ and CIN groups are quite distinct from each other in terms of genomic stability and frequency of copy number change. Cancers belonging to a particular pathological stage display significant clinical heterogeneity, which may reflect underlying molecular genetic heterogeneity.

Genetic markers could provide a useful refinement to existing pathological staging systems to more accurately identify patients with a worse prognosis [88] who would benefit from adjuvant treatment.

It is therefore highly desirable to identify any genetic markers from the many published studies that are of value in the management of CRC patients. The most precise assessment is achieved by systematic review of published studies using a standard meta-analysis technique with hazard ratio (HR) as the principle outcome measure. This has been undertaken by the Institute of Cancer Research (Royal Marsden, Sutton) on markers with the most consistent prognostic data, microsatellite-instability (MSI) [89], chromosomal-instability (CIN)[90] and thymidylate synthase (TS) [91].

1.5.1 Microsatellite instability

Fifteen percent of sporadic CRCs have microsatellite instability (MSI) [18, 19], they are predominantly located in the proximal colon, are poorly differentiated, mucinous, and show marked lymphocyte infiltration [92].

Thirty-two published studies are suitable to assess the relationship between MSI status and survival in CRC patients, reporting survival in a total of 7,642 cases, including 1,277 with MSI. Although many studies report better survival with MSI, there is considerable variation in the definition of MSI status, the pathological staging groups included and estimates of the prognostic value of MSI between studies.

A National Cancer Institute (NCI) workshop in 1997 recommended a reference panel, the Bethesda panel of 5 microsatellites for defining MSI status [93]. According to these guidelines, high frequency microsatellite instability (MSI-H) is defined as instability at two or more markers, and microsatellite stable (MSS) when all markers are normal. Only seven studies (7/32) have used this NCI recommended Bethesda panel of markers [94-99]. Patients were enrolled in clinical trials in only a minority of studies (5/32) [94, 99-101] raising the possibility of selection bias.

The largest pathological subset represented is the important stage II and III CRCs (13/32, 41%) [95, 98, 100-109]. Two studies [101, 105] present data suitable for examining the benefit of adjuvant 5-fluorouracil (5FU) chemotherapy in stage II or III CRCs by MSI status.

The overall survival (OS) for MSI+ CRCs is longer, confirmed by the combined HR estimate 0.65 (95% CI:0.59-0.71). This benefit is maintained restricting analyses to patients with stage II and III CRC and those enrolled in clinical trials, with no evidence of heterogeneity between studies.

There is also a disease free survival (DFS) advantage for MSI+ CRCs [96-98, 100, 102, 110, 111] , (HR 0.67, 95% CI:0.53-0.83).

Data is limited but MSI+ CRCs derive no benefit from adjuvant 5FU (HR=1.24 95% CI:0.72-2.14) [101, 105].

CRCs with MSI therefore have a significantly improved prognosis and are resistant to adjuvant 5FU.

1.5.2 Chromosomal instability

Eighty-five percent of sporadic CRCs are MSS, developing on an alternate pathway from the MSI group and demonstrate chromosomal-instability. They are characterised by loss or addition to their chromosomal set (aneuploidy), as well as more frequent chromosomal segmental loss or gain [112] and therefore allelic imbalance (AI) [20, 21].

Most studies focus on chromosome 18 as a marker of survival, as it is the region with the highest frequency of allelic imbalance (section 1.7) and appears to provide the best evidence for associated shorter survival in this CIN group. Alongside chromosome 18 allelic loss as a prognostic marker, published studies have also assessed loss of DCC expression as a prognostic marker but not SMAD 4 or SMAD 2 expression; this is probably historical, DCC being identified as a candidate tumour suppressor gene before SMAD 2 and SMAD 4.

Twenty-seven studies assessing survival by chromosome 18 allelic loss and *DCC* loss of expression are eligible for meta-analysis.

Only a minority of studies (5/27, 19%) [100, 101, 108, 112, 113] included patients enrolled in clinical trials the remainder being open to selection bias.

The largest group studied are the important stage II and III CRCs [20, 100, 101, 103, 108, 114-120].

There is little consistency in the methods used to assess chromosome 18 allelic loss and *DCC* loss of expression. Four methods are used to assess allelic loss in 24 studies, LOH is the most commonly used analysis [100, 101, 103, 108, 114, 115, 117, 120-126]. There are inconsistencies even within this method in the number, distribution and choice of markers used. The median number of markers assessed is only three (range 2 [101, 114, 117, 120, 122] -7 [121]). The range in the size of the region of chromosome 18 represented by the markers in each study varies enormously (4.6 to 43.6 Mb, with a median of 35.9Mb); in the majority of studies (9/13, 69%) markers map to a region that included *DCC* [103, 108, 114, 115, 121-123, 125, 126].

Most of the LOH studies (9/14, 64%) do not specify the number of markers required to display loss, to define the tumour as having allelic loss.

All of the *DCC* loss of expression studies used immunohistochemistry [113, 119, 122, 127] but with different antibodies and different expression thresholds.

Seven hundred and thirteen of the 1,322 stage II or III CRCs [20, 100, 101, 108, 115-117, 119, 120], have chromosome 18 allelic loss or DCC loss of expression; they have a shorter OS, (pooled HR=1.69, 95%CI:1.13-2.54) and DFS (HR=2.30, 95%CI:1.21-4.39) [108, 111, 118, 126]. There is insufficient evidence to examine the benefit of adjuvant 5FU in this group.

This shorter OS was maintained in CRCs with allelic loss that included the *DCC* gene or *DCC* loss of expression [108, 113, 115, 116, 119, 121, 122, 127] (HR=1.95, 95%CI:1.40-2.72) but this was also observed in CRCs where allelic loss did not include the DCC gene location (HR=1.84, 95%CI:0.74-4.58).

Allelic loss in the DCC region does not appear to be a discriminator of different survival in CRCs. It may be another tumour suppressor gene on chromosome 18 that is important, or that allelic loss on chromosome 18 is just a reflection of genome wide CIN.

Irrespective of the mechanism of action chromosome 18 allelic loss is associated with a shorter OS, but there is significant heterogeneity between studies.

1.5.3 Thymidylate synthase

Thymidylate synthase is (TS) the rate-limiting enzyme in the *de novo* synthesis of 2'-deoythymidine-5'-monophosphate (dTMP) [128], which is required for DNA synthesis. The antitumor effect of 5FU has been ascribed to a number of mechanisms [129-131], including competitive inhibition of thymidylate synthase. Thymidylate synthase may therefore play a role in determining tumour sensitivity to 5FU [132] . Although most studies have reported poorer OS and DFS with higher TS levels, estimates of the HR between studies differ wildly. As with many molecular marker studies, the clinical data, methods of assessing and defining TS differ markedly.

In most studies, patients were treated within the setting of clinical trials using unselected cases [113, 133-145].

In 2610 cases with stage II or III CRC [133, 135, 136, 140-142, 146], those with higher levels of TS have a shorter OS (1.35, 95% CI: 1.07-1.80), but with significant heterogeneity between studies.

Shorter OS is maintained in patients receiving surgery alone [133, 140, 146] but not in patients that receive adjuvant 5FU [136, 140, 146], indicating that adjuvant 5 FU has the ability to abrogate the poorer prognosis determined by high TS.

The prognostic value of TS remains uncertain because of the high level of heterogeneity in HR between relatively few studies. Additional studies with consistent methodology are needed before it could be considered for use in clinical studies.

1.5.4 *KRAS2* gene mutations

Conflicting study results [116, 125, 147-158] precluded clear conclusions being reached on the prognostic value of *K-RAS2* mutations, until the publication of the RASCAL I meta-analysis in 1998 [72, 84] and RASCAL II meta-analysis in 2001 [159]. The RASCAL II meta-analysis [159] of 3439 patients, identified only one *K-RAS2* gene mutation that was associated with a significantly worse DFS (HR 1.3, 95%CI: 1.09-1.54) and OS (HR 1.29, 95%CI: 1.08-1.55); a glycine to valine substitution on codon 12, present in only 8.6% of CRCs. This mutation is association with a reduced survival in stage III but not stage II CRCs. Poor response to adjuvant 5FU has been reported in stage III CRC with *K-RAS2* mutations compared to those without [151]. As a larger proportion of stage III than stage II CRCs received adjuvant chemotherapy this is a possible explanation for the shorter survival reported exclusively in stage III CRCs with *K-RAS2* mutations, rather than an inherently shorter survival.

The prognostic value of *K-RAS2* mutations in the management of CRC patients is uncertain; only one, uncommon *K-RAS2* mutation has been identified as predictive of poor survival and this is limited to stage III CRCs alone. Two recent publications have however provided further evidence for *K-RAS2* mutations predicting poor response to chemotherapeutic agents, both studies reporting no clinical response to cetuximab (anti-epidermal growth factor receptor) in metastatic CRC with K-RAS2 mutations [160, 161].

1.5.5 *TP-53* gene mutations

There is no clear evidence that *TP-53* mutations in CRC have prognostic value. Both protein levels as a surrogate for detecting *TP-53* mutations and direct mutation analysis of different regions of this very large gene have been utilised.

Half of studies report no association between TP-53 protein levels and survival [142, 143, 155, 156, 162-166] the other half report conflicting associations of both reduced and prolonged survival [167-175]. The results of studies assessing *TP-53* mutations directly proved no more conclusive, more reporting no association with survival [122, 142, 152, 176-178] than those reporting a significant association with reduced survival [153, 179-181]. Inter-study variability in the assessment of mutations may have contributed to the conflicting results. Multivariant analysis has identified 2 specific mutation locations of prognostic value [180, 182] and it is possible, that as for *K-RAS2* mutations in the RASCAL meta-analysis, assessing the impact of different mutations at specific locations on the *TP-53* gene may reveal a more consistent prognostic marker.

There is no evidence that *TP-53* mutations predict response to adjuvant chemotherapy despite TP-53 encoding a protein involved in cellular response to DNA damage [178].

1.5.6 Summary of colorectal cancer genetic markers of survival

The molecular profile of a tumour is likely to be a key determinant of clinical outcome [183, 184]. This has become evident in breast cancer [185, 186] and its exploitation to improve diagnostic accuracy in stage II and III CRC would be beneficial as locally advanced disease may be cured by surgery alone, with only a proportion deriving benefit from adjuvant therapy [187]. The identification of robust molecular prognostic markers to supplement conventional pathological staging systems is highly desirable, with the potential to improve survival, reduce morbidity and focus finite health resources; of

particular relevance with the emergence of new chemotherapeutic agents in the treatment of CRC [188, 189].

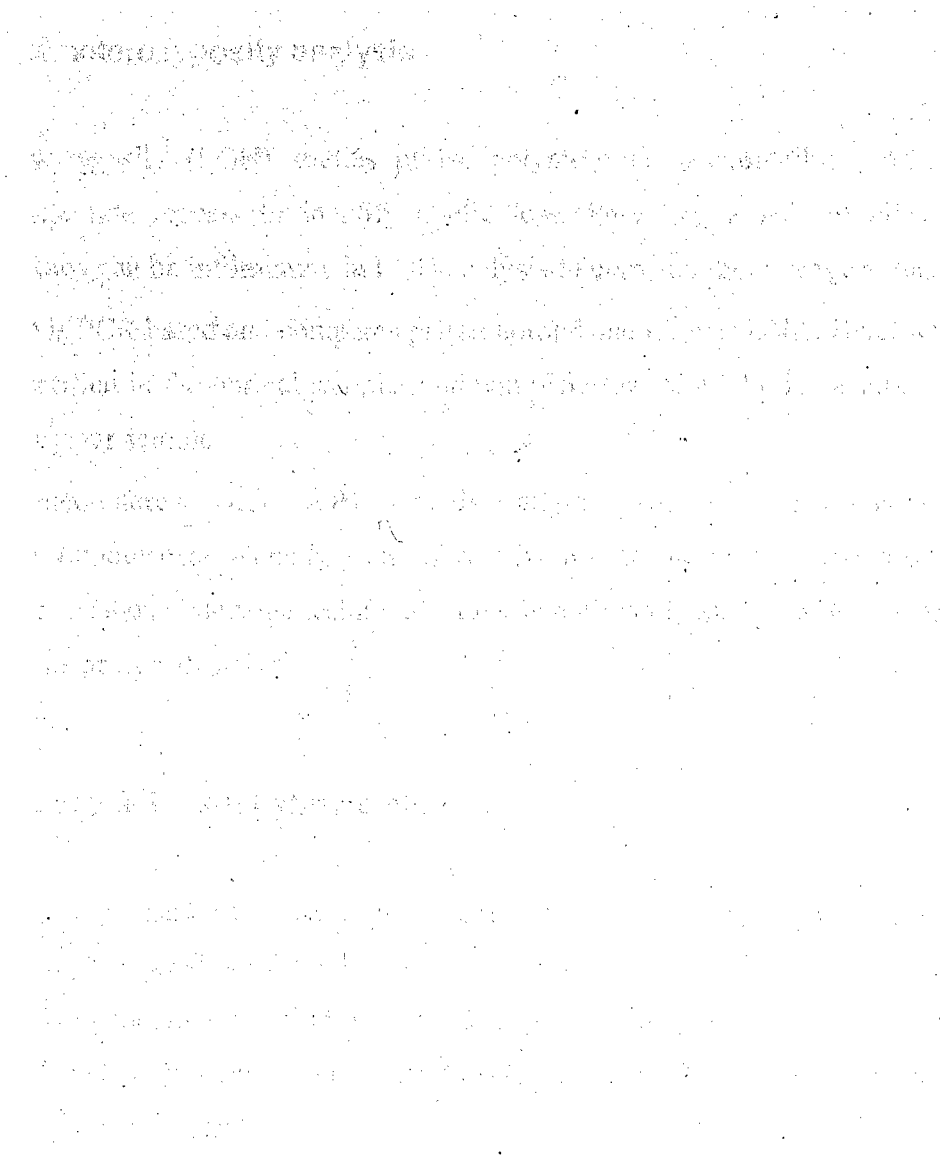
Numerous markers have been assessed in the published literature, but only MSI, chromosome 18 allelic loss and thymidylate synthase come close to consistent results. One of the major problems with many molecular marker studies is limited power, because of the small cohorts of cancers used in studies. For example only a third of MSI prognostic studies had 90% power to demonstrate a HR of 0.65 in favour of MSI, even stipulating a conservative significance level of 5%. To address the issue of power and to derive more robust estimates of prognostic value without substantial bias, meta-analysis of published studies have been undertaken for MSI [89], chromosome 18 allelic loss [90] and thymidylate synthase [91].

A meta-analysis from 1998 updated in 2001 of the prognostic value of *K-RAS2* mutations, revealed only one infrequent mutation to be of prognostic value and even this was restricted to stage III CRCs. Data on *TP-53* mutations is far too inconsistent, an overall pattern of its prognostic value is not evident from published studies. Relatively little interest remains in it as a prognostic marker and a meta-analysis would be of no value. High thymidylate synthase expression, from the meta-analysis, was found to be a marker for shorter OS and probably a good response to adjuvant 5 FU but because of the high level of heterogeneity between relatively few studies its value as a prognostic marker remains uncertain. A consistent methodology in a larger number of studies is required; but an appropriate technique has yet to clearly emerge.

The meta-analysis data confirm that CRCs with MSI have a significantly longer survival and are resistant to adjuvant treatment with 5FU, with no inter-study heterogeneity. It is therefore the only marker with sufficient current evidence for its inclusion in prospective clinical trials. However as only 15% of sporadic CRCs are MSI positive a large number of patients would need to be enrolled.

Chromosome 18 allelic loss was found to be associated with a worse OS and DFS in stage II and III CRCs but there is significant heterogeneity between studies and the critical region of loss is unknown. The higher resolution of micro-array CGH could perhaps more accurately define the critical region of loss and provide the opportunity for homogeneity between studies, if it was uniformly adopted. This is necessary to determine if

chromosome 18 allelic loss has a genuine role in clinical trials. It is eminently plausible that cancers with CIN, of the alternate pathway from the MSI chromosomally stable group in CRC development, should be markers of shorter survival. Sporadic CRCs have a relatively high frequency of chromosome 18 allelic loss (50-60%), which would provide a reasonable level of power in a clinical trial.



1.6 METHODS OF ASSESSING ALLELIC COPY NUMBER CHANGE

Genomic copy number changes are frequently found in cancers and are believed to contribute to their development and progression through inactivation of tumour suppressor genes and amplification of oncogenes. Identifying copy number change is a frequently used strategy to locate these genes [190-193]; it is also used to characterise different tumour types in terms of pathology, progression and survival.

1.6.1 Loss of heterozygosity analysis

Loss of heterozygosity (LOH) studies utilise polymorphic microsatellite markers of di/tri/tetra nucleotide repeats to identify allelic loss. Providing a pair of alleles are heterozygous they can be informative in LOH analysis of copy number change at that loci.

The technique is PCR based and compares paired normal and tumour DNA. Heterozygous alleles are identified in the normal samples and complete loss or allelic imbalance sought in the paired tumour sample.

Difficulties encountered with LOH include, allelic drop out, non-tumour cell contamination, tumour heterogeneity, non-informative alleles and marker resolution. It is a relatively slow, labour intensive technique which is not readily amenable to automation [194] limiting its power of analysis.

1.6.2 Fluorescence in situ hybridisation

Copy number change can also be detected using fluorescence in situ hybridisation (FISH). This is based on the sensitive detection of fluorescently labelled DNA probes that can hybridise to fixed metaphase chromosomes. The probes are generally genomic DNA clones labelled with a flurochrome that detects complementary DNA sequences in fixed metaphases from cells or tissues.

FISH was introduced in the early 1990s with the development of fluorescence techniques to replace isotope detection, increasing the sensitivity and resolution of in situ hybridisation [195].

The disadvantages of FISH as a screening tool is the requirement of prior knowledge of the type and location of expected aberrations and only a limited number of chromosomal loci can be analysed simultaneously.

Although a low through-put screening tool compared to more recently developed techniques, FISH can be useful in the confirmation of detected deletions or amplifications and clinical applications include monitoring disease progression in urothelial tumours [196, 197].

1.6.3 Comparative genomic hybridisation

Comparative genomic hybridisation (CGH), introduced in 1992, involves the direct comparison of tumour (test) and normal (reference) DNA to detect differences in copy number [198]. Tumour and normal DNA are labelled with different fluorescent dyes, co-hybridised to metaphase chromosomes and the ratio of colour intensities measured to give an estimate of copy number.

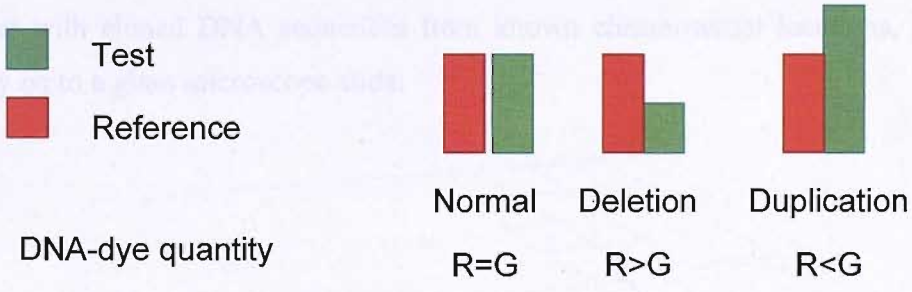


Fig. 1.4 Colour intensity ratios related to difference in test and reference copy number

CGH has developed into a very useful genome scanning technique because it enables the analysis of all chromosomes in a single experiment. CGH has been widely used to detect constitutional and somatic copy number change [199-202]. It has been utilised to identify chromosomal aberrations in foetal and neonatal genomes [203]; often proving more reliable than karyotyping in the diagnosis of complex structural abnormalities [204, 205]. CGH has been utilised in cancers to detect copy number change, identifying key genes [206] and providing diagnostic as well as prognostic information [207, 208].

The resolution of CGH using metaphase chromosomes can typically detect changes in the order of 5-10 mega bases (Mb) or at the highest resolution to 3-5Mb [199, 201, 209]; which is a major problem, limiting its clinical application [210].

A higher resolution technique is required; FISH was being used with resolutions down to 5-10 Kb but requires prior knowledge of the target. Microarray CGH combines the higher resolution of FISH with the whole genome screening potential of CGH, without the requirement of metaphases from dividing cells.

1.6.3.1 Microarray comparative genomic hybridisation

Microarray comparative genomic hybridisation (microarray CGH) overcomes the low resolution of conventional CGH by replacing metaphase chromosome as the hybridisation target with cloned DNA sequences from known chromosomal locations, gridded as an array on to a glass microscope slide.

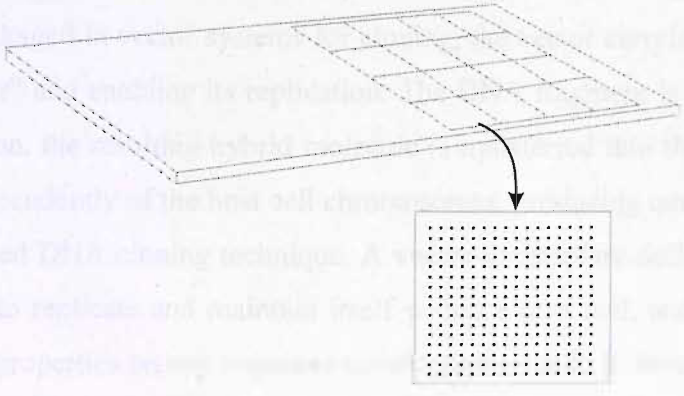


Fig. 1.5 Representation of microarray spots printed in a grid in 16 blocks on a glass slide

Resolution is determined by the size of the clones, typically 150 to 250kb, from the Human Genome and distance between the consecutive clones (Fig. 1.6), which is determined by the number of clones that can be processed for the selected region of genome coverage.

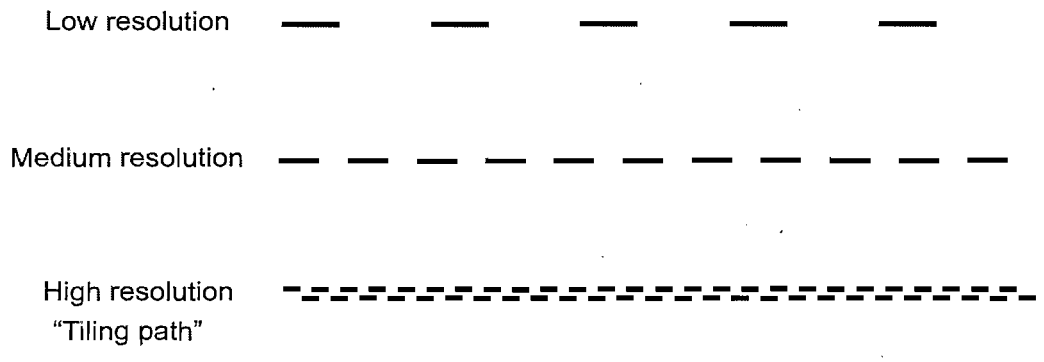


Fig. 1.6 Schematic representation of resolution, determined by size of clones and distance between consecutive clones.

The horizontal black lines represent clone length, resolution is low when the clones are longer and more widely spaced. The highest resolution to detect the smallest changes is achieved by smaller more densely packed overlapping clones.

The most commonly used genome wide clone set consists of approximately 3,200 clones of 150-250 kb in length at 1Mb intervals. A complete genome wide "tiling path" can be achieved with approximately 32,000 overlapping clones.

The relatively large 150-250 kb DNA sequences available following the Human Genome Project are packaged in vector systems for cloning, the vector carrying the DNA fragment as a "passenger" and enabling its replication. The DNA fragment is attached in vitro to a purified replicon, the resulting hybrid molecule is transferred into the host cells, where it replicates independently of the host cell chromosomes, producing quantities of pure DNA, in this cell based DNA cloning technique. A vector is therefore defined as a nucleic acid which is able to replicate and maintain itself within a host cell, and that can be used to confer similar properties on any sequence covalently linked to it; they have specific inserts including an antibiotic resistance gene, to allow selection of cells containing the vector.

Vectors with high copy number replicons will produce higher yields of DNA, a major disadvantage is that such vectors often show structural instability of inserts, resulting in deletion or rearrangement of portions of the cloned DNA. There are a variety of vector systems for cloning which have different sequence compositions and vary according to the size of DNA insert they can accept (table 1.2).

Cloning vector	Size of insert
Plasmid vectors	0-10 kb
Cosmid vectors	30-44 kb
Bacteriophage P1	70-100 kb
PAC (P1 artificial chromosome) vectors	130-150 kb
BAC (bacterial artificial chromosome) vectors	Up to 300kb
YAC (yeast artificial chromosome) vectors	0.2-2 Mb

Table 1.2 Cloning vectors

Vector systems that can incorporate the required 150-250 kb sequences include bacterial artificial chromosomes (BACs) which are based on E.coli fertility plasmid and P1 artificial chromosomes (PACs) based on bacteriophage P1. Bacteriophage P1 has a relatively large genome thereby affording the potential to accommodate large foreign DNA fragments, P1 has a protein coat, the P1 vector can be cleaved to generate 2 vector arms to which foreign DNA can be easily packaged. This can be incorporated into a protein coat, which has a high infection efficiency, into the host cell where it circularises and can be amplified. Part of the reason the BAC and PAC vector systems are able to accept the large inserts from the human genome and still maintain their structural stability is because they have low copy number replicons but therefore only produce relatively small yields of DNA.

The large-scale cell based clone cultures that would be required for the preparation of DNA for micro-arrays from low copy number replicons would be far too cumbersome because of the large number of clones and the large volumes that need to be produced

[211]. Partially degenerate oligonucleotide primer (DOP) polymerase chain reaction (PCR) is more suitable for this purpose, providing efficient amplification of small quantities of BAC DNA on a large scale.

DOP primers are used that specifically amplify human DNA efficiently but are poor at amplifying contaminating E. coli DNA. This DOP PCR technique maximises the capacity of each arrayed spot to bind human DNA and minimises non-specific background binding. Three primers are utilised with a 6 base degenerate sequence panel within each to minimize any sequence amplification bias (Fig. 1.7). The 3' sequence of the DOP-primers consists of 6mer panels selected for high occurrence in the human genome and low occurrence in the E. coli genome [212].

6MW (CCGACTCGAGNNNNNNATGTGG)
DOP 1 (CCGACTCGAGNNNNNNCTAGAA)
DOP 2 (CCGACTCGAGNNNNNNTAGGAG)
DOP 3 (CCGACTCGAGNNNNNNTTCTAG)

Fig. 1.7 Sequences of primers used for DOP-PCR amplification
N represents bases in a degenerate sequence panel

Amino linking PCR with a 5'amine modified primer is undertaken on the DOP-PCR product, the amine group forms a covalent bond with a glass surface, enabling it to act as a supporting platform for the array. The BAC/PAC clones are spotted in grids on to glass microscope slides using accurate high-density robotic printing, their location maintained by the amino covalent bond. The position of each unique DNA clone is known and can be reliably referenced to a chromosomal location [211, 213, 214]

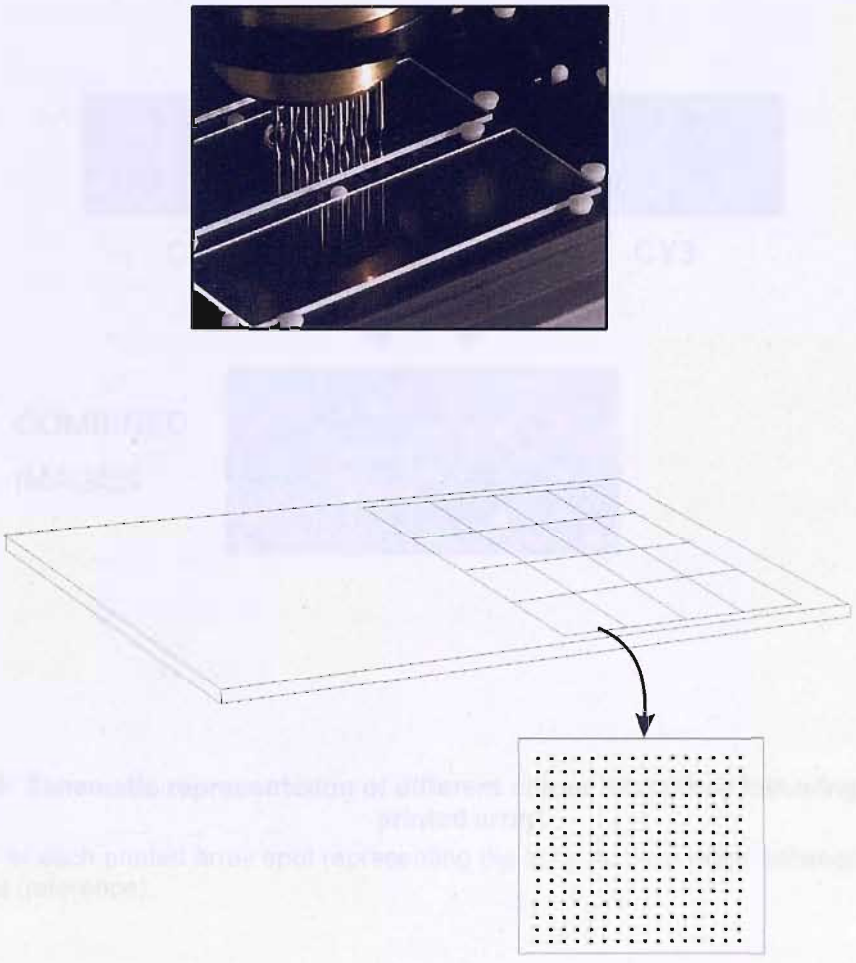


Fig. 1.8 Photograph of printing head of arrayer consisting of 16 pins. Representation of microarray spots printed by each pin in a corresponding block

As with conventional CGH, test and reference DNA are differentially labelled with fluorochrome dyes Cy5 (red) and Cy3 (green); the labelled DNA is then co-hybridised onto the microarray glass slide after hybridisation to repetitive DNA sequences has been blocked by the addition of Cot-1DNA. The slide is washed and laser scanned to quantitate emission from each of the fluorochrome dyes. The intensity ratio between the fluorochromes can then be utilised to determine loss or gain for each clone (Fig. 1.9).

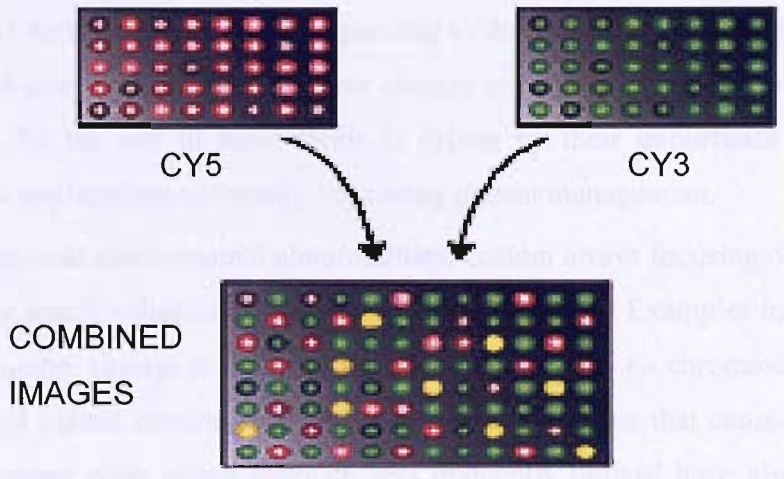


Fig. 1.9: Schematic representation of different colour intensities following CGH to the printed array.

The colour of each printed array spot representing the copy number ratios between tumour (test) and normal (reference).

The high quality fluorescent signal from the hybridisation targets is possible because of the technical ability to produce high concentrations of high molecular weight DNA in microscopic spots covalently bonded to a surface which has inherently low background fluorescence, a glass microscope slide [215, 216].

1.6.3.2 The clinical application of array comparative genomic hybridisation

The use of Array CGH is rapidly expanding to determine somatic copy number change in cancer and constitutional copy number change in congenital anomalies. The development of arrays for the use in these fields is driven by their importance in diagnostic and prognostic applications ultimately improving patient management.

In constitutional chromosomal abnormalities, custom arrays focusing on specific genomic regions for specific diseases have mostly been used [217]. Examples include the detection of copy number change in DiGeorge Syndrome, deletions on chromosome 22q11 causing cardiac and palatal abnormalities [217] and 1p36 deletions that cause mental retardation [218]. Genome wide arrays although less frequently utilised have also proved valuable, detecting a previously unknown deletion on chromosome 5q11 in a patient with features of DiGeorge Syndrome [219].

Array CGH has been used extensively to identify copy number change in cancers, being employed as both region specific and genome wide arrays. Continuous region specific tiling path arrays have located oncogenes [220] in breast cancer [221, 222] and oesophageal squamous cell cancer [223]. Lower resolution genome wide arrays have identified diagnostic sub-groups in prostate [224] and renal cell cancer [225] and have provided prognostic data linking copy number change in gastric cancer to a significantly different lymph node status and survival [226].

1.6.4 Alternative new technologies for investigating cancer molecular genetics

1.6.4.1 High density single-nucleotide polymorphic allele array.

High-resolution genome wide copy number analysis is also being developed using single nucleotide polymorphisms (SNPs).

More than 1.4 million SNPs have been identified; they are distributed throughout the human genome at an average density of one SNP per kilo base. Genome wide SNP analysis is possible using a commercially available chip based approach (Affymetrix HuSNP gene chip containing 1494 SNPs). These have already been used to detect LOH in breast, bladder and small-cell lung cancer [227-230].

SNP analysis has the potential to provide genotype information, in addition to detecting hemizygosity, SNP typing would indicate isodisomy (loss of one allele followed by duplication of the remaining allele) [231].

1.6.4.2 Expression arrays

Expression arrays utilise the same microarray technology to assess relative mRNA abundance. Instead of genomic DNA clones being arrayed onto glass slides, complementary DNA (cDNA) produced by reverse transcription from mRNA or oligonucleotides corresponding to an mRNA molecule (expression sequence tags) are used.

Expression arrays are capable of analysing genome wide transcripts simultaneously but data interpretation can be challenging with reproducibility and significance often variable between studies [232, 233].

Elevated expression may not identify all candidate oncogenes involved in tumourigenesis as only about 60% of genes with significant copy number gain show moderate or high levels of elevated expression [222]. Despite this, expression arrays have proved useful in characterising tumour sub-groups that relate to clinically relevant information, a good

example of this is diffuse large B-cell lymphomas (DLBCL). This lymphoma can be sub-classified by expression profile using a tailored gene chip into two groups, germinal centre B-like and activated B-like DLBCLs, with markedly different 5 year overall survival of 76% and 16% respectively [234]. The differences in expression profile between these 2 types of lymphoma is therefore accompanied by a remarkable divergence in clinical behaviour suggesting that these 2 subgroups of DLBCL should be regarded as distinct diseases.

Solid tumours have proved more difficult to sub-classify by expression profile; possibly because of heterogeneity both genetically and in the cell type within the tumour. One study of melanoma cells revealed expression profiles that correlated with progression to a metastatic phenotype [235] but other studies have not identified subgroups linked to clinically useful information.

In colonic cancers distinctive gene expression profiles have been identified for normal colonic mucosa and colonic adenocarcinomas [232, 233, 236, 237]. A similar proportion of genes were differentially up or down regulated in each study, ranging from 66/3400 (1.78%) to 235/9216 (2.5%) expressed sequence tags [232, 233, 237]; however only a few of the differentially expressed gene products were present in more than one study, transforming growth factor B [232, 233, 237], cytokeratin 20 and P-53 targets [233, 237]. This is probably because different gene chips were used in each of the studies but gives an indication of the vast number of genes apparently demonstrating differential expression in tumourigenesis.

In breast cancer, there are gene expression clusters associated with clinical data, a cluster including P-53 correlated with oestrogen receptor status and a cluster including HSP-90, which plays a role in oestrogen receptor signal transduction, correlated with tumour stage [238].

Recent work has revealed the existence of a class of small non-coding RNA species, known as microRNAs that regulate mRNA translation, there are 217 known microRNAs in the human genome. MicroRNA expression distinguishes tumours of different developmental origins which is useful in the classification and therefore treatment of poorly differentiated tumours [239].

1.7 COLORECTAL CANCER GENOME WIDE ALLELIC COPY NUMBER CHANGE

1.7.1 Loss of heterozygosity analysis using microsatellite markers

LOH studies utilising microsatellite markers have been used to assess genome wide copy number change in CRC. Several studies have reported chromosome 18 as the most frequent region of copy number loss, followed by chromosomes known to contain tumour suppressor genes (chromosome 17 and 5) as well as those that have no identified tumour suppressor genes (chromosome 8,4 and 14). There are however only a small number of these studies and they are limited by relatively infrequent markers irregularly distributed throughout the genome [240, 241].

	Chromosome	Frequency of loss/gain (%) [240]
LOSSES	18	69
	17	57
	8	50
	4	33
	14	31
	5	26
GAINS	20	67
	7	48
	8	45
	13	38
	20	38

Table 1.3: LOH analysis of genome wide allelic copy number change in colorectal cancer

1.7.2 Comparative genomic hybridisation analysis

Comparative genomic hybridisation (CGH) [209] to metaphase chromosomes, with the resolution to detect copy number changes of approximately 10Mb and over, has revealed consistent gains and losses in CRC. [242-245]. Confirming chromosome 18 as the most frequent region of loss, followed by chromosomes 8, 9 and 5. At this level of resolution significant loss of chromosome 17 has not been detected.

	Chromosome	Frequency of loss/gain (%)			
Study		**Ried[242]**	**Korn[243]**	**Aragane[245]**	**De Angelis[246]**
Study size		16	27	30	45
LOSSES	18	Y	89%	36%	49%
	8	Y	59%	-	Y (>10%)
	9	-	-	23%	Y (>10%)
	5	-	-	13%	Y(>10%)
	1	-	56%	-	Y(>10%)
GAINS	20	Y	85%	63%	67%
	8	Y	44%	43%	Y(>10%)
	16	-	-	40%	-
	13	Y	48%	-	Y(>10%)
	7	Y	44%	36%	Y(>10%)
	12	-	-	30%	Y(>10%)
	14	-	-	33%	Y(>10%)
	16	-	-	40%	-
	21	-	-	36%	-

Table 1.4: Metaphase CGH analysis of genome wide allelic copy number change in colorectal cancer. Ranked in descending order of loss/gain frequency.

Array CGH which offers greater resolution has been used in more recent copy number change analysis. Nakao et al [247] analysed 125 primary CRCs utilising 2463 genome wide BAC clones at a 1-2Mb resolution, reporting chromosome 18 as the most frequent region of loss, followed by chromosomes 17 and 8. The most frequent regions of gain were chromosome 20 and 8 (Table 1.5).

Study	Chromosome	Nakao K et al[247] Frequency loss/gain Primary CRCs
Study size		125
LOSSESS	18	60 (48%)
	17	46 (37%)
	8	37 (30%)
GAINS	20	65 (52%)
	8	42 (34%)

Table 1.5: 1-2 Mb array CGH analysis of genome wide allelic copy number change in colorectal cancer. Ranked in descending order of loss/gain frequency.

Thus LOH and metaphase CGH analysis identified 18q as the most frequent chromosomal arm of copy number loss but are limited by irregular marker position and poor resolution respectively. Array CGH has, with a higher level of resolution and larger sample numbers provided consistent supporting evidence; reporting 18 as containing clones with the highest frequency of copy number loss. Chromosome 17 is elevated in the rankings to the second commonest region of loss by array CGH which is not unexpected as it is the location of the TP53 tumour suppressor gene.

1.8 CHROMOSOME 18 ALLELIC COPY NUMBER CHANGE IN COLORECTAL CANCER

Thiangalingham in 1996 [241] using loss of heterozygosity analysis was the first to attempt to delineate a minimally lost sub-region on chromosome 18q, with 13 microsatellite markers. Fifty-five of 100 CRC cell lines showed allelic loss of 18q, of these, 38 had loss of all markers, leaving only 17 with incomplete marker loss to delineate a MLR. The consistently lost region was flanked by markers, D18S535 and 20C03 encompassing a region of 37 MB. The author then suggested that homozygous loss of markers D18S535 and D18S858 in 2 cell lines within this region was 'indicative' of the target tumour suppressor gene lying within a 16 MB region in 18q21 bounded by these two markers. This 16MB region contains the genes, DCC and SMAD 4.

Terdiman in 2001 [248], utilising 11 markers, identified 18q loss in 25 of 29 (86%) ulcerative colitis related CRCs. Allelic imbalance at all informative markers was detected in 14 (48%) tumours, the other 11 (38%) showing partial loss. The marker most commonly lost (D18S363) corresponding to SMAD4. Quantitative PCR was undertaken on these 29 ulcerative colitis related CRCs to assess relative loss of copy number of 3 candidate tumour suppressor genes on chromosome 18, SMAD2, SMAD4 and DCC; corresponding marker loss was detected in 40%, 57% and 53%, respectively.

This method of analysis reveals 18q loss as a common event in CRC tumourigenesis and can be used to identify regions of deletion within 18q but these regions are relatively large and do not accurately locate candidate tumour suppressor genes.

1.9 CHROMOSOME 18 CANDIDATE TUMOUR SUPPRESSOR GENES

Chromosome 18 copy number analysis has identified deleted regions that include the candidate tumour suppressor genes, *DCC, SMAD 2* and *SMAD 4* [248].

	SMAD 2	SMAD 4	DCC
Chromosomal Position	q21.1 43621608 to 43711221bps	q21.1 46810611 to 46860139bps	q 21.2 48121156 to 49311021bps
Exons	11	13	29
Transcript length	2,098 bp	3,196 bp	4,344 bp
Protein length	467 residues	552 residues	1,447 residues

Table 1.6: The chromosomal position, number of exons, transcript and protein length for SMAD 2, SMAD 4 and DCC.

All 3 candidate tumour suppressor genes map to the 16Mb region of deletion suggested by LOH analysis using microsatellite markers [241].

DCC (OMIM 120470: DELETED IN COLORECTAL CARCINOMA; DCC) is a large gene, it codes for what appears to be a transmembrane protein whose extra cellular domain has similarities to that of the cell-adhesion molecule N-CAM, suggesting that it may be involved in cell-cell or cell-matrix adhesion or in the reception of signals from the cell's environment. *DCC* expression is rarely lost in CRCs [249, 250] and mutations are evident in less than 2% of CRCs [103].

SMAD2 (OMIM 601366: MOTHERS AGAINST DECAPENTAPLEGIC, DROSOPHILA, HOMOLOG OF, 2; SMAD2) and *SMAD4* (OMIM 600993: MOTHERS AGAINST DECAPENTAPLEGIC, DROSOPHILA, HOMOLOG OF, 4; SMAD4) are members of the SMAD family of transcription factors, that are downstream regulators of tumour growth factor (TGF) β. There are three types of SMAD proteins: receptor-regulated SMADS (R-SMADS), CO-SMADS, and inhibitory or antagonistic SMADS.

Residues near the C-terminus of R-SMADS are phosphorylated by activated type I TGFβ receptors they then dimerize with CO-SMADS. The resulting heterodimers translocate to the nucleus and cooperate with other transcription factors to activate transcription of specific target genes (Fig. 1.10).

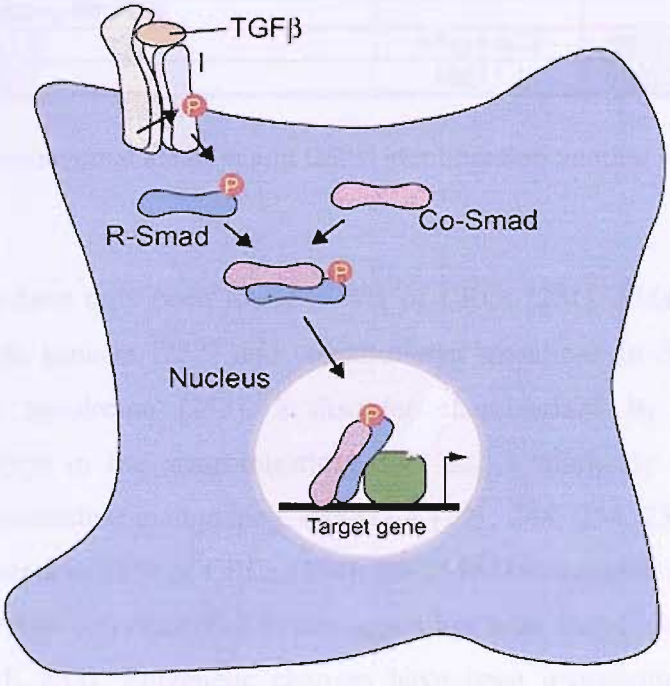


Fig. 1.10: The TGFβ signal-transduction pathway

The TGFβ signal-transduction pathway. Ligand binding induces activation of TGFβ receptor complexes, which then directly phosphorylate (P) specific R-SMADS. These R-SMADS associate with CO- SMAD and then translocate to the nucleus, where they interact with transcriptional factors to regulate gene expression.

Genome Biol. 2001; 2(8): reviews3010.1–reviews3010.8.
Published online 2001 August 2.

Name	Chromosomal Location	OMIM ID
R-SMADS		
SMAD2	18q21.1	601366
SMAD3	15q21-q22	603109
SMAD 1	4q28	601595
SMAD 5	15q31	603110
SMAD 8	13q12-q14	603295
CO-SMADS		
SMAD 4	18q21.1	600993
Inhibitory SMADS		
SMAD 6	15q21-q22	602931
SMAD 7	18q21.1	602932

Table 1.7: Chromosomal location and OMIM identification number of the SMADS

SMAD 2 mutations have only been found in 8% of CRCs [251]. *SMAD 4* is frequently deleted in pancreatic cancers [252] and constitutional mutations in *SMAD4* can cause Juvenile Polyposis Syndrome [253], a disorder characterized by the presence of hamartomatous polyps in the gastrointestinal tract and a markedly increased risk of developing a gastrointestinal malignancy. SMAD 4 [241, 248, 254, 255] expression has been found to be absent in 38% of CRCs [254]; the SMAD 4 mutation rate in CRCs with allelic copy number loss corresponding to this region has been found to be lower, reported as 15 to 20% [241, 255]. Epigenetic changes have been investigated, transcriptional silencing can be the result of methylation of CpG islands located in the promoters of genes [256] as evident in MSI+ CRCs. SMAD4 has a clearly identified promoter region including the non coding exon 1, with CpG islands as well as TATAA and CCAAT boxes [257] but there is no evidence of hypermethylation of this promoter [258].

DCC and SMAD 2 are unlikely to be causal tumour suppressor genes in CRC, as they have very low mutation rates.

There is more evidence in favour of SMAD 4 as a candidate tumour suppressor gene; expression is often absent, it has a higher mutation rate but there is still no clearly identified mechanism of loss of function in the majority of CRCs. It is still unclear if it plays a key role in colorectal tumourigenesis.

It is probable that the critical region of deletion on chromosome 18 can be more accurately defined by using a higher resolution technique. This should more accurately locate the key tumour suppressor gene in this chromosome 18 region.

1.10 HYPOTHESIS

Inactivation of a key tumour suppressor gene by deletion in chromosome 18 is critical to colorectal cancer tumourigenesis.

1.11 AIMS

-Investigate CRC samples for deletions within chromosome 18.
-Exploit high resolution array-CGH to localise key genes involved in CRC tumourigenesis.

CHAPTER 2: METHODS

2.1 MICROARRAY CONSTRUCTION

The constructed array provided complete tiling path coverage of chromosome 18.

Two types of control were included, genome wide positive controls spaced at approximately 5 Mb intervals and negative controls in the form of BAC Drosophila clones.

2.1.1 Identification of the target clones

DNA was acquired for the 860 clones providing tiling path coverage of chromosome 18 from the Children's Hospital Oakland Research Institute (CHORI) 32,000 genome wide clone set. Clone DNA is available commercially from the BAC and PAC Resource Centre(BPRC), CHORI, California (PI: Pieter De Jong PhD) (www.bacpac.chori.org/libraries) and was purchased for utilisation in this project by the Molecular and Population Genetics Laboratory, Cancer Research UK, Lincoln's Inn Fields.

Target BAC and PAC genomic clones constituting 600 genome wide positive controls at 5Mb intervals were selected from the Sanger Centre 1 Mb clone set (www.ensemble.co.uk). These clones were obtained from the Wellcome Trust Sanger Institute Resource as agar stabs, prepared as inoculated cultures according to standard protocols (sections 2.1.2 to 2.1.4) prior to DNA extraction using standard miniprep methodologies (sections 2.1.5 and 2.1.6).

2.1.2 Testing BAC/PAC clones for phage contamination.

Agar plates were prepared (see section 2.5) and placed into a 37°C incubator for at least an hour. A DH10B *E.coli* culture was grown up over-night in 10ml Luria Bertani (LB) (see section 2.5) broth at 37°C; 4g of agarose was added to 500ml of LB broth, microwaved until the agarose dissolved and placed in a 45°C water bath. A 10ml alliquot of *E. coli* culture was added to the agarose/LB broth and thoroughly mixed. A 10 ml aliquot of broth/culture mix was then poured on to the incubated agar plates. Once the plates had set, genomic clones were streaked out using a sterile tip and placed

in a 37°C incubator overnight. Plates were checked the following day to see if the bacterial lawn had been lysed. Phage contamination results in lysis of the bacterial lawn and would appear as translucent areas on the bacterial lawn.

2.1.3 BAC/PAC growth on agar plate and in broth

LB agar was prepared, with appropriate antibiotic added (chloramphenicol 20mg/ml if BAC inoculation, kanamycin 25mg/ml if PAC inoculation) and poured as plates. A Sterile tip was used to transfer clones and streak them onto an agar plate. Inoculated plates were placed in an incubator at 37°C overnight.

LB broth was prepared, with appropriate antibiotic added, 500 ul aliquoted to each well of a deep 96 well plate. Single colonies were picked in turn from LB plates and inoculated into a single well, the deep 96 well plate sealed with a gas permeable seal and placed in a 37°C shaker overnight at 100 g (260rpm).

2.1.4 DNA extraction/ Micro-prepping procedure

250ul of culture from each well was mixed and then transferred to the corresponding position in a 96 well round bottomed plate with a lid; spun for 4minutes at 500g and growth confirmed by presence of pellets. The supernatant was discarded; to each of the pellets, 25µl of solution I (See section 2.5) was added (vortexed to resuspend), 25µl of solution II (See section 2.5) (prepared immediately prior to commencing the extraction procedure, mixed and left at room temp. for 5 mins.) and 25µl of solution III (See section 2.5) (mixed and left at room temperature for 5 mins.). The content of each well was transferred to a 96 well filter plate covered with a Dynex plate seal, located on top of a 96 well round bottomed plate containing 100ul of isopropanol in each well and spun at 2500rpm for 2 mins. The filter plate was discarded, the isopropanol plate left at room temperature for 30 minutes then spun at 600g (3200 rpm) for 20 minutes. The supernatant was discarded and the visible pellets dried for 15 minutes. 100µl of 70% ethanol was added to each well, the plate gently tapped to wash the pellets and then spun at 600g (3200 rpm) for 10 minutes, then dried for 5 minutes; this process was repeated a second time drying for 30 minutes. The DNA pellets were then resuspended in 50 µl of TO.1E with RNase (10ul of 1mg/ml RNase per 1ml of T0.1E), covered and left at room temperature for 30 minutes and stored at –20°C.

2.1.5 Amplification of BAC and PAC DNA using Degenerate oligonucleotide primer (DOP)-PCR.

Degenerate oligonucleotide primer-PCR was carried out on genomic DNA extracted from BACs and PACs using three DOP-PCR primers in separate reactions (the design of the DOP-PCR primers is discussed in more detail in section 1.4.3.1). To ensure no genomic contamination, the dH2O, buffer, W1, tips and pipettes were exposed to UV light in a light proof box for 30 minutes and the DOP-PCR carried out in the same box accessed via a removable side panel. DOP-PCR amplification was carried out in a 50µl reaction mix of 5µl DOP-PCR buffer (see section 2.5), 5µl DOP primer (20µm), 4µl dNTPs (2.5mM), 5ul Mg Cl_2, 2.5µl 1% W1, 0.5µl, 5u/µl Taq polymerase, 25.5µl H_20, 2.5µl (5ng) BAC/PAC DNA.

PCR programme:

94°C for 3 minutes,
94°C for 1.3 minutes,
30°C for 2.30 minutes,
ramp at 0.1°C per second to72°C,
72°C for 3 minutes,
go to step 2 for 9 cycles,
94°C for 1 minutes, 62°C for $1^1/_2$ minutes,
72°C f or 2 minutes,
go to step 7 for 29 cycles,
72°C for 8 minutes, 12°C for ever.

DOP PCR products of 8 Drosophila clones were provided by the Molecular and Population Genetics Laboratory, Cancer Research UK, Lincoln's Inn Fields

2.1.6 DOP–PCR product visualisation

5µl of each product was run with 3µl of loading dye on a 2 % agarose gel, visualised and a hard copy taken on a fluorimager.

2.1.7 Attachment of 5' amino group using amino-linked PCR primers.

Amino PCR was carried out to attach an amino group to the 5' terminus of the DOP-PCR product to facilitate attachment to the amine-binding glass slide. This was carried out using PCR primers with an amino-group at their 5' end (see section 1.4.3.1). Amino-PCR was carried out on combined DOP 1,2 and 3-PCR amplified BAC and PAC DNA. Conditions were as follows: 6µl aminolinking buffer, 6µl dNTPs (2.5mM), 3µl aminoprimer (20µM), 0.6µl 5u/µl *Taq* polymerase, 42.4µl H_2O, 2µl DOP-PCR product. PCR programme using MJ thermocyclers:

95°C for 10 minutes,
35 cycles of 95°C for 1 minutes,
60°C for $1^1/_2$ minutes,
72°C for 7 minutes,
72°C for 10 minutes, held at 10°C.

2.1.8 Amino–PCR product visualisation

5µl of each product was run with 3ul of loading dye on a 2 % agarose gel, visualised and a hard copy taken on a fluorimager.

2.1.9 Printing of the array-slides

Ten 96 well plates containing DOP/ amino PCR products for the clones constituting the chromosome 18 tiling path array, 15 96 well plates (comprising of some empty wells) containing the DOP/ amino PCR products for the control clones and 8 96 well plates (comprising of some empty wells) containing the DOP/ amino PCR products of 8 Drosophila clones were transferred to 5 384 well plates.

The 384 well plates were divided into 4 columns, 6 wells wide, 2 columns containing chromosome 18 clones interspersed with 2 columns containing control and Drosophila clones. 15 µl of the amino-linked products for each clone were combined with 15µl 4x microarray spotting buffer in each well.

An excel spread sheet was generated identifying the content of each well of the five 384 well plates by clone name as ID, the chromosomal location of each clone was then imported during data analysis using clone ID and the most up to date positional data. This excel spread sheet was converted to a text file, by saving as a CSV file type,

R clicking to open, opening with notepad. This text file was then imported into "datatraker programme" on the Genetix arrayer as a Q soft library manager file using "import process file" on the "tools" tab.

The genetix arrayer print head was set up with sixteen pins in a 4x4 configuration to print 16x16 spot blocks; the clones were printed in duplicate in a random pattern onto coated amine-binding slides (Amersham, codelink). Protocols for the slide printing can be found at http://www.sanger.ac.uk/Projects/Microarrays/.

The 384 well plates were stored at –20 C, defrosted for 1 hour at 37°C and centrifuged at 250g (1000 rpm) for 30 seconds prior to removal of plate seals. The individual wells within the 384 plates were thoroughly mixed and water added to each well to achieve a volume of between 15 and 25µl in preparation for printing.

2.1.10 Denaturing slides

Residual reactive groups were blocked using a pre-warmed blocking solution at 50°C for 30 minutes. The blocking solution was discarded and the slides rinsed twice with de-ionised water. The slides were washed with 4xSSC (see section 2.5), 0.1 % SDS (pre-warmed to 50°C) for 30 min on a shaker. The wash solution was discarded and the slides briefly rinsed in de-ionised water, then placed in boiling water for 2 minutes. The slides were rinsed twice with de-ionised water and then centrifuged at 200g (800 rpm) for 3 min in a separate 50ml Falcon tube containing Whatman paper for drying. The slides were stored at ambient temperature in a slide box (which included a desiccant for long term storage).

2.1.11 Assessment of printed spot quality

The quality of each batch of printed slides were assessed by determining the spot morphology and intensity by hybridising fluorescently labelled oligonucleotide mix and then scanning in the array scanner.

For a printed array area of 22x22 mm, 20ul of Spot Check (Genetix) was used. The Spot

Check mixture was denatured by heating to 95°C for 5 minutes, then immediately cooled on ice. The cooled mixture was pipetted onto the array and a cover slip placed

on top without the introduction of bubbles. The slide was kept in the dark at room temperature for 20 minutes.

The slide was washed at room temperature in, 1xSSC, 0.1 % SDS for 10 minutes, 1xSSC for 5 minutes and 0.1xSSC for 30 seconds. The slide was then dried in a 50 ml Falcon tube, centrifuged at 1000 rpm for 1 minute and a scanned.

2.2 PREPARATION OF COLORECTAL CANCER SAMPLES

Resected colorectal cancer specimens were prepared in the routine manner by a pathologist. Blocks of tumour and normal tissue approximately 10x6x6mm were provided. The tissue blocks were collected from the Pathology Department (Southampton University Hospitals NHS Trust) by the Tissue Library Supervisor, wrapped in parafilm, placed in cryo-vials and transported to the Cancer Sciences Building. The samples were snap frozen, anonymised, catalogued and stored in liquid nitrogen. A copy of the signed consent form and routine pathology report were securely retained by the Tissue Library Supervisor for each linked anonymised sample.

2.2.1 DNA Extraction

The tissue was placed in a 2 stop tube, 750µl of Tail buffer (see section 2.5) and then 40µl of proteinase k (see section 2.5) added. The tube was vortexed and incubated overnight in a water bath at 56°C.

The tube was vortexed then rocked for 5 minutes and the contents transferred to an eppendorf. This was centrifuged at 11,000g (11,000 rpm) for 10 minutes, 750µl of supernatant transferred to a new eppendorf, 500µl of isopropanol added and the tube inverted several times. The DNA pellet was added to a new tube containing 500µl of 70% ethanol, the ethanol was removed and the pellet dried until gelatinous around the edges. Approximately 1 ml of dH_20 was added depending on the quantity of DNA.

2.3 MICROARRAY COMPARATIVE GENOMIC HYBRIDISATION

2.3.1 Random-prime labelling of genomic DNA

0.6μg (600ng) of DNA was labelled using an Invitrogen Bioprime array CGH genomic labelling system. 0.6μg DNA in 21μl dH_20 was added to 20μl X2.5 Random Primers Solution (Invitrogen kit). The DNA was denatured in a heat block for 5 minutes at 95°C, and immediately cooled on ice for 5 minutes. On ice 5μl X10 dCTP mix, 3μl Cy3 or Cy5 labelled dCTP (1mM), 1μl Klenow fragment were added (Invitrogen kit). The reaction was incubated at 37°C for 2 hours. The reaction was stopped by adding 5μl of stop buffer (Invitrogen kit).

2.3.2 Purification of labelled probes

45μl of TE pH 8.0 (see section 2.5) and 400μl of purification buffer A (Invitrogen purification module) was added to each tube and vortexed for 30 seconds. The sample was loaded into a purification (Micro-spin G50) column within a 2ml collection tube. The columns were centrifuged at 11,000g (11,000 rpm) for 1 minute at room temperature and the flow through discarded. 600 μl of purification buffer B (Invitrogen purification module) was added to the column and centrifuged at 11,000 rpm for 1 min at room temperature and the flow through discarded. 200 μl of purification buffer B (Invitrogen purification module) was added to the column and centrifuged at 11,000g (11,000 rpm) for 1 minute at room temperature and the flow through discarded. The purification columns were placed in new sterile 1.5ml amber collection tubes, 50μl of H_2O added to the column, incubated at room temperature for 1 min and then centrifuged at 11,000 rpm for 1min (the flow through then contains the purified labelled DNA probes). 140μl of H_2O added to increase volume of hybridisation solution.

2.3.3 Precipitation of pre-hybridisation solutions

Pre-hybridisation solution (T2) was made with 80μl of precipitated herring sperm (10mg/ml), 135μg of human Cot1 DNA, 23μl of NaAc (3M; pH5.2) and 400μl of 100% cold ethanol.

Human Cot1 DNA was used as a blocking agent. Highly repetitive short DNA sequences within genomic DNA are in large excess over the labelled control and test DNA. As a result such sequences will readily associate with the complementary strands in the labelled DNA, effectively blocking their hybridisation to target sequences. This was overcome using a fraction of genomic DNA that is enriched for highly repetitive DNA sequences, DNA with a cot value of 1.0 (Cot-1 DNA).

Hybridisation solution (T1) was made with 180µl of each labelled DNA, 135µg of human Cot1 DNA (Roche), 55µl of NaAc (3M; pH5.2) and 1000µl of 100% cold ethanol. The T1 and T2 tubes were gently mixed and precipitated overnight at -20°C.

2.3.4 Resuspension of DNA

Precipitated DNA was spun for 15 minutes at 14,000g (14,000rpm). The supernatant was discarded. 500µl of 80% ethanol was added and re-spun at 14,000rpm for 5 minutes. Supernatant was removed and the samples were re-spun at 13,000rpm for 1 minute. The remaining supernatant was removed using a pipette.

Tube T1 containing the labelled DNA was resuspended in 6µl yeast tRNA (100µg/µl) and then 60µl of pre-warmed (75°C) hybridisation solution. Tube T2 containing the pre-hybridisation DNA was re-suspended in 140µl pre-warmed hybridisation buffer. Both tubes were placed for 3 minutes in a 75°C heat block and ensured that fully resuspended

The DNA in both tubes was denatured for 10 minutes at 75°C (mixing again after 5mins), both tubes were pulse spun. Tube T1 containing the labelled DNA was then incubated at 37°C for 60 minutes in the dark.

2.3.5 Microarray pre-hybridisation

Three layers of rubber cement were applied around the array grids. Pre-hybridisation solution (tube T2, denatured Herring sperm/Cot1 mix) was applied to the well, rocked to ensure even coverage and the slide transferred to a humidified chamber, containing Whatman paper soaked in 2x SSC/40% formamide, and placed on a rocking table for 60 minutes (5 rpm).

2.3.6 Microarray hybridsation

The pre-hybridisation solution was removed with a pipette after tilting the slide. The hybridisation mix was applied to the slide, which was rocked to ensure even coverage and to mix with the remaining prehybridisation solution. The slide was then transferred to a slide mailer humidified with 20% formamide/2xSSC and sealed with para-film placed in a 43°C incubator for 48 hours in the dark.

2.3.7 Washing array slides

The slide was removed from the slide mailer, the rubber cement removed and the slide washed in PBS/0.05% Tween 20. The slide was transferred to fresh PBS/0.05 Tween 20 solution shaking for 10 minutes, after which it was placed in pre-heated (43°C) 50% formamide/ 2x SSC solution shaking for 30 minutes in a 43°C heat block. The slides were rewashed in PBS/0.05 Tween 20 shaking for 10 minutes and dried by spinning at 2000 rpm for 5 minutes in a 50 ml Falcon tube containing a strip of Whatmann paper. Slide stored in light proof box until ready to scan

2.3.8 Scanning of array slides and data analysis

Slides were scanned using an Agilant scanner, the images captured as separate TIFF files for Cy3 and Cy5. The acquired images were analysed by use of Bluefuse software producing raw intensity ratios for each spot.

Data was analysed Blue fuse quantification/ post processing.

Positional data was imported an excel spread sheet (cfg file) with the most up to date clone positional data available from the ensemble and CHORI websites (March 2005 release) for positive controls and chromosome 18 tiling path clones respectively.

Blue fuse analysis involved post processing analysis by block normalisation and spots excluded according to a function of the Standard deviation around the mean value and confidence grading (Standard Deviation >3, confidence less than 0.4). Significance of amplifications and deletions was assessed using array-CGH smooth (utilising standards log thresholds of -0.6 for deletions and 0.5 for amplifications)

2.4 FLOURESCENT IN SITU HYBRIDISATION (FISH)

2.4.1 DNA EXTRACTION (Mini Scale Preparation)

200ml of LB broth 37°C overnight culture was transferred to a 50 ml Falcon tube and centrifuged at 6,000g (6,000 rpm) for 15 minutes. The supernatant was removed and Falcon tube inverted for 5 minutes. The harvested cells were resuspended by vortexing until the pellet dispersed, then 2ml of resuspension solution was added and complete resuspension ensured. The cells were lysed by adding 2 ml of lysis solution, inverting 4-6 times to mix and incubated at room temperature for 5 minutes.

2ml of chilled neutralisation solution was added to the lysate, which was immediately mixed by inverting 6-8 times and incubated on ice for 5-10 minutes. The tube was centrifuged at 15,000g for 20 minutes at 2-8C°. The clear lysate was transferred to a glass centrifuge tube. Nucleic acids were precipitated by adding 3.6ml of isopropronol to the cleared lysate, which was thoroughly mixed by gentle swirling and centrifuged at 12,500 for 20 minutes at 2-8°C. The supernatant was poured off, the pellet washed with 2ml of 70% ethanol, centrifuged at 12,500g for 10 minutes at 2-8°C and the remaining supernatant removed. The pellet was air dried for 5 minutes

The pellet was resuspended with 650µl of elution solution; 1ul of RNase cocktail was added to the tube and incubated at 60°C for 10 minutes.

The salt concentration was adjusted by adding 50µl of sodium acetate buffer solution (3M, pH 7).120µl of Endotoxin Removal Solution was added; the tube chilled on ice for 5 minutes, the solution then became light blue and clear.

The tube was then warmed at 37°C for 5 minutes, the solution then became cloudy. The phases were separated by centrifuging at 16,000g for 3 minutes. The clear upper phase containing the BAC DNA was transferred to a new microcentrifuge tube and the process repeated.

The clear upper phase was transferred to a 2ml microcentrifuge tube. 700µl of DNA Precipitation Solution added and mixed by gentle inversion. The tube was centrifuged at max speed for 20 minutes at 4°C, the supernatant removed, the pellet washed with 500µl of 70% ethanol and then the process repeated with 100µl of 70% ethanol. The supernatant was removed and the pellet air dried for 5 minutes. The DNA was then

dissolved in 100µl of sterile deionised water, mixed for 1 hour and concentration measured by spectrophotometry.

2.4.2 BAC labelling by nick translation

5µl of 10x nick translation buffer, 5µl dNTP mix (0.5mM dACGTP/0.4mM dTTP), 5µl 0.1M beta-mercaptoethanol (14.4M), 5µl nick translation mix (3u/ml DNAseI and 1000u/ml DNA polymeraseI) and 2.5µl of fluorescent dye (BAC texas green and marker texas red) were added to 27.5µl of dH2O containing 1ug of DNA. This was incubated at 15°C for 2 hrs then stored at 4°C prior to checking the product on an electrophoresis gel (1% agarose, 0.5 TBE)

2.4.3 Probe preparation

The labelled DNA was precipitated with 8µg of human Cot-1 DNA, 2.5µl of sonnicated herring sperm DNA, 2.5µl of tRNA, $1/10^{th}$ vol 3M Na acetate pH 5.2 and X3vol cold (-20°C) 100% ethanol and then left at –20°C for 15 minutes. It was then centrifuged at 14,000 rpm for 25 minutes, the supernatant discarded, the pellet washed with 400µl of 70% ethanol and again centrifuged at 14,000g (14,000rpm) for 10 minutes before discarding the supernatant. The pellet was then dried in an incubator at 37°C. The pellet was then resuspended in 50µl of hybridisation mix (7.5ul deionised formamide, 6ul dextran sulphate (50%) and 1.5µl 20xSSC) for 30 minutes at 45°C.

2.4.4 Preparation of Metaphases

10µl of colcemid was added per 1ml of actively growing cell culture, incubated for 20 minutes, then following trypsinisation transferred to a 15ml Falcon tube and centrifuged at 1500 rpm for 7 minutes. The supernatant was removed and the pellet carefully resuspended in a few drops of KCl (75mM, 0.560g/100ml). KCl was then added slowly to a volume of 8 ml, gently mixed and incubated at 37°C for 10 minutes. Several drops of methanol fix (3:1 methanol:acetic acid) were added gently, mixed and then made up to a volume of 10ml. It was then centrifuged at 1500g (1500 rpm) for 7 minutes, the supernatant removed and the pellet resuspended in 10ml of fixative for 30 minutes at 4°C. The cells were then centrifuged at 1500g (1500rpm) for 7 minutes and

washed 3 times in fixative. Three drops of fixative were placed on a microscope slide, followed by 2 drops of the metaphase solution, held vertically and presence of metaphases confirmed under a microscope.

2.4.5 Slide preparation

Metaphase slides were placed on 80°C hot plate for 1 hr and then incubated in 0.005% pepsin/0.01M HCl at 37C for 30 mins. The slides were then placed in 1xPBS at room temperature for 5 mins, washed 3 times in an ethanol series for 5 mins (70%, 90% and 100%) and then air dried.

2.4.6 Hybridisation

10µl (20-100ng) of labelled DNA and marker was applied to the slide a cover slip located over it and sealed with rubber cement. The slide was placed on an 80°C heat block for 3 mins. to denature the DNA and then quickly transferred to 37°C.

2.4.7 Post Hybridisation Washing

Each slide was washed 3 times in 0.1xSSC at 60°C for 5mins, counterstained with Vectar Shield containing DAPI and the cover slip sealed with nail varnish. The slides were then viewed under a microscope and images captured

2.5 General laboratory solutions, media and buffers

Aminolinking buffer: 25ml of KCl (1M), 1.25ml $MgCl_2$ (1M), 5ml Tris (0.5M) pH8.5.
Chloramphenicol: 0.100g in 5ml of 100% ethanol.
DOP-PCR buffer: 2ml MgCl2 (1M), 10ml Tris HCl (1M) pH 8.4, 50ml KCl, (1M),10ml gelatin (10mg/ml) and 28ml ddH2O. Autoclaved
Luria Bertani (LB) broth preparation: 10g bacto-tryptone, 10g NaCl, 5g bacto-yeast extract, ddH20 added to 1 litre. Sterilised immediately using autoclave, antibiotic added when cooled.
LB agar preparation: 10g bacto-tryptone, 10g NaCl, 5g bacto-yeast extract, 15g agar, ddH20 added to 1 litre. Sterilised immediately using autoclave, antibiotic added when

cooled. Agar poured into plates, flask opening was continually sterilising with Bunsen flame

Microarray hybridisation buffer: 50% formamide, 10% dextran sulphate, 0.1% Tween 20, 2x SSC, 10mM Tris pH 7.4.

4x Microarray spotting buffer: 1M sodium phosphate buffer, pH8.5, 0.001 sarkosyl.

Proteinase k:100mg of proteinase k was dissolved in 10ml of H_2O (stock 10mg/ml, stored at -20°C

Random prime labelling dNTP mix: 0.5 mM dCTP, 2 mM dATP, 2mM dGTP, 2mM dTTP in TE buffer

Saturated NaCl: 35g of NaCl was added to 100ml of H_2O and vigorously agitated.

Solution I: 4.504g glucose (50mM final conc.), 10ml 0.5M EDTA (10mM final conc.), 12.5ml 1M Tris pH8.0 (5mM final conc.). Made up to 500ml with d H_2O, filter sterilized. (Stored at 4°C)

Solution II: 8.6ml H2O, 400 µl 5M NaOH (2N final conc.), 1ml 10% SDS (2N final conc.)

Solution III: 3M KOaC pH 5.5. (Stored at 4°C)

SSC: 175.3g of NaCl and 88.2g of sodium citrate was added to 800ml of H_2O, pH was adjusted to 7.0 with NaOH/ HCl. Volume was made up to 1 litre.

SDS: 100g of SDS was dissolved in 900ml of H_2O. Heated to 68°C. pH was adjusted to 7.2 with NaOH/ HCl. Volume was made up to 1 litre.

Tail buffer: 1ml of 5M NaCl, 2.5ml of 1M Tris-HCl (pH 8), 10ml of 0.5M EDTA (pH 8), 5ml of 10% SDS and 31.5ml of H_2O added to make a total volume of 50ml.Tris: 121.1g of Tris was dissolved in 800ml of H_2O. Adjusted to required pH with NaOH/ HCl. Volume was made up to 1 litre

EDTA 186.1q of disodium ethylenediaminetetra acetate was added to 500ml of H_2O. pH was adjusted to 8.0 with a few drops of NaOH

TE: 1ml of 1M Tris-HCl (pH 7.5), 0.2ml 0.5M EDTA (pH 8) and 98.8ml of sterile H_2O was added to make a volume of 100ml.

CHAPTER 3: RESULTS

It is clear that the most commonly deleted chromosome in CRC carcinogenesis is chromosome 18 but conventional techniques have failed to delineate a small region that is likely to contain a key tumour suppressor gene. A high resolution technique to investigate the whole of chromosome 18 could be provided by microarray CGH.

Commercially produced slides were available but too expensive to allow analysis of a meaningful number of CRC samples. We therefore produced our own specifically designed array featuring 3 components, a tiling path of overlapping DNA clones to provide complete coverage of chromosome 18, genome wide positive controls and a smaller number of negative controls.

We ensured the accuracy of our technique in both production of the microarray slide and CGH technique. All PCR amplified DNA for the microarray was confirmed free of contamination and print quality regularly assessed. The CGH technique was verified by analysing normal versus normal DNA and a CRC cell line with a known amplification and large deletion.

Having validated our technique we proceeded initially to analyse 47 CRC cell lines.

3.1 CONSTRUCTION OF MICROARRAY

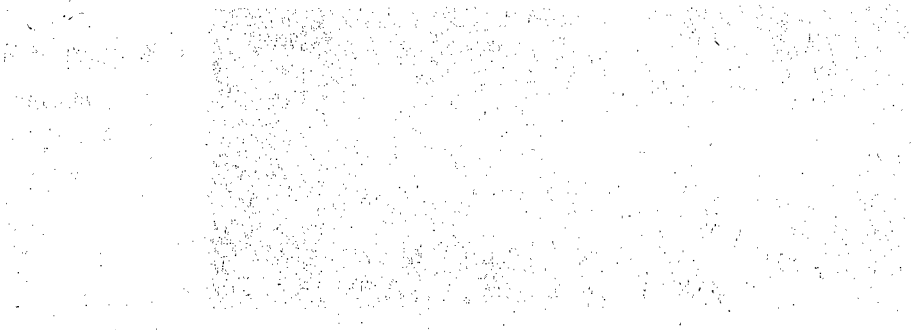
The microarray was designed to target chromosome 18 at the highest resolution with a complete "tiling path" of 860 overlapping BAC clones to represent this region. To normalise the array CGH intensity ratios from the chromosome specific tiling path; low-resolution genome wide positive controls, spaced at 5 Mb intervals, were incorporated in the array. Negative array controls in the form of Drosophila DNA were also included to assess background signal from non-specific probe binding.

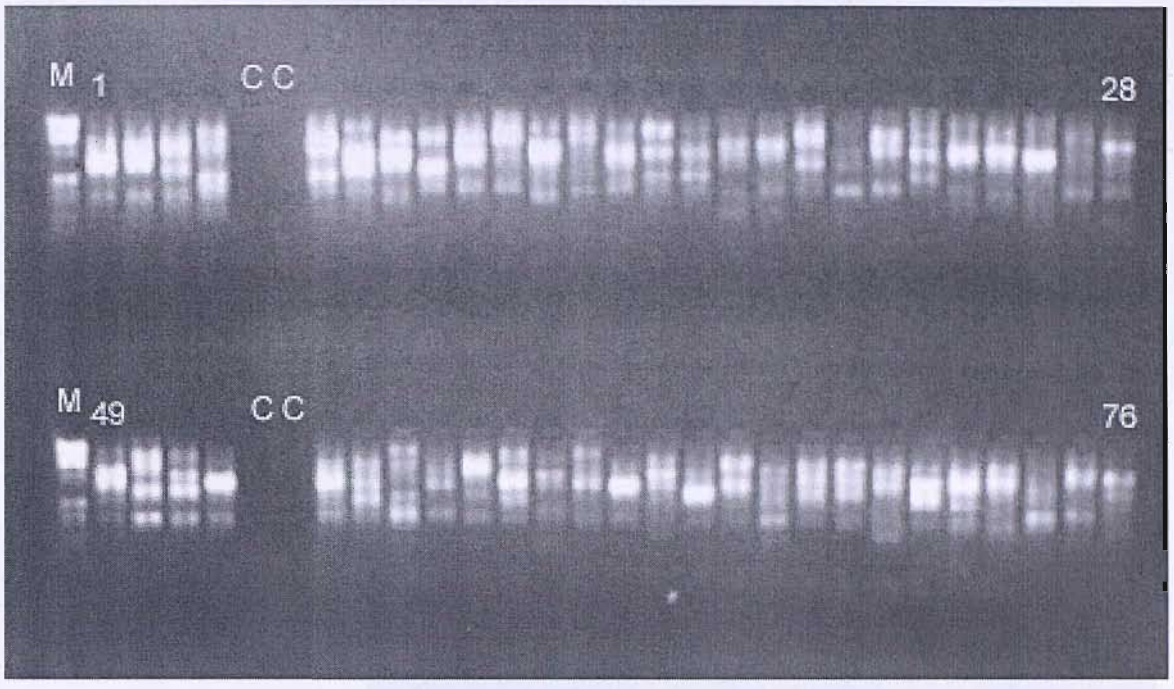
Extracted DNA from the 860 BAC clones representing the chromosome 18 tiling path were purchased in 10 96 well plates from the Children's Hospital Oakland Research Institute (CHORI), California. Their 32,000 genome wide clone set is available commercially as BAC or PAC clones or the corresponding extracted DNA, from the CHORI BAC and PAC Resource Centre (BPRC), (www.bacpac.chori.org/libraries). All the extracted genomic BAC/PAC DNA for the tiling path array was amplified

separately in triplicate by 3 different DOP-PCR primers, DOP 1, DOP 2 and DOP 3 (section 2.1.5) (Fig 3.7).

The 600 genome wide BAC/PAC positive array control clones were selected from the Sanger Centre 1Mb clone set (www.ensembl.co.uk). DNA extraction (section 2.1.4). and DOP-PCR of the positive controls was performed by Noel Wortham at the Molecular and Population Genetics Laboratory, Cancer Research UK, Lincoln's Inn Fields.

It was important to confirm successful DOP-PCR amplification of each clone by all three DOP-PCR primers without contamination. Negative control wells were incorporated in the DOP-PCR plates with BAC DNA excluded from the DOP-PCR reaction mix. The DOP products from each of the DOP1, DOP 2 and DOP 3, 96 well plates were loaded onto 3 electrophoresis gels (100 lanes), with markers incorporated (1 kb ladder) to confirm product size (section 2.1.6)(Fig. 3.1).





M: Marker

C: negative control

Loading method, see section 2.1.6

Correct DOP 1 products and 8 clean negative controls confirmed.

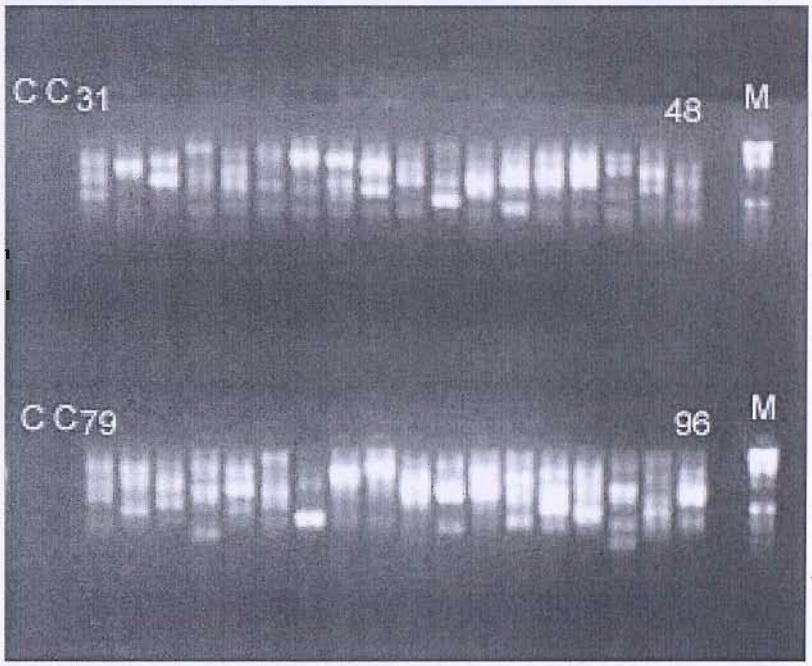


Fig 3.1: Electrophoresis gel of DOP1 PCR products from plate Ch 18-RP1-A1 captured in 2 images.

The three DOP-PCR products from each BAC clone were combined. Amino PCR was carried out (Fig. 3.7) to incorporate an amino group to the 5' terminus of the DOP-PCR product to facilitate attachment to the amine-binding glass slides (section 2.1.7). The Amino-PCR product was confirmed and contamination excluded on an electrophoresis gel incorporating markers (1 kb ladder) and negative controls (section 2.1.8) (Fig. 3.2).

M: Marker

C: negative control

Loading method, see section 2.1.8

Correct amino products and 5 clean negative controls confirmed.

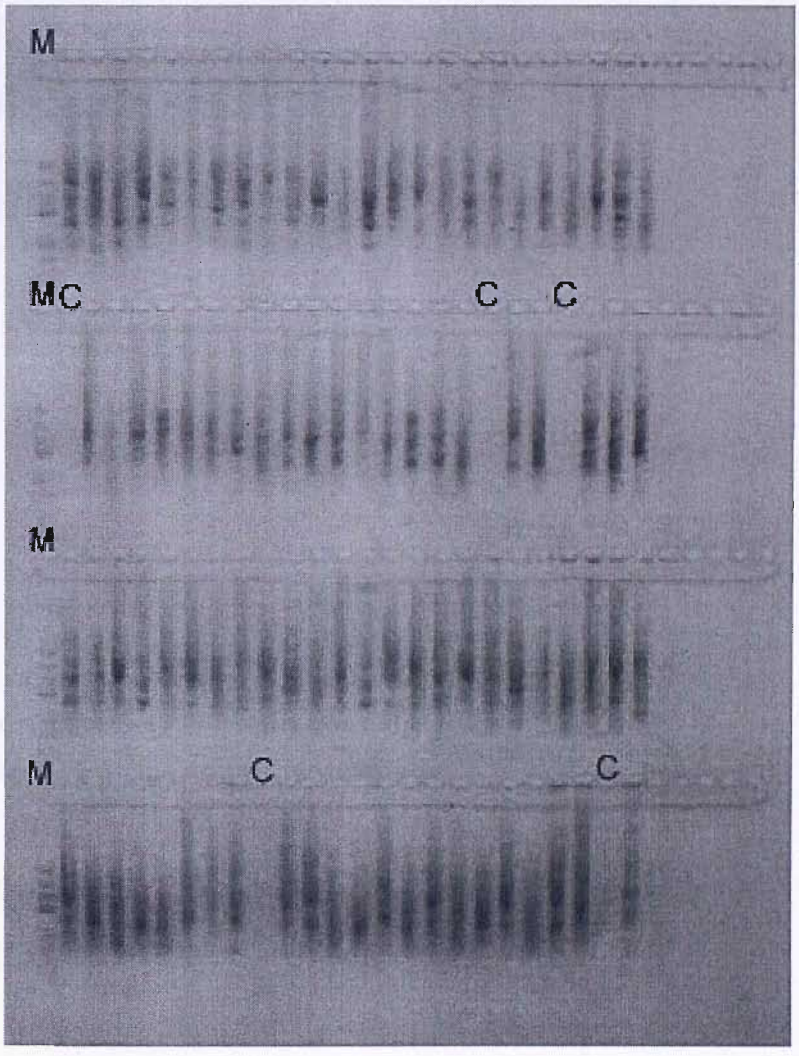


Fig 3.2: Electrophoresis gel of amino PCR products from plate Ch 18-RP1-A2.

For the array negative controls, DOP/amino PCR products of 8 Drosophila clones were provided by the Molecular and Population Genetics Laboratory, Cancer Research UK, Lincoln's Inn Fields.

The 96 well plates containing DOP/amino PCR products for the 860 BAC clones constituting the chromosome 18 tiling path array, the 600 positive and 8 negative array control clones were transferred to 5 384 well plates.

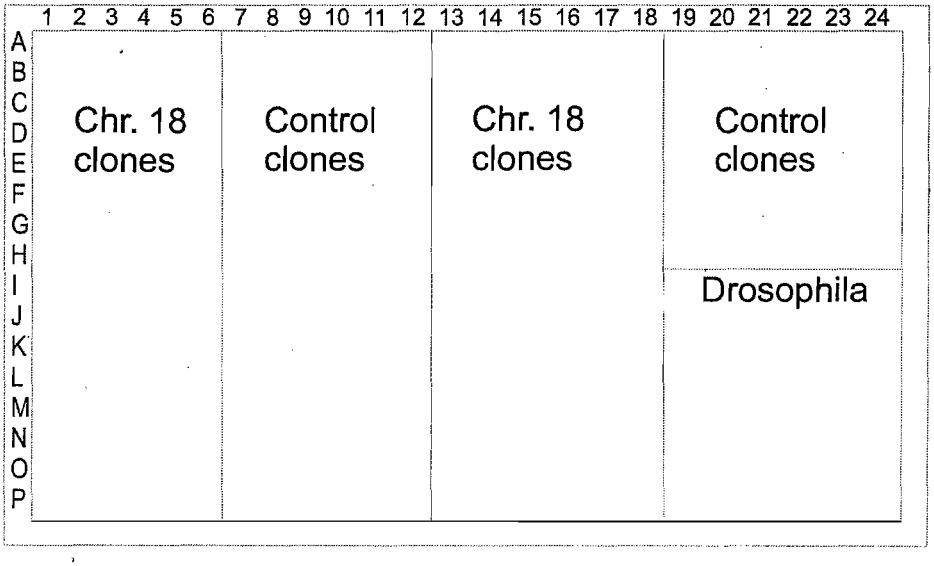


Fig. 3.3: Arrangement of chromosome 18 and array control clones in the 384 well plates loaded into the arrayer for printing.

The amino-linked product for each clone (15μl) was combined with 15μl of 4x microarray spotting buffer (15μl) in each well (section 2.1.9).

An excel spread sheet was generated identifying the content of each well of the 5 384 well plates by clone name as ID, this was imported as a text file into the Q array print program.

The genetix arrayer print head was set up with sixteen pins in a 4x4 configuration to print 16x16 spot blocks; the clones were printed in a random duplicate pattern onto coated amine-binding slides (Amersham, codelink).

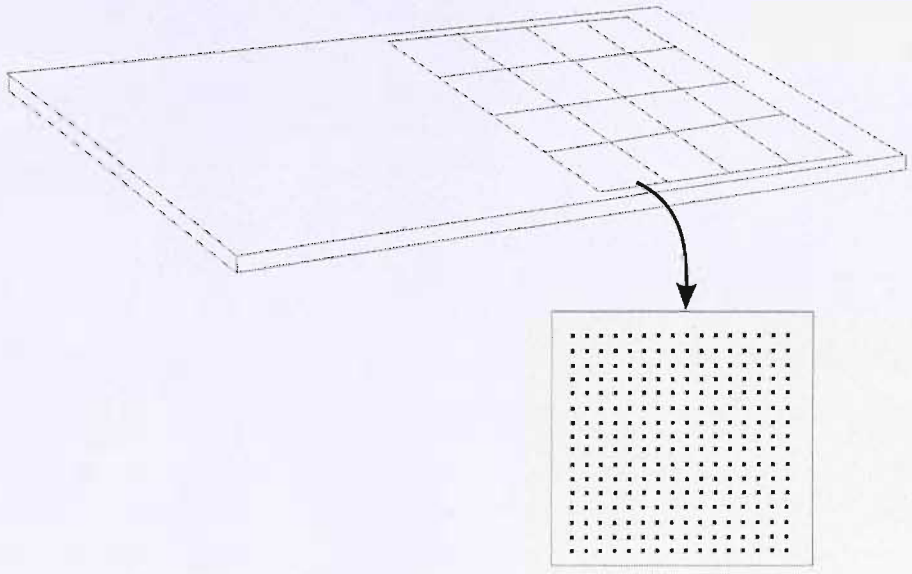


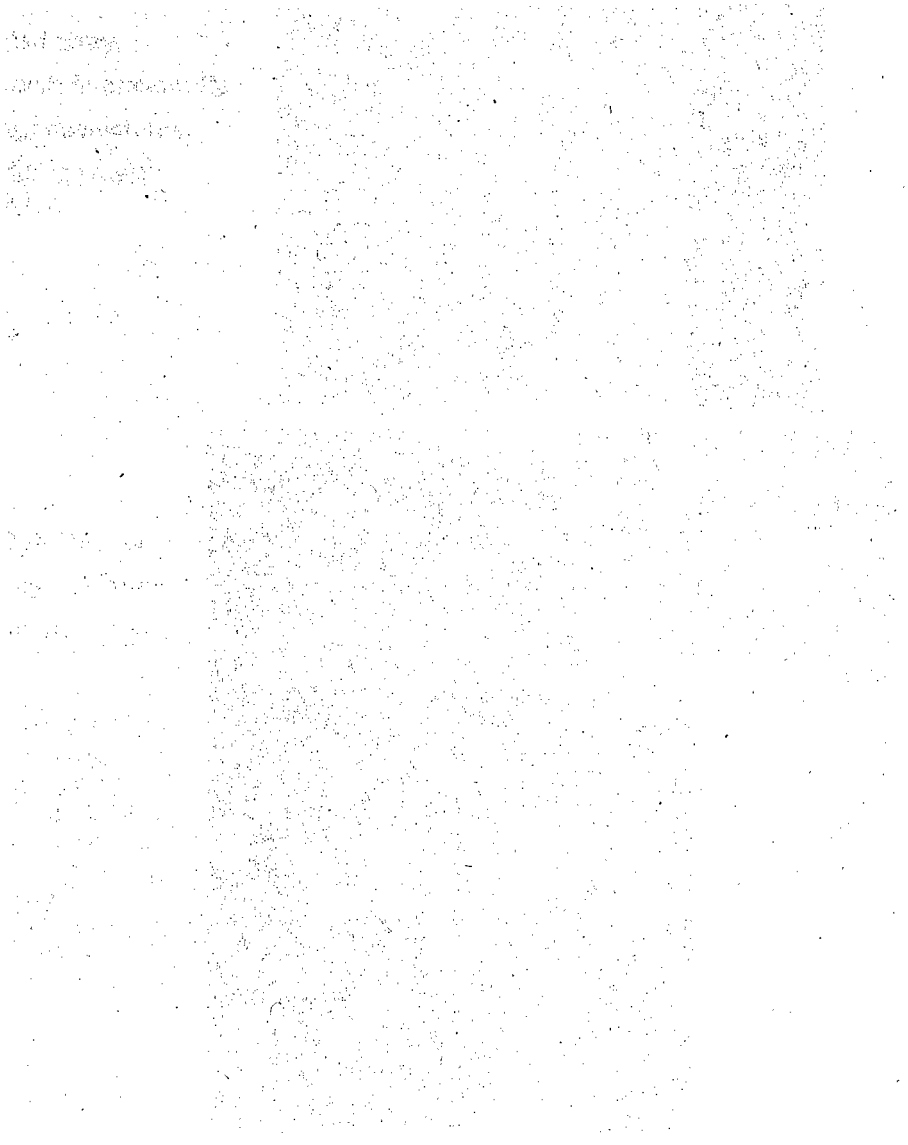
Fig. 3.4: Representation of a glass microscope slide with a 4x4 block configuration printed on it.

One of the blocks is enlarged to show the 16x16 spot pattern printed within each block

The slides were stored at ambient temperature in a slide box.

The microarray slides were prepared for comparative genomic hybridisation by blocking any residual reactive groups, denaturing, washing and drying (section 2.1.10).

On completion of the print run the original data of clone well position had been translated by the Q array printing software to account for randomisation and duplication, providing clone position within the printed array as a Gal file. The Gal file provided clone identification in the post-hybridisation data analysis. The chromosomal position of each clone was subsequently imported during post-hybridisation data analysis using the Gal file data for clone identification, allowing the most up to date "build" of genomic clone position to be utilised.



3.2 ASSESSMENT OF PRINTED SPOT QUALITY

The quality of printed arrays was assessed using a mixture of labelled oligonucleotides to visualise the morphology, alignment and completeness of printing. The acceptable standard required complete circular spots of equal size, symmetrical in morphology and alignment, with low levels of background fluorescence and artefact, an example is shown in Fig. 3.5a and 3.5b (section 2.1.11).(Gaps without printed spots is expected, a consequence of wells deliberately left empty and the random print pattern)

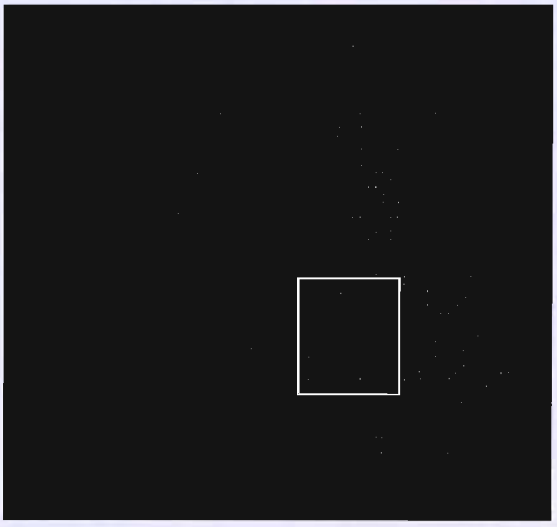


Fig 3.5a: Outline view of whole printed array visualised with fluorescently labelled oligonucleotides.

White box defines region viewed in detail

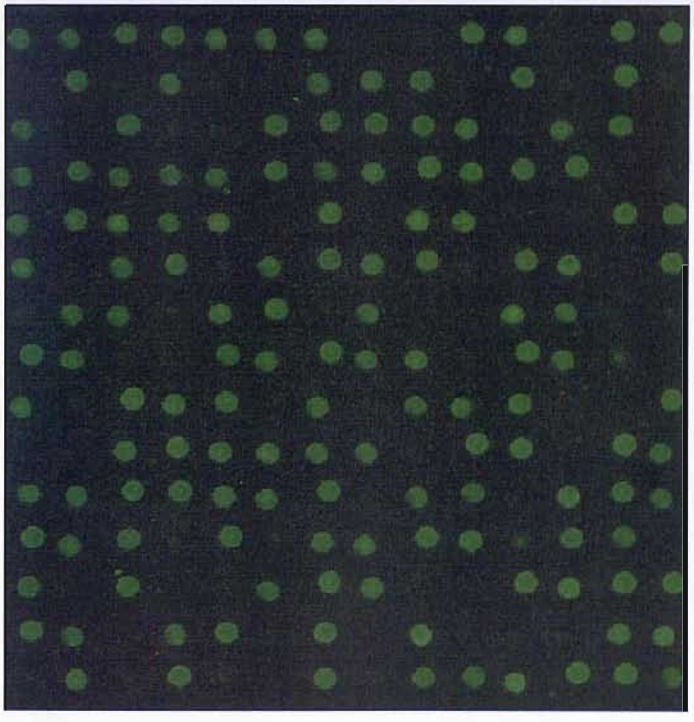


Fig 3.5b: Detailed view of sub-region of array of acceptable print quality

This is an example of an acceptable quality of spot printing. Complete symmetrical spots and low level background artefact. Slides from this 'print run' were utilised for array-CGH analysis and data collection in this thesis

Unacceptable printing quality with incomplete asymmetric spots and background artefact can be seen in Fig. 3.6

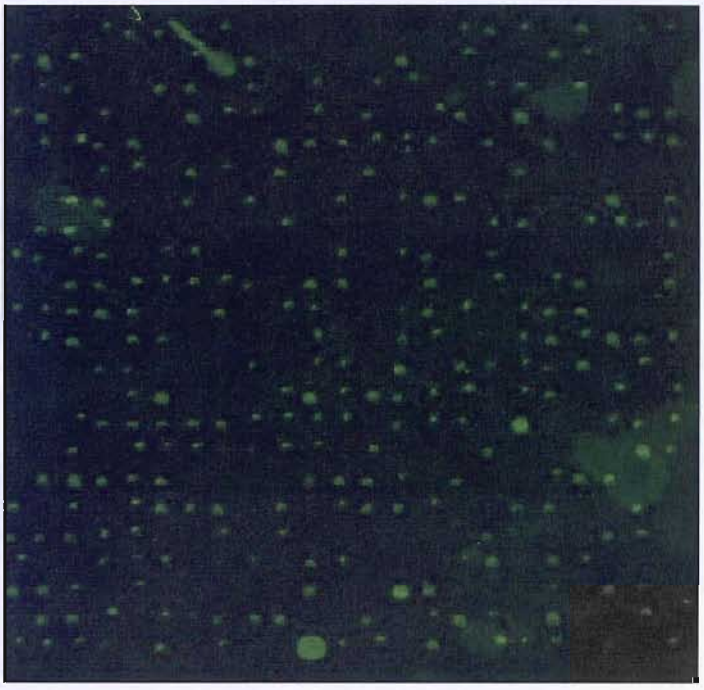


Fig 3.6: Detailed view of unacceptable print quality.

This is an example were the print quality was obviously not acceptable for use in array CGH analysis.

There are incomplete small asymmetric spots, with background artefact

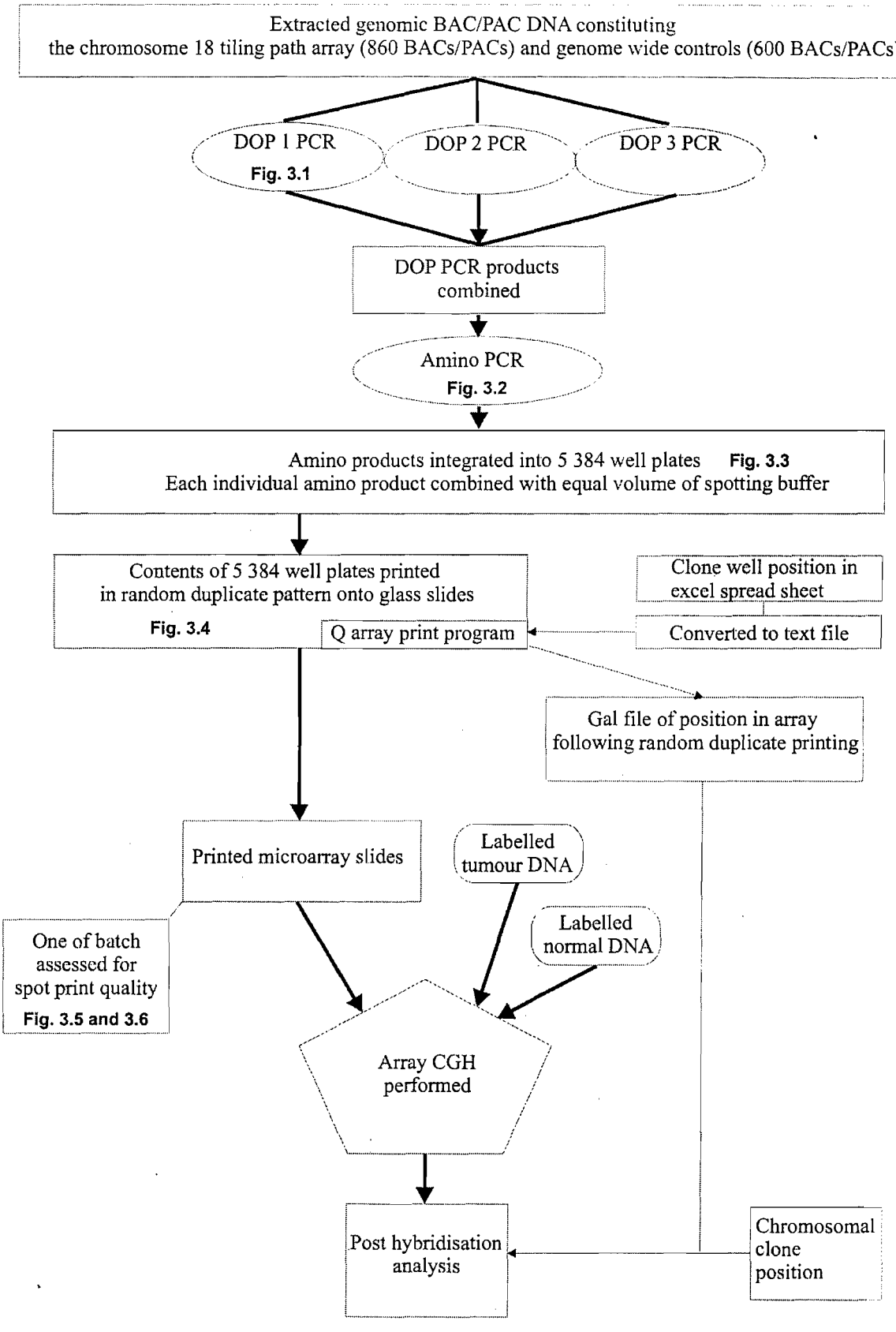


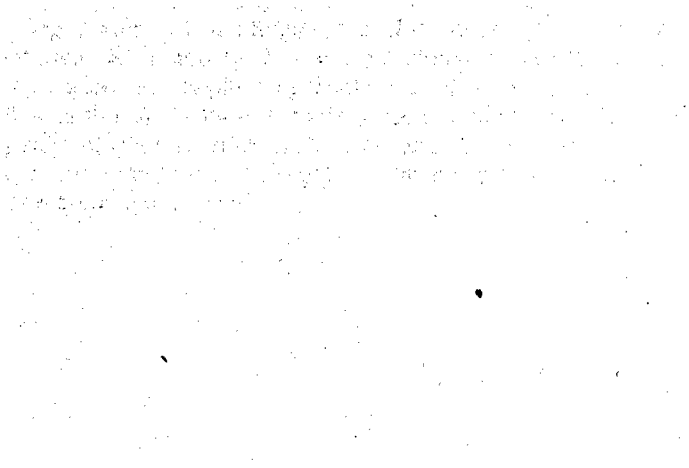
Fig 3.7: Flow chart summarising array construction, comparative genomic hybridisation and processing of clone positional data.

3.3 COMPARATIVE GENOMIC HYBRIDISATION TO THE CONSTRUCTED CHROMOSOME 18 ARRAY

Tumour and normal DNA were labelled with different fluorescent dyes (Cy3, Cy5) and then purified to remove the non-incorporated dyes (section 2.3.1 and 2.3.2).

A well was created around the array field on the glass slide with rubber cement. A prehybridisation mixture containing Cot1 DNA was evenly spread over the array to block highly repetitive DNA sequences from involvement in the co hybridisation. The pre-hybridisation mixture was replaced by the hybridisation mixture of labelled DNAs and incubated for 48hrs (section 2.3.3 to 2.3.6). The rubber cement was then removed, the slide washed and dried (section 2.3.7).

The slide was then laser scanned and an image for each Cy3 and Cy5 dye captured as a separate TIFF image. The spot intensities were then analysed, rejected on the basis of quality and normalised by dividing each raw Cy3/Cy5 intensity ratio by the median intensity ratio for each block. Duplicate spots were excluded if they differed by >3 standard deviations from mean. The log2 ratio was calculated and the mean value of duplicate spots plotted according to genome position (section 2.3.8).



3.3.1 Comparative hybridisation of normal versus normal DNA

The inherent error of the constructed array and CGH technique was assessed by co-hybridising normal against normal DNA (10 pooled normal DNAs) (Fig. 3.8 and 3.9).

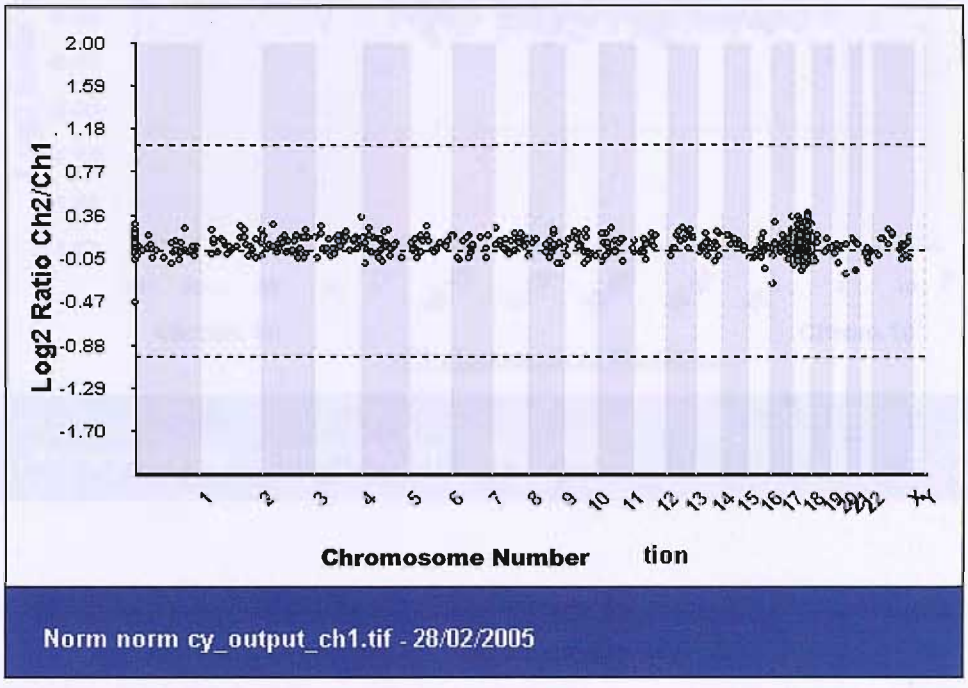


Fig 3.8: Genome wide plot of log2 intensity ratio for each clone, from a normal versus normal array comparative genomic hybridisation, against genome location.

This is a standard graphical output for the analysed results of array comparative genomic hybridisation. The Y axis is the intensity ratio for the test/normal expressed as a log2 ratio, therefore a 1/1 ratio is plotted at 0, a perfect single allelic deletion a 0.5/1 ratio would be plotted at -1 and a perfect single amplification a 3/2 ratio plotted at +0.6. The X-axis is the genome position of the BAC clone. A perfect normal versus normal hybridisation would result in an intensity ratio of 0 for each clone. The spread of the plotted intensity ratios, ranges from +0.4 and -0.2 in this standard hybridisation; outlining the range of error for the printed array and hybridisation technique.

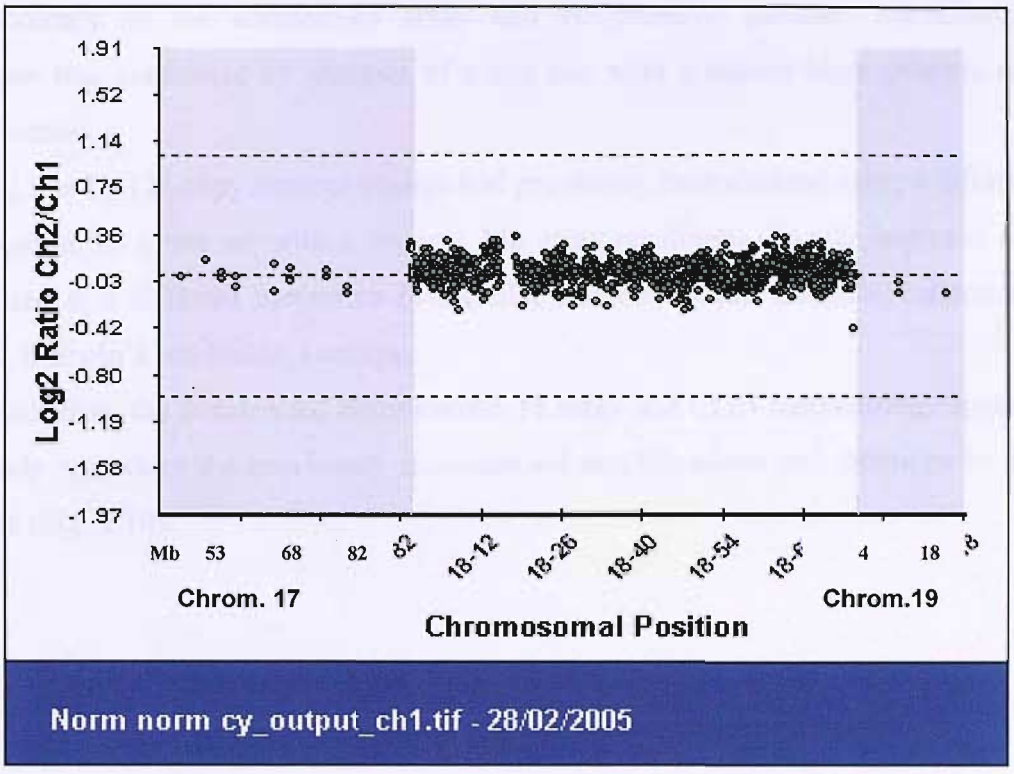


Fig 3.9: Chromosome 18 plot of log2 intensity ratio for each clone, from a normal versus normal hybridisation, against chromosome 18 location.

This shows the dense overlapping chromosome 18 tiling path array with a break in continuity corresponding to the centromere. The chromosome 18 tiling path array is bordered by the lower resolution clones of a small section of the adjacent chromosomes 17 and 19.
The spread of the plotted intensity ratios for the chromosome 18 clones ranges from +0.3 to -0.2

3.3.2 Analysis of cell lines with known copy number change

The accuracy of the constructed array and comparative genomic hybridisation technique was confirmed by analysis of a cell line with a known large deletion and amplification.

The cell line Ls 123 copy number change had previously been defined using a different chromosome 18 clone set with a lower 1 Mb array resolution (Sanger Institute) and hybridised at a different institution (Molecular and Population Genetics Laboratory, CRUK, Lincoln's Inn Fields, London).

The result from the constructed chromosome 18 array and CGH methodology appears to closely reproduce the previously characterised amplifications and deletions in this cell line (Fig. 3.10).

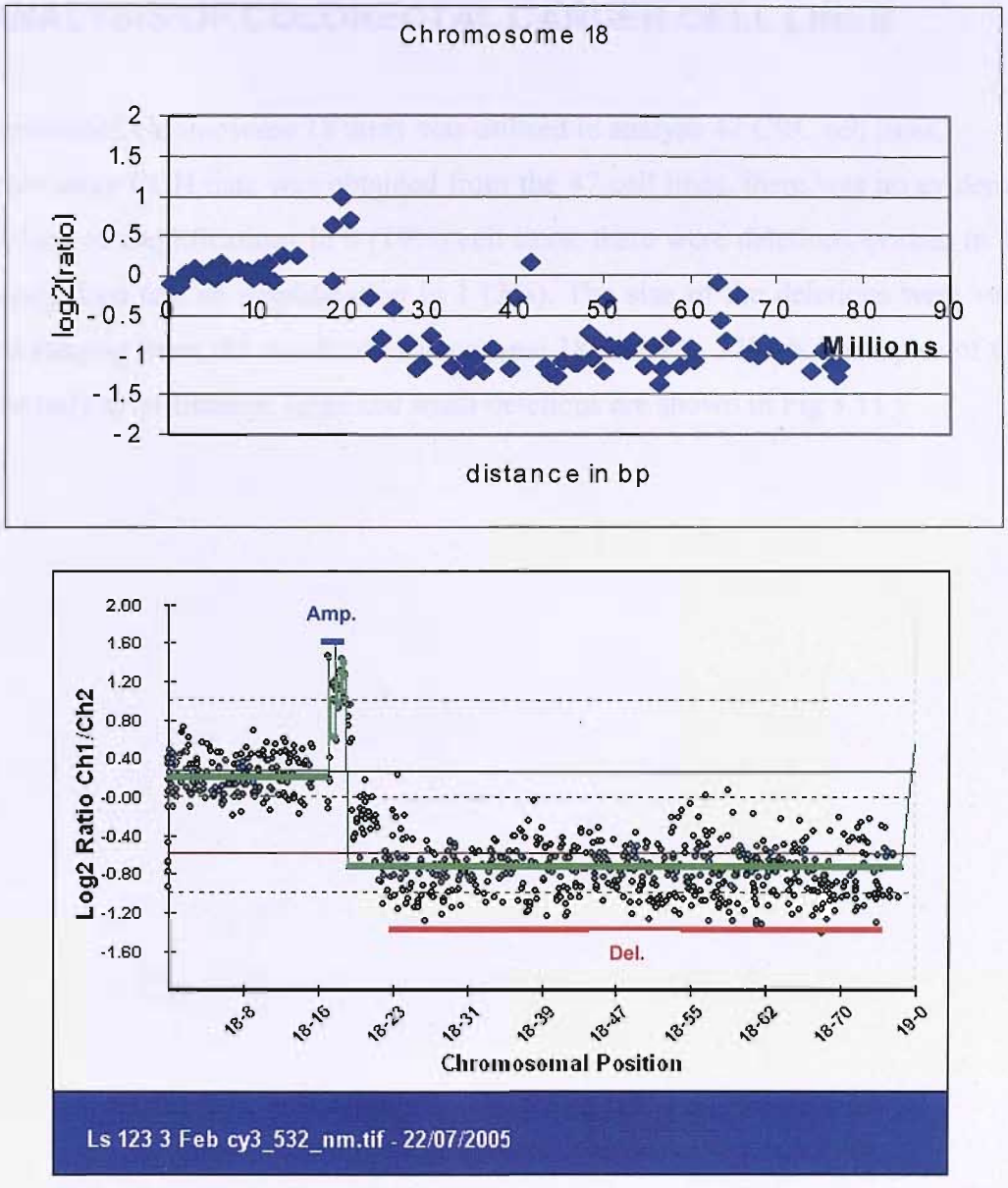
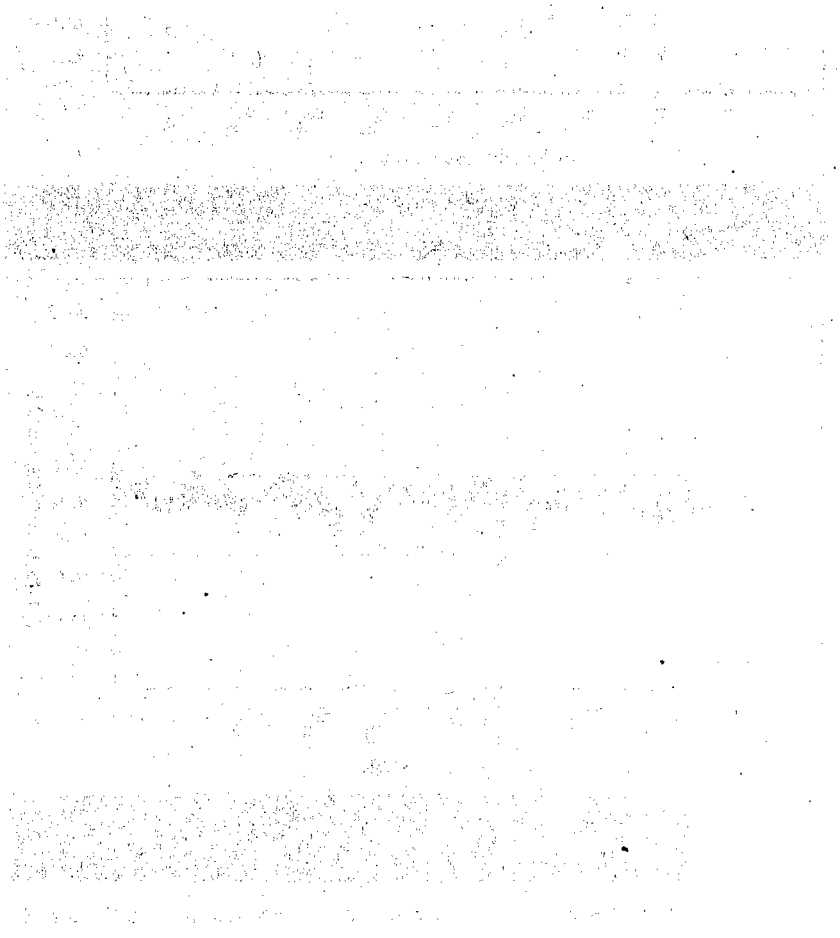


Fig 3.10: Parallel plots of array CGH analysis of chromosome 18 in the Ls 123 cell line. Top graph (a): Log ratio of intensity defined by a 1Mb array. Bottom graph (b): Log ratio of intensity defined by the chromosome 18 tiling path array.

The high-density chromosome 18 tiling path array intensity ratios (b) similar small region of amplification at 18-16 to 18-18 and the large deletion from 18-21 to the telomere detected by the lower resolution 1Mb Sangar clone set (a).
The light green line of "best fit" is generated by the Bluefuse software package, it represents the mean intensity ratio value for consecutive sub-sections of clones, only deviating when an important consecutive segment of clones "step away" from the mean. A thin horizontal red line has been generated by the array CGH smooth statistical software package utilising standard log parameters, an intensity ratio plotted below this threshold line (at a log2 intensity ratio of -0.667) represents a deletion.

3.4 ANALYSIS OF COLORECTAL CANCER CELL LINES

The constructed chromosome 18 array was utilised to analyse 47 CRC cell lines. The microarray CGH data was obtained from the 47 cell lines, there was no evidence of deletions or amplifications in 9 (19%) cell lines, there were deletions evident in 38 (81%) cell lines and an amplification in 1 (2%). The size of the deletions were very variable ranging from the whole of chromosome 18 down to 371 kb. Examples of the data, the only amplification, large and small deletions are shown in Fig 3.11



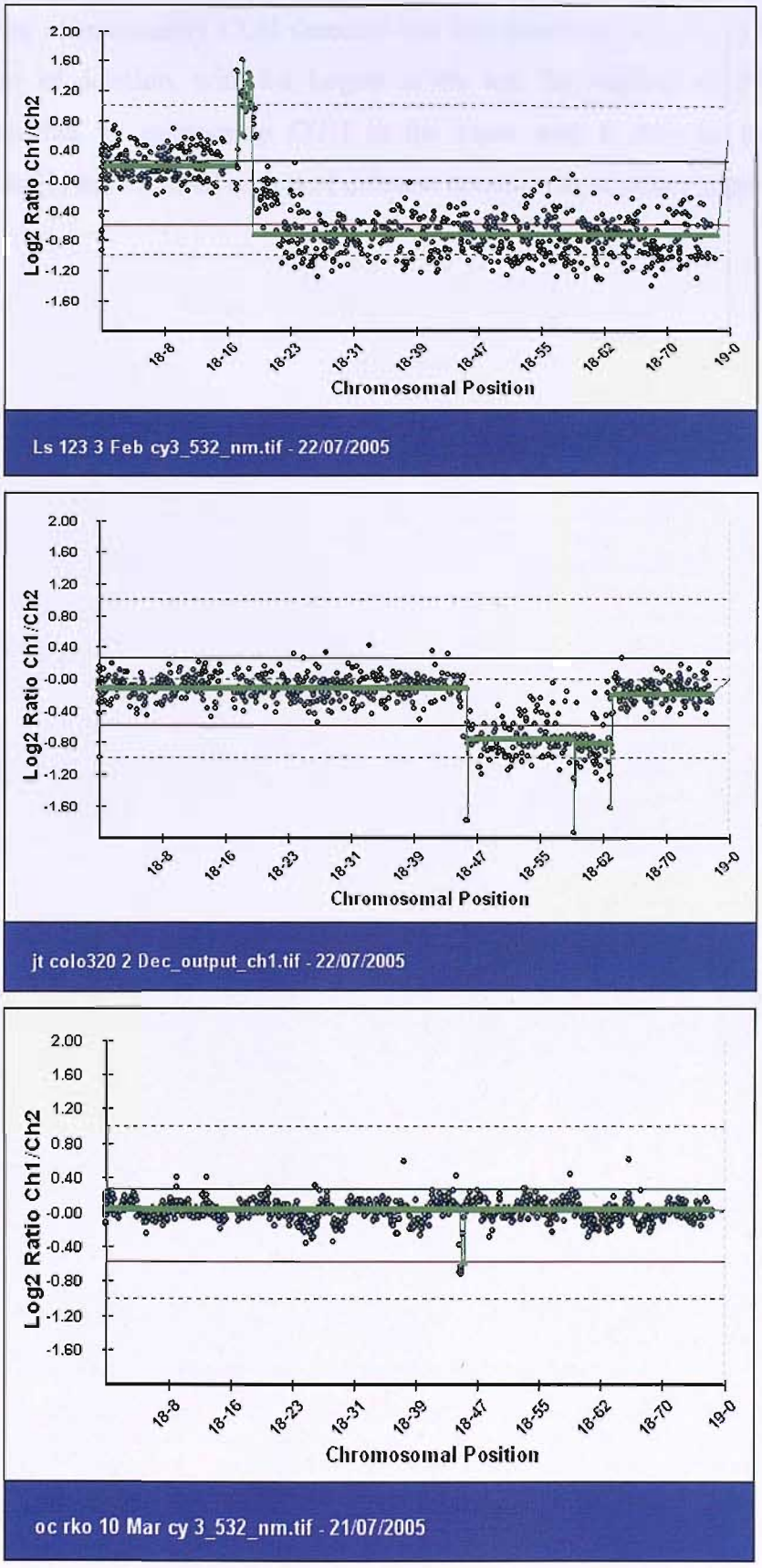
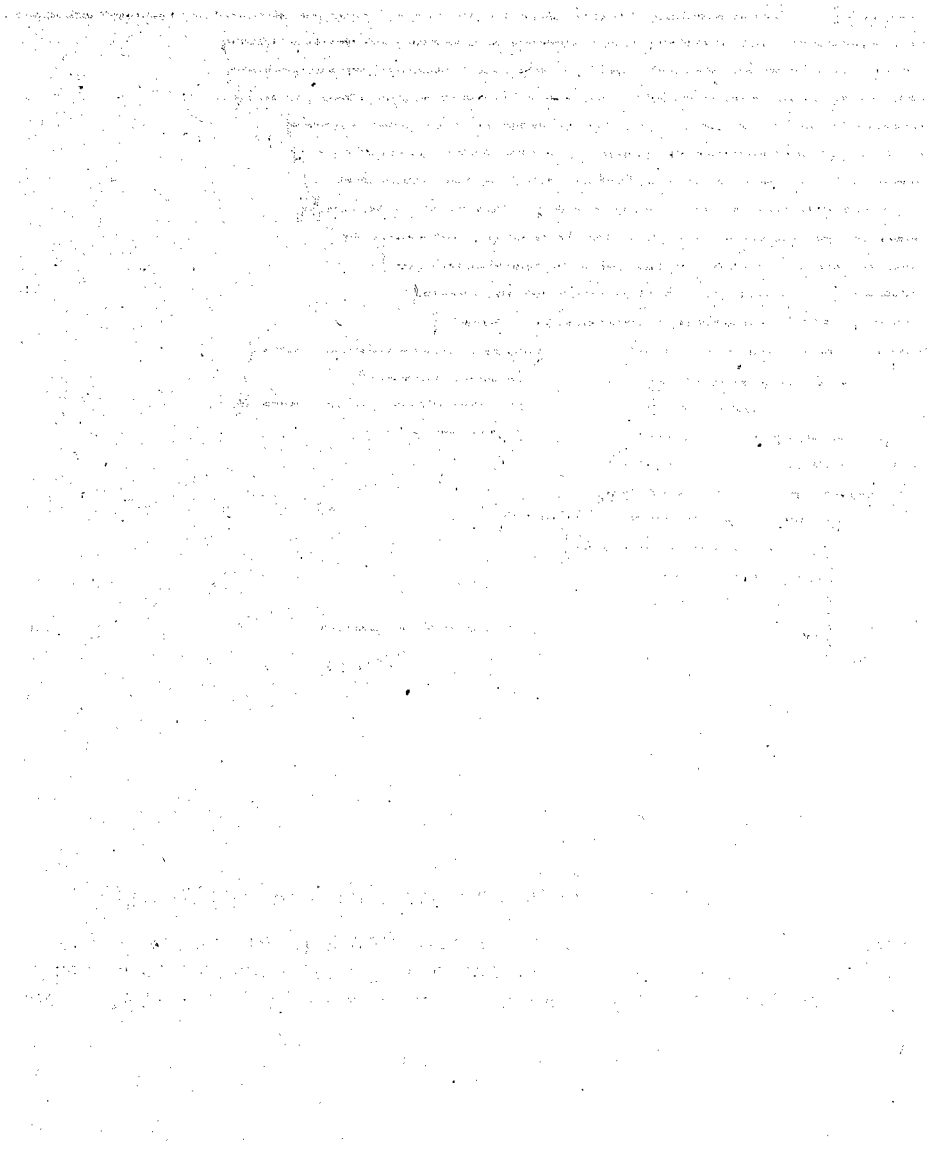


Fig 3.11 Examples of size range of identified cell line deletions

The 3 examples of microarray CGH detected cell line deletions (Fig. 3.11) have been ranked by size of deletion, with the largest at the top. By ranking all 38 cell line deletions identified by microarray CGH in the same way it may be possible to determine which is the minimal region of deletion common to all other larger deletions (Fig. 3.12).



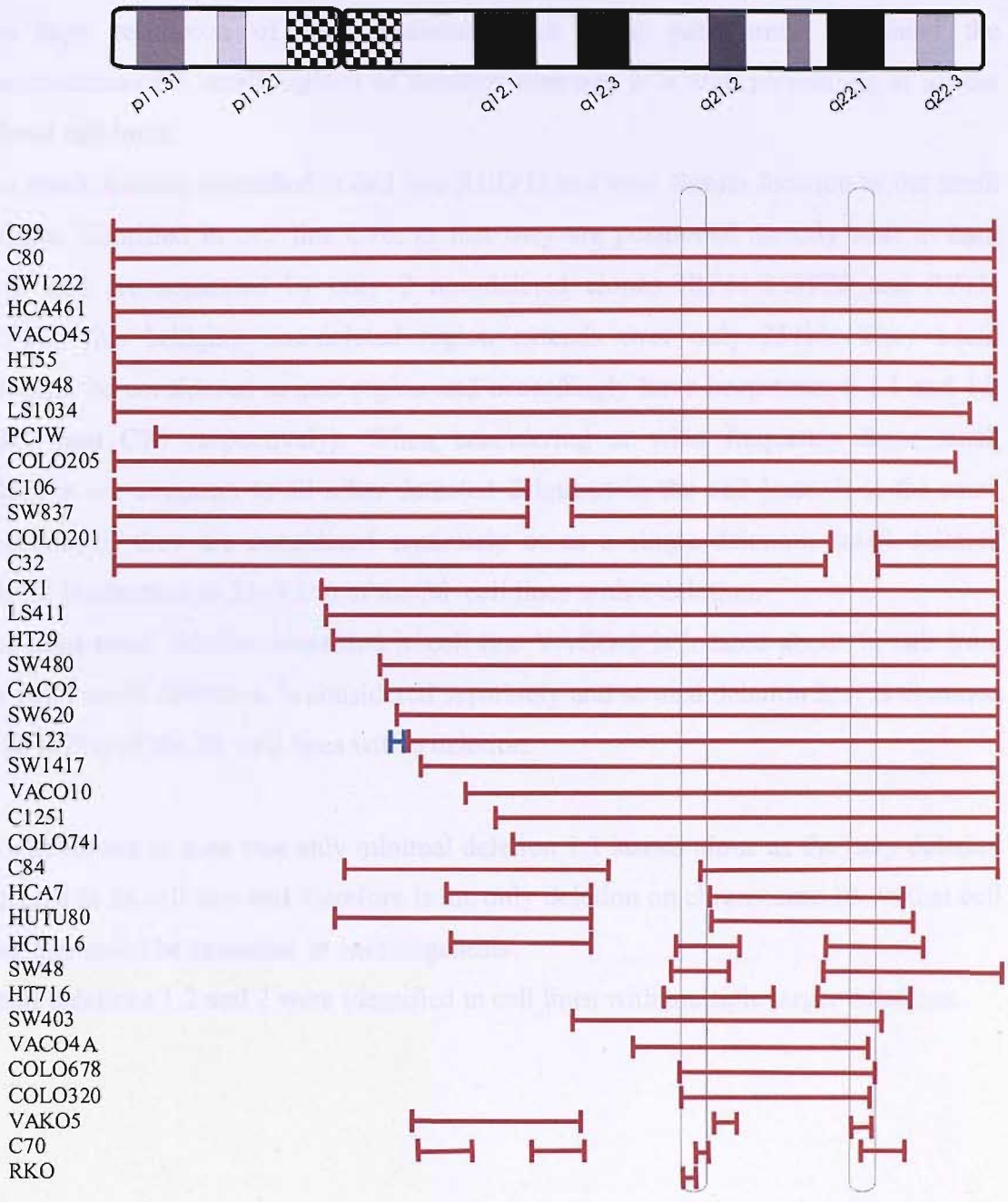


Fig 3.12: Ranked chromosome 18 cell line deletions

Cell line deletions are represented by a horizontal red bar and amplifications by a blue bar, below an ideogram of chromosome 18. The deletions are ranked with the largest at the top and the smallest at the bottom. The 2 vertical bars represent common regions of deletion.

The high resolution of the chromosome 18 tiling path array facilitated the identification of 3 small regions of deletion common to a high percentage of all the deleted cell lines.

The small deletion identified in cell line RKO is in a very similar location to the small deletion identified in cell line C70, in fact they are positioned directly next to each other and are separated by only 2 non-deleted clones (RP11-813F20 and RP11-879D5), this bridging non-deleted region extends over only 251kb. They could therefore be considered as one region and accordingly have been termed 1.1 and 1.2 (RKO and C70 respectively). When considering at what frequency these small deletions are common to all other detected deletions in the cell lines, it is the same frequency if they are considered separately or as a single deletion; small deletion 1.1/1.2 is common to 35 (92%) of the 38 cell lines with a deletion.

The third small deletion identified in cell line VACO 5 is located about 12 Mb from the other small deletions, is considered separately and termed deletion 2, it is common to 35 (92%) of the 38 cell lines with a deletion.

It is important to note that only minimal deletion 1.1 stands alone as the only deletion detected in its cell line and therefore is the only deletion on chromosme 18, in that cell line, that could be causative in carcinogenesis.

Small deletions 1.2 and 2 were identified in cell lines with multiple larger deletions.

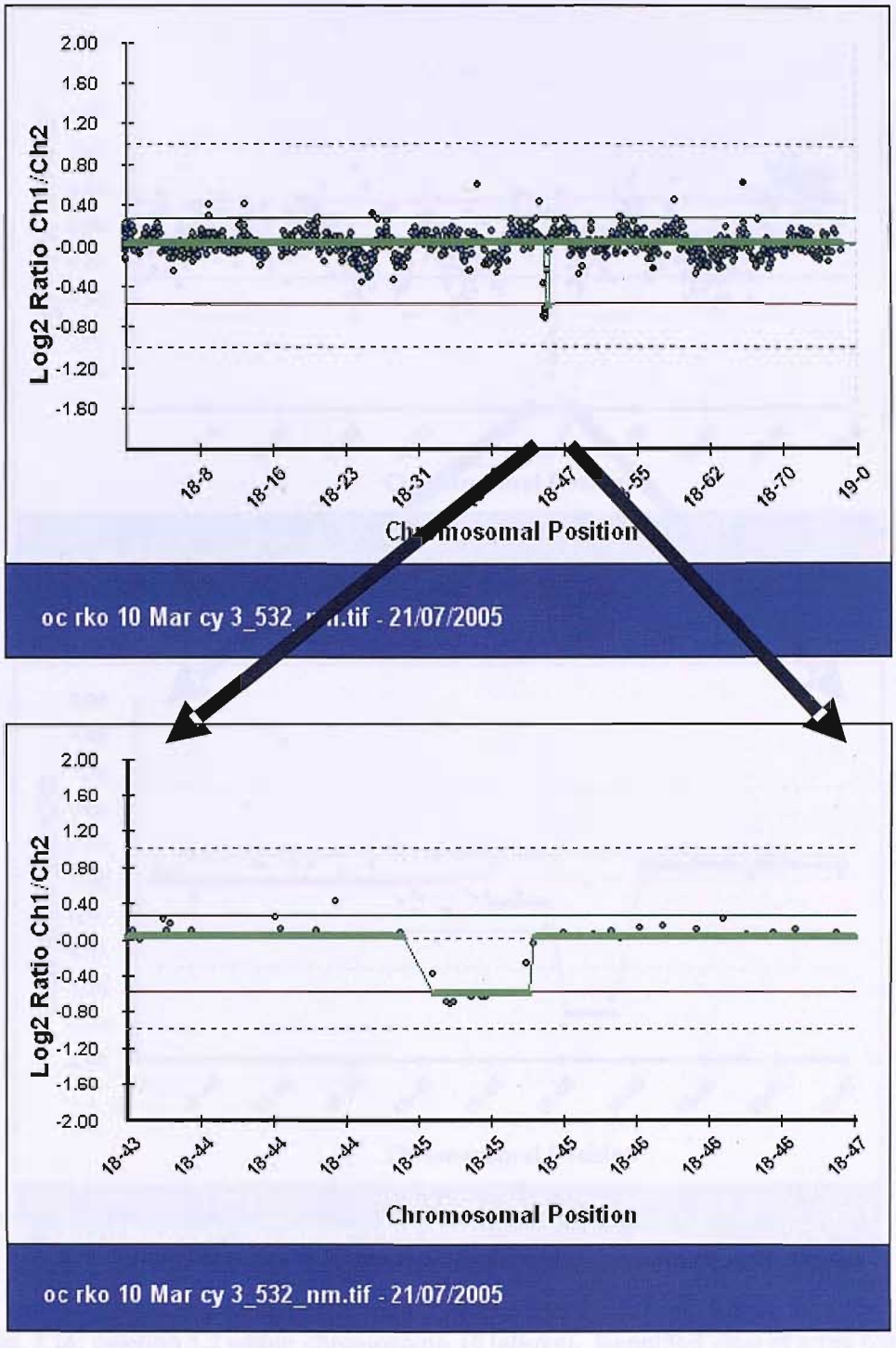


Fig. 3.13: deletion 1.1 within chromosome 18 (above). Magnified view of array CGH intensity ratios in a 4Mb sub-region, centred on deletion 1.1 (below).

Deletion 1.1 is identifiable as the only region of step down within this 4Mb sub region of chromosome 18.

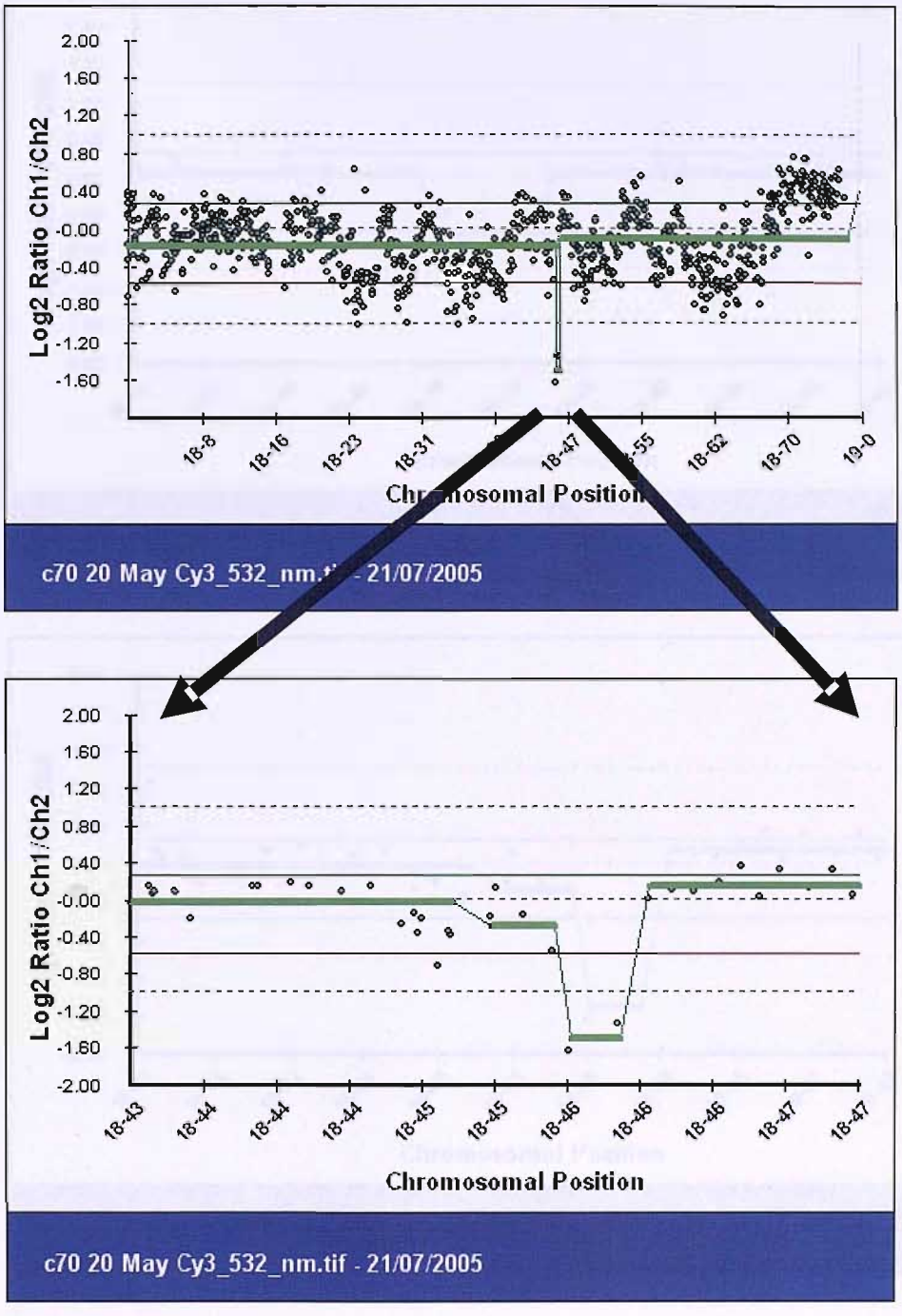


Fig. 3.14: deletion 1.2 within chromosome 18 (above). Magnified view of array CGH intensity ratios in a 4Mb sub-region, centred on deletion 1.2 (below).

There are multiple deletions. Deletion 1.2 is identified by the small number of consecutive clones located at 18-46 with intensity ratios of less than -1.

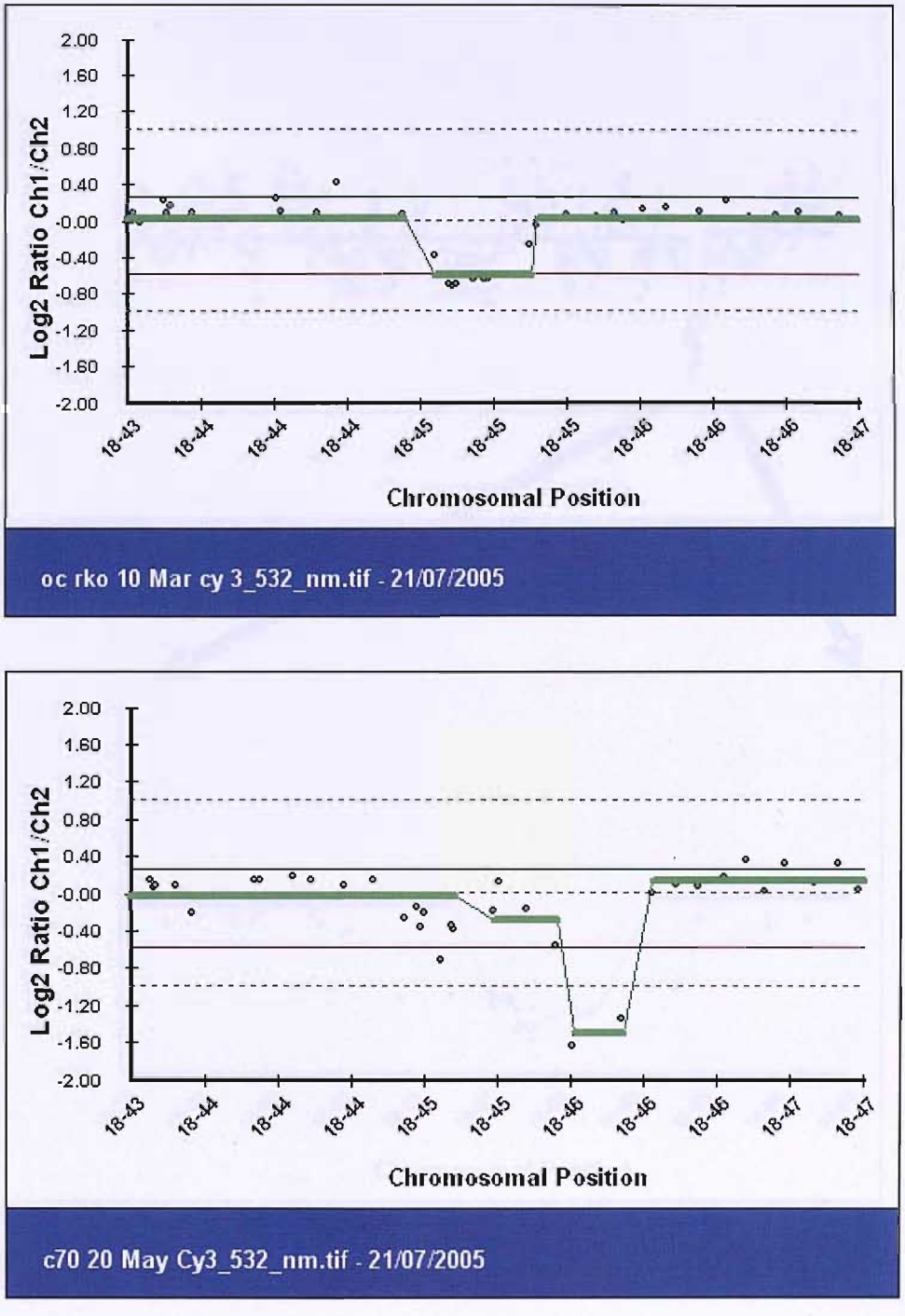


Fig. 3.15: Magnified views of array CGH intensity ratios in the same 4Mb sub-region, deletion 1.1 (above) and deletion 1.2 (below).

This illustrates that small deletions 1.1 and 1.2 are positioned directly next to each other and are separated by only 2 non-deleted clones (RP11-813F20 and RP11-879D5).

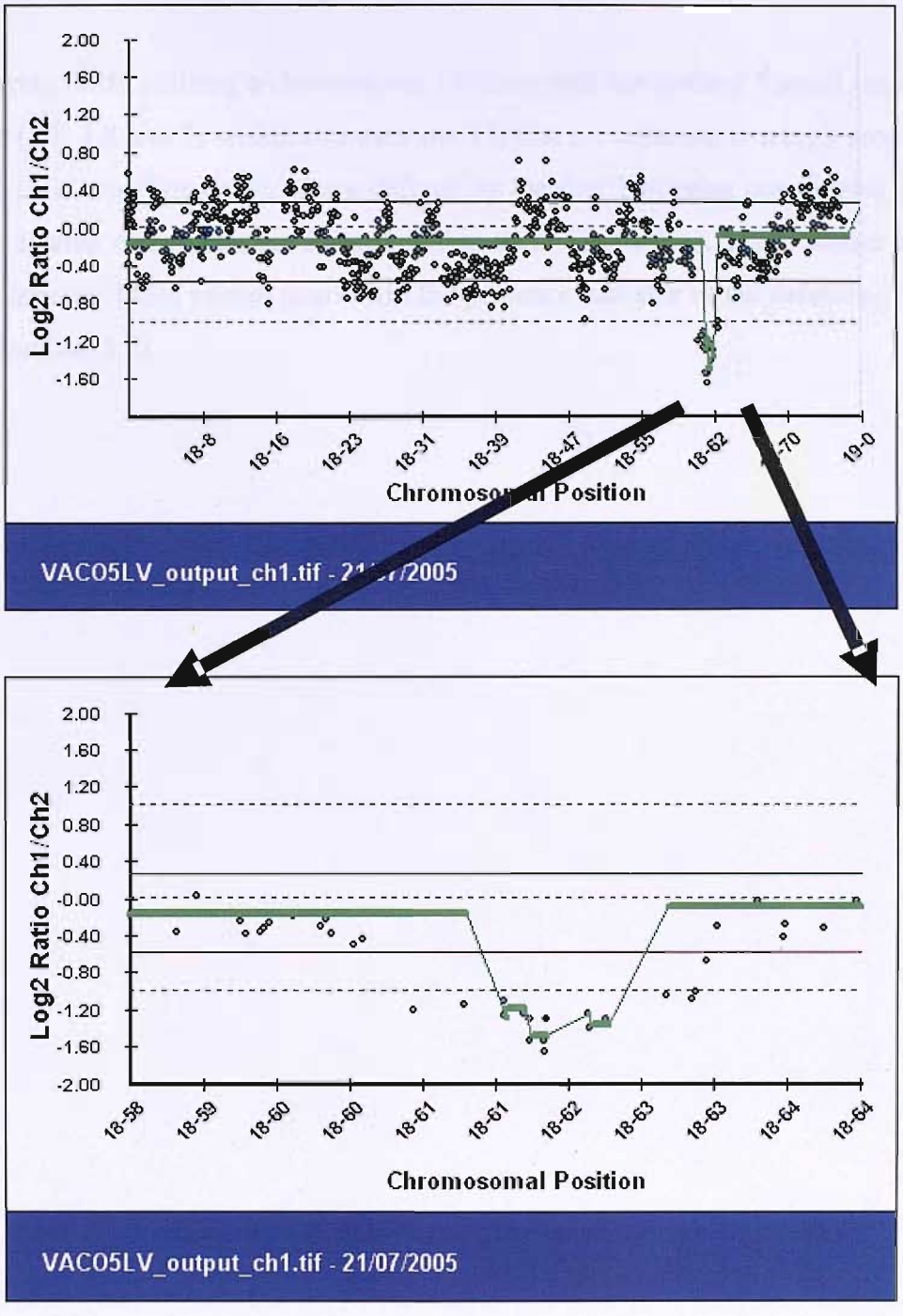
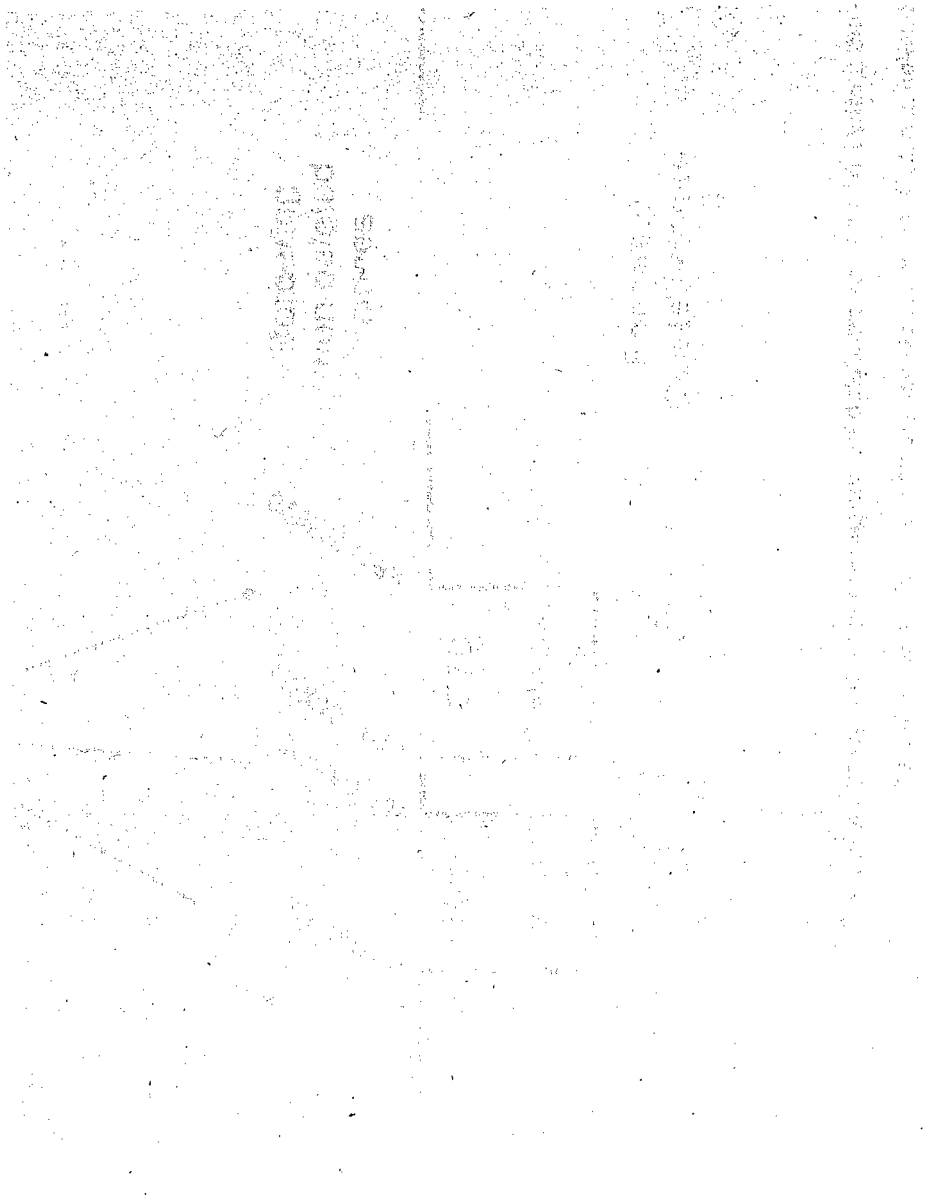


Fig. 3.16: deletion 2 within chromosome 18 (above). Magnified view of array CGH intensity ratios in a 6Mb sub-region, centred on deletion 2 (below).

There are multiple deletions. Deletion 2 is identified by the small number of consecutive clones located at 18-60 to 18-63 with intensity ratios of less than -1

Micro-array CGH utilising a chromosome 18 tiling path has defined 3 small regions of deletion (1.1, 1.2 and 2) within chromosome 18, that are common to a high proportion of CRC cell lines. The deletions are defined by the first bordering non-deleted clones and the deleted clones that are located within them (Table 3.1). These clones can be used as labelled FISH probes to confirm the presence and size of the deletions, (Table 3.2 and section 3.5).



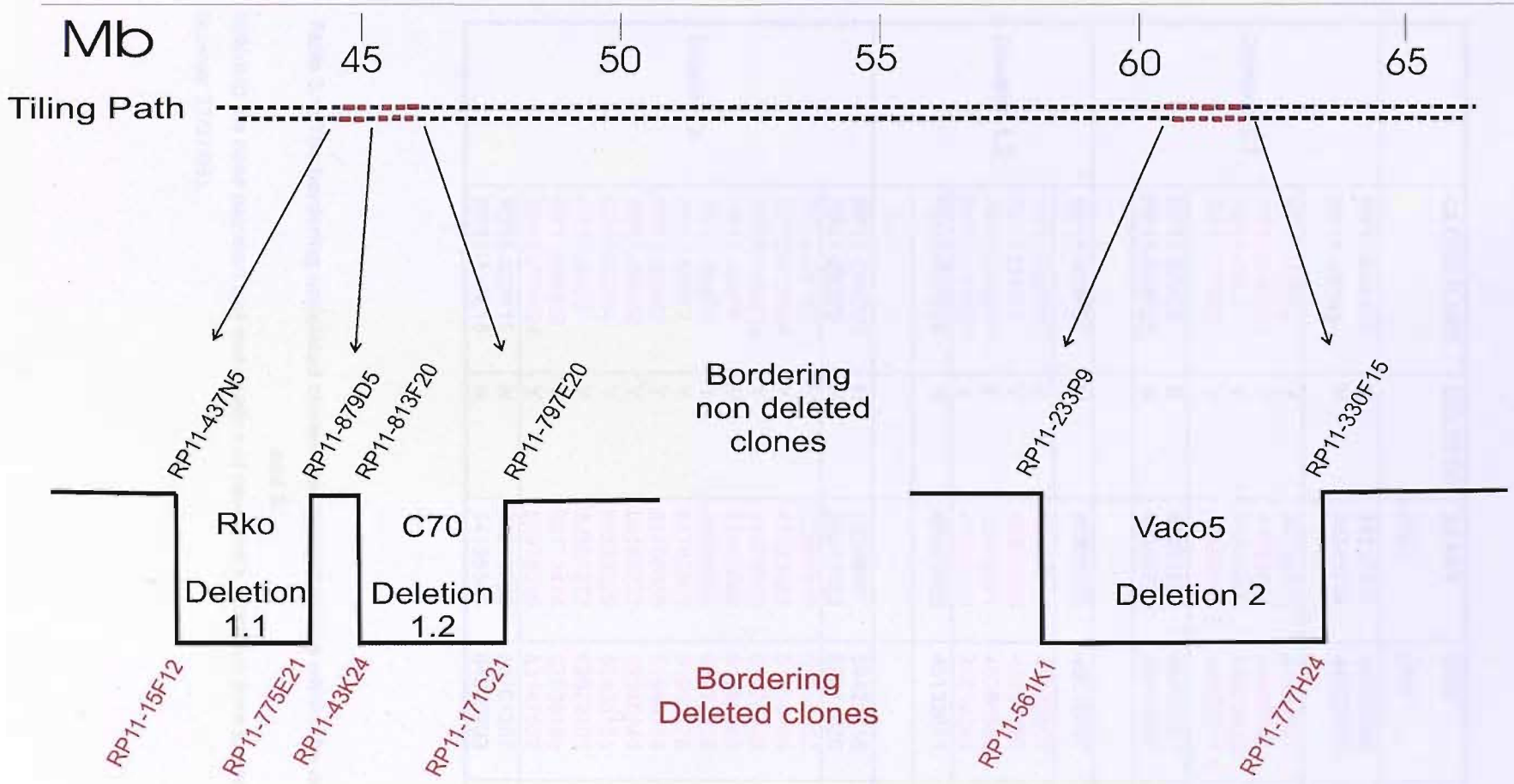


Fig. 3.17 Representation of the location of deletions 1.1, 1.2 and 2 within the chromosome 18 tiling path.

The bordering or "flanking" non deleted clones are in black and deleted clones are in red.

	CLONE NAME	DELETED	START (bp)	END (bp)	DELETION SIZE
Deletion 1.1	RP11-594A17	N	44137937	44305300	
	RP11-437N5	N	44549759	44723913	
	RP11-15F12	Y	44770510	44941621	
	RP11-745D21	Y	44791561	44965941	
	RP11-30L3	Y	44890987	45070820	478 kb
	RP11-775E21	Y	45158592.5	45357811.5	
	RP11-879D5	N	45201782	45400367	
	RP11-813F20	N	45385722	45610087	
Deletion 1.2	RP11-813F20	N	45385722	45610087	
	RP11-727G13	Y	45487817	45670802	
	RP11814I23	Y	45570558	45750720	
	RP11-43K24	Y	45609221	45786550	371kb
	RP11-171C21	Y	45822849	45992284	
	RP11-797E20	N	45981002	46132611	
Deletion 2	RP11-745G22	N	59249097	59421715	
	RP11-233P9	N	59815023	59971920	
	RP11-561K1	Y	60895917	61045057	
	RP11-910H24	Y	61121023	61310464	
	RP11-453M23	Y	61095109	61276475	
	RP11-223P4	Y	61445086	61590960	
	RP11-284D6	Y	61561805	61774929	
	RP11-539E2	Y	61782801	61959178	3.5 Mb
	RP11-505N2	Y	61766655	61959178	
	RP11-784M8	Y	62150222	62340941	
	RP11-727B4	Y	62332329	62498151	
	RP11-90O2	Y	62501272	62672101	
	RP11-45M20	Y	62773474	62936169	
	RP11-777H24	Y	62975316	63147932	
	RP11-330F15	N	63510386	63701257	
	RP11-177K16	N	64395345	64590793	

***Table* 3.1: The bordering undeleted clones and deleted clones within the deletions 1.1, 1.2 and 2.**

Including the base pair start and end location of each clone (Location from Ensemble genome browser 22/07/05).

3.5 FLOURESCENT IN SITU HYBRIDISATION (FISH)

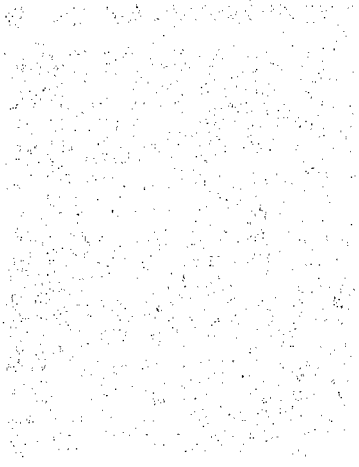
FISH was utilised to confirm the critical regions of deletion identified by the array CGH analysis.

The BAC clones delineating the critical regions of deletion were obtained from the BAC and PAC Resource Centre (BPRC) (www.bacpac.chori.org/libraries) as agar stabs. They were then prepared as inoculated cultures according to standard protocols (sections 2.4.1) prior to DNA extraction using standard methodologies (section 2.4.1). The BAC clones and a chromosome 18 marker (centromeric) were then labelled by nick translation with texas green and red respectively (section 2.4.2).

The cell lines containing the critical regions of deletion 1.1 and 2, RKO and VACO 5 respectively were utilised. The cell line defining deletion 1.2, C70 continues to be unavailable because of technical problems with its culture.

Normal metaphases and metaphases from the selected cell lines were prepared and placed on microscope slides (section 2.4.4 and 2.4.5).

Hybridisation of the labelled chromosome 18-marker and BAC clones to the metaphases was performed (section 2.4.6) and the slides washed (2.4.7). The slides were then viewed under a microscope and images captured.



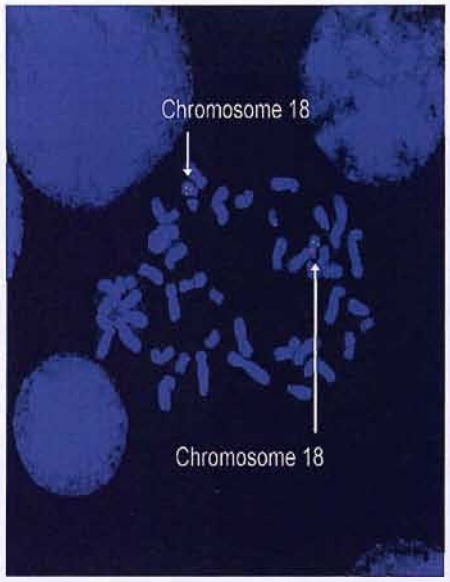


Fig. 3.18a: RP11-437N5
no deletion

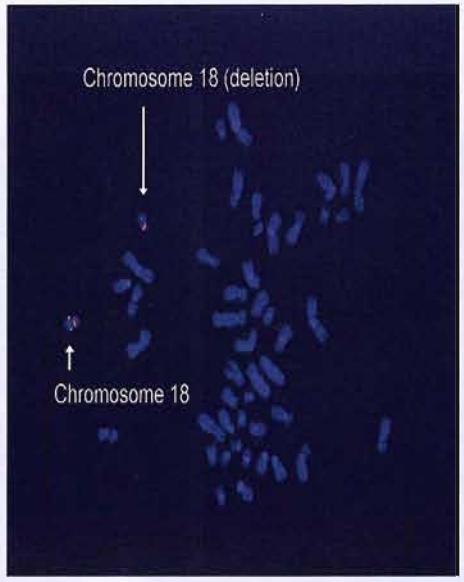


Fig. 3.18 b: RP11-15F12
deletion

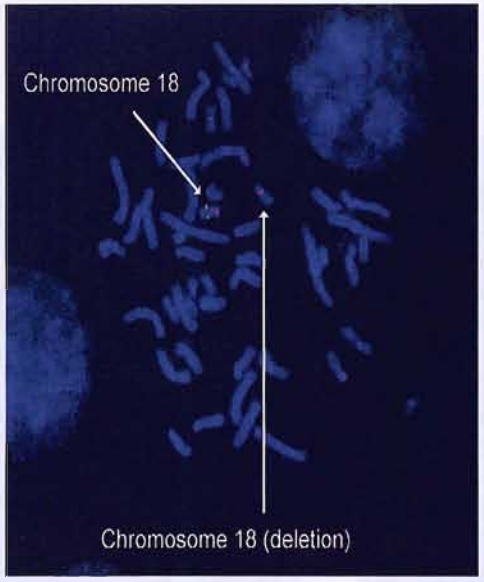


Fig. 3.18 c: RP11-775E21
deletion

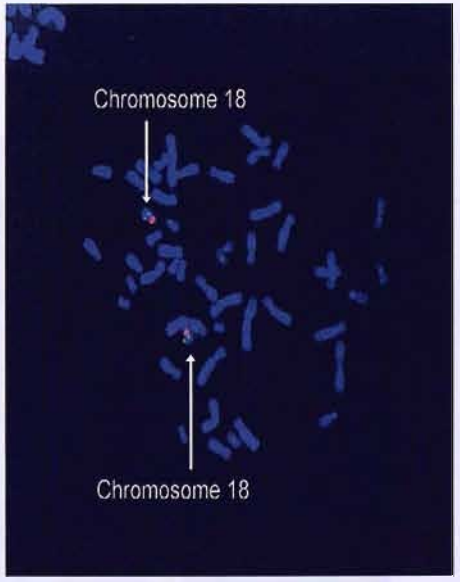


Fig. 3.18 d: clone RP-879D5
no deletion

Fig. 3.18: FISH metaphase, cell line rko,

Chromosome 18 is identified by the red marker in each of the 4 metaphases shown above. Clones are identified by green markers. In the metaphases shown in Fig 3.18 a and d, green markers are present on both chromosome 18 alleles, confirming that there is no deletion. In 3.18 b and c a deletion is confirmed, as a green marker is absent on one chromosome 18q allele.

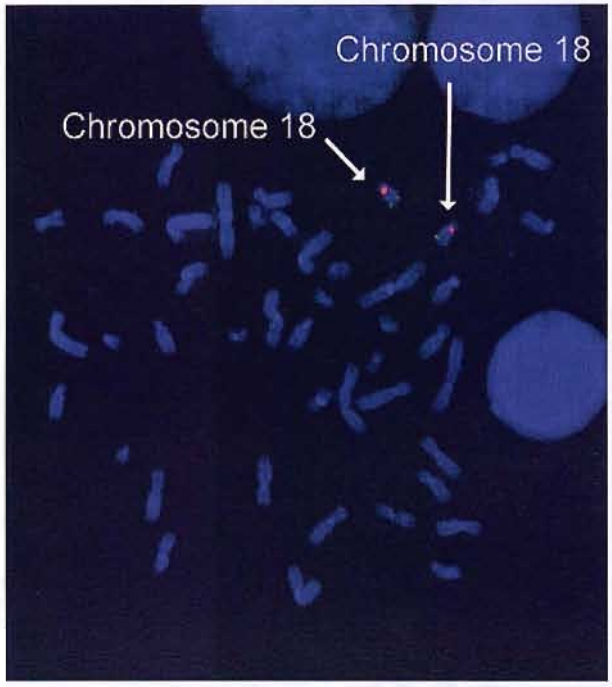


Fig 3.19a clone RP11-233P9, no deletion

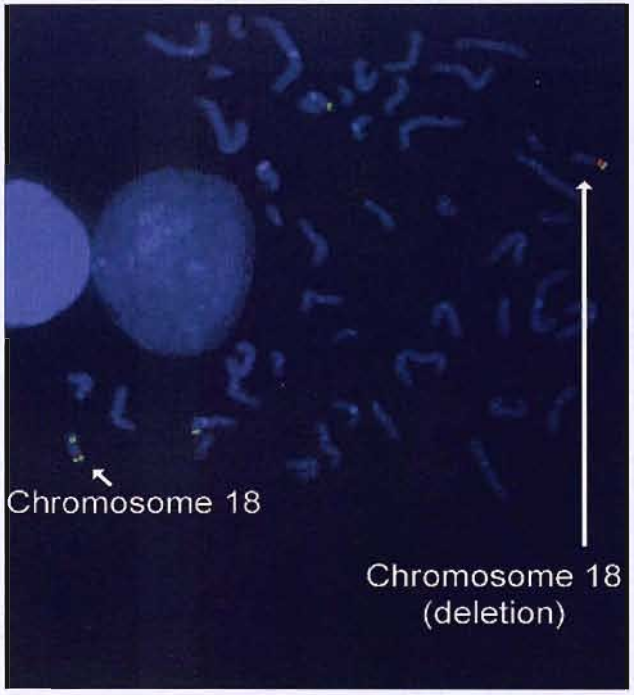


Fig. 3.19 b clone RP11-784M8, a deletion

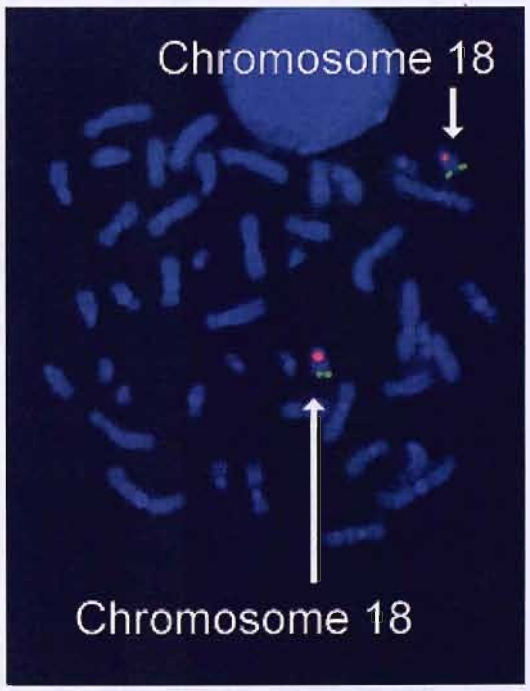


Fig 3.19 c clone RP11-330F15, no deletion

Fig. 3.19: FISH metaphases, cell line vako 5.

Chromosome 18 is identified by the red marker in each of the 3 metaphases shown above. Clones are identified by green markers. In Fig 3.19 a and c, green markers are present on both chromosome 18 alleles, confirming that there is no deletion. In Fig 3.19 b a deletion is confirmed, as a green marker is absent on one chromosome 18q allele. This metaphase result is complicated by the green labelled clone hybridising not only to 18q but also 18p and one other chromosome.

	CLONE NAME	DELETED	START (bp)	END (bp)	FISH
Deletion 1.1	RP11-594A17	N	44137937	44305300	
	RP11-437N5	N	44549759	44723913	Non deletion confirmed
	RP11-15F12	Y	44770510	44941621	Deletion confirmed
	RP11-745D21	Y	44791561	44965941	
	RP11-30L3	Y	44890987	45070820	
	RP11-775E21	Y	45158592.5	45357811.5	Deletion confirmed
	RP11-879D5	N	45201782	45400367	Non deletion confirmed
	RP11-813F20	N	45385722	45610087	
Deletion 1.2	RP11-813F20	N	45385722	45610087	
	RP11-727G13	Y	45487817	45670802	
	RP11814I23	Y	45570558	45750720	
	RP11-43K24	Y	45609221	45786550	
	RP11-171C21	Y	45822849	45992284	
	RP11-797E20	N	45981002	46132611	
Deletion 2	RP11-745G22	N	59249097	59421715	
	RP11-233P9	N	59815023	59971920	Non deletion confirmed
	RP11-561K1	Y	60895917	61045057	
	RP11-910H24	Y	61121023	61310464	
	RP11-453M23	Y	61095109	61276475	
	RP11-223P4	Y	61445086	61590960	
	RP11-284D6	Y	61561805	61774929	
	RP11-539E2	Y	61782801	61959178	
	RP11-505N2	Y	61766655	61959178	
	RP11-784M8	Y	62150222	62340941	Deletion confirmed
	RP11-727B4	Y	62332329	62498151	
	RP11-90O2	Y	62501272	62672101	
	RP11-45M20	Y	62773474	62936169	
	RP11-777H24	Y	62975316	63147932	
	RP11-330F15	N	63510386	63701257	Non deletion confirmed
	RP11-177K16	N	64395345	64590793	

Table 3.2: A copy of table 3.1 containing the array-CGH identified non deleted clones and deleted clones defining the deletions 1.1, 1.2 and 2.

This table has an additional column on the right containing the deletion 1.1 and 2 clones that have also been consistently confirmed by FISH as non deleted and deleted

The FISH findings were clearly consistent in all but one case, confirming deletions and non-deletions as expected from the array CGH analysis (Fig.3.18 and Fig. 3.19). The exception was clone RP11-784M8 which was identified in cell line vaco5 as a deletion with an intensity ratio of –1.44; this clone as expected co-hybridised to the long arm of chromosome 18 on one allele but not the other, confirming a deletion. The inconsistency was that RP11-784M8 also co-hybridised to the short arm of chromosome 18 and one additional chromosome. A chromosome 18q deletion was confirmed as expected by FISH but with this clone mapping to 2 other regions without deletions, an array CGH intensity ratio closer to 0 would be expected.

3.6 ANALYSIS OF PRIMARY COLORECTAL CANCERS

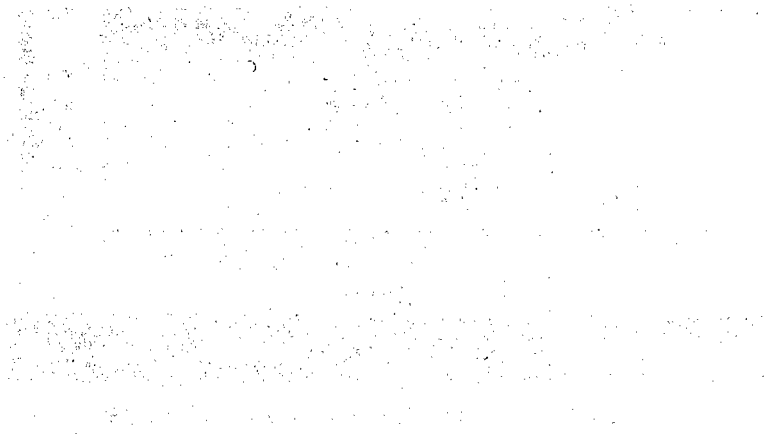
Having identified a high frequency of deletions in the cell lines including 2 minimal regions of deletion confirmed by FISH it was important to confirm the presence of these deletions in primary colorectal cancers. This is important as non causative deletions can occur de novo in cell lines.

DNA extracted from 69 consecutive stage II and III primary colorectal cancers was provided by William Chambers from the Department of Colorectal Surgery, John Radcliffe Hospital, Oxford. The samples had been micro-dissected to minimise normal and adenoma cell contamination then stored in liquid nitrogen prior to DNA extraction.

Microarray CGH analysis and data was obtained from all 69 primary cancer DNA samples (section 2.2).

There was no evidence of deletions in 31 (45%) primary cancers, there were deletions evident in 38 (55%) of the primary cancers. The size of the deletions were very variable, ranging in size from the whole of the long arm of chromosome 18 down to 0.7 Mb.

Examples of the data, large and small deletions are shown in Fig. 3.20.



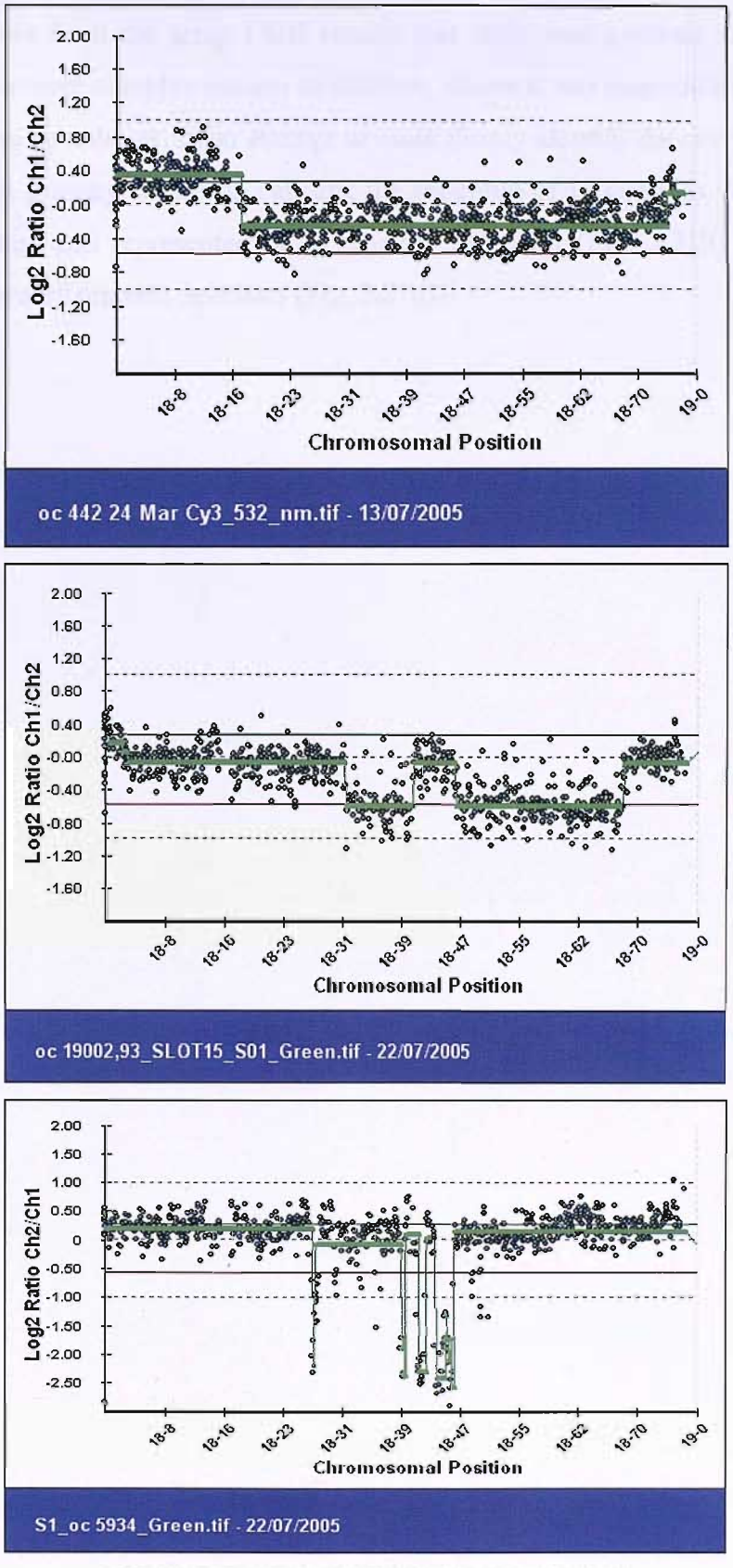
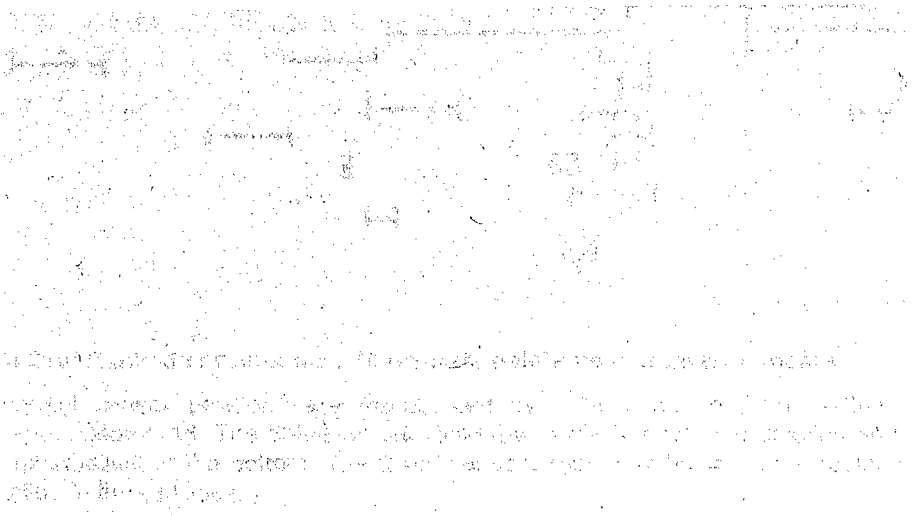


Fig. 3.20 examples of primary cancer deletions

It was apparent from the array CGH results that there was a subset of 14 primary cancers with a more complex pattern of deletion, where it was more difficult to clearly define deletion boundaries. In an attempt to more clearly identify the critical regions of deletion in the primary colorectal cancers; the subgroup of 14 complex deletions have been segregated and represented on a separate bar chart (Fig. 3.21b) from the 24 clearly demarcated discrete deletions (Fig. 3.21a).



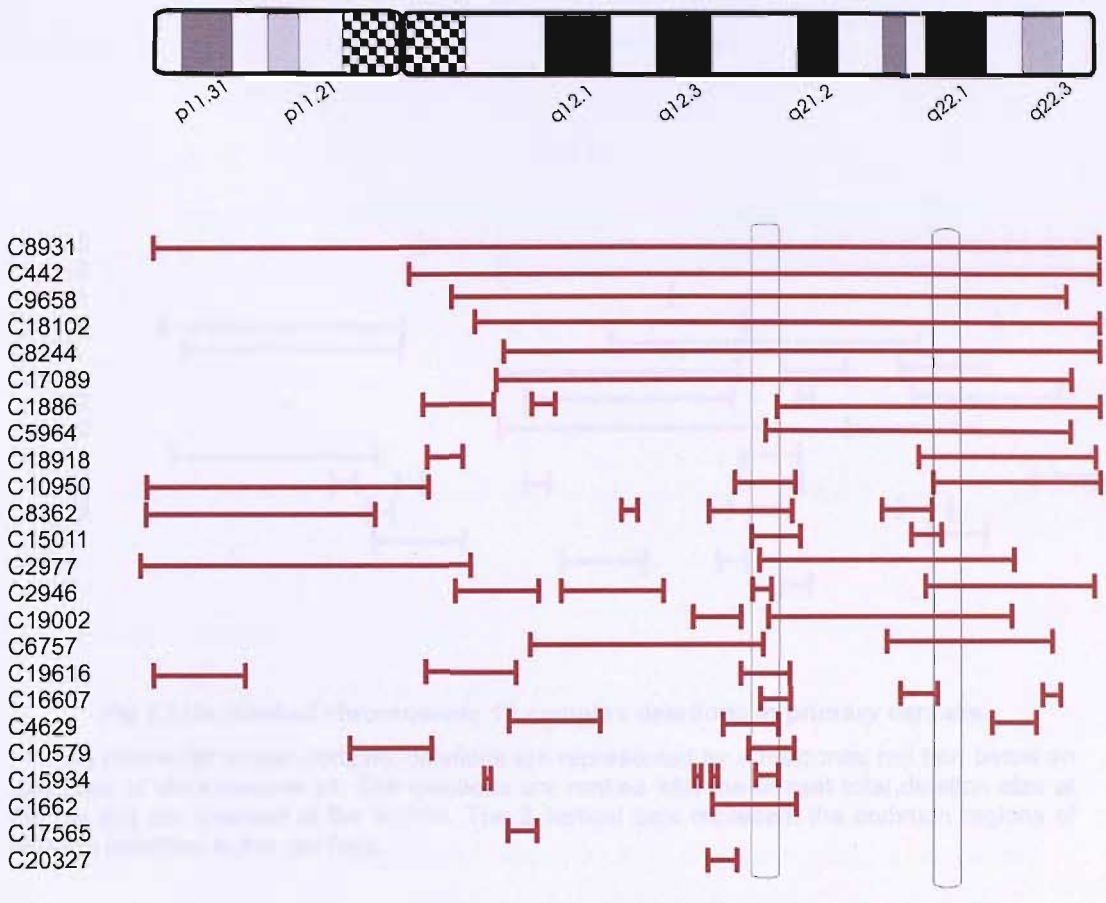


Fig 3.21a: Ranked chromosome 18 discrete deletions in primary cancers

Primary colorectal cancer deletions are represented by a horizontal red bar, below an ideogram of chromosome 18. The deletions are ranked with the largest total deletion size at the top and the smallest at the bottom. The 2 vertical bars represent the common regions of deletion identified in the cell lines.

Fig 3.21b: Ranked chromosome 18 complex deletions in primary cancers

Primary colorectal cancer complex deletions are represented by a horizontal red bar, below an ideogram of chromosome 18. The deletions are ranked with the largest total deletion size at the top and the smallest at the bottom. The 2 vertical bars represent the common regions of deletion identified in the cell lines.

The primary cancer deletions are more complex and when ranked in size order less clearly identify a common region of deletion compared to the cell lines.

The small regions of deletion, (1.1, 1.2 and 2) identified in the cell lines are however common to a high proportion of the larger deletions detected in the primary cancers (Table 3.3). The small deletion 1.1/1.2 is common to 30 (79%) of the 38 primary cancers with a deletion and deletion 2 common to 24 (63%) of the 38 primary cancers with a deletion.

	Primary Cancers (n=69)
Frequency of all deletions	55% (38/69)
Percentage of all deletions incorporating deletions 1.1/1.2	79% (30/38)
Percentage of all deletions incorporating deletion 2	63% (24/38)

Table 3.3: The frequency of chromosome 18 deletions in primary cancers.
The percentage of primary cancer deletions that incorporate minimal regions of deletions identified in the cell lines

Small deletions were identified in the primary cancers which closely replicate 2 of the minimal regions of deletion identified in the cell lines but they were not exclusive in each of these primary cancers, with other deletions evident in other regions of chromosome 18.

Deletion 1.1 and 1.2 are almost exactly incorporated in a deleted region evident in the primary cancer 15934 (Fig. 3.22).

Deletion 1.2 is immediately adjacent but not directly incorporated by a small deletion in the primary cancer 16607 (Fig 3.23).

Deletion 2 is not closely replicated by a small deletion in the primary cancers.

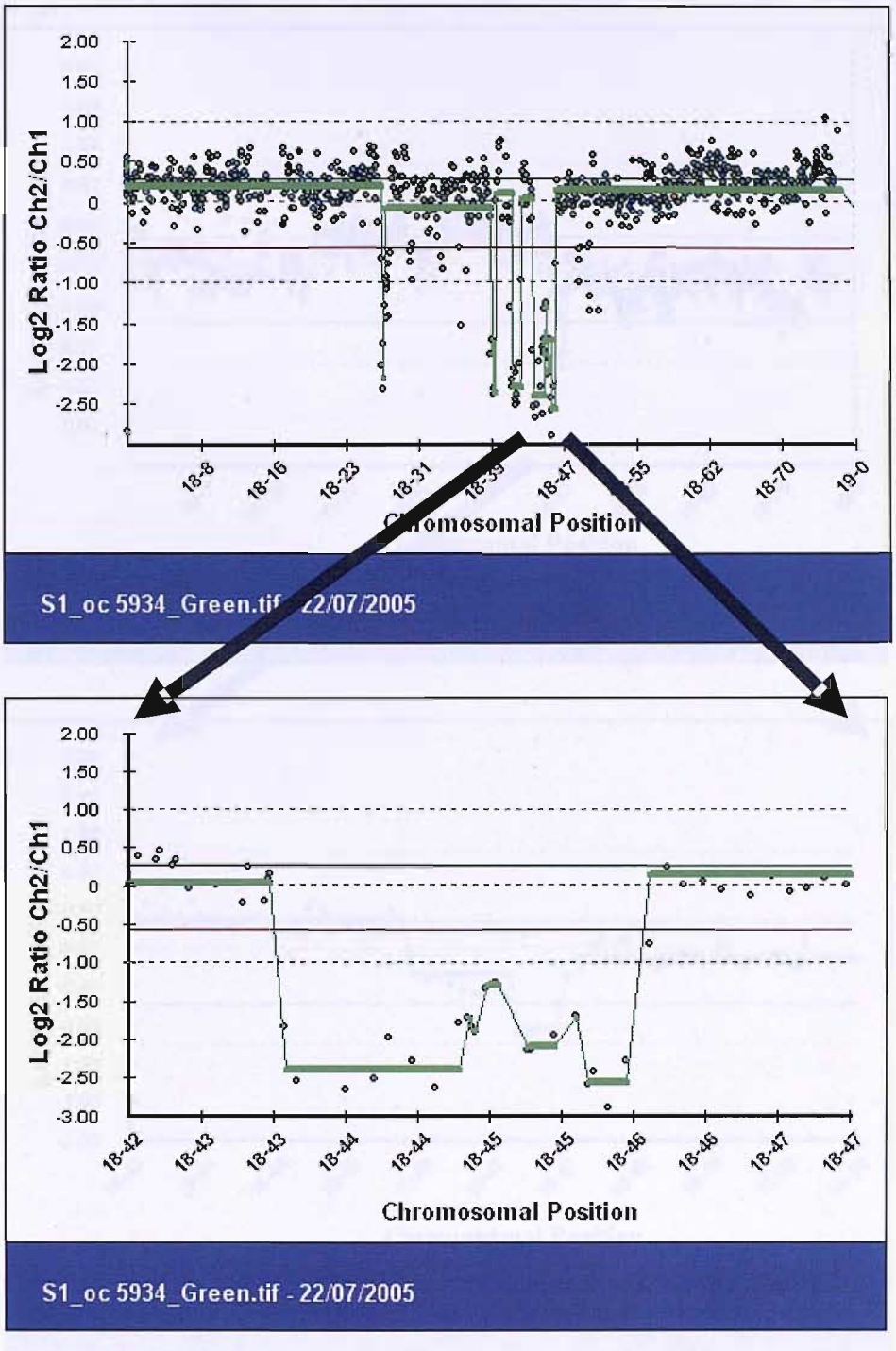


Fig. 3.22: chromosome 18 deletions within a primary cancer (above). Magnified view of array CGH intensity ratios in a 5Mb sub-region, centred on the largest deletion (below).

There are multiple deletions. The largest deletion magnified below closely resembles the combined cell line minimal deletions of 1.1/ 1.2, which span the region 18-44 to 18-46.

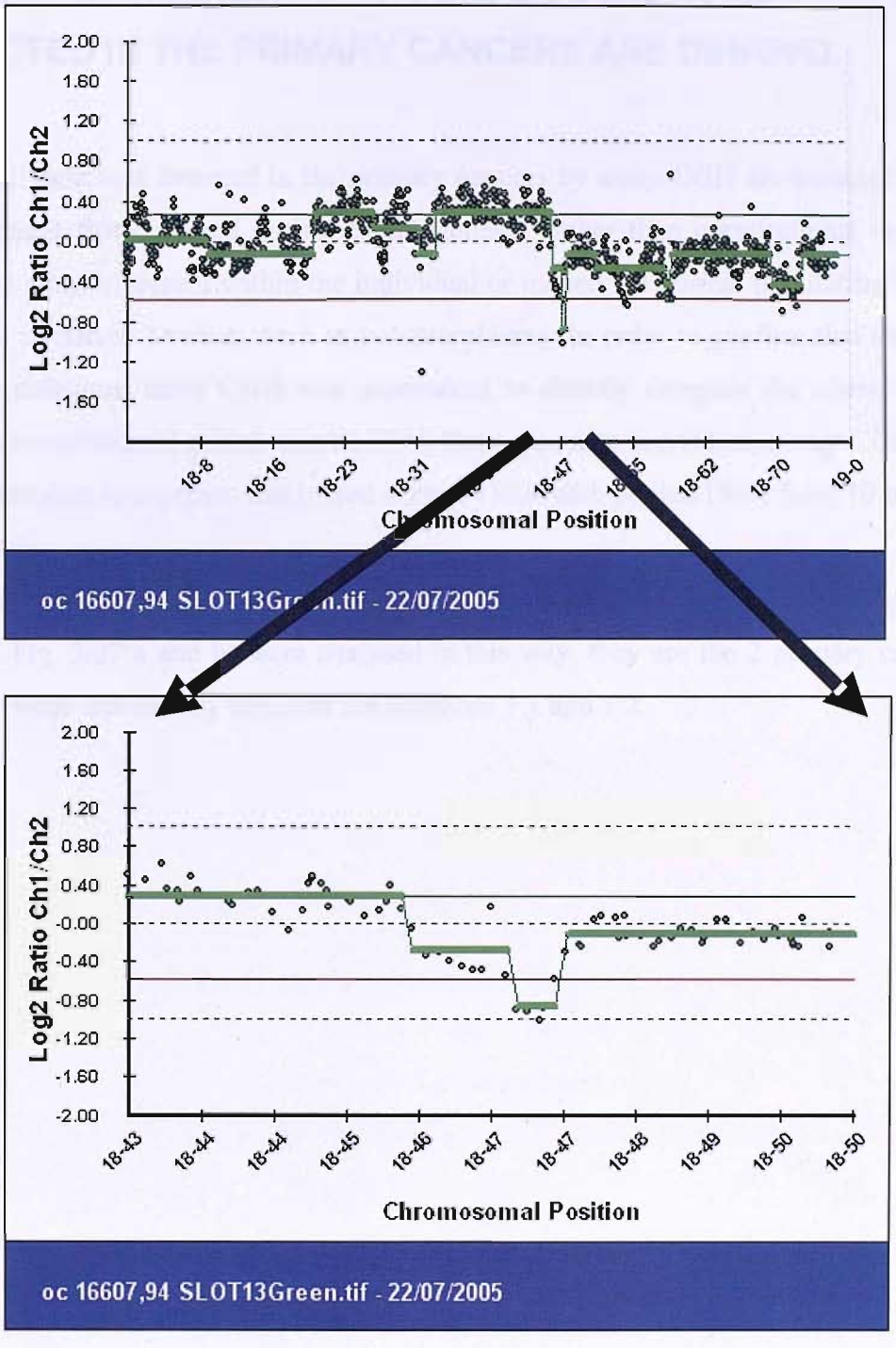


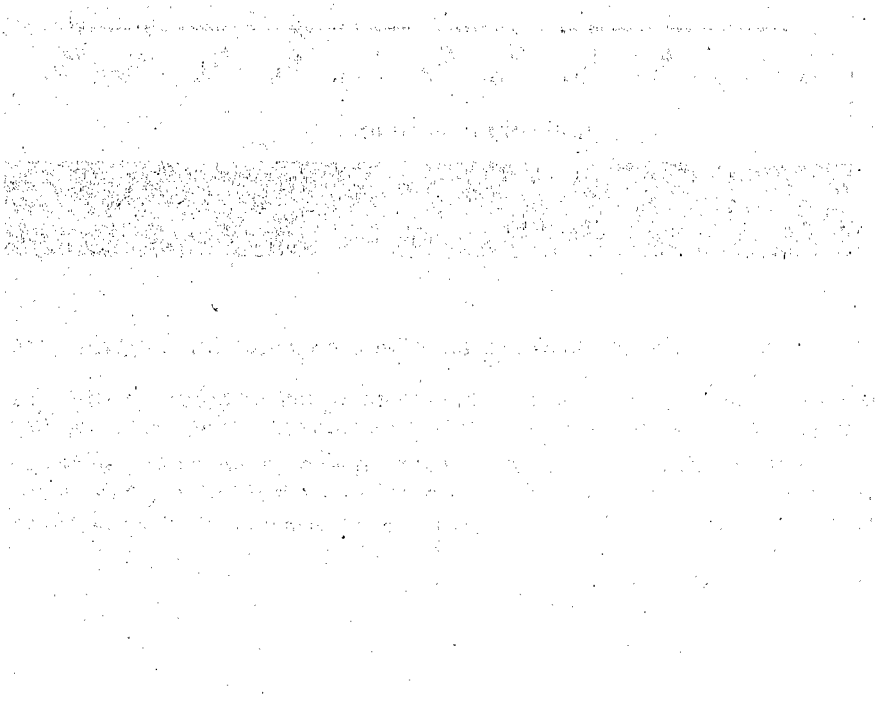
Fig. 3.23: chromosome 18 deletions in a primary cancer (above). Magnified view of array CGH intensity ratios in a 7Mb sub-region, centred on the clearest deletion (below).

There are multiple deletions. The largest deletion magnified below. A clear deletion is present with intensity ratios below -0.6 at 18-47. On the centromeric side are consecutive clones from 18-45 to 18-47 with intensity ratios close to a threshold level. This latter region from 18-45 to 18-47 in this primary cancer incorporates the cell line deletion 1.2 from 18-45.6 to 18-46.

3.7 CONFIRMATION THAT THE SMALL DELETIONS DETECTED IN THE PRIMARY CANCERS ARE DENOVO.

The small deletions detected in the primary cancers by array CGH are assumed to be new changes that occurred during carcinogenesis, rather than constitutional changes that were already present within the individual or indeed the normal population before the CRC started to develop, such as polymorphisms. In order to confirm that they are de-novo deletions, array CGH was undertaken to directly compare the cancer DNA with the constitutional paired normal DNA from the same individual. Array CGH was also undertaken to compare this paired normal DNA with pooled DNA from 10 normal individuals.

The primary cancers 15934 (Fig. 3.24 a and b. Fig. 3.25 a and b) and 16607 (Fig. 3.26 a and b. Fig. 3.27 a and b) were analysed in this way, they are the 2 primary cancers with deletions that closely replicate the deletions 1.1 and 1.2.



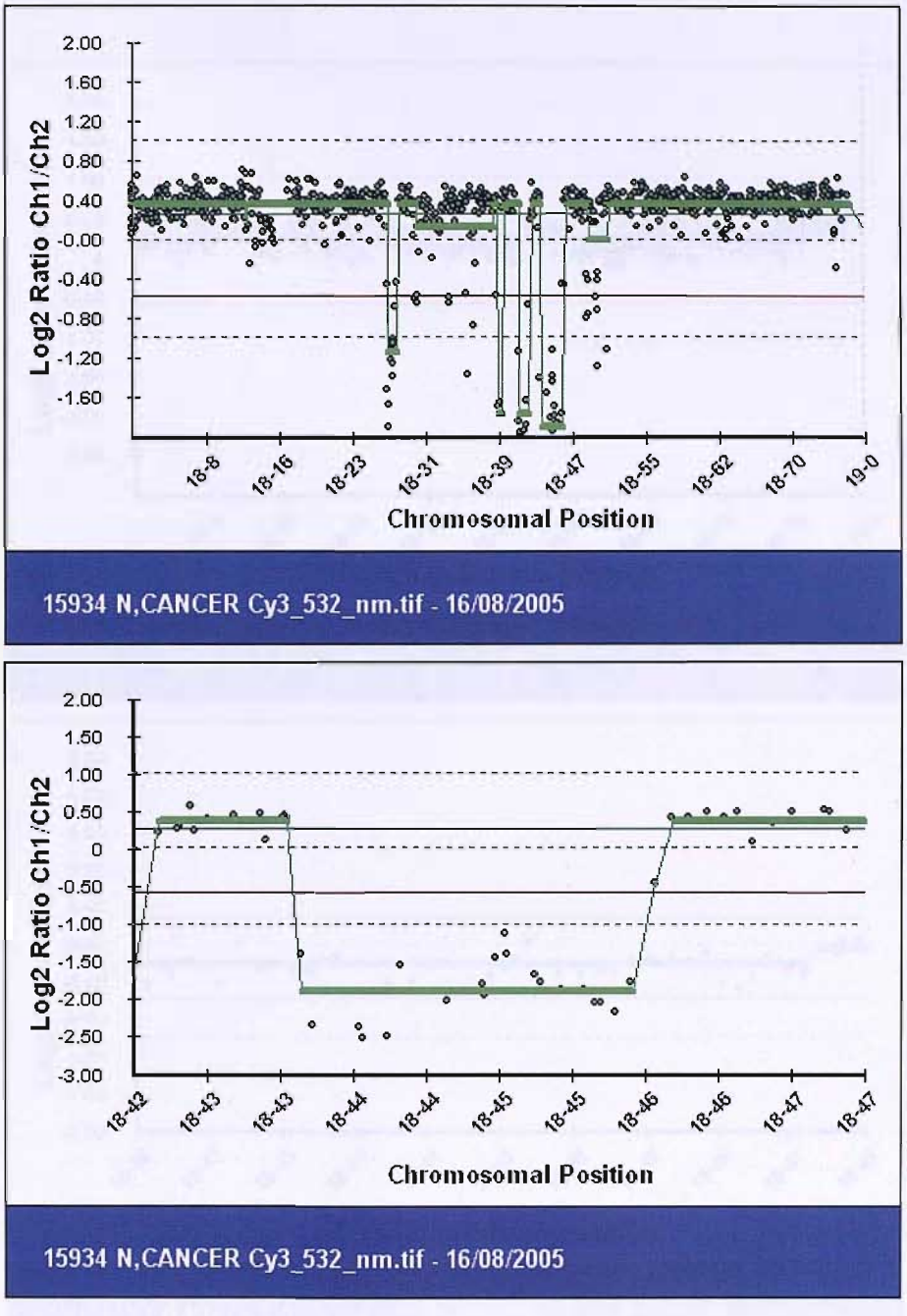


Fig. 3.24a/b: Array CGH intensity ratios for cancer DNA vs. paired normal DNA

Fig 3.24a: shows intensity ratios for the whole of chromosome 18. Fig 3.24b: shows intensity ratios for the 5Mb region centred on the deletion that closely replicates deletions 1.1 and 1.2.

There is an obvious copy number change from 18-43 to 18-46.This confirms that the previously detected deletions were not present in the paired normal constitutional DNA but occurred during carcinogenesis to be present within the cancer DNA from the same individual.

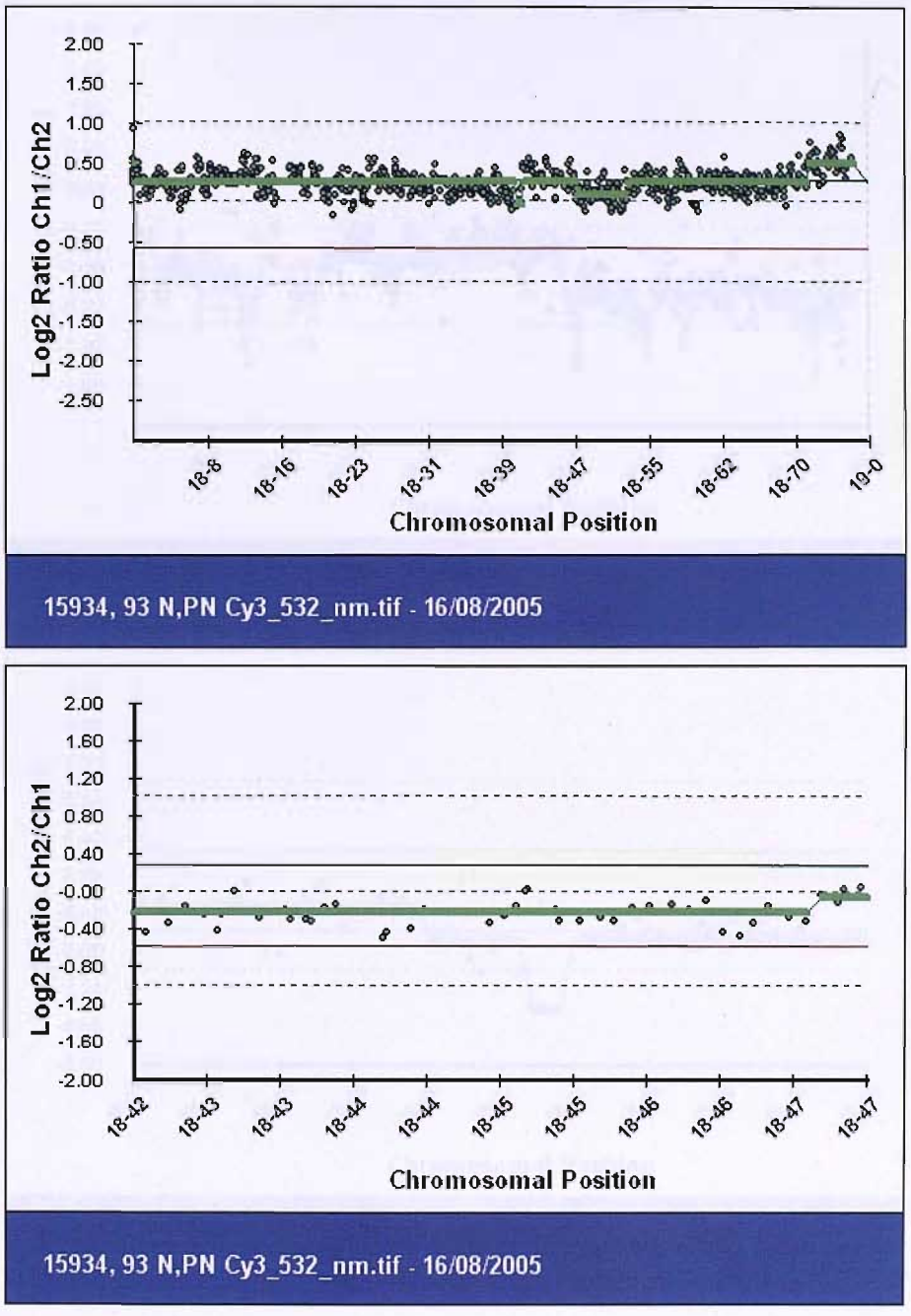


Fig. 3.25a/b: Array CGH intensity ratios for paired normal DNA vs. pooled normal DNA

Fig 3.25a: shows intensity ratios for the whole of chromosome 18. Fig 3.25b: shows intensity ratios for the 5Mb region centred on the deletion that closely replicates deletions 1.1 and 1.2. The intensity ratio is approximately 0 for the whole of chromosome 18 including the 18-44 to 18-46 region of deletions 1.1 and 1.2. This indicates that there is no significant copy number difference for this region between the 15934 constitutional paired normal DNA and DNA pooled from 10 normal individuals.

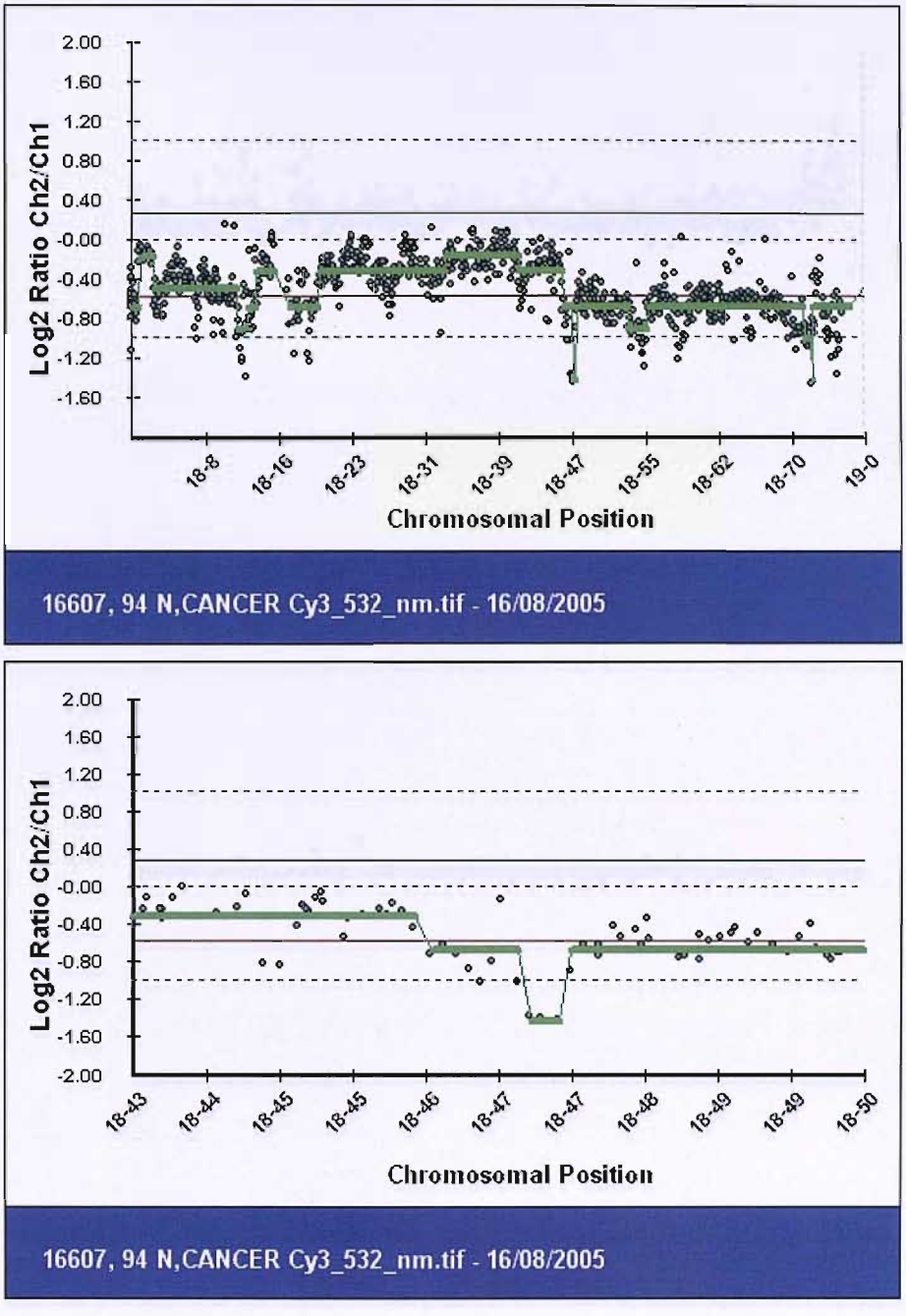


Fig. 3.26 a/b: Array CGH intensity ratios for cancer DNA vs. paired normal DNA

Fig 3.26a: shows intensity ratios for the whole of chromosome 18. Fig 3.26b: shows intensity ratios for the 7Mb region centred on the location of deletion 1.2.

There is an obvious copy number change from 18-45 to 18-47.This confirms that the previously detected deletions were not present in the paired normal constitutional DNA but occurred during carcinogenesis to be present within the cancer DNA from the same individual.

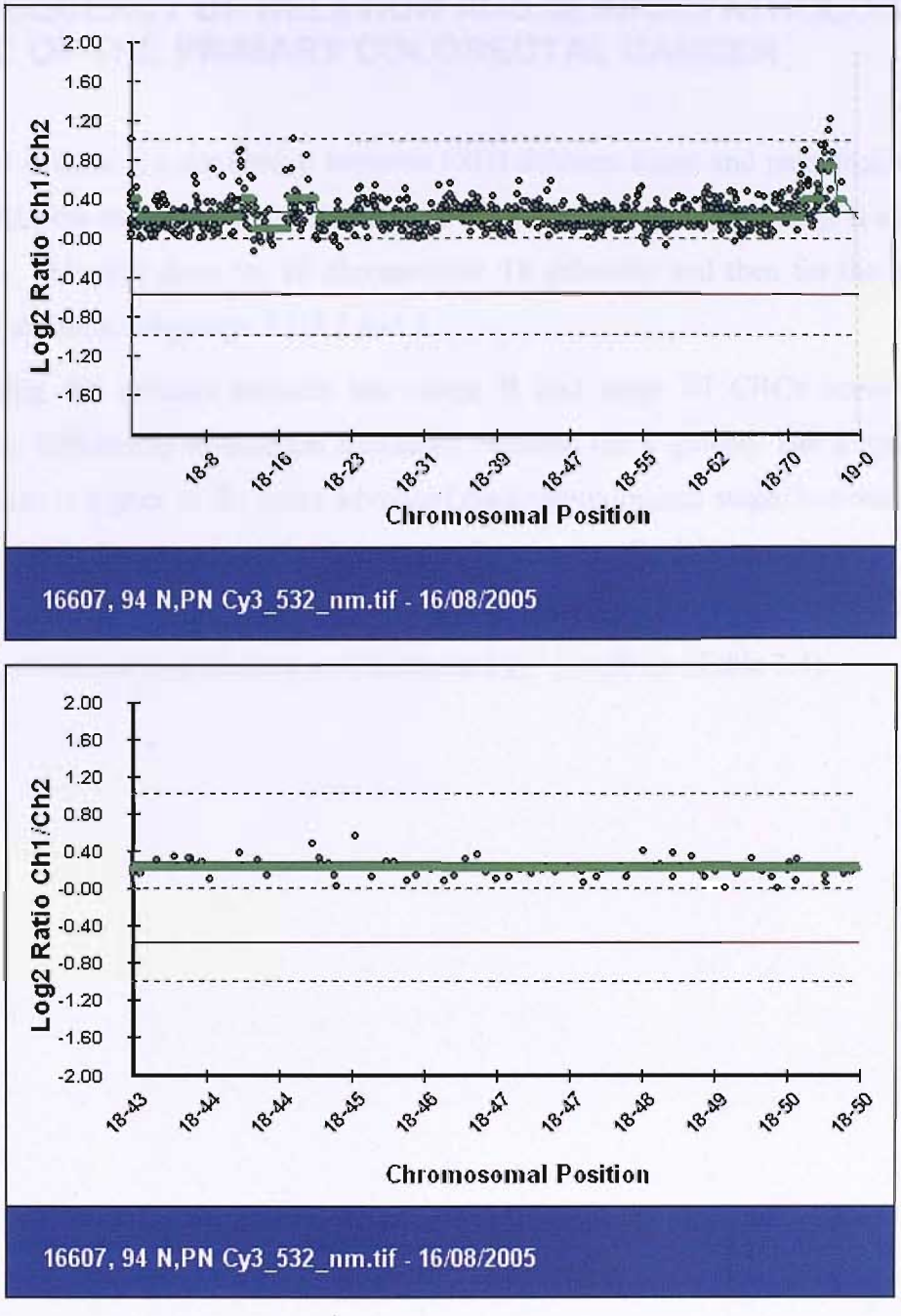


Fig. 3.27 a/b: Array CGH intensity ratios for paired normal DNA vs. pooled normal DNA

Fig 3.27a: shows intensity ratios for the whole of chromosome 18. Fig 3.27b: shows intensity ratios for the 7Mb region centred on the deletion 1.2.

The intensity ratio is approximately 0 for the whole of chromosome 18 including the 18-45 to 18-46 region of deletion 1.2. This indicates that there is no significant copy number difference for this region between the 16607 constitutional paired normal DNA and DNA pooled from 10 normal individuals.

3.8 FREQUENCY OF DELETION AND CLINICOPATHOLOGICAL STAGE OF THE PRIMARY COLORECTAL CANCER

To assess if there is a correlation between CGH deletion status and pathological stage of the CRC, the frequency of deletions were calculated separately for stage II and stage III CRCs. This was done for all chromosome 18 deletions and then for the minimal region of deletion subgroups 1.1/1.2 and 2.

Subdividing the primary cancers into stage II and stage III CRCs reveals some interesting differences in deletion frequency between the 2 groups. The frequency of all deletions is higher in the more advanced clinicopathological stage, increasing from 50% to 61%. The increase in frequency for the small deletions is even larger, especially 1.1/1.2, 33% to 55% between stage II and stage III and this is the difference closest to statistical significance on Chi-squared (χ^2) analysis (Table 3.4).

	Primary Cancers (n=69)		Stage II cancers (n=36)	Stage III cancers (n=33)	Degree freedom	Chi-squared χ^2	Stage II vs. stage III p value
Frequency of all deletions	55% (38/69)		50%(18/36)	61%(20/33)	1	0.783	0.376
Frequency of deletions 1.1/1.2	44% (30/69)		33%(12/36)	55%(18/33)	1	3.152	0.0758
Frequency of deletion 2	35% (24/69)		28%(10/36)	42%(14/33)	1	1.628	0.201

Table 3.4: Frequency of deletions in all primary colorectal cancers and two subgroups of pathological stage.

CHAPTER 4: DISCUSSION

Array CGH has proved to be effective in high resolution analysis of chromosome 18 copy number change in colorectal cancer. A chromosome 18 tiling path array has been successfully constructed and an established array CGH technique replicated. The accuracy of this technique has been verified by the analysis of pooled normal DNA and a cell line with a previously defined large deletion within 18q. The technique has then been utilised to analyse 47 cell lines and 69 primary cancers, identifying deletions in 81% of the cell lines and 55% of the primary cancers. The high resolution of the technique has enabled the identification of 3 small common regions of deletion in the cell lines, spanning 371kb to 3.5 Mb. The presence and location of these deletions was confirmed with FISH analysis.

The primary CRCs have a more complex pattern of deletion which less clearly defines a minimal region of deletion from the cell lines when ranked in size order but it does importantly confirm the presence of similar minimal regions of deletions in the primary cancers to those identified in the cell lines.

Despite identification of these deletions in cell lines and primary CRCs it is still possible that they are not causative in colorectal carcinogenesis. It was important to verify that the small deletions identified by array-CGH had occurred during tumourigenesis and were not simply polymorphisms. The increased resolution of techniques to identify genome wide copy number changes has resulted in the detection of previously unknown constitutional deletions and amplifications [259, 260], some of which may be important to disease susceptibility [261] whilst others may be polymorphisms found in healthy individuals and not thought to be pathogenic. One strategy to demonstrate that the deletions are not normal variants is to screen a normal population using the chromosome 18 tiling path array. An alternative more time efficient strategy was chosen; hybridisation of the CRC DNA with a detected small deletion was performed against the DNA extracted from blood of the same individual and this in turn was hybridised against DNA pooled from 10 normal individuals. This strategy confirmed that the small deletions were de-novo in the CRCs as they were not present in the DNA extracted from blood of the same individual or pooled normal DNA.

Array CGH has continued to be productive in the identification of copy number change not just in CRC in this thesis but in other cancers as well.

Specific regions within the genome have been investigated, identifying copy number change in prostate cancer (13q33-44, 49-51 and 74-76) [262] and sporadic phaeochromocytoma (22q and 1p) [263]. Identification of regions containing possible causative genes in T cell lymphoma (2q34-37, 8q23-24 and 20p) [264] and accurately locating breakpoints in acute lymphoblastic leukaemia (9p13.2 and 20q11.2) [265].

Genome wide arrays have identified genomic complexity in acute lymphoblastic leukaemia [266] and specific copy number change which may be associated with target genes in uterine cervical carcinoma (3q21-22, 5p15.2, 5p13 and 3p14.2) [267]. Genome wide array CGH has also identified different diagnostic sub-groups in prostate [224], pancreatic [268], renal (Wilms tumours) [269] and hepatocellular carcinoma [270]; as well as differentiate between different pre-malignant conditions in breast cancer [271].

4.1 MINIMAL REGION OF DELETION

Previously LOH analysis had only narrowed down the important region of deletion on chromosome 18 to 37 Mb, with only limited evidence from LOH data to narrow it further to a 16Mb region [241] (Fig 4.1). This 16Mb region may however have significance as it overlaps at least part of all the cell line deletions detected by array-CGH in this thesis, except for 1 minimal region of deletion. The smallest region that does overlap at least part of all other deletions including both minimal regions of deletion is accurately represented by the deletion detected in the cell line colo 320 (Fig 4.1, red bar). This region is about 18Mb in size and contains a large number of genes. A higher level of deletion resolution would help to define a small number of candidate tumour suppressor genes.

The smallest regions of deletion identified by the chromosome 18 tiling path array are 478kb, 371kb and 3.5Mb in size (deletions 1.1, 1.2 and 2 respectively) (Fig 4.1) and are common to a high percentage of all detected deletions, inferring they are perhaps critical (and therefore causative) regions of deletion. Both deletions 1.1/1.2 and 2 were found to be common in 92% of all deletions in the cell lines (Table 3.2). The frequency

was lower in primary cancers; deletion 1.1/1.2 common in 79% and deletion 2 in 63% of all deletions in the primary tumour tissue (Table 3.2).

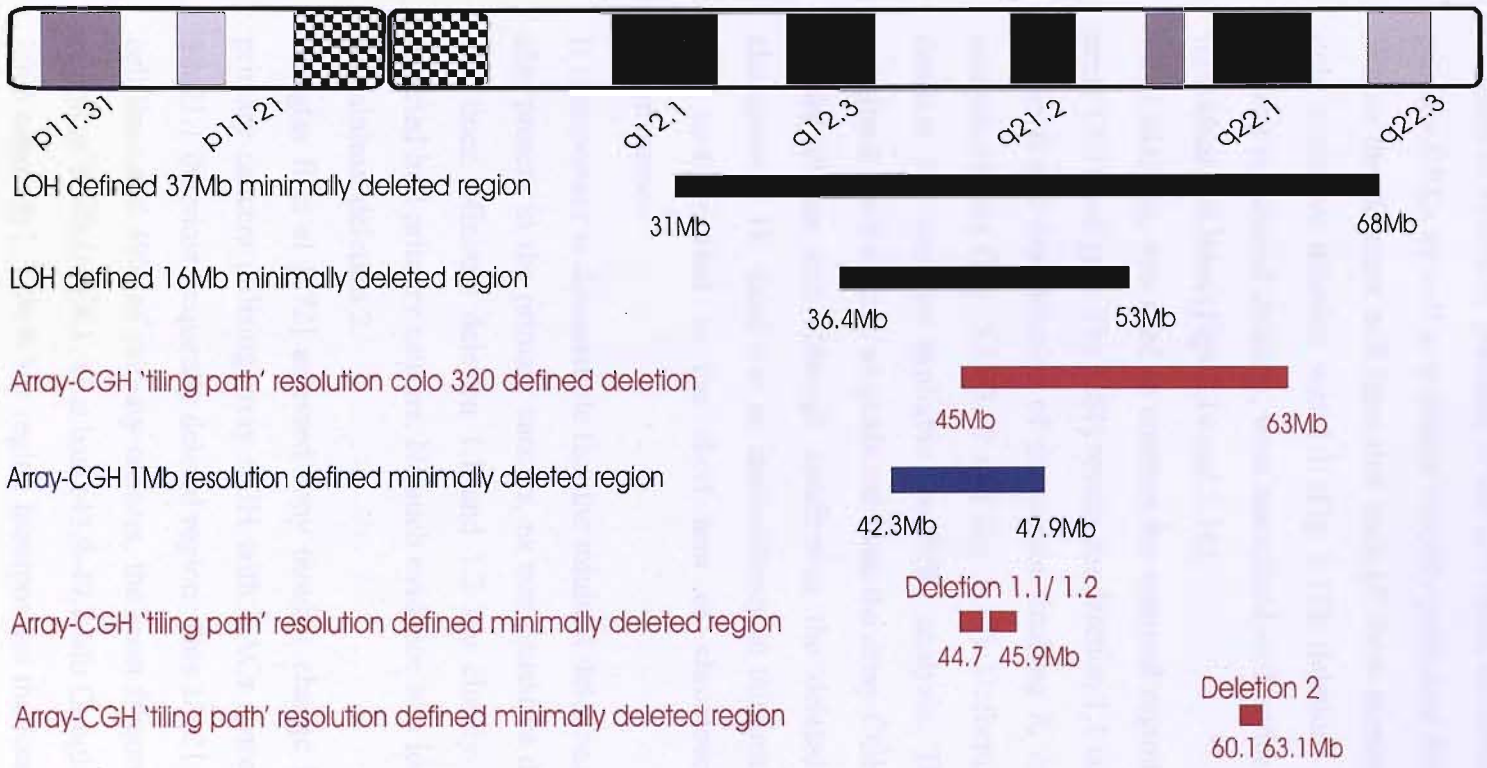


Fig. 4.1: Minimal region of deletion, defined by LOH, array-CGH at a 1 Mb resolution and by array-CGH 'tiling path' resolution in this thesis.

In size order from the top down. Black horizontal bars represent the minimal regions of deletion identified by LOH analysis. Red horizontal bar (3rd from the top) is the smallest region of deletion identified by the array-CGH 'tiling path' in this thesis that overlaps at least part of all other deletions. It is accurately represented by the deletion detected in the cell line colo 320. It is a similar size to the minimal region of deletion identified by LOH but is about 10Mb closer to the telomere. The horizontal blue bar represents the minimal region of deletion identified by array-CGH at a 1 Mb resolution (work published during the laboratory work of this thesis [272]). The smallest horizontal red bars represent the minimal regions of deletion identified by the array-CGH 'tiling path' in this thesis (common to 92% of all deletions in the cell lines)

When estimating which of these minimal regions of deletion is more likely to be critical (causative) it is important to consider the cumulative evidence from this thesis in terms of exclusive presence in the cell lines, confirmation by FISH and presence in primary CRCs as well as evidence recently published from other studies.

Within the different cell lines that each of these minimal regions were identified the only exclusive deletion was 1.1 (Fig 3.13); deletions 1.2 and 2 although the most obvious regions of deletion, were associated with other regions of deletion within the individual cell lines (Figs 3.14 and 3.16).

FISH analysis, was used to confirm the minimal regions of deletion identified by the array CGH analysis. The FISH results for deletion 1.1 confirmed deletion of the clones within it and non-deletion of the clones flanking it, consistent with the array CGH analysis results (Fig. 3.13-3.17 and Fig. 3.19). Unfortunately the cell line to confirm deletion 1.2 was not available for FISH analysis. The FISH results for minimal deletion 2 were exactly as predicted from the array-CGH analysis for the flanking non deleted clones and although confirming the deleted clone on the long arm of chromosome 18 there was an inconsistency in this last metaphase result as the clone also co-hybridised to the short arm of chromosome 18 and one additional chromosome.

It is important to demonstrate that the minimal deletions identified in the cell lines are also present in the primary cancers, as non causative deletions can occur *de novo* in cell lines. Minimal deletion 1.1 and 1.2 are closely replicated by small deletions detected in 2 primary cancers. No such evidence was identified in the primary cancers for minimal deletion 2.

Douglas E et al [272] assessed copy number change for 48 CRC cell lines and 37 primary cancers utilising array CGH with BACs spaced at 1MB intervals, reporting 18q21.1 the most frequently deleted region. This 18q21.1 loss was common to 60% of cell lines and 49% of primary cancers, the most frequently deleted clones covering a region of 6 Mb (Fig 4.1, blue bar) (43.9-49.7 Mb Douglas et al/ 42.35-47.94 Mb, April 2005 assembly). This 6 Mb region incorporates the location of deletion 1.1/1.2 (44.7-45.9 Mb) but not deletion 2 (Fig. 4.1).

The array-CGH and FISH evidence therefore indicates deletion 1.1 as causative in CRC carcinogenesis.

The Douglas paper [272] reported that chromosome 18 loss was more frequent in CIN+ tumours (71% of cell lines and 42% of primary CRCs) than MSI+ tumours (12% of cell lines and 1% of primary CRCs); which is to be expected as CIN+ tumours tend to have a more unstable karyotype. Interestingly they also noted that within the MSI+ tumour group the frequency of deletion of the specific 6Mb region at 18q21.1 (Fig 4.1, blue bar) was at a much higher frequency than the average (21% of cell lines and 29% of primary CRCs). This suggests that these deletions are more likely to be causative in CRC tumourigenesis rather than a reflection of a generally unstable karyotype.

The results reported by Douglas et al were consistent with those obtained in this thesis, a higher overall frequency for chromosome 18 loss identified in the cell lines than the primary cancers (Table 3.2). This may be an accurate assessment of a genuinely lower deletion rate in the primary cancers; it may also be a reflection of a lower detection rate because of normal cell contamination. The other consideration is the case mix, the cell lines analysed included all CRC stages, the primary cancers analysed did not include the advanced stage IV CRCs; more aberrations occur as cancers progress to a more advanced stage [179].

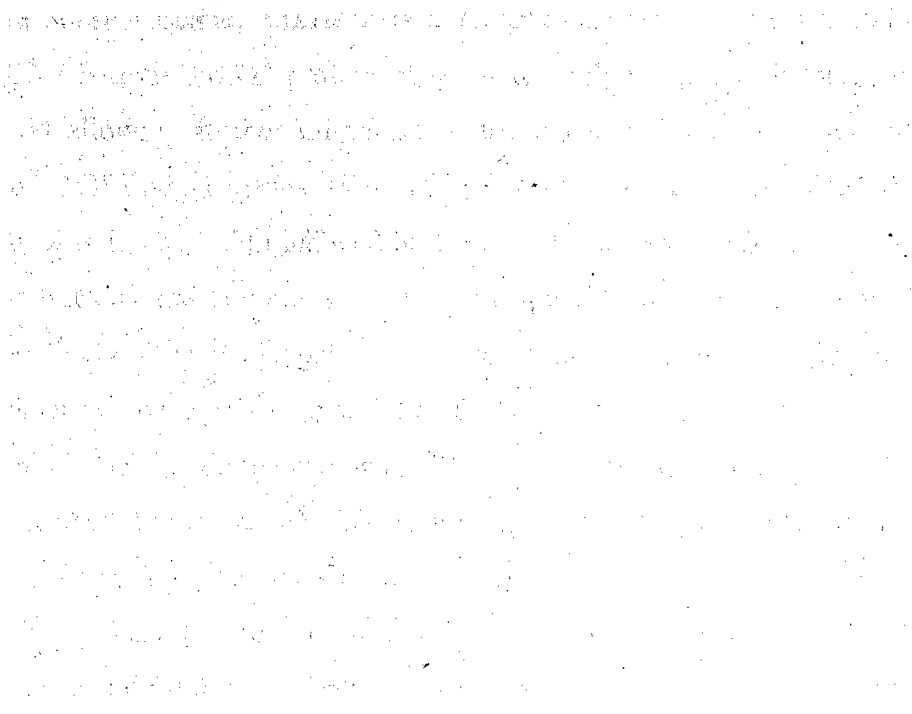
The observed frequency of deletions is consistent with this, with a higher rate in the more advanced stage III cancers compared to the stage II cancers. This increase is however relatively small when including all detected deletions from 50% in stage II cancers to 61% in stage III cancers (χ^2, p=0.376). The relative increase in frequency is greater when limiting analysis to the sub-group of deletions incorporating the minimal deletions; the frequency for deletion 2 increases from 28% to 42% (χ^2, p=0.201), with the greatest relative increase in frequency is for deletion 1.1/1.2 from 33% in stage II cancers to 55% in stage III cancers (χ^2, p=0.076) (Table 3.4). This is interesting as it suggests deletion 1.1/1.2 may more closely reflect disease stage and therefore survival and may potentially prove to be a useful prognostic marker.

The evidence points to deletion 1.1/1.2 as the most likely causative minimal region of deletion but do the genes in this region match up as more likely candidate tumour suppressor genes?

A limitation to the power of defining a minimal region of deletion in this thesis is the relatively high frequency of large deletions in the sample group. Samples with a lower frequency of genome wide copy number loss and an identified chromosome 18 deletion are more likely to be informative in defining a minimal region of deletion.

The evidence that MSI status and stage of tumour is linked to frequency of copy number loss could be utilised to select tumours for analysis. If this study were to be repeated with access to a large sample population such molecular and clinical data would be useful to select a sub-group that would provide more informative results.

Relating frequency of deletion identified in this thesis to pathological stage of tumours has suggested some regions of deletion may be more significant and relating the results to other data such as MSI and tumour suppressor gene mutation status would be useful. Some of this data is available but incomplete (appendix 1). The MSI and ploidy status of the cell lines is relatively well known. It is interesting to observe that 2 of the cell lines defining minimal regions of deletion RKO and VACO 5 are MSI +ve (only 10 of the 47 cell lines are MSI +ve). These cell lines also appear to be relatively chromosomally stable in terms of ploidy status, VACO 5 being diploid and C70 hypotriploid.



4.2 CANDIDATE GENES IN DELETION REGIONS

4.2.1 Deletion 1.1

SMAD 7

Within the minimal regions of deletion identified in this study a strong candidate for the causative tumour suppressor gene is *SMAD 7* as it is within deletion 1.1. SMAD 7 is a member of the SMAD family of transcription factors, that is part of the tumour growth factor (TGF) β signalling pathway, a potent inhibitor of cell growth, inducer of apoptosis [273] and important in colorectal carcinogenesis.

One previous study investigated SMAD copy number, using quantitative PCR with primers designed specifically for *SMAD 7, SMAD 2, SMAD 4* and *DCC*, this reported deletions in 48%, 64%, 66% and 59% of CRCs respectively. Interestingly in the subgroup of cancers with deletions of only part of the region spanned by all 4 genes investigated (excluding multiple discontinuous deletions and amplifications), *SMAD 7* had the highest deletion rate of 70%, *SMAD 4* 68%, *SMAD 2* 43% and *DCC* 30% [274].

Chromosome 18q21 contains a gene for each of the three SMAD sub-families involved in the TGFβ response cascade. *SMAD 2* is a receptor activated SMAD protein (r-SMAD); *SMAD 4* is a co-SMAD protein which heteropolymerizes with the receptor activated SMAD allowing further migration to the nucleus where they can induce transcription of TGFβ target genes (Fig 4.2). *SMAD 7* is a target gene of the TGFβ cascade, with transcription induced by the receptor/co-SMAD complex when it migrates to the nucleus and complexes with tissue specific transcription factors (Fig 4.2) [273]. Relatively little investigation has been undertaken into the function of SMAD 7 but what evidence exists reports 2 different functions both an inhibition and induction effect of TGFβ-mediated apoptosis. When associated with inhibition, SMAD 7 has been described as an anti-SMAD, inhibiting receptor SMAD activation and therefore acting as a negative feedback on the TGFβ cascade [274]; if this was the sole function of *SMAD 7* loss of function through deletion would in theory increase TGFβ mediated apoptosis and inhibition of cell growth. There is however contrary evidence that shows, higher levels of SMAD7 increasing TGFβ mediated apoptosis [275]; if this is the case loss of function through deletion would reduce TGFβ mediated apoptosis

and inhibition of cell growth and so promote carcinogenesis. In the only published prognostic study, *SMAD 7* deletions were associated with a better prognosis in colorectal cancer [276]. To date no sequence mutation analysis has been reported on SMAD 7 in CRC.

The published literature has yet to clearly establish the function of SMAD 7; it is possible that it promotes apoptosis and inhibition of cell growth with a secondary negative feedback function.

This study has detected a cell line in which the only deletion on chromosome 18 was 478 kb in size, incorporating SMAD 7 and 2 other unlikely candidate genes, this deletion was also identified in a primary cancer although not as an exclusive deletion. This study has also identified a high percentage of SMAD 7 deletions, 35 of 47 (74%) cell lines and 31 of 69 (45%) primary cancers. It adds to the weight of evidence from previous publications that the area of chromosome 18 in the immediate vicinity of SMAD 7 is an important region of deletion. What can not be concluded from deletion data even with a tiling path array is which of the genes in this region is more likely to be causative, SMAD 4 is located within 3Mb, SMAD 2 and DCC within 6Mb of SMAD 7; because of the relatively large size of the majority of detected deletions and the proximity of these candidate genes they are all found to be deleted at a relatively high frequency in this study.

The region containing all 4 of these candidate genes was deleted in 28 of 47 (60%) cell lines and 17 of 69 (25%) primary cancers in this study.

SMAD 4 was deleted in 35 of 47 (74%) cell lines and 28 of 69 (41%) primary cancers in this study, nearly as high a frequency as SMAD 7, but not all the SMAD 4 deletions were in the same tumours and there was not an example of a single discrete SMAD 4 deletion in any tumour.

The is some evidence from this study to suggest that SMAD 7 is a more likely causative gene than SMAD 4 but no clear conclusion can be made. What is clear is the importance of the TGF signalling pathway and perhaps all the SMAD genes in this region.

Importantly a recent publication in a high impact journal reporting a genome-wide association study, identified polymorphisms within SMAD 7 that caused predisposition to CRC [277].This study, genotyping SNPs in 940 individuals with familial colorectal tumours and 965 controls, identified 3 SNPs in SMAD 7 associated

with CRC, 1 of these SMAD 7 SNPs was the most highly significant along with a SNP mapping to 8q 24.21. Twelve sequence changes (6 known SNPs and 6 previously unknown changes) were identified on resequencing the coding region of SMAD 7 in 65 individuals (35 carrying high risk haplotypes and 30 non carriers). This genome wide association study using SNPs identifies in addition to our microarray CGH study an important association between SMAD 7 and CRC but via a different method and identifying a sequence change rather than a deletion. This clearly provides further strong evidence for SMAD 7 being a causative tumour suppressor gene in CRC. This evidence from this thesis and the genome-wide association study (published in Nov 2007) is original; only 1 other study had linked SMAD 7 with CRC and this was SMAD 7 deletions associated with a better prognosis in colorectal cancer [276].

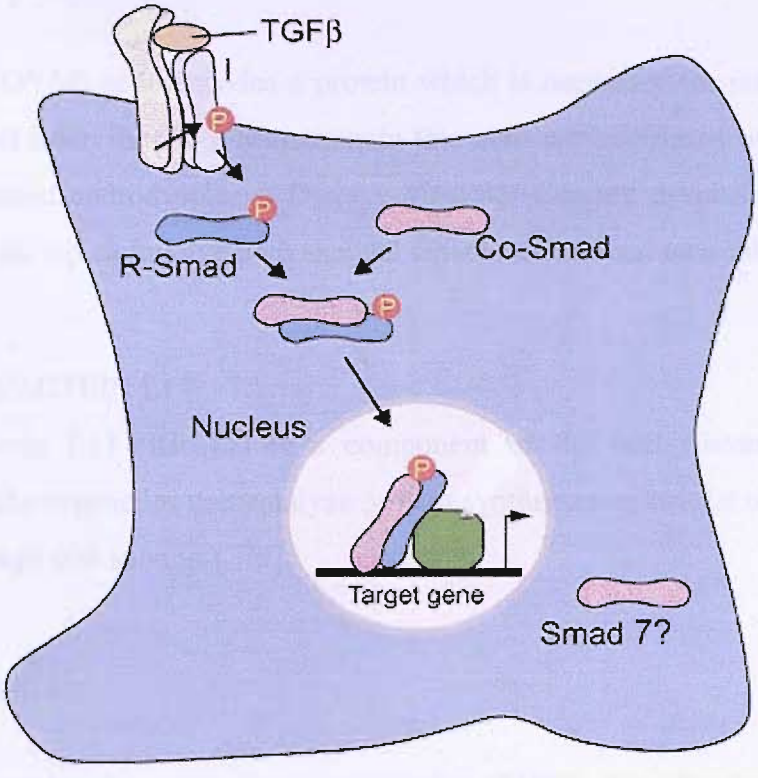


Fig. 4.2: SMAD 7 and the TGFβ signal-transduction pathway.

Activation of the TGFβ receptor complex, phosphorylates (P) specific r-SMADS (eg. SMAD 2). These r-SMADS associate with the co- SMADS (eg. SMAD4) and then translocate to the nucleus, where they regulate gene expression. SMAD 7 is one of these target genes and therefore is potentially involved in apoptosis and inhibition of cell growth. SMAD 7 may also have a negative feedback role, inhibiting r-SMAD activation.

Deletion 1.1 in addition to *SMAD 7* contains the genes dymeclin (*DYM*) and ribosomal protein l17 (RPL17) (Table 4.2).

DYMECLIN

The dymeclin (DYM) gene encodes a protein which is necessary for normal skeletal development and brain function. Mutations in this gene are associated with two types of recessive osteochondrodysplasia, Dyggve-Melchior-Clausen dysplasia and Smith-McCort dysplasia, which involve both skeletal defects and mental retardation [278].

RIBOSOMAL PROTEIN L17

Ribosomal protein L17 (RPL17) is a component of the 60S ribosomal subunit. Ribosomes are the organelles that catalyze protein synthesis and consist of a small 40S subunit and a large 60S subunit [279].

4.2.2 Deletion 1.2

Deletion 1.2 contains the genes lipase endothelial (LIPG), mitochondrial acetyl-coa acyltransferase (ACAA2) and myosin VB (MYO5B) (Table 4.2).

LIPASE ENDOTHELIAL

The lipase endothelial (LIPG) gene encodes a protein that is located in the cytoplasm, it has substantial phospholipase activity and may be involved in lipoprotein metabolism [280].

ACETYL-COA ACYLTRANSFERASE

Acetyl-coa acyltransferase (ACAA2) catalyzes the last step of the mitochondrial fatty acid beta-oxidation spiral, the 397-amino acid ACAA2 protein is present in human liver, fibroblasts, and intercostal muscle. It is 1 of 5 known types of thiolase in mammalian tissues, other functions of the thiolases include cholesterol biosynthesis and ketone body metabolism [281].

MYO5B plays a role in plasma membrane recycling [282]. Myosins are molecular motors that, upon interaction with actin filaments, utilize energy from ATP hydrolysis to generate mechanical force [283].

4.2.3 Deletion 2

Cadherin 7 and cadherin 19

Deletion 2 contains 2 cadherin genes, cadherin 7 (CDH7) and cadherin 19 (CDH19) (Table 4.2). The cadherins are a family of cell surface molecules involved in the structural and functional organisation of cells in various tissues. Epithelial cell-cell tight junctions are composed of adherens junctions and desmosomes [284-286]. Cadherins are the major adhesion molecules in the adherens junctions which are important in regulating the activity of the entire junctional complex [285-287]; including selective cell-cell adhesion [288, 289]. As disturbance of intercellular adhesion is a prerequisite for invasion and metastasis of tumour cells, cadherins are considered prime candidates for tumour suppressor genes [52].

Cadherin genes have been mapped to regions of LOH in tumours; chromosome 16q shows LOH in a variety of solid tumours. Five cadherin genes (CDH1,3, 5, 8 and 11) have been mapped to 16q21-q22.1 which demonstrates LOH in breast [290] and prostate cancers [291, 292] (Table 4.1).

Three cadherin genes have been mapped to chromosome 18q (CDH 2, 7 and 19) but not previously directly linked to CRC. CDH 7 LOH is evident in head and neck squamous cell carcinoma (Table 4.1)

CDH7 encodes a membrane protein containing 5 extra cellular calcium-binding repeats, a transmembrane domain, and a cytoplasmic tail. CDH7 is classified as an atypical (type II) cadherin as it lacks the specific cell adhesion recognition sequence of classic (type I) cadherins [293]. Expression of CDH7 has been detected in the testis, brain, prostate but not in mesenchymal or epithelial cell lines [293].

CDH19 is also an atypical (type II) cadherin. CDH19 expression has been detected in a wide variety of tissues, with the exception of the uterus [293, 294].

Gene symbol	Alias	Chromosome location	Associated tumour LOH
CDH1	E-cadherin	16q22.1	Breast/prostate
CDH2	N-cadherin	18q12.1	
CDH3	P-cadherin	16q22.1	Breast/prostate
CDH4	R-cadherin		
CDH5	VE-cadherin	16q22.1	Breast/prostate
CDH6/7	K-cadherin	18q22-q23	Squamous cell cacinoma
CDH8		16q22.1	Breast/prostate
CDH11	OB-cadherin	16q22.1	Breast/prostate
CDH12	Br-cadherin	5p13-p14	
CDH13	H-cadherin	16q24.2-q24.3	
CDH15	M-cadherin	16q24.3	
CDH16	LI-cadherin	8q22.1	Acute myeloid leukaemia
CDH19		18q22-q23	

Table 4.1: The cadherin genes, their chromosomal location and tumours in which they have LOH.

Modified from kremmidiotis, G; 1998 Genomics [294]

CDH 1 or E-cadherin (expressed in the epithelium) is known to be associated with the CRC tumour suppressor gene APC. The APC protein degrades the oncogene B-catenin which is involved with cell-cell adhesion by linking the cell adhesion molecule CDH1 (E-cadherin) to the actin cytoskeleton via other cadherins [50-52].

Transforming growth factor β (TGF-β) is known for suppressing the growth of many normal epithelial cells. It is interesting to note that TGF-β also has tumour-promoting activity, owing to its ability to influence cell adhesion [273, 295-297]. TGF-β has been observed to have some influence on the cadherins, it is associated with a decreased expression of CDH1 (E-cadherin) and an increased expression of CDH 2 (N-cadherin) [298, 299]. It is interesting to note the link between the TGF-β pathway and chromosome 18 via both the SMAD and the CADHERIN genes.

The minimal regions of deletion identified by the tiling path array-CGH on chromosome 18 contain some good candidate tumour suppressor genes primarily SMAD 7 but also CDH 7 and CDH 19.

In order to confirm whether these or any other gene in the deletion regions are involved in CRC tumourigenesis, further investigation would be needed such as mutation screening of individual genes.

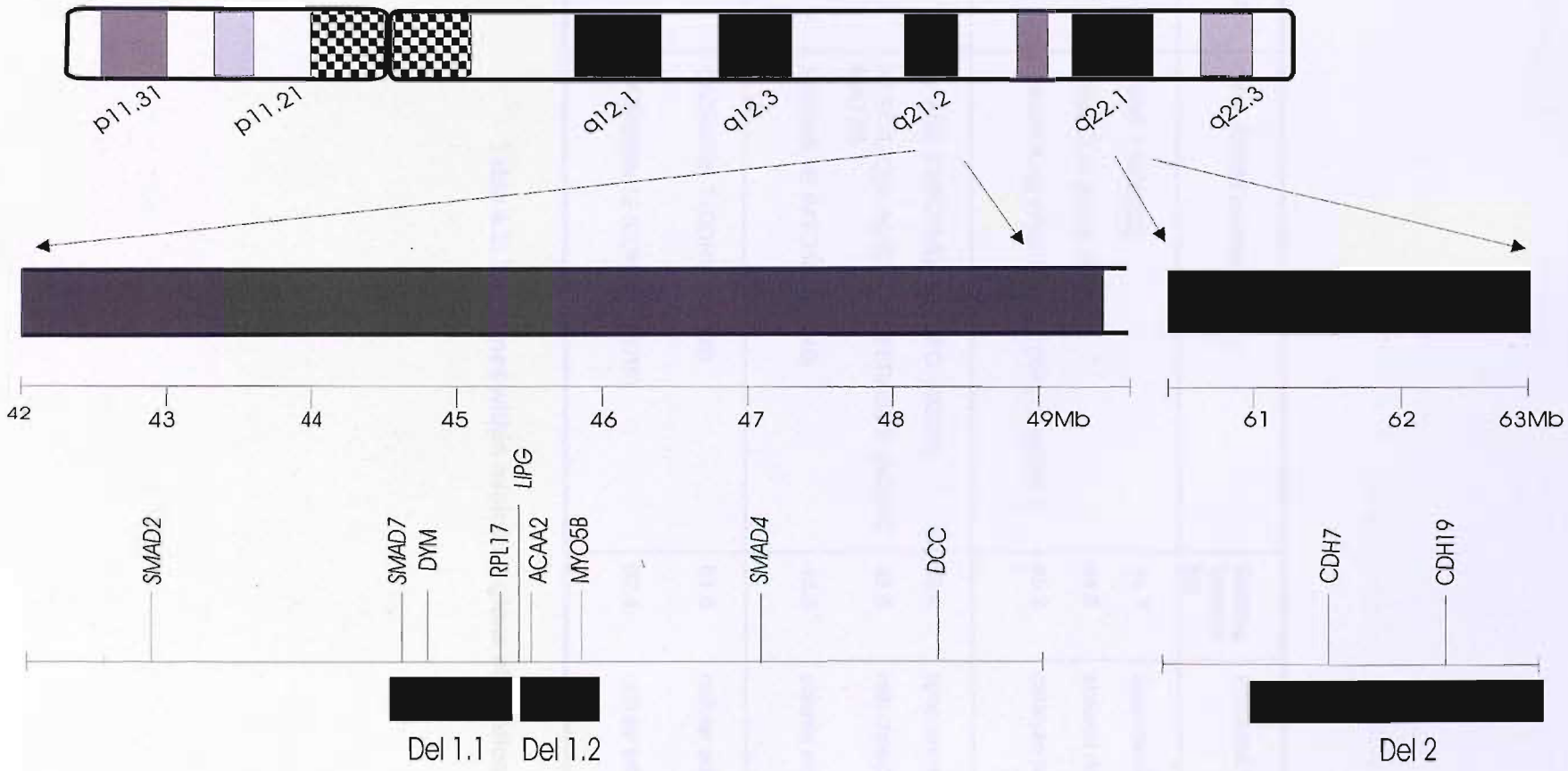


Fig. 4.3: Summary of minimal regions of deletion and corresponding location of candidate tumour suppressor genes

The regions of chromosome 18 containing the minimal regions of deletion have been enlarged in this diagram to enable the location of candidate tumour suppressor genes to be shown. All genes located within the minimal regions of deletion identified by this thesis are shown in grey (Table 4.2 includes details of full names, OMIM number, location and function) and other candidate tumour suppressor genes proposed from the minimal region of deletion previously defined by LOH are shown in green.

Deletion	Gene (OMIM number)	Starting location Mb	Predicted Function
Del. 1.1	SMAD 7 (602932)	44.7	apoptosis/inhibition of cell growth
	DYMECLIN (DYM, 607461)	44.8	skeletal development and brain function
	RIBOSOMAL PROTEIN L17 (RPL17, 603661)	45.2	catalyze protein synthesis
Del 1.2	LIPASE, ENDOTHELIAL (LIPG, 603684)	45.4	lipoprotein metabolism
	ACETYL-CoA ACYLTRANSFERASE 2; (ACAA2, 604770)	45.5	mitochondrial fatty acid synthesis
	MYOSIN VB (MYO5B, 606540)	45.6	plasma membrane recycling
Del 2	CADHERIN 7 (CDH7, 605806)	61.6	cellular adhesion
	CADHERIN 19 (CDH19, 603016)	62.4	cellular adhesion

Table 4.2: The genes within minimal regions of deletion 1.1, 1.2 and 2

4.3 FUTURE WORK

4.3.1 Direct continuation of work

Further supporting evidence for the minimal regions of deletion identified could be achieved by analysing a larger number of cell lines and primary CRCs utilising the constructed chromosome 18 tiling path array. However large numbers of suitable samples are not readily available, especially as the tissue from primary CRCs needs to be fresh frozen in liquid nitrogen. Alternatively a variation of the same technique, SNP array could be utilised to identify minimal regions of deletion [227-230].

The obvious continuation of this work is mutation screening, sequencing the genes that are located within the minimal regions of deletion, the high resolution of the array CGH tiling path array has made this a realistic objective because it has narrowed down the number of candidates. There are only 8 genes located within these regions and in terms of function SMAD 7 and CADHERINS 7 and 19 would be the first genes to be sequenced. The detection of *de novo* nucleotide changes in tumour DNA that predict altered function would be considered very strong evidence in favour of them being causative genes in CRC tumourigenesis. An additional approach would be to investigate the affect in an animal model, removing gene function, most commonly achieved in 'knock out' mice.

4.3.2 Role of tiling path array CGH to refine prognostic data

The chromosome 18 array from this project offers a vastly greater resolution with 860 markers compared to the 3 or 4 that have previously been used with LOH techniques.

Conventional LOH techniques have assessed chromosome 18 LOH as a prognostic marker, 9 of 14 studies [20, 100, 101, 108, 115-117, 119, 120] have linked chromosome 18 LOH with shorter survival; a meta-analysis has reported an overall association with shorter survival but with significant heterogeneity between the studies [90]. More robust data is required for chromosome 18 deletions to be incorporated as a prognostic marker into prospective trials in Europe. The higher resolution of a chromosome 18 tiling path array might provide a useful refinement in linking chromosome 18 allelic loss with survival.

4.3.3 Genome wide tiling path array

Little change is required to this technique to extend it to a genome wide tiling path array, the major limitation is the high initial cost of producing the large numbers of clonal BAC DNA to represent the genome. Such a genome wide tiling path could be utilised to screen the constitutional DNA of high risk individuals who are excluded from carrying the major inherited autosomal dominant conditions, FAP and HNPCC.

CONCLUSIONS

The work for this thesis has resulted in the successful construction of a CGH chromosome 18 tiling path array, effective in screening the entire chromosome and identifying a large number of deletions. The resolution of the tiling path array has defined 2 minimal regions of deletion, which have been confirmed by a second method (FISH) and contain 3 strong candidate tumour suppressor genes which may be causative in colorectal carcinogenesis, SMAD 7, CADHERIN 7 and CADHERIN 19.

REFERENCES

1. Jass, J.R., et al., *Emerging concepts in colorectal neoplasia.* Gastroenterology, 2002. **123**(3): p. 862-76.
2. Knudson, A.G., Jr., *Hereditary cancer, oncogenes, and antioncogenes.* Cancer Res, 1985. **45**(4): p. 1437-43.
3. Spigelman, A.D., S.V. Hodgson, and J.P. Thomson, *Management of familial adenomatous polyposis.* Baillieres Clin Gastroenterol, 1992. **6**(1): p. 75-94.
4. Bussey, H.J., *Familial polyposis coli.* Pathol Annu, 1979. **14 Pt 1**: p. 61-81.
5. Groves, C.J., et al., *Duodenal cancer in patients with familial adenomatous polyposis (FAP): results of a 10 year prospective study.* Gut, 2002. **50**(5): p. 636-41.
6. McKay, D.L., *Congenital hypertrophy of the retinal pigment epithelium and familial adenomatous polyposis.* Aust N Z J Ophthalmol, 1993. **21**(1): p. 3-6.
7. Sturt, N.J. and S.K. Clark, *Current ideas in desmoid tumours.* Fam Cancer, 2006. **5**(3): p. 275-85; discussion 287-8.
8. Clark, S.K., et al., *Desmoid tumours complicating familial adenomatous polyposis.* Br J Surg, 1999. **86**(9): p. 1185-9.
9. Clark, S.K. and R.K. Phillips, *Desmoids in familial adenomatous polyposis.* Br J Surg, 1996. **83**(11): p. 1494-504.
10. Brett, M. and M. Hershman, *Transanal endoscopic microsurgery.* Br J Hosp Med, 1994. **52**(8): p. 386, 389-94.
11. Wu, J.S., et al., *APC genotype, polyp number, and surgical options in familial adenomatous polyposis.* Ann Surg, 1998. **227**(1): p. 57-62.
12. Vasen, H.F., et al., *Molecular genetic tests as a guide to surgical management of familial adenomatous polyposis.* Lancet, 1996. **348**(9025): p. 433-5.
13. Niessen, R.C., et al., *Identification of mismatch repair gene mutations in young patients with colorectal cancer and in patients with multiple tumours associated with hereditary non-polyposis colorectal cancer.* Gut, 2006. **55**(12): p. 1781-8.
14. Jass, J.R., *Hereditary Non-Polyposis Colorectal Cancer: the rise and fall of a confusing term.* World J Gastroenterol, 2006. **12**(31): p. 4943-50.
15. Peltomaki, P. and A. de la Chapelle, *Mutations predisposing to hereditary nonpolyposis colorectal cancer.* Adv Cancer Res, 1997. **71**: p. 93-119.
16. Papadopoulos, N. and A. Lindblom, *Molecular basis of HNPCC: mutations of MMR genes.* Hum Mutat, 1997. **10**(2): p. 89-99.
17. Moore, J. and P. Cowled, *Hereditary non-polyposis colorectal cancer syndrome.* Aust N Z J Surg, 1999. **69**(1): p. 6-13.
18. Herman, J.G., et al., *Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma.* Proc Natl Acad Sci U S A, 1998. **95**(12): p. 6870-5.
19. Wheeler, J.M., et al., *Mechanisms of inactivation of mismatch repair genes in human colorectal cancer cell lines: the predominant role of hMLH1.* Proc Natl Acad Sci U S A, 1999. **96**(18): p. 10296-301.
20. Jen, J., et al., *Allelic loss of chromosome 18q and prognosis in colorectal cancer.* N Engl J Med, 1994. **331**(4): p. 213-21.
21. Rodrigues, N.R., et al., *p53 mutations in colorectal cancer.* Proc Natl Acad Sci U S A, 1990. **87**(19): p. 7555-9.
22. Muto, T., H.J. Bussey, and B.C. Morson, *The evolution of cancer of the colon and rectum.* Cancer, 1975. **36**(6): p. 2251-70.
23. Fearon, E.R. and B. Vogelstein, *A genetic model for colorectal tumorigenesis.* Cell, 1990. **61**(5): p. 759-67.
24. Diaz-Llopis, M. and J.L. Menezo, *Congenital hypertrophy of the retinal pigment epithelium in familial adenomatous polyposis.* Arch Ophthalmol, 1988. **106**(3): p. 412-3.
25. Jagelman, D.G., J.J. DeCosse, and H.J. Bussey, *Upper gastrointestinal cancer in familial adenomatous polyposis.* Lancet, 1988. **1**(8595): p. 1149-51.
26. Herrera, L., et al., *Gardner syndrome in a man with an interstitial deletion of 5q.* Am J Med Genet, 1986. **25**(3): p. 473-6.
27. Bodmer, W.F., et al., *Localization of the gene for familial adenomatous polyposis on chromosome 5.* Nature, 1987. **328**(6131): p. 614-6.

28. Leppert, M., et al., *The gene for familial polyposis coli maps to the long arm of chromosome 5.* Science, 1987. **238**(4832): p. 1411-3.
29. Bos, J.L., et al., *Prevalence of ras gene mutations in human colorectal cancers.* Nature, 1987. **327**(6120): p. 293-7.
30. Forrester, K., et al., *Detection of high incidence of K-ras oncogenes during human colon tumorigenesis.* Nature, 1987. **327**(6120): p. 298-303.
31. Muleris, M., et al., *Consistent deficiencies of chromosome 18 and of the short arm of chromosome 17 in eleven cases of human large bowel cancer: a possible recessive determinism.* Ann Genet, 1985. **28**(4): p. 206-13.
32. Reichmann, A., P. Martin, and B. Levin, *Chromosomal banding patterns in human large bowel adenomas.* Hum Genet, 1985. 70(1): p. 28-31.
33. Smith, G., et al., *Mutations in APC, Kirsten-ras, and p53--alternative genetic pathways to colorectal cancer.* Proc Natl Acad Sci U S A, 2002. **99**(14): p. 9433-8.
34. Georgiades, I.B., et al., *Heterogeneity studies identify a subset of sporadic colorectal cancers without evidence for chromosomal or microsatellite instability.* Oncogene, 1999. **18**(56): p. 7933-40.
35. Curtis, L.J., et al., *Specific patterns of chromosomal abnormalities are associated with RER status in sporadic colorectal cancer.* J Pathol, 2000. **192**(4): p. 440-5.
36. Lin, J.K., et al., *Loss of heterozygosity and DNA aneuploidy in colorectal adenocarcinoma.* Ann Surg Oncol, 2003. **10**(9): p. 1086-94.
37. Hawkins, N.J., et al., *Microsatellite-stable diploid carcinoma: a biologically distinct and aggressive subset of sporadic colorectal cancer.* Br J Cancer, 2001. **84**(2): p. 232-6.
38. Kinzler, K.W. and B. Vogelstein, *Lessons from hereditary colorectal cancer.* Cell, 1996. **87**(2): p. 159-70.
39. Aberle, H., et al., *beta-catenin is a target for the ubiquitin-proteasome pathway.* Embo J, 1997. **16**(13): p. 3797-804.
40. Korinek, V., et al., *Two members of the Tcf family implicated in Wnt/beta-catenin signaling during embryogenesis in the mouse.* Mol Cell Biol, 1998. **18**(3): p. 1248-56.
41. Barker, N., P.J. Morin, and H. Clevers, *The Yin-Yang of TCF/beta-catenin signaling.* Adv Cancer Res, 2000. **77**: p. 1-24.
42. Peifer, M. and P. Polakis, *Wnt signaling in oncogenesis and embryogenesis--a look outside the nucleus.* Science, 2000. **287**(5458): p. 1606-9.
43. Barker, N. and H. Clevers, *Catenins, Wnt signaling and cancer.* Bioessays, 2000. **22**(11): p. 961-5.
44. Korinek, V., et al., *Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC-/- colon carcinoma.* Science, 1997. **275**(5307): p. 1784-7.
45. He, T.C., et al., *Identification of c-MYC as a target of the APC pathway.* Science, 1998. **281**(5382): p. 1509-12.
46. Marcu, K.B., S.A. Bossone, and A.J. Patel, *myc function and regulation.* Annu Rev Biochem, 1992. **61**: p. 809-60.
47. Sherr, C.J., *Cancer cell cycles.* Science, 1996. **274**(5293): p. 1672-7.
48. Bartkova, J., et al., *Cell cycle-related variation and tissue-restricted expression of human cyclin D1 protein.* J Pathol, 1994. **172**(3): p. 237-45.
49. Arber, N., et al., *Increased expression of cyclin D1 is an early event in multistage colorectal carcinogenesis.* Gastroenterology, 1996. **110**(3): p. 669-74.
50. Ben-Ze'ev, A. and B. Geiger, *Differential molecular interactions of beta-catenin and plakoglobin in adhesion, signaling and cancer.* Curr Opin Cell Biol, 1998. **10**(5): p. 629-39.
51. Willert, K. and R. Nusse, *Beta-catenin: a key mediator of Wnt signaling.* Curr Opin Genet Dev, 1998. **8**(1): p. 95-102.
52. Behrens, J., *Cadherins and catenins: role in signal transduction and tumor progression.* Cancer Metastasis Rev, 1999. **18**(1): p. 15-30.
53. Miyoshi, Y., et al., *Somatic mutations of the APC gene in colorectal tumors: mutation cluster region in the APC gene.* Hum Mol Genet, 1992. **1**(4): p. 229-33.
54. Powell, S.M., et al., *APC mutations occur early during colorectal tumorigenesis.* Nature, 1992. **359**(6392): p. 235-7.
55. Potter, J.D., *Colorectal cancer: molecules and populations.* J Natl Cancer Inst, 1999. **91**(11): p. 916-32.
56. Soengas, M.S., et al., *Inactivation of the apoptosis effector Apaf-1 in malignant melanoma.* Nature, 2001. **409**(6817): p. 207-11.

57. Papadopoulos, N., et al., *Mutation of a mutL homolog in hereditary colon cancer.* Science, 1994. **263**(5153): p. 1625-9.
58. Buermeyer, A.B., et al., *Mammalian DNA mismatch repair.* Annu Rev Genet, 1999. **33**: p. 533-64.
59. Thomas, D.C., A. Umar, and T.A. Kunkel, *Microsatellite instability and mismatch repair defects in cancer.* Mutat Res, 1996. **350**(1): p. 201-5.
60. Modrich, P. and R. Lahue, *Mismatch repair in replication fidelity, genetic recombination, and cancer biology.* Annu Rev Biochem, 1996. **65**: p. 101-33.
61. Aaltonen, L.A., et al., *Clues to the pathogenesis of familial colorectal cancer.* Science, 1993. **260**(5109): p. 812-6.
62. Thibodeau, S.N., G. Bren, and D. Schaid, *Microsatellite instability in cancer of the proximal colon.* Science, 1993. **260**(5109): p. 816-9.
63. Thibodeau, S.N., et al., *Altered expression of hMSH2 and hMLH1 in tumors with microsatellite instability and genetic alterations in mismatch repair genes.* Cancer Res, 1996. **56**(21): p. 4836-40.
64. Borresen, A.L., et al., *Somatic mutations in the hMSH2 gene in microsatellite unstable colorectal carcinomas.* Hum Mol Genet, 1995. **4**(11): p. 2065-72.
65. Bubb, V.J., et al., *Microsatellite instability and the role of hMSH2 in sporadic colorectalcancer.* Oncogene, 1996. **12**(12): p. 2641-9.
66. Herman, J.G., et al., *Inactivation of the CDKN2/p16/MTS1 gene is frequently associated with aberrant DNA methylation in all common human cancers.* Cancer Res, 1995. **55**(20): p. 4525-30.
67. Merlo, A., et al., *5' CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers.* Nat Med, 1995. **1**(7): p. 686-92.
68. Herman, J.G., et al., *Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma.* Proc Natl Acad Sci U S A, 1994. **91**(21): p. 9700-4.
69. Popescu, N.C., et al., *Chromosomal localization of three human ras genes by in situ molecular hybridization.* Somat Cell Mol Genet, 1985. **11**(2): p. 149-55.
70. Adjei, A.A., *Ras signaling pathway proteins as therapeutic targets.* Curr Pharm Des, 2001. **7**(16): p. 1581-94.
71. Bos, J.L., *ras oncogenes in human cancer: a review.* Cancer Res, 1989. **49**(17): p. 4682-9.
72. Andreyev, H.J., et al., *Kirsten ras mutations in patients with colorectal cancer: the multicenter "RASCAL" study.* J Natl Cancer Inst, 1998. **90**(9): p. 675-84.
73. Levine, A.J., *p53, the cellular gatekeeper for growth and division.* Cell, 1997. **88**(3): p. 323-31.
74. Harada, T., et al., *Detection of genetic alterations in pancreatic cancers by comparative genomic hybridization coupled with tissue microdissection and degenerate oligonucleotide primed polymerase chain reaction.* Oncology, 2002. **62**(3): p. 251-8.
75. Gorunova, L., et al., *Cytogenetic analysis of pancreatic carcinomas: intratumor heterogeneity and nonrandom pattern of chromosome aberrations.* Genes Chromosomes Cancer, 1998. **23**(2): p. 81-99.
76. Strong, L.C., M. Stine, and T.L. Norsted, *Cancer in survivors of childhood soft tissue sarcoma and their relatives.* J Natl Cancer Inst, 1987. **79**(6): p. 1213-20.
77. Li, F.P., et al., *A cancer family syndrome in twenty-four kindreds.* Cancer Res, 1988. **48**(18): p. 5358-62.
78. Schneider, G. and R.M. Schmid, *Genetic alterations in pancreatic carcinoma.* Mol Cancer, 2003. **2**: p. 15.
79. Evans, S.C. and G. Lozano, *The Li-Fraumeni syndrome: an inherited susceptibility to cancer.* Mol Med Today, 1997. **3**(9): p. 390-5.
80. Baker, S.J., et al., *p53 gene mutations occur in combination with 17p allelic deletions as late events in colorectal tumorigenesis.* Cancer Res, 1990. **50**(23): p. 7717-22.
81. Vogelstein, B. and K.W. Kinzler, *p53 function and dysfunction.* Cell, 1992. **70**(4): p. 523-6.
82. Maughan, N.J. and P. Quirke, *Pathology--a molecular prognostic approach.* Br Med Bull, 2002. **64**: p. 59-74.
83. Wingo, P.A., et al., *Long-term trends in cancer mortality in the United States, 1930-1998.* Cancer, 2003. **97**(12 Suppl): p. 3133-275.
84. *Efficacy of adjuvant fluorouracil and folinic acid in colon cancer. International Multicentre Pooled Analysis of Colon Cancer Trials (IMPACT) investigators.* Lancet, 1995. **345**(8955): p. 939-44.

85. Mamounas, E., et al., *Comparative efficacy of adjuvant chemotherapy in patients with Dukes' B versus Dukes' C colon cancer: results from four National Surgical Adjuvant Breast and Bowel Project adjuvant studies (C-01, C-02, C-03, and C-04).* J Clin Oncol, 1999. **17**(5): p. 1349-55.
86. *Efficacy of adjuvant fluorouracil and folinic acid in B2 colon cancer. International Multicentre Pooled Analysis of B2 Colon Cancer Trials (IMPACT B2) Investigators.* J Clin Oncol, 1999. **17**(5): p. 1356-63.
87. Ilyas, M., et al., *Genetic pathways in colorectal and other cancers.* Eur J Cancer, 1999. **35**(14): p. 1986-2002.
88. Obrand, D.I. and P.H. Gordon, *Incidence and patterns of recurrence following curative resection for colorectal carcinoma.* Dis Colon Rectum, 1997. **40**(1): p. 15-24.
89. Popat, S., R. Hubner, and R.S. Houlston, *Systematic review of microsatellite instability and colorectal cancer prognosis.* J Clin Oncol, 2005. **23**(3): p. 609-18.
90. Popat, S. and R.S. Houlston, *A systematic review and meta-analysis of the relationship between chromosome 18q genotype, DCC status and colorectal cancer prognosis.* Eur J Cancer, 2005. **41**(14): p. 2060-70.
91. Popat, S., A. Matakidou, and R.S. Houlston, *Thymidylate synthase expression and prognosis in colorectal cancer: a systematic review and meta-analysis.* J Clin Oncol, 2004. **22**(3): p. 529-36.
92. Lynch, H.T., et al., *Genetics, natural history, tumor spectrum, and pathology of hereditary nonpolyposis colorectal cancer: an updated review.* Gastroenterology, 1993. **104**(5): p. 1535-49.
93. Farrington, S.M., et al., *Evidence for an age-related influence of microsatellite instability on colorectal cancer survival.* Int J Cancer, 2002. **98**(6): p. 844-50.
94. Liang, J.T., et al., *High-frequency microsatellite instability predicts better chemosensitivity to high-dose 5-fluorouracil plus leucovorin chemotherapy for stage IV sporadic colorectal cancer after palliative bowel resection.* Int J Cancer, 2002. **101**(6): p. 519-25.
95. Wright, C.M., et al., *Prognostic significance of extensive microsatellite instability in sporadic clinicopathological stage C colorectal cancer.* Br J Surg, 2000. **87**(9): p. 1197-202.
96. Gryfe, R., et al., *Tumor microsatellite instability and clinical outcome in young patients with colorectal cancer.* N Engl J Med, 2000. **342**(2): p. 69-77.
97. Colombino, M., et al., *Prevalence and prognostic role of microsatellite instability in patients with rectal carcinoma.* Ann Oncol, 2002. **13**(9): p. 1447-53.
98. Guidoboni, M., et al., *Microsatellite instability and high content of activated cytotoxic lymphocytes identify colon cancer patients with a favorable prognosis.* Am J Pathol, 2001. **159**(1): p. 297-304.
99. Ward, R.L., et al., *Adverse prognostic effect of methylation in colorectal cancer is reversed by microsatellite instability.* J Clin Oncol, 2003. **21**(20): p. 3729-36.
100. Watanabe, T., et al., *Molecular predictors of survival after adjuvant chemotherapy for colon cancer.* N Engl J Med, 2001. **344**(16): p. 1196-206.
101. Barratt, P.L., et al., *DNA markers predicting benefit from adjuvant fluorouracil in patients with colon cancer: a molecular study.* Lancet, 2002. **360**(9343): p. 1381-91.
102. Chen, W.S., et al., *Microsatellite instability in sporadic-colon-cancer patients with and without liver metastases.* Int J Cancer, 1997. **74**(4): p. 470-4.
103. Choi, S.W., et al., *Genetic classification of colorectal cancer based on chromosomal loss and microsatellite instability predicts survival.* Clin Cancer Res, 2002. **8**(7): p. 2311-22.
104. Wang, C., et al., *Prognostic significance of microsatellite instability and Ki-ras mutation type in stage II colorectal cancer.* Oncology, 2003. **64**(3): p. 259-65.
105. Ribic, C.M., et al., *Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer.* N Engl J Med, 2003. **349**(3): p. 247-57.
106. Elsaleh, H. and B. Iacopetta, *Microsatellite instability is a predictive marker for survival benefit from adjuvant chemotherapy in a population-based series of stage III colorectal carcinoma.* Clin Colorectal Cancer, 2001. **1**(2): p. 104-9.
107. Liang, J.T., et al., *Clinicopathologic and carcinogenetic appraisal of DNA replication error in sporadic T3N0M0 stage colorectal cancer after curative resection.* Hepatogastroenterology, 1999. **46**(26): p. 883-90.
108. Halling, K.C., et al., *Microsatellite instability and 8p allelic imbalance in stage B2 and C colorectal cancers.* J Natl Cancer Inst, 1999. **91**(15): p. 1295-303.
109. Curran, B., et al., *Replication error phenotype, clinicopathological variables, and patient outcome in Dukes' B stage II (T3,N0,M0) colorectal cancer.* Gut, 2000. **46**(2): p. 200-4.

110. Kochhar, R., et al., *Allelic imbalance and microsatellite instability in resected Duke's D colorectal cancer.* Diagn Mol Pathol, 1997. **6**(2): p. 78-84.
111. Massa, M.J., et al., *Differential prognosis of replication error phenotype and loss of heterozygosity in sporadic colorectal cancer.* Eur J Cancer, 1999. **35**(12): p. 1676-82.
112. Hemminki, A., et al., *Microsatellite instability is a favorable prognostic indicator in patients with colorectal cancer receiving chemotherapy.* Gastroenterology, 2000. **119**(4): p. 921-8.
113. Aschele, C., et al., *Deleted in colon cancer protein expression in colorectal cancer metastases: a major predictor of survival in patients with unresectable metastatic disease receiving palliative fluorouracil-based chemotherapy.* J Clin Oncol, 2004. **22**(18): p. 3758-65.
114. Lindforss, U., et al., *Allelic loss is heterogeneous throughout the tumor in colorectal carcinoma.* Cancer, 2000. **88**(12): p. 2661-7.
115. Carethers, J.M., et al., *Prognostic significance of allelic lost at chromosome 18q21 for stage II colorectal cancer.* Gastroenterology, 1998. **114**(6): p. 1188-95.
116. Dix, B.R., et al., *The common molecular genetic alterations in Dukes' B and C colorectal carcinomas are not short-term prognostic indicators of survival.* Int J Cancer, 1994. **59**(6): p. 747-51.
117. Lanza, G., et al., *Chromosome 18q allelic loss and prognosis in stage II and III colon cancer.* Int J Cancer, 1998. **79**(4): p. 390-5.
118. O'Connell, M.J., et al., *Current status of adjuvant chemotherapy for colorectal cancer. Can molecular markers play a role in predicting prognosis?* Cancer, 1992. **70**(6 Suppl): p. 1732-9.
119. Shibata, D., et al., *The DCC protein and prognosis in colorectal cancer.* N Engl J Med, 1996. **335**(23): p. 1727-32.
120. Zauber, N.P., et al., *Ki-ras gene mutations, LOH of the APC and DCC genes, and microsatellite instability in primary colorectal carcinoma are not associated with micrometastases in pericolonic lymph nodes or with patients' survival.* J Clin Pathol, 2004. **57**(9): p. 938-42.
121. Jernvall, P., et al., *Microsatellite instability: impact on cancer progression in proximal and distal colorectal cancers.* Eur J Cancer, 1999. **35**(2): p. 197-201.
122. Diep, C.B., et al., *Genetic tumor markers with prognostic impact in Dukes' stages B and C colorectal cancer patients.* J Clin Oncol, 2003. **21**(5): p. 820-9.
123. Martinez-Lopez, E., et al., *Allelic loss on chromosome 18q as a prognostic marker in stage II colorectal cancer.* Gastroenterology, 1998. **114**(6): p. 1180-7.
124. Jen, J., et al., *Molecular determinants of dysplasia in colorectal lesions.* Cancer Res, 1994. **54**(21): p. 5523-6.
125. Font, A., et al., *Prognostic value of K-ras mutations and allelic imbalance on chromosome 18q in patients with resected colorectal cancer.* Dis Colon Rectum, 2001. **44**(4): p. 549-57.
126. Ogunbiyi, O.A., et al., *Confirmation that chromosome 18q allelic loss in colon cancer is a prognostic indicator.* J Clin Oncol, 1998. **16**(2): p. 427-33.
127. Saito, M., et al., *Expression of DCC protein in colorectal tumors and its relationship to tumor progression and metastasis.* Oncology, 1999. **56**(2): p. 134-41.
128. Santi, D.V., C.S. McHenry, and H. Sommer, *Mechanism of interaction of thymidylate synthetase with 5-fluorodeoxyuridylate.* Biochemistry, 1974. **13**(3): p. 471-81.
129. Danenberg, P.V., R.J. Langenbach, and C. Heidelberger, *Structures of reversible and irreversible complexes of thymidylate synthetase and fluorinated pyrimidine nucleotides.* Biochemistry, 1974. **13**(5): p. 926-33.
130. Kufe, D.W., et al., *5-Fluoro-2'-deoxyuridine incorporation in L1210 DNA.* J Biol Chem, 1981. **256**(17): p. 8885-8.
131. Danenberg, P.V., et al., *The incorporation of 5-fluoro-2'-deoxyuridine into DNA of mammalian tumor cells.* Biochem Biophys Res Commun, 1981. **102**(2): p. 654-8.
132. Longley, D.B., D.P. Harkin, and P.G. Johnston, *5-fluorouracil: mechanisms of action and clinical strategies.* Nat Rev Cancer, 2003. **3**(5): p. 330-8.
133. Lenz, H.J., et al., *p53 and thymidylate synthase expression in untreated stage II colon cancer: associations with recurrence, survival, and site.* Clin Cancer Res, 1998. **4**(5): p. 1227-34.
134. Lenz, H.J., et al., *p53 point mutations and thymidylate synthase messenger RNA levels in disseminated colorectal cancer: an analysis of response and survival.* Clin Cancer Res, 1998. **4**(5): p. 1243-50.
135. Allegra, C.J., et al., *Investigation of the prognostic and predictive value of thymidylate synthase, p53, and Ki-67 in patients with locally advanced colon cancer.* J Clin Oncol, 2002. **20**(7): p. 1735-43.

136. Nanni, O., et al., *Role of biological markers in the clinical outcome of colon cancer.* Br J Cancer, 2002. **87**(8): p. 868-75.
137. Berglund, A., et al., *Thymidylate synthase and p53 expression in primary tumor do not predict chemotherapy outcome in metastatic colorectal carcinoma.* Anticancer Res, 2002. **22**(6B): p. 3653-9.
138. Gonen, M., et al., *Thymidylate synthase expression in hepatic tumors is a predictor of survival and progression in patients with resectable metastatic colorectal cancer.* J Clin Oncol, 2003. **21**(3): p. 406-12.
139. Johnston, P.G., et al., *Thymidylate synthase protein expression in primary colorectal cancer: lack of correlation with outcome and response to fluorouracil in metastatic disease sites.* J Clin Oncol, 2003. **21**(5): p. 815-9.
140. Edler, D., et al., *Thymidylate synthase expression in colorectal cancer: a prognostic and predictive marker of benefit from adjuvant fluorouracil-based chemotherapy.* J Clin Oncol, 2002. **20**(7): p. 1721-8.
141. Yamachika, T., et al., *A new prognostic factor for colorectal carcinoma, thymidylate synthase, and its therapeutic significance.* Cancer, 1998. **82**(1): p. 70-7.
142. Allegra, C.J., et al., *Prognostic value of thymidylate synthase, Ki-67, and p53 in patients with Dukes' B and C colon cancer: a National Cancer Institute-National Surgical Adjuvant Breast and Bowel Project collaborative study.* J Clin Oncol, 2003. **21**(2): p. 241-50.
143. Paradiso, A., et al., *Thymidilate synthase and p53 primary tumour expression as predictive factors for advanced colorectal cancer patients.* Br J Cancer, 2000. **82**(3): p. 560-7.
144. Shirota, Y., et al., *ERCC1 and thymidylate synthase mRNA levels predict survival for colorectal cancer patients receiving combination oxaliplatin and fluorouracil chemotherapy.* J Clin Oncol, 2001. **19**(23): p. 4298-304.
145. Etienne, M.C., et al., *Prognostic value of tumoral thymidylate synthase and p53 in metastatic colorectal cancer patients receiving fluorouracil-based chemotherapy: phenotypic and genotypic analyses.* J Clin Oncol, 2002. **20**(12): p. 2832-43.
146. Tomiak, A., et al., *Thymidylate synthase expression in stage II and III colon cancer: a retrospective review.* Am J Clin Oncol, 2001. **24**(6): p. 597-602.
147. Benhattar, J., et al., *Prognostic significance of K-ras mutations in colorectal carcinoma.* Gastroenterology, 1993. **104**(4): p. 1044-8.
148. Lee, J.C., et al., *K-ras gene mutation is a useful predictor of the survival of early stage colorectal cancers.* Anticancer Res, 1996. **16**(6B): p. 3839-44.
149. Span, M., et al., *A detailed analysis of K-ras point mutations in relation to tumor progression and survival in colorectal cancer patients.* Int J Cancer, 1996. **69**(3): p. 241-5.
150. Pajkos, G., et al., *The prognostic value of the presence of mutations at the codons 12, 13, 61 of K-ras oncogene in colorectal cancer.* Anticancer Res, 2000. **20**(3A): p. 1695-701.
151. Ahnen, D.J., et al., *Ki-ras mutation and p53 overexpression predict the clinical behavior of colorectal cancer: a Southwest Oncology Group study.* Cancer Res, 1998. **58**(6): p. 1149-58.
152. Bouzourene, H., et al., *p53 and Ki-ras as prognostic factors for Dukes' stage B colorectal cancer.* Eur J Cancer, 2000. **36**(8): p. 1008-15.
153. Pricolo, V.E., et al., *Prognostic value of TP53 and K-ras-2 mutational analysis in stage III carcinoma of the colon.* Am J Surg, 1996. **171**(1): p. 41-6.
154. Laurent-Puig, P., et al., *Survival and acquired genetic alterations in colorectal cancer.* Gastroenterology, 1992. **102**(4 Pt 1): p. 1136-41.
155. Bell, S.M., et al., *Prognostic value of p53 overexpression and c-Ki-ras gene mutations in colorectal cancer.* Gastroenterology, 1993. **104**(1): p. 57-64.
156. Morrin, M., et al., *Mutations of Ki-ras and p53 genes in colorectal cancer and their prognostic significance.* Gut, 1994. **35**(11): p. 1627-31.
157. Markowitz, S., et al., *Mutant K-ras oncogenes in colon cancers Do not predict Patient's chemotherapy response or survival.* Clin Cancer Res, 1995. **1**(4): p. 441-5.
158. Andersen, S.N., et al., *K-ras mutations and prognosis in large-bowel carcinomas.* Scand J Gastroenterol, 1997. **32**(1): p. 62-9.
159. Andreyev, H.J., et al., *Kirsten ras mutations in patients with colorectal cancer: the 'RASCAL II' study.* Br J Cancer, 2001. **85**(5): p. 692-6.
160. Lievre, A., et al., *KRAS mutation status is predictive of response to cetuximab therapy in colorectal cancer.* Cancer Res, 2006. **66**(8): p. 3992-5.
161. Di Fiore, F., et al., *Clinical relevance of KRAS mutation detection in metastatic colorectal cancer treated by Cetuximab plus chemotherapy.* Br J Cancer, 2007. **96**(8): p. 1166-9.

162. Kressner, U., et al., *Immunohistological p53 staining is of limited value in the staging and prognostic prediction of colorectal cancer.* Anticancer Res, 1996. **16**(2): p. 951-7.
163. Grewal, H., et al., *p53 nuclear overexpression may not be an independent prognostic marker in early colorectal cancer.* Dis Colon Rectum, 1995. **38**(11): p. 1176-81.
164. Tanaka, M., et al., *Prognostic factors of colorectal cancer: K-ras mutation, overexpression of the p53 protein, and cell proliferative activity.* J Surg Oncol, 1994. **57**(1): p. 57-64.
165. Nathanson, S.D., et al., *Relationship among p53, stage, and prognosis of large bowel cancer.* Dis Colon Rectum, 1994. **37**(6): p. 527-34.
166. Bazan, V., et al., *DNA ploidy and S-phase fraction, but not p53 or NM23-H1 expression, predict outcome in colorectal cancer patients. Result of a 5-year prospective study.* J Cancer Res Clin Oncol, 2002. **128**(12): p. 650-8.
167. Remvikos, Y., et al., *Increased p53 protein content of colorectal tumours correlates with poor survival.* Br J Cancer, 1992. **66**(4): p. 758-64.
168. Diez, M., et al., *Prediction of recurrence in B-C stages of colorectal cancer by p53 nuclear overexpression in comparison with standard pathological features.* Eur J Surg Oncol, 1995. **21**(6): p. 635-9.
169. Lazaris, A.C., et al., *Prognostic significance of p53 and c-erbB-2 immunohistochemical evaluation in colorectal adenocarcinoma.* Histol Histopathol, 1995. **10**(3): p. 661-8.
170. Houbiers, J.G., et al., *Antibodies against p53 are associated with poor prognosis of colorectal cancer.* Br J Cancer, 1995. **72**(3): p. 637-41.
171. Zeng, Z.S., et al., *p53 nuclear overexpression: an independent predictor of survival in lymph node--positive colorectal cancer patients.* J Clin Oncol, 1994. **12**(10): p. 2043-50.
172. Yamaguchi, A., et al., *Expression of p53 protein in colorectal cancer and its relationship to short-term prognosis.* Cancer, 1992. **70**(12): p. 2778-84.
173. Buglioni, S., et al., *p53 nuclear accumulation and multiploidy are adverse prognostic factors in surgically resected stage II colorectal cancers independent of fluorouracil-based adjuvant therapy.* Am J Clin Pathol, 2001. **116**(3): p. 360-8.
174. Schwandner, O., et al., *p53 and Bcl-2 as significant predictors of recurrence and survival in rectal cancer.* Eur J Cancer, 2000. **36**(3): p. 348-56.
175. Adrover, E., et al., *Expression of high p53 levels in colorectal cancer: a favourable prognostic factor.* Br J Cancer, 1999. **81**(1): p. 122-6.
176. Haseba, M., et al., *Detection of p53 gene mutations by nonisotopic RNase cleavage assay as a predictor of poor prognosis in colorectal cancers.* Dig Dis Sci, 2003. **48**(10): p. 1984-9.
177. Elsaleh, H., et al., *p53 gene mutation, microsatellite instability and adjuvant chemotherapy: impact on survival of 388 patients with Dukes' C colon carcinoma.* Oncology, 2000. **58**(1): p. 52-9.
178. Soong, R., et al., *Prognostic significance of TP53 gene mutation in 995 cases of colorectal carcinoma. Influence of tumour site, stage, adjuvant chemotherapy and type of mutation.* Eur J Cancer, 2000. **36**(16): p. 2053-60.
179. Rebischung, C., et al., *Prognostic value of P53 mutations in rectal carcinoma.* Int J Cancer, 2002. **100**(2): p. 131-5.
180. Samowitz, W.S., et al., *Prognostic significance of p53 mutations in colon cancer at the population level.* Int J Cancer, 2002. **99**(4): p. 597-602.
181. Kahlenberg, M.S., et al., *p53 tumor suppressor gene mutations predict decreased survival of patients with sporadic colorectal carcinoma.* Cancer, 2000. **88**(8): p. 1814-9.
182. Russo, A., et al., *p53 mutations in L3-loop zinc-binding domain, DNA-ploidy, and S phase fraction are independent prognostic indicators in colorectal cancer: a prospective study with a five-year follow-up.* Cancer Epidemiol Biomarkers Prev, 2002. **11**(11): p. 1322-31.
183. van de Vijver, M.J., et al., *A gene-expression signature as a predictor of survival in breast cancer.* N Engl J Med, 2002. **347**(25): p. 1999-2009.
184. Rosenwald, A., et al., *The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma.* N Engl J Med, 2002. **346**(25): p. 1937-47.
185. van't Veer, L.J., S. Paik, and D.F. Hayes, *Gene expression profiling of breast cancer: a new tumor marker.* J Clin Oncol, 2005. **23**(8): p. 1631-5.
186. Jones, C., et al., *Molecular cytogenetic identification of subgroups of grade III invasive ductal breast carcinomas with different clinical outcomes.* Clin Cancer Res, 2004. **10**(18 Pt 1): p. 5988-97.
187. Moertel, C.G., et al., *Fluorouracil plus levamisole as effective adjuvant therapy after resection of stage III colon carcinoma: a final report.* Ann Intern Med, 1995. **122**(5): p. 321-6.

188. Mayer, R.J., *Two steps forward in the treatment of colorectal cancer.* N Engl J Med, 2004. **350**(23): p. 2406-8.
189. Andre, T., et al., *Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer.* N Engl J Med, 2004. **350**(23): p. 2343-51.
190. Lasko, D., W. Cavenee, and M. Nordenskjold, *Loss of constitutional heterozygosity in human cancer.* Annu Rev Genet, 1991. **25**: p. 281-314.
191. Gryfe, R., et al., *Molecular biology of colorectal cancer.* Curr Probl Cancer, 1997. **21**(5): p. 233-300.
192. Minamoto, T., M. Mai, and Z. Ronai, *K-ras mutation: early detection in molecular diagnosis and risk assessment of colorectal, pancreas, and lung cancers--a review.* Cancer Detect Prev, 2000. **24**(1): p. 1-12.
193. Mitchell, R.J., et al., *Mismatch repair genes hMLH1 and hMSH2 and colorectal cancer: a HuGE review.* Am J Epidemiol, 2002. **156**(10): p. 885-902.
194. Gruis, N.A., et al., *PCR-based microsatellite polymorphisms in the detection of loss of heterozygosity in fresh and archival tumour tissue.* Br J Cancer, 1993. **68**(2): p. 308-13.
195. Van Prooijen-Knegt, A.C., et al., *In situ hybridization of DNA sequences in human metaphase chromosomes visualized by an indirect fluorescent immunocytochemical procedure.* Exp Cell Res, 1982. **141**(2): p. 397-407.
196. Sokolova, I.A., et al., *The development of a multitarget, multicolor fluorescence in situ hybridization assay for the detection of urothelial carcinoma in urine.* J Mol Diagn, 2000. **2**(3): p. 116-23.
197. Bastian, B.C., et al., *Molecular cytogenetic analysis of Spitz nevi shows clear differences to melanoma.* J Invest Dermatol, 1999. **113**(6): p. 1065-9.
198. Kallioniemi, A., et al., *Detection of retinoblastoma gene copy number in metaphase chromosomes and interphase nuclei by fluorescence in situ hybridization.* Cytogenet Cell Genet, 1992. **60**(3-4): p. 190-3.
199. Lichter, P., et al., *Comparative genomic hybridization: uses and limitations.* Semin Hematol, 2000. **37**(4): p. 348-57.
200. Knuutila, S., K. Autio, and Y. Aalto, *Online access to CGH data of DNA sequence copy number changes.* Am J Pathol, 2000. **157**(2): p. 689.
201. Forozan, F., et al., *Genome screening by comparative genomic hybridization.* Trends Genet, 1997. **13**(10): p. 405-9.
202. Wells, D. and B. Levy, *Cytogenetics in reproductive medicine: the contribution of comparative genomic hybridization (CGH).* Bioessays, 2003. **25**(3): p. 289-300.
203. Yu, L.C., et al., *Objective aneuploidy detection for fetal and neonatal screening using comparative genomic hybridization (CGH).* Cytometry, 1997. **28**(3): p. 191-7.
204. Lapierre, J.M., et al., *Analysis of uncultured amniocytes by comparative genomic hybridization: a prospective prenatal study.* Prenat Diagn, 2000. **20**(2): p. 123-31.
205. Forster-Gibson, C.J., et al., *Cryptic duplication of 21q in an individual with a clinical diagnosis of Down syndrome.* Clin Genet, 2001. **59**(6): p. 438-43.
206. Trask, B.J., *Human cytogenetics: 46 chromosomes, 46 years and counting.* Nat Rev Genet, 2002. **3**(10): p. 769-78.
207. Weiss, M.M., et al., *Comparative genomic hybridisation as a supportive tool in diagnostic pathology.* J Clin Pathol, 2003. **56**(7): p. 522-7.
208. Weiss, M.M., et al., *Genomic alterations in primary gastric adenocarcinomas correlate with clinicopathological characteristics and survival.* Cell Oncol, 2004. **26**(5-6): p. 307-17.
209. Kallioniemi, A., et al., *Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors.* Science, 1992. **258**(5083): p. 818-21.
210. Kallioniemi, O.P., et al., *Optimizing comparative genomic hybridization for analysis of DNA sequence copy number changes in solid tumors.* Genes Chromosomes Cancer, 1994. **10**(4): p. 231-43.
211. Solinas-Toldo, S., et al., *Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances.* Genes Chromosomes Cancer, 1997. **20**(4): p. 399-407.
212. Fiegler, H., et al., *DNA microarrays for comparative genomic hybridization based on DOP-PCR amplification of BAC and PAC clones.* Genes Chromosomes Cancer, 2003. **36**(4): p. 361-74.
213. Pinkel, D., et al., *High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays.* Nat Genet, 1998. **20**(2): p. 207-11.

214. Veltman, J.A., et al., *High-throughput analysis of subtelomeric chromosome rearrangements by use of array-based comparative genomic hybridization.* Am J Hum Genet, 2002. **70**(5): p. 1269-76.
215. Southern, E., K. Mir, and M. Shchepinov, *Molecular interactions on microarrays.* Nat Genet, 1999. **21**(1 Suppl): p. 5-9.
216. Cheung, V.G., et al., *Making and reading microarrays.* Nat Genet, 1999. **21**(1 Suppl): p. 15-9.
217. Buckley, P.G., et al., *A full-coverage, high-resolution human chromosome 22 genomic microarray for clinical and research applications.* Hum Mol Genet, 2002. **11**(25): p. 3221-9.
218. Yu, W., et al., *Development of a comparative genomic hybridization microarray and demonstration of its utility with 25 well-characterized 1p36 deletions.* Hum Mol Genet, 2003. **12**(17): p. 2145-52.
219. Prescott, K., et al., *A novel 5q11.2 deletion detected by microarray comparative genomic hybridisation in a child referred as a case of suspected 22q11 deletion syndrome.* Hum Genet, 2005. **116**(1-2): p. 83-90.
220. Ishkanian, A.S., et al., *A tiling resolution DNA microarray with complete coverage of the human genome.* Nat Genet, 2004. **36**(3): p. 299-303.
221. Hyman, E., et al., *Impact of DNA amplification on gene expression patterns in breast cancer.* Cancer Res, 2002. **62**(21): p. 6240-5.
222. Pollack, J.R., et al., *Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors.* Proc Natl Acad Sci U S A, 2002. **99**(20): p. 12963-8.
223. Ishizuka, T., et al., *Gene amplification profiling of esophageal squamous cell carcinomas by DNA array CGH.* Biochem Biophys Res Commun, 2002. **296**(1): p. 152-5.
224. van Dekken, H., et al., *Evaluation of genetic patterns in different tumor areas of intermediate-grade prostatic adenocarcinomas by high-resolution genomic array analysis.* Genes Chromosomes Cancer, 2004. **39**(3): p. 249-56.
225. Wilhelm, M., et al., *Array-based comparative genomic hybridization for the differential diagnosis of renal cell cancer.* Cancer Res, 2002. **62**(4): p. 957-60.
226. Weiss, M.M., et al., *Genomic profiling of gastric cancer predicts lymph node status and survival.* Oncogene, 2003. **22**(12): p. 1872-9.
227. Lindblad-Toh, K., et al., *Loss-of-heterozygosity analysis of small-cell lung carcinomas using single-nucleotide polymorphism arrays.* Nat Biotechnol, 2000. **18**(9): p. 1001-5.
228. Schubert, E.L., et al., *Single nucleotide polymorphism array analysis of flow-sorted epithelial cells from frozen versus fixed tissues for whole genome analysis of allelic loss in breast cancer.* Am J Pathol, 2002. **160**(1): p. 73-9.
229. Dumur, C.I., et al., *Genome-wide detection of LOH in prostate cancer using human SNP microarray technology.* Genomics, 2003. **81**(3): p. 260-9.
230. Hoque, M.O., et al., *Genome-wide genetic characterization of bladder cancer: a comparison of high-density single-nucleotide polymorphism arrays and PCR-based microsatellite analysis.* Cancer Res, 2003. **63**(9): p. 2216-22.
231. Wong, K.K., et al., *Allelic imbalance analysis by high-density single-nucleotide polymorphic allele (SNP) array with whole genome amplified DNA.* Nucleic Acids Res, 2004. **32**(9): p. e69.
232. Notterman, D.A., et al., *Transcriptional gene expression profiles of colorectal adenoma, adenocarcinoma, and normal tissue examined by oligonucleotide arrays.* Cancer Res, 2001. **61**(7): p. 3124-30.
233. Kitahara, O., et al., *Alterations of gene expression during colorectal carcinogenesis revealed by cDNA microarrays after laser-capture microdissection of tumor tissues and normal epithelia.* Cancer Res, 2001. **61**(9): p. 3544-9.
234. Alizadeh, A.A., et al., *Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling.* Nature, 2000. **403**(6769): p. 503-11.
235. Clark, E.A., et al., *Genomic analysis of metastasis reveals an essential role for RhoC.* Nature, 2000. **406**(6795): p. 532-5.
236. Alon, U., et al., *Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays.* Proc Natl Acad Sci U S A, 1999. **96**(12): p. 6745-50.
237. Zhang, L., et al., *Gene expression profiles in normal and cancer cells.* Science, 1997. **276**(5316): p. 1268-72.
238. Martin, K.J., et al., *Linking gene expression patterns to therapeutic groups in breast cancer.* Cancer Res, 2000. **60**(8): p. 2232-8.

239. Lu, J., et al., *MicroRNA expression profiles classify human cancers.* Nature, 2005. **435**(7043): p. 834-8.
240. Aust, D.E., et al., *Chromosomal alterations in ulcerative colitis-related and sporadic colorectal cancers by comparative genomic hybridization.* Hum Pathol, 2000. **31**(1): p. 109-14.
241. Thiagalingam, S., et al., *Evaluation of candidate tumour suppressor genes on chromosome 18 in colorectal cancers.* Nat Genet, 1996. **13**(3): p. 343-6.
242. Ried, T., et al., *Comparative genomic hybridization reveals a specific pattern of chromosomal gains and losses during the genesis of colorectal tumors.* Genes Chromosomes Cancer, 1996. **15**(4): p. 234-45.
243. Korn, W.M., et al., *Chromosome arm 20q gains and other genomic alterations in colorectal cancer metastatic to liver, as analyzed by comparative genomic hybridization and fluorescence in situ hybridization.* Genes Chromosomes Cancer, 1999. **25**(2): p. 82-90.
244. Al-Mulla, F., et al., *Comparative genomic hybridization analysis of primary colorectal carcinomas and their synchronous metastases.* Genes Chromosomes Cancer, 1999. **24**(4): p. 306-14.
245. Aragane, H., et al., *Chromosomal aberrations in colorectal cancers and liver metastases analyzed by comparative genomic hybridization.* Int J Cancer, 2001. **94**(5): p. 623-9.
246. De Angelis, P.M., et al., *Chromosomal gains and losses in primary colorectal carcinomas detected by CGH and their associations with tumour DNA ploidy, genotypes and phenotypes.* Br J Cancer, 1999. **80**(3-4): p. 526-35.
247. Nakao, K., et al., *High-resolution analysis of DNA copy number alterations in colorectal cancer by array-based comparative genomic hybridization.* Carcinogenesis, 2004. **25**(8): p. 1345-57.
248. Terdiman, J.P., et al., *High resolution analysis of chromosome 18 alterations in ulcerative colitis-related colorectal cancer.* Cancer Genet Cytogenet, 2002. **136**(2): p. 129-37.
249. Zetter, B.R., *Adhesion molecules in tumor metastasis.* Semin Cancer Biol, 1993. **4**(4): p. 219-29.
250. Gotley, D.C., et al., *The deleted in colon cancer (DCC) gene is consistently expressed in colorectal cancers and metastases.* Oncogene, 1996. **13**(4): p. 787-95.
251. Takagi, Y., et al., *Somatic alterations of the SMAD-2 gene in human colorectal cancers.* Br J Cancer, 1998. **78**(9): p. 1152-5.
252. Hahn, S.A., et al., *DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1.* Science, 1996. **271**(5247): p. 350-3.
253. Howe, J.R., et al., *Mutations in the SMAD4/DPC4 gene in juvenile polyposis.* Science, 1998. **280**(5366): p. 1086-8.
254. Salovaara, R., et al., *Frequent loss of SMAD4/DPC4 protein in colorectal cancers.* Gut, 2002. **51**(1): p. 56-9.
255. Takagi, Y., et al., *Somatic alterations of the DPC4 gene in human colorectal cancers in vivo.* Gastroenterology, 1996. **111**(5): p. 1369-72.
256. Jones, P.A. and P.W. Laird, *Cancer epigenetics comes of age.* Nat Genet, 1999. **21**(2): p. 163-7.
257. Hagiwara, T., et al., *Genomic organization, promoter analysis, and chromosomal localization of the gene for the mouse glial high-affinity glutamate transporter Slc1a3.* Genomics, 1996. **33**(3): p. 508-15.
258. Roth, S., et al., *No SMAD4 hypermethylation in colorectal cancer.* Br J Cancer, 2000. **83**(8): p. 1015-9.
259. Tuzun, E., et al., *Fine-scale structural variation of the human genome.* Nat Genet, 2005. **37**(7): p. 727-32.
260. Sebat, J., et al., *Large-scale copy number polymorphism in the human genome.* Science, 2004. **305**(5683): p. 525-8.
261. Gonzalez, E., et al., *The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility.* Science, 2005. **307**(5714): p. 1434-40.
262. Wolf, M., et al., *High-resolution analysis of gene copy number alterations in human prostate cancer using CGH on cDNA microarrays: impact of copy number on gene expression.* Neoplasia, 2004. **6**(3): p. 240-7.
263. Jarbo, C., et al., *Detailed assessment of chromosome 22 aberrations in sporadic pheochromocytoma using array-CGH.* Int J Cancer, 2006. **118**(5): p. 1159-64.
264. Melendez, B., et al., *Identification of amplified and highly expressed genes in amplicons of the T-cell line huT78 detected by cDNA microarray CGH.* Mol Cancer, 2005. **4**(1): p. 5.

265. Schoumans, J., et al., *Characterisation of dic(9;20)(p11-13;q11) in childhood B-cell precursor acute lymphoblastic leukaemia by tiling resolution array-based comparative genomic hybridisation reveals clustered breakpoints at 9p13.2 and 20q11.2.* Br J Haematol, 2006. **135**(4): p. 492-9.
266. Strefford, J.C., et al., *Genome complexity in acute lymphoblastic leukemia is revealed by array-based comparative genomic hybridization.* Oncogene, 2007. **26**(29): p. 4306-18.
267. Hidalgo, A., et al., *Microarray comparative genomic hybridization detection of chromosomal imbalances in uterine cervix carcinoma.* BMC Cancer, 2005. **5**: p. 77.
268. Loukopoulos, P., et al., *Genome-wide array-based comparative genomic hybridization analysis of pancreatic adenocarcinoma: identification of genetic indicators that predict patient outcome.* Cancer Sci, 2007. **98**(3): p. 392-400.
269. Natrajan, R., et al., *Array CGH profiling of favourable histology Wilms tumours reveals novel gains and losses associated with relapse.* J Pathol, 2006. **210**(1): p. 49-58.
270. Moinzadeh, P., et al., *Chromosome alterations in human hepatocellular carcinomas correlate with aetiology and histological grade--results of an explorative CGH meta-analysis.* Br J Cancer, 2005. **92**(5): p. 935-41.
271. Mastracci, T.L., et al., *Genomic alterations in lobular neoplasia: a microarray comparative genomic hybridization signature for early neoplastic proliferationin the breast.* Genes Chromosomes Cancer, 2006. **45**(11): p. 1007-17.
272. Douglas, E.J., et al., *Array comparative genomic hybridization analysis of colorectal cancer cell lines and primary carcinomas.* Cancer Res, 2004. **64**(14): p. 4817-25.
273. Roberts, A.B. and L.M. Wakefield, *The two faces of transforming growth factor beta in carcinogenesis.* Proc Natl Acad Sci U S A, 2003. **100**(15): p. 8621-3.
274. Nakao, A., et al., *Identification of Smad7, a TGFbeta-inducible antagonist of TGF-beta signalling.* Nature, 1997. **389**(6651): p. 631-5.
275. Lallemand, F., et al., *Smad7 inhibits the survival nuclear factor kappaB and potentiates apoptosis in epithelial cells.* Oncogene, 2001. **20**(7): p. 879-84.
276. Boulay, J.L., et al., *SMAD7 is a prognostic marker in patients with colorectal cancer.* Int J Cancer, 2003. **104**(4): p. 446-9.
277. Broderick, P., et al., *A genome-wide association study shows that common alleles of SMAD7 influence colorectal cancer risk.* Nat Genet, 2007. **39**(11): p. 1315-7.
278. Cohn, D.H., et al., *Mental retardation and abnormal skeletal development (Dyggve-Melchior-Clausen dysplasia) due to mutations in a novel, evolutionarily conserved gene.* Am J Hum Genet, 2003. **72**(2): p. 419-28.
279. Mager, D.L. and J.D. Freeman, *A human gene related to the ribosomal protein L23 gene of Halobacterium marismortui.* Nucleic Acids Res, 1990. **18**(17): p. 5301.
280. Jaye, M., et al., *A novel endothelial-derived lipase that modulates HDL metabolism.* Nat Genet, 1999. **21**(4): p. 424-8.
281. Abe, H., et al., *Cloning and sequence analysis of a full length cDNA encoding human mitochondrial 3-oxoacyl-CoA thiolase.* Biochim Biophys Acta, 1993. **1216**(2): p. 304-6.
282. Lapierre, L.A., et al., *Myosin vb is associated with plasma membrane recycling systems.* Mol Biol Cell, 2001. **12**(6): p. 1843-57.
283. Zhao, L.P., et al., *Cloning and characterization of myr 6, an unconventional myosin of the dilute/myosin-V family.* Proc Natl Acad Sci U S A, 1996. **93**(20): p. 10826-31.
284. Tsukita, S., M. Furuse, and M. Itoh, *Multifunctional strands in tight junctions.* Nat Rev Mol Cell Biol, 2001. **2**(4): p. 285-93.
285. Wheelock, M.J. and K.R. Johnson, *Cadherin-mediated cellular signaling.* Curr Opin Cell Biol, 2003. **15**(5): p. 509-14.
286. Wheelock, M.J. and K.R. Johnson, *Cadherins as modulators of cellular phenotype.* Annu Rev Cell Dev Biol, 2003. **19**: p. 207-35.
287. Nagafuchi, A., *Molecular architecture of adherens junctions.* Curr Opin Cell Biol, 2001. **13**(5): p. 600-3.
288. Takeichi, M., *Cadherins: a molecular family important in selective cell-cell adhesion.* Annu Rev Biochem, 1990. **59**: p. 237-52.
289. Takeichi, M., et al., *Cadherin subclasses: differential expression and their roles in neural morphogenesis.* Cold Spring Harb Symp Quant Biol, 1990. **55**: p. 319-25.
290. Doggett, N.A., M.H. Breuning, and D.F. Callen, *Report of the fourth international workshop on human chromosome 16 mapping 1995.* Cytogenet Cell Genet, 1996. **72**(4): p. 271-93.
291. Latil, A., et al., *Loss of heterozygosity at chromosome 16q in prostate adenocarcinoma: identification of three independent regions.* Cancer Res, 1997. **57**(6): p. 1058-62.

292. Suzuki, H., et al., *Three distinct commonly deleted regions of chromosome arm 16q in human primary and metastatic prostate cancers.* Genes Chromosomes Cancer, 1996. **17**(4): p. 225-33.
293. Kools, P., G. Van Imschoot, and F. van Roy, *Characterization of three novel human cadherin genes (CDH7, CDH19, and CDH20) clustered on chromosome 18q22-q23 and with high homology to chicken cadherin-7.* Genomics, 2000. **68**(3): p. 283-95.
294. Kremmidiotis, G., et al., *Localization of human cadherin genes to chromosome regions exhibiting cancer-related loss of heterozygosity.* Genomics, 1998. **49**(3): p. 467-71.
295. Akhurst, R.J. and R. Derynck, *TGF-beta signaling in cancer--a double-edged sword.* Trends Cell Biol, 2001. **11**(11): p. S44-51.
296. Derynck, R., R.J. Akhurst, and A. Balmain, *TGF-beta signaling in tumor suppression and cancer progression.* Nat Genet, 2001. **29**(2): p. 117-29.
297. Wakefield, L.M. and A.B. Roberts, *TGF-beta signaling: positive and negative effects on tumorigenesis.* Curr Opin Genet Dev, 2002. **12**(1): p. 22-9.
298. Bhowmick, N.A., et al., *Transforming growth factor-beta1 mediates epithelial to mesenchymal transdifferentiation through a RhoA-dependent mechanism.* Mol Biol Cell, 2001. **12**(1): p. 27-36.
299. Piek, E., et al., *Expression of transforming-growth-factor (TGF)-beta receptors and Smad proteins in glioblastoma cell lines with distinct responses to TGF-beta1.* Int J Cancer, 1999. **80**(5): p. 756-63.

References not available on PUBMED

I Bodmer WF. The somatic evolution of cancer. The Harveian Oration of 1996. J R Coll Physicians Lond 1996; 31:82-89

Appendix 1 Cell line and primary cancer clinical and genetic data

Cell lines

CELL LINE	AGE	SEX	STAGE	SITE	KARYOTYPE	MSI	SMAD4	CGH 18 DEL
C1251					Hypotriploid (60)	-	None found	Y
C32	64	F	C	Colon	Hypertriploid (74)	-	None found	Y
C70	69	F	B	Sigmoid Colon	Hypotriploid (61)	-	None found	Y
C80	69	M	C	Rectum	Triploid (69)	-	?1051c	Y
C84	67	M	C	Caecum	Hyperdiploid(56)	-	None found	Y
C99	69	M	C	Rectum	Hyperdiploid (52)	-	None found	Y
C75	56	M	C	Sigmoid Colon	Hypertriploid (70)	-		N
CACO2	72	M		Colon	Hypertetraploid (96)	-	?1051c	Y
COLO201	70	M	D	Colon-Ascites	Near-triploid (68)	-	Del ex1-ex4	Y
COLO205	70	M	D	Colon- Ascites	Hypertriploid (78)	-	Del ex1-ex4	Y
COLO320DM	55	F	C	Sigmoid colon	Hyperdiploid (53)	-	None found	Y
COLO678	69	M	C	Met to lymph node	Hyperdiploid (?)	-	Del ex1 to ex11	Y
COLO741	69	F	D	Pelvic wall met	Not done	(+)	None found	Y
CX-1	44	F			Hypertriploid (72)	-	c931t, g1237t	Y
DLD1/HCT15			C	Colon	Near-diploid (46)	+	None found	N
HT716						-		Y
HCA461	53	F	C	Sigmoid colon	Not done	-	None found	Y
HCA7	58	F	B	Colon	Hypodiploid (43)	+	None found	Y
HCT116		M		Colon	Near-diploid (45)	+	None found	Y
HCT8				Colon	Near-diploid (46)	+	None found	N
HUTU80						+		Y
LIM1863	74	F	C	Ileocaecum	Hypertriploid (80)	-	None found	N
LS1034	54	M	C	Caecum	Hypertriploid (77)	-	None found	Y
LS123	65	F			Hyperdiploid (63)	-		Y
LS411	32	M	B	Caecum	Hypertriploid (75)	+	None found	Y
PC/JW	30	F		Colon	Not done	-	?	Y
RKO					Not done	+		Y
SW1222	44	M	C2		Not done	(+)	t1619g	Y
SW1417	53	F	C2	Colon	Near-triploid (70)	-	?	Y
SW403	51	F	C1	Colon	Near-triploid (68)	-	Del ex 10+11	N
SW48	82	F	C2	Colon	Near-diploid (47)	+	None found	Y
SW480	50	M	B	Colon	Hypotriploid (57)	-	5th base intron 7, g->c	Y
SW620	51	M	C	Colon l. node	Hyperdiploid (50)	-	5th base intron 7, g->c	Y
SW837	53	M	C2	Rectum	Hypodiploid (40)	(+)	None found	Y
SW948	81	F	C	Colon	Hypotriploid (67)	-	g1609t	Y
VACO10	72	F	D	Colon omentum	Pentaploid (115)	-	c709t	Y
VACO5	78	F	C2	Caecum	Diploid (47)	+	None found	Y
VACO4A	59	M	D	Rectum	Hypotriploid (60)	-	None found	Y
VACO45	59	M	D	Rectum	Hypotriploid (64)	-	None found	Y
HT55	54	F	C	Rectum	Hypertriploid	-	?	Y
C106	78	F	A	Lower Rectum	Hypertriploid (79)	-		Y

HT29	44	F	N/D	Colon	Hypertriploid (71)	-	c931t, g1237t	Y
C10	71	M	B	Left colon	Hyperdiploid (49)	-	Nt 898-1,g->t	N
LS174T	58	F	B	Colon	Near-diploid (45)	+	None found	N
VACO16								N
HT 716								Y
VACO 13								N

Primary colorectal cancers

PRIMARY COLORECTAL CANCER	SEX	AGE	OPERATION	STAGE	APC LOH	BAT26	CGH 18 DEL
9779/93C	F	69	r hemi	C			N
19395/93C	F	81	recto sig	C			N
11064/93C	M	75	r hemi	C			N
19616/93C	F	59	R hemi	B			Y
17089/90C	F	70	high ar	C	+		Y
13068/93C	F	55	Sig col	B			N
14609/90C	M	47	AR	B			N
8943/93C	M	38	Rhemi	C			N
9069C	M	63	AR	C			Y
2946/93C	M	69		C			Y
2428/94C	M	56	R hemi	B			N
18921/93C	M	70	high ar	C			N
4890/93C	M	71	low ar	C	+		Y
17690/93C	M	81	r hemi	C		+	N
546/94C	F	78		C			N
19546/90C	M	70		C		+	N
5964/93C	M	67	AR	B			Y
7131/92C	F	83	AR	B			N
19213/93C	F	70	APR	B	+		Y
4629/93C	F	75	subtotal	B			Y
442/94C	F	72	Sigmoid colectomy	B			Y
17309/93C	F	67	AR	B	+		N
598/91C	M	52	AR	B			N
19092/94C	F	75	R hemi	C			Y
4610/93C	F	55	AP	B			Y
16622/90C	F	51	r hemi	C			Y
5446/93C	M	57	AR	B	+		N
3822/93C	F	81	r hemi	C	+	+	N
1593493C	M	36	AR	C			Y
1057990C	F	81	AP	B	+		Y
17433/90C	F	54	R hemi	B	+		N
19002/90C	M	77	ar	C			Y
17565C	M	83	r hemi	B	+	+	Y
13733/93C	M	84	Sig col	B			N
19774/93C	M	67	R hemi	B			N
11351/93C	F	86	AR	B	+		Y
5172/93C	M	75	AR	B	+		N
6972/90C	F	71	AR	B			N
2973/92C	M	77	R hemi	B	+	+	N
1886/94C	M	44	r hemi	C			Y
17800/90C	M	67	APR	B			N
824493C	M	59	sig ca	B			Y

5444/93C	M	76	AR	B			N
468/93C	F	53		C			N
16597/90C	M	71	AR	B			Y
968/90C	M	85	R hemi	B			Y
16462/90C	F	66	AR	C			N
9803/90C	F	75	AR	B	+		Y
18102/93C	F	70	AR	B			Y
15011/90C	M	75	Sig col	B			Y
8931/93C	M	55	sig col	B			Y
4007/93C	M	63	AP	B			Y
12851/90C	F	56	sig colon	C			N
19248/93C	F	39	recto sig	C			Y
3424/93C	M	68	APR	B			Y
18918/93C	F	71	R hemi	C			Y
6737/93C	M	57	low ar	C			N
11974/93C	M	77	l hemi	B			N
11361/90C	M	67	Sig col	B	+		N
16607/94C	M	56	Rhemi	C	+		Y
8930/90C	M	80	R colon	B			N
8362/90C	M	79	AP	C			Y
20327/94C?P	F	78		C		+	Y
2977/92C	F	86		?C		+	Y
6757/93C	M	53		C			Y
10950	F	67		C			Y
15935	M	78		C			Y
10574	F	69		C			Y
11538	M	84		C			Y