Dukes Stage B Colorectal Cancers Analysed with Array Comparative Genomic Hybridisation

William Chambers

University College London

This thesis is submitted for the degree of MD (Res) University College London

The research was conducted between 2003-2005

I, William Chambers, confirm that the work presented in this thesis is my own. When information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Introduction: The main aim was to investigate Dukes B cancer prognosis using array comparative genomic hybridisation (aCGH). The cancers were typed for microsatellite instability (MSI), *APC* loss of heterozygosity (LOH) and ploidy. A secondary aim was the selection, on the basis of aCGH results, of potential colorectal cancer genes using In Situ Hybridisation (ISH). Sporadic adenomas were also investigated with aCGH to investigate the timing of chromosomal instability in tumourigenesis.

Methods: Dukes B cancers collected in Oxford, Leeds and Harrow were examined using aCGH based on a set of 3452 BAC clones at ~1Mb spacing. MSI was assessed using BAT26 and D5S346. D5S346 gave LOH at the *APC* gene locus. Ploidy was assessed using flow cytometry (FACS). From the aCGH data genes in regions of copy number gain were chosen for ISH analysis.

Results: 79 Dukes B cancers (43 good outcome, 36 bad outcome on the basis of 5 year survival) were investigated with aCGH. The most commonly gained chromosomes across all cancers were 13, 20 and 7, and the most commonly lost were 22, 18 and 14. Comparing survival groups; chromosome 6 was more often lost in cancers associated with good outcome, chromosome 16 was more often gained in microsatellite stable cancers associated with good outcome, and chromosome 22 more often gained in microsatellite stable cancers associated with poor outcome. Chromosome 13 showed greater than single copy number change significantly more often in bad outcome cancers. Several new areas of small genomic gain and loss were detected. Four candidate genes were identified (*CDX2*, *RHOA*, *FLT1* and *ARHGEF1*). None showed a relationship with outcome. Large scale chromosomal changes were found in 10 of14 sporadic adenomas.

Conclusions: Array CGH did identify some difference between good and bad outcome Dukes B cancers. However on the basis of this data, the technique could not be used as a useful clinical tool to predict prognosis.

Acknowledgements

I would like to thank the following people who have made this thesis possible:

The CRUK Molecular and Population Genetics Laboratory, Ian Tomlinson, Angela Jones, Andrew Rowan, Oliver Sieber and Patricia Gorman.

The Department of Colorectal Surgery, John Radcliffe Hospital, Oxford, Neil Mortensen.

The Oxford Medical Genetics Laboratory, Churchill Hospital, Oxford, Anneke Seller.

The Wellcome Laboratory, Roosevelt Way, Headington, Oxford, Regina Regan, Sam Knight.

The Department of General Surgery, Wexham Park Hospital, Slough, Jonathan Gilbert.

The Department of Pathology, John Radcliffe Hospital, Oxford, Mohammad Ilyas.

The Nuffield Department of Pathology, Oxford University, Steven Fox.

St Mark's Hospital, Harrow, London, John Northover.

The Department of Pathology Leeds General Infirmary, Phillip Quirke.

The Department of Pathology, University College London, Marco Novelli.

Particular thanks to Professor Ian Tomlinson for all his help, Mr Jonathan Gilbert for his role in the development of the project, Professor Neil Mortensen for his supervision, Professor Marco Novelli for his help with submission and to the Lady Sobell Trust, Wexham Park and the Oxford Genetics Knowledge Park for their financial support.

Accreditation of work

The list below details the work performed by me and by other researchers. The array CGH work was by far the most time consuming element of the research. In terms of work not carried out by me this can be seen to be the previous collection of the Oxford tissue samples, the histological analysis of these samples, the collection of the Leeds samples and the extraction of DNA from these,and the collection of follow up data related to these and finally the collection of St Marks samples and the experimentation on these.

(William Chambers unless specified)

1. Ethics approval	William Chambers
2. **Oxford samples**	
Collection-	Past Oxford researchers
Archiving	William Chambers
DNA extraction	William Chambers
MSI and LOH	William Chambers
FACS	William Chambers
aCGH	William Chambers
Clinicopathological data	William Chambers
and follow up data collection	
Histological analysis-	Professor M. Ilyas
3. **Leeds samples**	
Collection-	Professor P. Quirke's team
DNA extraction-	Professor P. Quirke's team
MSI and LOH	William Chambers
aCGH	William Chambers

Clinicopathological data	
and follow up data collection-	Professor P. Quirke's team
4. **St Marks samples**	
Collection- researchers	Past St Marks Hospital
DNA extraction-	Cancer Research UK London
MSI and LOH-	Cancer Research UK London
FACS-	Cancer Research UK London
aCGH-	Cancer Research UK London
Clinicopathological data	
and follow up data collection	William Chambers
Data processing	William Chambers
5. In situ hybridisation	
Production of probes	William Chambers
In situ hybridisation-	Professor Richard Poulsom, Cancer Research UK
Data processing	William Chambers

Abbreviations

aCGH	Array Comparative Genomic Hybridisation
APC	Adenomatous polyposis coli
ATP	Adenosine triphosphate
BFB	Break fusion break cycles
CEA	Carcinoembryonic antigen
CGH	Comparative Genomic Hybridisation
CIN	Chromosomal instability
CRA	Colorectal adenoma
CRC	Colorectal cancer
CRUK	Cancer Research United Kingdom
CTP	Cytosine triphosphate
DEPC	Diethylpyrocarbonate
DI	DNA index
DOP-PCR reaction	Degenerate oligonucleotide primed polymerase chain
DNA	Deoxyribonucleic acid
DSB	Double strand DNA breaks
EDTA	Ethylenediaminetetraacetic acid
FACS	Flow
FAP	Familial adenomatous polyposis
FISH	Flourescent in situ hybridisation
GTP	Guanine triphosphate
HNPCC	Hereditary non polyposis colorectal cancer

ISH	In situ hybridisation
LOH	Loss of heterozygosity
LVI	Lymphovascular invasion
MSI	Microsatellite instability
OD	Optical density
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PFA	Paraformaldehyde
PMT	Photon multiplier tube
RNA	Ribonucleic acid
SNP	Single nucleotide polymorphism
SSC	Saline sodium citrate
SSCP	Single strand conformation polymorphism
TNE	Tris NP40 EDTA

Publications including work performed by William Chambers in the course of this research:

CDC4 mutations occur in a subset of colorectal cancers but are not predicted to cause loss of function and are not associated with chromosomal instability.
Kemp Z, Rowan A, Chambers W, Wortham N, Halford S, Sieber O, Mortensen N, von Herbay A, Gunther T, Ilyas M, Tomlinson I. Cancer Res. 2005; 65: 11361-11366.

Analysis of copy number changes suggests chromosomal instability in a minority of large colorectal adenomas.
Jones A, Thirlwell C, Howarth K, Graham T, Chambers WM, Segditsas S, Page K, Philips R, Thomas H, Sieber O, Sawyer E, Tomlinson I. J. Pathol 2007; 213: 249-256.

APC and the three hit hypothesis.
Segditsas S, Rowan AJ, Howarth K, Jones A, Leedham S, Wright NA, Gorman P, Chambers WM, Domingo E, Roylance RR, Sawyer EJ, Sieber OM, Tomlinson IP. Oncogene 2009; 28: 146-155.

Table of Contents.

12. Appendices

Chapter 1.

Introduction

1.1 Colorectal Cancer

Colorectal cancer (CRC) accounted for 16,000 deaths in the UK in 2003 (Steele R, 2005) and constitutes a major health problem in the Western world. The 5-year survival rate is approximately 45%. The disease has a lifetime risk in the general population of 5%. The incidence of the disease is equal in men and women. Rectal cancers account for about 30% of CRC. Within the colon 25% of cancers arise in the right side of the colon and 45% arise in the left. It is generally accepted that most colonic cancers arise from pre-existing adenomatous polyps (Morson et al 1966, Bussey et al 1975, Muto et al 1975, Granquist et al 1981, Chu et al 1986, Eide et al 1986, Winawer et al 1987, Vogelstein et al 1988), the evidence for this being that adenomatous tissue often accompanies cancer, sporadic adenomas are identical histologically to the adenomas of familial adenomatous polyposis, large adenomas are more likely to display cellular atypia than small lesions, the distribution of adenomas in the colon is similar to that for cancers, adenomas are found in up to a third of specimens resected for cancer, the incidence of CRC falls with screening programmes involving colonoscopy and polypectomy and finally that patients present with adenomas on average 5 years before those patients presenting with cancers.

Ninety-five percent of CRC arise sporadically with approximately 5% related to the inherited syndromes FAP (Familial Adenomatous Polyposis) and HNPCC (Hereditary Non Polyposis Colorectal

Cancer). Other inherited conditions such as Peutz-Jeghers, Juvenile Polyposis, Cowden disease and mixed polyposis syndromes are also associated with a high risk of CRC.

Colorectal cancers mainly exhibit one of two types of genomic instability, chromosomal instability (CIN), characterised by an unstable chromosome number and an aneuploid/polyploid karyotype (Lengauer et al 1998, Fodde et al 2001) or microsatellite instability (MSI) due to a defective DNA mismatch repair pathway, resulting in sequence changes but with maintenance of a diploid karyotype (Modrich et al 1991, Ilyas et al 1999). A small number of cancers show either both or neither type of instability (Jones et al 2005). Most sporadic cancers and those associated with FAP are CIN+ whereas tumours associated with HNPCC and ~10% of sporadic CRC are MSI+.

Broadly speaking colorectal cancers may be divided into three groups in terms of genetic factors. First, a chromosomally unstable microsatellite stable group showing FAP related gene mutation. Secondly, a microsatellite unstable chromosomally stable group of inherited cancers. Thirdly, a group of sporadic, microsatellite unstable, chromosomally stable cancers showing CPG island methylation.

1.2 **Genetics of colorectal cancer**

1.2.1 **Chromosomal Instability (CIN)**

Since this thesis is primarily involved in investigating genomic stability I will begin the detailed discussion of the genetics of colorectal cancer by considering this aspect.

Colorectal cancers may be divided into those showing chromosomal instability (CIN) and those showing microsatellite instability with much smaller numbers showing either both types of instability or neither

type (Ilyas et al 1999). CIN describes cancers that show aneuploidy or polyploidy with large scale gains and losses of genetic material at the chromosomal level.

Debate exists as to the nature of CIN both in terms of its cause and its importance in tumourigenesis. It is argued by some that CIN is initiated early in tumourigenesis and is essential in generating a mutation rate necessary for tumour development (Loeb et al 1991). Others argue that it is a by product of the processes essential for tumourigenesis and not essential per se (Tomlinson et al1999).

Therefore a number of questions exist in relation to CIN. At what stage in tumorigenesis does it become evident? What causes it to develop? What is its importance in relation to tumorigenesis and once initiated how does it take place? How can it be measured?

1.2.2 **Timing of CIN**

By experimenting on adenomas one can attempt to discover how early CIN is evident.

Sieber et al (2002) studied evidence of genomic instability in 55 small adenomas from 18 FAP patients with 2 *APC* hits. They found little evidence of instability using a variety of techniques including LOH analysis, flow cytometry (FACS) and comparative genomic hybridisation (CGH). CGH analysis of 3 polyps which had shown LOH at *APC* showed normal profiles indicating that LOH at *APC* was not due to physical loss of material but more likely due to somatic recombination.

Shih et al (2001) assessed allelic imbalance using digital PCR in 32 sporadic adenomas with an average size of 2mm at 1p, 5q, 8p, 15 q and 18q. They found allelic imbalance in respectively, 10%, 55%, 19%, 28%, and 28% of cases. Over 90% of tumours exhibited allelic

imbalance at at least one site. These findings showed that allelic imbalance was a common event in very small adenomas and that these events occurred at an early stage of colorectal neoplasia.

While these results may seem to some degree contradictory they both show evidence to a greater or lesser extent of a degree of early genomic instability without evidence necessarily of large scale chromosomal instability. Shih et al did not test their samples using FACS or CGH and therefore we cannot know this for certain, but allelic imbalance itself does not imply large scale chromosomal change as shown by the CGH experiments carried out on Sieber et al's samples that had shown LOH.

Previous experiments have looked at the timing of chromosomal instability as evidenced by aneuploidy using FACS (Bauer et al 1987, Giaretti et al 1988, Quirke et al 1986). These show that about 10% of low grade polyps smaller than 1cm show aneuploidy and 12-37% of high grade polyps greater than 1cm. Levels for frank adenocarcinomas are between 50% and 85%. Thus it may be seen that a few early adenomas demonstrate aneuploidy and that the percentage of lesions demonstrating aneuploidy increases with advancing stages of tumorigenesis.

1.2.3 **LOH in relation to CIN**

It may be seen that while aneuploidy is not widespread early in tumorigenesis LOH is more evident. Such LOH may be achieved in a number of ways: mitotic non-disjunction, loss of a segment of chromosome resulting from a deletion event, mitotic recombination between two homologous chromosomes, break-induced replication, recombination between two non-homologous chromosomes (translocation) and gene conversion. Thiagalingam et al (2001) investigated likely mechanisms of LOH. The team examined the five chromosomes most often lost in CRC using 88 markers and FISH in 62 CRC cell lines which were microsatellite stable. Mechanisms of

loss were chromosome specific. Some chromosomes displayed complete loss as might be predicted by mitotic non-disjunction. However, more than half of the losses were associated with losses of part of the chromosome rather than the whole chromosome. These losses were largely due to structural alterations rather than mitotic recombination, break-induced replication or gene conversion.

Haigis et al (2002, 2004) described several different possible mechanisms mediating functional loss of APC/Apc: mutation, non-disjunction, homologous somatic recombination and epigenetic silencing. Using a mouse model they demonstrated that loss of Apc function could occur by LOH through somatic recombination between homologs. Robertsonian translocation RB(7.18)9Lub(Rb9) suppressed the multiplicity of adenomas in a mouse model. Their evidence showed that homologous recombination was the main pathway for LOH in adenomas in B6 min mice.

Tischfield and Shao (2003) in a commentary of the above paper wrote that homologous recombination was a key process in the initiation of cancer. It is a general mechanism that provides genomic sequence integrity in the repair of double strand breaks and rescues stalled DNA-replication forks. But when mitotic recombination occurs between homologs one of which has a mutation in a tumour suppressor allele the normal allelic sequence may be lost resulting in homozygosity for the mutant allele. Mitotic recombination seems to be modulated by genetic background and the degree of homology between homologous chromosome pairs. Regional proximity of homologous chromosome regions is required for the process. One might expect that any feature of somatic nuclear architecture that increases or decreases the interphase proximity of homologs is likely to affect the frequency of LOH due to mitotic recombination.

Gaasenbeek et al (2006) investigated colorectal cancer cell lines using array-comparative genomic hybridization (CGH) for copy

number changes and single-copy number polymorphism (SNP) microarrays for allelic loss (LOH). Many array based CGH changes were not found by LOH because they did not cause true reduction-to-homozygosity. Conversely, many regions of SNP-LOH occurred in the absence of copy number change, comprising an average per cell line of 2 chromosomes with complete LOH; 1-2 terminal regions of LOH (mitotic recombination); and 1 interstitial region of LOH. Microsatellite unstable (MSI+) lines infrequently showed gains/deletions or whole-chromosome LOH, but their near-diploid karyotypes concealed mitotic recombination frequencies similar to those of MSI- lines. These data suggest that CIN is not synonymous with copy number change and some cancers have a specific tendency to whole-chromosome deletion and regain or to mitotic recombination.

An enhanced rate of mitotic recombination is seen in hereditary syndromes like Bloom's, Fanconi anaemia and Werner's syndrome. The pattern of chromosome instability in Bloom's syndrome is characterized by sister-chromatid exchange and homologous chromatid interchanges reflected in a gain of homozygosity for polymorphic loci. The genes responsible for these syndromes have in part been cloned and protein products of both *BLM* and *WRN* genes are DNA helicases. However these patients do not seem to develop bowel cancer.

In summary gross aneuploidy and/or polyploidy does not seem to be a common early event in tumourigenesis. It has been proposed that a two step model for aneuploidy might exist in which early instability does not result in aneuploidy directly but results in mutations in genes that then result in aneuploidy (Jones et al 2004). LOH is detectable in early lesions and a process resulting in supranormal levels of LOH at an early stage might be seen as a potential initiating

factor. Homologous recombination seems the most attractive explanation for this. However the mechanisms for early increases in homologous recombination remain obscure.

1.2.4 **APC and CIN**

APC mutation is widely considered the initial event in a large proportion of colorectal cancers and there is evidence relating this to CIN.

In relation to *APC* as a cause of CIN Fodde et al (2001) examined mouse embryonic cells homozygous for Min or Apc1638T alleles and showed that mutant ES cells displayed extensive chromosome and spindle aberrations. APC accumulated at kinetochores during mitosis. *Apc* mutant cells formed mitotic spindles with microtubules that inefficiently connected to kinetochores. They concluded that loss of *Apc* sequences that lay C-terminal to the β-catenin regulatory domain contributed to CIN in CRC.

Kaplan et al (2001) showed that APC bound to and stabilized microtubules in vivo and in vitro, localized to clusters at the ends of microtubules near the plasma membrane of interphase cells and was an important regulator of cytoskeleton function. They showed that cells carrying a truncated *APC* gene were defective for chromosomal segregation in vitro. During mitosis APC localized to the end of microtubules imbedded in kinetochores and formed a complex with the checkpoint proteins BuB1 and BuB3. In vitro, APC was a high affinity substrate for BuB kinases. They suggested that truncations in the APC that eliminate microtubule binding may contribute to chromosomal instability.

However this in vitro evidence is at odds with what is seen in vivo, with both Sieber (2002) and Giaretti (2004) failing to find evidence for aneuploidy in all cancers with truncating *APC* mutations. Those that did demonstrate aneuploidy often showed a near diploid DNA index

whereas in the in vitro experiments detailed above the CIN seen was often of a polypoid kind usually, tetraploid. Further Giaretti found that the type of *APC* mutation most likely to be associated with aneuploidy caused less loss of β-catenin binding and degradation sites when compared with mutations not associated with aneuploidy. Both types of mutations led to loss of the EB1 binding sites. EB1 proteins mediate the interaction of APC proteins with kinetochores. They concluded that the differential loss of β-catenin binding and degradation sites appeared to affect CIN indirectly probably by disregulating apoptosis and survival.

More recently Tighe et al (2004) expressed N-terminal APC fragments in HCT-116 cells (a diploid cell line with 2 wild-type *APC* alleles) in an attempt to prove *APC* mutation could initiate CIN. Cells expressing N-APC mutants exited mitosis prematurely in the presence of spindle toxins consistent with a spindle checkpoint defect. Also N-APC cells showed enhanced survival following prolonged spindle damage. The N-APC survivors frequently contained dicentric chromosomes and went on to become highly aneuploid. These observations suggested that truncating mutations of *APC* can exert dominant effects which could initiate CIN. The *APC* mutation not only compromised tumour suppressor function but may also have oncogenic properties. The initial mutation may act as a 'double whammy' destabilising the genome and setting the stage for deregulated proliferation following the loss of the second *APC* allele.

Dikovskaya et al (2004) showed that depletion of APC from cryostatic factor Xenopus extracts led to a decrease in microtubule density and changes in tubulin distribution in spindles and asters formed in such extracts. Addition of full length protein to a large N-terminally depleted fragment or APC depleted extracts restored normal spindle morphology. They suggested that APC was important for centrosomally driven spindle formation. Lack of microtubule binding may lead to mis-segregation of chromosomes.

1.2.5 **Mechanisms of CIN**

Gollin (2005) in a review article described mechanisms leading to chromosomal instability. Structural chromosomal instability frequently results from breakage-fusion-bridge (BFB) cycles. A chromatid break occurs exposing an unprotected chromosomal end which after replication is thought to fuse with either another broken chromatid or its sister chromatid to produce a dicentric chromosome. During the anaphase stage of mitosis the two chromosomes are pulled to opposite poles forming a bridge that breaks resulting in more unprotected chromosomal ends and thus the cycle continues. The basis of these cycles is unclear.

Aberrations in the processes of chromosome segregation also lead to aneuploidy. Several factors can result in such instability, including abnormal kinetochore-spindle interactions, premature chromatid separation, centrosome amplification, multipolar spindles and abnormal cytokinesis. Gollin et al (2005) showed that chromosomal segregation defects (multipolar spindles, lagging chromosomes at metaphase and anaphase and anaphase bridging) in cancer cell lines are intrinsic heritable traits in the general tumour cell population.

Centrosomes (Saunders W 2005) are important for normal chromosomal segregation and they also play a role in controlling cell division including initiating cytokinesis and the entry into S-phase of the cell cycle. Centrosome defects were originally proposed to lead to aneuploidy and cancer by Boveri in 1914. Centrosome changes are strongly linked to aneuploidy and cancer in many studies. Under some conditions centrosomal replication can become uncoupled from the cell cycle allowing multiple centrosomes to form in a single cell. In tumour cell lines such over-replication corresponds to reduced activity of the p53 pathway. Centrosomal amplification often correlates with increased ploidy which in the absence of p53

inactivation leads to apoptosis in cell lines. Many observations link DNA damage to changes in centrosome number. Evidence shows that double strand breaks may be the signal that induces the centrosomal changes. Spindle multipolarity is strongly correlated to anaphase bridges which result in DNA breaks. Centrosome amplification could proceed through either over-replication of the centrosomes or by a cell division or cytokinesis defect leading to amplified centrosome numbers. Examples of over-replication in the absence of cytokinesis defects are rare.

The main impact of centrosome amplification on cancer cells is most likely to be the formation of multipolar spindles. Both centrosomal multipolarity and centrosome amplification are often seen in tumour tissue. However the number of mitotic cells in such tissue is generally low and strong correlations are difficult to make. Additional changes in the cell other than centrosomal amplification may be necessary to induce spindle multipolarity.

Several centrosomal proteins have been investigated. One family of proteins, the Aurora kinases, regulate centrosome function, bipolar spindle assembly, chromosomal segregation and cytokinesis (Carmena et al 2003, Meraldi et al 2004, (Katayama et al 2003). Aurora A kinases localize to the centrosome from the time that centrosomes duplicate until the end of mitosis and to a region of the microtubules proximal to the centrosomes during mitosis. Aurora B kinases are associated with the proteins Survivin and Inner Centromere Protein (INCENP) and localize to heterochromatin early in mitosis, the central spindle at anaphase, the cell cortex where the contractile ring will form and the midbody during cytokinesis. Aurora C localizes to the centrosome from anaphase to telophase. All are over expressed in cancer cells. Aurora A over expression is associated with centrosome amplification and multipolar spindles. Over expression can be as a result of regulatory alterations or gene amplification. The gene *AURKA* or *STK15* is located at 20q13.2-13.3

and this region is amplified in many cancers. Amplification of a common genetic variant of *AURKA* has been shown to be associated with CRC (Ewart-Toland et al 2003). Aurora A over expression has been shown to override spindle assembly checkpoint, resulting in arrested mitosis with incomplete cytokinesis leading to multinucleation (Anand et al 2003).

Aurora B is involved in destabilization of improper microtubule attachments and also plays a role in maintaining assembly checkpoints (Hauf et al 2003). Inhibition of Aurora B results in misaligned chromosomes, syntelic attachments of chromosomes to the spindle poles (in which both chromatids are attached to the same pole), cell division failure and endoreduplication. VX-680 inhibits all three kinases and results in an accumulation of cells with ≥4 times the normal amount of genomic DNA. It also inhibits cell proliferation leading to apoptotic cell death in many tumour types (Harrington et al 2004). Aurora C is the least well understood member of the group. It is essential for mitosis.

'Chromosome breakage' syndromes can lead to chromosomal instability. Double strand DNA breaks (DSB) represent one type of damage that is usually repaired. Failure of that response can lead to genetic alteration and chromosomal instability (Mills et al 2003) and neoplastic transformation. The DNA damage response involves the sensing of DNA damage followed by transduction of the damage signal to a network of cellular pathways from those involved in the cellular survival response, including cell cycle checkpoints, DNA repair and stress responses to telomere maintenance, and the apoptotic pathway.

The relationship between the DNA damage response genes and chromosomal instability is not clear. Haploinsufficiency of damage response genes has been associated with instability (Bassing et al 2003). ATR duplication also is shown to result in abnormal

centrosome amplification and aneulpoidy. This is the case for several other damage response genes (*BRAC1* and *MRE11A*) which are reviewed in (Bharadwaj et al 2004).

O'Hagan et al (2002) hypothesised that telomere based crisis and associated breakage-fusion-bridge cycles drive CIN in cancer cells and age related epithelial carcinogenesis. It is suggested that telomeres may be the connecting factor between mitotic instability and chromosome aberrations. Dysfunctional telomeres may be a source of mitotic instability as such dysfunction gives rise to a high rate of chromosome bridges at anaphase. These bridges can lead to structural chromosome rearrangements through chromatin fragmentation or to whole-chromosome losses through kinetochore-spindle detachment.

Telomerase positive tumours seem to maintain some degree of constant genomic flux. Possibly this base line instability may facilitate clonal evolution and adaptation to micro-enviromental challenges (Gisselsson D 2005).

It is unusual to find only one type of cell division disturbance in a tumour. Either the cell division machinery appears normal or there are a plethora of abnormalities.

Multiple mitoses and anaphase bridging are seen at variable rates indicating that there might be a gradual transition from normal bipolar mitoses to mitotic figures exhibiting various abnormalities at a high frequency. During transition from colorectal adenomas to carcinomas in Min mice the anaphase bridging frequencies increase gradually up to the stage of high-dysplasia and then decrease again in metastatic tumours (Rudolph et al 2001). This model of telomere dysfunction and anaphase bridging has the following limitations. Neither telomere dysfunction nor anaphase bridging induced in normal cells by ionising radiation will produce ongoing genomic instability at levels comparable to neoplastic cells. Disruption of DNA damage sensitive

checkpoints is necessary to prevent cells from undergoing proliferative arrest or apoptosis. A combination of driving factors is necessary to produce lasting instability.

In CRC, mutations of *CDC4* (a putative G1-S checkpoint gene) have been demonstrated in 10% of CRC. Disruption of *CDC4* in stable cell lines triggered abnormal chromosome transmission at mitosis and aneuploidy. *CDC4* mutations can be seen in the adenoma stage before telomere dysfunction. A study by Rajagopalan et al (2004) has shown that mutational inactivation of human *CDC4* and consequent up regulation of *Cyclin E* occurs at an early stage in colorectal cancer and is an important cause of chromosomal instability in these tumours. We (Kemp Z 2005), however, found that *CDC4* mutations occurred in a subset of colorectal cancers, but were not predicted to cause loss of function and were not associated with CIN.

1.2.6 **Genomically stable cancers**

Evidence exists for genomically stable cancers. Georgiades et al (1999) – using CGH, flow cytometry and MSI with 2 out of 4 markers called MSI+ – examined and found a substantial subset of tumours with neither microsatellite nor multiple major chromosomal abnormalities. Hawkins et al (2001) examined 46 sporadic cancers for CIN using flow cytometry and MSI assessed with PCR using standard markers. Immunohistochemistry was also preformed for P53 expression. 25 (54%) of tumours were aneuploid, 14 (30%) were diploid and microsatellite stable and 7 (15%) were diploid and MSI+. From our group Jones et al (2004) screened 23 MSI-CIN- colorectal cancers for gains and losses using array-based comparative genomic hybridization and compared their findings with those from a small set of MSI+CIN+ cancers, and data from MSI-CIN+ and MSI+CIN- cancers. They found little evidence of any form of genomic instability in MSI-CIN- cancers. At the level of the

chromosome arm, the MSI-CIN- cancers had significantly fewer gains and losses than MSI-CIN+ tumours. The chromosomal-scale changes found in MSI-CIN- cancers generally involved the same sites as those in MSI-CIN+ tumours, and in both cancer groups, the best predictor of a specific change was the total number of such changes in that tumour. A few chromosomal-scale changes did, however, differ between the MSI-CIN- and MSI-CIN+ pathways. MSI-CIN- cancers showed: low frequencies of gain of 9p and 19p; infrequent loss of 5q and a high frequency of 20p gain. Overall, their data suggested that the MSI-CIN- group was heterogeneous, one type of MSI-CIN- cancer having few (≤6) chromosomal-scale changes and the other with more (≥10) changes resembling MSI-CIN+ cancers. At the level of individual clones, frequent and/or discrete gains or losses were generally located within regions of chromosomal-scale changes in both MSI-CIN- and MSI-CIN+ cancers, and fewer losses and gains were present in MSI-CIN- than MSI-CIN+ tumours. No changes by clone, which were specific to the MSI-CIN- cancers, were found. Their data supported the suggestion that some MSI-CIN- carcinomas formed a qualitatively different group from the other cancer types, and also suggested that the MSI-CIN- group was itself heterogeneous.

1.2.7 **HNPCC and MSI**

Hereditary Non-Polyposis Colorectal Cancer (HNPCC) accounts for approximately 2-4% of familial CRC. In addition to early-onset CRC, presenting at a median age of 42, HNPCC patients are also prone to the development of tumours in the endometrium, ovary, stomach, small bowel, pancreas and ureter (Watson et al 1993). HNPCC patients have germline mutations in one of four DNA mismatch repair (MMR) genes, *MLH1, MSH2* (together accounting for 70% of cases), *MSH6* and *PMS2* (Fishel et al 1993, Leach et al 1993, Peltomaki et al 1993, Bronner et al 1994, Nicolaides et al 1994, Yu et al 1997). *MSH6* mutations often occur in an atypical subset of HNPCC

patients, characterised by a later age of onset and a more frequent occurrence of endometrial cancer (Peltomaki et al 2001). The MMR genes function to ensure the correct repair of DNA base pair mismatches and insertion-deletion loops arising as a consequence of slippage by DNA polymerase during replication (Modrich et al 1991). Replication slippage is most likely to occur in regions containing nucleotide repeat sequences such as microsatellites and if the mismatch repair pathway is impaired these errors will not be corrected, resulting in insertion or deletion of nucleotides. The presence of this repair defect is evidenced by microsatellite alleles of different lengths and hence is termed a microsatellite instability positive (MSI+), or mutator phenotype (Ilyas et al 1999). Loss of mismatch repair does not just affect microsatellites; all nucleotide repeat sequences are potential targets and loss of function frameshift mutations have been found in cancer associated genes which contain nucleotide repeats in their exons. Thus, through MSI, growth homeostasis and normal cell functioning is impaired and tumour progression is promoted. Approximately 10-15% of sporadic colorectal tumours are also found to be MSI+ (Peltomaki et al 1995) and 95% lack expression of either *MLH1* or *MSH2*; *MLH1* is silenced by promoter methylation in 95% of these cases (Thibodeau et al 1998).

As in CIN+ cancers, mutations in *APC* and *KRAS2* are a frequent occurrence in MSI+ cancers (Aaltonen et al 1993). *KRAS2* is mutated in HNPCC and *BRAF* in sporadic MSI+ cancers (Deng et al 2004). The prevalence of *P53* mutations is decreased (Grady et al 2002) and even those MSI+ cancers which do have a *P53* mutation, which is usually associated with aneuploidy, still maintain their diploid karyotype (Eshleman et al 1998). Mutations in a mononucleotide repeat sequence in *APC* are particularly prevalent in HNPCC patients (Huang et al 1996), suggesting that loss of MMR function precedes *APC* mutation, but in sporadic MSI+ cancers, mutations are

not concentrated in the mononucleotide repeat, suggesting that loss of MMR does not occur until after *APC* mutation in the sporadic form of the disease (Homfray et al 1998).

It appears that the similar pathways may be affected in CIN+ cancers as are affected in MSI+ cancers, with deregulation achieved by instability of sequence rather than structure. For instance, frameshift mutation in the A10 tract of *TGFBR2*, resulting in protein truncation, is a frequent event in MSI+ cancers, whereas inactivation of *SMAD4*, which function downstream in the TGFβ pathway, occurs in CIN+ cancers (Davies et al 2002, Rajagopalan et al 2002, Yuen et al 2002). While mutation of *p53* occurs at a lower frequency, mutation of *BAX*, which encodes a *p53* dependent inducer of apoptosis, is a frequent occurrence in MSI+ cancers. Furthermore, mutation of *BRAF,* encoding a Ras-regulated kinase has been associated with MSI+ cancers and is particularly prevalent in cases where *KRAS2* mutations are absent, again implicating abrogation of the same pathway but by alternative mechanisms.

1.2.8 **Methylator phenotype**

It has been noted that a subset of CRCs, largely sporadic MSI+ cancers (Ahuja et al 1997), display a CpG island methylator phenotype (CIMP) in which gene inactivation by promoter hypermethylation is a common event (Toyota et al 1999, Yamamoto et al 2002). For example, mutations in the MMR genes are rarely detected in sporadic MSI+ CRC; rather, hypermethylation of the *MLH1* promoter region and consequent loss of transcription and protein expression appears to be the major mechanism of microsatellite instability (Kane et al 1997, Cunningham et al 1998, Herman et al 1998, Veigh et al 1998, Wheeler et al 1999, Kuismanen et al 2000). Epigenetic silencing of a range of tumour associated

genes has now been identified. For example, hypermethylation of *P16* is found in 60% of MSI+ cancers as compared to 22% of CIN+ cancers (Ahuja et al 1997); P16 regulates G_1 cell cycle progression and stabilises P53 by sequestering MDM1. Furthermore, aberrant promoter methylation of *P14*, an alternatively spliced form of *P16*, has been found to be associated with microsatellite instability and also with absence of *P53* mutation (Shen et al 2003), providing another possible alternative to mutation of *P53* in MSI+ cancers. An increased frequency of inactivation by methylation in MSI+ cancers is also detected in the cases of Thrombosin-1 (TSP-1), an angiogenesis inhibitor (27% vs 0%) and the Insulin Growth Factor II (IGFII*)* (60% vs 6%) (Ahuja et al 1997). Overall, 60% of MSI+ cancers display hypermethylation of two or more gene loci compared to only 9% of CIN+ cancers.

1.2.9 **FAP and the APC gene**

We now leave the discussion of genomic instability to consider in more detail the gene related basis of colorectal cancer. Familial Adenomatous Polyposis (FAP) is an autosomally dominantly inherited disease affecting 1 in 8000 individuals, accounting for ~1% of CRC and is caused by germline mutations in the Adenomatous Polyposis Coli (*APC*) gene. FAP patients develop hundreds of colorectal adenomatous polyps in the second and third decades of life. They may also develop duodenal polyps and multiple extraintestinal manifestations. These may be divided into changes of ectodermal origin (epidermoid cysts, pilomatrixoma, central nervous system tumours and congenital hypertrophy of the retinal pigment epithelium), mesodermal origin (desmoids tumours, bone tumours and dental malformations) and those of endodermal origin (as well as gut lumen adenomas adenomas of the biliary tract, thyroid and adrenal cortex, and hepatoblastomas). Patients unless treated have a 100% chance of developing CRC (Inherited bowel cancer. Clark S, Colorectal Surgery Third edition (2005) Elsevier Saunders).

A mutation in the *APC* gene can only be identified in 80% of individuals. A milder form of FAP- Attenuated FAP also exists in which patients demonstrate fewer polyps than in classical FAP. Study of FAP has provided useful insights into sporadic CRC since mutation of *APC* is also the earliest detected genetic event in sporadic CRC, occurring in up to 80% of cases (Ilyas et al 1999).
APC is located on chromosome 5q21 and encodes a large, multifunctional 312kD protein that participates in several cellular processes, such as cell adhesion and migration, signal transduction, microtubule assembly and chromosome segregation (Fodde et al 2001). However the main tumour suppressing function of *APC* resides in its capacity to regulate intracellular β-catenin levels.

β-catenin functions as a WNT pathway transducer and therefore *APC* has a function in controlling this pathway. The majority of *APC* mutations are nonsense or frameshift mutations that result in a truncated protein product with abnormal function (Fearnhead et al 2002). As well as the initial germline *APC* mutation loss or mutation of the second *APC* allele is necessary for tumorigenesis (Knudsen's two hit hypothesis (Knudson 1996)).The type of germline *APC* mutation in FAP appears to determine the nature of the somatic hit. If the germline mutation occurs between codons 1194 and 1392 there is a strong selection for allelic loss of *APC* as the second hit (Lamlun et al 1999). If the second hit occurs outside this region the second hit is likely to be a mutation in the somatic mutation cluster region.

Phenotypes of FAP can vary according to the location of the inherited *APC* mutation with a particularly severe phenotype seen in those with mutation between codons 1250 and 1464, especially at codon 1309, (Nagase et al 1992). Attenuated polyposis is attributed to mutations at the extreme 5' or 3' ends of the gene (Spirio et al 1993). Mutation selection appears to favour a sufficient degree of impairment of *APC*

function to allow nuclear accumulation of β-catenin and consequent downstream signalling but is selective against complete loss of regulation, perhaps because excessive accumulation of β-catenin in the nucleus can result in apoptosis (Kim et al 2000). Approximately 6% of the Ashkenazi Jewish population carry a germline I1307K polymorphism in *APC* and have approximately twice the risk of developing CRC (Laken et al 1997). Lastly, the *APC* promoter may be hypermethylated, leading to loss of transcription.

1.2.10 **WNT signalling**

APC functions as a negative regulator of the WNT signalling transduction pathway through its ability to control the levels and intracellular location of β-catenin, as depicted in Figure 1.1.

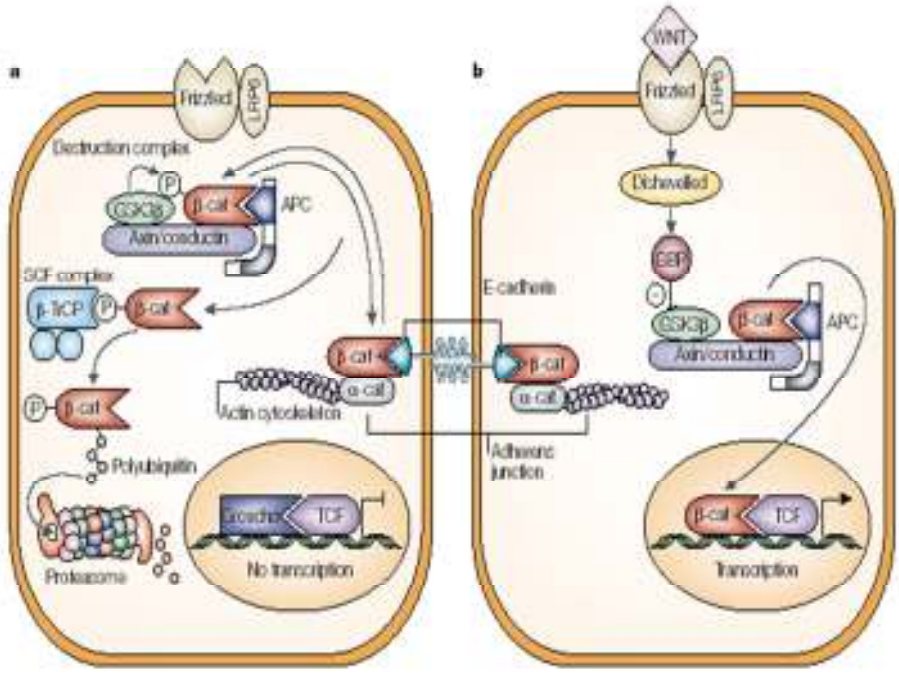


Figure 1.1. The WNT signalling pathway (Fodde et al 2001) a) in the absence of WNT signalling β-catenin is targeted for degradation b) in

the presence of WNT signalling β-catenin shuttles to the nucleus and activates gene expression.

In the absence of WNT signalling, a "destruction complex" comprising *APC*, the scaffolding proteins Axin and Conductin and Glycogen synthase kinase 3β (GSK3β) binds to β-catenin in the cytoplasm (Behrens et al 1998, Hart et al 1998, Fagotto et al 1999, Kishida et al 1999). GSK3β phosphorylates β-catenin (Ikeda et al 2000), which allows it to be recognised by an SCF complex with E3-ubiquitin ligase activity which transfers Ubiquitin to β-catenin, labelling it for degradation by the proteasome (Jiang et al 1998, Marikawa et al 1998). Axin may function by facilitating the phosphorylation of β-catenin by GSK3β (Hart et al 1998, Ikeda et al 2000).

In the presence of WNT the Frizzled receptor is activated, which in co-operation with Low-density-lipoprotein-related protein 6 (LRP6), leads to the inactivation of GSK3β in the destruction complex. GSK3β inactivation is thought to occur via Dishevelled, which can activate GBP, an inhibitor of GSK3β. In addition, it has also been suggested that LRP6 may bind and inhibit Axin (Kishida et al 1999). Stabilised β-catenin transfers to the nucleus where, in complex with the T-cell factor (TCF) family of DNA binding proteins, it activates the transcription of TCF-responsive genes (Behrens et al 1996, Molenaar et al 1996). When no WNT signal is present, TCF/LEF are bound and sequestered by the Groucho family of transcriptional repressors (Roose et al 1998, Brantjes et al 2001). The TCF/LEF family of transcription factors consists of four members: TCF1, LEF1, TCF3 and TCF4, the latter being the major binding partner of β-catenin in the colon.

When *APC* is mutated, it loses its ability to bind and inactivate β-catenin , therefore, β-catenin is constitutively active even in the

absence of WNT signalling and free to continue to transactivate its target genes, resulting in uncontrolled cell proliferation.

Additional components of the pathway are still being identified, such as the Casein kinases I and II, which are thought to dephosphorylate and inactivate Dishevelled (Polakis et al 2000) and the SFRP family, which are found to silence WNT signalling by binding the WNT proteins and preventing their interaction with the Frizzled receptors (Suzuki et al 2002, 2004).

1.2.11 **Beta-catenin**

The induction of β-catenin/TCF signalling can induce transcription of a range of genes influencing cell fate, proliferation and tumour progression. For example, the first identified transactivation targets were *MYC* (He et al 1998) and *CCND1* (Shtutman et al 1999), the over-activity of these genes resulting in uncontrolled cell cycle progression (Tetsu et al 1999). In addition the expression of *AF17*, a fusion partner of *MLL* in acute leukaemias, which promotes cell cycle progression at the G_2-M transition, is also induced. Two components of the AP-1 transcription complex, *C-JUN* and *FRA-1* have been identified as targets, along with the AP-1 activated urokinase-type plasminogen activator receptor (*uPAR*), which has a predicted role in invasive growth and metastasis formation (Mann et al 1999). Genes which code for the multidrug resistance 1 protein (*MDR1*), which promotes tumorigenesis by suppressing programmed cell death (Yamada et al 2000) and *WISP-1*, which can attenuate *P53* mediated apoptosis in response to DNA damage (Xu et al 2000), are also downstream targets of β-catenin/TCF signalling. In addition, increased expression of the genes *ENC1*, *MET*, *CD44*, *MMP-7* and the γ2 chain of *Laminin-5* have been demonstrated in response to β-catenin induction. *ENC1* encodes the ectodermal-neural-cortex 1 protein, which can inhibit differentiation through reorganisation of the

actin cytoskeleton (Fujita et al 2001); the receptor tyrosine kinase *MET* is a regulator of cell growth, motility and survival through its ligand HGF and a crucial player in tumour invasion and metastasis (Rong et al 1994, Boon et al. 2002); CD44 promotes *MET* signalling (Mann et al 1999); the matrix metalloproteinase matrilysin *MMP-7* functions in the breakdown of the extracellular matrix (Brabletz et al 1999, Crawford et al 1999) and the γ2 chain of *Laminin-5* induces epithelial cell migration (Hlubek et al 2001).

In addition to its role as a transactivator of transcription, β-catenin is an essential component of adherins junctions, which tightly seal the space between epithelial cells. β-catenin links the cell adhesion protein E-cadherin to α-catenin and the actin cytoskeleton (Kemler et al 1993). The activity of E-cadherin can affect cell migration and morphogenesis and reduced expression levels have been noted in CRC (Ilyas et al 1997). Thus, through β-catenin, *APC* can also modulate intercellular adhesion.

It is thought that the role of *APC* as a negative regulator of the WNT signalling pathway constitutes its major function as a tumour suppressor. This is supported by the fact that ~50% of CRCs with wild type *APC* contain activating mutations in *β-catenin* (Morin et al 1997, Sparks et al 1998). Most *β-catenin* mutations affect the GSK3β phosphorylation sites and thus provide resistance to degradation (Morin et al 1997, Polakis et al 1999). However, adenomas containing *β-catenin* mutations are not as likely to progress to carcinomas as those with mutant *APC*, suggesting that *APC* and *β-catenin* mutations are not functionally equivalent and that *APC* is likely to have other tumour suppressor functions (Samowitz et al 1999). For example, in addition to the regulation of *β-catenin*, it

has been demonstrated that *γ-catenin*, which can also function as an oncogene, is a target of *APC* regulation. *β-* and *γ-catenin* have analogous structures and functions and it was found that *γ-catenin*, in co-operation with TCF/LEF factors, activates *C-MYC* expression more strongly than *β-catenin* (Kolligs et al 2000). Thus, *γ-catenin* provides an additional route by which the impairment of *APC* function results in the deregulation of WNT signalling.

1.2.12 **KRAS2, SMAD 2, SMAD4, p53**

In addition to *APC*, mutations are also frequently found in the *KRAS2* oncogene, located on chromosome 12p. *KRAS2* couples growth factors to the MAP kinase signal transduction pathway, promoting cell proliferation. Gain of function mutations are found in up to 50% of sporadic CRCs and in synergism with mutant *APC*, *KRAS2* is thought to promote tumour progression (Forrester et al 1987, Bos et al 1988).

Loss of heterozygosity (LOH) of chromosome 18q and in particularly 18q21-qtel is also a frequent event in CRC, occurring in approximately 70% of cases (Thiagalingam et al 1996). However, the incidence of LOH of 18q is only ~10% in early adenomas suggesting the importance of this event at later stages of cancer progression (Vogelstein et al 1988). *SMAD2* and *SMAD4*, two members of the TGF-β signal transduction pathway, which has a pronounced growth inhibitory effect in the colon, map to this region and mutations in these genes have been found in 16% (*SMAD4*) (Takagi et al 1996) and 5% (*SMAD2*) (Takenoshita et al 1998) of CRC.

The transcription factor p53, located on chromosome 17p, is involved in maintaining genomic stability and has been called the “guardian of the genome” (Lane et al 1993) due to its role in control of the cell cycle and apoptosis in response to genotoxic stress. p53 activates target genes such as *p21*, *BAX*, and *GADD45,* stalling DNA replication while lesions in the DNA are repaired or, if the degree of damage is severe, inducing apoptosis (Somasundaram et al 2000). p53 also activates the expression of MDM2, which targets p53 for ubiquinylation and degradation in a negative feedback loop (May et al 1999). Mutations in *P53* have been found in up to 70% of sporadic CRC, however mutations are rarely found in benign lesions and it has been suggested that loss of p53 function is required for tumour progression rather than initiation (Vogelstein et al 1988, Rodrigues et al 1990).

1.3 **Copy number changes**

The maintenance of normal cell growth depends upon the fine balance of expression of growth promoting and growth suppressing genes. In cancer, growth promoting genes, or “oncogenes”, can become over-expressed through mutation or an increase in number, leading to increased cell proliferation. Alternatively, a decrease in expression of growth suppressing genes can also lead to inappropriate cell growth, or a reduction in cell death, such genes being termed “tumour suppressors” (Kinzler et al 1997). While the over-activity of one copy of an oncogene is sufficient to disrupt the homeostatic balance, loss of both copies of a growth suppressing gene is necessary to alter the growth rate. Inactivation of tumour suppressor genes commonly follows Knudson’s “two-hit” hypothesis

(Knudson 1996); for instance, by the deletion of one allele and inactivation of the other by mutation or promoter hypermethylation. The accumulation of multiple gene alterations results in the progressive change from normal to uncontrolled growth, malignancy, invasion into the surrounding tissue and metastasis.

One way in which the activity of oncogenes and tumour suppressors can be disrupted is by alterations in DNA copy number, ranging from gain or loss of entire chromosomes to amplification or deletion of discrete sections of DNA (Albertson et al 2003). An increase in the number of copies of a region of the genome can lead to increased expression of the genes contained within the region; conversely, deletion can lead to reduction or loss of expression of genes located within the region of loss (Sawyer W 1983). Haploinsufficency describes the position when one copy of a dosage sensitive gene has been lost and such loss leads to phenotypic effects; this is to say that the expression products of a single allele are not sufficient for normal function. Not all loss of a single allele will lead to this and therefore not all genes are haploinsufficient. Dominant loss of function mutations are the result of mutations in haploinsufficient genes. Parallel measurement of DNA copy number and mRNA expression levels in breast tumours has revealed that 62% of genes mapping to regions of amplification show elevated expression levels and that, on average, a 2-fold change in copy number is associated with a 1.5 fold change in mRNA levels (Pollack et al 2002). For example, in colorectal cancer, gain of chromosome 17q has been found to correlate with over-expression of the epidermal growth factor receptor (ERBB2) (Knosel et al 2002). The importance of deletions in tumour genomes has also been shown, for instance in the inactivation of tumour suppressor genes such as *PTEN* in prostate cancer (Vogel F 1979). Furthermore, the identification of copy number changes has directly led to the discovery of genes important in tumorigenesis such as the retinoblastoma susceptibility

gene (*RB1*), which was identified as a consequence of its location within a region of loss on chromosome 13q14 (211-213). Overall, however, the functional significance of copy number changes without amplification, homozygous deletion, mutation or methylation are unknown.

1.3.1 **Comparative Genomic Hybridisation**

In terms of measuring genomic instability a number of techniques exist. Detection of LOH using probes at specific loci will reveal LOH but says nothing about the genesis of that LOH which may, as we have seen, be due to chromosomal loss, part chromosomal loss of chromosomal loss and replication. FACS will reveal gross alteration in the amount of DNA while revealing nothing of the site of that change within the genome. The identification of copy number changes at pre-defined loci has historically been achieved by Fluorescent In Situ Hybridisation (FISH), using specific DNA probes of known chromosomal location hybridised to metaphase chromosomes. For example, amplification of *CCND1* at chromosome 11q13, detected by FISH, is associated with metastasis and reduced survival in squamous cell carcinoma of the head and neck (Alavi et al 1999). However, the development of the technique of Comparative Genomic Hybridisation (CGH) (Kallioniemi et al 1992) has enabled screening for copy number changes on a genome wide scale without the need for metaphase tumour chromosomes, which are difficult to obtain in large numbers, or prior knowledge of specific regions of interest. In CGH, genomic DNA from a tumour sample and a normal control are labelled with two different fluorochromes and co-hybridised to normal metaphase chromosome spreads. The differences in tumour and control fluorescence intensities along each reference chromosome are reflections of the copy number status of the corresponding

sequences in the tumour genome. If a region is present in equal copy number in the normal and tumour genomes, an equal contribution from each fluorochrome is seen, however if additional or fewer copies of a region are present in the tumour genome, this will be reflected in an increased or decreased ratio of fluorescence signal from the tumour DNA compared to the normal control.

CGH investigations of tumour genomes have led to the identification of genes important in the development and progression of cancer (Lichter et al 2000). For example, CGH studies in patients with Peutz-Jeghers syndrome, which confers an increased risk of gastrointestinal tumours, resulted in the localisation of a susceptibility locus to chromosome 19p13.3 and the consequent identification of the *STH11* tumour suppressor gene (Hemminki et al 1998). In addition, a CGH study of oesophageal cancer identified chromosome 9p23-p24 as a common region of amplification, leading to the cloning of a novel oncogene, *GASC1*, which encodes a transcriptional regulator involved in chromatin remodelling (Yang et al 2000).

1.3.2 **Array Comparative Genomic Hybridisation**

A major disadvantage of conventional CGH is the limited resolution, which at its best is approximately 3-5Mb (Kirchhoff et al. 2001). Thus, any unbalanced region detected by conventional CGH is likely to contain many potential candidate genes for the pathogenesis of the disease. In matrix (Solinas-Toldo et al. 1997) or array (Pinkel et al. 1998) CGH, large insert clones (BACs, PACs and cosmids) spotted onto glass slides are used as the hybridisation target in the place of metaphase chromosomes. In this way, resolution can be greatly improved and is determined by the insert size of the clones and the density with which they are selected from the genome. More recent developments include platforms that use smaller clones and oligonucleotides (Dutt et al, 2007).

The use of genomic clones for array construction is not straightforward as large insert clones produce a low DNA yield, making extraction of sufficient quantities of DNA labour intensive. Therefore, strategies have been developed to amplify small amounts of purified clone DNA, to give a reliable representation of the template DNA sequence and produce high quantities of DNA for spotting. Firstly, ligation-mediated PCR, which utilises a frequently cutting restriction enzyme (producing fragments of 200bp to 2kb) and ligation of a universal oligonucleotide adaptor to provide a priming site for PCR amplification, was used to generate arrays and accurately detect copy number changes (Pinkel et al 1998). The use of cDNA arrays, developed for the analysis of gene expression, has also been demonstrated for the investigation of genomic copy number changes (Pollack et al 1999). This approach allows the parallel analysis of copy number and expression levels, but analysis is limited to well characterised genes, excluding regulatory elements and as yet unidentified transcripts. In addition, while this type of array can identify high level copy number changes, low level changes are difficult to reliably detect.

Array CGH on a genome-wide scale was first demonstrated with the assembly of a genomic microarray comprising ~2400 BAC clones, providing an average resolution of approximately 1.4Mb and enabling the reliable detection of copy number changes, ranging from single copy losses to homozygous deletions (Snijders et al 2001).

The identification of copy number changes by array CGH has been used to investigate genomic imbalances present in a variety of cancers. For example, clustering of copy number changes has enabled the successful diagnosis of different types of renal cell cancer (Wilhelm et al 2002), discrimination between liposarcoma subtypes (Fritz et al 2002) and has been used to distinguish gastric cancer patients with a high risk of lymph node metastasis and consequently a decreased likelihood of survival (Weiss et al 2003).

In addition, copy number changes containing genes which may be of relevance to the pathogenesis of the disease have been identified by array CGH such as gains of chromosome 12q13-q15, including *MDM2,* in pulmonary artery intimal sarcomas (Zhao et al 2002), chromosome 3q22-q26, containing *PIK3CA,* in non-small cell lung cancer (Massion et al 2002) and loss of 10q23.2 and *PTEN* in bladder cancer (Hurst et al 2004). The capability of the technique in providing even higher resolution mapping of variation in copy number has been demonstrated in the analysis of breast tumours (Albertson et al 2000). A previously reported 2Mb region of copy number gain on chromosome 20q13.2 was analysed using an array of overlapping large insert clones covering the region. Two distinct regions of amplification (of ~50kb) were resolved, centred upon two potential oncogenes, *ZNF217* and *CYP24*. Thus, the improved resolution of array CGH enables the fine structure of copy number aberrations to be resolved and can identify potential candidate genes, likely to be directly involved in the development of the tumour.

However, two problems in the analysis of copy number change have become apparent (Freeman et al 2006). One is the appreciation that copy number change is a feature of normal individuals who are essentially healthy and show no obvious genetic disorders. These changes represent germline copy number variation. The second is that different allelic dosages may produce phenotypic effects and that array CGH measures the total number of alleles and cannot distinguish between copy number changes of different alleles of the same gene. Further in some cases where copy number change is observed modification of the phenotypic effects of that dosage change by changes at other genomic sites may mean that such dosage change will not be consistent phenotypically across all individuals in whom such genomic change is observed. Efforts have been made to document common sites of copy number variation in normal individuals and also to develop techniques which allow the assessment of individual allelic copy number change.

An important consideration in the analysis of cancer material by array CGH is that often only limited amounts of DNA are available for analysis. However, it has been demonstrated recently that, using a modified ligation-mediated PCR protocol and labelling the DNA using two nucleotides conjugated to the same fluorophore (rather than the traditional one), single copy number changes can be detected with only 1ng of genomic DNA as starting material. In addition, a remaining limitation in the use of genomic clones for array construction is that the complete sequence of the clone insert can contain common repetitive elements such as Alu and LINE repeats and regions of homology to other parts of the genome, potentially leading to difficulty in the interpretation of array data (Buckley et al 2002, Fiegler et al 2003). Recently, arrays have been developed which allow the pre-selection of the exact sequence which is spotted on the array by using PCR primers designed to amplify only the unique, repeat free sequence present in the clone (Mantripragada et al 2003). This approach is especially useful in the analysis of aberrations located in regions rich in low copy repeats and has been used to identify deletions at a resolution of 15kb in the Di George syndrome region of chromosome 22 (Mantripragada et al 2003). Detection of copy number changes using oligonucleotide arrays, hybridised with genomic “representations” created by cleavage of the genome with restriction enzymes and amplification of the resulting fragments, has also been demonstrated (Lucito et al 2003, Bignell et al 2004). Lucito et al(2003) employed ROMA (representational oligonucleotide microarray analysis) methodology using an array of 70mer oligonucleotides, giving an average genome-wide resolution of ~30Kb, while Bignell et al (2004) utilised oligonucleotide arrays originally designed to detect single nucleotide polymorphisms (SNPs), allowing the parallel identification of copy number changes and loss of heterozygosity (LOH) in cancer genomes. Genomic microarrays have also been used in the analysis of breakpoints of aberrant chromosomes (Fiegler et al 2003).

Various commercially available oligonucleotide array CGH (oaCGH) platforms exist (Ylstra et al 2006). Affymetrix is a commercial oaCGH platform, which contains short 25mer oligonucleotides photo-lithograhically synthesized on the arrays (http://www.affymetrix.com/). These are single channel arrays, which means that only test DNA needs to be labelled and hybridized. Another commercial oaCGH platform was introduced by Agilent Technologies (http://www.agilent.com/). They evaluated their original expression arrays for this purpose as well as arrays designed specifically for aCGH, which include oligonucleotides covering intergenic regions. Both Agilent array platforms consist of 60mer oligonucleotides. The labelling protocol is similar to the one used for the cDNA arrays and requires 1 mg of input DNA which hampers the use of small clinical samples. To overcome this problem, a PCR amplification procedure was developed allowing as little as 10 ng of input DNA. A third oligonucleotide platform offered commercially is by NimbleGen (http://www.nimblegen.com/). They provide arrays containing 385 K oligonucleotides photolithograhically synthesized on the array. The array production is extremely flexible such that each array produced can have a different set of oligonucleotides on it.

1.4 **Prognosis in colorectal cancer.**

Colon cancer is staged pathologically on the basis of the extent of invasion of the primary cancer and the metastatic spread to lymph nodes or distant organs. The Dukes classification is commonly used to describe the stage of colorectal cancers. A Dukes B cancer describes one that has invaded through the muscularis propria of the bowel wall but shows no evidence of lymph node involvement or metastatic spread. 25-30% of Dukes B patients are expected to develop tumour relapse. It would be desirable to identify at the time of treatment which cancers will progress and target adjuvant treatment at this group sparing patients from adjuvant treatment whose cancers are unlikely to progress.

In a consensus statement from the College of American Pathologists in 1999 (Compton et al 2000) prognostic factors for colorectal cancer were presented as Categories I-IV according to strength of evidence. Category I factors were the local extent of tumour assessed pathologically; regional lymph node metastasis; blood and lymphatic vessel invasion; residual tumour following surgery with curative intent and preoperative carcinoembryonic antigen (CEA) elevation. Category IIA factors were tumour grade; radial margin status; and residual tumour in the specimen following neoadjuvant therapy. Category IIB factors were histologic type; histologic features associated with microsatellite instability (MSI) (host lymphoid response, and medullary or mucinous type); high degree of MSI; loss of heterozygosity (LOH) of 18q; tumour border configuration. Category III factors were DNA content; all other molecular markers except LOH of 18q and MSI-H; perineural invasion; microvessel density; tumour associated proteins or carbohydrates; peritumoural fibrosis; peritumoural inflammatory response; focal neuroendocrine differentiation; nuclear organising regions and proliferation indices. Category IV factors were those shown to have no relation to prognosis and were tumour size and gross tumour configuration. It may be seen that the highest placed molecular markers (LOH 18q and MSI-H) ranked below the standard pathological stage in this hierarchy.

1.4.1 **Clinicopathological factors**

The gold standard of prognostication remains clinicopathological stage (Maughan et al 2002). This is to say invasion of the tumour through the bowel wall, peritoneal involvement, the presence of lymph node involvement or distant spread of cancer.

Ratto et al (1998) reviewed the literature relating various clinicopathological variables to outcome. Patient gender was extensively evaluated (Wied et al 1985, Fielding et al 1986, Scott et

al 1987, Shepherd et al 1989, Hermanek et al 1989, Albe et al 1990, Karenaga et al 1991, Garcia-Peche et al 1991, Ponz de Leon et al 1992, Lasser et al 1993, Hermanek et al 1994, Deans et al 1994, Ikeda et al 1994, Crucutti et al 1991), in the majority of studies, it was of no significance in predicting survival independently of other factors. In only 4 of 18 studies did patient gender reach statistical significance. In three of these studies (Chapuis et al 1985, Griffin et al 1987, Newland et al 1994), prognosis was better in females than in males, and in the fourth study (Schmitzd-Moormann et al 1987), the prognosis was not specified. Gender was found to be significantly related to survival in the groups own patients showing a better prognosis for females.

Considering patient age results are more discordant. In 10 of 17 studies (Fielding et al 1986, Hermanek et al 1989, Albe et al 1990, Karenaga et al 1991, Garcia-Peche et al 1991, Ponz de Leon et al 1992, Ikeda et al 1994, Crucitti et al 1991, Wiggers et al 1988, Laurent-Puig et al 1992) when this parameter was evaluated, it was not found to be an independent prognostic variable. In 6 studies it was found to be of importance (Wied et al 1985, Scott et al 1987, Lasser et al 1993, Chapuis et al 1985, Griffin et al 1987, Newland et al 1994).

Most of the older studies documented a worse five year survival rate for patients with rectosigmoid or rectal cancers than for those with tumours located elsewhere (Spratt et al 1967, Godwin et al 1975, Copeland et al 1968, Wotmark et al 1983, Dwight et al 1969). However, a large amount of more recent evidence from multivariate analyses (Fielding et al 1986, Scott et al 1987, Shepherd et al 1989, Albe et al 1990, Karenaga et al 1991, Garcia-Peche et al 1991, Ponz de Leon et al 1992, Lasser et al 1993, Hermanek et al 1994, Deans et al 1994, Ikeda et al 1994, Crucutti et al 1991, Chapuis et al 1985, Griffin et al 1987, Schmitzd-Moormann et al 1987, Wiggers et al 1988, Laurent-Puig et al 1992, Steinberg et al 1986) suggests no

significant influence of tumour site on long-term outcome. In only 3 of 21 reports was location of the tumour found to be a prognostic factor; 2 of these reports (Hermanek et al 1989, Fredman et al 1984) considered rectal cancer exclusively, reporting a decreased survival rate in lower-third rectal locations. In the third study, (Newland et al 1994) colonic localization of tumours had a better prognosis than rectal.

A large number of studies suggest that tumour stage, based on intramural spread, lymph node involvement, and presence of distant metastases, is the most important independent prognostic factor (Scott et al 1987, Hermanek et al 1989, Albe et al 1990, Ponz de Leon et al 1992, Hermanek et al 1994, Deans et al 1994, Crucitti et al 1991, Chapuis et al 1985, Griffin et al 1987, Wiggers et al 1988, Laurent-Puig et al 1992). In only 3 of 15 multivariate analyses (Wied et al 1985, Korenaga et al 1991, Kohal et al 1989) in which tumour stage was considered was it not independently related to outcome. However, the elements composing the TNM classification have been considered individually only in a few studies analyzing tumour stage with conflicting results. Among nine studies in which T stage was evaluated it was found to be of prognostic significance in six of them (Fielding et al 1986, Hermanek et al 1989, Garcia-Peche et al 1991, Newland et al 1994).

Blood vessel invasion, particularly referring to venous vessels, has been found to be significantly related to tumour stage and grading (Minsky et al 1989, Khankhanien et al 1977, Heald et al 1986) and an independent prediction factor in many analyses, performed on colonic (Wied et al 1985, Minsky et al 1989) rectal (Freedman et al 1984, Horn et al 1991) or colorectal tumours (Garcia-Peche et al 1991, Ikeda et al 1994, Chapuis et al 1985, Newland et al 1994) Lymphatic vessel invasion also increases with stage and grade and was shown to be of prognostic significance in most analyzed studies

(Hermanek et al 1989, Ponz de Leon et al 1992, Minsky et al 1989, Michelassi et al 1990) both in colonic and rectal carcinomas.
The role of tumour differentiation degree in outcome prediction has been widely investigated. However, conflicting results have been reported, most probably
because of the complex evaluation criteria that are not well standardized. Among 24 studies included in the multivariate analysis to assess the value of tumour grading in outcome prediction, only 10 studies (Wied et al 1985, Scott et al 1987, Albe et al 1990, Garcia-Peche et al 1991, Deans et al 1994, Chapuis et al 1985, Griffin et al 1987, Newland et al 1994, Wiggers et al 1988, Selby et al 1992) found it to be an independent prognostic factor. In comparison with tumour stage, grading did not add any significant advantage in survival prediction. Indeed, when both
parameters were considered only in one study (Wied et al 1985) was grading found to be an independent factor, but stage was not; in four studies (Ponz de Leon et al 1992, Crucitti et al 1991, Laurent-Puig et al 1992, Quirke et al 1987) stage was independently related to outcome, whereas grading was not; in two studies (Karenaga et al 1991, Kohal et al 1989) neither grading nor stage were prognostic factors; and finally, in seven studies (Scott et al 1987, Albe et al 1990, Deans et al 1994, Chapuis et al 1985, Griffin et al 1987, Wiggers et al 1988, Selby et al 1992), both variables were able to predict outcome.
The Petersen index (Petersen et al 2002) is a prognostic model consisting of clinicopathological features which has been shown to be of use in relation to survival for patients with Dukes B cancers. This model uses tumour perforation, peritoneal involvement, venous spread, and surgical margin involvement. Dukes B cancers with none of these characteristics had a comparable prognosis to Dukes A cancers while the presence of high risk factors reduced the five year survival to 49.8%. This index has been independently verified in a further study (Morris et al 2007).

Other clinical factors have been found to be independent prognostic markers. These include presentation type, for example, obstruction (Ratto et al 1998), intraoperative factors such as the surgeon and the use of total mesorectal excision (MacFarlane et al 1993) and whether blood transfusion was required. The quality of clinical care, surgery and pathology taken together impacts on survival with the hospital the patient is treated in effecting survival (McArdle et al 2002). The quality of the surgical specimen, as assessed by circumferential involvement and the macroscopic appearance of the specimen (Birbeck et al 2002), is an important audit role of the pathologist. Such pathology must be carried out to a high standard and the pathologist must correctly diagnose nodal involvement (Tepper et al 2001).

Clearly in order that novel prognostic factors are correctly assessed the quality of the surgical resection, pathological examination and general care must be high so that these variables do not unduly influence survival and skew survival results.

1.4.2 **Genetic factors.**

Many somatic genetic factors have been investigated in relation to prognosis, the principle ones of significance being 18q LOH and MSI. Prognostic information relating to other genetic factors known to be related to CRC is given below. The prognostic data relating to aneuploidy, chromosome 5 LOH and MSI is given in more detail later in the appropriate discussion sections.

18q LOH has been associated with poor prognosis. In a systematic review and meta-analysis of the relationship between chromosome 18q genotype, DCC status and colorectal cancer prognosis Popat and Houlston (Popat et al 2005) found after reviewing 27 eligible studies that 17 studies showed cancers with chromosome 18q loss

appear to have a poorer prognosis. However they did note a lack of consistent methodology.

It has been previously shown that more than 75 percent of colorectal cancers show 17p LOH (Allen et al 2005). The 17p segment that is lost contains the tumour suppressor gene *P53*. A number of studies have assessed p53 expression in relation to prognosis and found no prognostic role in early stage CRC (Allegra et al 2003, Giatromanolaki et al 1999, Soong et al 1997, Bhatavdekar et al 2001). P53 overexpression demonstrated by immunohistochemistry has been related to poor prognosis, but other studies, especially those using PCR-SSCP, do not always confirm this relationship (Allen et al 2005). Ilyas (Ilyas et al 1996) compared Dukes B cancers and found those with recurrence to have 56.9% positive p53 staining compared to 53.3% for those cancers not associated with recurrence.

P53 mutation has been associated with poorer survival. Mutations which lengthen its half life have been shown to lead to worse outcome (Mcleod et al 1999). If both *KRAS* and *P53* mutations are found within a tumour these may have an additive effect on reducing survival.

Specific *KRAS2* mutations have been shown to be linked to poorer survival (Kirsten ras in Colorectal Cancer Collaborative Group 2001). The RASCAL I and II meta-analyses examined the importance of *KRAS2* mutation and found that different gene mutations had different effects on survival even if mutations occurred at the same site. Recently it has been shown that mutations of codon 12 and 13 of exon 2 of the gene render a cancer resistant to the anti epidermal growth factor receptor drugs Cetuximab and Panitumumab (Plesec et al 2009).

MSI+ve CRCs may have an improved prognosis (Haydon et al 2002, Elsaleh et al 2001). Genes carrying short repetitive sequences tend to be mutated in MSI+ tumours: *CTNNB1, BAX, TCF4, CDX2, E2F4, TGFBR2* and some other mismatch repair genes (Haydon et al 2002). None of these genes has been investigated for prognostic significance.

Aneuploidy has been shown to be correlated with worse prognosis (Barratt et al 2002, Quirke et al 1987, Schutte et al 1987).

Increased Bcl-2 expression detected by immunohistochemistry may be associated with better prognosis (McLeod et al 1999). Studies (Ilyas et al 1996, Meterissian et al 2001) have shown a relationship between enhanced Bcl-2 expression and improved survival. Ilyas (Ilyas et al 1996) found positive Bcl-2 staining in 69% of non-recurrent Dukes B cancers and 33% of recurrent cancers (In this work 37.5% of >5 year survival patients showed positive Bcl-2 staining).However, other research is less convincing in terms of this relationship (Barratt et al 2002).

Angiogenesis promoting factor VEGF is over expressed in 50% of tumours (Papamichael et al 2001). VEGF+ve tumours may have a worse prognosis than VEGF–ve tumours (McLeod et al 1999, Papamichael et al 2001).

E-cadherin expression is lost in tumour cells leading to decreased adhesion and facilitating metastasis. A recent study from Lugli et al (2007) examined the relationship between E-cadherin and β-catenin expression and prognosis in a large number of tumours (1420). For MSI- cancers, increased nuclear β-catenin expression and loss of membranous E-cadherin were independently associated with higher N stage, vascular invasion, and worse survival. Additionally there was a relationship between loss of membranous E-cadherin expression and higher T stage. Previous research relating β-catenin

to prognosis had been less conclusive (Hao et al 1997, Kobayashi et al 2000, Iwamoto et al 2000, Maruyama et al 2000). Factors such as small study size, differences in antigen retrieval and staining procedures and lack of a standard evaluation system for declaring a case as positive or negative for aberrant β-catenin expression have probably been responsible for differences in results.

Lytic enzymes such as matrilysin have been seen in increased amounts on cDNA arrays (Notterman et al 2001) and matrix metalloproteinase (MMP) production is increased in CRCs (McLeod et al 1999). These enzymes degrade extracellular matrix allowing tumour invasion. The presence of MMP1 is associated with poorer prognosis.

In terms of the sort of changes one might expect to determine by CGH, analysis has been done relating LOH at various loci to outcome. 18q and 17p have already been discussed. Those patients with moderate 20q copy number increase had longer survival compared to those without the changes or those who had high increase in copy number. Loss of 18q in conjunction with 8p may be a poor prognostic sign (Gerdes et al 1995). 1p has been examined in three further studies (Gerdes et al 1995, De Angelis et al 2001, Ogunbiyi et al 1997). Two of these studies showed that loss of 1p was associated with shorter survival time. The third study found no such relationship. De Angelis et al (2001) showed a relationship between deletion of 14q and survival. Al-Mulla et al (2006) also found 14q loss to be associated with metastatic recurrence additionally finding loss of 4p significantly associated with metastatic disease and 4p and 14q loss jointly to be an independent prognostic indicator on multivariate analysis. Arribas et al (1999) also found 4p 14-16 allelic loss to be associated with poor survival. Choi et al (2002) found that high levels of LOH in cases without MSI correlated with earlier onset,

lymphatic invasion and rectal position whereas lesser levels of chromosomal arm involvement were associated with cancers of the proximal colon and stage I and II. For stage II cancers high level chromosomal loss was a significant predictor survival.

CGH has been used to examine colorectal cancers with some work relating to outcome. De Angelis et al (2001) found that losses of 1p, 4q, 8p, 14q and 18q and gains of 20q were associated with shorter survival times, as were >6 chromosomal aberrations per cancer. 1p and 8p loss were independent prognostic markers. Al-Mulla et al (2006) found that cancers associated with metastasis exhibited more complex genomic aberrations than non-metastasising cancers. Loss of 4p was an independent prognostic factor as was loss of arms 8p and 18q. Loss of chromosomes 4 and 14q bestowed poorer prognosis compared with those with only one of these losses. This study used array CGH for a small number of samples. In these two studies patients were not stratified for MSI status. This represents a flaw in their methodology since MSI + and MSI- differ markedly in terms of genomic gains and losses.

Gene expression array research has been used in relation to prognosis. Wang et al (2004) used RNA from 74 patients with Dukes B cancer. 31 patients had developed tumour relapse in less than 3 years and 43 had remained disease free for greater than 3 years. Gene expression profiling identified a 23 gene signature that predicted recurrence in Dukes B cancers. This was used to predict survival in a further 36 patients correctly predicting outcome in 28. The markers conferred a 13-fold increase in risk of relapse.

1.5 **Aims of thesis**

The primary aim of this thesis was to investigate the use of aCGH in prognosis in colorectal cancer by examining differences in aCGH results for Dukes B cancers with good and bad outcome, the null hypothesis being that there is no difference between the two in terms of these types of changes in relation to array CGH. Other test results- namely for MSI, *APC* LOH and ploidy- were also to be compared between these two groups.

A secondary aim was to identify new oncogenes related to CRC.

A tertiary aim of the thesis was to use aCGH to investigate the timing of chromosomal instabitility by using the technique to study genomic gains and losses in adenomas.

At the time of planning this research (2003) no publications existed relating array CGH or simple CGH to outcome for Dukes B cancers. It was unknown whether copy number change would be linked to outcome and no alternative technique existed at this point for assessing copy number change over the whole genome in a single experiment at high resolution(1 Mb), although results from LOH experiments performed on series of different stage cancers and those comparing good and bad outcome cancers, suggested that there might be differences between outcome groups in terms of these factors.

Further by analysing small gains of genomic material revealed by array CGH and comparing such areas with the human genome map it was hoped that it would be possible to determine candidate oncogenes of as yet uncertain importance in colorectal cancer and further investigate whether such genes exhibited unusual expression in cancer tissue with In Situ Hybridisation (ISH) analysis. At the time the research was conducted very little was known about the importance of the candidate genes investigated in terms of their importance in CRC tumorigenesis.

Additionally, to investigate the timing of genomic instability we planned to analyse sporadic adenomas with aCGH. As has been discussed, the point at which CIN develops in tumorigenesis is a subject of debate and at the time of performing this research no papers existed that had used array CGH to investigate adenomas.

Thus, in terms of overview the work took the following course. Cancer samples collected at the John Radcliffe Hospital, the Leeds General Infirmary and St Mark's Hospital were used for the work. We needed to gather samples from three centres in order to maximise the number of Dukes B cancer DNA samples with associated long term follow up data. Although a power calculation had suggested a need for 45 samples, the power calculation was made on the basis of LOH data and we thought it advisable to maximise the amount of samples analysed . Dukes B cancer tissue from patients with long term follow up data is hard to get and multiple centres were needed to increase sample numbers.79 Dukes B cancers were divided along survival lines into those from patients with 5 year survival and those from patients who died within 5 years of surgery. For these cancers, clinicopathological data including age, gender, tumour location, lymphovascular invasion, T stage and pathological grade were determined. These were related to outcome. All cancers were tested for MSI status and *APC* LOH. These data were compared with outcome. The Oxford cancers were also tested for aneuploidy/polyploidy using flow cytometry. aCGH experiments were conducted on all Dukes B cancers and results given for the whole group and the different survival groups. Differences in genomic gains and losses between the survival groups were determined.

Four candidate genes were then selected on the basis of these and other aCGH results and ISH experiments undertaken on a tissue microarray (composed of a different set of cancers) and the results compared with known outcome for the tissue array samples. The tissue array was used as it allowed ISH experimentation on a large

number of cancers in a single experiment. For clarity I will restate that all cancers comprising the tissue array were different to those used in the main body of research.

Lastly aCGH was performed on adenomas and gains and losses for that set of tissue recorded.

In terms of summary by chapter:

Chapter 1 gives the background to the subject of colorectal cancer genetics, prognosis in Dukes B cancer and aCGH.

Chapter 2 outlines methods and materials used in the research.

Chapter 3 outlines sample collection and processing from the three centres and relates simple patient and tumour characteristics to outcome.

Chapter 4 describes the experiments allowing the classification of the samples in terms of their gross genomic stability.

Chapter 5 describes the aCGH findings and comparison of these between different outcome and instability groups.

Chapter 6 describes the selection on the basis of aCGH results of candidate genes and ISH work on these candidates. The chapter also describes the use of a tissue array comprised of a separate set of cancers to investigate possible links between candidate gene expression and outcome in this separate cancer group.

Chapter 7 describes aCGH experiments on a group of adenomas to investigate timing of genomic instability.

Chapter 2.

Methods and materials.

2.1 **DNA extraction from fresh-frozen tissue.**

DNA was extracted from fresh-frozen tissue using the DNEasy Kit (Qiagen) according to the manufacturer's instructions. This method permits isolation of 10-30µg of DNA from 25mg of tissue. Briefly, the tissue was cut into small pieces, lysed with proteinase K, treated with RNAase to degrade RNAs, precipitated with ethanol and added to a spin column to bind the DNA. After a series of washes to remove protein debris, the DNA was eluted from the column with dH_2O.

2.2 **Spectrophotometry**

To assess DNA quantity and quality, aliquots were diluted in dH_20 and analysed by spectrophotometry at 260nm and 280nm using a SPECTRAmax PLUS spectrophotometer. As an optical density (OD) of 1 at 260nm corresponds to about 50µg/ml double-stranded DNA, sample concentrations in µg/ml were calculated as 'dilution factor x 50 x OD 260. OD 260/OD 280 ratios were calculated to assess the DNA quality. Pure DNA samples have an OD 260/OD 280 ratio of 1.8. RNA contaminationl increased ratio towards 2.0 while protein conyamination decreased the ratio towards 1.4.

2.3 **Polymerase chain reaction**

The polymerase chain reaction (PCR) is an in vitro technique for DNA synthesis, which allows selective amplification of a specific DNA sequence. The reaction requires two oligonucleotide primers, which hybridise to opposite strands and flank the target region in the template DNA. In addition, the reaction requires the four

deoxyribonucleoside triphosphates dATP, dTTP, dGTP and dCTP, and a thermostable DNA polymerase. A PCR reaction typically consists of 35 cycles, and each cycle has three steps: (i) Denaturation of the DNA template, (ii) annealing of the primers to the separated template strands, and (iii) extension of the primers by DNA polymerase. The first PCR cycle creates two new strands of variable length per DNA template, which can act as targets in the second cycle. The third cycle produces double-stranded DNA molecules which comprise precisely the target region defined by the two primers. The following cycles result in an exponential doubling of this target fragment which soon becomes the predominant reaction product. PCR amplification is not infinite, and the desired fragment gradually stops to accumulate. Several factors determine the point at which the reaction plateau is reached including the utilisation of substrates (dNTPs or primers), the stability of DNA polymerases, and the competition for reactants by non-specific products or primer-dimers.

PCR primers were obtained from public sources using the Primer3 programme of the Whitehead Institute of Biomedical Research (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_ www.cgi). A typical 25µl PCR reaction contained 30-50ng of template DNA, 0.4pmoles of each primer, 1x Mg2+-free PCR buffer (Promega), 2mM MgCl2 (Promega), 0.2mM of each dNTP (Amersham) and 1 unit of PIC Taq DNA polymerase (Cancer Research UK). DNA samples were aliquoted into a 96-well plate (ABgene), and a PCR master mix was made up with the remaining reaction components. The master mix was vortexed and added to the samples. The plate was sealed with Thermowell Sealers (Corning) and immediately run on a Tetrad PCR Machine (MJ Research). A standard PCR reaction consisted of an initial denaturation at 94°C for 5min, 35 cycles of 94°C for 1min, 55°C for 1min, and 72°C for 1min, and a final extension at 72°C for 10min. Depending on the primer pair and target

size, variations of these conditions (Mg2+ concentration, additives, annealing temperature, extension time) were necessary. Products were run on the ABI377 sequencer, and results were analyzed using Genotyper software

2.4 **Loss of Heterozygosity analysis**

For LOH analysis, matched DNA samples from normal colonic mucosa and tumour tissue were genotyped at a set of microsatellite markers close to the tumour suppressor gene under investigation. Microsatellite markers are short tandem repeats of mono-, di-, tri- or tetra-nucleotides which are frequently polymorphic in the population. The forward primer of each PCR reaction was labelled with a fluorochrome (FAM, HEX or TET). The PCR products were run on a semi-automated ABI Prism 377XL DNA sequencer (PE-Applied Biosystmes) together with GENESCAN-350 TAMRA size marker (PE-Applied Biosystems). The results were analysed for allelic loss using GENESCAN and GENOTYPER software (PE-Applied Biosystems). For heterozygous samples, the area under each allele peak was determined, and the ratio of the areas of the two alleles was calculated. The LOH value was calculated. The LOH value was calculated by dividing the allelic ratio of the normal tissue by the allelic ratio of the matched tumur tissue. LOH values ≤ 0.5 or ≥ 2.0 were scored as allelic loss, assuming that the tumour contained more than 50% neoplastic material as confirmed by histopathological examination.

2.5 Flow Cytometry

Flow cytometry (FACS) techniques offer a means of assessing the amount of DNA contained within the cell nucleus. It can be applied to any tissue that can be reduced to a single-cell suspension. After disaggregation, the cells are treated with RNase (to degrade RNAs)

and stained with a flouresecent dye that binds stoichometrically to DNA. As individual cells are passed through a laser beam, dye-emission can be measured which is proportional to the DNA content. Typically, 10,000 or more cells can be analysed in minutes, and the results can be expressed as a frequency curve of DNA content. From this data estimates of the number of cells with 2C (C=haploid genome), intermediate and 4C DNA content can be obtained, corresponding to G0/G1, S and G2+M phases of the cell cycle. When comparing normal and tumour samples, aneuploidy or polyploidy can be identified as extra peaks, provided that more than 5% of chromosomal material have been gained or lost (approximately 1 large or two medium/small sized chromosomes). The ratio of medians of G0/G1 peaks of tumour and normal sample is referred to as DNA index (DI).

FACS was performed on formalin fixed, paraffin-embedded tissue which was prepared by cutting a 50mm section from each tissue block, placing it into a histopathology cassette between two sheets of 3MM filter paper, and dewaxing it in 100% xylene (BDH) overnight. The section was rehydrated through an ethanol series (100%, 95%, 90%, 70%, 50%) and rinsed twice in water. The tissue was removed from the slide with a scalpel and digested with (0.5% pepsin (Sigma), 0.9% NaCl, pH 1.5) for 30 min at 37°C. After centrifugation at 1500 r.p.m for 5 min at room temperature cells were washed twice in PBS before flow cytometric analysis and detection of near-diploid and polyploidy cell populations. Following washing with PBS, the cells were treated with 40 µg/ml propidium iodide (PI) and 10 µg/ml RNase for 30 min at room temperature then analysed on a FACSCalibur (Becton Dickinson).
Cells were excited by the argon laser emitting 488nm. PI fluorescence was detected using a 670nm long pass filter. Forward and right angle light scatter were used to set gate including all cells, but excluding debris. A second gate set on area and width of PI

fluorescence were collected in linear mode, and acquisition was stopped after 8000 gated events had been acquired. Data were analysed for aneuploidy (DNA index or D1) using dedicated Modfit software (Verity Software House). The DNA diploid peak was set using the normal samples.

2.6 Array Comparative Genomic Hybridisation

The genesis and function of CGH has been described in the introduction.The array construction and hybridisation were performed as described by Fiegler et al (2003) and Douglas et al (2004). In brief, DNA samples were isolated from 3452 large insert genomic clones (BACs, PACs and cosmids), at an average spacing of about 1Mb throughout the genome, and were amplified using DOP-PCR. The DOP-PCR products were subjected to a further amino-linking PCR and the products were arrayed onto amino-binding glass slides in duplicate. Six Drosophila clones were arrayed in duplicate alongside the human clones.

Test and control DNAs were differentially labelled using Cy-dye modified dCTPs (Cy3 and Cy5) (NEN Life Science Products, Boston, MA) and a Bioprime Labelling Kit (Invitrogen) with a modified dNTP mix. For array CGH, 450ng of test and reference genomic DNAs were labelled. The DNA was added to H_2O to give a final volume of 66µl, 60µl of 2.5x random priming solution was added and the DNA denatured at 100°C for 10 minutes. 15µl of 10x dNTP mix (1mM dCTP, 2mM dATP, 2mM dGTP, 2mM dTTP in TE buffer), 1.5µl of 1M Cy3-dCTP or Cy5-dCTP and 3µl Klenow fragment were added on ice and the reaction incubated overnight at 37°C. The reaction was terminated by the addition of 15µl stop buffer and unincorporated nucleotides were removed using microspin G50 columns (Pharmacia). Three columns per sample were prepared for use by vortexing briefly to resuspend the resin, loosening the cap, snapping off the bottom closure and centrifuging at 735g for 1 minute whilst

placed in a 1.5ml Eppendorf tube. The column was placed in a fresh 1.5ml Eppendorf tube and 55µl of the reaction mix was loaded onto the resin bed and centrifuged at 735g for two minutes. Spin through from the 3 columns was pooled and 5µl run on a 2.5% agarose gel.

For the 1Mb array (6cm^2), 180µl of Cy3 labeled DNA, 180µl Cy5 labeled DNA, 135µl human Cot1 DNA (Roche), 54µl of 3M sodium acetate pH 5.2 and 1ml 100% ethanol were precipitated together for 30 minutes at -70°C. In a separate tube, 80µl Herring sperm DNA (10mg/ml Sigma), 135µl Cot1 DNA (Roche), 24µl 3M sodium acetate pH 5.2 and 1ml 100% ethanol were precipitated together for 30 minutes at -70°C. Both tubes were centrifuged at 7700g for 15 minutes and the supernatant removed. Pellets were washed with 500µl of 80% ethanol, centrifuged at 7700g for 5 minutes and the supernatant removed.

The labeled DNA/Cot1 mix was resuspended in 60µl of hybridisation buffer , (50% formamide (Fluka), 10% dextran sulphate (Amersham), 0.1% Tween 20 (BDH), 2xSSC, [10mM Tris-Cl pH 7.4]), which had been prewarmed to 70°C, along with 6µl of yeast tRNA (100µg/ul, Invitrogen), to create a hybridisation solution. The Herring sperm/Cot1 mix was resuspended in 160µl of hybridisation buffer to create a prehybridisation solution. Both solutions were denatured at 70°C for 10 minutes.

Two layers of rubber cement were placed around the area covered by the array on the slide to create a hybridisation well. The prehybridisation solution was applied within the well and the slide was agitated to ensure that the entire array was covered with prehybridisation solution. The slide was placed in a humidity chamber, which contained 3MM paper (Whatman) saturated with 2xSSC/40% formamide and incubated for 60 minutes at 37°C whilst rocking (5rpm). The hybridisation solution was incubated at 37°C in the dark for 60 minutes. As much prehybridisation solution as

possible was removed from the hybridisation well using a p10 pipette and the hybridisation solution was applied to the array. The slide was placed in a second hybridisation chamber, which contained 3MM paper (Whatman) saturated with 2xSSC/20% formamide. The chamber was sealed with parafilm and incubated for 48 hours at 37°C whilst rocking (5rpm).

The slide was removed from the hybridisation chamber and the rubber cement removed with forceps. The slide was briefly rinsed in PBS/0.05% Tween 20 (BDH) and then washed in PBS/0.05% Tween 20 (BDH) on a rocking platform (70rpm) for 10 minutes at room temperature. The slide was incubated whilst rocking in 50% formamide/2xSSC at 42°C for 30 minutes, followed by a second wash in fresh PBS/0.05% Tween 20 (BDH) on a rocking platform (70rpm) for 10 minutes at room temperature. The slide was dried by centrifugation at 150g for 5 minutes and stored in the dark until scanning

Slides were scanned on an Axon 4000B scanner (Axon Instruments) or an Agilent G2565BA scanner (Agilent Technologies). The photon multiplier tube (PMT) levels used for detection were manually (Axon) or automatically (Agilent) adjusted and tailored for each array. Images were quantified using Genepix 4.0 software (Axon instruments) or the “Spot” analysis program.

Sanger arrays were analysed using an excel spreadsheet, proceeding through the following steps. To control for non-specific hybridisation signals, the raw Cy3 and Cy5 intensity signals produced by each spot on the array were compared to the average Cy3 and Cy5 intensity signals produced by the *Drosophila* spots. Only spots with greater than twice the average *Drosophila* spot intensity in at least one fluorescence channel were accepted for analysis. A raw ratio of intensities for each accepted spot was calculated by dividing the test intensity by the reference intensity. To adjust for any

imbalances in the labelling efficiency or scanning of the fluorochromes, raw ratios were normalised by dividing by the median raw ratio for all autosomal spots. The median normalised ratio for each pair of duplicate spots was calculated and compared to the individual ratios reported by each duplicate spot. If the individual ratios deviated by greater than 10% from the median, the spots were rejected from further analysis. The final linear ratio for each clone was taken as the mean of the normalised ratios of the duplicates spots. This ratio was then converted into a $\log^2$ ratio and plotted against the mapped position of the clone in the genome. A previously constructed excel macro programme provided different presentations of the data including plots of the raw data and normalised data both for the whole genome and for each chromosome.

For Wellcome arrays processing of the glycerols, extracting DNA and DOP-PCR was conducted in the same way as described by the Sanger group (http://www.sanger.ac.uk/HGP/methods/cytogenetics/DOPPCR.shtml). 3452 Sanger clones were used. Processed DNA was spotted onto a Corning GAPS 11 Amino-silane coated slide with 3 replicates per slide using a Lucidea Array Spotter (GE Healthcare). In terms of hybridisation, 1 microgram of cancer DNA and 1 microgram of pooled normal female DNA, were used for each hybridisation. These were labelled as for the Sanger arrays described above.

The hybridisation of the Wellcome arrays differed from the method used for the Sanger arrays and was as follows. Labelled target DNA was confirmed as being in base of an Eppendorf tube and then boiled at >95°C for 5 minutes with foil over the top of the tube. The Eppendorf tube was transferred to a 37°C hot block, covered with foil and left for 55 minutes. The required number of cover slips was prepared one at a time: cover slips were removed from 100% EtOH using forceps and placed in clean 50ml tubes (Falcon) and the tubes placed in a microwave oven for 55 seconds. A hotplate was

preheated to 37°C. When there were 15 minutes of DNA heating left the CMT hybridisation chamber (Corning 231) was prepared. This chamber was washed with water and dust particles removed. The chamber and black holding clamps were lightly greased with Vaseline. When DNA incubation had 5 minutes to go, the wells at each end of chamber were filled with 3.4X SSC. It was also necessary to add drops of 3.4X SSC at each end of the chamber next to the wells. Then the slides and cover slips were prewarmed by resting them on a 37°C hot block. Enough space should be left between cover slips for slides to be lowered on to them. The labelled target DNA was then removed from the bottom of the Eppendorf tube and pipetted either side of the central area of the clean cover slip placed at the edge of the hot block. The array slide was carefully lowered onto the cover slip so that the arrayed DNAs were in contact with the cover slip. The slide was placed in the hybridisation chamber and the lid put on the hybridisation chamber which was sealed at the sides with clamps. The chamber was placed in a 65°C water bath for 48 hours. Washing was performed as for the Sanger arrays.

Wellcome arrays were scanned using an Axon GenePix 4000B scanner and data analysed using GenePix Pro 4.1and Microsoft Excel. Using the GenePix Pro proceeded as follows. Scanned images were opened. The settings for a grid that matched the layout of the slide were chosen. A preliminary visual examination of the slide was made and features (spots) flagged that were poor or absent. Analysis was then performed and the results file saved. This output file was then saved as an Excel file and data sorted by Chromosome and ID. To normalise the data a median of the values “Ratio of Means (635/532)” for all autosomal clones was calculated. First “Ratio of Means (635/532)” values for flagged clones were removed from the median calculation data set. The value of “Ratio of Means (635/532)” for all clones was divided by the calculated median to get a normalised ratio for each clone. The mean and standard

deviation were calculated, using the normalised ratios, for each clone (mean and SD calculated for the 3 replicates of that clone). If the calculated SD was above 0.2 for the 3 replicates, the most outlying value of the Normalised ratios was removed. If more than 1 replicate was excluded the clone was excluded from further analysis. A graph to visualise the result of the hybridisation was then drawn: Average Normalised ratios vs. Chromosome (Genomic position).

Deciding whether a clonal gain or loss is statistically significant has been outlined previously by Roylance et al (2006). In this method a confidence interval for the variability in fluorescence intensity ratios was produced by performing normal versus normal hybridisations. It then becomes possible to identify clones differing significantly from a ratio of 1. While the method described by Douglas et al (2004) is applicable to cell lines derived DNA, the differing proportions of normal tissue contaminating cancer tissue mean that no single threshold for gain and loss can be applied generally across all experiments. For experiments involving tissue derived DNA threshold values for each hybridisation need to be calculated. To do this a modal area of a minimum of 50 Mb was taken and compared to the area showing the smallest gain of probable significance within the same hybridisation. The baseline area was defined as the longest contiguous region in which 95% of the clones fell within the 99% confidence intervals calculated from the linear ratios of the autosomal clones in the normal versus normal hybridisations produced by Douglas et al (2004). A Student's t test comparing the log2 ratios for clone in the two areas could then be performed and if the presumed area of change was shown to be significantly different this difference could be taken as the threshold limit of change for that hybridisation and all larger changes could be taken to be significant. The significance level for the difference between the tumour and pooled female normal log 2 ratios was set at $p<0.0001$. Copy number changes for sex chromosomes were not assessed. In practice a

minimum of 3 clones was accepted as showing a meaningful area of gain or loss and gains and losses composed of one or two clones were not considered of significance unless they were supported by similar changes in other experiments. Areas of gain or loss that were obviously markedly larger that the threshold level set to represent likely single copy change were noted but no arbitrary threshold level for such change was set. A BAC clone having been defined as showing no significant change, gain or loss it was assigned a value of 0, 1 or -1. Using this scoring it was then possible to compare gains and losses at each BAC location across all the cancers. Results were produced to compare whole chromosomal gain and loss, and chromosomal arm changes for the survival groups.

2.7 In Situ Hybridisation

This procedure assesses the levels of RNA relating to a specific gene in tissue and so investigates gene expression. A labelled probe is hybridised against RNA in tissue sections mounted on glass slides. A single stranded RNA probe (riboprobe) complementary to the target RNA is produced and labelled with a radioisotope. The hybridised probe is visualized using autoradiographic procedures. The localization of the silver grains is often visualized using dark field microscopy.

ISH was carried out according to an isotope method previously described by Jeffrey (Jeffrey et al 2005).

Riboprobes are single-stranded RNA molecules synthesized by in vitro transcription
using a DNA-directed RNA polymerase (SP6, T3, or T7), sufficient nucleoside triphosphates (including UTP), and a suitable DNA template. Templates are lengths of double-stranded DNA bearing the sequence for RNA polymerase binding and initiation, followed by a sequence that is specific to the target RNA. A riboprobe capable of binding to an mRNA must be a strand complementary to the target

mRNA, i.e., an RNA version of the non-coding strand. Usually, templates are made from plasmid DNA that has been linearized with a restriction endonuclease. Large amounts of good-quality template can be produced from a cloned cDNA fragment in a plasmid vector. Probes were prepared according to the following method. To a microfuge tube, 50 µg plasmid DNA (approx 1 mg/mL solution in TE or water), 35 µL of 10X restriction endonuclease buffer appropriate to the chosen restriction enzyme, and nuclease-free water to 340 µL were added. These were mixed, and then 3.5 µL removed as a "pre-digest" sample to be kept at –20°C for later analysis. 10 µL of chosen restriction enzyme (10 U/µL) was then added, mixed, and incubated at the temperature specified by the enzyme supplier for 3–4 h. Ingredients were mixed, and then 3.5 µL removed as a "post-digest" sample to be kept at –20°C for later analysis. 350 µL of phenol/chloroform were added and the mixture shaken gently to emulsify for 2 min, then centrifuged at 13,000*g* for 5 min in a bench-top microcentrifuge. The upper (aqueous) phase was then transferred to a fresh microfuge tube, taking care to avoid the interphase. The last two steps were then repeated. 300 µL of chloroform/isoamyl alcohol were added, the mixture shaken gently to emulsify for 2 min, and then centrifuged at 13000g for 5 minutes. The upper (aqueous) phase was transferred to a fresh microfuge tube, taking care to avoid the interphase. 100 µL of 7.5 *M* ammonium acetate were added and then 750 µL of ethanol added. The mixture was allowed to stand for approximately 30 minutes at room temperature. The DNA was pelleted at 13,000g for 15 minutes. The supernatant was discarded, and then 1 mL of 70% ethanol added. The DNA was pelleted at 13,000g for 15 minutes. All of the supernatant was discarded by careful pipeting, and the pellet allowed to air-dry for about 15 minutes.

25 µL of 10 m*M* Tris-HCl and 0.1 m*M* EDTA (pH 7.6) were added to dissolve the

template DNA (37°C, 10 min).1 µL was removed as a “final template sample.” The stock template solution can be stored at –20°C for several years. The efficiency of plasmid cleavage production was assessed by conventional 1% agarose ethidium bromide gel electrophoresis of the three samples taken. The concentration of the final template solution was measured spectrophotometrically, or estimated by comparing the fluorescence of its band with those in the pre- and post digestion samples. If necessary, the concentration of the final template solution may be adjusted to approximately 1 mg/mL.

To label the probes, the following were added in order to a microfuge tube at room temperature: 2.5 µL of 5X transcription buffer, 1.0 µL of RNase inhibitor, 0.7 µL of DTT (100 m*M*), 2.0 µL of AGC mix, 2.4 µL of stock template solution/nuclease-free water (to give 1 µg template DNA), 3.5 µL of 35S UTP. g. 0.4 µL (6–8 U) of appropriate RNA polymerase (Promega). The mixture was incubated at 37–40°C for 60 minutes following this the template was destroyed by adding DNase I (1 U, 1 µL) and incubating at 37°C for 15 min. During this time a Chromaspin-30 column was prepared by centrifugation. DTT (10 m*M*, 25 µL) was added as well as carrier Ribosomal RNA (10 mg/mL, 1.5 µL) to the reaction mixture tube, and mixed well. A 1-µL sample was removed and diluted with 50 µL water, and 3 mL scintillant added to estimate the initial 35S present. The bulk of the reaction mixture was added to the top of the gel in a Chromaspin-30 column and centrifuged at 700g for 3 minutes at 15°C, the eluate was collected in a new tube containing DTT (100 m*M*, 4 µL) and RNase inhibitor (2 µL). This was mixed well and 1 µL removed for dilution with 50 µL water and 3 mL scintillant to estimate the total 35S incorporated. Riboprobe quality was assessed by standard 6% polyacrylamide denaturing gel electrophoresis (cat. no. EC 6865580X; Invitrogen, UK) of duplicate aliquots of 106 cpm riboprobes (and standard riboprobes) for 55 minutes, dry gel, and

exposed to X-ray film for 15 minutes. Meanwhile, riboprobes were stored at –20°C until use within a few days. Columns and vials were discarded as appropriate for 35S waste.

Wearing gloves, 4 µm sections were cut, using a microtome with disposable blades. The sections were floated using a clean brush or forceps onto DEPC-treated water (fresh daily), degassed before use if necessary. Sections were collected onto treated microscope slides. Sections were dried overnight in 40°C oven in racks or card slide trays protected from dust and then dewaxed in fresh xylene for 8 minutes (two changes) and rehydrated by sequential immersion in the rehydrating alcohols plus DEPC for 5 minutes each and then rinsed in PBS+. The tissue was permeablised with proteinase K at 37°C for 10 minutes and rinsed in glycine/2X PBS for 5 minutes to block the proteinase then rinsed in PBS+ for 5 minutes and postfixed in 4% PFA in PBS for 10 minutes and rinsed in PBS+ a further three times, for 5 minutes each rinse. In a fume hood, slides were immersed in 500 mL of acetylation buffer for 10 minutes and then washed in PBS+, three times, for 5 minutes each wash then dehydrated by sequential immersion in the dehydrating alcohols plus DEPC for 5 minutes each and air-dried. The specimens were now ready for hybridization. For hybridisation riboprobe/water were added to the hybridization buffer in a microfuge tube and heated at 80°C for 1 minute to “denature” the probe then centrifuged briefly to reduce aerosols, and then chilled on ice. 20 µL were pipetted onto each section (the mix was viscous, so volumes were approximate). An RNase-free glass cover slip was gently lowered onto each slide until it made contact with the hybridization buffer, and then the buffer allowed to spread under the weight of the cover slip. Working in a fume hood, the slides were placed horizontally in, for example, plastic slide-mailing boxes or Sakura slide racks, and placed in lunch boxes humidified with blotting paper saturated with 1X salts and 50% formamide. The lunch boxes were sealed with PVC tape and incubated overnight at 55°C. For washing all solutions were

prewarmed before use. The first wash buffers were placed in water baths ready for use the following day: per 25–50 slides, 5 L of TNE buffer at 37°C, and 2 L of formamide wash buffer at 55°C. In a fume hood, the lunch box containing the slides was opened. The slides were removed, and each cover slip removed by soaking in buffer if necessary, and placed in a slide rack immersed in 500 mL of formamide wash buffer at 55°C on a rocking table, and four changes made over a total of 3–4 hours. All traces of the formamide wash buffer were removed using nine changes of 500 mL of TNE buffer at 37°C, shaking over approx 45 minutes. During the TNE washes, 1 mL of stock RNase A was thawed, and then added to 500 mL of TNE. The slides with the RNase A solution were incubated at 37°C for 1 hour in a plastic lunch box. Specific containers were kept for this step. RNase-contaminated gloves were disposed of. Slides were washed in 2X SSC for 30 minutes at 65°C with agitation, twice. Slides were then washed in 0.5X SSC for 30 minutes at 65°C with agitation. Slides were then passed through graded ethanols (30, 50, 70, 90, and 100% ethanol), all containing 0.3 *M* ammonium acetate and air-dried and covered to avoid dust. The specimens were now ready for autoradiography. In a darkroom, using a 902 filter and a 15-W bulb, a water bath was heated to 42°C. A metal plate was cooled on ice. 25 mL of water was added to a cut-down, 100-mL measuring cylinder or beaker in the 45°C bath. Ilford K5 emulsion was added until the volume was 40 mL and left for at least 10 minutes to allow the emulsion to melt. A test slide was dipped in the solution and the back wiped on a paper towel, and then placed on the cooled plate to chill the emulsion layer. The slide was held up to a safelight; there should be no bubbles (if so, the emulsion was allowed to rest for another 5 minutes) and the layer of emulsion should be smooth and even with no streakiness (if not, it was mixed again and retested). The slides were dipped one at a time into the warm emulsion until the tissue section was immersed. In one smooth movement, the slide was lifted clear of the emulsion and the end rested on the dipping

vessel for approximately 10 seconds to recover the excess. The end was wiped and the back of the slide wiped with a paper towel before placing it on the cooled plate. The slides were allowed to dry for about 2.5 hours, in total darkness if possible, until the surface of the emulsion was hard when scratched with a fingernail. The dry slides were placed in a wooden (or plastic but not metal) slide rack/holder/box and sealed into a light-proof, black plastic bag and stored in a 4°C fridge to expose the emulsion. Sets of exposed slides were developed in safelight conditions by immersion in preprepared Kodak D-19 developer at approx 18°C for 4 minutes, agitating each
minute (using glass or plastic slide carriers). The slides were then immersed in 400 mL of stop solution for 30 seconds then immersed in tap water for 30 seconds then immersed in 400 mL of sodium thiosulfate fixer, with two changes of 4 minutes each. The slides could then be exposed to light. Counterstaining of tissues was performed by immersing the racked slides in diluted Giemsa's stain for 3–4 minutes. The excess stain was washed off with tap water for 30 seconds and the slides air-dried and then mounted in DPX under glass cover slips. When the DPX had set, the back of the slides was cleaned with a hard-backed blade, then front and back with 70% ethanol and paper tissue to remove all traces of emulsion and grease. If possible, it was arranged to view the sections using a microscope that allowed rapid switching between conventional illumination and reflected light. Under conventional illumination with 100× overall magnification and moderate exposure times, only the most intense clusters of black-silver grains could be seen, unless much higher magnification is used and the field of view was limited to perhaps 100 or 200 cells. In contrast, looking at the same slide in reflected light, much lower densities of silver grains were seen easily. More importantly, the quality of background obtained can be assessed in moments. Sections that in conventional illumination appear to have no labelling, may, in fact, reveal significant patterns of

mRNA expression in reflected light. With Giemsa's counterstain, the nuclei of cells without nearby grains appear greenish-yellow or golden in reflected light, and the threshold for deciding whether a cell or cell type is labelled can be chosen while looking at thousands of cells and associated extracellular matrices. Switching between conventional illumination and reflected-light, dark-field conditions was an extremely useful way to assess whether there were any significant patterns of silver grains on the section, or if significant expression was occurring in specific regions.

2.8 **Tissue array**

A tissue array is a collection of tissue cores typically 1-3mm in size set in paraffin in a grid pattern according to the normal dimensions of a pathology slide. A typical array might include 100 such cores. The array so constructed allows staining experiments to be performed on multiple tissue samples at the same time, a key to the grid allowing the identification of the different tissue samples.

The array used in this work was constructed in the Nuffield pathology laboratory in Oxford. It was comprised of colorectal cancer tissue cores from patients with clinical follow up measurable in years. Sections were taken from the tissue array block as one would normally section paraffin embedded tissue with a microtome. The cancers comprising the tissue array formed a completely different group to those cancers used in the aCGH work.

2.9 **Statistical statement**

Simple proportions were compared using Fisher's Exact test or the Chi-squared test as appropriate. Non-parametric continuous data was compared using the Mann-Whitney U test. A power calculation as to the number of specimens required for aCGH analysis was made on the basis of previous results for 18q LOH. 70% of bad

outcome cancers could be expected to show such LOH and 35% of good outcome cancers (Houlston et al 1997). For a one sided test with a threshold p value of 0.05 and a power to detect difference of >95%, a sample size of at least 45 cancers was required.

Chapter 3.

Sample Collection and Basic Clinicopathological Data Analysis

3.1 Introduction

The initial aim was to collect Dukes B colorectal cancer tissue, extract DNA from it, gather related clinicopathological and follow up data and to relate the clinicopathological data to the follow-up data. Dukes B cases with documented 5 year survival or documented death within 5 years of surgery were selected. Patients who had died directly post-operatively, cases of known inherited cancer (stated in the notes as FAP or HNPCC) and cases of cancer related to inflammatory bowel disease were excluded.

3.2 **Statistical considerations**

A power calculation as to the number of specimens required for aCGH analysis was made on the basis of previous results for 18q LOH. 70% of bad outcome cancers could be expected to show such LOH and 35% of good outcome cancers (Houlston et al 1997). For a one sided test with a threshold p value of 0.05 and a power to detect difference of >95%, a sample size of at least 45 cancers was required.

3.3 **DNA samples**

Cancer DNA came from tissue collected at three centres: John Radcliffe Hospital, Oxford, Leeds General Infirmary and St Mark's

Hospital Harrow. Post-exclusions, 79 Dukes B cases remained, with 7 cases excluded (2 post operative deaths, 3 cases of inflammatory bowel disease related cancer and 2 cases of FAP related cancer).

Tissue from Oxford consisted of 45 Dukes B cancer samples with paired normal mucosal tissue (these were drawn from a collection of 142 cancers and 34 adenomas with paired normal tissue, previously collected in Oxford, which had been archived by me and from which (for the whole set) I had extracted DNA). 26 DNA samples from Dukes B cancers with paired normal mucosal DNA were received from Leeds and 9 samples of Dukes B tumour were received from St Mark's Hospital.

All tissue had been snap frozen in liquid nitrogen on collection and stored at -70 C. Tissue usage was covered by appropriate ethics committee approval which I obtained myself (C03.137/C03.077). The Oxford tissue was processed by me. Leeds sample DNA was extracted by local researchers. The St Mark's tissue was collected locally and processed by researchers in the CRUK Molecular and Population Genetics Laboratory who supplied data which were used by me. To our knowledge no patient included in this study received adjuvant therapy.

3.4 **Processing of tissue**

The fresh-frozen Oxford samples were cut into pieces, one piece being placed in formalin and blocked in paraffin wax and the second piece being frozen and used for DNA extraction by standard methods. Samples of extracted DNA were archived for future use along with the residual tissue. DNA was extracted (Methods 2.1) and DNA concentration and quality assessed (Methods 2.2). Pathological examination of all Oxford tissue was undertaken by Professor M. Ilyas to confirm the nature of the tissue.

Tissue from Leeds was subjected to micro-dissection and DNA extracted by standard methods.
Tissue from St Mark's tumours was treated in the same way as that from Oxford. This processing was performed in the CRUK laboratory.

3.5 **Clinicopathological data**

Follow up data and clinicopathological data was obtained from patient records in the case of the Oxford samples and the St Mark's samples according to ethics committee approval. Follow up data and clinicopathological data for the Leeds samples was kindly supplied by Professor Quirke's team.

3.6 **Clinicopathological data results**

For the 79 Dukes B patients with adequate follow up, 43 (54.4%) had at least 5 year survival and 36 (45.6%) less than 5 years survival from the time of operation. For the Leeds samples this was the only information available (i.e no exact time to death just 5 year survival +/-).
Comparison of clinicopathological data for the two survival groups can be seen in Table 2.1. There was a significant difference between the two groups for T4 stage (> 5 year 16.3% versus < 5 year 36.1% p= 0.043). The > 5 year survival group contained significantly more cancers from Leeds (44.2% versus 16.7% p=0.014) and significantly less cancers from Oxford (39.5% versus 77.8% p=0.0007).

TABLE 2.1 Clinicopathological variables in relation to 5 year survival for 79 Dukes B cancers.

Survival	**>5 year (%)**	**<5 year (%)**	**P**
Sample source			
Oxford	17(39.5)	28(77.8)	0.0007
Leeds	19(44.2)	6(16.7)	0.014
St Marks	7(16.3)	2(5.6)	0.17
Gender (female)	19(44)	16(44)	1.0
Mean Age (years)	66+/-5.2	70.6+/-6.1	0.47
Tumour site			
Right	13(30.2)	10(27.8)	1.0
Left	16(37.2)	9(25)	0.33
Rectal	14(32.5)	17(47.2)	0.25
Grade (differentiation)			
Well	7(16.2)	3(8.3)	0.33
Mod	31(72.1)	32(88.9)	0.09
Poor	5(11.6)	1(2.8)	0.21
T stage			
T4	7(16.3)	13(36.1)	0.043*
LVI	6(13.9)	3(8.3)	0.49

LVI=Lymphovascular invasion

3.7 Discussion

Although the power calculation suggested that 45 samples would be adequate for this project, we thought that analysing as many samples as possible was sensible. This was because the LOH test on which the power calculation was based was a very different technique to aCGH and while providing a guideline to numbers was unlikely to have produced a reliable figure on which to base project numbers. In practice the limit of numbers in this work was placed by scarcity of arrays for the aCGH and scarcity of DNA from colorectal cancers derived from patients for whom long term follow up data was available. In order to maximise the number of samples in the study we collected samples from 3 centres (Oxford, Leeds and St Marks).

There were differences between the centres in how the samples had been processed and the amount of follow up data available. However, it was felt that these differences were more than made up for by the increase in samples involved in the work.

The method of processing the Oxford samples introduced uncertainty as to the percentage of cancer in the analysed sample. While tumour content was at least 60% in all analysed samples according to the tissue on which histological analysis was performed, this assumes a similar proportion of tumour in the sample used for DNA extraction which may not have been the case.

Ideally cause of death, time of death and disease progression would have been available for all cases, allowing comparison of cases according to cancer specific survival, disease progression and local recurrence. However, for the Leeds samples only information relating to overall 5 year survival was available. Strenuous efforts were made to obtain more data but this proved impossible.

The 5 year survival of 54% for Dukes B cancer was poor but the cancers examined here do not represent a cohort drawn from a

consecutive experience in a single centre, simply those Dukes B cancers available for analysis with adequate additional data. It is a bonus that the cancers were split more evenly than would be expected as it gives a greater number of cases in the poor survival group. It is likely in the case of the Oxford samples that previous experiments had exhausted collected tissue in an uneven way leading, perhaps, to a disproportionate use of tissue from good outcome cases.

The Leeds cancer DNA was produced from microdissected tissue samples and might be expected to give better results for some experiments than the other samples because of less normal tissue contamination. There was a significant difference between the survival groups for the percentage of Leeds cancers (>5 years 44.2% versus <5 years 16.7% p=0.014). Previous experiments from our laboratory had, however, shown very comparable results between DNA samples with 60% and 100% cancer tissue.

In terms of the poor survival group having a higher number of T4 cancers, T stage is one of the most powerful prognostic markers (Introduction 1.4.1) and this difference is expected. However, the difference between the two groups in terms of T4 proportion represents a problem in terms of comparing molecular markers between the groups since outcome difference may be dictated more by the stage of the cancer than the biology of the cancer.

Chapter 4
Comparison of MSI, Ploidy, and *APC* LOH for the Dukes B Cancer Survival Groups.

4.1 **Introduction**

The Dukes B cancers were assessed for MSI status, ploidy and *APC* LOH. It was of interest to test the prognostic significance of MSI and ploidy. Microsatellite stable and aneuploid cancers have been shown in previous work to have worse outcome (Introduction 1.4.2).
It was also important to subdivide the survival groups by chromosomal instability status before comparison of the aCGH results was made since these groups are known to have marked aCGH difference (Douglas et al 2004, Jones et al 2005).

4.2 **MSI and LOH analysis**

For MSI and LOH analyses, the tumour DNA was PCR-amplified alongside the normal mucosal DNA from the patient using microsatellite markers (BAT 26, D5S356) -Methods 2.3 and 2.4. Description of primers and reaction conditions can be found in Appendix 1.

4.4 **MSI results**

There was no difference between the survival groups for MSI+ (> 5 year survival 18.6% versus < 5 year survival 16.7% p=1.0) (Appendix 2).

4.5 **LOH results (Appendix 2)**

There was no significant difference between the survival groups shown by D5S346 (*APC*) LOH (for informative cases, >5 years survival group 55.6% versus < 5 years survival group 32% p=0.17).

4.6 **FACS analysis**

Samples from Oxford and St Marks underwent flow cytometry analysis for ploidy. The work on the Oxford samples was performed by me (Methods 2.5) and the work on the St Marks samples performed previously at CRUK.

Figure 4.1 Example of D5S346 MSI+ (Normal mucosa above and cancer tissue below)

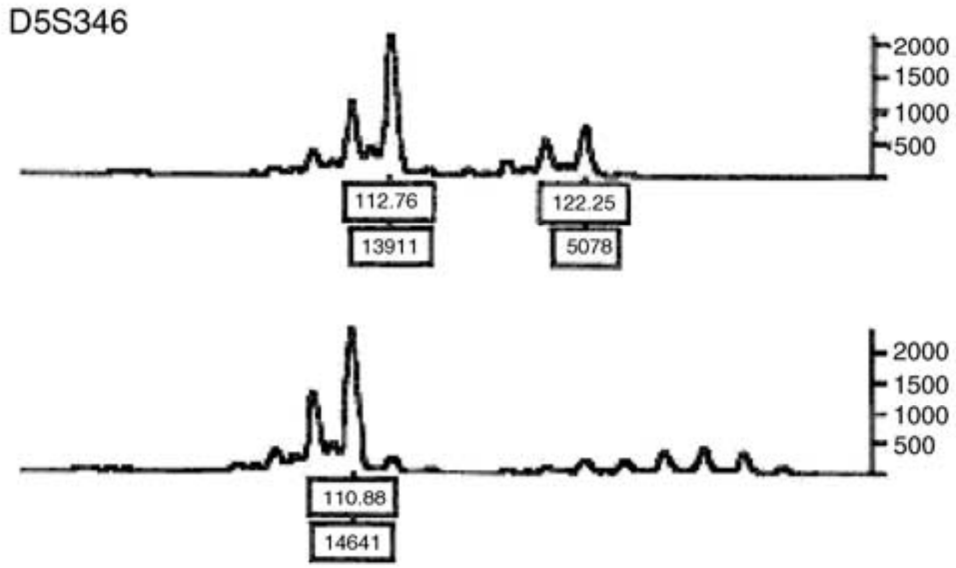


Figure 4.2 Example of D5S346 LOH. Normal sample on left and tumour on right. Ratio of peaks 80713/65407 x 13395/83657 = 0.2 (<0.5 and thus demonstrating LOH)

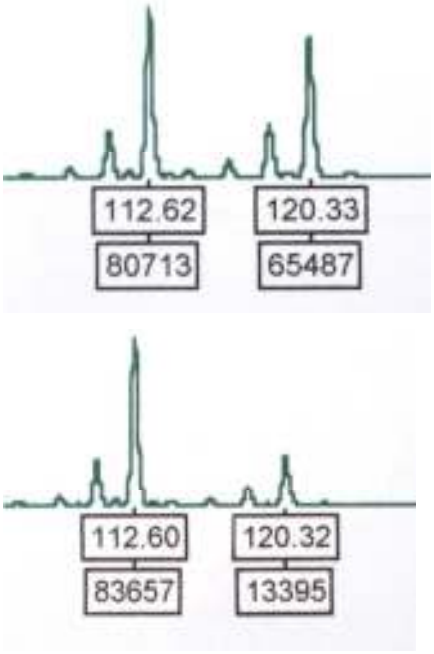
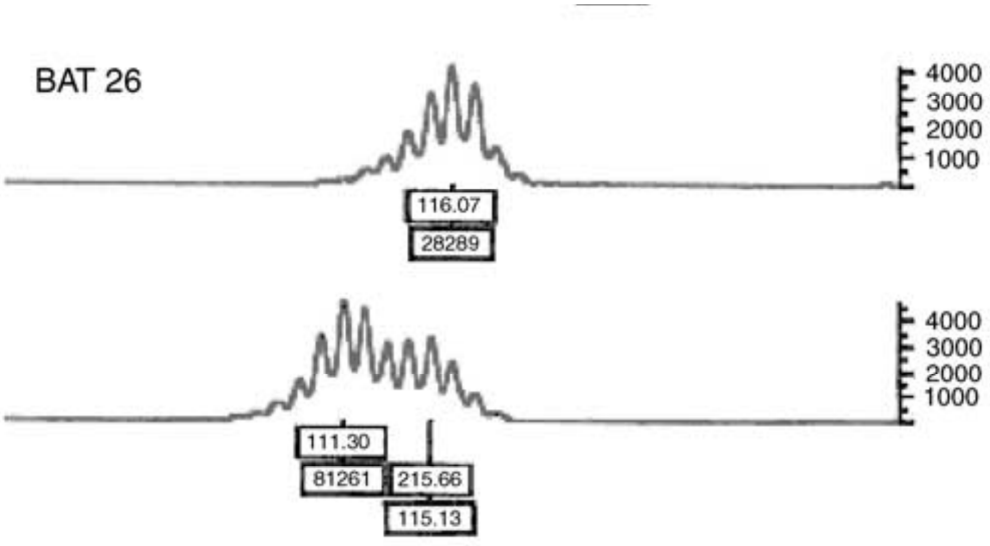


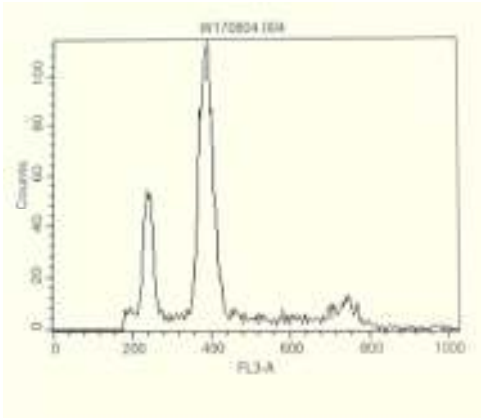
Figure 4.3 Bat 26 MSI+ (Normal tissue above and cancer tissue below). Notice additional peaks on left.

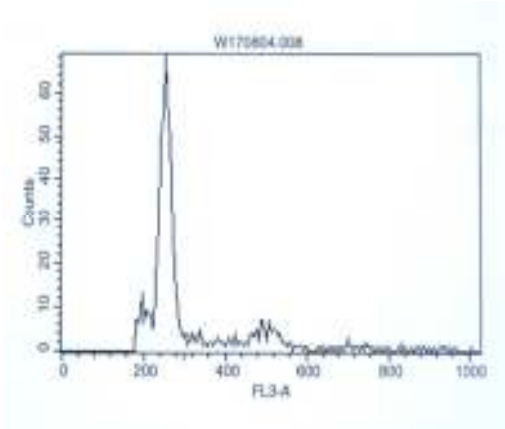


4.7 **FACS results (Appendix 3)**

50 results were available with 3 failures of the tested samples. There was no significant difference in aneuploidy between the survival groups (>5 year survival 43.5% versus < 5 year survival 55.6% p=0.57), although there were a greater number of aneuploid cancers in the poor survival group.

Figure 4.4. Examples of FACS for aneuploid cancer (left) and diploid cancer (right).





↑ 2N

↑ 2N

4.8 **Discussion**

MSI and LOH

Although no significant difference in rates of MSI was seen between the two survival groups MSI has been shown in a number of studies to be associated with improved survival (Allen et al 2005). Studies have confirmed a consistent and independent association between MSI-high phenotype and improved survival in stage II and III CRC patients. A systematic analysis of 215 sporadic colorectal cancer patients showed that those with microsatellite instability had a survival advantage over patients without it, independent of other prognostic factors. Microsatellite instability was found in 16.4% of colorectal cancers. The hazard ratio of patients with tumours showing microsatellite instability to those without was estimated to be 0.39 (95% CI 0.19 to 0.82) (Bubb et al 1996). More recently a systemic review (Popat et al 2005) showed a hazard ratio for overall survival associated with MSI of 0.65 (95% CI 0.59-0.71).

5q harbours the APC gene thus LOH at this locus is more likely to be related to tumour initiation than cancer progression and hence prognosis. Unsurprisingly then no studies have demonstrated a link between LOH at this locus and prognosis (Houlston et al 1997).

Ploidy

Although the difference was not statistically significant more aneuploidy was seen in the poor outcome group.
A recent meta-analysis (Araujo et al 2007) has studied the relationship between aneuploidy and prognosis. The main outcome measure was the five-year overall mortality rate after surgical

resection. For patients with Stage II colon cancer there was an absolute reduction of 14.3 percent in five-year overall mortality favouring diploid tumours. Additionally Walther et al (2008) produced a meta-analysis showing aneuploidy to be associated with poor prognosis (hazard ratio 1.45 (95% CI 1.35-1.55)).

Chapter 5.

Comparison of Array Comparative Genomic Hybridisation Results for the Dukes B Cancer Survival Groups

5.1 Introduction

As outlined in the introductory Chapter, aCGH provides a method of assessing genomic gains and losses in DNA samples. To examine differences in Dukes B cancers with good and bad survival, we examined 79 cancers using aCGH. The samples used were those already detailed in the previous chapters. The cancers were analysed using 2 different arrays, based on the same platform and with very similar content, one provided by the Sanger Laboratory in Cambridge and the other by the Wellcome group based in Oxford. 53 samples were examined using the Sanger array and 26 using the Wellcome array. These arrays used the same BAC clones. The values for all clones for the 79 samples compose too large a data set to be given here, but can be provided in disc form on request.

5.2 aCGH Results

To compare results, I calculated whole chromosomal changes, part chromosomal changes and changes per individual BAC. These results are provided for the whole group of 79 cancers, for the > 5

year/<5 year survival groups and for the survival/MSI subgroups. Small changes (<10Mb) and areas of very marked copy number gain are also given. Appendix 7 shows all aCGH results. In Appendix 7 large gain signifies gain of a length of chromosome >10Mb. Small gain signifies gain <10Mb. Supernormal gain represented a log2 ratio twice as large as would qualify as a simple gain.

5.3 Results for all Dukes B cancers

Considering all chromosomal scale changes (gains and losses of any part of the chromosome > 10Mb) 20.3% of all chromosomes showed change, with 9.7% showing gains and 10.7% losses.

The most commonly gained whole chromosomes were 13 (64.5%), 20 (45.6%), 7 (38%), 21 (24%), 10 (21.6%). The most commonly lost whole chromosomes were 22 (60.7%), 18 (51.9%), 14 (32.9%), 21 (27.8%), 1 (26.6%), 15 (26.6%), 19 (25.3%), 9 (22.8%).

Considering any gains or losses of p arm sections > 10Mb most common gains were for chromosomes 20 (49.4%), 7 (43%), 6 (25.3%), 10 (22.8%), 19 (21.5%) and 11 (20.2%). Most common losses were for chromosomes1 (69.6%), 18 (58.2%), 17 (62%), 8 (46.8%), 19 (26.6%), 20 (26.6%), 9 (22.8%).

Considering all q arm section changes >10 Mb most common gains were seen for chromosomes 20 (69.6%), 13 (59.5%), 8 (51.9%), 7 (40%), 21 (24%), 10 (22.8%), 11 (22.8%). Commonest losses were for chromosomes 22 (62%), 18 (54.4%), 15 (31.6%), 19 (30.4%), 14 (29.1%), 1 (27.8%), 21 (27.8%), 9 (25.3%) and 4 (20.25%).

5.5 Comparison of aCGH results for 5 year survival groups (Appendix 8)

For the >5 year survival group 20.4% of chromosomes showed chromosomal scale changes with 9.8% gains and 10.6% losses. For the <5 year survival group 20.2% (p=0.84) of chromosomes showed chromosomal scale changes with 9.4% (p=0.6) showing gains and 10.7% (p=0.83) losses.

Whole chromosome

Largest differences in terms of gain were for chromosomes 10, 16, 19 and 22, However none of these differences was statistically significant. Comparing whole chromosomal losses the largest difference were seen for chromosomes 4,6,14,19,21,22. Here the only statistically significant difference was for chromosome 6 (7/43 versus 0/36 p<0.014).

p arm section >10Mb

Largest differences were seen for chromosomes 5, 6, 12 and 16 with differences for chromosomes 5 and 16 approaching significance (11/43 versus 3/36 p=0.07 and 9/43 vs 2/36 p=0.06). Comparing losses largest differences were seen for chromosomes 4,5,6,10 and 20 with difference for chromosomes 4 and 5 reaching statistical significance (3/43 versus 9/36 p=0.03 and 0/43 versus 6/36 p=0.007).

q arm section >10Mb

Comparing gains between survival groups most difference was seen for chromosomes 6, 9 and 16. For losses most difference was seen for chromosomes 5, 19 and 21. None of these differences reached significance.

5.5 Comparison of aCGH results for MSI- survival groups (Appendix 8)

Considering the MSI- group as a whole 22.1% of chromosomes showed change, 10.6% gain and 11.5% loss. For the >5 year

survival group 22.4% of chromosomes showed change, 10.8% gain and 11.6% loss. For the <5 year survival group 21.7% (p= 0.52) of chromosomes showed change, 10.3% (p= 0.61) gain and 11.3% (p= 0.72) loss.

Whole chromosome

Most difference was seen for chromosomes10 and 16. Differences for 16 reached significance (>5 year survival 6/35 versus <5 year survival 0/30 p=0.027). Comparing whole chromosome loss most difference was seen for chromosomes 4, 6 and 10 with chromosome 6 difference reaching significance (6/35 versus 0/30 p=0.027).

p arm section >10 Mb

Largest difference was seen for chromosomes 5, 6, 10 and 16 with chromosome 16 differences reaching statistical significance (9/35 versus 1/30 p<0.016). Comparing losses most difference was seen for chromosomes 1,3,4,5,6 and 10 with differences for chromosomes 4 and 5 reaching statistical significance (3/35 versus 9/30 p<0.05 and 0/35 versus 5/30 p<0.017).

q arm section >10Mb

Largest difference was seen for chromosomes 9,13 and 16. Comparing losses chromosome 15 showed most difference. None of these differences was statistically significant.

5.6 Comparison of aCGH results for MSI+ survival groups (Appendix 8)

Considering the MSI+ group as whole 12.1% of chromosomes showed changes, 5.2% gain and 6.8% loss. For the >5 year survival group 11.5% of chromosomes showed change, 5.4% gain and 6.1% loss. For the <5 year survival group 12.6% (p= 0.57) of chromosomes showed change, 4.9% (p= 0.78) gain and 7.9% (p= 0.28) loss.

Whole chromosome

Comparing whole chromosome gain for microsatellite unstable cancers most difference was seen for chromosome 20 with no changes reaching significance. Comparing whole chromosome loss most difference was seen for chromosomes 20, 21 and 22 with no differences reaching significance.

p arm section >10 Mb

For microsatellite unstable cancers the largest difference for gains was for chromosomes 5,6 and 18. For losses largest difference was seen for chromosomes 1,6,8,18,19 and 20. None of these differences reached statistical significance.

q arm section >10Mb

Comparing gains in microsatellite unstable cancers most difference was seen for chromosomes 6,9 and 17. Comparing losses most difference was seen for chromosomes13,16,18,19,20 and 21. None of these differences reached statistical significance.

5.7 Results at the level of individual BAC clones

Appendices 9 and 10 show the most commonly gained and lost BAC clones. For all cancers the percentage of clones showing gain was 11.4%. Those cancers from patients with 5 year survival showed an average of 12.7% gain. Those without 5 years survival showed 9.9% of clones showing gain. In patients with 5 years survival and microsatellite stability 14.7% of clones showed gains. Those patients with microsatellite stable cancers without 5 year survival showed 10.8% of clones showing gain. Patients with microsatellite unstable cancers with 5 year survival showed 3.8% clonal gains as against 5% in those with microsatellite unstable cancers without 5 year survival.

For all cancers the percentage of clones showing losses was 9.8%. Those cancers associated with 5 year patient survival showed 9%

clonal losses compared with 10.7% for cancers associated with less than 5 year survival. For microsatellite stable cancers associated with 5 year patient survival 11.4% of clones showed losses compared with 11.4% for those not associated with 5 year survival. For patients with microsatellite unstable cancers associated with 5 year survival 5.6% of clones showed losses compared with 7.3% in those not associated with 5 year survival.

Looking at the largest gains for microsatellite stable cancers associated with >5 year survival these were 13, 23-27.9 Mb (77.1%), chr 20 49.3-63.4 Mb (74.4%) and chromosome 8 41.8-145.7 Mb (60%). For microsatellite stable cancers associated with < 5 year survival the highest percentage gains were for chromosomes 20 34.9-63.4 Mb (79.3%), 13 0-113.8 Mb (65.5%) and 8 90.6-128.4 and 138.9-145 Mb (62%).

For microsatellite unstable cancers associated with >5 year patient survival largest gains were seen for chromosomes 7 0-14.7 Mb and 98.5-101 Mb (50%) and 8 144.2-145 Mb (50%). For those microsatellite unstable cancers not associated with 5 year patient survival the greatest gains were seen for chromosomes 8 40.9-145.7 Mb (66.7%), 6 56.7%-75.9 Mb (50%), 20 33.8-63.4 Mb (50%).

The greatest losses for microsatellite stable cancers associated with > 5 year patient survival were chromosomes1 8.9-24.2 Mb (77.1%), 8 0-3.7 Mb (48.6%) and 17 0-20.1 Mb (68.6%). For those not associated with 5 year patient survival largest losses were seen at chromosomes1 0-23.6 Mb (70%), 8 0-30.7% (50%), 14 0-105.1 Mb (40%) and 17 0-20.1 Mb (70%).

For microsatellite unstable cancers associated with >5 year patient survival chromosome 20 showed an area (0-63.4 Mb) with 50% loss. For microsatellite cancers not associated with 5 year patient survival the greatest losses were chromosomes1 0-45 (83.3%), 8 0-34.1 Mb (50%) and 9 119.9-134 Mb (66.6%).

5.8 Small Areas of gain and loss

These areas of gain and loss (Appendix 11 and 12) of <10Mb were recorded for the outcome groups stratified by microsatellite status. It may be seen that such losses and gains are not common and are seldom repeated more than a few times in our sample. As such comparison of the outcome groups does not show obvious difference that is useful in a prognostic sense. Total numbers of such changes were highest in the poor outcome microsatellite stable group but not significantly so.

5.9 Areas of greater than single copy number change (Appendix 13)

These are areas of greater than single copy number change as defined above. Again it may be seen that for the most part such changes are uncommon and rarely repeated. However for chromosome 13 gain significant difference was found between the poor outcome group (8/36) and the good outcome group (0/43) (p=0.001).

Figure 5.1- 5.5. Examples of array CGH derived data shown as plots of log2 ratio of signal intensity (cancer compared with pooled normal DNA) against chromosomal position in Mb. Results are shown by chromosome. Examples are shown of CIN+ type change, MSI+ type change and specific small length gains and losses.

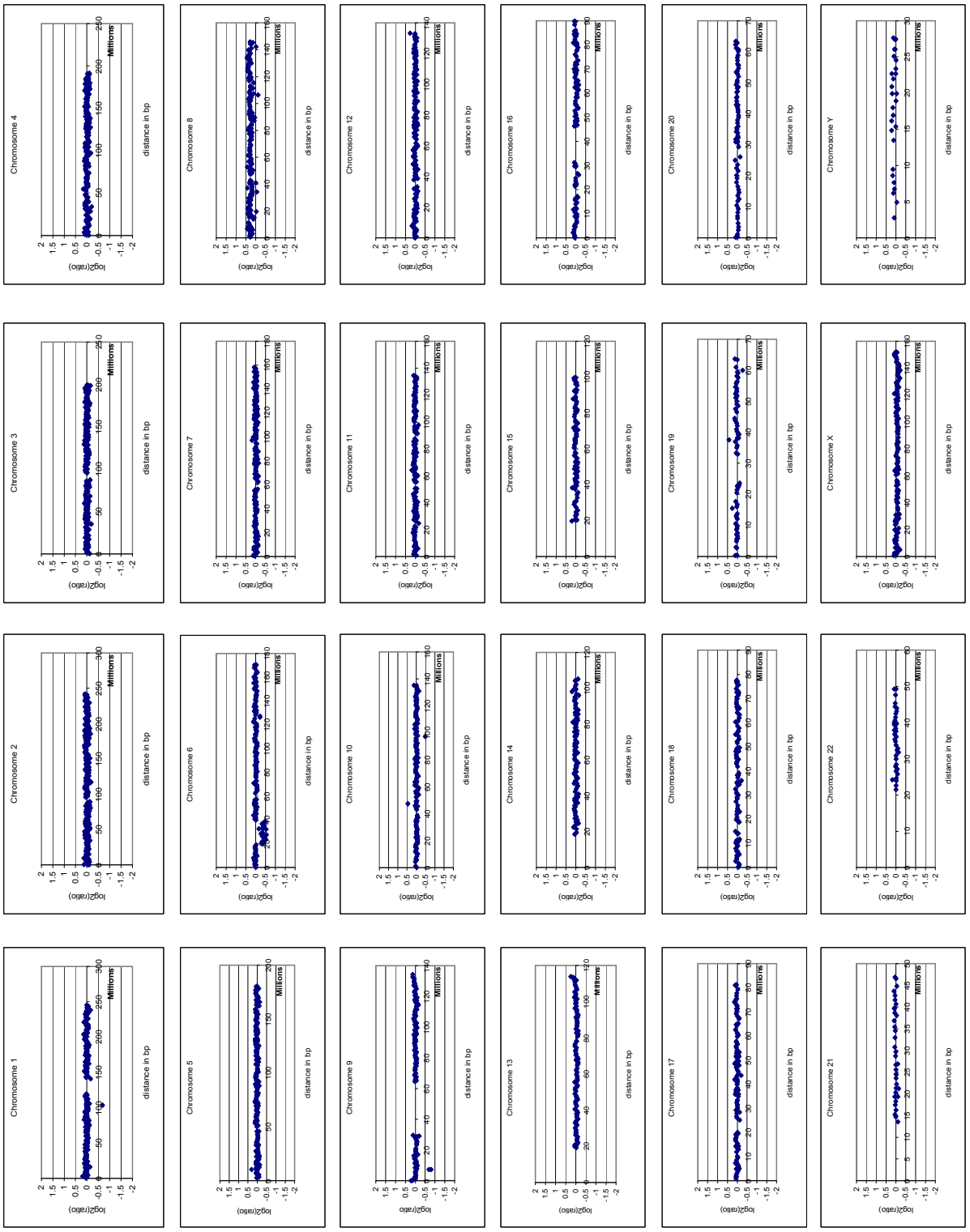


Figure 5.1 CGH result- MSI+ cancer showing minimal change. However, sub arm length change is seen for chr 6 and gain is seen for the whole of chromosome 8

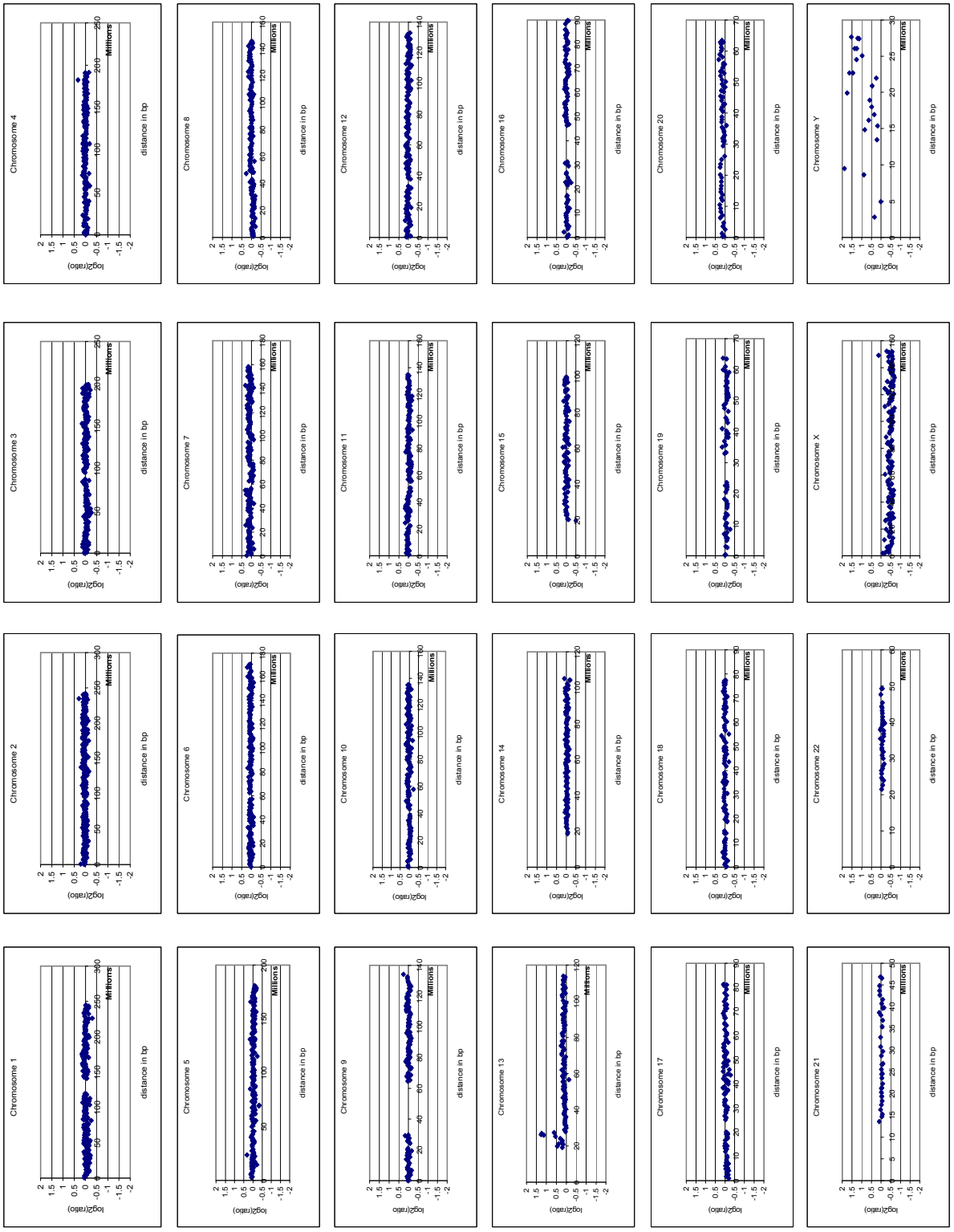


Figure 5.2. CGH result showing gain of a small section of chr 13

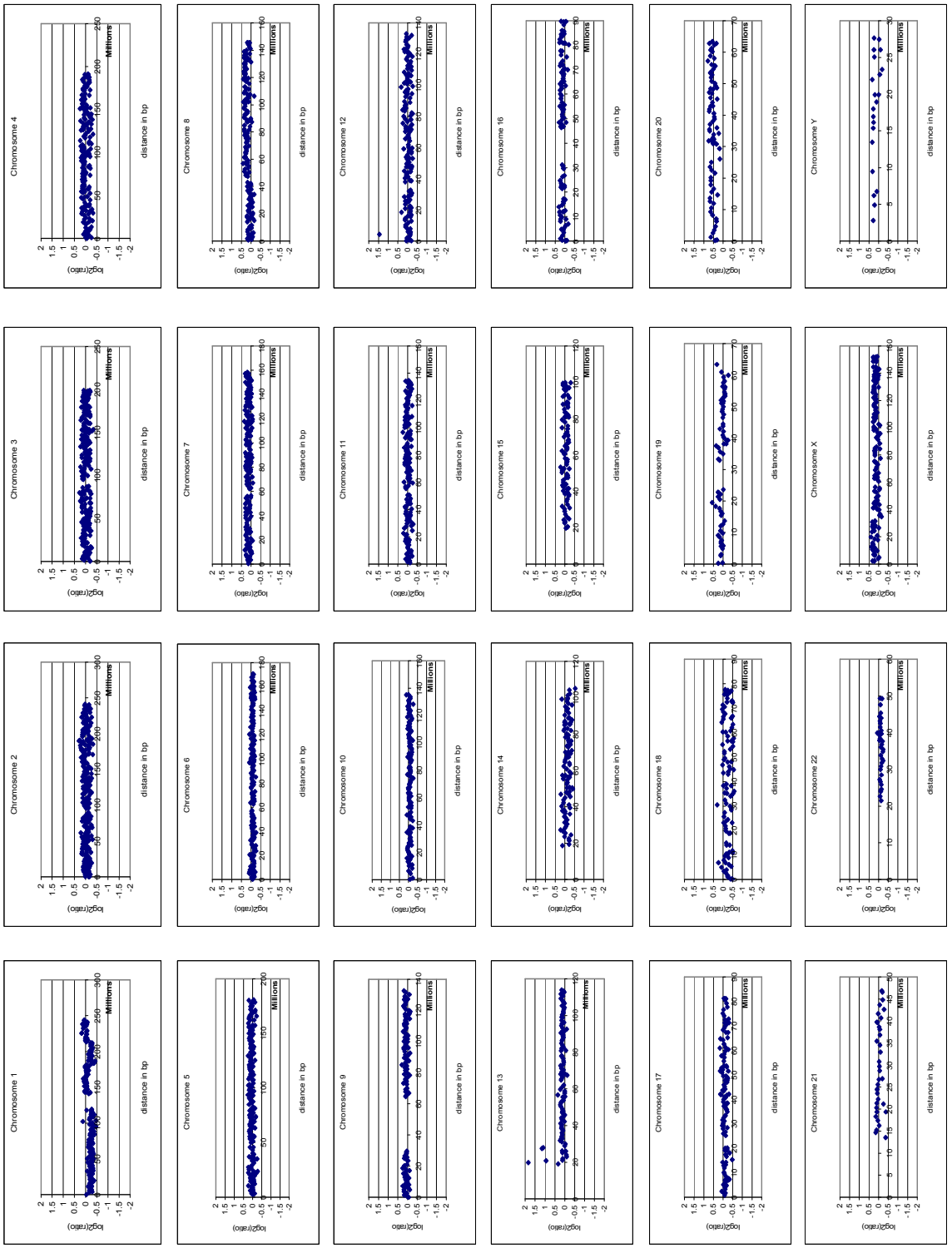


Figure 5.3 Second CGH result showing gain of same area of chr 13

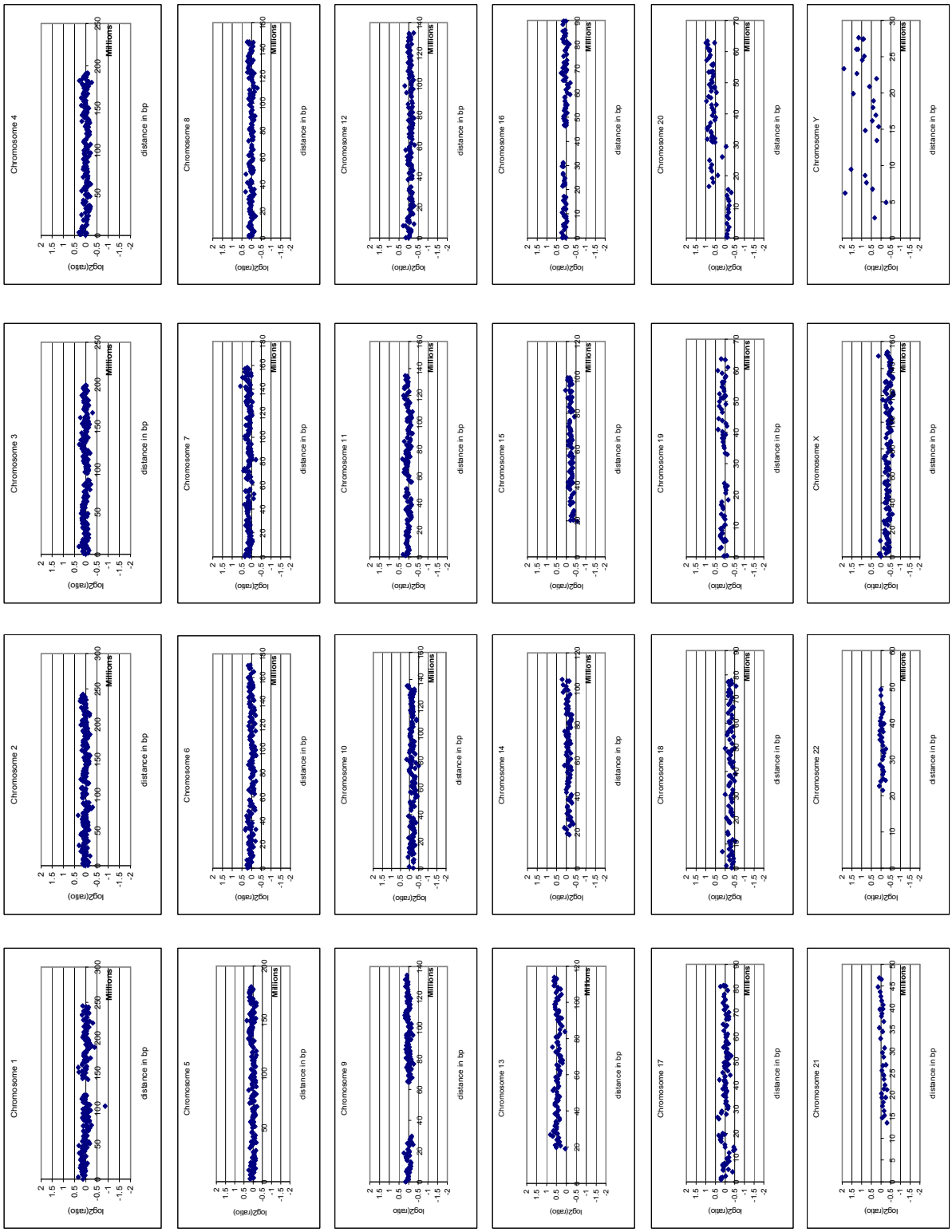


Figure 5.4 CGH result showing typical CIN+ type changes (gain chr 13, 20q(++), loss chr 18)

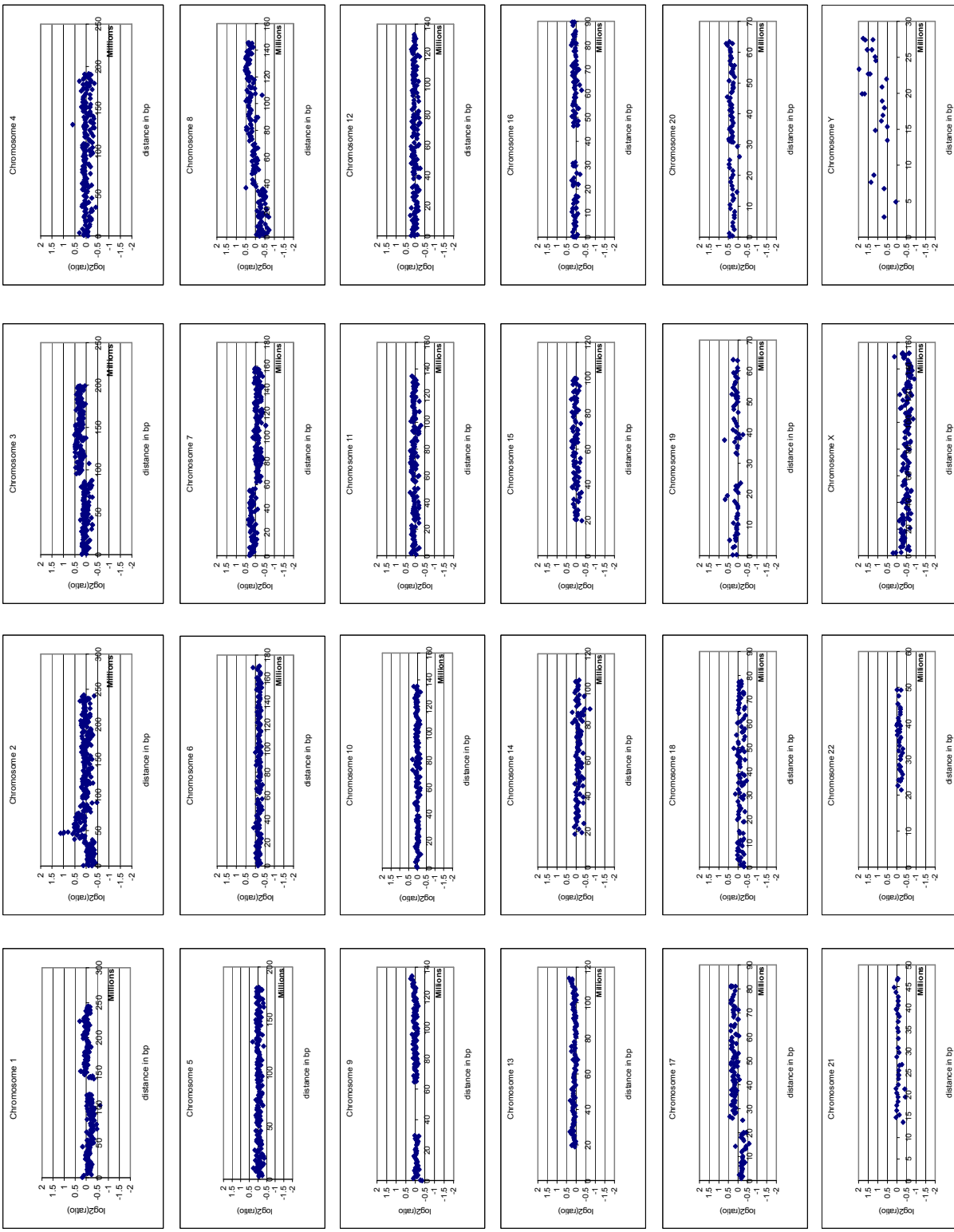


Figure 5.5. Second CGH result showing multiple genomic changes (note gain of small region of chr 2).

5.10 aCGH Discussion.

Array CGH has not been previously used in relation to prognosis for Dukes B colorectal cancer alone. We were unable to find marked difference between the outcome groups.

Comparing survival groups, chromosome 6 was more often lost in good outcome cancers, chromosome 16 was more often gained in good outcome microsatellite stable cancers.

Considering part p arm changes, cancers associated with good patient prognosis showed more gain of 5 and 16 and cancers associated with poor patient outcome showed more loss of 4 and 5. Considering microsatellite stable cancers, cancers associated with good patient outcome showed more gain of 16 and cancers associated with bad patient outcome showed more loss of 4 and 5.

Considering high amplitude areas of gain, whole chromosome 13 gain was seen significantly more often in cancers associated with bad patient outcome.

Clearly these differences may simply be the result of multiple testing. We have performed no formal Bonferroni correction, but if one takes a more stringent level for statistical significance ($p<0.01$) then only 5p loss and high amplitude gain of chromosome 13 retain significance. 5p loss was seen in only one sixth of bad outcome cancers and high amplitude gain of chromosome 13 was seen in only 8 of 36 bad outcome cancers, and therefore neither can be suggested as a useful predictor of outcome.

On a more positive note several novel areas of gain and loss were detected namely gains of chromosome 1 82-86 Mb, 3 66.7-74 Mb, 5 40.1-49.9 Mb, 6 149.5-156.9 Mb, 7 20-23.5, 13 19.1-27.9, 15 25-34

Mb, 16 84.6-89.7 Mb and losses of chromosome 7 62-67 Mb, 9 25.7-29.1, 12 74.7-82.3 Mb.

The samples from Leeds were microdissected and results from these cancers are likely to be of higher quality than those from Oxford and St Mark's. Without microdissection and using the histopathological analysis the estimated percentage content of the cancer within the specimen may not accurately reflect the percentage of cancer in the sample used to obtain DNA for analysis. In all cases the percentage of tumour is likely to be less than for the microdissected samples. The Oxford and St Mark's samples made up 40% of the microsatellite stable 5 year survival group (5+M-), 37.5% of the microsatellite unstable 5 year survival group (5+M+), 73.3% of the microsatellite stable < 5 year survival group (5-M-), and 100% of the microsatellite unstable <5 year survival group (5-M+). It may be seen that these samples are unequally spread across the 4 groups thus introducing possible bias. The Oxford samples were likely to contain a larger amount of normal tissue than other samples and this was likely to result in fewer chromosomal changes than undiluted samples.

Considering analysis of the whole group this is in line with previous analysis of colorectal cancer with array CGH. Hughes et al (2006) performed a meta-analysis of all previous CGH results for CRC. Gains of 20, 13q, 8q and 7p and loss of 18, 17p, 8p and 4p were the most frequent changes. Pooling identified less frequent, but significant changes, including gain of 1q and 3, and losses from 6q, 9p and 21q.

Considering publications showing series of CRC analysed with array CGH Nakao et al (2004) examined 125 CRC and found an overall an average of 17.3% (8.5% gain and 8.8% lost) of the genome altered . Common losses involved 8p, 17p, 18p or 18q occurring in respectively 37%, 46%, 49% and 60% of cases. Common gains were

8q and 20q gained in 42% and 65% of cases. Significant difference was seen between MSI +/- CRC with average genomic alteration being respectively 5% and 20%. MSI+ showed some copy number change most often at 8q. Douglas et al (2004) investigated copy number changes in 48 colorectal cancer (CRC) cell lines and 37 primary CRCs. The samples were divided for analysis according to the type of genomic instability that they exhibited; microsatellite instability (MSI) or chromosomal instability (CIN). Consistent copy number changes were identified, including gain of chromosomes 20, 13, and 8q and smaller regions of amplification such as chromosome 17q11.2-q12. Loss of chromosome 18q was a recurrent finding along with deletion of discrete regions such as chromosome 4q34-q35. A greater number of aberrations were detected in CIN+ than MSI+ samples as well as differences in the type and extent of change reported. For example, loss of chromosome 8p was a common event in CIN+ cell lines and cancers but was often found to be gained in MSI+ cancers. In addition, the target of amplification on chromosome 8q appeared to differ, with 8q24.21 amplified frequently in CIN+ samples but 8q24.3 amplification a common finding in MSI+ samples.

In terms of prognosis a small number authors have used array CGH to address this (although none had done so at the time of planning this research (2003)) Nakao et al (2004) did not find copy number changes to be associated with tumour stage, location, age or sex. Al-Mulla et al (2006) examined copy number change in relation to survival for 56 patients with stage I and II cancers. All were examined using conventional CGH with 10 cases being further examined with array CGH. The cancers were not stratified according to stability type but the majority were reported as being CIN+MSI-. Metastatic cancers had more complex genomic alterations than non-metastatic cancers. Loss of chromosome 4p was an independent prognostic marker for poor outcome on multivariate analysis. Loss of both 8p and 18q had a statistically significant negative effect on disease free

survival. Loss of both 4p and 14q was associated with poorer prognosis than loss of either one of these genes singly. Multivariate analysis showed loss of 4p and 5q to be independent prognostic factors.

Mehta et al (2005) used aCGH to try and predict cancer recurrence following the resection of hepatic metastases. They performed aCGH on 50 cancers and compared the results of aCGH data and a clinicopathological scoring system with outcome results. In general they found on average 30% of the genome to be altered (14% gain and 16% loss). The clinicopathological scoring system included nodal status of primary, length of disease free survival between primary cancer detection and detection of metastases, the number of hepatic tumours, size of the largest tumour and pre-operative CEA levels. Such a system had been shown if low (1-2) to predict a 60% chance of long term survival and if high to predict a 15% long term survival. The most commonly altered regions of the 50 cancers analysed were gains of 7p,7q, 8q, 13q and 20q and losses of 4p, 4q, 8p, 17p, 18p, 18q and 22q (all >50% analysed samples). Statistical analysis showed that the total fraction of the genome altered (FGA) was significantly associated with overall survival as was the clinically derived risk score. For those samples with >20% of the genome altered there was a median 38 month survival as compared with 18 months for those with <20% FGA. Differences for disease free survival did not reach significance. If the total number of chromosomal areas altered was substituted for FGA this also was significantly related to overall survival. No individual chromosomal arm or clone was an independent predictor of overall or disease free survival when analysed together with the clinical risk score or FGA. It was increased FGA that was associated with improved survival and they postulate that too great a rate of genomic instability may render a cell unviable. It may also be that MSI-CIN- cancers have a poor prognosis. Risques et al (2003) reported a subset of diploid MSI stable cancers with poor prognosis. Previous reports have generally

shown greater genomic alteration to be associated with poor outcome (Arribas et al 1997, Weber et al 2001, De Angelis et al 2001, Zhou et al 2002, Bajan et al 2002, Choi et al 2002, Diep et al 2003, Knosel et al 2003). However, other reports do not confirm this (Kochar et al 1997, Crowe et al 2001, Scott et al 1998, Flyger et al 1999, Geido et al 2002). Rooney et al (2001) also showed an increased number of chromosomal arm abnormalities to correlate with improved survival. Scott et al (1988) did not observe an increase in aneuploidy in higher stage cancers and observed a decrease in DNA index in lymph node metastases.
Diep et al (2006) performed a meta-analysis in order to try and ascertain the relation between different chromosomal changes and the point in tumorigenesis that such changes developed. Losses of 17p and 18 and gains of 8q, 13q and 20 were determined to be early changes. Loss of 4p was associated with progression of Dukes A to B/C/D. Loss of 8p and gain of 7p and 17q were associated with liver metastases. Loss of 14q and gains of 1q,11,12p and 19 were late events.

Cardoso et al (2006) in a systemic review of the available aCGH data made the following observations. The percentage of aneuploid events observed at different stages increased from benign to more malignant histological states. Common losses occurred at chromosomes 4,5,7,8,18 and 20 and gains at 7,8,13 and 20. These seemed to be found fairly early in progression but increased in their relative representation as tumours progressed from adenomas to cancers to metastatic cancers. Some changes, however, such as those related to chromosome 1 seemed to become less frequent as the cancer progressed.

In summary, ideally a larger sample size of Dukes B cancers would have been tested as after dividing into groups according to MSI

status the groups, despite our power calculation result, may be too small to show differences. There was difficulty in obtaining Dukes B DNA with the necessary length of follow up, also difficulty in procuring sufficient arrays for a larger analysis. And further, a lack of sufficient follow-up data for the Leeds samples as earlier described. Alternatively, despite the findings of differences between outcome groups for 18q LOH in previous studies, there may not be great difference between outcome groups in terms of copy number assessable by aCGH.

Chapter 6.

In Situ Hybridisation (ISH) Analysis of Candidate Genes

6.1 Introduction

One method of searching for new genes related to colorectal cancer consists of using array CGH to identify small areas of genomic gain and then using the human genome map to identify genes within those areas that might be candidates for tumour involvement. Expression of these genes can then be assessed using in situ hybridisation (ISH).

Our laboratory had analysed more than 200 colonic cancers and colonic cancer cell lines with aCGH and determined many small areas of gain and loss (<10Mb). This combined with analysis of individual BAC gain and loss allowed the identification of those small areas most often gained. Sometimes small areas of gain in a specific cancer corresponded to larger more frequent areas of gain in the set as a whole. For example we identified a small area of 13q which was gained in only 2 cancers, but gain of the whole of 13q was a common finding. Many of these areas proved to be too gene rich for easy selection of candidate genes for in situ analysis. Therefore a balance was struck in choosing regions for further analysis between the frequency of change and the gene density of the area. The four genes ultimately chosen from the long list were chosen in an arbitrary way by Professor Ian Tomlinson, all of the long list choices would have been reasonable choices for research but we only had resources to research four.

A long list of genes for possible investigation was made (Table 6.1) and four genes chosen for further analysis with ISH: *CDX2*, *FLT1*, *ARHGEF4* and *RHOA*.

Having identified candidate genes ISH experiments were performed on various tissue types to gauge whether those genes were likely to be abnormally expressed in colorectal cancer. The probes were prepared by me and the in situ work undertaken by the Histopathology Unit (Professor Richard Poulsom) at Cancer Research UK, Lincoln's Inn Field, London.

We had access to a tissue array- an array composed of 119 2mm cores of colon cancer tissue- this allowed investigation of multiple cancers' expression of a gene in a single experiment.

Table 6.1 'Long list' of possible candidate genes (actual choices underlined)

2: 131-133 especially 130.7-132.3 RAB6C, MAP3K2, CFC-1, PTPN18, **ARHGEF4**, TUBA2

2: 188-221 esp.189-191.5 COL3A1, DIRC1, PMSI, GDF8, INPPI

7: 99.3-99.7 COPS6, MCM7, TAF6, PILRB

8: 144.2-145.3 RIGE, GLI4, **RHOA**

9: 130-132.6 DDX31, CRSP8, BARHL, TSC1, RPL7A

10: 76-82 esp. 79.8-80.7 DLG5, PPIF

13: 25-27.5 esp. 26-26.2 GTF3A, POLRID, GSH1, **CDX2**, FLT3, **FLT1**

17: 35-38 esp. 35.8-37.8 AATF, TBCID3, PIP5K2B

6.1.1 *CDX2*

CDX2 is a homeobox transcription regulation gene and plays a major role in development especially in the gut and may also acts as a tumour suppressor in the adult colon. Chawengsaksophak et al (1997) made *Cdx2* null mutant mice. Homozygous *Cdx2-/-* embryos died at implantation. Heterozygotes had multiple polyp-like lesions with the highest frequency in the caecum, decreasing in incidence both proximally and distally but involving the whole of the intestinal region which expresses CDX2 during development. Lesions did not occur elsewhere in the intestinal tract. Histologically, the polyps contained normal gastric mucosa with Paneth cells interposed proximally and distally between the stomach mucosa and the surrounding colonic epithelium. The gastric mucosa was arranged in an orderly array passing from stratified squamous epithelium of forestomach-type through mucous glands of the cardia to gastric glands of the corpus and finally to mucous antral-type cells as one passes both proximally and distally. This phenotype represented an anterior homeotic shift in which intestinal epithelium had the character of a more rostral phenotype due to localised areas of *Cdx2* haplo-insufficiency. The “default” state was forestomach epithelium which in the normal animal did not express CDX2. Intercalary growth subsequently resulted in the orderly appearance of the appropriate tissue types to “fill in” the histological discontinuity between gastric and colonic epithelia.The *Cdx2* “knock out” studies detailed above showed that this gene is central to differentiation of midgut endoderm. They led to the conclusion that decreased levels of *Cdx2* expression resulted in rostralisation of gut differentiation with gastric mucosa constituting the “default” condition in which there was no expression of CDX2.

While *CDX2* mutations predisposing to sporadic CRC have not been identified (Woodford-Ritchens et al 2001, Sivagnanasundaram et al 2001) some polymorphism related to this gene has been (Rozek et al 2005). As stated, most evidence suggests a role for *CDX2* as a tumour suppressor gene, if it plays a role at all in cancer development. Expression of CDX2 does not seem to correlate to cancer stage, but does seem to correlate with tumour location (right sided), poor differentiation, high microsatellite instability status, and a positive first-degree family history (Rozek et al 2005). There is some evidence to suggest the importance of the microenviroment in relation to expression (Benahmed et al 2007). The regulation of *CDX2* by the microenvironment might be relevant during the process of metastatic dissemination when the gene is transiently turned down in invasive cells.

To understand the functional contributions of *CDX2* to colon cancer, one group disrupted *CDX2* in LOVO and SW48 human colon cancer cell lines by targeted homologous recombination. Consistent with the literature, disruption of *CDX2* enhanced anchorage-dependent cell proliferation. However, homozygous loss of *CDX2* led to significant inhibition of anchorage-independent growth in LOVO cells, and cell lethality in SW48 cells. Further analyses revealed that disruption of *CDX2* led to anchorage-independent G1 to S growth arrest and anoikis. In vivo xenograft studies confirmed that disruption of *CDX2* inhibited LOVO tumour growth. These data demonstrated that *CDX2* mediates anchorage-independent growth and survival. Thus, *CDX2* may have tumorigenic potential in the human colon cancer cell lines LOVO and SW48 (Dang et al 2006).

Witek et al (2005) showed over expression of *CDX2* by human colorectal tumours compared with matched normal mucosa. They found that >80% of colorectal tumours over expressed mRNA and protein compared with normal mucosa.

6.1.2 *FLT1*

FLT1 codes for a vascular endothelial growth factor (VEGF) receptor. Vascular endothelial growth factor (VEGF) is associated with tumour angiogenesis and poor prognosis in human colorectal cancer (CRC). VEGF receptor-1 (VEGFR-1 or FLT-1) is a high-affinity receptor for VEGF and is typically considered specific to endothelial cells. Fan et al (2005) reported the expression and function of FLT1 in CRC cell lines. FLT1 was expressed in all CRC cell lines studied as determined by RT-PCR, Western blot analysis, FACS, and ELISA. Treatment of the human CRC cell lines HT-29 and SW480 with VEGF-A (a ligand for both FLT1 and -2) or VEGF-B (a ligand specific for FLT1) led to activation of *Erk-1/2, SAPK/JNK*, and translocation of the p65 subunit of nuclear factor-kappaB into the nucleus. Both VEGF-A and -B led to significant induction of cell motility and invasiveness of CRC cells. Stimulation of cells with VEGF-A or -B also led to larger and more numerous colonies in soft agar. However, activation of *FLT1* did not increase CRC cell proliferation. In contrast to the previous paradigm that VEGFRs are not present on tumour cells of epithelial origin, they found that FLT1 was present and functional on CRC cells, and activation by VEGF family ligands could activate processes involved in tumour progression and metastasis.

Lesslie et al (2006) investigated the role of Src family kinases (SFKs) in VEGF-mediated signalling in CRC cell lines. Their results suggested that FLT1 promoted migration of tumour cells through a Src-dependent pathway linked to activation of focal adhesion components that regulate this process.

Further research (Yamaguchi et al 2007) based on sFLT1 levels in colorectal cancer without distant metastases showed patients with higher sFLT1 levels demonstrated significantly longer recurrence-free survival than patients with lower sFLT1 levels. The *FLT1* gene encodes for both the full-length receptor FLT1 (VEGFR-1) and a

soluble form designated sFLT1. SFLT1 carries the VEGF-binding domain of FLT1 as well as a 31-amino-acid stretch derived from an intron and tightly binds VEGF, suppressing its angiogenic activity. Multivariate analysis showed that the sFLT1 levels in cancer tissue were an independent prognostic indicator of disease progression. SFLT1 expression was significantly elevated in colorectal cancer tissue compared with normal mucosa and the intratumoral balance between sFLT1 and VEGF was significantly different between tumour tissue and normal controls. Furthermore, sFLT1 levels showed a significant prognostic value.

6.1.3 *ARHGEF4*

ARHGEF4 - also called APC-stimulated guanine nucleotide exchange factor (*ASEF)* and *GEF4*- codes one of the Rho family guanine nucleotide exchange factors (GEFs) which are essential links between extracellular signalling events and the activation of Rho family GTPases, acting as the direct facilitators of GDP displacement in these molecular switches (Hamann et al 2007). The GTP-loaded and activated Rho family GTPases, such as RHOA, RAC1, and CDC42, have classically been appreciated for their effects on cytoskeletal reorganization and the establishment of cellular polarity but also are known to induce proliferative responses through the binding and activation of proteins such as p21-activated kinase (PAK) and Rho associated kinase. Despite the potential for direct, unregulated cellular proliferation and metastasis through constitutive activation of Rho family GTPases, activating mutations similar to those established for Ras have not been discovered in human cancers. Therefore, Rho family GEFs are frequently investigated in terms of their potential as oncogenic triggers, since they are the first upstream activators of Rho-GTPases and potentially

misregulate GTPases when over expressed or mutated to constitutively active forms.

ARHGEF4 was first identified through a yeast two-hybrid screen using the *APC* armadillo repeat region (APCARM) as bait. APCARM is an important segment of *APC* and is retained in *APC* truncation mutations found in colorectal cancers and FAP. Although its exact function remains elusive, the APCARM interaction localized to an *APC* binding region (ABR) within ARHGEF4, a region lying immediately N-terminal to the protein's Src-homology 3 (SH3) domain. The most profound outcome of the APC-ARHGEF4 interaction is that it stimulates GEF activity, leading to Rac1 activation, lamellipod formation, and increased cell migration.

Kawasaki et al (2003) found evidence suggesting that the APC-ARHGEF4 complex functions in cell migration as well as in E-cadherin-mediated cell-cell adhesion, and that truncated APC present in colorectal tumour cells contributes to their aberrant migratory properties. It has therefore been suggested that the truncated forms of APC often found in colorectal cancer and FAP are not only devastating due to unregulated cellular β-catenin accumulation but may also enhance cellular metastasis due to constitutive *ARHGEF4* activation.

Mitin et al (2007) showed that binding of the armadillo repeats of APC to a 'core APC-binding' (CAB) motif within ARHGEF4, or truncation of the SH3 domain of ARHGEF4, relieves auto inhibition, allowing the specific activation of CDC42. Structural determination of auto inhibited ARHGEF4 revealed that the SH3 domain forms an extensive interface with the catalytic DH and PH domains to obstruct binding and activation of CDC42, and the CAB motif is positioned adjacent to the SH3 domain to facilitate activation by APC. In colorectal cancer cell lines, full-length, but not truncated, APC

activates CDC42 in an ARHGEF4-dependent manner to suppress anchorage-independent growth. The authors therefore proposed a model in which ARHGEF4 acts as a tumour suppressor when activated by APC and inactivation of ARHGEF4 by mutation or APC truncation promotes tumorigenesis.

6.1.4 *RHOA*

RHOA codes a GTPase which is related to rearrangement of the actin cytoskeleton. This process is primarily controlled by the members of Rho small GTPase family (Chang 2006), including *RHOA*, *RAC1*, and *CDC42*. Like other small GTPases, Rho GTPases cycle between inactive GDP-bound and active GTP-bound forms. Activation of Rho GTPases is stimulated by guanine nucleotide exchange factors and inhibited by GTPase-activating proteins. On activation, Rho GTPases recruit effector proteins to regulate the actin cytoskeleton. In epithelial cells, Rho GTPases are implicated in regulating morphology and adhesion because interactions between the actin cytoskeleton and adherens junctions determine cell shape and motility. Formation of adherens junctions promotes cell-cell adhesion and is important in organizing normal epithelial sheets. Adherens junctions consist of the transmembrane protein E-cadherin, whose cytoplasmic domain interacts with h-catenin, which binds α-catenin. Because formation of adherens junctions is associated with actin dynamics, RAC1 and CDC42 activity are required for the formation and maintenance of E-cadherin mediated adherens junctions. Although a basal level of RHOA activity is also necessary for adherens junction formation, high RHOA activity disrupts the formation of adherens junctions. On the other hand, E-cadherin-mediated formation of adherens junctions strongly inhibits RHOA activity but increases activities of both RAC1 and CDC42.

Up-regulation of small GTPases, *RHOA* and *RHOC*, is associated with tumour
progression in ovarian carcinoma (Horiuchi et al 2003). Analysis of mRNA levels of the RHO family genes revealed that levels of both RHOA and RHOC were significantly higher in carcinomas than in benign tumours.

Arango et al (2005) showed that RHOA expression was significantly related to survival for Dukes C cancers. Using immunohistochemistry and a tissue array containing 137 Dukes C cancer samples they found that reduced RHOA expression was associated with significantly shorter survival.

6.2 In Situ Hybridisation Results

6.2.1 ISH results in general

ISH was performed on sections of various tissue types to assess the behaviour of the probes. In order to establish that there was adequate preservation of mRNA in the blocks from which the sections were derived, the levels and patterns of expression of β-actin mRNA were assessed. Tissues from all cases revealed expression of β-actin mRNA, with expected patterns of expression; increasing in strength towards the luminal surface in the epithelium of normal crypts and with the strongest signals in lymphoid aggregates where present.

ISH using the 4 test probes gave expression results consistently only for *CDX2*. There was some expression seen for the *FLT1* and *RHOA* probes but to a lesser degree. *ARHGEF4* did not show meaningful levels of expression (detectable above background).

To assess CDX2 expression, a variety of tissue sections were probed: sections of Crohn's disease small bowel tissue revealed expression in the normal crypt epithelium, and, as expected , not in the regions of pseudopyloric metaplasia (Figure 6.1 a). Hybridisation to Barrett's oesophageal/gastric tissue with colonic type metaplasia revealed expression of CDX2 mRNA in the metaplastic tissue alone (Figure 6.1 b,c). For human villiform intestinal foetal tissue (12.4 weeks) strong CDX2 expression was seen (Figure 6.1 f). In colonic cancer tissue (Figure 6.1 d) fairly homogenous expression was seen within the cancer epithelium as well as in adjacent normal crypts. However, there seemed to be some loss of CDX2 mRNA expression at the invasive edge of the cancer. Normal adjacent crypts showed no weakness of expression at the base, although analysis of normal colonic mucosa (Figure 6.1 e) showed some weakness of expression at the bottom of crypts.

Drawing on the results of Witek (Witek et al (2005)) who showed over expression of CDX2 by human colorectal tumours compared with matched normal mucosa we probed the tissue array to see if we could relate CDX2 expression to outcome. We found variable CDX2 mRNA expression strength for the cancers analysed and therefore results for this array were scored (Figure 6.1 j, k,l) (see next section). RHOA expression was seen for foetal kidney tissue (Figure 6.1 i). Foetal gut did not show convincing RHOA expression. Expression on the tissue array was highly variable but very low level and scoring was not attempted.

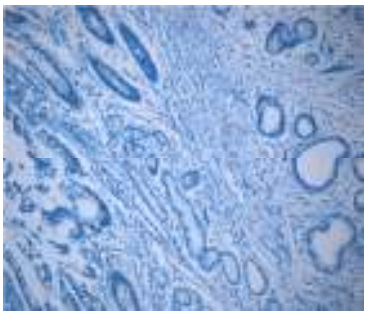
ARHGEF4 expression was weakly positive for foetal gut but 'spotty' in a distribution suspected to be due to eosinophils. For the tissue array ARHGEF4 expression was spotty and this again suggested expression due to eosinophils.

FLT1 expression was seen clearly on kidney tissue (Figure 6.1 h) and for joint tissue (Figure 6.1 g). FLT1 expression was always

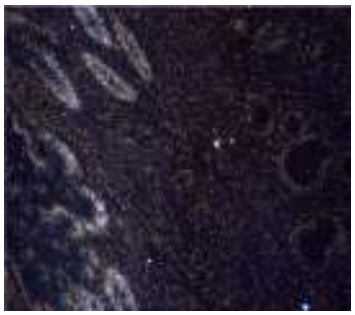
related to endothelial tissue. FLT was also seen related to clusters on small vessels in foetal gut (not shown). Although the tissue array produced expression data such expression related to endothelial tissue and was not cancer specific.

Figures 6.2 a-l . Results of ISH experiments

a).



Small bowel Crohns CDX2 x10 br



Small bowel Crohns CDX2 x10 dk

b).

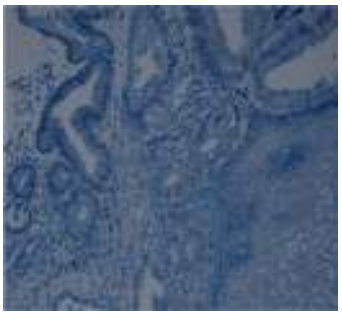
Barretts oesophageal/gastric tissue with colonic type metaplasia (i)

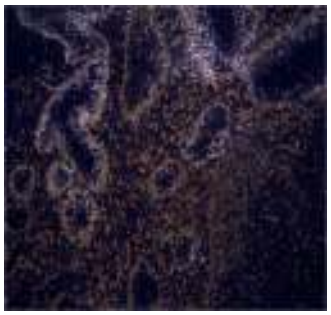
Barretts oesophageal/gastric tissue with colonic type metaplasia (i)

CDX2 x20 bright

CDX2 x20 dark

c).

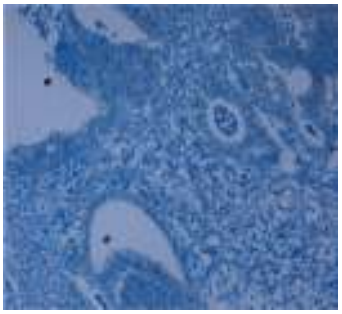




Barretts oesophageal/gastric tissue with colonic type metaplasia (ii) CDX2 x20 br

Barretts oesophageal/gastric tissue with colonic type metaplasia (ii) CDX2 x20 dk

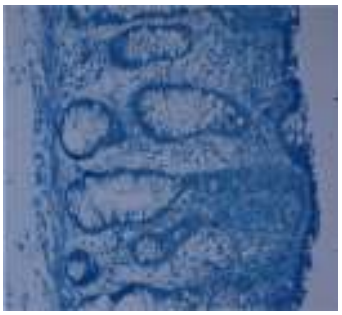
d).



Colonic cancer CDX2 x20 br

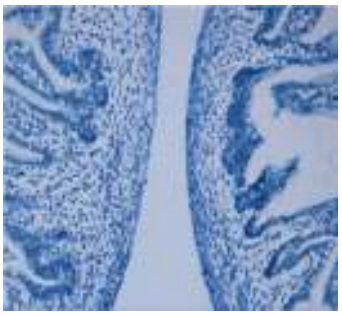
Colonic cancer CDX2 x20 dk

e).



Normal colonic mucosa CDX2 x20 br Normal colonic mucosa CDX2 x20 dk

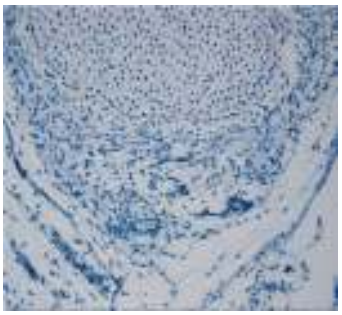
f).



Normal villiform intestinal foetal tissue Normal villiform intestinal foetal tissue

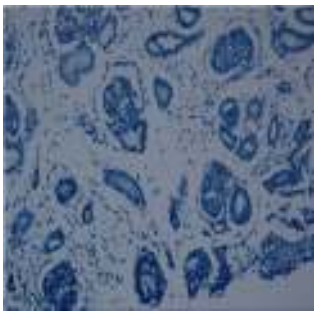
12.4 weeks CDX2 x20 br 12.4 weeks CDX2 x20 dk

g).

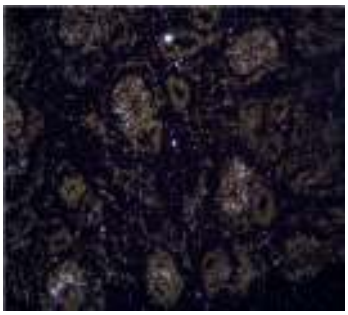


Joint FLT1 x20 br Joint FLT1 x20 dk

h).

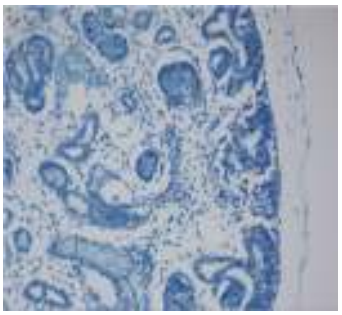


Kidney FLT1 x20 br



Kidney FLT1 x20 dk

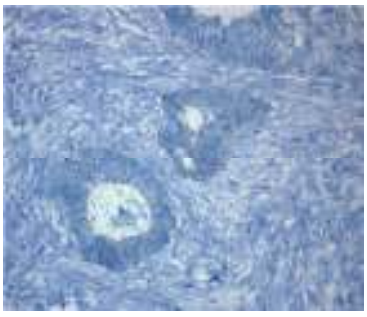
i).



Foetal kidney RHOA x20 br

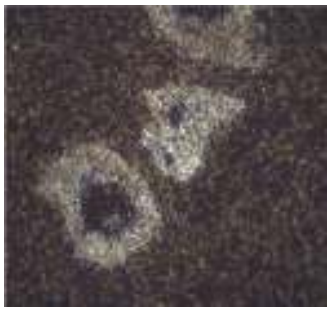
Foetal kidney RHOA x20 dk

j).



CDX2 grade 3 scoring example

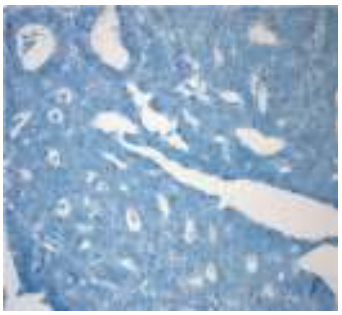
colonic cancer array x 20 br



CDX2 grade 3 scoring example

colonic cancer array x 20 dk

k).



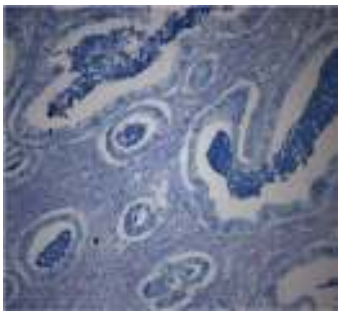
CDX2 grade 2 scoring example

colonic cancer array x 20 br

CDX2 grade 2 scoring example

colonic cancer array x 20 dk

l).



CDX2 grade 1 scoring example

colonic cancer array x 20 br

CDX2 grade 1 scoring example

colonic cancer array x 20 dk

6.2.2 Tissue array CDX2 expression results.

The tissue array was composed of 119 cancers of which 106 had associated patient outcome data. I reiterate that this collection of cancers was completely different to that used in previous experiments. We used the tissue array for the opportunistic reason that it had been already constructed , was available and allowed testing of multiple colorectal cancers in a single ISH experiment. 84(79.2%) were associated with >5 year patient survival and 22(20.7%) with patient death within 5 years of operation. Comparing background clinicopathological data for these groups (Table 5.2)

showed no statistically significant difference between the groups for age, gender, Dukes stage and grade. The proportion of T4 cancers was significantly different between the groups (>5 year survival 33.3% versus < 5 year survival 63.6% p=0.014).

Due to the failure of appropriate expression for the other three genes only CDX2 expression was scored using data from the tissue array experiment. Scoring of expression was 1, 2 or 3 with 3 being the strongest expression (Figure 6.1 j,k,l). Scoring was under taken by me with Professor Richard Poulsom verifying the scoring system used. For those cancers from patients with < 5 years survival scores were 1=4 (18.2%), 2=10 (45.4%), 3=8 (36.4%) versus 1=24 (28.6%), 2=29 (34.5%), 3=31 (36.9%) cancers related to >5 year survival. There was no significant relationship found between CDX2 expression and survival. Also, using a Chi-squared test for trend, the relationship between Dukes stage and CDX2 expression was tested and no significant relationship found (p = 0.52).

Table 6.2 Comparison of clinopathological factors for tissue array outcome groups

		5 year survival	< 5 year survival	p
Number		84	22	
Age		69.7	66.8	0.78
Gender	(Female)	44 (52.3)	11 (50)	1.0
Stage	T1	2 (2.4)	0 (0)	1.0
	T2	9 (10.7)	0 (0)	0.2
	T3	45 (53.6)	8 (36.4)	0.23

T4	28 (33.3)	14 (63.6)	0.014*
Dukes Stage			
A	9 (10.7)	0 (0)	0.2
B	53 (63)	12 (54.5)	0.47
C	22 (26.2)	10 (45.4)	1.0
Grade			
Well	32 (38)	5 (22.7)	0.22
Mod	47 (56)	16(72.7)	0.22
Poor	5 (5.9)	1 (4.5)	1.0
CDX2 Expression			
1	24 (28.6)	4 (18.2)	0.42
2	29 (34.5)	10 (45.4)	0.46
3	31 (36.9)	8 (36.4)	1.0

6.3 Discussion

For *CDX2* it remains unclear from our work whether the finding of reduced *CDX2* expression related to the invading edge in cancer represents an important finding suggestive of a role for reduced *CDX2* expression in relation to local invasion.

The gene was selected on the basis of an area of gain found on chromosome 13 with aCGH so we were expecting a gain in expression related to a gain in copy number rather than an effect related to reduced expression.

The loss of expression from the deepest invasive cellular layer might suggest the commencement of a possible change of cell type. An alternative explanation of this finding is provided by the fact that some normal crypts seem to have lost expression of CDX2 at the bottom of the crypts and that this loss of expression seen in cancer tissue simply mirrors that seen in normal tissue and is of no pathological significance. As the introduction to this chapter shows there is data suggesting that *CDX2* may be important in maintaining the commitment to small and large intestinal lineages, we are unable to make definite conclusions. The attempt to correlate CDX2 expression to outcome did not produce a significant result.

RHOA and *ARHGEF4* produced no results of significance. It may be that there is increased expression of either gene in CRC but we did not show this with the probes used. This is particularly true of *ARHGEF4* were eosinophilic expression was detected; this might be due to the presence of particular 'sticky' probe domains. Given the finding of RHOA expression correlating with Dukes C survival when expression was assessed using immunohistochemistry one might have hoped our experiments to have produced more conclusive results. The probe undoubtedly worked but we could not replicate the relation to prognosis found by Arango (Arango et al 2005).

As *FLT1* codes for a receptor the expression results are difficult to interpret. It is difficult to know what purpose an isolated increase in *FLT1* gene expression would serve in the context of tumour progression. One could conceive of an observed increase in receptor numbers related to an increase in endothelial cell numbers. That could be checked by counting vessels in sections stained with for example CD31.

7. Analysis of Sporadic Adenomas with Array Comparative Genomic Hybridisation

7.1 **Introduction.**

As discussed in the introduction (1.2.1 and 1.2.2) the timing and form of CIN in CRC have not yet been determined. In this part of the study, the aim was to examine the evidence for chromosomal-scale changes in colorectal adenomas (CRA) using aCGH. We deliberately analysed lesions that had grown and/or progressed from the very small, mildly dysplastic stage, so as to increase the chances of detecting genetic changes (this work was included in the publication Jones et al (2007). At the time of designing this work no work existed using aCGH to investigate adenomas.

All adenomatous tissue came from Oxford and was processed in the same way as the cancer tissue described earlier. Adenomas were snap frozen and stored with paired normal tissue at -70 C. DNA was extracted as previously described (2.1). Adenomas were examined histopathologically and for MSI status and ploidy (2.3, 2.4, 2.5).Then examined with aCGH (2.6).

7.2 **Results.**

Adenoma characteristics.

All 14 polyps were greater than 1cm in size. 12 were mild or moderately dysplastic with 2 severely dysplastic. 11 polyps were tubular adenomas and 3 were tubulovillous adenomas. All tissue

from which DNA was derived was confirmed as containing at least 60% adenomatous tissue by Professor Mohammad Ilyas.

MSI and ploidy.

No adenomas showed MSI on examination with the markers BAT 26 and D5S346. All adenomas were diploid on FACS analysis.

aCGH

Appendix 11 gives a summary of the aCGH findings. Changes were found in 10 cases. 6 cases had 1-3 changes, 2 had 4 changes and 1 had 12 changes and 1 had 13 changes. Of 44 chromosmal scale changes 12 involved the whole chromosome. There were more losses than gains (38 versus 6). The median number of gains was 0 (range 0-4). The median number of losses was 1 (range 0-12). Commonest changes seen were for 1p loss (57%), 17p loss (50%), 19 loss (28.6%), 13 gain (21.4%), and 22 loss (21.4%).

7.3 **Discussion**

We have found that a small number of large-scale genetic changes can be detected by aCGH in the majority of adenomas from sporadic cases. Whilst our adenomas were not ‘early’ lesions based on their size, all were scored as near-diploid by flow cytometry.The most frequent regions of change that we have found in our adenomas are similar to those found by Cardoso et al (2006), with the exception that chromosome 7 was not commonly gained in our series. Interestingly some changes - such as the deletion of chromosome 18q – that are thought to be present in early carcinomas (Hermsen et al 2002) were almost absent in the adenomas. Also most aCGH changes involved the smaller chromosomes, suggesting that these may have been tolerated as ‘background’ changes that produced no

great selective advantage or disadvantage. We would have liked to investigate more adenomas but the number of aCGH arrays available was limited.

8. Conclusions and Future Research Directions

Whilst we have not found differences between good and bad outcome Dukes B cancers that could be usefully used to predict prognosis because such differences as have been detected are only seen in few cancers, and not discovered evidence for a new gene of significance in relation to CRC, the techniques used remain valid in the search for answers to the questions addressed.

I have completed the analysis that was originally planned, collecting sufficient Dukes B cancers to meet the power calculation and performing all planned experiments on these with technical success. The work has yielded new areas of small gain and loss that could be investigated for new genes of importance in colorectal cancer. Also I have identified areas of copy number change in sporadic adenomas.

In terms of finding new genes of interest, the areas we have investigated contain many genes and continued investigation of these areas with immunohistochemical staining, were antibodies against proteins related to genes of interest are available, and ISH may yield results of interest.

The fact that 18q LOH has been shown to have prognostic ability gives rise to optimism that there is a difference between good and bad outcome cancers in terms of the factors measurable with aCGH. It may also be seen that outcome difference has been shown between MSI positive and negative cancers and CIN positive and negative cancers. It is probable that if a sample set associated with sufficiently high quality pathology, clinical and follow up data could be collected in sufficient numbers results of value might be produced. The use of the newer commercially available oligonucleotide arrays

will be helpful in terms of array availability and increased ease of usage.

Genes within the small areas of gain could be correlated with the results of large gene linkage studies such as the CORGI trial to ty and identify the true genes of interest in these areas.

In terms of the timing of CIN the adenoma work could be repeated using SNP arrays in the hope a identifying those genes exhibiting copy number change at the earliest stage. This would involve the microdissection of individual aberrant crypts and the extraction of DNA from these small entities which constitute the earliest histologically observable changes.

9.List of Tables

10. List of Figures

i. Foetal kidney RHOA

j. CDX2 grade 3 example colonic cancer tissue array

k. CDX2 grade 2 example colonic cancer tissue array

l. CDX2 grade 1 example colonic cancer tissue array

11. References

Aaltonen L.A., Peltomaki P., Leach F.S. et al. Clues to the pathogenesis of familial colorectal cancer. Science1993; 260: 812-816.

Ahuja N., Mohan A.L., Li Q. et al. Association between CpG island methylation and microsatellite instability in colorectal cancer. Cancer Res 1997; 57: 3370-3374.

Albe X, Vassilakos P, Helfer-Guamori K, el al. Independent prognostic value of ploidy in colorectal cancer. Cancer 1990; 66:1168-1175.

Allegra CJ, Paik S, Colangelo LH 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: 241-250.

Allen W, Johnston PG. Role of genetic markers in colorectal cancer treatment. J Clin Oncol 2005; 23: 4545-4552.

Al-Mulla F, AlFadhli S, Al-Hakim AH, Going JJ, Bitar MS. Metastatic recurrence of early stage colorectal cancer is linked to loss of heterozygosity on chromosomes 4 and 14q. J Clin Path 2006; 59: 624-630.

Al-Mulla F, Behbehani AI, Bitar MS, Varadharaj G, Going JJ. Genetic profiling of stage I and II colorectal cancer may predict metastatic relapse. Mod Pathol 2006; 19: 648-58.

Alavi S., Namazie A., Calcaterra T.C., Wang M.B. & Srivatsan E.S. Clinical application of fluorescence in situ hybridization for chromosome 11q13 analysis in head and neck cancer. Laryngoscope 1999; 109: 874-879.

Albertson D.G, Pinkel D. Genomic microarrays in human genetic disease and cancer. Hum Mol Genet 2003; 12:145-52.

Albertson D.G., Ylstra B., Segraves R. et al. Quantitative mapping of amplicon structure by array CGH identifies CYP24 as a candidate oncogene. Nat Genet 2000; 25: 144-146.

Anand S, Penrhyn-Lowe S, Venkitaraman AR. AURORA-A amplification overrides the mitotic spindle assembly checkpoint, inducing resistance to Taxol. Cancer Cell 2003; 3: 51-62.

Arango D, Laiho P, Kokko A et al. Gene expression profiling predicts recurrence in Dukes' C colorectal cancer. Gastroenterology 2005; 129: 874-884.

Araujo SE, Bernardo WM, Habr-Gama A et al. DNA ploidy status and prognosis in colorectal cancer: a meta-analysis of published data. Dis Colon Rectum. 2007; 50:1800-1810.

Arrabas R, Ribas M, Risques R et al. Prospective assessment of allelic losses at 4p14-16 in colorectal cancer: Two mutational patterns and a locus associated with poorer survival. Clin Can Res 1999; 5; 3454-3459.

Arribas R, Capella G, Tortola S, et al. Assessment of genomic damage in colorectal cancer by DNA fingerprinting: prognostic applications. J Clin Oncol 1997; 15: 3230– 3240.

Bala S, Peltomaki P. CYCLIN D1 as a genetic modifier in hereditary nonpolyposis colorectal cancer. Cancer Res 2001; 61: 6042-6045.

Barratt PL, Seymour MT, Stenning SP et al. Molecular markers predicting benefit from adjuvant 5-flourouracil in colon cancer patients. Lancet 2002; 360 :1381-1391

Bassing CH, Suh H, Ferguson DO et al. Histone H2AX: a dosage-dependent suppressor of oncogenic translocations and tumors. Cell 2003; 114 : 359-370.

Bauer BF, Chacho MS, Roseman DL et al. Multiparameter flow cytometric analysis of colon polyps. Am J Ch Pathol 1987; 87:313-318.

Bazan V, Migliavacca M, Zanna I, 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: 650–658.

Behrens J., von Kries J.P., Kuhl M. et al. Functional interaction of beta-catenin with the transcription factor LEF-1. Nature 1996; 382: 638-642.

Behrens J, Jerchow B, Wurtele M et al. Functional interaction of an axin homolog, conductin, with beta-catenin, APC, and GSK3beta. Science 1998; 280: 596-599.

Bell SM, Scott N, Cross D et al. Prognostic value of p53 overexpression and c-Ki-ras gene mutations in colorectal cancer. Gastroenterology 1993; 104: 57-64.

Bellovin DI, Simpson KJ, Danilov T et al. Reciprocal regulation of RhoA and RhoC characterizes the EMT and identifies RhoC as a prognostic marker of colon carcinoma. Oncogene. 2006; 25: 6959-6967.

Benahmed F, Gross I, Guenot D et al. The microenvironment controls CDX2 homeobox gene expression in colorectal cancer cells. Am J Pathol. 2007 ; 170: 733-44.

Bharadwaj R, Yu H. The spindel checkpoint, aneuploidy and cancer. Oncogene 2004; 23: 2016-2027.

Bhatavdekar JM, Patel DD, Chikhlikar PR et al. Molecular markers are predictors of recurrence and survival in patients with Dukes B and C colorectal adenocarcinoma. Dis Colon Rectum 2001; 44: 523-533.

Bignell G.R., Huang J., Greshock J. et al. High-resolution analysis of DNA copy number using oligonucleotide microarrays. Genome Res 2004; 14: 287-295.

Birbeck KF, Macklin CP, Tiffin NJ et al. Rates of circumferential resection margin involvement vary between surgeons and predict outcomes in rectal cancer surgery. Ann Surg 2002; 235: 449-457.

Boon E , van der Neut R, van de Wetering M et al. Wnt signaling regulates expression of the receptor tyrosine kinase met in colorectal cancer. Cancer Res 2002; 62: 5126-5128.

Bos JL. The ras gene family and human carcinogenesis. Mutat Res 1988; 195: 255-271.

Bouzourene H, Gervaz P, Cerottini JP et al. P53 and Ki-ras as prognostic factors for Dukes stage B colorectal cancer. Eur J Cancer 2000; 36: 1008-1015.

Brabletz T, Jung A, Dag S, Hlubek F, Kirchner T. beta-catenin regulates the expression of the matrix metalloproteinase-7 in human colorectal cancer. Am J Pathol 1999; 155: 1033-1038.

Brantjes H, Roose J, van De Wetering M et al. All Tcf HMG box transcription factors interact with Groucho-related co-repressors. Nucleic Acids Res 2001; 29: 1410-1419.

Bravou V, Klironomos G, Papadaki E et al. ILK over-expression in human colon cancer progression correlates with activation of beta-catenin, down regulation of E-cadherin and activation of the Akt-FKHR pathway. J Pathol 2006; 208: 91-99.

Broderick P, Carvajal-Carmona L, Pithan AM et al. A genome wide association study shows that common alleles of SMAD 7 influence colorectal cancer risk. Nat Genet. 2007; 39: 1315-1317.

Bronner C.E., Baker S.M., Morrison P.T. et al. Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. Nature 1994; 368: 258-261.

Bubb VJ, Curtis LJ, Cunningham C, Dunlop MG et al. Microsatellite instabilitv and the role of hMSH2 in sporadic colorectal cancer. Oncogene 1996; 12: 2641-2649.

Buckley P.G., Mantripragada K.K., Benetkiewicz M. et al. A full-coverage, high-resolution human chromosome 22 genomic microarray for clinical and research applications. Hum Mol Genet 2002; 11: 3221-3229.

Buglioni S, D'Agnano I, Vasselli 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: 360-368.

Bussey HJ. Familial polyposis coli. Baltimore: Johns Hopkins University Press, 1975. Cardoso J, Boer J, Morreau H, Fodde R.

Expression and genomic profiling of colorectal cancer. Biochim et Biophys Acta 2007; 1775: 103-137.

Cardoso J, Molenaar L, de Menezes RX, et al. Chromosomal instability in MYH- and APC-mutant adenomatous polyps. Cancer Res 2006; 66: 2514-2519.

Carmena M, Earnshaw WC. The cellular geography of Aurora kinases. Nat Rev Mol Cell Biol 2003; 4: 842-854.

Celeste A, Peterson S, Romanienko PJ et al. Genomic instability in mice lacking histone H2AX. Science 2002; 296:922-927.

Chang Y, Marlin J, Chance T et al. RhoA mediates cyclooxygenase-2 signaling to disrupt the formation of adherens junctions and increase cell motility. Cancer Res. 2006; 66: 11700-11708.

Chapuis PH, Dent OF, Fisher R, et al. A multivariate analysis of clinical and pathological variables in prognosis after resection of large bowel cancer. Br J Surg 1985; 72: 698-702.

Chawengsaksophak K, James R, Hammond VE et al. Homeosis and intestinal tumours in Cdx2 mutant mice. Nature 1997; 386: 84–87.

Choi SW, Lee KJ, Bae YA, et al. Genetic classification of colorectal cancer based on chromosomal loss and microsatellite instability predicts survival. Clin Cancer Res 2002; 8: 2311 –22.

Chu DZ, Glacco G, Martin RG, Gu et al. The significance of synchronous carcinoma and polyps in the colon and rectum. Cancer 1986; 57: 445-450.

Chun SY, Chen F, Washburn JG et al. CDX2 promotes anchorage-independent growth by transcriptional repression of IGFBP-3. Oncogene. 2007; 26: 4725-4729.

Chung D. The genetic basis of colorectal cancer: insights into critical pathways of tumorigenesis. Gastroenterology 2000; 119: 854-865.

Clark S. Inherited bowel cancer. Colorectal surgery Third edition 2005 Elsevier Saunders

Compton C, Fielding P, Burgart L et al. Prognostic factors in colorectal cancer. College of American Pathologists Consensus Statement 1999. Arch Pathol Lab Med 2000; 124: 979-994.

Copeland EM, Miller LD, Jones RS et al. Prognostic factors in carcinoma of the colon and rectum. Am J Surg 1968; 116: 875-881.

Crawford H.C, Fingleton B.M., Rudolph-Owen L.A. et al. The metalloproteinase matrilysin is a target of beta-catenin transactivation in intestinal tumors. Oncogene 1999; 18: 2883-2891.

Crowe PJ, Yang JL, Berney CR, et al. Genetic markers of survival and liver recurrence after resection of liver metastases from colorectal cancer. World J Surg 2001; 25: 996– 1001.

Crucitti F, Solo L, Doglietto GB, et al. Prognostic factors in colorectal cancer: current status and new trends. J Surg Oncol 1991; 2: 76-82.

Cui H, Horon I, Ohlsson R et al. Loss of imprinting in normal tissue of colorectal cancer patients with microsatellite instability. Nat Med 1998; 4: 1276-1280.

Cunningham J, Christensen E, Tester D, et al. Hypermethylation of the hMLH1 promoter in colon cancer with microsatellite instability. Cancer Res 1998; 58: 3455-3460.

Dang LH, Chen F, Ying C et al. CDX2 has tumorigenic potential in the human colon cancer cell lines LOVO and SW48. Oncogene. 2006; 25 :2264-2272.

Davies H., Bignell G.R., Cox C. et al. Mutations of the BRAF gene in human cancer. Nature 2002; 417: 949-954.

De Angelis PM, Stokke T, Beigi M et al. Prognostic significance of recurrent chromosomal aberrations detected by comparative genomic hybridisation in sporadic colorectal cancer. Int J Colorectal Dis 2001; 16: 38-45.

Deans GT, Patterson CC, Parks TG, et al. Colorectal carcinoma: importance of clinical and pathological factors in survival. Ann R Coll Surg Engl 1994; 76: 59.

Deng G, Bell I, Crawley S, Gum J, Terdiman JP, Allen BA, Truta B, Sleisenger MH, Kim YS. BRAF mutation is frequently present in sporadic colorectal cancer with methylated hMLH1, but not in hereditary nonpolyposis colorectal cancer. Clin Cancer Res 2004; 10:191-195.

Diep CB, Kleivi K, Ribeiro FR, et al. The order of genetic events associated with colorectal cancer progression inferred from meta-analysis of copy number changes, Genes Chromosomes Cancer 2006; 45: 31–41.

Diep CB, Thorstensen L, Meling GI et al. Genetic tumor markers with prognostic impact in Dukes B and C colorectal cancer patients. J Clin Oncol 2003; 21: 820-829.

Dikovskaya D, Newton IP, Nathke IS. The adenomatous polyposis coli protein is required for the formation of robust spindles formed in CSF Xenopus extracts. Mol Biol Cell 2004; 15: 2978-2991.

Douglas EJ, Fiegler H, Rowan A, et al. Array comparative genomic hybridization analysis of colorectal cancer cell lines and primary carcinomas, Cancer Res 2004; 64: 4817–4825.

Duesberg P, Li R, Rasnick D. Aneuploidy approaching a perfect score in predicting and preventing cancer: highlights from a conference held in Oakland, CA in January, 2004. Cell Cycle. 2004; 3: 823-8.

Dutt A, Beroukhim R. Single nucleotide polymorphism array analysis of cancer. Curr Opin Oncol. 2007; 19: 43-49.

Dwight RW, Higgins GA, Keehan RJ. Factors influencing survival after resection in cancer of the colon and rectum. Am J Surg 1969; 117: 512- 522.

Eide TJ. Prevalence and morphological features of adenomas of the large intestine with and without colorectal carcinoma. J Histopathol 1986; 10: 111-118.

Elsaleh H, Powell B, McCaul K et al. P53 alteration and microsatellite instability have predictive value for survival benefit from chemotherapy in stage III colorectal carcinoma. Clin Cancer Res 2001; 7: 1343-1349.

Eshleman J.R., Casey G., Kochera M.E. et al. Chromosome number and structure both are markedly stable in RER colorectal cancers and are not destabilized by mutation of p53. Oncogene 1998; 17: 719-25.

Esteller M, Corn PG, Baylin SB et al. A gene hypermethylation profile of human cancer. Cancer Res 2001; 61: 3225-3229.

Ewart-Toland A, Briassouli P, de Koning JP et al. Identification of Stk6/STK15 as a candidate low-penetrance tumor susceptibility gene in mouse and human. Nat Genet 2003; 34: 403-412.

Fagotto F., Jho E., Zeng L. et al. Domains of axin involved in protein-protein interactions, Wnt pathway inhibition, and intracellular localization. J Cell Biol 1999; 145: 741-756.

Fan F, Wey JS, McCarty MF et al. Expression and function of vascular endothelial growth factor receptor-1 on human colorectal cancer cells. Oncogene. 2005; 24: 2647-53.

Fearnhead NS, Wilding JL, Bodmer WF. Genetics of colorectal cancer: hereditary aspects and overview of colorectal tumorigenesis. Brit Med Bull 2002; 64: 27-43.

Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell 1990; 61: 759-767.

Fiegler H, Gribble S, Burford D, et al. Array painting: a method for the rapid analysis of aberrant chromosomes using DNA microarrays. J Med Genet 2003; 40: 664-670.

Fiegler H., Carr P., Douglas E.J. et al. DNA microarrays for comparative genomic hybridization based on DOP-PCR amplification of BAC and PAC clones. Genes Chromosomes Cancer 2003; 36: 361-374.

Fielding LP, Phillips RK, Fry JS et al. Prediction of outcome after curative resection for large bowel cancer. Lancet 1986; 18: 904-906.

Fishel R., Lescoe M.K., Rao M.R. et al. The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. Cell 1993; 75: 1027-1038.

Flyger HL, Larsen JK, Nielsen HJ et al. DNA ploidy in colorectal cancer, heterogeneity within and between tumors and relation to survival. Cytometry 1999; 38: 293–300.

Fodde R, Smits R, Clevers H. APC, signal transduction and genetic instability in colorectal cancer. Nature Reviews 2001; 1: 55-67.

Fodde R, Kuipers J, Rosenberg C. et al. Mutations in the APC tumour suppressor gene cause chromosomal instability. Nat Cell Biol 2001; 3: 433-438.

Fodde R, Kuipers J, Rosenberg C et al. Mutations in the APC tumor suppressor gene cause chromosomal instability. Nature Cell Biol 2001; 3: 433-438.

Forrester K., Almoguera C., Han K et al. Detection of high incidence of K-ras oncogenes during human colon tumorigenesis. Nature 1987; 327: 298-303.

Freeman JC, Penny GH, Feuk L et al. Copy number variation; new insights in genome diversity. Genome Res. 2006; 16: 949-961.

Freedman LS, Macaskill P, Smith AN. Multivariate analysis of prognostic factors for operable rectal cancer. Lancet 1984; 2: 733-736.

Fritz B., Schubert F., Wrobel G. et al. Microarray-based copy number and expression profiling in dedifferentiated and pleomorphic liposarcoma. Cancer Res 2002; 62: 2993-2998.

Fujita M., Furukawa Y., Tsunoda T. et al. Up-regulation of the ectodermal-neural cortex 1 (ENC1) gene, a downstream target of the beta-catenin/T-cell factor complex, in colorectal carcinomas. Cancer Res 2001; 61: 7722-7726.

Gaasenbeek M, Howarth K, Rowan A, et al. Combined array-Comparative Genomic Hybridization and Single Nucleotide Polymorphism Loss of Heterozygosity analysis reveals complex changes and multiple forms of chromosomal instability in colorectal cancers. Cancer Res 2006; 66: 3471-3479.

Garcia-Peche P, Vasquez-Prado A, Fabra-Ramis R et al. Factors of prognostic value in long-term survival of colorectal cancer patients. Hepatogastroenterology 1991; 38: 438-443.

Georgiades IB, Curtis LJ, Morris RM et al. Heterogeneity studies identify a subset of sporadic colorectal cancers without evidence for chromosomal or microsatellite instability. Oncogene 1999; 23: 7933-7940.

Gerdes H, Chen Q, Elahi AH et al. Recurrent deletions involving chromosomes 1, 5, 17 and 18 in colorectal carcinoma: Possible role in biological and clinical behaviour of tumours. Anticancer Res 1995; 15: 13-24.

Gervaz P, Bouzourene H, Cerottini JP et al. Dukes B colorectal cancer: Distinct genetic categories and clinical outcome based on proximal or distal tumor location. Dis Colon Rectum 2001; 44: 364-373.

Giaretti W, Sciallev S, Bruno S et al. Cytometry of endoscopically examined colorectal adenomas and adenocarcinomas. Cytometry 1998; 9: 238-244.

Giaretti W, Venesio T, Prevosto C, Lombardo F, Ceccarelli J, Molinu S, Risio M. Chromosomal instability and APC gene mutations in human sporadic colorectal adenomas. J Pathol 2004; 204: 193-199.

Giatromanolaki A, Strathopoulos GP, Tsiobanou E et al. Combined role of tumor angiogenesis, bcl-2, and p53 expression in prognosis of patients with colorectal carcinoma. Cancer 1999; 86: 1421-1430.

Gisselsson D, Hoglund M. Connecting mitotic instability and chromosome aberrations in cancer-can telomeres bridge the gap? Seminars in Cancer Biology 2005; 15: 13-23.

Godwin JD, Brown CC. Some prognostic factors in survival of patients with cancer of the colon and rectum. J Chronic Dis 1975; 28: 441-454.

Gollin SM. Mechanisms leading to chromosomal instability. Seminars in Cancer Biology 2005; 15: 33-42.

Gotthardt K, Ahmadian MR. Asef is a Cdc42-specific guanine nucleotide exchange factor. Biol Chem. 2007; 388 :67-71.

Grady W.M, Markowitz S.D. Genetic and epigenetic alterations in colon cancer. Annu Rev Genomics Hum Genet 2002; 12: 3101-28.

Granquist S. Distribution of polyps in the large bowel in relation to age. A colonoscopic study. Scand J Gastroenterol 1981; 16:1025-1031.

Griffin MR, Bergstralh EJ, Coffey RJ et al. Predictors of survival after curative resection of carcinoma of the colon and rectum. Cancer 1987; 60: 2318-24.

Grizzle WE, Manne U, Weiss HL et al. Molecular staging of colorectal cancer in African-American and Caucasian patients using phenotypic expression of p53, bcl-2, MUC 1 and p27kip1. Int J Cancer 2002; 97: 403-409.

Geido E, Sciutto A, Rubagotti A, et al. Combined DNA flow cytometry and sorting with k-ras2 mutation spectrum analysis and the prognosis of human sporadic colorectal cancer. Cytometry 2002; 50: 216– 24.

Haigis KM, Caya JG, Reichelderfer M et al. Intestinal adenomas can develop with a stable karyotype and stable microsatellites. Proc Natl Acad Sci USA 2002; 99: 8927-8931.

Haigis KM, Hoff PD, White A et al. Tumor regionality in the mouse intestine reflects the mechanism of loss of APC function. Proc Natl Acad Sci USA 2004; 101:9769-9773.

Halling KC, French AJ, McDonnell SK et al. Microsatellite instability and 8p allelic imbalance in stage B2 and C colorectal cancers. J Nat Can Inst 1999; 91: 1295-1303.

Hamann MJ, Lubking CM, Luchini DN, Billadeau DD. Asef2 functions as a Cdc42 exchange factor and is stimulated by the release of a autoinhibitory module from a concealed C-terminal activation element. Mol Cell Biol. 2007; 27: 1380-93.

Hanahan D. and Weinberg R.A. The hallmarks of cancer. Cell 2000; 100: 57-70.

Hanrahan V, Currie M, Gunningham S. et al. The angiogenic switch for vascular endothelial growth factor (VEGF)-A, VEGF-B, VEGF-C, and VEGF-D in the adenoma-carcinoma sequence during colorectal cancer progression. J Pathol 2003; 200: 183-94.

Harrington EA, Bebbington D, Moore J et al. VX-680, a potent and selective small-molecular inhibitor of the Aurora kinases, suppresses tumor growth in vivo. Nat Med 2004; 10: 262-267.

Hart M, de los Santos R, Albert I.N et al. Down regulation of beta-catenin by human Axin and its association with the APC tumor suppressor, beta-catenin and GSK3 beta. Curr Biol 1998; 8: 573-581.

Hauf S, Cole RW, LaTerra S et al. The small molecule Hesperadin reveals a role for Aurora B in correcting kinetochore-microtubule attachment and in maintaining the spindle assembly checkpoint. J Cell Biol 2003; 161: 281-294.

Hawkins NJ, Tomlinson I, Meagher A, Ward RL. Microsatellite-stable diploid carcinoma: a biologically distinct and aggressive subset of sporadic colorectal cancer. B J Cancer 2001; 84: 232-236.

Haydon AM, Jass J. Emerging pathways in colorectal cancer development. Lancet 2002; 3: 83-88.

He T, Sparks A, Rago C. et al. Identification of c-MYC as a target of the APC pathway. Science 1998; 281: 1509-1512.

Heald RJ, Ryall RD. Recurrence and survival after total mesorectal excision for rectal cancer. Lancet 1986 ; 1: 1479-1482.

Hemminki A, Markie D, Tomlinson I. et al. A serine/threonine kinase gene defective in Peutz-Jeghers syndrome. Nature 1998; 391: 184-187.

Herman J.G., Umar A., Polyak K. et al. Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. Proc Natl Acad Sci U S A 1998; 95: 6870-6875.

Hermanek P, Guggenmoons-Holzmann I, Gall FP. Prognostic factors in rectal carcinoma. Dis Colon Rectum 1989; 32: 593-599.

Hermanek P, Wiebelt H, Riedl S et al. Langzeitergebnisse der chirurgischen therapie des coloncarcinoma. Chirurg 1994; 65: 287-297.

Hlubek F, Jung A, Kotzor N et al. Expression of the invasion factor laminin gamma2 in colorectal carcinomas is regulated by beta-catenin. Cancer Res 2001; 61: 8089-8093.

Hao X, Tomlinson I, Ilyas M et al. Reciprocity between membranous and nuclear expression of beta-catenin in colorectal tumours. Virchows Arch 1997; 431: 167-172.

Hermsen M, Postma C, Baak J et al. Colorectal adenoma to carcinoma progression follows multiple pathways of chromosomal instability. Gastroenterology 2002; 123:1109-19.

Homfray TF, Cottrell SE, Ilyas M. et al. Defects in mismatch repair occur after APC mutations in the pathogenesis of sporadic colorectal tumours. Hum Mutat 1998; 11: 114-120.

Horiuchi A, Imai T, Wang C et al. Up-regulation of small GTPases, RhoA and RhoC, is associated with tumor progression in ovarian carcinoma. Lab Invest. 2003; 83: 861-70.

Horn A, Dahl O, Morild I. Venous and neural invasion as predictors of recurrence in rectal adenocarcinoma. Dis Colon Rectum 1991; 34: 798-804.

Hosoi Y, Watanabe T, Nakagawa K. et al. Up-regulation of DNA-dependent protein kinase activity and Sp1 in colorectal cancer. Int J Oncol 2004; 25: 461-468.

Houlston RS. What we could do now: molecular pathology of colorectal cancer. J Clin Pathology 2001; 54:206-214.

Houlston RS, Tomlinson I. Genetic prognostic markers in colorectal cancer. J Clin Pathol 1997; 50: 281-288.

Huang J, Papadopoulos N, McKinley A.J. et al. APC mutations in colorectal tumors with mismatch repair deficiency. Proc Natl Acad Sci U S A 1996; 93: 9049-9054.

Hughes S, Williams RD, Webb E et al. Meta-analysis and pooled re-analysis of copy number changes in colorectal cancer detected by comparative genomic hybridization. Anticancer Res. 2006; 26: 3439-44.

Hurst C.D., Fiegler H., Carr P. et al. High-resolution analysis of genomic copy number alterations in bladder cancer by microarray-based comparative genomic hybridization. Oncogene 2004; 23: 2250-2263.

Ikeda S, Kishida M, Matsuura et al. GSK-3beta-dependent phosphorylation of adenomatous polyposis coli gene product can be modulated by beta-catenin and protein phosphatase 2A complexed with Axin. Oncogene 2000; 19: 537-545.

Ikeda Y, Mori M, Adachi Y et al. Prognostic value of the histochemical expression of helix pomatia agglutinin in advanced colorectal cancer: a univariate and multivariate analysis. Dis Colon Rectum 1994; 37: 181-194.

Ilyas M, Tomlinson IPM, Hanby A et al. Bcl-2 expression in colorectal tumours: evidence of different pathways in sporadic and ulcerative colitis-associated carcinomas. Am J Pathol 1996; 149: 1719- 1726.

Ilyas M, Straub J, Tomlinson IP et al. Genetic pathways in colorectal and other cancers. Eur J Cancer 1999; 35: 335-351.

Ilyas M, Hao X, Wilkinson K et al. Loss of Bcl-2 expression correlates with tumour recurrence in colorectal cancer. Gut 1998; 43: 383–387.

Ilyas M, Tomlinson I.P, Hanby A et al. Allele loss, replication errors and loss of expression of E-cadherin in colorectal cancers. Gut 1997; 40: 654-659.

Iwamoto M, Ahnen DJ, Franklin WA et al. Expression of beta-catenin and full length APC protein in normal and neoplastic colonic tissues. Carcinogenesis 2000; 21: 1935-1940.

Jaeger E, Webb E, Howarth K et al. Common genetic variants at the CRAC1 (HMPS) locus on chromosome 15q 13.3 influence colorectal cancer risk. Nat Genet 2008; 40: 26-28.

Jaffrey RG, Pritchard SC, Clark C et al. Genomic instability at the BUB1 locus in colorectal cancer, but not in non-small cell lung cancer. Cancer Res 2000; 60: 4349-4352.

Jass JR, Atkin WS, Cuzick J, et al. The grading of rectal cancer: historical perspectives and a multivariate analysis of 447 cases. Histopathology 1986; 10: 437-459.

Jen J, Kim H, Piantadosi S et al. Allelic loss of chromosome 18q and prognosis in colorectal cancer. N Eng J Med 1994; 331: 213-221.

Jeffrey R, Hunt T, Poulsom R. In situ hybridisation combined with immunohistochemistry to localise gene expression. Methods in Molecular Medicine. Chapter 23; 323-346. Editors; Brookes SA, Harris AL. Humana Press Inc, Totowa NJ 2005.

Jessup JM, Loda M. Prognostic markers in rectal carcinoma. Semin Surg Oncol 1998; 15: 131-140.

Jiang J, Struhl G. Regulation of the Hedgehog and Wingless signalling pathways by the F-box/WD40-repeat protein Slimb. Nature 1998; 391: 493- 496.

Jones A, Douglas E, Halford S, et al. Array-CGH analysis of microsatellite-stable, near-diploid bowel cancers and comparison with other types of colorectal carcinoma. Oncogene 2005; 24: 118-129.

Jones A, Thirlwell C, Chambers W et al. Analysis of copy number changes suggests chromosomal instability in a minority of large colorectal adenomas. J. Pathol 2007; 213: 249-256.

Jubb AM, Bell SM, Quirke P. Methylation and colorectal cancer. J Pathol 2001; 195: 111-134.

Kallioniemi A, Kallioniemi O, Sudar D. et al. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science 1992; 258: 818-821.

Kane M.F., Loda M., Gaida G.M. et al. Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines. Cancer Res 1997; 57: 808-811.

Kaplan KB, Burds AA, Swedlow JR et al. A role for the adenomatous polyposis coli protein in chromosomal segregation. Nature Cell Biol 2001; 3: 429-433.

Katayama H, Brinkley WR, Sen S. The Aurora kinases: role in cell transformation and tumorigenesis. In: Kumar R, editor. Cancer and metastasis reviews, vol. 22.Netherlands: Kluwer Academic Publisher; 2003. p. 451-464.

Kawasaki Y, Sato R, Akiyama T. Mutated APC and Asef are involved in the migration of colorectal tumour cells. Nat Cell Biol. 2003; 5: 211-5.

Kemler R. From cadherins to catenins: cytoplasmic protein interactions and regulation of cell adhesion. Trends Genet 1993; 9: 317-321.

Kemp Z, Rowan A, Chambers W et al. CDC4 Mutations occur in a subset of colorectal cancers but are not predicted to cause loss of function and are not associated with chromosomal instability. Cancer Res 2005; 65; 11361-11366.

Khankhanien N, Maulight GM, Russel WO. Prognostic significance of vascular invasion in colorectal cancer of Dukes' B class. Cancer 1977; 39: 1195-1200.

Kim N.K., Ahn J.Y., Song J. et al. Expression of the DNA repair enzyme, N-methylpurine-DNA glycosylase (MPG) in astrocytic tumors. Anticancer Res 2003; 23: 1417-1423.

Kim K, Pang K, Evans M et al. Overexpression of beta-catenin induces apoptosis independent of its transactivation function with LEF-1 or the involvement of major G1 cell cycle regulators. Mol Biol Cell 2000; 11: 3509-3523.

Kinzler K, Vogelstein B. Cancer-susceptibility genes. Gatekeepers and caretakers. Nature 1997; 386: 761-763.

Kirsten ras in Colorectal Cancer Collaborative Group. Kirsten ras mutations in patients with colorectal cancer: the 'RASCAL II' study. Br J Cancer 2001; 85: 692-696.

Kishida M., Koyama S., Kishida S. et al. Axin prevents Wnt-3a-induced accumulation of beta-catenin. Oncogene 1999; 18: 979-985.

Knosel T, Yu Y, Stein U et al. Overexpression of c-erbB-2 protein correlates with chromosomal gain at the c-erbB-2 locus and patient survival in advanced colorectal carcinomas. Clin Exp Metastasis 2002; 19: 401-407.

Knosel T, Schluns K, Stein U,et al. Genetic imbalances with impact on survival in colorectal cancer patients. Histopathology 2003; 43: 323– 31.

Knudson A.G. Hereditary cancer: two hits revisited. J Cancer Res Clin Oncol 1996; 122: 135-140.

Kobayashi M, Honma T, Matsuda Y et al. Nuclear translocation of beta-catenin in colorectal cancer. Br J Cancer 2000; 82: 1689-1693.

Kochhar R, Halling KC, McDonnell S et al. Allelic imbalance and microsatellite instability in resected Duke's D colorectal cancer. Diagn Mol Pathol 1997; 6: 78– 84.

Kokal WA, Gardine RL, Sheibani K, et al. Tumor DNA content in resectable, primary colorectal carcinoma. Ann Surg 1989; 209: 188-193.

Kolligs F, Kolligs B, Hajra K.M. et al. Gamma-catenin is regulated by the APC tumor suppressor and its oncogenic activity is distinct from that of beta-catenin. Genes Dev 2000; 14: 1319-1331.

Kong S., Amos C.I., Luthra R. et al. Effects of cyclin D1 polymorphism on age of onset of hereditary nonpolyposis colorectal cancer. Cancer Res 2000; 60: 249-52.

Korenaga D, Ueo H, Mochida K, et al. Prognostic factors in Japanese patients with colorectal cancer: the significance of large bowel obstruction—univariate and multivariate analyses. J Surg Oncol 1991; 47: 188-192.

Kuismanen S, Holmberg M, Salovaara R et al. Genetic and epigenetic modification of MLH1 accounts for a major share of microsatellite-unstable colorectal cancers. Am J Pathol 2000; 156: 1773-1779.

Laken S, Petersen G, Gruber S. et al. Familial colorectal cancer in Ashkenazim due to a hypermutable tract in APC. Nat Genet 1997; 17: 79-83.

Lamlun H, Ilyas M, Rowan A et al. The type of somatic mutation at APC in familial adenomatous polyposis is determined by the site of the germline mutation: a new facet to Knudson's 'two hit' hypothesis. Nat Med 1999; 5: 1071-1075.

Lane D.P. Cancer. A death in the life of p53. Nature 1993; 362: 786-787.

Lasser P, Mankarios H, Elias D, et al. Etude pronostique uni et multi-factorielle de 400 adenocarcinomes rectaux reseques. J Chir (Paris) 1993; 2: 57-65.

Laurent-Puig P, Olschwang S, Delattre O, et al. Survival and acquired genetic alterations in colorectal cancer. Gastroenterology 1992; 102: 1136-1141.

Le Marchand L, Donlon T, Seifried A. et al. Association of a common polymorphism in the human GH1 gene with colorectal neoplasia. J Natl Cancer Inst 2002; 94: 454-460.

Lengauer C, Kinzler K, Vogelstein B. Genetic instabilities in human cancers. Nature 1998; 396: 643-649.

Lennon GG, Auffray C, Polymeropoulos M et al. The I.M.A.G.E Consortium: An Integrated Molecular Analysis of Genomes and their Expression. Genomics 1996; 33: 151-152.

Lesslie DP, Summy JM, Parikh NU et al. Vascular endothelial growth factor receptor-1 mediates migration of human colorectal carcinoma cells by activation of Src family kinases. Br J Cancer. 2006; 94: 1710-1717.

Li A.W, Murphy P. Expression of alternatively spliced FGF-2 antisense RNA transcripts in the central nervous system: regulation of FGF-2 mRNA translation. Mol Cell Endocrinol 2000; 162: 69-78.

Li J, Yen C, Liaw D. et al. PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. Science 1997; 275: 1943-1947.

Lichter P, Joos S, Bentz M et al. Comparative genomic hybridization: uses and limitations. Semin Hematol 2000; 37: 348-357.

Loeb LA. Mutator phenotype may be required for multistage carcinogenesis. Cancer Res 1991; 51: 3075-3079.

Leach F, Nicolaides N, Papadopoulos N et al. Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. Cell 1993; 75: 1215-1225.

Lichtenstein P, Holm N, Verkasalo P et al. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. N Engl J Med 2000; 343: 78-85.

Lucito R, Healy J, Alexander J. et al. Representational oligonucleotide microarray analysis: a high-resolution method to detect genome copy number variation. Genome Res 2003; 13: 2291-2305.

Lugli A, Zlobec I, Minoo P et al. Prognostic significance of wnt signalling pathway molecules APC, beta-catenin and E-cadherin in colorectal cancer- a tissue microarray-based analysis. Histopathology 2007; 50: 453-464.

McArdle CS, Hole DJ. Outcome following surgery for colorectal cancer: analysis by hospital after adjustment for case-mix and deprivation. Br J Cancer 2002; 86: 331-335.

MacFarlane JK, Ryall RD, Heald RJ. Mesorectal excision for rectal cancer. Lancet 1993; 341: 457-460.

Mandel JS, Church TR, Bond JH et al. The effect of faecal occult-blood screening on the incidence of colorectal cancer. N Engl J Med 2000; 343: 1603-1607.

Mann B, Gelos M, Siedow A et al. Target genes of beta-catenin-T cell-factor/lymphoid-enhancer-factor signaling in human colorectal carcinomas. Proc Natl Acad Sci U S A 1999; 96: 1603-1608.

Mantripragada K, Buckley B, Jarbo C et al. Development of NF2 gene specific, strictly sequence defined diagnostic microarray for deletion detection. J Mol Med 2003; 81: 443-451.

Marikawa Y, Elinson R. Beta-TrCP is a negative regulator of Wnt/beta-catenin signaling pathway and dorsal axis formation in Xenopus embryos. Mech Dev 1998; 77: 75-80.

Maruyama K, Ochiai A, Akimoto S et al. Cytoplasmic beta-catenin accumulation as a predictor of hematogenous metastasis in human colorectal cancer. Oncology 2000; 59: 302-309.

Massion P, Kuo W, Stokoe D, et al. Genomic copy number analysis of non-small cell lung cancer using array comparative genomic hybridization: implications of the phosphatidylinositol 3-kinase pathway. Cancer Res 2002; 62: 3636-3640.

Maughan N, Quirke P. Pathology- a molecular prognostic approach. British Medical Bulletin 2002; 64: 59-74.

May P, May E. Twenty years of p53 research: structural and functional aspects of the p53 protein. Oncogene 1999; 18: 7621-7636.

McLeod HL, Murray GI. Tumour markers of prognosis in colorectal cancer. Br J Cancer 1999; 79: 191-203.

Mehta KR, Nakao K, Zuraek MB, et al. Fractional genomic alteration detected by array-based comparative genomic hybridization independently predicts survival after hepatic resection for metastatic colorectal cancer, Clin Cancer Res 2005; 11: 1791–1797.

Meraldi P, Honda R, Nigg EA. Aurora kinases link chromosome segregation and cell division to cancer susceptibility. Curr Opin Genet Dev 2004; 14: 29-36.

Meterissian SH, Kontogiannea M, Al-Sowaidi M et al. Bcl-2 Is a Useful Prognostic Marker in Dukes' B Colon Cancer. Annals of Surgical Oncology 2001; 8: 533-537.

Michelassi F, Vannucci L, Ayala JJ et al. Local recurrence after curative resection of colorectal adenocarcinoma. Surgery 1990; 108: 787-793.

Middeldorp A, Jagmohan-Changur S, Helmer Q et al. A procedure for the detection of linkage with high density SNP arrays in a large pedigree with colorectal cancer. BMC Cancer 2007; 7: 6.

Mills KD, Ferguson DO, Alt FW. The role of DNA breaks in genomic instability and tumorigenesis. Immunol Rev 2003; 194: 77-95.

Minsky BD, Mies C, Rich TA et al. Lymphatic vessels invasion is an independent prognostic factor for survival in colorectal cancer. IntJ Radiat Oncol Biol Phys 1989 ; 17: 311-318.

Mitin N, Betts L, Yohe ME et al. Release of autoinhibition of ASEF by APC leads to CDC42 activation and tumor suppression. Nat Struct Mol Biol. 2007; 14: 814-823.

Modrich P. Mechanisms and biological effects of mismatch repair. Annu Rev Genet 1991; 25229-25253.

Moinova H, Chen W, Shen L et al. HLTF gene silencing in human colon cancer. Proc Natl Acad Sci U S A 2002; 99: 4562-4567.

Molenaar M., van de Wetering M., Oosterwegel M. et al. XTcf-3 transcription factor mediates beta-catenin-induced axis formation in Xenopus embryos. Cell 1996; 86: 391-399.

Moolgavkar S, Knudson A. Mutation and cancer: a model for human carcinogenesis. J Natl Cancer Inst 1981; 66: 1037-1052

Morin P, Sparks A, Korinek V. et al. Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. Science 1997; 275: 1787-1790.

Morris EJ, Maughan NJ, Forman D, Quirke P. Who to treat with adjuvant therapy in Dukes B/stage II colorectal cancer? The need for high quality pathology. Gut 2007; 56: 1419-1425.

Morson BC. Factors influencing the prognosis of early cancer of the rectum. Pros R Soc Med 1966; 59: 607-608.

Muto T, Bussey HJ, Morson BC. The evolution of cancer of the colon and rectum. Cancer 1975; 36: 2251-2270.

Nagase H, Miyoshi Y, Horii A et al. Correlation between the location of germ-line mutations in the APC gene and the number of colorectal polyps in familial adenomatous polyposis patients. Cancer Res 1992; 52: 4055-4057.

Nakao K, Mehta KR, Fridlyand J, et al. High-resolution analysis of DNA copy number alterations in colorectal cancer by array-based comparative genomic hybridization. Carcinogenesis 2004; 42: 1345-1357.

Nicolaides N, Papadopoulos N, Liu B. et al. Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. Nature 1994; 371: 75-80.

Newland RC, Dent OF, Lyttle MN et al. Pathologic determinants of survival associated with colorectal cancer with lymph node metastases. Cancer 1994; 73: 2076-2082.

Notterman DA, Alon U, Sierk AJ et al. Transcriptional gene expression profiles of colorectal adenoma, adenocarcinoma, and normal tissue examined by oligonucleotide arrays. Cancer Res 2001; 61: 3124-3130.

Ogunbiyi OA, Goodfellow PJ, Gagliardi G et al. Prognostic value of chromosome 1p allelic loss in colon cancer. Gastroenterology 1997; 113: 761-766.

O'Hagan RC, Chang S, Maser RS et al. Telomere dysfunction provokes regional amplification and deletion in cancer genomes. Cancer Cell 2002; 2: 149-155.

Palmqvist R, Hallmans G, Rinaldi S et al. Plasma insulin-like growth factor 1, insulin-like growth factor binding protein 3, and risk of colorectal cancer: a prospective study in northern Sweden. Gut 2002; 50: 642-646.

Papamichael D. Prognostic role of angiogenesis in colorectal cancer. Anticancer Res 2001; 21:4349-4354.

Peltomaki P. Deficient DNA mismatch repair: a common etiologic factor for colon cancer. Hum Mol Genet 2001; 10: 735-740.

Peltomaki P. Microsatellite instability and hereditary non-polyposis colon cancer. J Pathol 1995; 176: 329-330.

Peltomaki P, Aaltonen L, Sistonen P et al. Genetic mapping of a locus predisposing to human colorectal cancer. Science 1993; 260: 810-812.

Petersen VC, Baxter KJ, Love SB et al. Identification of objective pathological prognostic determinants and models for prognosis in Dukes B colon cancer. Gut 2002; 51: 65-69.

Plesec TP and Hunt JL. Kras mutation testing in colorectal cancer. Adv Anat Pathol 2009; 16: 196-203.

Polakis P. Wnt signaling and cancer. Genes Dev 2000; 14: 1837-1851.

Pollack J, Sorlie T, Perou C. 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: 12963-12968.

Ponz de Leon M, Sant M, Micheli A, et al. Clinical and pathologic prognostic indicators in colorectal cancer. Cancer 1992; 69: 626-635.

Popat S, Houlston RS. A systematic review and meta-analysis of the relationship between chromosome 18q genotype, DCC status and colorectal cancer prognosis. Eur J Cancer. 2005; 41: 2060-2070.

Popat S, Hubner R, Houlston RS. Systemic review of microsatellite instability and colorectal cancer prognosis. J Clin Oncol 2005; 20: 609-618.

Quirke P, Dixon MF, Clayden AD et al. Prognostic significance of DNA aneuploidy and cell proliferation in rectal carcinomas. J Pathol 1987; 151: 285-291.

Quirke P, Fozard JB, Dixon M et al. DNA aneuploidy in colorectal adenomas. Br J Cancer 1986; 53: 477-481.

Rajagopalan H, Bardelli A, Lengauer C. et al. Tumorigenesis: RAF/RAS oncogenes and mismatch-repair status. Nature 2002; 418: 934.

Rajagopalan H, Jallepalli PV, Rago C, et al. Inactivation of hCDC4 can cause chromosomal instability. Nature 2004; 428: 77–81.

Ratto C, Sofo L, Ippoliti M et al. Prognostic factors in colorectal cancer. Dis Colon Rectum 1998; 41:1033-1049.

Richards F, McKee S, Rajpar M. et al. Germline E-cadherin gene (CDH1) mutations predispose to familial gastric cancer and colorectal cancer. Hum Mol Genet 1999; 8: 607-610.

Risques RA, Moreno V, Ribas M et al. Genetic pathways and genome-wide determinants of clinical outcome in colorectal cancer. Cancer Res 2003; 63 : 7206– 7214.

Roca F, Mauro LV, Morandi A et al. Prognostic value of E-cadherin, beta-catenin, MMPs (7 and 9), and TIMPs (1 and 2) in patients with colorectal carcinoma. J Surg Oncol 2006; 93: 151-160.

Rodrigues N, Rowan A, Smith M et al. p53 mutations in colorectal cancer. Proc Natl Acad Sci U S A 1990; 87: 7555-7559.

Rong S, Segal S, Anver m et al. Invasiveness and metastasis of NIH 3T3 cells induced by Met-hepatocyte growth factor/scatter factor autocrine stimulation. Proc Natl Acad Sci U S A 1994; 91: 4731-4735.

Rooney PH, Boonsong A, McKay JA et al. Colorectal cancer genomics: evidence for multiple genotypes which influence survival. Br J Cancer 2001; 85: 1492– 1498.

Roose J, Molenaar M, Peterson J et al. The Xenopus Wnt effector XTcf-3 interacts with Groucho-related transcriptional repressors. Nature 1998; 395: 608-612.

Rowe D.W., Moen R.C., Davidson J.M. et al. Correlation of procollagen mRNA levels in normal and transformed chick embryo fibroblasts with different rates of procollagen synthesis. Biochemistry 1978; 17: 1581-1590.

Roylance R, Gorman P, Papior T et al. A comprehensive study of chromosome 16q in invasive and lobular breast carcinoma using array CGH. Oncogene 2006; 25: 6544-6553.

Rozek LS, Lipkin SM, Fearon ER et al. CDX2 polymorphisms, RNA expression, and risk of colorectal cancer. Cancer Res 2005; 65: 5488-5492.

Rudolph KL, Milliard M, Bosenberg MW, DePinho RA. Telomere dysfunction and evolution of intestinal carcinoma in mice and humans. Nat Genet 2001; 28:155-159.

Sahai E, Garcia-Medina R, Pouysségur J et al. Smurf1 regulates tumor cell plasticity and motility through degradation of RhoA leading to localized inhibition of contractility. J Cell Biol 2007; 176: 35-42

Samowitz W.S., Powers M.D., Spirio L.N. et al. Beta-catenin mutations are more frequent in small colorectal adenomas than in larger adenomas and invasive carcinomas. Cancer Res 1999; 59: 1442-1444.

Sandmeyer S., Smith R., Kiehn d et al. Correlation of collagen synthesis and procollagen messenger RNA levels with transformation in rat embryo fibroblasts. Cancer Res 1981; 41: 830-838.

Sanger W.G. Chromosomes and neoplasia. Ann Clin Lab Sci 1983; 13: 366-370.

Saunders W. Centrosomal amplification and spindle multipolarity in cancer cells. Seminars in Cancer Biology 2005; 15: 25-32.

Schmitzd-Moormann P, Himmelmann GW, Baum U et al. Morphological predictors of survival in colorectal carcinoma: univariate and multivariate analysis. J Cancer Res Clin Oncol 1987; 113: 586-92.

Schutte B, Reynders MM, Wiggers T et al. Retrospective analysis of the prognostic significance of DNA content and proliferative activity in large bowel carcinoma. Cancer Res 1987; 47: 5494-5496.

Scott CA, Desinan L, Avellini C, et al. DNA index shift with disease progression in colorectal adenocarcinoma: a morphological and flow cytometric study. Hum Pathol 1998; 29: 482–90.

Scott NA, Wieand HS, Moertel CG et al. Colorectal cancer: Dukes stage, tumor site, preoperative plasma CEA level, and patient prognosis related to tumor DNA ploidy pattern. Arch Surg 1987; 122:1375-1379.

Segditsas S, Rowan AJ, Howarth K, Jones A, Leedham S, Wright NA, Gorman P, Chambers W et al. APC and the three-hit hypothesis. Oncogene. 2009; 28 :146-155.

Selby JV, Friedman GD, Quesenberry CP et al. A case-control study, of screening sigmoidoscopy and mortality from colorectal cancer. N Engl J Med 1992; 326: 653-657.

Sengupta P, Smith E, Kim K et al. DNA hypermethylation near the transcription start site of collagen alpha2(I) gene occurs in both cancer cell lines and primary colorectal cancers. Cancer Res 2003; 63: 1789-1797.

Shepherd NA, Saraga E-P, Love SB et al. Prognostic factors in colonic cancer. Histopathology 1989; 14: 613-620.

Shen L, Kondo Y, Hamilton et al. P14 methylation in human colon cancer is associated with microsatellite instability and wild-type p53. Gastroenterology 2003; 124: 626-633.

Shih I, Wei Z, Goodman SN et al. Evidence that genetic instability occurs at an early stage of colorectal tumorigenesis. Cancer Res 2001; 61: 818-822.

Shtutman M, Zhurinsky J, Simcha I et al. The cyclin D1 gene is a target of the beta-catenin/LEF-1 pathway. Proc Natl Acad Sci U S A 1999; 96: 5522-5527.

Sieber OM, Heinimann K, Gorman P et al. Analysis of chromosomal instability in human colorectal adenomas with two mutational hits at APC. Proc Natl Acad Sci USA. 2002; 99: 16910-16915.

Sivagnanasundaram S, Islam I, Talbot I. The homeobox gene CDX2 in colorectal carcinoma: a genetic analysis. Br J Cancer 2001; 84: 218–225.

Smith G., Carey F.A., Beattie J. et al. Mutations in APC, Kirsten-ras, and p53--alternative genetic pathways to colorectal cancer. Proc Natl Acad Sci U S A 2002; 99: 9433-9438.

Somasundaram K. Tumor suppressor p53: regulation and function. Front Biosci 2000; 5D: 424-437.

Soong R, Grieu F, Robbins P et al. P53 alterations are associated with improved prognosis in distal colonic carcinomas. Clin Cancer Res 1997; 3: 1405-1411.

Sparks A, Morin P, Vogelstein B et al. Mutational analysis of the APC/beta-catenin/Tcf pathway in colorectal cancer. Cancer Res 1998; 58: 1130-1134.

Sparkes R, Sparkes M, Wilson M et al. Regional assignment of genes for human esterase D and retinoblastoma to chromosome band 13q14. Science 1980; 208: 1042-1044.

Sparkes R, Murphree A, Lingua R et al. Gene for hereditary retinoblastoma assigned to human chromosome 13 by linkage to esterase D. Science 1983; 219: 971-973.

Spirio L, Olschwang S, Groden J et al. Alleles of the APC gene: an attenuated form of familial polyposis. Cell 1993; 75: 951-957.

Spratt JS Jr, Spjut HJ. Prevalence and prognosis of individual clinical and pathologic variables associated with colorectal carcinoma. Cancer 1967; 20: 1976-1985.

Steele R. Colonic cancer. Colorectal surgery. Third Edition 2005, Elsevier Saunders.

Steinberg SM, Barkin JS, Kaplan RS, Stablein DM. Prognostic indicators of colon tumors. Cancer 1986; 57: 1866-1870.

Suh E, Ha C, Rankin E et al. DNA methylation down-regulates CDX1 gene expression in colorectal cancer cell lines. J Biol Chem 2002; 277: 35795-35800.

Suzuki H, Gabrielson E, Chen W et al. A genomic screen for genes upregulated by demethylation and histone deacetylase inhibition in human colorectal cancer. Nat Genet 2002; 31: 141-149.

Suzuki H, Watkins D, Jair K et al. Epigenetic inactivation of SFRP genes allows constitutive WNT signalling in colorectal cancer. Nat Genet 2004; 36: 417-422.

Takagi Y, Kohmura H, Futamura M. et al. Somatic alterations of the DPC4 gene in human colorectal cancers in vivo. Gastroenterology 1996; 111: 1369-1372.

Takano Y, Shiota G, Kawasaki H. Analysis of genomic imprinting of insulin-like growth factor 2 in colorectal cancer. Oncology 2000; 59: 210-216.

Takenoshita S, Tani M, Mogi A et al. Mutation analysis of the SMAD2 gene in human colon cancers using genomic DNA and intron primers. Carcinogenesis 1998; 19: 803-807.

Tepper JE, O'Connell MJ, Niedzwiecki D et al. Impact of number of nodes retrieved on outcome in patients with rectal cancer. J Clin Oncol 2001; 19: 157-163.

Tetsu O, McCormick F. Beta-catenin regulates expression of cyclin D1 in colon carcinoma cells. Nature 1999; 398: 422-426.

Thiagalingam S, Lengauer C, Leach F et al. Evaluation of candidate tumour suppressor genes on chromosome 18 in colorectal cancers. Nat Genet 1996; 13: 343-346.

Thiagalingam S, Laken S, Willson J et al. Mechanisms underlying losses of heterozygosity in human colorectal cancers. Proc Natl acad Sci USA 2001; 98: 2698-2702.

Thibodeau S.N, French A, Cunningham J et al. Microsatellite instability in colorectal cancer: different mutator phenotypes and the principal involvement of hMLH1. Cancer Res 1998; 58: 1713-1718.

Tighe A, Johnson VL, Taylor SS. Truncating APC mutations have dominant effects on proliferation, spindle checkpoint control, survival and chromosomal stability. J Cell Sci 2004 ; 117: 6339-6353.

Tischfield JA, Shao C. Somatic recombination redux. Nat Genet 2003; 33: 5-6.

Tomita S, Kawamata H, Imura J. et al. Frequent polymorphism of peroxisome proliferator activated receptor gamma gene in colorectal cancer containing wild-type K-ras gene. Int J Mol Med 2002; 9: 485-488.

Tomlinson I, Bodmer W. Selection, the mutation rate and cancer: ensuring the tail does not wag the dog. Nat Med 1999; 5: 11-12.

Tomlinson I, Webb E, Carvajal-Carmona L et al. A genome wide association scan of tag SNPs identifies a susceptibility variant for CRC at 8q24.21. Gastroenterology 2007; 113: 2063-2065.

Toyota M, Ahuja N, Ohe-Toyota M et al. CpG island methylator phenotype in colorectal cancer. Proc Natl Acad Sci USA 1999; 96: 8681-8686.

Veigl M, Kasturi L, Olechnowicz J. et al. Biallelic inactivation of hMLH1 by epigenetic gene silencing, a novel mechanism causing human MSI cancers. Proc Natl Acad Sci U S A 1998; 95: 8698-8702.

Vogelstein B, Fearon ER, Hamilton SR et al. Genetic alterations during colorectal-tumour development. N Engl J Med 1988; 319: 525-532.

Walther A, Houston R, Tomlinson I. Association between chromosomal instability and prognosis in colorectal cancer: a metaanalysis. Gut 2008; 57: 941-950.

Wang Y, Jatkoe T, Zhang Y et al. Gene expression profiles and molecular markers to predict recurrence of Dukes B colon cancer. J Clin Oncol 2004; 22: 1564-1571.

Watson P, Lynch H. Extracolonic cancer in hereditary nonpolyposis colorectal cancer. Cancer 1993; 71: 677-685.

Weber JC, Schneider A, Rohr S, et al. Analysis of allelic imbalance in patients with colorectal cancer according to stage and presence of synchronous liver metastases. Ann Surg 2001; 234: 795– 802.

Weiss M.M., Kuipers E.J., Postma C et al. Genomic profiling of gastric cancer predicts lymph node status and survival. Oncogene 2003; 22: 1872-1879.

Wheeler JM. Epigenetics, mismatch repair genes and colorectal cancer. Ann R Coll Surg Engl. 2005; 87: 15-20.

Wheeler J, Beck N, Kim H. 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: 10296-102301.

Wied U, Nilsson T, Knudsen JB, Sprechler M, Johansen A. Postoperative survival of patients with potentially curable cancer of the colon. Dis Colon Rectum 1985; 28: 333-335.

Wiesner G, Daley D, Lewis S et al. A subset of familial colorectal neoplasia kindreds linked to chromosome 9q22.2-31.2. Proc Natl Acad Sci U S A 2003; 100: 12961-12965.

Wiggers T, Arends JW, Schutte B et al. A multivariate analysis of pathologic prognostic indicators in large bowel cancer. Cancer 1988; 61: 386-395.

Wilhelm M, Veltman J, Olshen A et al. Array-based comparative genomic hybridization for the differential diagnosis of renal cell cancer. Cancer Res 2002; 62: 957-960.

Winawer SJ, Zauber A, Diaz B. Temporal sequence of evolving colorectal cancer from the normal colon (abstract). Gastrointest Endosc 1987; 33:167.

Witek ME, Nielsen K, Walters R et al. The putative tumor suppressor Cdx2 is overexpressed by human colorectal adenocarcinomas. Clin Cancer Res. 2005; 11: 8549-8556.

Wong NA, Morris RG, McCondichie A et al. Cyclin D1 overexpression in colorectal carcinoma in vivo is dependent on beta-catenin protein dysregulation, but not k-ras mutation. J Pathol 2002; 197: 128-135.

Woodford-Ritchens KL, Rowan AJ, Bodmer WF et al. SMAD 4 mutations in colorectal cancer probably occur before chromosomal instability, but after divergence of the microsatellite instability pathway. Proc Natl Acad Sci USA 2001; 98: 9719-9723.

Woodford-Richens KL, Halford S, Rowan A, et al. CDX2 mutations do not account for juvenile polyposis or Peutz-Jeghers syndrome and

occur infrequently in sporadic colorectal cancers. Br J Cancer 2001; 84: 1314–1316.

Wotmark N, Wieand HS, Rockette HE et al. The prognostic significance of tumor and location and bowel obstruction in Dukes B and C colorectal cancer: findings from the NSABP clinical trials. Ann Surg 1983; 198 :743-752.

Vogel F. Genetics of retinoblastoma. Hum Genet 1979; 52: 1-54.

Vogelstein B, Fearon ER, Kern SE et al. Allelotype of colorectal carcinomas. Science 1989; 244: 207-212.

Xu L, Corcoran R, Welsh J et al. WISP-1 is a Wnt-1- and beta-catenin-responsive oncogene. Genes Dev 2000; 14: 585-595.

Yamada T, Takaoka A, Naishiro Y et al. Transactivation of the multidrug resistance 1 gene by T-cell factor 4/beta-catenin complex in early colorectal carcinogenesis. Cancer Res 2000; 60: 4761-4766.

Yamamoto H, Min Y, Itoh F. et al. Differential involvement of the hypermethylator phenotype in hereditary and sporadic colorectal cancers with high-frequency microsatellite instability. Genes Chromosomes Cancer 2002; 33: 322-325.

Yang J, Yu Y, Markovic B et al. Overexpression of c-erbB-2 mRNA and/or c-neu oncoprotein is a predictor for metastases from colorectal cancer. Anticancer Res 1997; 17: 1023-1026.

Yang Z, Imoto I, Fukuda Y et al. Identification of a novel gene, GASC1, within an amplicon at 9p23-24 frequently detected in esophageal cancer cell lines. Cancer Res 2000; 60: 4735-4739.

Yin J, Kong D, Wang S et al. Mutation of hMSH3 and hMSH6 mismatch repair genes in genetically unstable human colorectal and gastric carcinomas. Hum Mutat 1997; 10: 474-478.

Yamaguchi T, Bando H, Mori T et al. Overexpression of soluble vascular endothelial growth factor receptor 1 in colorectal cancer: Association with progression and prognosis. Cancer Sci. 2007; 98: 405-10

Ylstra B, van den IJessel P, Meijer G. Bac to the future! Or oligonucleotides a perspective for micro array CGH (array CGH). Nucleic Acids Research 2006; 34: 445-450.

Yuen S, Davies H, Chan T et al. Similarity of the phenotypic patterns associated with BRAF and KRAS mutations in colorectal neoplasia. Cancer Res 2002; 62: 6451-6455.

Zandvliet DW, Hanby AM, Austin CA et al. Analysis of foetal expression sites of human type II DNA topoisomerase alpha and beta mRNAs by in situ hybridisation. Biochim Biophys Acta. 1996; 1307: 239-247.

Zhao J, Roth J, Bode-Lesniewska B et al. Combined comparative genomic hybridization and genomic microarray for detection of gene amplifications in pulmonary artery intimal sarcomas and adrenocortical tumors. Genes Chromosomes Cancer 2002; 34: 48-57.

Zhou W, Goodman SN, Galizia G et al. Counting the alleles to predict recurrence of early stage colorectal cancers. Lancet 2002; 359: 219-225.

12. Appendices.

1 . Sequences, and reaction conditions for the PCR primers used in PCR experiments.

Name	Primer sequence	Product	[MgCl]	Annealing temp
D5S346 F D5S346 R	[6-FAM]actcactctagtgataaatcggg agcagataagacagtattactagtt	96-122 bp	2.5 uM	55°C
BAT26 F BAT26 R	[6-FAM]tgactacttttgacttcagcc aaccattcaacatttttaaccc	80-100 bp	2.5 uM	60°C

2. MSI and LOH results compared for 5 year survival groups

Survival	>5 year	<5 year	p
MSI (%)	8/43(18)	6/36 (16.7)	1.0
BAT 26 (%)	8/43(18)	3/36 (8.8)	0.31
D5S346 LOH	20	8	
MISMATCH	1	2	
MSI	0	3	
NI	6	4	
NL	16	17	
% INFORMATIVE	55.6	32	0.17

3. Comparison of ploidy for survival groups.

Survival	>5 years	<5 years	p
Aneuploid (%)	10(43.5)	15(55.6)	0.57
Tetraploid (%)	1(4.3)	1(3.7)	1.0
Diploid (%)	12(52.2)	11(40.7)	0.57

4. Summary of genomic gain/loss by chromosome for the different survival/MSI subgroups.

	5yr+/MSI-	5yr+/MSI+	5yr+	5yr-/MSI-	5yr-/MSI+	5yr-
Chr1	0	0	0	0	0	0
Chr arm	1	0	1	1	1	2
Large gain	6	0	6	4	0	4
Small gain <10Mb	3	0	3	3	0	3
Gain++	1	0	1	2	0	2
Chr2	5	0	5	4	0	4
Chr arm	2	0	2	0	0	0
Large gain	4	0	4	2	0	2
Small gain <10Mb	0	0	0	2	0	2
Gain++	2	0	2	2	0	2
Chr3	2	0	2	1	0	1
Chr arm	1	0	1	1	0	1
Large gain	4	0	4	0	0	0
Small gain <10Mb	1	0	1	1	0	1
Gain++	1	1	2	0	0	0
Chr4	2	0	2	0	0	0
Chr arm	1	0	1	0	0	0
Large gain	1	0	1	1	0	1
Small gain <10Mb	0	0	0	0	0	0
Gain++	0	0	0	1	0	1

Chr5	3	1	4	0	0	0
Chr arm	5	2	7	2	0	2
Large gain	1	0	1	3	0	3
Small gain <10Mb	0	0	0	1	0	1
Gain++	1	0	1	0	0	0
Chr6	6	0	6	5	2	7
Chr arm	1	0	1	2	0	2
Large gain	1	0	1	3	1	4
Small gain <10Mb	0	0	0	1	0	1
Gain++	0	0	0	2	0	2
Chr7	16	2	18	11	1	12
Chr arm	2	0	2	1	0	1
Large gain	1	1	2	0	0	0
Small gain <10Mb	0	1	1	3	0	3
Gain++	1	2	3	1	0	1
Chr8	3	0	3	2	1	3
Chr arm	17	2	19	14	3	17
Large gain	0	0	0	2	0	2
Small gain <10Mb	0	0	0	1	0	1
Gain++	3	1	4	1	1	2
Chr9	4	0	4	6	0	6
Chr arm	0	0	0	0	0	0
Large gain	4	1	5	1	0	1
Small gain <10Mb	1	0	1	1	0	1
Gain++	0	0	0	0	0	0
Chr10	6	1	7	9	1	10
Chr arm	1	2	3	0	0	0
Large gain	2	0	2	1	0	1
Small gain <10Mb	1	1	2	2	0	2
Gain++	1	1	2	2	0	2
Chr11	6	0	6	3	0	3
Chr arm	3	0	3	1	0	1
Large gain	3	1	4	4	1	5
Small gain <10Mb	0	0	0	0	0	0
Gain++	3	0	3	0	0	0

Chr12	4	0	4	4	0	4
Chr arm	1	0	1	2	0	2
Large gain	1	0	1	1	1	2
Small gain <10Mb	1	0	1	2	0	2
Gain++	0	0	0	1	1	2
Chr13	26	2	28	20	2	22
Chr arm	0	0	0	1	0	1
Large gain	0	0	0	0	0	0
Small gain <10Mb	1	0	1	1	0	1
Gain++	0	0	0	6	2	8
Chr14	1	0	1	2	0	2
Chr arm	0	0	0	0	0	0
Large gain	0	0	0	0	0	0
Small gain <10Mb	0	0	0	0	0	0
Gain++	1	0	1	0	0	0
Chr15	2	0	2	2	0	2
Chr arm	0	1	1	1	0	1
Large gain	0	0	0	0	0	0
Small gain <10Mb	1	0	0	0	0	0
Gain++	0	0	0	0	0	0
Chr16	6	0	6	0	1	1
Chr arm	2	0	2	1	0	1
Large gain	1	0	1	1	0	1
Small gain <10Mb	0	2	2	0	0	0
Gain++	0	1	1	0	0	0
Chr17	0	0	0	0	0	0
Chr arm	4	3	7	6	0	6
Large gain	2	0	2	0	0	0
Small gain <10Mb	1	1	2	0	0	0
Gain++	0	1	1	0	0	0
Chr18	1	1	2	0	1	1
Chr arm	1	2	3	1	1	2
Large gain	0	0	0	1	0	1
Small gain <10Mb	0	1	1	0	0	0
Gain++	0	1	1	1	0	1

Chr19	6	0	6	8	1	9
Chr arm	1	0	1	0	0	0
Large gain	1	0	1	1	0	1
Small gain <10Mb	0	0	0	0	0	0
Gain++	0	0	0	1	0	1
Chr20	16	3	19	17	1	18
Chr arm	10	0	10	7	1	8
Large gain	2	0	2	0	1	1
Small gain <10Mb	1	0	1	0	2	2
Gain++	11	1	12	10	2	12
Chr21	8	2	10	7	2	9
Chr arm	0	0	0	0	0	0
Large gain	0	0	0	0	0	0
Small gain <10Mb	0	0	0	0	0	0
Gain++	0	0	0	0	0	0
Chr22	2	0	2	1	0	1
Chr arm	0	0	0	0	0	0
Large gain	1	0	1	0	0	0
Small gain <10Mb	0	0	0	0	0	0
Gain++	0	0	0	0	0	0

Loss

Chr 1	8	2	10	9	2	11
Chr arm	4	0	4	5	1	6
Large loss	17	0	17	10	2	12
Small loss <10Mb	0	0	0	0	0	0
Loss++	1	0	1	1	0	1
Chr 2	1	0	1	0	0	0
Chr arm	0	0	0	0	0	0
Large loss	3	0	3	1	0	1
Small loss <10Mb	0	0	0	2	0	2
Loss++	0	0	0	0	0	0
Chr 3	0	0	0	0	0	0

Chr arm	0	0	0	3	0	0
Large loss	4	1	5	2	0	2
Small loss	0	0	0	0	0	0
<10Mb						
Loss++	0	0	0	0	0	0
Chr 4	3	0	3	8	0	8
Chr arm	1	0	1	1	0	1
Large loss	4	0	4	0	0	0
Small loss	0	0	0	0	0	0
<10Mb						
Loss++	0	0	0	0	0	0
Chr 5	1	0	1	2	1	3
Chr arm	3	0	3	2	0	2
Large loss	0	0	0	3	0	3
Small loss	0	0	0	0	0	0
<10Mb						
Loss++	0	0	0	0	0	0
Chr 6	6	1	7	0	0	0
Chr arm	1	0	1	3	0	3
Large loss	3	1	4	3	1	4
Small loss	0	0	0	0	0	0
<10Mb						
Loss++	0	0	0	0	0	0
Chr 7	1	1	2	0	0	0
Chr arm	1	0	1	2	0	2
Large loss	3	0	3	0	0	0
Small loss	1	0	1	0	0	0
<10Mb						
Loss++	0	0	0	0	0	0
Chr 8	1	1	2	0	0	0
Chr arm	15	1	16	14	3	17
Large loss	0	0	0	1	0	1
Small loss	0	0	0	0	0	0
<10Mb						
Loss++	0	0	0	0	0	0
Chr 9	8	1	9	7	2	9
Chr arm	0	0	0	1	1	2
Large loss	2	0	2	0	1	1

Small loss	0	0	0	1	0	1
<10Mb						
Loss++	0	0	0	0	0	0
Chr 10	5	0	5	1	0	1
Chr arm	1	2	3	0	0	0
Large loss	2	0	2	0	0	0
Small loss	1	0	1	0	0	0
<10Mb						
Loss++	0	0	0	0	0	0
Chr 11	0	0	0	0	0	0
Chr arm	3	0	3	1	0	1
Large loss	3	1	4	3	0	3
Small loss	0	0	0	2	0	2
<10Mb						
Loss++	0	0	0	1	0	1
Chr 12	2	0	2	1	0	1
Chr arm	1	0	1	1	0	1
Large loss	0	0	0	1	1	2
Small loss	0	0	0	1	0	1
<10Mb						
Loss++	0	0	0	0	0	0
Chr 13	3	1	4	3	1	4
Chr arm	0	0	0	0	0	0
Large loss	1	1	2	0	0	0
Small loss	0	0	0	0	0	0
<10Mb						
Loss++	0	0	0	0	0	0
Chr 14	12	0	12	12	1	13
Chr arm	0	0	0	0	0	0
Large loss	1	0	1	0	0	0
Small loss	0	0	0	0	0	0
<10Mb						
Loss++	0	0	0	0	0	0
Chr 15	13	0	13	6	1	7
Chr arm	0	0	0	0	0	0
Large loss	1	0	1	1	0	1
Small loss	0	0	0	0	0	0
<10Mb						
Loss++	0	0	0	0	0	0

Chr 16	3	0	3	2	1	3
Chr arm	0	0	0	1	0	1
Large loss	1	0	1	0	1	1
Small loss	0	0	0	0	0	0
Loss++	0	0	0	0	0	0
Chr 17	7	0	7	4	0	4
Chr arm	15	2	17	17	2	19
Large loss	1	0	1	0	0	0
Small loss <10Mb	1	0	1	0	0	0
Loss++	0	0	0	0	0	0
Chr 18	22	0	22	18	1	19
Chr arm	1	1	2	1	2	3
Large loss	1	0	1	1	0	1
Small loss <10Mb	0	0	0	0	0	0
Loss++	0	0	0	0	0	0
Chr 19	8	1	9	9	3	12
Chr arm	0	0	0	0	0	0
Large loss	0	0	0	0	0	0
Small loss <10Mb	0	0	0	0	0	0
Loss++	0	0	0	0	0	0
Chr 20	2	3	5	1	1	2
Chr arm	7	0	7	6	0	6
Large loss	1	0	1	0	0	0
Small loss <10Mb	0	0	0	0	0	0
Loss++	0	0	0	0	0	0
Chr 21	12	2	14	8	0	8
Chr arm	0	0	0	0	0	0
Large loss	1	0	1	0	0	0
Small loss <10Mb	0	0	0	0	0	0
Loss++	0	0	0	0	0	0
Chr 22	22	3	25	20	4	24
Chr arm	0	0	0	0	0	0
Large loss	1	0	1	0	0	0
Small loss <10Mb	0	0	0	0	0	0

Loss++	0	0	0	0	0	0

5. Comparison of gains and losses by survival/MSI groups

>5 year survival versus < 5 year survival

Chromosome gain	**>5(43)**	**%**	**<5(36)**	**%**	**p**	**Sig'**
10	7	16.3	10	27.8	0.27	
16	6	14.0	1	2.8	0.11	
19	6	14.0	9	25.0	0.25	
p arm section gain > 10Mb						
5p	11	25.6	3	8.3	0.07	
6p	8	18.6	12	33.3	0.19	
12p	4	9.3	7	19.4	0.2	
16p	9	20.9	2	5.6	0.06	
q arm section gain > 10Mb						
6q	6	14	9	25	0.27	
9q	4	9.3	7	19.4	0.2	
16q	7	16.3	2	5.6	0.17	
Chromosome loss						
4	3	7.0	8	22.2	0.1	
6	7	16.3	0	0.0	0.014	*
14	12	27.9	13	36.1	0.47	
19	9	20.9	12	33.3	0.3	
21	14	32.6	8	22.2	0.33	
22	25	58.1	24	66.7	0.49	
p arm section loss >10 Mb						
4p	3	7	9	25	0.03	*
5p	0	0	6	16.7	0.007	*
6p	10	23.3	3	8.3	0.12	
10p	7	16.3	2	5.6	0.17	
20p	14	32.6	7	19.4	0.2	

q arm section loss >10 Mb						
5q	4	9.3	7	19.4	0.2	
19q	11	25.6	13	36.1	0.34	
21q	14	32.6	8	22.2	0.32	
5+M- versus 5-M-						
Chromosomal gain						
10	6	17.1	9	30.0	0.25	
16	6	17.1	0	0.0	0.027	*
p arm section gain >10 Mb						
5p	8	22.9	3	10	0.2	
6p	8	22.9	10	33.3	0.4	
10p	7	20	10	33.3	0.57	
16p	9	25.7	1	3.3	0.016	*
q arm section gain >10 Mb						
9q	3	8.6	7	23.3	0.17	
13q	25	71.4	18	60	0.43	
16q	6	17.1	1	3.3	0.23	
Chromosomal loss						
4	3	8.6	8	26.7	0.09	
6	6	17.1	0	0.0	0.027	*
10	5	14.3	1	3.3	0.21	
p arm section loss >10 Mb						
1p	28	80	21	70	0.4	
3p	2	5.7	5	16.7	0.23	
4p	3	8.6	9	30	0.05	*
5p	0	0	5	16.7	0.017	*
6p	8	22.9	3	10	0.2	
10p	7	20	2	6.7	0.16	
q arm section loss >10 Mb						
15q	15	42.9	9	30	0.31	
5+M+ versus 5-M+						
Chromosome gain						
20	3	37.5	1	16.7	0.58	

p arm section gain >10Mb					
5p	3	37.5	0	0	0.2
6p	0	0	2	33.3	0.16
18p	2	25	0	0	0.47
q arm section gain >10Mb					
6q	0	0	2	33.3	0.16
8q	2	25	3	50	0.58
17q	2	25	0	0	0.47
Chromosome loss					
20	3	37.5	1	16.7	0.58
21	2	25	0	0.0	0.47
22	3	37.5	4	66.7	0.16
p arm section loss >10Mb					
1p	2	25	4	66.7	0.28
6p	2	25	0	0	0.47
8p	2	25	3	50	0.58
18p	2	25	3	50	0.58
19p	1	12.5	3	50	0.24
20p	4	50	0	0	0.08
q arm section loss >10Mb					
13q	2	25	0	0	0.47
16q	0	0	2	33.3	0.16
18q	0	0	2	33.3	0.16
19q	2	25	3	50	0.58
20q	3	37.5	0	0	0.2
21q	2	25	0	0	0.47

6. Gains at specific BAC clone sites

>5yrMSI-		>5yrMSI+		<5yrMSI-		<5yrMSI+	
chr 7	%	**7**	%	**7**	%	**8q**	%
0-42.46	54.3	0-14.7	50	0-2.4	55.2	34.3-37	33.3
42.46-50.3	51.4	15.8-97.1	37.5	3.3-23.5	51.7	38.40.2	33.3
50.3-56.46	48.6	98.5-101	50	24.3-49.6	48.3	40.9-145.7	50
56.46	45.7	101-158.3 end	37.5	50.2-51.4	44.4	**6**	66.7
62.65-70.97	42.85	**8q**		52.3-56.4	41.4	0-54.7	
70.97-76.8	45.7	48.3-65	25	56.4-67.4	37.9	56.7-75.9	33.3
77.5-158.28	48.6	65-142.3	37.5	68.1-158.3	41.4	76.5-170.5	50
8q		144.2-145.7	50	**8q**		**20**	33.3
0-36.17	20	145.7	37.5	40.2	24.2	0-9.2	
36.17-37.14	25.7	**5p**		40.2	27.6	10.4-16.6	16.7
37.14	37.14	0-45.6	37.5	41.2-42.2	34.4	17.8-20.2	50
39.5-41.86	40	45.6-53.3	25	42.2-42.9	44.8	21.5-22.5	16.7
41.86	45.7	**9**		47.5-61.2	48.5	23.5-29.4	33.3
41.86-145.7	60	120.6-129.25	25	61.3-69.9	51.7	30.8-32.8	16.7
13		130-134	37.5	72.2-89.7	58.6	33.8-63.4	33.3
0-22.1	74.3	**10**		90.6-128.4	62	**13**	
23-27.9	77.1	131.9-135	25	128.4-137.5	55.2	0-113.8	33.3
27.9-113.9	74.3	**13**		138.9-145,7	62		
20		0-113.9 end	25	**13**			
0-8.2	54.3	**16**		0-113.8	65.5		
8.2-14.6	57.1	85.6-89.7	25	**20**			
14.6-15.5	60	**17**		0-11.9	58.6		
16.6-23.5	65.7	26-81.2 end	25	12.5-20.2	62.1		
25-26.1	68.6	**20**		21.5-26.1	65.5		
26.1	71.4	0-63.4 end	25	29.4	68.9		
29.4-48.7	74.28	**21**		30.8-34.2	75.9		
49.3-63.4	74.4	0-46.9	25	34.9-63.4	79.3		
2				**6**			
19.5-34.4	20			0-40.8	31		
34.4-74.5	22.8			42-62.5	27.6		
74.5-88.3(end)	20			63.7	31		
5p				65-69.9	24.1		
0-43	22.8			71-143.4	20.7		
43-45.6	20			144.4-149.5	17.2		
6				149.5-156.9	20.7		
0-29.3	25.7			158.2-170.5	17.2		
29.3-58.6	28.6			**9**			
58.6-71	25.7			0-29.2	20.7		
71-170.5	22.85			29.6-129	17.2		
9				130-134.2	20.7		
0-29.17	20			**10**			
10				0-37.2	37.9		
0-38.8	20			37.8-132.6	31		
38.8-87.4	22.85			132.6-135.2	20.7		
88.28-104.9	20			**11**			
106.1-135(end)	22.85			0-42	20.7		
11				43-57.4	24.1		
0-50.5	22.85			59.5-60	20.7		

65.1-97.7	20
97.7-112.4	22.85
112.4-134.6	25.7
16p	
0-31.1	25.7
31.1-56.6	20
17	
51.5-66.7	20
19	
0-63.7end	22.85
21	
0-41.6	25.7
41.6-46.9	22.85
61.4-73	24.1
74.3-74.39	20.7
75.7-134.6	17.2
17	
45.2-66.7	20.7
69.2-81.2	20.7
19	
0-51.7	34.5
52.1-63.7	31
21	
0-46.9 end	31

7. Losses at specific BAC clone sites

>5yrMSI-	%	>5yrMSI-	%	<5yrMSI-	%	<5yrMSI+	%
1		**20**		**1**		**1**	
0-7.4	74.3	0-63.4	50	0-23.6	70	0-45	83.3
8.9-24.2	77.1	**1**		23.8-32	66.7	45.2-45.9	66.6
26.6-32.9	74.3	0-245 end	25	32.6-34.6	63.3	46.2-118.9	50
33.8-43.6	71.4	**13**		36.3-40	60	140.9-245	33
44.1-45.9	62.8	0-53.8	25	40.5-50.3	56.7	**8**	
46.6	57.1	**17**		51.3-97.4	53.3	0-34.1	50
46.6-52.7	60	0-20.1	37.5	98.2-110.2	50	34.3-37	33.3
53.2-53.4	54.3	**21**		111.6-113.7	46.7	38-40.1	16.7
54.5-69.5	51.4	0-46.8	25	113.7-118.9	43.3	**9**	
70.1-81.4	48.6	**22**		140.9-161.9	33.3	0-29.5	33.3
82	45.7	0-49.3	37.5	162.2-196.8	30	64.7-119.9	50
83.1-84.5	40			198.7-245	33.3	119.9-134.2	66.6
85.4-118.9	42.8			**8p**		**17**	
141-142.7	31.4			0-30.7	50	0-20.2	50
143.2-145.2	28.6			31.7-34.3	46.6	25.2-26.7	16.7
145.2-175.3	25.7			35-38	43.3	**19**	
176.3-180.2	28.6			39.5	40	0-63.6	50
180.2-192.8	31.4			40.1	33.3	**16**	
193.9-200.9	34.3			40.1	30	65.2-89.7	33.3
200.9-218.3	31.4			40.9	26.7	**18**	
218.3-220.4	34.3			41-41.8	23.3	18.8-77.6	33.3
220.4-224	31.4			42.2-42.9	13.3	**20**	
224-245	28.6			**14**		0-20.1	33.3
8p				0-105.1	40	**22**	
0-31.7	48.6			**17**		21.4-49.2	66.7
31.7-32.7	45.7			0-20.1	70		
34.1-34.3	42.8			25.2-28.4	33.3		
34.3-36.2	40			29.2-34.9	30		
36.2	37.1			35.8-40.8	26.7		
37	28			40.9-41.8	23.2		
38	25.7			42.1-43.3	20		
39.5-41.9	22.8			**18**			
17				0-18.9	63.3		
0-20.1	68.6			19.7-24.4	60		
25.2	40			25.7-42.9	63.3		
25.2-26	31.4			43.5-77.6 end	66.7		
26.7-44.1	28.6			**22**			
45.2-48.5	25.7			0-49.2 end	66.7		
48.9-50.9	28.6						
51.5-66.7	25.7			**4**			
67.3-81.2	28.6			0-8.3	33.3		
18				10.3-55	30		
0-14.9	60			56-191.3	26.7		
18.8-48.1	62.8			**5**			
48.4-65.9	65.7			49.9-53.3	20		
66.5-77.6	62.8			55.4-71.7	23.3		
22				72.8-89.6	20		
0-36.8	62.8			90.8-109.4	23.3		

37.6-49.3	65.7
4	
179.7-191.3	22.8
9	
0-29.6	28.6
64.8-117.4	25.7
10	
69.7-77	20
88.2-104.9	20
14	
0-25.6	34.3
26.5-60.9	31.4
62.3-105.1	34.3
15	
0-53.2	40
54.8-99.7	37.1
19	
0-49.1	25.7
20	
0-14.6	25.7
15.5-	22.8
16.6-17.8	17.1
21	
0-30.7	34.3
32.9-41.6	31.4
42.6-46.9 end	34.3

110.1-118.3	20
9	
0-25.7	26.7
26.5-29.1	30
29.5	26.7
64.7-134.2	23.3
10	
85.7-101.1	20
15	
0-47.5	30
48.8-62.2	26.7
62.9-99.7	23.3
19	
0-51.7	30
51.7-63.3	33.3
20	
0-11.8	23.3
12.5-20.1	20
21	
0-46.8	23.3

8. Small areas of gain (<10Mb)

	5 yr +/MSI-	5 yr +/MSI+	5yr-/MSI-	5yr-/MSI+
Chr1	82-86		80-86	
	0-7.4		183.8-192.8	
	81.4-84.5		113.9-118	
Chr2			31.5-34.4	
			189-209	
Chr 3	66.7-74			
Chr5			40.1-49.9	
Chr6			149.5-156.9	
Chr7		98.5-101	0-2.4	
			20-23.5	
			67.4-end	
Chr8			142.3-145.7	
Chr9	129-134		130-134.2	
Chr10	41-48	130.9-135.2	75-83	
			132.6-135.2	
Chr12	0-4.1		0-4.1	
			66-74.7	
Chr13	19.1-27.9		19.1-27.9	
	22-27.9			
Chr15	25-34			
Chr16		0-7		
		84.6-89.7		
Chr17	44.1-48.5	80-83		
Chr18		77-80		
Chr20	15.5-17.8		0-9.3	

			16.1-22.5	

9. Small areas of loss (<10Mb)

	5yr+/MSI-	5yr+/MSI+	5yr-/MSI-	5yr-/MSI+
Chr2			129.9-136	
			236.4-242	
Chr4			0-8.3	
Chr7	62-67			
Chr9			25.7-29.1	
Chr10	68-77			
Chr11			73-81	
			104.4-107.9	
Chr12			74.7-82.3	
Chr17	48.5-50.9			

10. Areas of greater than single copy number change

	5yr+/MSI-	5yr+/MSI+	5yr-/MSI-	5yr-/MSI+
Chr1	140-152.8		161-209	
			113.9-118	
Chr2	34.4-74.5		31.5-34.4	
	187-225		129.9-136	
Chr3	66.7-74.5	90-94		
Chr4			0-37	
Chr5	126.1-135.8			
Chr6			2x p arm	
Chr7	p arm	0-14.7	31-40	
		98.5-101		
Chr8	3x q arm	143-146	q arm	80-145
	102-146			
Chr10	41-48	130.9-135	73-83	
			52-70	
Chr11	5-20			
	43-48			
	77-81			
Chr12			1x p arm	0-38
Chr13	19.1-27.9		6x whole chr	2x whole chr
Chr 18	77-80	18.9-42.9		
Chr 20	3x whole chr	60-66	6x whole chr	16.6- 22.5
	5 x q arm		2x q arm	29.4-47
	42-46		30-65	
	33-50		50-65	

29.4-63.4

11. Summary of adenoma aCGH changes.

Adenoma	Regions of change in Mb
1	1, 0-54,del; 6,33-47,del; 8,all,gain; 9,122-ter,del; 13,all,gain; 16,0-31,del; 16,67-ter,del; 17,0-50,del; 19,all,del; 20,all,gain; 21,all,gain; 22, all,del
2	1,0-51,del; 2,123-152,del; 17,all,del; 19,all,del
3	1, 0-57, del; 9,104-111, del ; 17,0-50,del
4	1, 0-47, del; 17, 0-50, del; 22,all,del
5	
6	1, 0-46, del; 17, 44-50,del
7	1, 0-46,del; 10,91-103,del; 18,51-64,del; 19,all,del
8	1,0-46,del; 7,62-74,del; 9,123-ter,del; 12,111-ter,del; 13,53-96,gain; 16,0-31,del; 16,66-ter,del; 17,0-50,del; 17,72-ter,del; 19,0-18,del; 19,38-ter,del; 20,31-36,del; 22,all, del
9	
10	1, 0-51, del
11	13, all, gain
12	
13	17, 0-14, del
14	