Meta-analysis and Pooled Re-analysis of Copy Number Changes in Colorectal Cancer Detected by Comparative Genomic Hybridization

SIMON HUGHES[1], RICHARD D. WILLIAMS[2], EMILY WEBB[1] and RICHARD S. HOULSTON[1]

Sections of [1]Cancer Genetics and [2]Paediatric Oncology,
Institute of Cancer Research, 15, Cotswold Road, Surrey, SM2 5NG, U.K.

Abstract. *Background: Comparative genomic hybridization (CGH) provides an insight into chromosomal changes associated with colorectal cancer (CRC) development. However, a problem with many studies is the limited cohort size, making the significance of some findings unclear. Materials and Methods: To derive a better insight into the chromosomal changes associated with CRC, we performed a meta-analysis and pooled re-analysis of published metaphase CGH data. Results: In addition to recurrent alterations, gains of 20 13q, 8q and 7p and loss of 18, 17p, 8p and 4p, pooling identified less frequent, but significant changes, including gain of 1q and 3, and losses from 6q, 9p and 21q. Conclusion: These additional alterations may be characteristic of some tumors and thus have relevance to CRC biology. Meta-analysis not only has the potential to detect novel changes, present at low frequency in several independent studies, but can provide greater reliability for their detection than single studies alone.*

Colorectal cancer is a common malignancy with around 110,000 cases in the United States each year (1) and despite improvements in surgical and medical management disease, prognosis remains poor (2). Chromosome copy number abnormalities (CNAs) are a feature of CRC development and progression (3). Allelic imbalance at tumor suppressor gene or oncogene loci may be important to the tumorigenic process, whilst other CNAs may reflect global genomic instability.

Investigation of CNAs in CRC has included both molecular and cytogenetic techniques (4-6). Metaphase (m) CGH has the advantage of not requiring *in vitro* culturing

of tumor cells, whilst still providing genome-wide analysis of chromosomal changes in a single experiment (7). Recurrent CNAs, as detected by mCGH, include gain of chromosome 20q and 13q and losses of chromosome 18q and 4q (3, 8, 9). Other changes, such as gain of 1q and 3, have been reported in some (10, 11), but not all studies (12, 13). As different reports have examined varying numbers of tumors, it is likely that less frequently occurring CNAs will not have been observed in all studies. To determine whether meta-analysis of mCGH data from several studies can clarify the prevalence of specific CNAs and better evaluate the significance of rare chromosomal abnormalities in CRC, we have reassessed the published mCGH literature. Using this approach, we have identified distinct and consistent regions of copy number change, furthermore, additional, less common CNAs have also been identified, which may be of biological relevance.

Materials and Methods

Identification of published studies. Reports detailing mCGH data on colorectal tumors were identified using the PubMed (January 2006) database (http://www.ncbi.nlm.nih.gov). The search terms were "colorectal cancer", "colon cancer", "rectal cancer", "adenoma", "CGH", "comparative genomic hybridization" and "comparative genomic hybridisation". Primary data papers (published in English in peer-reviewed journals) matching these criteria were retrieved and the bibliographies checked for other relevant publications. Data were also obtained from the Progenetix database (http://www.progenetix.net), SKY/M-FISH and CGH database (http://www.ncbi.nlm.nih.gov/sky) and Charité Online CGH Tumour Database (http://amba.charite.de/~ksch/cghdatabase/index.htm).

Database construction. Patient age, gender, Dukes stage, histological grade, DNA index and tumor microsatellite instability (MSI) status were extracted from the published articles and summarized in a consistent manner to aid comparison. Where possible, and when specified in the studies, cases with MSI and no chromosomal alterations were excluded from the analysis. Each

Correspondence to: Richard S. Houlston, Institute of Cancer Research, 15, Cotswold Road, Surrey, SM2 5NG. U.K. e-mail: richard.houlston@icr.ac.uk

Key Words: Meta-analysis, CGH, colorectal cancer, chromosome, genetic hybridization.

case was given a unique identification number in the meta-analysis database. Chromosomes were divided into individual bins defined by the 862 cytogenetic bands. Alterations were defined as the sequence of cytogenetic bands that they covered, for example, a copy number change on chromosome 8p would encompass the cytobands 8p23.3 to 8p11.1. This approach enabled the analysis of aberrations of differing size. Thresholds for the definition of losses and gains, between the papers, ranged from 0.5 to 0.875 and 1.15 to 1.5, respectively. Chromosomal alterations were transformed into loss –1, no change 0 and gain +1. The sex of the control DNA used for the separate studies was not always provided, therefore, though the data from the X and Y chromosomes was extracted from reports, it was not included in the meta-analysis. No discrimination was made between high copy number or standard gain, as it was of concern that such a differentiation would potentially bias the data.

Statistical analysis. To assess heterogeneity between studies in the rate of loss or gain at each cytogenetic band, a Monte Carlo approach was adopted with the program CLUMP being used to generate 10,000 replicates (14). The resulting *p*-values were adjusted using the Bonferroni correction for multiple testing, and compared with an individual test significance level of 0.05.

Threshold values for chromosomal alterations were assigned for the three tumor types, adenomas, primary tumors and liver metastases, and corresponded to the overall 95% upper confidence limits for the rates of gain and loss of copy number. Differences between adenomas and primary tumors, and between primary tumors and metastasis were assessed by means of the Chi-square test or Fisher exact test. A *p*-value of 0.05 or less was determined as being statistically significant.

Results

Systematic review and Meta-analysis. Forty-one studies were identified, which reported on the mCGH profiles of colorectal tumors (3, 5, 8-13, 15-47). Four of the studies were excluded, as they only examined single tumors (19), or they were a re-analysis of previously published data (32, 33, 35). Of the remaining studies, 14 did not present raw mCGH data in a form suitable for extraction and pooling (8, 9, 11-13, 17, 24, 25, 27, 34, 36, 38, 39, 41). However, 23 studies providing extractable data on 538 colorectal tumors, were suitable for pooling and provided an overview of allelic imbalances (3, 5, 10, 15, 16, 18, 20-23, 26, 28-31, 34, 37, 40, 42-47). Assuming partial independence of CNAs at each cytoband, there was no significant heterogeneity between observations from different studies for the three tumor types, as all adjusted *p* values were greater than 0.05. For liver metastases, 5% of adjusted *p* values were less than 0.05 and hence significant, but it was felt that this was not a high enough proportion to reject the assumption of no heterogeneity between study observations. On this basis, we felt it was permissible to pool mCGH data from individual studies.

The tumors comprised 65 adenomas, 308 primary tumors and 165 metastases (123 liver, 11 lymph node, 7 peritoneal and 24 lung). Several of the excluded papers presented

their mCGH data as summary figures, which made incorporation into this meta-analysis difficult. For the metastasis group only, the data from liver metastasis were further analyzed due to the small numbers of lymph, peritoneal and lung metastases. MMR status and DNA index were infrequently reported; hence it was not possible to make meaningful sub-analyses of the pooled data using such information.

The overall CNA frequencies were significantly higher in liver metastases compared with primary lesions and in primary lesions compared with adenomas ($p<0.0001$) (Figure 1). For all cytogenetic bands, the mean frequency of copy number gain were 8%, 11% and 14%, and for losses 6%, 9% and 11%, respectively for adenomas, primary tumors and metastases.

Analysis of the chromosomal differences with increasing Dukes' stage identified a significant increase ($p<0.05$) in the frequency of gain of 1q, 7p, 8q, 13q, 17q and 20q, in addition to loss of 2q, 4, 8p, 10q and 18q.

Five of the studies reported mCGH findings for 43 matched primary tumors and liver metastases (9, 15, 16, 23, 31). Pooled analysis of these samples identified a series of CNAs and although not statistically significant, gains of chromosomes 6, 7 and 8q, and losses of chromosomes 4, 8p and 18q were more frequent in metastatic tumors. When analyzing the non-matched pooled data the frequency of 11 (3 gains, 8 losses) CNAs were significantly different between adenomas and primary tumors (Figure 1a). Though notably higher frequencies of 20q11.1q13.3 and 13q11q34 and loss of 18q11.1q23 were observed in primary lesions, some aberrations appeared more common in adenomas. A comparison of primary tumors and liver metastases identified 19 (15 gains, 4 losses) significant differences in the frequency of CNAs between specific chromosomal regions (Figure 1b).

Discussion

This meta-analysis has provided an unbiased estimate of the frequency of CNAs across the genome and has demonstrated that pooling of data can detect commonly reported CNAs, as well as further common changes (present in greater than 20% of tumors) not determined as such by the original contributing mCGH studies. This is likely due to the relatively limited numbers of samples examined in some of the studies, which would make it difficult to interpret the significance of low frequency CNAs.

Analysis of the 43 matched primary tumors and liver metastasis allowed for an assessment of the validity of inferences derived from pooled unpaired data (9, 15, 16, 23, 31). Although the observed changes were not statistically significant, they showed an increase in CNA frequency, similar to that observed for unmatched tumors.
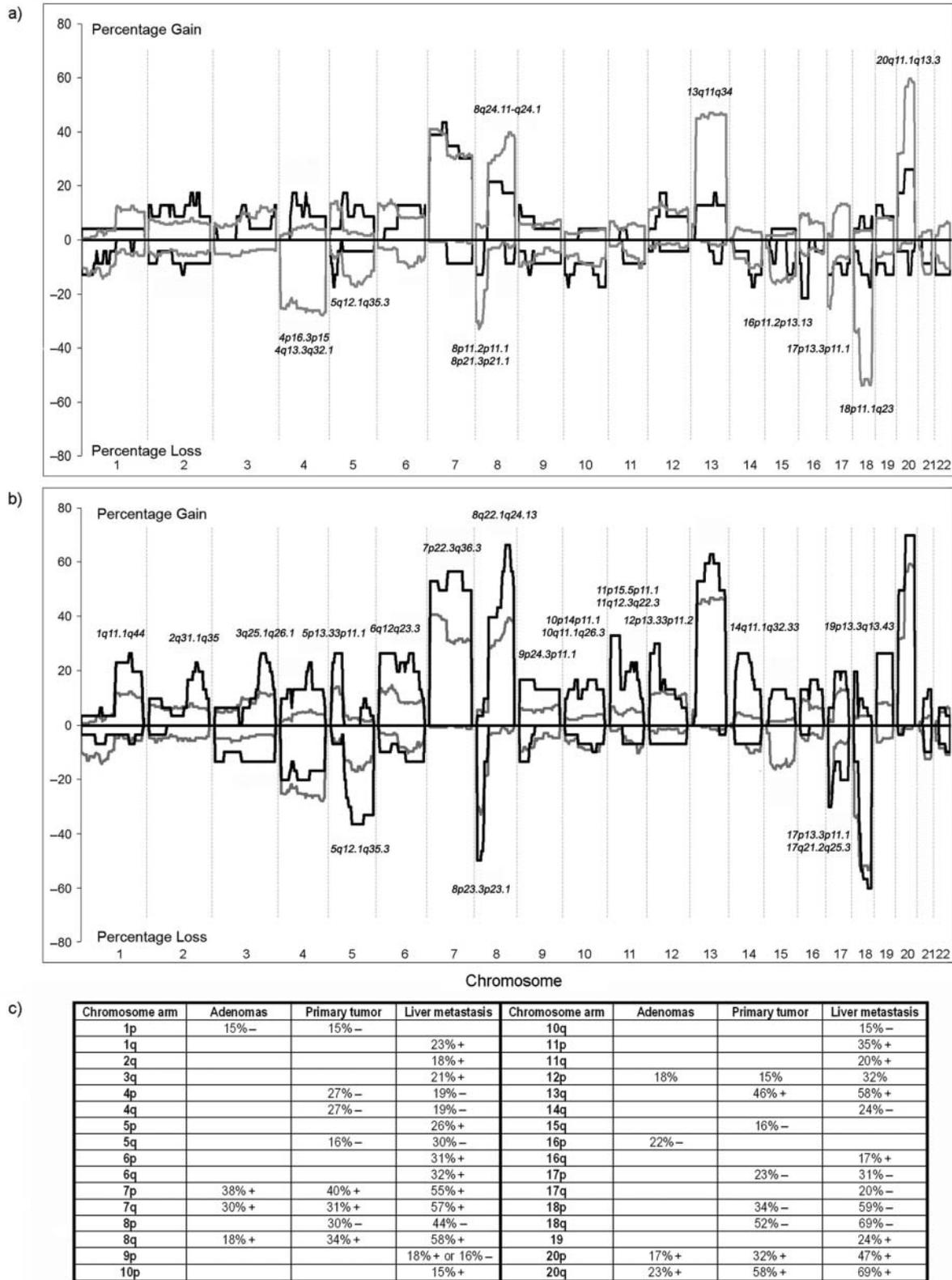
Figure 1. *a) Comparison of adenomas (indicated by black lines) with primary tumors (indicated by grey lines). b) Comparison of primary tumors (indicated by grey lines) with metastatic tumors (indicated by black lines). The frequency of copy number abnormalities at each chromosome region is expressed as a percentage. Dashed vertical lines indicates the boundary between chromosomes. c) A summary of the percentage CNA frequencies in adenomas, primary tumors and liver metastases. + signifies gain, and – signifies loss. Percentage values correspond to the averaged value over the region or gain or loss.*

Chromosome arm	Adenomas	Primary tumor	Liver metastasis	Chromosome arm	Adenomas	Primary tumor	Liver metastasis
1p	15%−	15%−		10q			15%−
1q		23%+		11p			35%+
2q		18%+		11q			20%+
3q		21%+		12p	18%	15%	32%
4p		27%−	19%−	13q		46%+	58%+
4q		27%−	19%−	14q			24%−
5p		26%+		15q		16%−	
5q		16%−	30%−	16p	22%−		
6p		31%+		16q			17%+
6q		32%+		17p		23%−	31%−
7p	38%+	40%+	55%+	17q			20%−
7q	30%+	31%+	57%+	18p		34%−	59%−
8p		30%−	44%−	18q		52%−	69%−
8q	18%+	34%+	58%+	19			24%+
9p			18%+ or 16%−	20p	17%+	32%+	47%+
10p			15%+	20q	23%+	58%+	69%+

Whilst the three tumor types share common alterations, additional CNAs are acquired during disease progression (Figure 1c). A summary of published data displays a step wise accumulation of CNAs involved in the adenoma-carcinoma sequence (48). Alterations observed on chromosomes 1, 6p, 9p and 21q, although reported in independent studies, have rarely been referred to as recurrent changes.

An analysis of chromosomal differences between Dukes' stages identified an increased frequency of copy number gain from stage B to D, confirming previous observations (20). As there is strong relationship between chromosomal gain and/or loss and modification of gene expression, it is likely the changes observed will reflect programmed cycle of events that promote tumor development and survival.

Although mCGH does not permit specific gene identification, it does point to regions of particular interest. Furthermore, several proven or proposed tumor suppressor genes or oncogenes have been reported to reside on effected chromosomal arms. Therefore, the summarizing of several papers, as performed here, can help to direct researchers to particular sets of genes that may be of interest.

A point not discussed in the reviewed articles is the effect of sequence dependent differences on the detection of imbalances by metaphase CGH (49). The centromeric, pericentromeric and possibly telomeric regions of chromosomes 1, 9, 16 and 19, contain a large amount of repetitive sequence elements, which may be more prone to exhibiting elevated background signals by metaphase CGH (50). Furthermore, results based on indirect fluorochrome labelling are more prone to locus specific artefacts than results from direct labelling (49). As a result, it is difficult to predict their effect on a single experiment. Meta-analysis of four recent studies, which applied array CGH for the analysis of colorectal tumors (51-54) (data not shown) has produced results similar to those identified by the mCGH meta-analysis. This concordance increases the confidence with which the results from the meta-analysis can be viewed and shows the usefulness of such analyses for identifying significant CNAs following the pooling of data from multiple studies.

By employing a meta-analysis approach, mCGH data from different studies can be summarized. Furthermore, the data obtained can provide a refinement of the results obtained from independent studies. The combined database described here has identified distinct and consistent regions of copy number change, ranging from specific cytobands to whole chromosomes. Moreover, it has identified additional CNAs associated with colorectal tumors that may have biological relevance yet to be elucidated. The pooling of data, as described here, has the potential to provide greater reliability of detecting genomic aberrations than single studies alone.

References

1 Smith RA, Cokkinides V and Eyre HJ: American Cancer Society guidelines for the early detection of cancer, 2004. CA Cancer J Clin *54(1)*: 41-52, 2004.

2 Bendardaf R, Lamlum H and Pyrhonen S: Prognostic and predictive molecular markers in colorectal carcinoma. Anticancer Res *24(4)*: 2519-2530, 2004.

3 Meijer GA, Hermsen MA, Baak JP *et al*: Progression from colorectal adenoma to carcinoma is associated with non-random chromosomal gains as detected by comparative genomic hybridisation. J Clin Pathol *51(12)*: 901-909, 1998.

4 Couturier-Turpin MH, Bertrand V and Couturier D: Distal deletion of 1p in colorectal tumors: an initial event and/or a step in carcinogenesis? Study by fluorescence *in situ* hybridization interphase cytogenetics. Cancer Genet Cytogenet *124(1)*: 47-55, 2001.

5 Ried T, Knutzen R, Steinbeck R *et al*: Comparative genomic hybridization reveals a specific pattern of chromosomal gains and losses during the genesis of colorectal tumors. Genes Chromosomes Cancer *15(4)*: 234-245, 1996.

6 Yashiro M, Carethers JM, Laghi L *et al*: Genetic pathways in the evolution of morphologically distinct colorectal neoplasms. Cancer Res *61(6)*: 2676-2683, 2001.

7 Kallioniemi A, Kallioniemi OP, Sudar D *et al*: Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science *258(5083)*: 818-821, 1992.

8 Aragane H, Sakakura C, Nakanishi M *et al*: Chromosomal aberrations in colorectal cancers and liver metastases analyzed by comparative genomic hybridization. Int J Cancer *94(5)*: 623-629, 2001.

9 Knosel T, Petersen S, Schwabe H *et al*: Incidence of chromosomal imbalances in advanced colorectal carcinomas and their metastases. Virchows Arch *440(2)*: 187-194, 2002.

10 He QJ, Zeng WF, Sham JS *et al*: Recurrent genetic alterations in 26 colorectal carcinomas and 21 adenomas from Chinese patients. Cancer Genet Cytogenet *144(2)*: 112-118, 2003.

11 Leslie A, Pratt NR, Gillespie K *et al*: Mutations of APC, K-ras, and p53 are associated with specific chromosomal aberrations in colorectal adenocarcinomas. Cancer Res *63(15)*: 4656-4661, 2003.

12 Paredes-Zaglul A, Kang JJ, Essig YP *et al*: Analysis of colorectal cancer by comparative genomic hybridization: evidence for induction of the metastatic phenotype by loss of tumor suppressor genes. Clin Cancer Res *4(4)*: 879-886, 1998.

13 Richter H, Slezak P, Walch A *et al*: Distinct chromosomal imbalances in nonpolypoid and polypoid colorectal adenomas indicate different genetic pathways in the development of colorectal neoplasms. Am J Pathol *163(1)*: 287-294, 2003.

14 Sham PC and Curtis D: Monte Carlo tests for associations between disease and alleles at highly polymorphic loci. Ann Hum Genet *59(Pt 1)*: 97-105, 1995.

15 Alcock HE, Stephenson TJ, Royds JA *et al*: Analysis of colorectal tumor progression by microdissection and comparative genomic hybridization. Genes Chromosomes Cancer *37(4)*: 369-380, 2003.

16 Al-Mulla F, Keith WN, Pickford IR *et al*: Comparative genomic hybridization analysis of primary colorectal carcinomas and their synchronous metastases. Genes Chromosomes Cancer *24(4)*: 306-314, 1999.

17 Aust DE, Willenbucher RF, Terdiman JP *et al*: Chromosomal alterations in ulcerative colitis-related and sporadic colorectal cancers by comparative genomic hybridization. Hum Pathol *31(1)*: 109-114, 2000.

18 Chan TL, Curtis LC, Leung SY *et al*: Early-onset colorectal cancer with stable microsatellite DNA and near-diploid chromosomes. Oncogene *20(35)*: 4871-4876, 2001.

19 Clausen OP, Andersen SN, Stroomkjaer H *et al*: A strategy combining flow sorting and comparative genomic hybridization for studying genetic aberrations at different stages of colorectal tumorigenesis in ulcerative colitis. Cytometry *43(1)*: 46-54, 2001.

20 De Angelis PM, Clausen OP, Schjolberg A *et al*: Chromosomal gains and losses in primary colorectal carcinomas detected by CGH and their associations with tumour DNA ploidy, genotypes and phenotypes. Br J Cancer *80(3-4)*: 526-535, 1999.

21 De Angelis PM, Stokke T, Beigi M *et al*: Prognostic significance of recurrent chromosomal aberrations detected by comparative genomic hybridization in sporadic colorectal cancer. Int J Colorectal Dis *16(1)*: 38-45, 2001.

22 Diep CB, Parada LA, Teixeira MR *et al*: Genetic profiling of colorectal cancer liver metastases by combined comparative genomic hybridization and G-banding analysis. Genes Chromosomes Cancer *36(2)*: 189-197, 2003.

23 Diep CB, Teixeira MR, Thorstensen L *et al*: Genome characteristics of primary carcinomas, local recurrences, carcinomatoses, and liver metastases from colorectal cancer patients. Mol Cancer *3(1)*: 6, 2004.

24 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 *18(56)*: 7933-7940, 1999.

25 Ghadimi BM, Grade M, Liersch T *et al*: Gain of chromosome 8q23-24 is a predictive marker for lymph node positivity in colorectal cancer. Clin Cancer Res *9(5)*: 1808-1814, 2003.

26 Gunawan B, Schulten HJ, von Heydebreck A *et al*: Site-independent prognostic value of chromosome 9q loss in primary gastrointestinal stromal tumours. J Pathol *202(4)*: 421-429, 2004.

27 Hawkins NJ, Gorman P, Tomlinson IP *et al*: Colorectal carcinomas arising in the hyperplastic polyposis syndrome progress through the chromosomal instability pathway. Am J Pathol *157(2)*: 385-392, 2000.

28 Habermann JK, Upender MB, Roblick UJ *et al*: Pronounced chromosomal instability and multiple gene amplifications characterize ulcerative colitis-associated colorectal carcinomas. Cancer Genet Cytogenet *147(1)*: 9-17, 2003.

29 Hermsen M, Postma C, Baak J *et al*: Colorectal adenoma to carcinoma progression follows multiple pathways of chromosomal instability. Gastroenterology *123(4)*: 1109-1119, 2002.

30 Iwamoto M, Banerjee D, Menon LG *et al*: Overexpression of E2F-1 in lung and liver metastases of human colon cancer is associated with gene amplification. Cancer Biol Ther *3(4)*: 395-399, 2004.

31 Korn WM, Yasutake T, Kuo WL *et al*: Chromosome arm 20q gains and other genomic alterations in colorectal cancer metastatic to liver, as analyzed by comparative genomic hybridization and fluorescence *in situ* hybridization. Genes Chromosomes Cancer *25(2)*: 82-90, 1999.

32 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 *19(5)*: 401-407, 2002.

33 Knosel T, Schluns K, Stein U *et al*: Genetic imbalances with impact on survival in colorectal cancer patients. Histopathology *43(4)*: 323-331, 2003.

34 Knosel T, Schluns K, Stein U *et al*: Chromosomal alterations during lymphatic and liver metastasis formation of colorectal cancer. Neoplasia *6(1)*: 23-28, 2004.

35 Knosel T, Yu Y, Stein U *et al*: Overexpression of cyclooxygenase-2 correlates with chromosomal gain at the cyclooxygenase-2 locus and decreased patient survival in advanced colorectal carcinomas. Dis Colon Rectum *47(1)*: 70-77, 2004.

36 Li L, Kim N, Kim S *et al*: Chromosomal imbalances in the colorectal carcinomas with microsatellite instability. Am J Pathol *163(4)*: 1429-1436, 2003.

37 Leung SY, Yuen ST, Chan TL *et al*: Chromosomal instability and p53 inactivation are required for genesis of glioblastoma but not for colorectal cancer in patients with germline mismatch repair gene mutation. Oncogene *19(35)*: 4079-4083, 2000.

38 Nakao K, Shibusawa M, Tsunoda A *et al*: Genetic changes in primary colorectal cancer by comparative genomic hybridization. Surg Today *28(5)*: 567-569, 1998.

39 Nakao K, Shibusawa M, Ishihara A *et al*: Genetic changes in colorectal carcinoma tumors with liver metastases analyzed by comparative genomic hybridization and DNA ploidy. Cancer *91(4)*: 721-726, 2001.

40 Rooney PH, Boonsong A, McKay JA *et al*: Colorectal cancer genomics: evidence for multiple genotypes which influence survival. Br J Cancer *85(10)*: 1492-1498, 2001.

41 Schlegel J, Stumm G, Scherthan H *et al*: Comparative genomic *in situ* hybridization of colon carcinomas with replication error. Cancer Res *55(24)*: 6002-6005, 1995.

42 Willenbucher RF, Zelman SJ, Ferrell LD *et al*: Chromosomal alterations in ulcerative colitis-related neoplastic progression. Gastroenterology *113(3)*: 791-801, 1997.

43 Willenbucher RF, Aust DE, Chang CG *et al*: Genomic instability is an early event during the progression pathway of ulcerative-colitis-related neoplasia. Am J Pathol *154(6)*: 1825-1830, 1999.

44 Tarafa G, Prat E, Risques RA *et al*: Common genetic evolutionary pathways in familial adenomatous polyposis tumors. Cancer Res *63(18)*: 5731-5737, 2003.

45 Platzer P, Upender MB, Wilson K *et al*: Silence of chromosomal amplifications in colon cancer. Cancer Res *62(4)*: 1134-1138, 2002.

46 Poeaim S, Rerkamnuaychoke B, Jesdapatarakul S *et al*: Chromosome alterations in colorectal cancer in Thai patients. Cancer Genet Cytogenet *160(2)*: 152-159, 2005.

47 Jiang JK, Chen YJ, Lin CH *et al*: Genetic changes and clonality relationship between primary colorectal cancers and their pulmonary metastases – an analysis by comparative genomic hybridization. Genes Chromosomes Cancer 43(1): 25-36, 2005.

48 Fearon ER and Vogelstein B: A genetic model for colorectal tumorigenesis. Cell *61(5)*: 759-767, 1990.

49 Kallioniemi OP, Kallioniemi A, Piper J *et al*: Optimizing comparative genomic hybridization for analysis of DNA sequence copy number changes in solid tumors. Genes Chromosomes Cancer *10(4)*: 231-243, 1994.

50 Larramendy ML, El-Rifai W and Knuutila S: Comparison of fluorescein isothiocyanate- and Texas red-conjugated nucleotides for direct labeling in comparative genomic hybridization. Cytometry *31(3)*: 174-179, 1998.

51 Douglas EJ, Fiegler H, Rowan A *et al*: Array comparative genomic hybridization analysis of colorectal cancer cell lines and primary carcinomas. Cancer Res *64(14)*: 4817-4825, 2004.

52 Jones AM, Douglas EJ, Halford SE *et al*: Array-CGH analysis of microsatellite-stable, near-diploid bowel cancers and comparison with other types of colorectal carcinoma. Oncogene *24(1)*: 118-129, 2005.

53 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 *11(5)*: 1791-1797, 2005.

54 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 *25(8)*: 1345-1357, 2004.