

Genomic and transcriptomic prognostic factors in R0 Dukes B and C colorectal cancer patients

ELISENDA VENDRELL¹, MARIA RIBAS¹, JOAN VALLS², XAVIER SOLÉ², MÒNICA GRAU^{2,3},
VÍCTOR MORENO², GABRIEL CAPELLÀ² and MIGUEL A. PEINADO¹

¹Institut d'Investigació Biomèdica de Bellvitge (IDIBELL), ²IDIBELL-Institut Català d'Oncologia,
L'Hospitalet, Barcelona, Catalonia, Spain

Received November 17, 2006; Accepted January 30, 2007

Abstract. The advent of various 'omic' technologies has increased expectations in the field of biomarkers. In an attempt to clarify how different strategies may contribute to improving prognostic classification and to identify new predictors of patient outcome we analyzed genomic and transcriptomic profiles in a series of R0 Dukes B and C colorectal carcinomas. We have compared the predictive capability of each approach against conventional clinicopathological and molecular parameters. At a genomic level, gains at 11q including amplification at 11q13 were an indicator of poorer outcome. In transcriptomic analyses we identified 68 genes whose expression levels correlated with survival ($p < 0.01$) and included overexpression of WASF1, NFE2L2, and MMP9, and underexpression of ITGAL, TSC2, and SDF2. Gene expression levels paralleled chromosomal changes only in 56% of the genes, suggesting that, as a general trend, the direct effect of chromosomal copy number changes on gene expression levels is minimal. Classification of tumors by genomic and transcriptomic signatures resulted in non-overlapping subgroups and was not of prognostic value. We conclude that genomic and transcriptomic profiling of colorectal carcinomas may contribute as novel prognostic markers, but it does not improve outcome prediction when global profiles or signatures are considered.

Introduction

Colorectal cancer is the third most common type of neoplasia in both men and women and the second-leading cause of death by cancer in occidental countries (1). The extent of

tumor bowel wall infiltration and lymph node metastases, both included in Dukes' stage and TNM classification systems, are the most important prognostic factors in colorectal cancer (2). Nevertheless, traditional morphologic criteria based upon pathologist's evaluation are accurate for predicting recurrence only in 50-75% of the patients with non-metastatic invasive colon carcinoma. Therefore there is a need for additional, less subjective, independent factors to better predict outcome.

Multiple genetic aberrations are required for tumor initiation and progression of colorectal cancer, which is one of the best studied systems of multistage human carcinogenesis. Besides the advances in the understanding of the molecular factors involved in this process, the heterogeneity and complexity of the disease make it difficult to apply molecular information to predict the evolution of an individual patient's disease (3). A major challenge is to integrate information that can describe this complexity so as to facilitate an understanding of the disease mechanisms as well as to guide the development and application of therapies (4). The advent of various 'omic' technologies has increased expectations in the field of biomarkers, but they have not yet produced widely applicable approaches in prognostic assessment and patient treatment. Four levels of analyses can be considered: genomic, transcriptomic, epigenomic, and proteomic, the first two being the most often applied due to the availability of appropriate methodologies.

Chromosomal aberrations in the form of aneuploidy and structural rearrangements are early markers and probably the most prevalent genetic alteration in colorectal carcinogenesis (5,6). Recurrent chromosomal abnormalities often clustered in association patterns are also observed and may be used to classify colorectal cancers (6,7). Furthermore, a subset of colorectal tumors with few or no chromosomal alterations are characterized by ubiquitous somatic mutations at repeated sequences (6-8). These tumors represent a distinctive pathway of tumor progression in which defects in the DNA mismatch repair machinery underlie the genetic instability expressed as an exacerbated microsatellite instability (MSI).

Conventional G-banding cytogenetics has been instrumental in the identification of the chromosomal alterations associated with malignancy and has provided potential prognostic markers in colorectal cancer (9). The availability of comparative genomic hybridization (CGH) (10) as an alternative to classic cytogenetics has facilitated karyotyping and nowadays is the

Correspondence to: Dr Miguel A. Peinado, Institut d'Investigació Biomèdica de Bellvitge (IDIBELL), Av. Granvia km 2.7, 08907 L'Hospitalet, Barcelona, Catalonia, Spain
E-mail: mpeinado@idibell.org

Present address: ³Microarrays Unit, Centre de Regulació Genòmica, Barcelona, Catalonia, Spain

Key words: colorectal cancer, prognostic factor, chromosomal alterations, transcriptomic profiles

Table I. Molecular and clinicopathological characteristics of the 50 CRC cases.

Variable	No. of cases
Sex	
Female	24
Male	26
Dukes' stage	
B	30
C	20
Location	
Right	15
Left	15
Rectum	20
Outcome	
Alive without disease	29
With relapse	21
Therapy	
Untreated	20
Chemotherapy	20
Radiotherapy	10
p53	
Wild-type	15
Mutated	35
K-ras	
Wild-type	33
Mutated	17

technique most frequently used. It allows the investigation of specific chromosomal alterations together with global patterns of chromosomal disruption. Most recurrent chromosomal alterations previously detected by G-banding and allelotyping analyses, have been confirmed by CGH. Different studies have reported an enhanced number of chromosomal aberrations in advanced stages or associated with worst survival in colorectal cancer (11-14). CGH studies have also revealed multiple single chromosomal alterations that are potential prognostic markers (12,13), and have contributed to defining patterns of chromosomal aberrations associated with progression (11,15).

In the last decade technical and bioinformatic advances have made it practical to quantify the expression of thousands of genes using microarrays. This approach has been used to analyze gene expression profiles in a variety of human tumors. One of the most popular applications of transcriptomics has been the generation of biomarkers with predictive competency (16,17). Classification of colorectal cancers according to gene expression profiles has been reported to improve prognostic assessment (18-22) and prediction of response to adjuvant therapy (23). Besides, these promising results, application of large-scale genomic analyses to the prediction and management of human disease remains excluded from routine clinical practice. Confounding and biased study designs, technology-

related limitations and statistical and analytical problems pose critical obstacles to the development of reliable predictors using these technologies (16,17).

In an attempt to clarify how different strategies may contribute to improve prognostic classification and to identify new predictors of patient outcome we have obtained comprehensive profiles at genomic and transcriptomic levels in a series of R0 Dukes B and C colorectal carcinomas. We have compared the predictive capability of each approach among them and against conventional clinicopathological and molecular parameters. Moreover, we have also investigated the overlapping between each classification.

Materials and methods

Samples. A series of 50 patients preoperatively diagnosed with colorectal cancer and operated upon with curative or palliative intention between 1996 and 1998 at the Ciutat Sanitària i Universitària de Bellvitge was used in this study. This series included R0 Dukes B or Dukes C colorectal cancer and was part of a larger series of patients prospectively included in a study designed to evaluate the prognostic value of genetic and epigenetic alterations (24). Clinical and pathological features of the patients are summarized in Table I. The patient and tumor characteristics did not differ significantly between the relapse and no relapse patients, or between treated or non-treated patients.

Surgical specimens were collected at the operating room and immediately taken to the pathology department in ice. Carcinomas and paired normal samples were snap frozen within 2 h after removal and then stored at -80°C. Tumor sections were cut from microdissected regions of the tumor infiltration front with approximately 75% tumor cell content, as assessed by visual examination of hematoxylin and eosin-stained preparations. Special care was taken to discard areas with necrotic tissue or harboring a high inflammatory component. DNA and RNA amenable for genetic analyses were obtained using standard procedures.

Main clinicopathological features and molecular data. Mutations in the tumor suppressor gene p53, the K-ras oncogene and the presence of microsatellite instability were assessed as previously described (24).

Comparative genome hybridization (CGH). This technique was performed as described elsewhere (10). Briefly, normal and tumor-cell DNA were labeled with SpectrumRed-dUTP and SpectrumGreen-dUTP and equal amounts of normal and tumor labeled probes hybridized to normal metaphase spreads. Chromosomes were counterstained with 4,6-diamino-2-phenylindole (DAPI) for chromosome identification. Slides were analyzed by using a Cytovision Ultra workstation (Applied Imaging, Sunderland, UK). The software calculated the tumor DNA versus normal DNA fluorescent ratios along the length of each chromosome. Ratio values obtained from at least 10 metaphase spreads were averaged, and the resulting profiles were plotted next to the chromosome ideograms.

Gene expression analysis. To generate enough material for technical replicate hybridizations and to increase the representation

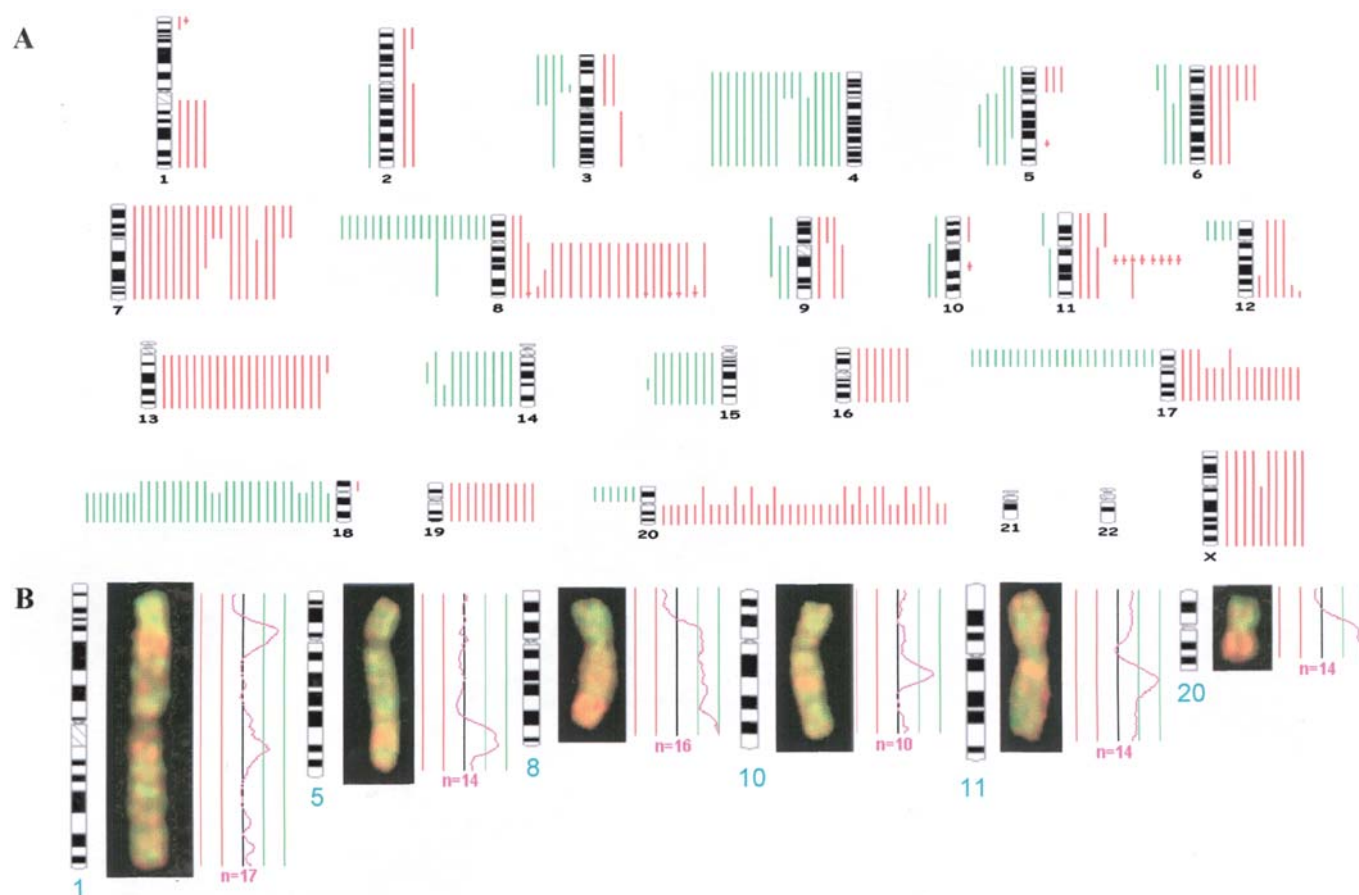


Figure 1. (A) Summary of the genomic alterations detected by CGH in 50 colorectal carcinomas. Bars on the left side of the chromosome ideogram denote a loss, bars on the right a gain. Green bars indicate losses, red bars indicate gains. Dots indicate regions of high-level copy number gains (amplifications). (B) Illustrative examples of gene amplification as detected by CGH.

of rare transcripts, total RNA was amplified by RNA Arbitrarily Primed PCR (RAP-PCR) and hybridized to cDNA arrays as previously described (25). Briefly, RAP-PCR consisted of a reverse transcription step using primer pU6 (5'-GCTTCTGACTTATTCTTGCTCTGAG-3') followed by a PCR using the same primer pU6 consisting of five low-stringency cycles (annealing at 40°C) and 35 high-stringency cycles (1 annealing at 60°C). Performance and quality of the reaction were assessed by gel electrophoresis in denaturing sequencing gels and silver staining. The most variable step in using RAP-PCR as a probe was the amplification by PCR (25). Therefore, we performed two independent RT and from each one, three RAP-PCRs. After checking their quality by gel electrophoresis, we pooled them before labeling. A pool of 12 RAP-PCRs generated from 12 independent normal colon mucosas was used as control.

RAP-PCR products were purified and labeled with dCTP-Cy3 or dCTP-Cy5 (Amersham Biosciences UK Limited). The tumor sample was labeled with fluorochrome Cy5 and the normal pool was labeled with fluorochrome Cy3. The pooled RAP-PCR from each tumor tissue was cohybridized in duplicate with the pool of normal tissues in poly-L-lysine cDNA microarrays (Hu 4.6K; W.M. Keck Foundation Biotechnology Resource Laboratory, <http://keck.med.yale.edu>) containing 4608 human cDNAs spotted in duplicate as previously described (25). A list of the genes covered can be

obtained at http://keck.med.yale.edu/dnaarrays/genelists/QAHU-4_6k_cy.htm. Quality controls, reproducibility and individual gene validation for this design were performed as described in a previous study (25).

Slides were scanned with a GSI Lumonics ScanArray 4000 and images were analyzed by use of Spot software (CSIRO, Mathematical and Informatics Sciences), and the resulting data were processed to filter out low-quality spots based on spot area and similarity of readings between the two replicates of each gene. Replicate intensities were averaged and normalized. To represent relative intensities, a reference value obtained from the mean of 13 independent hybridizations of the normal pool was used.

Statistical analysis. The effect of the clinicopathological and molecular characteristics on the total number of chromosomal alterations was assessed using a t-test or an ANOVA-test as appropriate. To classify the tumors based on the CGH data (-1 loss, 0 no change and 1 gain) we applied a hierarchical classification algorithm using the Ward criterion. The prognostic value of the variables on survival was evaluated using the log-rank test. Kaplan-Meier estimates were computed to generate the survival curves. Multivariate Cox models that included Dukes' stage were further used to evaluate the prognostic value of the variables. Gene expression data was used to detect differentially expressed genes across all samples,

Table II. Recurrent chromosomal alterations in 50 colorectal carcinomas.

Chromosome	No. of cases	Frequency (%)
Losses		
18q	33	66
17p	24	48
18p	20	40
8p	19	38
4p	16	32
4q	15	30
14	11	22
15	9	18
Gains		
20q	37	74
8q	23	46
13	22	44
7p	19	38
7q	16	32
17q	16	32
20p	13	26
11q	11	22
19	11	22
Xq	10	20
Xp	9	18

using for this purpose a t-test. Furthermore, the effect of gene expression on the survival was assessed using Cox models allowed to compute hazard ratios (HR). To evaluate significance we computed the score (log-rank) test. Again, multivariate Cox models that included Dukes' stage were used and a Wald test was computed to evaluate significance. Moreover, the effect of the chromosomal alterations on the gene expression were evaluated using a t-test.

The significance level was set to 0.05, or to 0.01 when multiple analyses were performed. All analyses were performed using the statistical package R (<http://www.R-project.org>).

Results

Chromosomal alterations. The CGH profiles of the 50 colorectal carcinomas are shown in Fig. 1A. To determine the number of chromosomal alterations, losses and gains at p and q arms were scored independently, except for telocentric chromosomes 13, 14, and 15, and chromosome 19, which always showed gains of the whole chromosome. Chromosomes 21, 22 and Y were not considered in this study. The number of alterations per tumor ranged from 0 to 20 (mean 8.8 ± 6.3). Nine tumors (18%) had no alterations, and the 41 remaining (82%) had at least one alteration. The most recurrent alterations were losses of chromosome arms 18q and 17p and gains of 20q and 8q (Table II). DNA amplification was observed in 19 tumors and affected 6 different loci: 20q (n=10), 11q13 (n=8), 8q24 (n=5), 5q31-q33, 10q22 and 1p36

Table III. Number of chromosomal alterations in reference to clinicopathological and molecular characteristics.

Variable	N	No. of alterations	p-value ^a
Sex			
Male	26	9.9±6.5	0.206
Female	24	7.6±5.9	
Dukes' stage			
B	30	8.3±6.7	0.495
C	20	9.5±5.6	
Localization ^b			
Left ^a	35	8.5±6.0	0.560
Right	15	9.6±7.0	
Disease-free survival			
Yes	29	8.1±5.8	0.361
No	21	9.8±6.9	
Overall survival			
Yes	34	8.3±5.8	0.439
No	16	9.8±7.3	
p53			
Wt	15	5.1±6.1	0.006
Mutated	35	10.4 ±5.7	
K-ras			
Wt	33	8.1±6.1	0.249
Mutated	17	10.2±6.6	
MSI			
Yes	3	1.7±1.5	0.041
No	47	9.3±6.2	
Amplifications ^c			
Yes	19	12.4±4.9	0.001
No	31	6.6±6.0	
CGH group			
MSI	3	1.7±1.5	<0.001
NT	8	0	
CIN-M	22	12.8±4.3	
CIN-R	17	9.0±5.0	

^aStudent's t-test or ANOVA as appropriate. ^bLeft includes tumors of the rectum. ^cAmplifications were not included in the computation of the number of chromosomal alterations.

(in one tumor each). Examples of gene amplification are illustrated in Fig. 1B.

When tumors were classified according to different clinicopathological and molecular data (Table III), a clear association between the presence of mutation in the p53 gene and the number of chromosome alterations was observed. Even when only one type of alteration was considered: i.e. gains, losses, structural, numerical, amplifications (data not shown), the

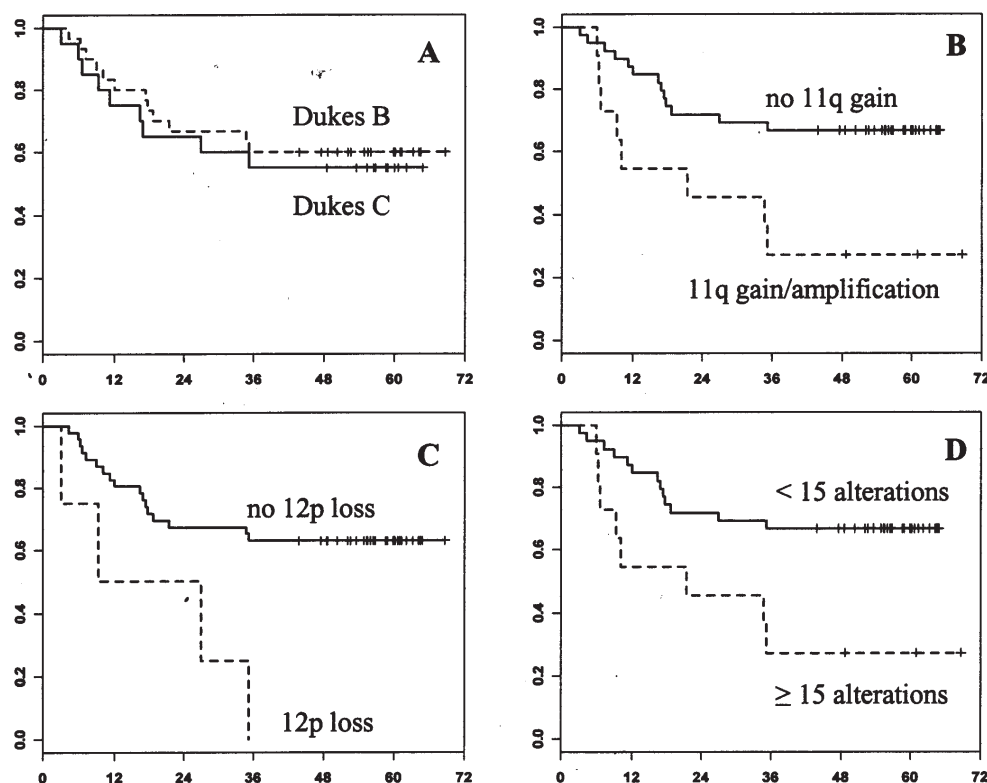


Figure 2. Kaplan-Meier disease-free survival curves for colorectal cancer patients according to (A) Dukes' stage (log-rank, $p=0.655$), (B) gains at 11q/amplification at 11q13 (log-rank, $p=0.012$), (C) losses at 12p (log-rank, $p=0.006$), and (D) number of alterations in CIN-M group (log-rank = 0.016).

statistical association was significant, suggesting the principal role of p53 inactivation in chromosomal instability. Tumors with microsatellite instability showed less chromosomal alterations and the presence of amplifications also correlated with a higher number of chromosomal alterations other than amplifications (Table III).

Subgrouping of tumors according to genomic profiles. Colorectal cancer is a heterogeneous disease and interactions and confounding factors are likely to mask associations restricted to a homogeneous subgroup of tumors. To circumvent this problem different studies have considered it appropriate to subclassify tumors according to putative genetic pathways. Several strategies and criteria were considered to subgroup tumors. All types of alterations (numerical, structural, losses, gains and amplifications) correlated (data not shown), therefore it was considered inappropriate to classify tumors depending on the prevailing type of chromosomal alteration.

Hierarchical cluster analysis identified two main groups: tumors without alterations and others (data not shown). Tumors with alterations could be also classified into two subgroups essentially differentiated by the number of alterations but no specific signatures of each group were observed (data not shown). Finally, we used criteria adapted from classic cytogenetic studies reported by Dutrillaux (7). Accordingly, tumors were classified into four groups: microsatellite instability (positive for MSI), normal type (NT) with no MSI and no chromosomal alterations, chromosomal instability monosomic-like (CIN-M) (concomitant loss of 17p and 18q) and chromosomal instability remaining (CIN-R). The main

factor differentiating groups was, in addition to the number of chromosomal alterations (Table III), the frequency of p53 mutations, which was 33% in the MSI and NT groups and 77 and 88% in the CIN-M and CIN-R groups respectively (ANOVA, $p=0.005$).

Genomic alterations and survival. None of the clinico-pathological and molecular parameters previously considered had prognostic value, including Dukes' stage (Fig. 2A). The association between chromosomal alterations and patient outcome was investigated. Gains at 11q ($n=11$), including tumors with amplification in this region, were an indicator of recurrence of the disease ($p=0.012$) (Fig. 2B). When only amplifications at 11q13 were considered ($n=8$), the differences were stressed ($p=0.0003$), suggesting the principal involvement of this region in conferring poor prognosis. Treatment received (adjuvant chemotherapy and/or radiotherapy) did not appear to modify the outcome. Loss at 12p occurred in four cases and all presented recurrence of the disease ($p=0.006$) (Fig. 2C). The statistical significance was maintained when multivariate Cox analysis including the Dukes' stage in the model was performed (Table IV). No differences in survival were observed when tumors were classified according to the number of chromosomal alterations or grouped by genomic signatures (data not shown). Nevertheless the number of chromosomal alterations appeared to have different associations in CIN-M and CIN-R tumors. In CIN-M tumors a higher number of chromosomal alterations was an indicator of poor prognosis (HR=1.2, CI 95%: 1.0-1.5, $p=0.036$). Receiver Operating Characteristic Curve analysis offered an optimal cut-off point

Table IV. Multivariate Cox analysis of disease-free survival.

	Categories	Hazard ratio	95% CI	P-value
Dukes' stage	C (vs. B)	1.9	0.7-5.0	0.198
11q	Gain (vs. rest)	3.9	1.4-10.5	0.008
Dukes' stage	C (vs. B)	0.8	0.3-2.2	0.653
12p	Loss (vs. rest)	4.9	1.3-18.5	0.019
Dukes' stage	C (vs. B)	0.9	0.2-3.7	0.850
No. of alterations ^a	≥15 (vs. <15)	5.0	1.2-21.4	0.031
Dukes' stage	C (vs. B)	1.4	0.6-3.3	0.502
WASF (expression) ^a	Intermediate (vs. low)	5.2	1.1-24.9	0.040
	High (vs. low)	8.2	1.8-36.9	0.007
Dukes' stage	C (vs B)	1.1	0.5-2.8	0.752
MMP9 (expression) ^b	Intermediate (vs. low)	2.9	0.7-11.3	0.129
	High (vs. low)	5.6	1.5-20.2	0.008

^aOnly CIN-M tumors (n=22) were included in this analysis. ^bGene expression levels were categorized in tertiles corresponding to low, intermediate and high values.

of ≥15 alterations (ROC area under the curve: 0.741, CI 95%: 0.520-0.9629). Kaplan-Meier survival curves are shown in Fig. 2D. The differences were statistically significant in log-rank analysis (p=0.016) and multivariate Cox analysis including the Dukes' stage (Table IV). In CIN-R tumors there were no statistically significant differences (data not shown).

Gene expression profiles. For each of the 50 samples, we finally obtained a ratio that is assumed to measure the gene expression level of the CRC relative to the pool of normal tissues. We obtained the lists of the most overexpressed and underexpressed genes (data not shown). The most underexpressed genes in tumor samples were PGM5 (mean log ratio = -3.42±0.81), FUBP3 (mean log ratio = -2.91±0.95), and PMVK (mean log ratio = -2.81±0.67), and the most overexpressed were TPX2 (mean log ratio = 1.38±0.91), GNB1 (mean log ratio = 1.02±0.97) and PDCD5 (mean log ratio = 0.97±0.71).

One of the applications of large scale transcriptomic analyses is the classification of tumors with similar gene expression signatures. The platform used in this study contained 4608 different transcripts. To generate simplified signatures it was considered appropriate to work only with a subset of the genes. Genes exhibiting the highest variability are likely to have a weighty impact in tumor class classification and therefore this criteria was used for gene selection. A standard deviation >0.8 was arbitrarily chosen as cut-off point, resulting in 128 genes available for analysis. Non-supervised hierarchical clustering resulted in four main subgroups of tumors according to the similarities of their expression profiles (data not shown). No differential features in these subgroups were observed in regard to clinicopathological, molecular and genomic information available, suggesting that the resulting transcriptomic signatures did not have a direct relationship with any of the parameters considered.

Gene expression profiles and survival. Using the Cox model we identified 68 genes whose expression levels correlated with disease-free survival (p<0.01) (Fig. 3). Interestingly, 74% of associations were overexpression (50 out of 68 genes). To illustrate the association Kaplan-Meier survival curves have been traced for a subset of genes after patient classification in tertiles according to the levels of expression (data not shown). The representation includes WASF1, NFE2L2, and MMP9 (in which overexpression was associated with poor outcome) and ITGAL, TSC2, and SDF2 (in which low expression was an indicator of a shorter disease-free survival). Most of the correlations were maintained in multivariate Cox analyses including the Dukes' stage in the model. Results for WASF1 and MMP9 are shown in Table IV.

Classification of patients according to the tumor expression profile (see above) showed different survival rates, although differences did not reach statistical significance (log-rank p=0.07) (Fig. 4).

Transcriptomic versus genomic profiles. Using ANOVA tests we compared the differential expression of the 4608 genes analyzed among the groups of tumors classified according to genomic profiles (CIN-M, CIN-T, and NT). MSI tumors were not included in this analysis due to their low occurrence (n=3). No statistically significant differences were observed, suggesting that genomic and transcriptomic profiles offer non-overlapping information.

For the most recurrent chromosomal alterations, we also investigated the correlation between copy number changes as detected by CGH and the expression of the genes mapping to the involved genomic region. T-test analyses resulted in a list of genes with p-values <0.05 (data not shown). It is of note that gene expression levels paralleled chromosomal changes only in 56% of the genes, suggesting that, as a general trend,

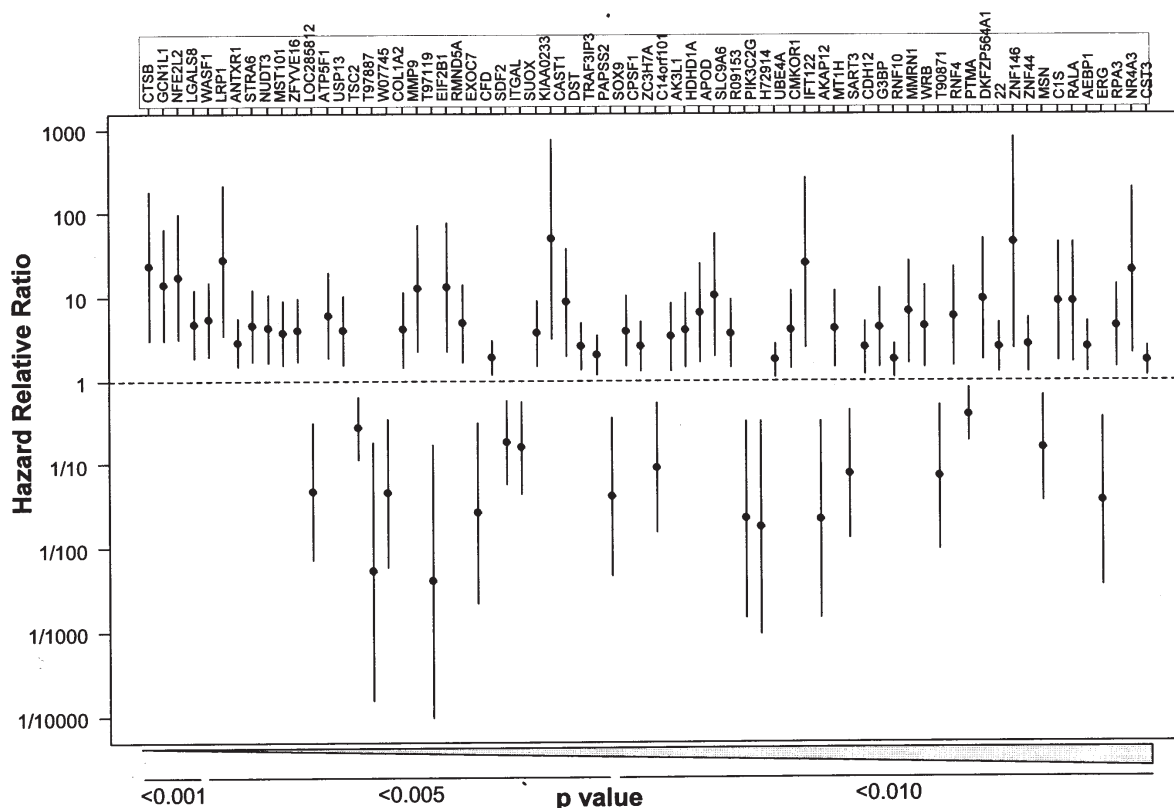


Figure 3. Univariate hazard relative ratio (mean \pm CI 95%) calculated using Cox models of gene expression ratio in regard to disease-free survival in colorectal cancer patients.

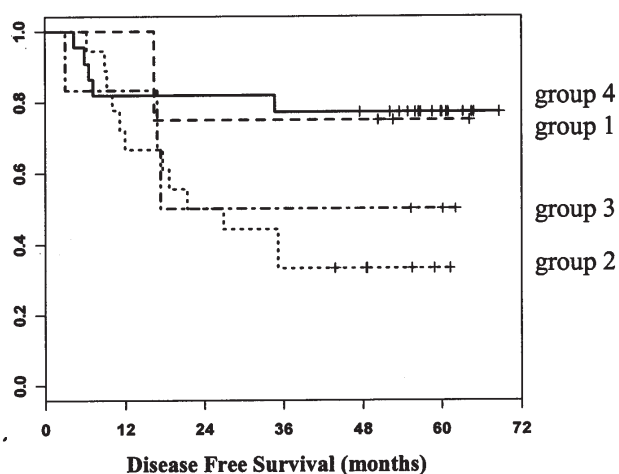


Figure 4. Kaplan-Meier disease-free survival curves for colorectal cancer patients according to gene expression clustering.

the direct effect of chromosomal copy number changes on gene expression levels is negligible.

Discussion

Colorectal cancer is probably the most studied of human solid malignancies. Besides an excellent knowledge of the genetic alterations and the molecular processes involved in carcinogenesis, this effort has not translated into a better prognostic assessment and molecular markers are not yet

included in routine clinical settings (26). Heterogeneity of the disease with multiple progression pathways, the affectation of multiple genes in each of the molecular pathways related with carcinogenesis and the involvement of genome wide disruptive mechanisms (namely, genetic and epigenetic instability) are critical factors that explain this poor outcome. The advent of various 'omic' technologies has increased expectations in the field of biomarkers. Large-scale screening strategies are more likely to generate specific markers applicable to different subsets of tumors.

Thirty to 50% of R0 colorectal cancer patients (Dukes B and C undergoing curative surgery) present recurrence of the disease. In this group of patients the staging system does not allow a precise prediction of recurrence at the individual patient level. Thus, additional prognostic factors are specially needed in this group of patients in which improvement in therapeutic and follow-up strategies may have a deeper impact on survival and quality of life. Here we have investigated genomic and transcriptomic profiles in a series of R0 Dukes B and C colorectal cancer patients. Conventional parameters are ineffectual predictors of outcome in this set of patients. At a molecular level we have identified several alterations that appear as independent prognostic markers.

At a genomic level the distribution of alterations was similar to others found in G-banding (7,27) and CGH karyotyping (data from various studies summarized in ref. 15). The alteration most clearly associated with survival in our series was gain at 11q (in most cases involving amplification at 11q13). Gain at 11q was observed in 11 tumors and was associated with recurrence of the disease (in 8 patients). In

the meta-analysis performed by Diep and coworkers (15), gains at 11q occurred in approximately 5% of primary carcinomas but were frequent in liver metastases (up to 16%), indicating that this is a late alteration that could be related with distant invasions. Amplification at 11q13 has been reported in many types of epithelial tumors (28) and is a common breakpoint in colorectal cancer (27). CCDN1 lies in this region and has been found to be amplified in different types of tumors, but does not appear to be the target in colorectal cancer (29). Therefore, other candidates, possibly EMSY (30), and RBM14 (31), deserve particular study. Losses at 12p were also associated with poor survival but only affected 4 tumors.

Unsupervised clustering of colorectal tumors according to their transcriptomic signatures often results in groups with different survival rates (18-22). Nevertheless, the signatures are quite different among the different investigations even in studies using the same microarray platform. In our analysis transcriptomic profiles also differentiated subgroups with different outcome, but bootstrap analyses indicated the low confidence of the signatures in predicting survival (data not shown).

At a transcriptomic level a large number of predictors of bad outcome were generated. Nevertheless, taking into account the high number of variables considered and the limited number of cases included in this study, the results must be taken very cautiously. A tendency toward overexpression as predictor of poor prognosis was observed. We think that this bias is probably due to the technical limitations of the approach that is likely to have a wider range of change detection for genes that overexpress than genes that underexpress, which very often may be near the limit of detection.

Many of the genes showing correlation with survival have been previously associated with malignant transformation. For instance, WASF1, whose overexpression correlates with a worse survival, plays a critical role downstream of Rac, a Rho-family small GTPase, in regulating the actin cytoskeleton required for membrane ruffling. It has been suggested that WASF1 may be critical for oncogenic transformation downstream of MAP kinase through phosphorylation in response to various external stimuli (32). NFE2L2, also overexpressed in more aggressive tumors, regulates the cellular oxidative response and constitutive activation of NFE2L2 might predispose to malignant disease in lung (33). MMP9, overexpressed in tumors of patients with poor survival, is a matrix-degrading enzyme and has been implicated in invasive cell behavior, and a study identified elevated levels of MMP9 mRNA as a predictor of relapse and short survival in colorectal cancer (34).

ITGAL and TSC2 are genes with lower expression which are associated with poorer survival. ITGAL encodes the integrin α L chain and plays a central role in leukocyte intercellular adhesion through interaction with its ligands, and also functions in lymphocyte co-stimulatory signaling. Therefore it plays a role in immune defense. Regarding cancer, it is more expressed in colorectal Dukes A and B class tumors as compared with Dukes C and D (35). TSC2 (tuberous sclerosis 2), also known as tuberlin, is underexpressed in tumors with poor survival. Its gene product is believed to be a tumor suppressor and is able to stimulate specific GTPases. The protein associates with hamartin (TSC1) in a cytosolic

complex. The TSC proteins regulate cell size control due to their involvement in the insulin signaling pathway. Furthermore they are potent positive regulators of the cyclin-dependent kinase inhibitor p27. It seems that TSC genes could trigger deregulation of the stability and localization of the tumor suppressor p27 (36).

Obviously a validation study using alternative techniques (i.e. real-time RT-PCR) and a different series of patients selected in a prospective setting should be performed to verify the usefulness of these candidate markers. This study is beyond the scope of this paper and the limitations of our setting (namely the size and the high number of candidates considered) preclude obtaining statistically significant conclusions. Future studies should address this issue appropriately.

In summary, CGH and transcriptome microarrays appear as powerful tools to uncover novel candidate predictors of relapse and shorter survival in colorectal cancer of Dukes B and C stages. Nevertheless, global signatures do not appear to constitute faithful markers of patient outcome.

Acknowledgements

We thank Rosa-Ana Risques, Gemma Tarafa and Lauro Sumoy for advice and help in microarray analysis, Gemma Aiza and Mar Muñoz for technical help. E.V. was a fellow of the Fondo de Investigación Sanitaria (FIS). This study was supported by a grant from the Ministry of Education and Science (SAF03/5821, AGL2004-07579-04, and SAF2006/351).

References

1. Jemal A, Siegel R, Ward E, *et al*: Cancer statistics, 2006. *CA Cancer J Clin* 56: 106-130, 2006.
2. Deans GT, Parks TG, Rowlands BJ and Spence RA: Prognostic factors in colorectal cancer. *Br J Surg* 79: 608-613, 1992.
3. Ransohoff DF: Rules of evidence for cancer molecular-marker discovery and validation. *Nat Rev Cancer* 4: 309-314, 2004.
4. West M, Ginsburg GS, Huang AT and Nevins JR: Embracing the complexity of genomic data for personalized medicine. *Genome Res* 16: 559-566, 2006.
5. Giaretti W: A model of DNA aneuploidization and evolution in colorectal cancer. *Lab Invest* 71: 904-910, 1994.
6. Remvikos Y, Muleris M, Salmon RJ and Dutrillaux B: Colorectal carcinogenesis: from chromosomal evolution pathways to molecular pathogenesis. *Cancer Genet Cytogenet* 93: 63-73, 1997.
7. Dutrillaux B: Pathways of chromosome alteration in human epithelial cancers. *Adv Cancer Res* 67: 59-82, 1995.
8. Ionov Y, Peinado MA, Malkhosyan S, Shibata D and Perucho M: Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* 363: 558-561, 1993.
9. Bardi G, Fenger C, Johansson B, Mitelman F and Heim S: Tumor karyotype predicts clinical outcome in colorectal cancer patients. *J Clin Oncol* 22: 2623-2634, 2004.
10. Kallioniemi A, Kallioniemi OP, Sudar D, *et al*: Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* 258: 818-821, 1992.
11. 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: 234-245, 1996.
12. 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: 623-629, 2001.
13. De Angelis PM, Stokke T, Beigi M, Mjaland O and Clausen OP: Prognostic significance of recurrent chromosomal aberrations detected by comparative genomic hybridization in sporadic colorectal cancer. *Int J Colorectal Dis* 16: 38-45, 2001.

14. Hermsen EH, Postma C, Baak J, *et al*: Colorectal adenoma to carcinoma progression follows multiple pathways of chromosomal instability. *Gastroenterology* 123: 1109-1119, 2002.
15. Diep CB, Kleivi K, Ribeiro FR, Teixeira MR, Lindgjaerde OC and Lothe RA: The order of genetic events associated with colorectal cancer progression inferred from meta-analysis of copy number changes. *Genes Chromosomes Cancer* 45: 31-41, 2006.
16. Jensen EH, McLoughlin JM and Yeatman TJ: Microarrays in gastrointestinal cancer: is personalized prediction of response to chemotherapy at hand? *Curr Opin Oncol* 18: 374-380, 2006.
17. Tinker AV, Boussioutas A and Bowtell DD: The challenges of gene expression microarrays for the study of human cancer. *Cancer Cell* 9: 333-339, 2006.
18. Croner RS, Peters A, Brueckl WM, *et al*: Microarray versus conventional prediction of lymph node metastasis in colorectal carcinoma. *Cancer* 104: 395-404, 2005.
19. Arango D, Laiho P, Kokko A, *et al*: Gene-expression profiling predicts recurrence in Dukes' C colorectal cancer. *Gastroenterology* 129: 874-884, 2005.
20. Bertucci F, Salas S, Eysteries S, *et al*: Gene expression profiling of colon cancer by DNA microarrays and correlation with histological parameters. *Oncogene* 23: 1377-1391, 2004.
21. 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* 22: 1564-1571, 2004.
22. Barrier A, Lemoine A, Boelle PY, *et al*: Colon cancer prognosis prediction by gene expression profiling. *Oncogene* 24: 6155-6164, 2005.
23. Mariadason JM, Arango D, Shi Q, *et al*: Gene expression profiling-based prediction of response of colon carcinoma cells to 5-fluorouracil and camptothecin. *Cancer Res* 63: 8791-8812, 2003.
24. Gonzalez-Garcia I, Moreno V, Navarro M, *et al*: Standardized approach for microsatellite instability detection in colorectal carcinomas. *J Natl Cancer Inst* 92: 544-549, 2000.
25. Grau M, Sole X, Obrador A, *et al*: Validation of RNA arbitrarily primed PCR probes hybridized to glass cDNA micro-arrays: application to the analysis of limited samples. *Clin Chem* 51: 93-101, 2005.
26. Weitz J, Koch M, Debus J, Hohler T, Galle PR and Buchler MW: Colorectal cancer. *Lancet* 365: 153-165, 2005.
27. Bardi G, Johansson B, Pandis N, *et al*: Cytogenetic analysis of 52 colorectal carcinomas-non-random aberration pattern and correlation with pathologic parameters. *Int J Cancer* 55: 422-428, 1993.
28. Knuutila S, Bjorkqvist AM, Autio K, *et al*: DNA copy number amplifications in human neoplasms: review of comparative genomic hybridization studies. *Am J Pathol* 152: 1107-1123, 1998.
29. Leach FS, Elledge SJ, Sherr CJ, *et al*: Amplification of cyclin genes in colorectal carcinomas. *Cancer Res* 53: 1986-1989, 1993.
30. Hughes-Davies L, Huntsman D, Ruas M, *et al*: EMSY links the BRCA2 pathway to sporadic breast and ovarian cancer. *Cell* 115: 523-535, 2003.
31. Sui Y, Yang Z, Xiong S, *et al*: Gene amplification and associated loss of 5' regulatory sequences of CoAA in human cancers. *Oncogene* 31: 31, 2006.
32. Miki H, Fukuda M, Nishida E and Takenawa T: Phosphorylation of WAVE downstream of mitogen-activated protein kinase signaling. *J Biol Chem* 274: 27605-27609, 1999.
33. Padmanabhan B, Tong KI, Ohta T, *et al*: Structural basis for defects of Keap1 activity provoked by its point mutations in lung cancer. *Mol Cell* 21: 689-700, 2006.
34. Zeng ZS, Huang Y, Cohen AM and Guillem JG: Prediction of colorectal cancer relapse and survival via tissue RNA levels of matrix metalloproteinase-9. *J Clin Oncol* 14: 3133-3140, 1996.
35. Hakansson L, Adell G, Boeryd B, Sjogren F and Sjodahl R: Infiltration of mononuclear inflammatory cells into primary colorectal carcinomas: an immunohistological analysis. *Br J Cancer* 75: 374-380, 1997.
36. Rosner M, Freilinger A and Hengstschlager M: The tuberous sclerosis genes and regulation of the cyclin-dependent kinase inhibitor p27. *Mutat Res* 613: 10-16, 2006.