Global alterations in mRNA polysomal recruitment in a cell model of colorectal cancer progression to metastasis

Alessandro Provenzani1,2, Raffaele Fronza1,2, Fabrizio Lorenzi3, Alessia Pascale4, Marialaura Amadio4 and Alessandro Quattrone1,2,5

1Fioren Foundation, Via Sacconi 6, 50019 Florence, Italy, 2Laboratory of Metabolomics and Systems Biology, Magnetic Resonance Center (CERM), University of Florence, Via Sacconi 6, 50019 Florence, Italy, 3Department of Biology, University Tor Vergata, 00133 Rome, Italy and 4Department of Experimental and Applied Pharmacology, University of Pavia, Via Taramelli 14, 27100 Pavia, Italy

Introduction

Colorectal cancer (CRC) is one of the leading causes of mortality in Western countries, and one of the preferred subjects of investigation in the molecular genetics of cancer progression. One remarkable result of this focus was the proposal by Fearon and Vogelstein (1) 15 years ago of a popular genetic model for CRC progression. This model foresees the temporary advancement from healthy mucosa to carcinoma in situ by a defined sequence of inactivating mutations in onc suppressor genes and gain of function alternations in oncogenes. Subsequent work has mapped the molecular pathways perturbed by these key lesions and identified additional genetic or epigenetic events, but in the meanwhile this core stepwise progression model has been generally validated by a number of clinical, histopathological and epidemiological studies, suggesting that it could be representative of about 90% of sporadic CRCs (2,3), all characterized by chromosome instability (CIN). The remaining sporadic CRCs are attributable to a different genetic route, initiated by alterations in DNA mismatch repair genes and characterized by a hypermutability phenotype (4) revealed by microsatellite instability (MIN).

While providing a solid genetic basis for the colorectal adenoma-carcinoma sequence, CRC molecular studies have instead generated very limited information for the subsequent stages of colorectal progression, from carcinoma in situ to invasive carcinoma to lymph node and visceral metastasis. Even the advent of high-density microarray analysis, allowing the production of transcriptome profiles of cancer specimens, has not proved until now to be really useful for the identification of specific genetic lesions promoting CRC late progression. For example, three different profiling studies conducted on primary CRC specimens of different histopathological stages provided signatures able to robustly distinguish between normal mucosa and CRCs (5,6), or...
between MIN and microsatellite stable (and therefore CIN) CRCs (5–7), or even between sporadic MIN or hereditary (hereditary non-polyposis CRC, HNPCC) CRCs (5), but were less efficient in generating primary tumour profile predictors of metastastic disease (7). An additional recent study aimed at specifically profiling CRC progression (with cases representative of all Duke’s stages and metastatic biopsies) distinguished normal mucosa from adenomas from the late stages of tumour development, but these stages clustered together (8). Moreover, microarray profiles on multiple individual tumour samples are affected by various confounding factors, among which are cell type heterogeneity due to the presence of non-tumour tissue in the sample (9), variability in the quality of the extracted messenger ribonucleic acids (mRNAs) and differences in basal gene expression due to the individual genetic variance (10). These sources of noise combine with tumour-derived alterations to produce complex patterns (11). While considering this intrinsic high variability in microarray analyses of tumour specimens, the differential resolutive power for the CRC early and late progression sequences could depend, at least in part, on limitations of a purely transcriptomic approach to gene profiling.

To address by a different perspective, that of global alterations in the control of translation, the identification of lesions promoting invasivity and metastastic spreading of CRCs, we choose to use a cellular, isogenic model of human CRC metastatic transition. The SW480 and SW620 cell lines were established from a 50-year-old CRC male patient, in a period of 6 months (12). The SW480 cells derive from the primary tumour, a poorly differentiated (grade 4) CRC invading the muscularis propria (Duke’s stage B). The SW620 cells derive from a lymph node metastasis taken from the same individual at the second laparotomy, with no intervening chemotherapy, when recurrent cancer with liver and mesenteric lymph node metastases was discovered. The extensive validation status of this model and its isogenic origin from the spontaneous progression of a human CRC render it ideal for analysis of gene expression changes in late CRC progression. The analysis was conducted by a ‘bidimensional’ transcriptomic approach, that is, we obtained a genome-wide measure not only of the level of cellular mRNAs but also of the degree of their engagement in the translation machinery at polysomes. The picture arising from this information-intensive survey is that of a profound change of the ability of metastatic cell to translate, more than to produce, cellular mRNAs, possibly related, among other lesions, to increased activation of the eIF4E translation initiation factor.

Materials and methods

Cell cultures and treatments

The human cell lines used in the study, SW480 colon carcinoma cells, and their relative lymph node metastatic SW620 cells were obtained from the European Collection of Cell Cultures (ECACC, Salisbury, UK). Cells were grown in RPMI 1640 medium with glutamine (2 mM), 10% fetal calf serum (FCS), penicillin (100 000 U/I), and streptomycin (100 mg/l). Cell cultures were maintained at 37°C in a humid atmosphere of 5% CO2 and 95% air. Cells were grown until 60–70% confluence, and the experiments performed after 4–8 passages. Puromycin (Sigma-Aldrich, St. Louis, MO), sparsomycin (Sigma-Aldrich) and cycloheximide (Sigma-Aldrich) were tested on the two cell lines with the following protocol: 10 000 cells per well were seeded in a 96-well plate; the second day after seeding the medium was changed and 0.1, 1, 10 or 100 μM of antibiotics were added. On Day 3 the cells were tested for viability (CellTiter-Glo® Luminescent Cell Viability Assay; Promega, Madison, WI). The levels of ATP was measured in a Victor3 (PerkinElmer, Wellesley, MA) multiplate reader. Experiments were run in quadruplicate; t-test (P < 0.05) was applied to validate differences.

Polysomal RNA extraction

Cells (1–2 × 10⁶) washed once with phosphate buffer saline (PBS) were treated directly on the plate with 300 μl lysis buffer [10 mM NaCl, 10 mM MgCl₂, 10 mM Tris–HCl, pH 7.5, 1% Triton X-100, 1% sodium deoxycholate, 0.2 U/μl RNase inhibitor (Fermenta, Burlington, CA) and 1 mM dithiothreitol] and centrifuged into an Eppendorf tube. After a few minute incubation on ice with occasional vortexing, the extracts were centrifuged for 10 min at 12000 g at 4°C. The supernatant was frozen in liquid nitrogen and stored at −70°C or loaded directly onto a 15–50% linear sucrose gradient containing 30 mM Tris–HCl, pH 7.5, 100 mM NaCl, 10 mM MgCl₂, and centrifuged in an SW41 rotor for 110 min at 180000 g. Fractions (polysomal and non-polysomal) were collected monitoring the absorbance at 260 nm and treated directly with proteinase K. After phenol–chloroform extraction and isopropanol precipitation, polysomal RNA was resuspended in 30 µl of water and then repurified with RNAeasy kit (Qagen, Hilden, Germany).

Soft agar cloning assay

Cells (1 × 10⁶) were resuspended in 1x RPMI 1640 plus 10% FCS in 0.7% (w/v) agarose, plated onto 60 mm dishes coated previously with 0.5% agarose, and maintained at 37°C. After 3 weeks, colonies were stained with crystal violet and manually counted. Colony assays were performed in three separate experiments.

Western blotting

Cells were lysed in Nonidet P-40 lysis buffer (20 mM Tris–HCl, pH 7.5, 150 mM NaCl, 10% glycerol, 1% Nonidet P-40, 10 mM NaF, 1 mM sodium pyrophosphate, 10 μg/ml leupeptin and aprotinin, 1 μg/ml sodium molybdate and 1 μM microcystin) containing protease inhibitors (Complete; Sigma-Aldrich). Total cell extracts were diluted in 2x SDS protein gel loading solution (Quality Biological), boiled for 5 min, separated on 12% SDS–polyacrylamide gel electrophoresis (SDS–PAGE) and processed following standard procedures. The goat polyclonal antibody anti-phospho-eIF4E (Ser-209; Santa Cruz Biotechnology, Santa Cruz, CA) was diluted at 1 : 1000; the rabbit anti-phospho-E-BPI (Thr-37/46; Cell Signaling Technology, Danvers, MA) at 1 : 3000 and the rat MAB1864 anti-α-tubulin (Chemicon International, Temecula, CA) at 1 : 1000. The nitrocellulose membrane signals were detected by chemiluminescence. Experiments were performed at least three times for each different cell preparation and α-tubulin was used to normalize the data. Statistical analysis of western blot data was performed on the densimetric values obtained with the NIH image software 1.61 (downloadable at http://rsb.info.nih.gov/nih-image).

High-density microarray analysis

Cells were harvested and washed twice with ice-cold PBS. For total RNA profiling, RNA was extracted using the RNAeasy kit (Qiagen). For polysomal RNA profiling, sucrose gradient fractions of polysomal RNA were collected and co-purified using RNAeasy kit. All the RNA samples were submitted to microarray analysis using the HG-U133A chip from Affymetrix, (Santa Clara, CA). Three biological replicates were carried out for each cell line and six hybridized arrays obtained. Biotinylated cRNA probe generation as along with array hybridization, washing and staining were carried out according to the standard Affymetrix GeneChip protocol. Fluorescence intensity for each array was captured with the Affymetrix GeneChip Scanner 3000.

Low- and high-level array data mining

Raw data were analysed using two microarray analysis software packages, dChip (13) and R- Robust Microarray Analysis (R- RMA) (14). Given the redundancy of Affymetrix probesets, their numbers correspond to an equal number of cellular mRNAs but to a minor number of cellular genes. From total RNA samples we obtained a significant signal change for 1153 probesets according to dChip and 1546 probesets according to R-RMA, of which 867 probesets (corresponding to 729 genes according to the NCBI GeneID identifier) were in common (see Supplementary Table I). For the polysomal RNA the same procedure gave 2018 common probesets (corresponding to 729 genes according to the NCBI GeneID identifier) were in common between the total RNA and polysomal RNA datasets (see Supplementary Table II). Three hundred and ninety-seven probesets (corresponding to 343 genes) are in common between the total RNA and polysomal RNA datasets (see Supplementary Table III). Polysomal RNA microarray data are deposited at the NCBI gene expression database, GEO (http://www.ncbi.nlm.nih.gov/projects/geo/), respectively, under the number GSE2509 and GSE1323. Biological over-representation analysis for the significantly changed genes on the polysomal RNA data and total RNA data was performed using EASE (15). Statistically significant categories of the ‘Biological Process’ Gene Ontology (GO) section (16) (<0.05) were grouped in clusters of the GO hierarchy in order to reduce the amount of information and summarize in a biologically meaningful way the results of the analysis. A cluster was
Quantitative real-time PCR

Quantitative real-time PCR (Q-RT–PCR) amplification reactions were carried out in triplicate. The PCR conditions were 95°C for 3 min, followed by 40 cycles of 95°C for 10 s, 60°C for 20 s and 72°C for 20 s. For each gene the comparative CT method was used according to User Bulletin # 2 (Applied Biosystems, Foster City, CA). As reference, genes cyclophilin and GAPDH were used. Seventeen genes were measured to confirm the microarray data; primer sequences are listed in Supplementary Table IV.

More detailed experimental procedures and tables are available in the Supplementary Materials and Methods section.

Results

Assessment of clonality and phenotype of the SW480/SW620 model of late CRC progression

The SW480 and SW620 cell lines represent an isogenic system of late CRC progression in which karyotypic analysis identified a shared background of common chromosomal abnormalities with few additional genomic lesions in the SW620 cell line, indicating further progression of the CIN phenotype (17,18). Phenotypically, the SW480/SW620 system has been characterized in vitro, with the measure for SW620 cells of an increased proliferation rate in serum starvation condition, of a loss of potential for TNFα- and Fas-dependent apoptosis (19,20) and of a higher chemoinvasive ability with enhanced gelatinase activity (20). In order to verify the clonality status and the differential phenotype of the SW480 and SW620 cell lines, we performed preliminary experiments of microsatellite fingerprinting, cell growth rate assessment by viability scoring and clonality potential measured by soft agar assay. Using 16 short highly polymorphic tandem repeats commonly employed in paternity exclusion tests, we obtained perfect matching between the two cell lines (data not shown), therefore confirming not only the origin from a single tumour but also the microsatellite stable nature of this tumour. In SW620 cells both the growth advantage (Figure 1A) and the tendency to autonomous clone growth was 3-fold higher (Figure 1B). These results ruled out any possible culture contamination or loss of phenotypic feature in the cells we subjected to bidimensional expression profiling.

Whole-genome profiling of primary and metastatic CRC cells shows that the majority of gene expression changes affect translation

We, therefore, performed a whole-genome comparative analysis of putative transcriptional and translational changes in the SW480/SW620 model of CRC progression. Starting from early passages of SW480 and SW620 cells, total RNA samples were extracted after cell lysis, while polysomal RNA samples were extracted after cell lysate centrifugation in a continuous sucrose gradient. Three independently grown cell cultures were used for each determination, to control for biological and instrumental variability. Each single RNA sample was subjected to microarray analysis on Affymetrix DNA chips bearing about 22 000 probesets, which interrogate collectively 14 500 human genes. Raw data were normalized and signal intensities calculated and averaged by way of two different series of widely used algorithms for microarray data treatment (implemented in the RMA and dChip packages; see Materials and methods). Given that the choice of the algorithms for microarray raw data processing can heavily bias the results either for total and polysomal RNA, we considered in the subsequent analysis only those genes that resulted as output of both series of algorithms to be significantly changed (P < 0.01). Surprisingly, while the genes changed at the mRNA level were 867, those altered for the polysomal occupancy of their mRNAs in the same conditions of analysis were 2018, almost 2.5 times more (Figure 2A). The two datasets were equally distributed in upregulated and downregulated genes (respectively, 1080 versus 968 in the polysomal mRNAs and 476 versus 391 in the cytoplasmic mRNAs). The genes changed purely at the mRNA level were 470, whereas those varied only for their polysomal occupancy were 1621. The number of genes varied in both datasets was 397, which, as...
respective. The solid line indicates a slope of 1.

Expected, displayed a very strong correlation in the sense of variation (Figure 2B), the genes discordant for the sign of change being only 16 (4%) and clustered in the central portion of the graph, where the variations are small (<2-fold change). For the 381 genes that were concordantly up- or downregulated, the linear fit (Figure 2B) gives a slope of around 0.5 in both directions, with meaningful correlation coefficients ($r = 0.63$ for the positive variations and $r = 0.69$ for the negative variations, $P < 0.0001$). This means that for the genes in which a differential expression between the two cell lines was detectable at both expression levels, the magnitude of the expected change was, on average, 4-fold greater at the translational than at the transcriptional level.

Transitionally silent changes of mRNA abundance are not expected to affect protein quantity. Therefore, in our case, the contribution to proteome composition should come only from the 2018 genes that underwent a differential polysomal loading, probably resulting in a change in the levels of produced proteins. Remarkably, since less than 400 transcriptional gene expression changes are also detected translationally, in our model of CRC late progression a classical microarray analysis would have missed more than 80% of the total expression variation responsible for proteome remodelling.

**Transition from primary to metastatic cells is associated with gene signatures specific of the level and the direction of gene expression alteration**

From the scoring of mRNA differential abundances in the whole-cell and polysomal compartments of the two cell lines we conclude that the major contribution to diversity is, unexpectedly, given by changes at the translational level. To give to these numbers the meaning of estimated biological activities, we annotated the polysomal and total RNA datasets of changed genes using the ‘biological process’ ontology of the GO, a system of hierarchical, controlled gene annotation suitable of quantitative analysis (16). Calculation of over-representation for each node of the ontology (15) allowed an unbiased, statistically controlled estimation of the cell functions that could be related to the differential behaviour of the primary and metastatic cell lines. Of the 76 significantly ($P < 0.05$) altered categories in the biological process ontology 14 were common to the two datasets, 26 emerged only in the polysomal RNA dataset and 36 were selective of the total RNA dataset (see Supplementary Table V). When clustered in the ontology graph to unravel the internal relationships, these categories showed a strong polarization toward cell cycle-related themes for both datasets (9 out of the 14 common categories), while the exclusively transcriptional changes were polarized toward purine nucleotide biosynthesis (16 out of 26 categories). Categories emerging as statistically significant from the translationally altered genes were prevalent in number and clustered in defined general biological themes. To better understand the biological impact of these translationally altered genes we performed the same ontological analysis dividing this group in up- and downregulated mRNAs. The emerging themes were RNA transport and metabolism, amino acid activation for translation and translation itself for the genes with increased polysomal occupancy, and programmed cell death for the genes with decreased polysomal occupancy. Figure 3A reports all these data as a function of the number of over-represented categories. We also performed an extensive real-time quantitative RT–PCR validation on the polysomal loading changes of 17 genes chosen between the above-mentioned ontological categories for their biological interest and detailed annotation status. As reported in Figure 3B, all the PCR validations confirmed the sense of change with only some difference in the magnitude of change. NGFR and FAS are members of the TNF/NGF receptor superfamily, and are both translationally downregulated in the metastatic clone. The expected decrease of signalling activity could be associated phenotypically with its documented decrease in activation-induced apoptosis in this clone (20). The CYR61 and RHOB genes, again markedly decreased in the translational fitness of their corresponding mRNAs, are both involved in processes of integrin adhesion and signalling, and therefore their translational downregulation could contribute to the loss of matrix anchorage typical of metastatic cells. Instead, the decrease in translational activity of CCND1, CCND2 and EGFR is not easily explainable in a contest of increased tumour
aggressiveness, even if downregulation of these genes in cancer specimens have been already documented in some cases.

We also verified for two of the validated changes in translational fitness, those for the NGFR and CYR61 genes, the correspondence with the steady-state levels of the corresponding proteins, detected by western blotting. The concordance suggests that polysomal profiling could represent an affordable quantitative projection of the proteome. A remarkable observation is also that among the genes whose expression change is purely due to a change in translational fitness (variation in polysomal loading without variation in total mRNA content) we find several genes, like VEGF, ribonucleotide reductase, C/EBP and FOS, that are known to be cancer-related and demonstrated to be susceptible of regulation at the level of translation initiation.

Considering the overall remarkable clusterization of inferred biological activities from the change in mRNA levels in the examined population of genes, and given the good level of phenotypic characterization of our model, we looked for matching phenotypic behaviour. The cell cycle signature, arising from changes in the transcriptome and conserved among polysomal changes, was paralleled by the described increased cell growth kinetic of SW620 cells (see Figure 1A), while the cell death signature characterizing polysomal downregulated mRNAs could match with the cited differential apoptotic behaviour in response to factors belonging to the TNF family (20).

The polysomal mRNA upregulation signatures of mRNA metabolism, translation and amino acid activation all point broadly to the cell activity of post-transcriptional regulation of gene expression, and suggest that the increased translational fitness of selected genes could be a major lesion driving late CRC progression. Amino acid activation resulted to be an especially interesting biological theme, since the over-represented ‘tRNA aminoacylation’ node of GO represents the convergence of two over-represented chains of nodes relating to tRNA metabolism and amino acid activation, the two processes leading to the formation of aminoacyl-tRNAs ready for translation. Moreover, in real-time RT-PCR experiments performed on selected genes to validate the microarray results (Figure 3B), members of the aminoacyl-tRNA synthetase complex appeared to be consistently and strongly upregulated. To look for a phenotypic correlate of this signature, we exploited the known differential mechanisms of activity of antibiotics targeted to the eukaryotic ribosome. Both cycloheximide and sparsomycin block translational elongation by acting at the P site of the ribosome—the first drug by inhibition of elongation factor-dependent translocation (21) and the second drug by stabilization of the interactions at the P site (22). On the other hand, puromycin, which structurally resembles an aminoacyl-tRNA, directly competes with the aminoacyl-tRNA loading at the A site (23). Therefore, a differential availability of aminoacyl-tRNAs between the two cell lines could affect primarily puromycin action. Dose-response curves in Figure 4 show that the two cell lines have indeed a growth inhibition profile indistinguishable for cycloheximide and sparsomycin but clearly different for puromycin, which is 5-fold more effective in SW620 cells at 10 and 100 μM. The increased susceptibility of SW620 cells to puromycin could be due to a decreased availability of free aminoacyl-tRNAs, recruited at the polysomes by an overall increased translational activity.

**Transition from primary to metastatic cells is associated with increased phosphorylation of eIF4E and 4E-BP1**

We therefore looked for possible changes of activity in the translational initiation machinery, which could account for the wide number of differentially translated mRNAs
between the two cell lines. Cap-dependent translation is mostly regulated at the level of initiation, the formation of the mRNA-loaded 48S ribosomal complex being the major point of involvement in human cancer (24). This route of translation initiation starts with the binding of the eIF4E factor to the mRNA cap structure and its subsequent association with the docking eIF4G1 or eIFG2 protein, which binds to the eIF4A RNA helicase to form the eIF4F complex with the mRNA (25). Phosphorylation of eIF4E is believed to be associated with eIF4F activity, even if its functional effects are uncertain (26), and could represent a point of regulation. The major regulative mechanism of eIF4F formation and function depends on the action of three eIF4E binding proteins (4E-BPs), designated 4E-BP1, 4E-BP2, and 4E-BP3 (27), where the predominant species is 4E-BP1. Each 4E-BP competes with each eIF4G for eIF4E using the same eIF4E binding motif, and inhibits formation of eIF4F by sequestering eIF4E in a translationally inactive complex (8). Hyperphosphorylated 4E-BP1 has a decreased affinity for eIF4E, resulting in its liberation and the formation of eIF4F. We therefore checked the SW480 and SW620 cells for the abundance of various phosphorylation variants of 4E-BP1, by western blotting with polyclonal antibodies recognizing the isoform phosphorylated at least in Thr37 and Thr46 or the full spectrum of different isoforms. We also checked for total and phosphorylated (Ser209) eIF4E. The results, reported in Figure 5A for paradigmatic gels (top) and for densitometric quantification (bottom), show that while the general tendency is of content increase in the metastatic versus the primary tumour clones, this tendency reaches statistical significance ($P < 0.05$) only in the phosphorylated forms. Since the increase at the protein level for the total eIF4E and the bona fide non-phosphorylated 4E-BP1 (alpha lane in the gel) is not significant, we carried out a real-time PCR analysis of the polysomal mRNA for eIF4E and 4E-BP1. Figure 5B shows that the amounts of the two translated mRNAs, with respect to the two housekeeping genes GAPDH and cyclophyllin, are equal in the two cell lines, confirming the absence of a substantial increase in the protein content. Therefore, the differential degree of phosphorylation should derive solely from upstream differential signalling activities. While the functional meaning of eIF4E phosphorylation is still in doubt, a body of evidence suggests that sequential increase in the 4E-BP1 phosphorylation content results in release of eIF4E and translational activation.

Translationally changed genes in metastatic cells have 5'-UTRs longer than transcriptionally changed genes

From the above results, the translational changes emerging from the comparison of polysomal profiles between the two cell lines could at least partly be due to altered signalling to the eIF4F complex, resulting in altered differential loading of mRNAs into the ribosome. The formation of eIF4F and the initiation of translation are believed to depend on the structure of the mRNA 5'-UTR. In particular, the 5'-UTR length of mRNAs is considered a determinant of translational fitness, because mRNAs with long, structured 5'-UTRs are expected to compete less effectively for translational initiation than mRNAs with shorter, structurally simpler, 5'-UTRs (28,29), even if conclusive evidence for this general rule is lacking. We therefore examined the distribution of the 5'-UTR lengths of the genes in the total and polysomal RNA datasets. We derived the annotated 5'-UTRs of all the human mRNAs from the assembly 16 of the human genome sequence, which after reduction of redundancy appeared to be 15 596. Seventy-five per cent of the genes of the total dataset and 70% of the genes of the polysomal dataset could be associated with an annotated 5'-UTR. In particular, the 5'-UTR length of mRNAs is considered a determinant of translational fitness, because mRNAs with long, structured 5'-UTRs are expected to compete less effectively for translational initiation than mRNAs with shorter, structurally simpler, 5'-UTRs (28,29), even if conclusive evidence for this general rule is lacking. We therefore examined the distribution of the 5'-UTR lengths of the genes in the total and polysomal RNA datasets. We derived the annotated 5'-UTRs of all the human mRNAs from the assembly 16 of the human genome sequence, which after reduction of redundancy appeared to be 15 596. Seventy-five per cent of the genes of the total dataset and 70% of the genes of the polysomal dataset could be associated with an annotated 5'-UTR. Figure 6A shows the distribution of the 5'-UTR lengths in both datasets, superimposed with the smoothed curves derived by calculating running averages. The corresponding mean and median values of the distributions are 171 and 104 for the total RNA dataset and 206 and 130 for the polysomal RNA dataset, respectively, indicating a differential composition of the lengths of the 5'-UTRs. To test this divergence for statistical significance, we calculated the difference between the two normalized distributions, which
gave rise to the curve in Figure 6B. We then generated a confidence interval (CI) obtained by permutation analysis of 10,000 randomly generated distributions generated by sampling all the genomic 5'-UTRs obtained, superimposed as the dashed area in Figure 6B. For short 5'-UTR lengths (up to about 100 bases) the polysomal dataset were depleted of 142 mRNAs, which are distributed among longer 5'-UTRs. In the polysomally altered genes the ratio between the number of mRNAs with 5'-UTRs >100 and <100 bases is 0.63; on the other hand the same ratio for genes altered in the total mRNA content is 0.91. This difference is statistically significant, and means that for genes with 5'-UTRs longer than 100 bases and up to 300 bases derangement in the translational control is the prevalent reason of their change in expression. No bias for these genes is present with respect to the sense of variation of ribosomal loading, that is, the number of polysomally upregulated and downregulated mRNAs is about equal. Therefore, in agreement with a differential activity of the eIF4F complex, the 5'-UTR length appears to be an mRNA feature influencing directly the level at which gene expression deregulation occurs during CRC late progression.

**Discussion**

We have applied a systems biology approach to the whole-cell characterization of gene expression changes occurring during the invasive and metastatic spread of CRCs. The nature of the technologies used and the need of minimizing the instrumental and biological variability impose the use of a cellular model. Obviously, the generality of the results emerging from the characterization of this model depend on its degree of representation of the common events involved in sporadic CRC late progression. Several indications suggest this degree to be high. The Fearon and Vogelstein proposal for the colorectal adenoma-carcinoma transition imply a sequence of genetic lesions, starting with early inactivation of the adenomatous polyposis coli (APC) tumour suppressor, followed by oncogenic KRAS mutations arising in the adenomatous stage, followed by TP53 inactivation and chromosome 18q deletions for
malignancy onset. The SW480 cells recapitulate all these key lesions: APC truncation of one allele and loss of heterozygosity of the other allele (30–32), KRAS activating mutations in both alleles (30,33), inactivating TP53 mutation in one allele and loss of heterozygosity in the other (30,34). Moreover, these cells are microsatellite stable (30), and so perfectly represent the canonical sporadic CIN CRC, blocked at the stage of invasive tumour. SW620 cells, clonally strictly related to SW480 cells, are taken from a lymph node metastasis of the same individual, and therefore truly reflect the natural history of CRC progression, as opposed to the experimental metastasis models, derived either by in vitro or in vivo cell selection, which often give misleading results (35). SW620 cells indeed retain an increased CRC metastatic potential, which has been assessed in vivo with the production of 9-fold more liver metastases after splenic injection (19) and 5-fold more lymph node metastases (20) after subcutaneous injection than the SW480 cells in two different nude mice xenotransplantation models. If, therefore, these cells are bona fide examples of an invasive-metastatic CRC sequence, this sequence is, from our analysis, heavily characterized by alterations at the translational level of gene expression. We measured translational alterations by the change in mRNA occupancy in polysomes, which is considered an indirect indication of the degree synthesis of the relevant protein better describing the proteome than cellular mRNA quantification (36–38). Moreover, use of the same genome-wide quantification technique for both total and translated mRNA populations assures the same level of resolution, sensitivity and error rate, allowing meaningful comparisons. In this framework, our analysis shows that 80% of the genes undergoing a gene expression change in the transition between SW480 and SW620 cells do it by varying their degree of polysomal loading, implying a dramatic subversion in the signalling control of translation and/or in the translational machinery itself. These changes are intrinsically undetectable by a standard transcriptome profiling approach. If they are representative of CRC progression, this could be a cause of the described lack of discriminative power for the microarray-based clustering of different CRC stages. An implication of this possibility would be a lack of correlation between transcriptomic and proteomic data in the same samples. Indeed, in the first attempt of simultaneous assessment of mRNA and proteins in CRC specimens representing all stages of progression (8), a dramatic difference between mRNA and the corresponding protein levels was found (26% of positive correlations toward 74% of non-significant or negative correlations). Even poorer results (17% of correlation) were obtained in a similar transcriptomic/proteomic survey in lung adenocarcinoma (39). These disappointing results strongly favour our proposal that translational mechanisms can be heavily involved in the genetic perturbation underlying CRC progression and cancer development in general, and highlight the need of a bidimensional (36) transcriptome profiling of gene expression. Attempts to obtain more efficient and high-throughput polysomal separations (40) combined with mRNA amplification procedures (41) should help in facing the technical drawbacks preventing the application of this approach to tumour specimens, and would allow epidemiological validation of the proposed model.

The contribution of transcriptional movements to the overall reshaping of gene expression is represented in our analysis by 35% of genes, of which 15% are ‘reasorbed’ by the absence of a corresponding change in the level of mRNA association to polysomes. As also a validation of the good technical quality of the analysis, all genes that undergo a change at both levels are tightly concordant in the sense of change, a phenomenon already noted in a combined transcriptional/translational analysis of induced gene expression changes in yeast (42), but never described for alterations deriving from ongoing chromosomal instability in cancer. In addition, for these mRNAs the shift in polysomal engagement is always of a higher magnitude than the variation in cellular mRNA content, as a sort of translational amplification of an induced change in transcriptional activity. This phenomenon awaits further study. A statistical analysis of over-representation of biological themes uncovers a marked polarization towards cellular programs, which is specific for the groups of commonly changed mRNAs in the cellular and polysomal compartments and for the polysomal upregulated and downregulated mRNAs. The protein products of the 400 homodirectionally changed mRNAs are involved mostly in cell proliferation, which therefore results to be the cell activity interested by lesions affecting synergistically both levels of gene expression regulation, probably reflecting a strong selective advantage in the transition to the successful metastatic clone from the invading tumour mass. The suppression of the transcriptional signature of nucleotide metabolism exerted at the translational level, due to the fact that 470 genes clearly changed in the total mRNA levels do not display any variation in terms of ribosomal loading, is particularly interesting, again in terms of selective advantage. The clustering in functions related to the biosynthesis of purine nucleotides is very high in this group of mRNAs, which could mean that CRC progression tends to suppress by a compensatory translational adjustment transcriptional alterations resulting by themselves in death or impaired growth of the bearing cancer cell clones. If of general interest, this mechanism of coherent interplay between transcriptional and translational regulation in the course of the natural history of a tumour could be, added to the homodirectional changes observed for cell cycle genes, the indication of a new level of cooperativity in cancer genetics. In other words, to shape the oncoproteome tumour progression could operate not only by combining synergistically different oncogenic and oncosuppressor lesions but also by combining the rate of nuclear production of transcripts and their translational fitness. The apoptotic signature present in the translationally downregulated mRNAs could match the phenotypically increased resistance of the metastatic clone to apoptosis induced by members of the TNF family, even if a weight of the pro-apoptotic and the anti-apoptotic contribution to the signature is unfeasible; in the same direction of genotypic/phenotypic matching the increase in translational fitness in components of the tRNA aminoacylation pathway could be related to the enhancement of SW620 cells sensitivity to puromycin. These correlations are always highly hypothetical and represent the fundamental limitation in the power of analysis of microarray-based studies; indeed, polysomal profiling is expected to produce a better description than transcriptome analysis of the phenotypic features of a cell (43). The tRNA activation pathway is part of a general signature of translational control and RNA metabolism extracted from the mRNAs that increase their translational fitness despite the absence of significant changes in cellular levels. Since it is known that many genes controlling translation are themselves regulated in expression at the translational level, this could represent a positive feedback loop contributing to the progressive derangement of the
translational machinery in the course of tumour progression: alteration in the polysomal mRNA occupancy of a translation factor drives alteration of the protein availability of this factor, which impairs the translational machinery and therefore the polysomal mRNA occupancy of other mRNAs of translation factors, and so on. For this process to take place, no transcriptional change is in principle necessary.

To our knowledge, the only other cancer-related study involving simultaneous cellular and polysomal mRNA assessment by microarray profiling was conducted on a model of mouse glial progenitor cells transformed by chronic activation of the Akt and RAS pathways, to mimic common signalling alterations present in human glioblastomas (44). Akt and KRAS are known to be part of the two fundamental signalling pathways controlling translation—the first through the rapamycin-sensitive PI3K/mTOR cascade to finally induce hyperphosphorylation of 4E-BP1, and the second to promote phosphorylation of eIF4E through a Raf/MEK/ERK way (25). Combined KRAS and Akt activities in glial progenitors indeed promoted these final phosphorylation events, and were pharmacologically blocked to study their effects on the cellular and polysomal mRNA levels. The authors found a few mRNAs that were differentially affected at the cellular mRNA level by blockade of the two cascades, but hundreds of mRNAs that changed their polysomal occupancy. Given the similarity of these results with those coming from our description of the perturbations induced by CRC progression, we evaluated the overall levels and degree of phosphorylation of eIF4E and 4E-BP1 and found that the transition from SW480 to SW620 cells is associated with an increase in the phosphorylated fractions of both proteins, without significant increase in the total protein content measured both by western blotting and ribosome-loaded corresponding mRNAs. eIF4E is a factor already known for being deregulated in a wide variety of human cancers (45), and for being involved in metastatic onset (28). Phosphorylation of eIF4E has recently been shown to enhance transformation in cell culture (46), even if there is controversy on the significance of this event with respect to eIF4E activity and downstream translation (26), while in CRC specimens eIF4E immunoreactivity is increased in the adenoma-carcinoma sequence (47,48). High levels of 4E-BP1 in gastrointestinal tumours correlate with the absence of lymph node and distant metastasis (49), and its hyperphosphorylated status, still awaiting evaluation in epidemiological studies, has been convincingly associated with tumourigenesis and maintenance of the transformed state by the use of phosphorylation site mutants in breast carcinoma cells (50). The point in our analysis is whether the imbalances we found in the phosphorylation state of eIF4E and 4E-BP1 in the SW480/SW620 model of late CRC progression could be associated to the genome-wide translational deregulation we have detected by a microarray-based bidimensional mRNA profiling approach. In this study we provide evidence in favour of this possibility. The eIF4F activated complex acts on the 5'-UTRs of eIF4E-bound mRNAs by scanning them and unwinding secondary structures to finally reveal the initiation codon, trigger polysomal engagement and promote translation (51). For the relatively low abundance of eIF4E and its being sequestered by eIF4BPs and other proteins (52), cellular mRNAs must compete for eIF4E binding. A popular, albeit not formally proven, model dictates that at low level of eIF4F activity the engagement of eIF4F in translational initiation is expected to favour mRNAs endowed with short, unstructured 5'-UTRs like those of many housekeeping genes, while only at high eIF4F activity levels mRNAs with long and structured 5'-UTRs, like those of many proliferation-related genes, would be efficiently translated (28). Following this scheme, we performed an mRNA population analysis between primary tumour and metastatic cells, demonstrating that translational loading of mRNAs is more likely to be changed than transcriptional activity in a window of relatively high (between 100 and 300 bases) 5'-UTR lengths. This relationship, established on population statistics basis, is expected to be modulated for individual mRNAs by additional lesions in specific mRNA cis sequences, trans factors and proteins involved in the translational machinery. For example, we have detected and verified in our translational signature an overexpression of components of the amynoacyl-tRNA synthetase complex, which could cooperate with an enhanced mRNA loading on the ribosome to favour translation of mRNAs with complex 5'-UTRs. Moreover, it is known that eIF4E cooperates with c-myc to induce tumours in transgenic mice (53), which could be possibly related to the newly discovered function of c-myc in activating RNA polymerase I transcription of ribosomal genes (54–56). In this context, our preliminary observation of a ribosomal gene expression imbalance in the same cellular model of CRC late progression (A.Provenzani and A.Aquitrone, unpublished data) could be of significance.

Overall, from our bidimensional transcriptome analysis it emerges that in the transition from the SW480 invasive carcinoma cells to the SW620 metastatic lymph node cells the CRC genome acquires changes impacting by far the translational more than the transcriptional control of gene expression. These changes would have been completely missed in a conventional transcriptome analysis detecting cellular mRNAs irrespective of their degree of polysomal loading, which emphasizes the value of this new approach in the systems-biology-based analysis of cancer progression. These changes can be related, at least partially, to a progressive activation of the eIF4F complex, and to the consequent massive changes in translational fitness of mRNAs involved in the metastatic phenotype onset. Certainly, other cooperating lesions are active in this cell mode, as well as in the natural history of CRC spreading, whose impact on translational deregulation should be, from this and other studies, considered as a primary subject of investigation.

Supplementary material
Supplementary material is available at: http://www.carcin.oxfordjournals.org/

Acknowledgements
We thank Drs Francesca Buccioliero and Irene Bova for their kind help with the experiments and the Genexpress facility (University of Florence) for use of the instrumentation. This work was supported by a grant from Ente Cassa di Risparmio di Firenze to FiorGen Foundation. We also thank the Associazione Sandro Pitigliani per La Lotta contro i Tumori for its support. A.P. and R.F. are recipients of two fellowships from FiorGen Foundation.

Conflict of Interest Statement: None declared.

References


*Received August 24, 2005; revised January 25, 2006; accepted February 19, 2006*