Subnuclear Proteomics in Colorectal Cancer

IDENTIFICATION OF PROTEINS ENRICHED IN THE NUCLEAR MATRIX FRACTION AND REGULATION IN ADENOMA TO CARCINOMA PROGRESSION*§

Jakob Abrethsen‡, Jaco C. Knol‡, Sander R. Piersma‡, Thang V. Pham‡, Meike de Wit§, Sandra Mongera§, Beatriz Carvalho§, Henk M. W. Verheul‡, Remond J. A. Fijneman§, Gerrit A. Meijer§, and Connie R. Jimenez‡¶

Abnormalities in nuclear phenotype and chromosome structure are key features of cancer cells. Investigation of the protein determinants of nuclear subfractions in cancer may yield molecular insights into aberrant chromosome function and chromatin organization and in addition may yield biomarkers for early cancer detection. Here we evaluate a proteomics work flow for profiling protein constituents in subnuclear domains in colorectal cancer tissues and apply this work flow to a comparative analysis of the nuclear matrix fraction in colorectal adenoma and carcinoma tissue samples. First, we established the reproducibility of the entire work flow. In a reproducibility analysis of three nuclear matrix fractions independently isolated from the same colon tumor homogenate, 889 of 1,047 proteins (85%) were reproducibly identified at high confidence (minimally two peptides per protein at 99% confidence interval at the protein level) with an average coefficient of variance for the number of normalized spectral counts per protein of 30%. This indicates a good reproducibility of the entire work flow from biochemical isolation to nano-LC-MS/MS analysis. Second, using spectral counting combined with statistics, we identified proteins that are significantly enriched in the nuclear matrix fraction relative to two earlier fractions (the chromatin-binding and intermediate filament fractions) isolated from six colorectal tissue samples. The total data set contained 2,059 non-redundant proteins. Gene ontology mining and protein network analysis of nuclear matrix-enriched proteins revealed enrichment for proteins implicated in “RNA processing” and “mRNA metabolic process.” Finally, an explorative comparison of the nuclear matrix proteome in colorectal adenoma and carcinoma tissues revealed many proteins previously implicated in oncogenesis as well as new candidates. A subset of these differentially expressed proteins also exhibited a corresponding change at the mRNA level. Together, the results show that subnuclear proteomics of tumor tissue is feasible and a promising avenue for exploring oncogenesis. Molecular & Cellular Proteomics 9:988–1005, 2010.

Clinical proteomics aims to comprehensively identify and quantify proteins in human samples to gain insight into cellular pathways in disease and to discover novel biomarkers for screening, early detection, diagnosis, prognosis, and prediction of response to specific treatments. However, a fundamental challenge in clinical proteomics is to overcome the vast dynamic range and complexity of biological proteomes (1, 2). Some proteomics approaches utilize a combination of multidimensional fractionation and/or depletion of abundant proteins for improved protein coverage. An alternative proteomics strategy is to focus on subproteomes of particular interest, e.g., a subcellular fraction. Subcellular fractionation of tissue and cells in combination with MS/MS has proven to be a powerful approach for the identification of proteins contained in specific organelles, such as the nucleus and mitochondria, as well as other subcellular structures, such as vesicles and the cytoskeleton (3, 4). Combining MS/MS and subnuclear fractionation provides an even more focused view of the proteomes in selected nuclear subdomains, such as the nuclear envelope (5), the nucleolus (6), and the nuclear matrix (7).

Genomic instability is an important feature of colon adenoma to carcinoma progression. About 85% of Colorectal cancer (CRC)¹ cases exhibit chromosomal instability (CIN),

¹ The abbreviations used are: CRC, colorectal cancer; CB, chromatin-binding; CV, coefficient of variance; FDR, false discovery rate; GO, gene ontology; IF, intermediate filament; MAR, matrix-associated region; NM, nuclear matrix; PPI, protein-protein interaction; PTM, post-translational modification; MIN, microsatellite instability; CIN, chromosome instability; 2DGE, two-dimensional gel electrophoresis; MIN +, MIN-positive; CIN +, CIN-positive; MLPA, multiplex ligation-dependent probe amplification; PIPES, 1,4-piperazineethanesulfonic acid; Bis-Tris, 2-[bis(2-hydroxyethyl)aminol]-2-(hydroxyethyl)-propane-1,3-diol; LTQ, linear trap quadrupole; IPI, International Protein Index; STRING, Search Tool for the Retrieval of Interacting Genes/Proteins; hnRNP, heterogeneous nuclear ribonucleoprotein; VWF, von Willebrand factor; MNDA, myeloid cell nuclear differentiation antigen; VCAN, versican V1; HMGB3, high mobility group protein B3; PDLIM, PDZ and LIM domain protein; PLL, periplakin; EVPL, envoplakin; NUP, nuclear pore complex protein; FUBP, far upstream element-binding protein; LTBP, latent transforming growth factor β-binding protein; CHMP, charged multivesicular body protein; PSPC1, parapelle protein 1; PURA, Pur-α; NUP1L1, nucleoprophobe-4-like 1 isoform c; TAF15, TATA-binding protein-associated factor 2N; LAMC1, laminin gamma 1; VCL,vinculin; NDF, neuregulin 1; hnRN-PUL1, heterogeneous nuclear ribonucleoprotein U-like 1;
whereas 15% of CRC cases exhibits microsatellite instability (MIN) (8). Progression of non-malignant precursor lesions, colorectal adenomas, into CIN+ CRC has been associated with a number of specific chromosomal alterations of which gain of additional copies of chromosome 20q is most prominent (9, 10). Tumor progression is accompanied by important phenotypic changes in neoplastic cells, in particular alterations of nuclear structure such as nuclear size and shape, numbers and sizes of nucleoli, and chromatin texture as observed by the pathologist under the microscope (9, 11, 12). Yet, still little is known about the protein constituents underlying these phenotypic characteristics (12). As such, proteins supporting nuclear structure and organization represent a relevant target for proteomics analysis of cancer tissue and precursor lesions. Here we evaluate a proteomics work flow for profiling of subnuclear fractions from early stage CRC tumors and apply this work flow to a comparative analysis of the nuclear matrix fraction of colorectal adenoma and carcinoma tissues.

A focused proteomics work flow for analysis of patient tumor tissue should effectively separate subproteomes of interest from interfering and highly abundant proteins and other tissue components. In addition, the method must be practical, reproducible, and compatible with relatively small amounts of tissue. Here we investigate a biochemical protocol that originally was developed by Fey and co-workers (13, 14) in the 1980s. In brief, tissue samples are homogenized in a solution with stabilizers of RNA-protein complexes, and soluble, cytoskeletal, and DNA-associated proteins are extracted. The remaining detergent-, salt-, and DNase-resistant pellet is cleared of filaments and is referred to as the “nuclear matrix” (NM) fraction. The NM has been postulated to be a dynamic RNA/protein network that spans the nucleoplasm and to provide support for higher order chromatin packaging and overall shape of the nucleus (15). Several lines of evidence support roles for the NM in the processing of genetic information and for establishing functional subdomains and macromolecule assemblies in the nucleus (15). Oncogenes or suppressor genes can function differentially in different cell types, and much of this cellular context for transformation may be determined, among others, by the NM (16). Interestingly, early studies using two-dimensional gel-based proteomics have indicated that the NM proteome varies significantly between different tissues and cells and in cancer versus healthy tissue (for reviews, see Refs. 7 and 17). However, only a few cancer-related NM proteins have been identified so far. A comprehensive LC-MS/MS-based proteomics analysis of the NM proteome in adenoma-carcinoma progression may shed light on the potential role of the NM in (tissue-specific) cancer pathogenesis; in particular, the NM proteins have been suggested to provide a promising source of protein biomarkers for cancer (17, 18). The NMP22/Bladdercheck (Matritech) test measures urine levels of the NM-associated protein nuclear mitotic apparatus protein (NUMA) and has been United States Food and Drug Administration-approved for follow-up measurements in bladder cancer patients (17).

In addition to the NM fraction, two hitherto uncharacterized fractions are also extracted with this protocol: the “chromatin-binding” (CB) and the “intermediate filament” (IF) fractions. Using differential analysis of the three fractions, we show that the NM fraction is enriched for proteins derived from distinct subnuclear domains. Importantly, this subcellular fractionation approach was reproducible, and a differential comparison of the nuclear matrix proteome in colorectal adenoma and carcinoma tissues uncovered many proteins previously implicated in cancer biology as well as novel candidates. Therefore, the proteomics work flow presented here provides relevant targets for comprehensive profiling of proteins involved in subnuclear structure and organization in early stage CRC tumors and is a promising avenue to explore oncogenesis.

**EXPERIMENTAL PROCEDURES**

**Materials**

All basic chemicals were obtained from Sigma. HPLC solvents, LC-MS grade water, acetonitrile, and formic acid were obtained from Biosolve (Valkenswaard, The Netherlands). Porcine sequence grade modified trypsin was obtained from Promega (Leiden, The Netherlands). Complete protease inhibitor mixture tablets (mini, EDTA-free) were purchased from Roche Diagnostics. Protein assay reagent was obtained from Bio-Rad. Precast gradient gels and gel buffers were acquired from Invitrogen.

**Clinical Material**

CRC tissue was collected at the VU University Medical Center, Amsterdam, The Netherlands, in compliance with institutional regulations for use of leftover material. The tissue samples (50–200 mg) were obtained from surgical resection specimens at the department of pathology, snap frozen in liquid nitrogen, and stored at −80 °C. All samples were inspected by a pathologist to ascertain a tumor content of at least 70% cancer cells using hematoxylin- and eosin-stained sections. Biological subtypes of carcinomas (MIN+ versus CIN+) were determined to include both carcinoma types in the analysis. In brief, copy number changes were determined by MLPA as described by Schouten et al. (19). To this end, we used previously designed probes for chromosomes arms 8q, 13q, and 20q (20). MLPA data analysis was performed using MLPAnalyzer version 7.0 (21). MIN analysis was performed using the MIN Analysis System (MSI Multiplex System Version 1.1, Promega) as described (22). Clinical details are described in supplemental Table 1.

**Biochemical Fractionation**

The tissue was prepared essentially by the protocol originally developed by Fey and co-workers (13, 14). The homogenization and fractionation were performed in a cold room (4 °C) and on ice. First, the frozen tissue was placed in a Petri dish with −1 ml of liquid nitrogen and cut into small pieces (−1 mm²) with a precooled razor
blade. The frozen tissue pieces were immediately placed in a precooled Potter-Elvehjem tube containing 600 μl of precooled homogenization buffer/100 mg of tissue (0.5% Triton X-100, 300 mM sucrose, 100 mM NaCl, 10 mM PIPES, pH 6.8, 3 mM MgCl₂, 2 mM vanadyl ribonucleoside, 1 mM EGTA, 1 mM PMSF, Complete protease inhibitor mixture (Roche Diagnostics)). The tissue was homogenized for 2 min (50 strokes) on ice using a precooled Potter-Elvehjem homogenizer and electrical drill. Next, the homogenate was filtered through three layers of 350-μm nylon mesh. The homogenate was left to mix on a turning wheel for 5 min (4 °C) and centrifuged for 5 min at 750 × g (4 °C), and the supernatant was collected (fraction 1, the “cytosolic” fraction containing lipids and soluble proteins). The pellet was resuspended in 600 μl of “extraction buffer”/100 mg of starting material (same as homogenization buffer but with 250 mM ammonium sulfate instead of 100 mM NaCl). The sample was mixed on a turning wheel for 5 min (4 °C) and centrifuged for 5 min at 750 × g (4 °C), and the supernatant was collected (fraction 2, the “cytoskeletal” fraction containing salt-extracted tissue elements and proteins). Third, the remaining pellet was resuspended in 400 μl of “digestion buffer”/100 mg of starting tissue (same as homogenization buffer but with 100 μg/ml RNase-free DNase I, and a reduced concentration of 50 mM NaCl). The sample was left at room temperature for 30 min for digestion of chromosomal DNA. The sample was brought back to high-salt conditions by addition of ammonium sulfate solution, left at room temperature for 5 min, centrifuged for 5 min at 1000 g (4 °C), and the supernatant was collected (fraction 3, the chromatin-binding fraction). The remaining detergent-, high salt-, and DNase-resistant pellet was resuspended in 400 μl of “disassembly buffer”/100 mg of starting material (8 mg urea, 20 mM PIPES, pH 6.8, 1 mM EGTA, 1 mM DTT, 1 mM PMSF), the suspended proteins were treated by overnight dialysis against >1,000 volumes of “reassembly buffer” (25 mM imidazole, pH 7.1, 150 mM KCl, 5 mM MgCl₂, 0.125 mM EGTA, 1.4 mM DTT, 0.26 mM PMSF). The repolymerized filaments were pelleted by centrifugation for 90 min at 217,000 × g for 90 min at 217,000 × g for 90 min at 217,000 × g for 90 min at 217,000 × g. The visible pellet was dissolved in Tris buffer (10 mM Tris-HCl, pH 7.0, 100 mM KCl), shaken vigorously, and incubated overnight at 4 °C after which the majority of protein was dissolved (fraction 4, the intermediate filament fraction). The 217,000 × g supernatant was collected as the final fraction (fraction 5, the nuclear matrix fraction). For the analysis of unfractionated CRC tissue, three CRC tissue samples (~75 mg each) were homogenized in a ratio of 100 mg/ml of buffer (7 mM urea, 2 mM thiourea, 4% CHAPS, and 10 μl/ml protease inhibitor mixture). Protein concentrations of the samples were determined with the Bradford assay (Bio-Rad). For proteomics analysis, 60 μg of CB, IF, and NM fractions and unfractionated CRC lysates were loaded on SDS-polyacrylamide gels.

**Gel Electrophoresis**

SDS-PAGE was performed with precast 4–12% gradient gels containing Bis-Tris buffer (NuPAGE MES system, Invitrogen). After electrophoresis, the gels were fixed in 50% ethanol containing 3% phosphoric acid and stained with Coomassie Blue R-250. After staining, the gels were washed in Milli-Q water and stored at 4 °C until processing for in-gel digestion.

**In-gel Digestion**

Coomassie-stained entire gel lanes were cut in 10 bands, and each band was processed for in-gel digestion according to the method of Shevchenko et al. (23). Briefly, bands were washed and dehydrated three times in 50 mM ammonium bicarbonate, pH 7.9 and in 50 mM ammonium bicarbonate, pH 7.9 + 50% ACN, respectively. Subsequently, cysteine bonds were reduced with 10 mM DTT for 1 h at 56 °C and alkylated with 50 mM iodoacetamide for 45 min at room temperature in the dark. After two subsequent wash/dehydration cycles, the bands were dried for 10 min in a vacuum centrifuge and incubated overnight with 0.06 μg/μl trypsin at 25 °C. Peptides were extracted once in 1% formic acid and subsequently two times in 50% ACN in 5% formic acid. The volume was reduced to 50 μl in a vacuum centrifuge prior to nano-LC-MS/MS analysis.

**Nano-LC-MS/MS and Database Searching**

**Nano-LC Separation**—Peptides were separated by an Ultimate 3000 nano-LC system (Dionex LC Packings, Amsterdam, The Netherlands) equipped with a 20 cm × 75-μm inner diameter fused silica column custom-packed with 3-μm 120-Å ReproSil Pur C18 aqua (Dr. Maisch GmbH, Ammerbuch-Entringen, Germany). After injection, peptides were trapped at 30 μl/min on a 5 μm × 300-μm inner diameter PepMap C18 cartridge (Dionex LC Packings) at 2% buffer B (buffer A, 0.05% formic acid in water; buffer B, 80% ACN + 0.05% formic acid in Milli-Q) and separated at 300 nL/min in a 10–40% buffer B gradient in 60 min.

**Injection Scheme Label-free Comparison**—Nano-LC-MS/MS data were acquired for the three subnuclear compartments (CB, IF, and NMP) in three batches with six colorectal tissues within 3 months and stable instrument performance. Per subnuclear fraction, the injection scheme of the 10 gel bands per sample was as follows: “adenoma1-band1,” “adenoma2-band1,” “MIN1-band1,” “MIN2-band1,” “CIN1-band1,” and “CIN2-band1” followed by the second gel band for all samples, the third, etc. until band 10. The injection scheme spreads instrumental drift over the biological samples and minimizes bias.

**Tandem Mass Spectrometry—**Eluting peptides were ionized at 1.7 kV in a Nanomate Triversa Chip-based nanospray source using a Triversa LC coupler (Advion, Ithaca, NY). Intact peptide mass spectra and fragmentation spectra were acquired on an LTQ-FT hybrid mass spectrometer (Thermo Fisher, Bremen, Germany). Intact masses were measured at resolution 50,000 in the ICR cell using a target value of 1 × 10⁶ charges. In parallel, following an FT prescan, the top five peptide signals (charge states 2+ and higher) were submitted to MS/MS in the linear ion trap (3-amu isolation width, 30-ms activation, 35% normalized activation energy, Q value of 0.25, and threshold of 5,000 counts). Dynamic exclusion was applied with a repeat count of 1 and an exclusion time of 30 s.

**Database Searching**—MS/MS spectra were searched against the human IPI database (6) (67,511 entries) using Sequest (version 27, revision 12), which is part of the BioWorks 3.3 data analysis package (Thermo Fisher, San Jose, CA). MS/MS spectra were searched with a maximum allowed deviation of 10 ppm for the precursor mass and 1 amu for fragment masses. Methionine oxidation and cysteine carbamidomethylation were allowed as variable modifications, two missed cleavages were allowed, and the minimum number of tryptic termini was 1. After database searching the DTA and OUT files were imported into Scaffold (versions 1.07 and 2.01) (Proteome Software, Portland, OR). Scaffold was used to organize the gel slice data and to validate peptide identifications using the PeptideProphet (24, 25) algorithm; only identifications with a probability >95% were retained. Subsequently, the ProteinProphet (24, 25) algorithm was applied to group peptide identifications to protein identifications. Protein identifications with a probability of >99% with two peptides or more in at least one of the samples were retained. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped.

Protein FDR was calculated in pilot experiments for three subnuclear fractions analyzed in a 10-gel band experiment each. For the CB fraction, 1,497 proteins were identified (two or more unique peptides and ProteinProphet score >99%) in the forward database search, and six were identified in the reversed database search, resulting in an FDR (reverse/reverse + forward) of 0.00399. Conversely, for the IF fraction, 893 were identified in the forward database.
search, and four were identified in the reversed database search, and for NMP fraction, 905 forward and six reversed database identifications were found, yielding an FDR of 0.00446 and 0.00659, respectively. The average FDR for these three fractions is 0.00501 or 0.50%.

For each protein identified, the number of spectral counts (the number of MS/MS spectra associated with an identified protein) was exported to Excel. Normalized spectral counts were calculated by dividing the spectral counts for an identified protein by the sum of the spectral counts per sample.

MALDI-TOF/TOF MS

Protein bands from Coomassie-stained SDS-polyacrylamide gels were excised manually using a clean scalpel or razor blade and digested as described above. The tryptic peptides extracted from gel slices were concentrated and desalted on reversed-phase C18 ZipTips (Millipore) and eluted with 3 µl of α-cyano-4-hydroxycinnamic acid matrix (6.2 mg/µl α-cyano-4-hydroxycinnamic acid in 56% (v/v) acetonitrile, 36% MeOH; Agilent Technologies). A sample volume of 0.8 µl from the peptide-matrix mixture was analyzed by MALDI-TOF/TOF MS (4800 MALDI-TOF/TOF, Applied Biosystems). The obtained mass spectra were searched against the human database using a MASCOT search engine (Matrix Science) with a mass tolerance set at 20 ppm for MS1 and 0.6 Da for MS/MS. The protonated trypsin autodigestion products at 842.510, 1,045,564, and 2,211.104 were used for internal calibration of the MALDI-TOF MS spectra. All the proteins listed were identified with a confidence interval of at least 95% from the combined MS and MS/MS data.

Statistical Analyses

Cluster Analysis—Cluster analysis was performed using hierarchical clustering in R. The protein abundances were normalized to zero mean and unit variance for each individual protein. Subsequently, the Euclidean distance measure was used for protein clustering. For sample clustering, the difference between two samples was defined as the sum of differences in individual proteins. Let the detection rates of a protein p in the two samples be $\lambda_1 > 0$ and $\lambda_2 > 0$. The difference in p between the two samples is as follows.

$$D_p = \lambda_1 \ln \lambda_2 + \lambda_2 \ln \lambda_1 \over \lambda_2$$

(Eq. 1)

This measure is the Jeffrey divergence, a symmetric variant of the Kullback-Leibler divergence, between two Poisson distributions with parameters $\lambda_1$ and $\lambda_2$. This measure has the advantage that it prevents highly abundant proteins from dominating others in contribution to the total sample difference.

Relative Enrichment of Proteins in Subnuclear CB, IF, and NM Fractions—Enrichment analysis of proteins in the NM fraction was performed by two pairwise comparisons with the CB fraction and the IF fraction. For each comparison, the Mantel-Haenszel method was used for significance analysis and estimation of a common fold change. A protein is enriched in the NM fraction when it is 2 times more abundant in the NM fraction than in both the CB and IF fractions with p values less than 0.05.

Differential Analysis of Nuclear Matrix Proteome in Colorectal Adenoma and Carcinoma Tissues—Differential analysis of colorectal adenoma, MIN+ CRC, and CIN+ CRC was performed using the β-binomial test. The β-binomial test takes into account the within-sample variation and the between-sample variation in a single statistical model (26). Three pairwise comparisons were performed. The threshold for statistical significance is 0.05. To select for the more robust changes, the regulated proteins with overlapping counts in each two-group comparison were removed. For the supervised cluster analysis, the results of the three comparisons were merged.

Gene Ontology and Pathway Analysis

Analysis of Enrichment of Nuclear Proteins—To assess the percentage of nuclear proteins in a list of differential proteins, their entries in the UniProt Knowledgebase were checked for any evidence that these proteins can be found in the nucleus. Entries at “General annotation (Comments) Subcellular location,” “Ontologies > Keywords > Cellular component,” and “Ontologies > Gene ontology (GO) > Cellular component” were all checked.

Relative Enrichment of GO Terms in Subnuclear Domains—IPI names of subnuclear fraction-enriched proteins were compared using the Functional profiling by functional enrichment methods (FatIGO+) tool version 3.2 on the Babelomics server (27). Enriched proteins in the CRC fractions and unfraccionated CRC tissue were compared for relative enrichment of GO terms in the GO categories “biological process,” “molecular function,” and “cellular component,” allowing for no redundant GO terms, and significant hits included only false discovery rate-adjusted p values <0.05. In addition, the FatIGO+ tool was used to search for enriched InterPro motifs in the amino acid sequences of the NM-enriched proteins.

Analysis of Enriched Protein Complexes—Gene symbol names of identified proteins that were significantly enriched in the CB, IF, and NM fractions and differential NM proteins of the adenoma-carcinoma comparison were analyzed with the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) tool version 8.0 (28). Protein-protein interaction (PPI) maps were created with default settings, allowing for experimentally verified and predicted PPIs; stronger associations are represented by thicker lines.

Analysis of Cancer-associated Proteins and Pathway Analysis—IPI names of identified proteins were searched with the Ingenuity Pathway analysis (IPA) tool version 7.0 for “Bio functions.” Putative cancer association was deduced from the Disorders reference list given in the Disorders and Mutations section of the pertinent protein entry in GeneCards. Manual literature mining was done with the Entrez Gene server.

RESULTS

Assessment of Reproducibility of Entire Work Flow

First, we investigated the reproducibility of the differential extraction protocol that yields the nuclear matrix fraction (13, 14) in conjunction with GeLC-MS/MS analysis and spectral counting. To this end, four CRC samples were pooled (total mass, 500 mg), the tissue was homogenized, and the homogenate was separated into three aliquots that were fractionated independently, resulting in three nuclear matrix fractions for analysis. First, 50, 60, or 70 µg of protein from one NM fraction was loaded onto an SDS-polyacrylamide gel. The maximum amount that could be loaded without smearing was 60 µg. Accordingly, 60 µg of each of the three replicates of NM fractions was separated by SDS-PAGE. The band pattern obtained in each NM fraction was very similar (Fig. 1A). For analysis of reproducibility of protein identification, the three lanes were cut into 10 pieces, aiming for an equal amount of total protein in each gel piece (Fig. 1A). The gel pieces were processed by in-gel digestion and subjected to nano-LC-MS/MS. The total number of identified proteins was 1,047 (with identification criteria of minimally two unique peptides at 95%
confidence at the peptide level and >99% confidence at the protein level in minimally one sample) in the total data set consisting of three replicate analyses. The average false discovery rate as revealed by reversed database searching was ~0.5% at the protein level, which is a low value. The number of proteins detected in all three replicates was 889 (85% of the total number of identified proteins), indicating good reproducibility (Fig. 1B). See supplemental Table 2 for documentation of protein identifications in the NM triplicate analysis and associated spectral count data.

To evaluate the combined technical variation including tissue fractionation, in-gel digestion, and nano-LC-MS/MS, we used protein spectral counting. The average CV of the number of spectral counts was 27% for the 889 proteins identified in all three replicates (supplemental Table 2). As expected, the CVs varied strongly between proteins with the proteins with the highest spectral counts producing the lowest CV% values (~10%; see supplemental Table 2).

Identification and Data Mining of NM-enriched Proteins

The proteomes of six colorectal tissues, i.e. two colorectal adenoma, two MIN+ CRC, and two CIN+ CRC, were fractionated each into five subproteomes by biochemical fractionation. Fig. 2A shows an example of the protein band patterns in the five fractions of a CRC tissue as analyzed by SDS-PAGE. The majority of the protein mass (~90%) was collected in the two first fractions using extraction by detergent and high salt removing membrane/soluble and cytoskeletal proteins, respectively. By subsequent treatment with DNase I, ~5% of the total protein mass was collected in the third
fraction, the so-called chromatin-binding fraction. The remaining pellet was dissolved in urea-containing buffer, the urea was dialyzed off, allowing intermediate filament proteins to reassemble into insoluble networks, and the sample was centrifuged at high speed in an ultracentrifuge. The final pellet (fraction 4) and supernatant (fraction 5) each contained ~1% of the total protein mass. The SDS-PAGE profiles of fractions 1 and 2 were very similar, whereas fractions 3–5 differed substantially (see Fig. 2). The profile of the third fraction was characterized by two strong bands of ~16 kDa, identified by MALDI-TOF/TOF MS as histones (H1d and H1a). The fourth fraction was characterized by a cluster of strong bands of 40–50 kDa identified as keratin (cytokeratin 8), actin (cytoplasmic 2), and tubulin (β2 and α3). The fifth fraction was characterized by two protein bands of ~70 kDa that were identified as lamins B1 and A/C, respectively (Fig. 2). In the following, fractions 3, 4, and 5 are referred to as the CB, the IF, and the NM fractions, respectively.

To identify proteins specifically enriched in the NM fraction of the six SDS-PAGE-separated colorectal tissue samples, the entire gel lanes of the CB, IF, and NM fractions were sliced in 10 pieces, and the proteins were subjected to in-gel digestion and analysis by nano-LC-MS/MS. A total of 2,059 proteins was identified in the whole data set with 753 detected in all three fractions of the six tissue samples (Fig. 2B). To obtain a global view of the data, we performed an unsupervised cluster analysis. The cluster analysis yielded three main clusters that corresponded to the three subnuclear fractions. Each fraction cluster showed further subgrouping, largely according to biological tissue types (see Fig. 3 for a heat map view of the cluster analysis).

To distinguish candidate core components of the NM from contaminants, we performed a statistical analysis to identify proteins that were significantly enriched in the NM fraction relative to the protein levels in the CB and IF fractions. The fraction enrichment analysis yielded 158 proteins enriched in the NM fraction of a total of 1,153 proteins identified in the NM fraction (criteria: p value <0.05 with minimum 2-fold change). Besides statistical significance, we added the extra filter that the numerical enrichment is consistently observed in at least three fractions. This allowed us to reduce the number of contaminants while retaining the majority of known NM proteins. See supplemental Table 3 for documentation of protein identifications in the unfractionated CRC tissue lysate. The different levels for each protein were reproducibly identified in unfractionated tissue homogenates of three CRC tissue samples (see supplemental Table 4 for documentation of protein identifications in the unfractionated CRC tissues). To this end, we used the GO mining tool FatiGO+ (27). Interestingly, the NM fraction was uniquely enriched for several GO terms associated with specific nuclear domains and functions. Table I lists the distribution of the significant (adjusted p < 0.05) biological process, molecular function, and cellular component GO terms that were associated with the proteins enriched in the NM fraction relative to unfractionated CRC tissue lysate. The different levels for each ontology are shown in Table I with the significantly different GO terms listed in order of decreasing significance. For example, the NM-enriched proteins were significantly associated with the cellular component term “nucleus” and with the biological func-
Nuclear Matrix Proteomics of Colorectal Cancer

**TABLE I**

Significant gene ontology terms associated with proteins enriched in nuclear matrix fraction

| Level | Biological process | Molecular function | Cellular component |
|-------|-------------------|--------------------|-------------------|
| 3     | Cell adhesion     | Protein binding    | Extracellular region part |
|       | Multicellular organismal development | Extracellular matrix structural constituent | Extracellular matrix part |
|       | Cellular component organization and biogenesis | Nucleic acid binding | |
|       | Anatomical structure development | | |
| 4     | RNA metabolic process | Cytoskeletal protein binding | Proteinaceous extracellular matrix |
|       | | RNA binding | |
| 5     | RNA processing | Actin binding | |
|       | mRNA metabolic process | Zinc ion binding | |
| 6     | mRNA processing | Vascular endothelial growth factor receptor activity | |
| 7     | RNA splicing | | |
| 8     | | | Nucleus |

Significant GO term association was assessed using Ingenuity Pathway Analysis software. GO terms are given in regular font when associated with an adjusted p value of \( p < 0.05 \); terms given in italic are associated with an adjusted p value of \( p < 0.1 \).

In view of a potential important role of the nuclear matrix in colorectal adenoma to carcinoma progression, we further analyzed the nuclear matrix proteome obtained from the two adenoma and four carcinoma tissues to identify cancer-associated proteins using \( \beta \)-binomial statistics (26). The results for the IF and CB fractions will be reported elsewhere. Included in the analysis of carcinoma tissues were the two major modes for genomic instability, CIN (85% of cases) and MIN (15% of cases), to further pinpoint candidate CIN-associated proteins.

An overview of the identification data and statistical analyses is presented in Table II, and the NM protein fractionation and a summary of the protein identification data are shown in supplemental Fig. 1. The total data set contained 1,153 proteins with about half of the proteins identified in four of six clinical samples (supplemental Fig. 1). We focus here on the comparisons adenoma versus all CRC tissues (Table III) and adenoma versus CIN+ CRC (Table IV). Tables III and IV display the lists of nuclear annotated and/or NM-enriched proteins as well as the differential NM proteins with corresponding mRNA change in the comparisons of adenoma versus all CRC tissues (Table III) and adenoma versus CIN+ CRC (Table IV).

Supplemental Tables 5 and 6 contain the complete lists and associated spectral count data of the significantly regulated proteins. Statistical comparisons revealed that, of a total of 1,153 identified NM proteins, 129 proteins were significantly regulated (28 up- and 101 down-regulated; \( p < 0.05 \)) in the comparison adenoma versus all CRC tissues (Table III and supplemental Table 5), 169 proteins were significantly regulated (87 up- and 82 down-regulated) in the comparison adenoma versus CIN+ CRC (Table IV and supplemental Table 6), and 237 proteins were significantly regulated (158 up- and 79 down-regulated) in the comparison adenoma versus MIN+ CRC (supplemental Table 7). Fig. 5 shows a heat map view of the supervised cluster analysis using the differential proteins from the above mentioned comparisons. The tree contains two main clusters consisting of one cluster with the two MIN tumors, and the other main cluster contains in one subcluster the two adenomas and one of the CIN tumors and the other CIN tumor in a separate branch.
Interestingly, the CIN tumor in the adenoma cluster that exhibited an “intermediate” expression profile (Fig. 5 and supplemental Fig. 1A) was a small lymph node-negative tumor (p8 in supplemental Table 1).

The majority of the significantly regulated NM proteins exhibited a -fold change >2 (Tables III and IV and supplemental Tables 5–7). About half of the proteins are of nuclear origin according to UniProt annotations, and about half of the proteins (57 of 129 and 81 of 169) have been previously associated with cancer according to literature references in the corresponding entries in GeneCards (see annotations in Tables III and IV and supplemental Tables 5–7). Integration with our previously generated transcriptomics data that are based on a large series of clinical samples (31 adenomas and 36 carcinomas; sets available at Gene Expression Omnibus under accession number GSE8067) reveals multiple candidates (26 of 129 and 37 of 169) that exhibited a corresponding change at the mRNA level (p < 0.1) (Tables III and IV and supplemental Tables 5–7). Differential proteins that exhibit a corresponding mRNA change include for example versican V1 (VCAN), myeloid cell nuclear differentiation antigen (MNDA), high mobility group protein B3 (HMGB3), proteins S100-A8 and -A9, isoform 1 of PDLIM4, periplakin (PPL), envoplakin (EVPL), ITGB4, nuclear pore complex protein 88 (NUP88), and isoform 1 of RNA-binding protein 14 (TMEM137/RBM14).

To explore the connectivity of the nuclear annotated NM proteins and/or NM-enriched proteins that were differential in the comparisons adenoma versus carcinoma (n = 6 samples) and adenoma versus CIN+ carcinoma (n = 4 samples), we used the STRING tool. The PPI map of the 68 regulated NM-enriched nuclear proteins from the adenoma versus all CRC comparison is shown in Fig. 6 and contains one large cluster of highly connected down-regulated proteins (cluster A) and two smaller, well connected clusters of which one cluster exclusively contains down-regulated proteins (cluster B). The dense cluster of down-regulated proteins (Fig. 6A) corresponds to the nuclear pore complex and contains

![Protein-protein interaction map of proteins enriched in nuclear matrix fraction. Squares A–C indicate functional clusters (see main text for description). Panels A, B, and C are magnifications of regions of interest in the main figure.](image-url)
Nuclear Matrix Proteomics of Colorectal Cancer

### DISCUSSION

Nuclear phenotypic alterations and chromosomal instability are associated with progression toward malignancy. Scaffolding proteins may play an important underlying role in this process. In the past, a specific differential biochemical extraction procedure has been devised to investigate such a nuclear scaffolding network, or matrix (13, 14). We are the first to provide a deep, MS/MS identification-based view of the NM fraction, the final product of this procedure. Importantly, through differential comparison against two earlier fractions that contain nuclear proteins, i.e. the chromatin-binding and the intermediate filament fractions, we are for the first time able to pinpoint NM-enriched proteins that may represent core constituents. Moreover, we show that a combination of this differential extraction protocol together with GeLC-MS/MS analysis and spectral counting as a means of quantification is reproducible (average CV% of <30%), allowing us in a proof-of-concept study to identify NM proteins associated with adenoma to chromosomal instable carcinoma progression. We propose that sequential cellular extraction of proteins in combination with nano-LC-MS/MS and protein-protein interaction network analysis

---

**Table II**

Summary of data on differential comparisons, protein identifications, and annotations

| Analysis summary | All | IPA mapped | IPA cancer |
|------------------|-----|------------|------------|
| Total number of identified proteins | 888 | 847 | 245 (29%) |
| Regulated protein 2-group comparisons | | | |
| Adenoma (n = 2) vs. carcinoma (n = 4) | p < 0.05 | >-2-fold | Unique | IPA cancer | UniProt nuclear |
| Up in adenoma | 129 | 122 | 35 | 58 (48%) | 50 (39%) |
| Up in carcinoma | 101 | 95 | 26 | 47 (49%) | 39 (39%) |
| Adenoma (n = 2) vs. 20q CIN (n = 2) | 28 | 27 | 9 | 14 (52%) | 11 (39%) |
| Up in adenoma | 169 | 157 | 57 | 68 (43%) | 73 (43%) |
| Up in carcinoma | 87 | 77 | 31 | 29 (38%) | 36 (41%) |
| Adenoma (n = 2) vs. MIN (n = 2) | 82 | 80 | 26 | 39 (49%) | 37 (45%) |
| Up in adenoma | 237 | 225 | 100 | 98 (44%) | 93 (39%) |
| Up in MIN carcinoma | 158 | 148 | 74 | 61 (41%) | 66 (42%) |
| Up in MIN carcinoma | 79 | 77 | 26 | 40 (52%) | 27 (34%) |

11 nucleoporins and the nuclear pore components nucleoprotein TPR and FIG1 (for details see Table III). The cluster that is directly connected to the nuclear pore protein cluster (Fig. 6C) consists of up-regulated RNA-binding proteins involved in translational regulation (heterogeneous nuclear ribonucleoprotein U-like protein 1 (hnRNPU1) and DHX9). RNA processing (SYNCRIP), and RNA trafficking (KSRP) as well as down-regulated RNA processing and splicing factors (DDX, BEF, splicing factor arginine/serine-rich 6 (SFRS6), and splicing factor U2AF 65-kDa subunit (U2AF2)). The third cluster at the opposing end of the network (Fig. 6B) consists of down-regulated (structural) proteins, including catenins β-1, α-1, and δ-1 (CTNB1, CTNB1A1, and p120); tight junction protein ZO-2 (TJP2); vinculin (VCL); and LIM domain and actin-binding protein 1 (LIM1). In addition, a number of proteins are poorly or unconnected (e.g. VCAN, actin-binding LIM protein 1 (ABLM1), and neuregulin 1 (NDF); see Fig. 6).

The PPI map of the 96 regulated NM-enriched nuclear proteins from the adenoma-chromosomal instable carcinoma comparison is shown in Fig. 7 and contains two highly connected clusters; one symmetric cluster consists of exclusively down-regulated proteins of the nuclear pore complex (NUPL1, NUP98, NUP133, NUP85, NUP205, NUP98, NUP214, FIG1, and TPR), and the other cluster consisting of four up- and six down-regulated proteins involved in RNA processing and splicing. The up-regulated proteins include hnRNPU1, Y-box-binding protein 1 (YBX1), the transcriptional regulator hnRNPD, and the transcriptional activator ATP-dependent RNA helicase A (DHX9) (Fig. 7). The down-regulated proteins in the second cluster include poly(rC)-binding protein 2 (PCBP2/hnRNPE2), hnRNPH2, protein mago nashi homolog (MAGOH), SFRS6, stalk cell-derived growth factor SF20 (SF2), and U2AF2 (Fig. 7). The nuclear pore and RNA processing clusters are connected through hnRNPU1 that is directly connected to NUP214 and NUP98 and through YBX1 that is indirectly linked to the nuclear pore cluster through its connection to (up-regulated) nucleolin (NCL) that in turn is connected to NUP98 (Fig. 7). Other up-regulated proteins in close vicinity of the RNA processing-splicing cluster include the up-regulated far upstream element-binding protein 2 (KSRP) and heterogeneous nuclear ribonucleoprotein Q (SYNCRIP) that are directly connected through hnRNPD. The up-regulated RNA-binding protein DAZAP1 is connected to YBX1, and peptidyl-prolyl cis-trans isomerase NIMA-interacting 4 (PIN4) is indirectly connected through KSRP (Fig. 7). In addition, several up-regulated histones (H2AFY, HIST1H1D, and HIST1H1C) are connected to nucleolin (Fig. 7) as well as MNDA and nucleolar protein NOP5 (NOL5). More diffuse areas of the network include mainly structural proteins such as PDLIM4, VCL, LAMC1, ITGB4, galectin-3-binding protein (90K), and galectin-3 (MAC-2).
| Differential Proteins Between Adenomas versus Colorectal Carcinomas (CRC) | Accession No. | Gene Symbol | Cancer associated | Subcellular Localization | RNA ADE-CR p-value | Fold Change ADE-CR | p-Value ADE-CR | X (Protein Enriched in XR MM tissues) |
|---|---|---|---|---|---|---|---|---|
| NPC1L1 | P02241 | NPC1L1 | NUC, CYT | NUC, CYT | 1.0 | 1.0 | 1.0 | 1.0 |
| NPC2 | P02242 | NPC2 | NUC, CYT | NUC, CYT | 1.0 | 1.0 | 1.0 | 1.0 |

### Table II

| Regulated NM proteins that are nuclear annotated and/or NM enriched and/or with corresponding mRNA change in the comparison of adenoma versus all CRC tissues |
|---|
| **HPBP3** HPBP3/Pi protein Interactor 1 binding protein 3 |
| **HNRNPU** HNRNPU isoform 1 of heterogeneous nuclear ribonucleoprotein U-like protein 1 |
| **SF3B1** SF3B1 Antibody - variant |
| **TIF1B** TIF1B Interferon Regulatory Factor 1 Homolog |
| **KHSRP** KHSRP kHSRP splicing regulatory protein1 | 
| **NUP98** NUP98 Nuclear protein 5 |
| **DIKH** D1KH ATP-dependent RNA helicase A |
| **CAML** CAML isoform 1 of Calmodulin |
| **HKG2** HKG2 calmodulin |
| **MYC** MYC |
| **RBM9** RBM9 |
| **SFRS3** SFRS3 |
| **SNAP25** SNAP25 |
| **NUP98** NUP98 |
| **VCP** VCP |
| **ERBB2** ERBB2 |
| **CDK5RAP3** CDK5RAP3 |
| **TP53** TP53 |
| **PRKACG** PRKACG |
| **DAXX** DAXX |
| **CCNB1** CCNB1 Isoform 1 of Cyclin B1 |
| **PIM1** PIM1 |
| **NR4A3** NR4A3 |
| **MALT** MALT |
| **NUP98** NUP98 |
| **CPSF6** CPSF6 Isoform 1 of Cleavage and polyadenylation specificity factor 6 |
| **VHL** VHL |
| **SCF** SCF |
| **UBE2D1** UBE2D1 |
| **IL10** IL10 |
| **MAF** MAF |
| **NUP98** NUP98 |

### Table III

| Nuclear Matrix Proteomics of Colorectal Cancer |
|---|
| **NPC1L1** NPC1L1 isoform 2 of Aloe sp. protein 3 |
| **NPC2** NPC2 isoform 2 of Aloe sp. protein 3 |
| **NPC3** NPC3 isoform 2 of Aloe sp. protein 3 |
| **NPC4** NPC4 isoform 2 of Aloe sp. protein 3 |
| **NPC5** NPC5 isoform 2 of Aloe sp. protein 3 |
| **NPC6** NPC6 isoform 2 of Aloe sp. protein 3 |

### Table IV

| Nuclear Matrix Proteins with Concurrent mRNA Changes |
|---|
| **COL1A1** COL1A1 isoform 4 of Collagen alpha1(I) chain precursor |
| **COL1A2** COL1A2 isoform 4 of Collagen alpha2(I) chain precursor |
| **COL1A3** COL1A3 isoform 4 of Collagen alpha3(I) chain precursor |
| **COL3A1** COL3A1 isoform 4 of Collagen alpha3(I) chain precursor |
| **COL4A2** COL4A2 isoform 4 of Collagen alpha2(IV) chain precursor |
| **COL5A1** COL5A1 isoform 4 of Collagen alpha1(V) chain precursor |
| **COL6A1** COL6A1 isoform 4 of Collagen alpha1(IV) chain precursor |
| **COL7A1** COL7A1 isoform 4 of Collagen alpha1(IV) chain precursor |
| **COL8A1** COL8A1 isoform 4 of Collagen alpha1(IV) chain precursor |
| **COL9A1** COL9A1 isoform 4 of Collagen alpha1(IV) chain precursor |
| **COL10A1** COL10A1 isoform 4 of Collagen alpha1(IX) chain precursor |
| **COL11A1** COL11A1 isoform 4 of Collagen alpha1(X) chain precursor |
| **COL12A1** COL12A1 isoform 4 of Collagen alpha1(XII) chain precursor |
| **COL13A1** COL13A1 isoform 4 of Collagen alpha1(XIII) chain precursor |
| **COL14A1** COL14A1 isoform 4 of Collagen alpha1(XIV) chain precursor |
| **COL15A1** COL15A1 isoform 4 of Collagen alpha1(XV) chain precursor |
| **COL16A1** COL16A1 isoform 4 of Collagen alpha1(XVI) chain precursor |
| **COL17A1** COL17A1 isoform 4 of Collagen alpha1(XVII) chain precursor |
| **COL18A1** COL18A1 isoform 4 of Collagen alpha1(XVIII) chain precursor |
| **COL19A1** COL19A1 isoform 4 of Collagen alpha1(XIX) chain precursor |
| **COL20A1** COL20A1 isoform 4 of Collagen alpha1(XX) chain precursor |
| **COL21A1** COL21A1 isoform 4 of Collagen alpha1(XXI) chain precursor |
| **COL22A1** COL22A1 isoform 4 of Collagen alpha1(XXII) chain precursor |
| **COL23A1** COL23A1 isoform 4 of Collagen alpha1(XXIII) chain precursor |
| **COL24A1** COL24A1 isoform 4 of Collagen alpha1(XXIV) chain precursor |
| **COL25A1** COL25A1 isoform 4 of Collagen alpha1(XXV) chain precursor |
| **COL26A1** COL26A1 isoform 4 of Collagen alpha1(XXVI) chain precursor |
| **COL27A1** COL27A1 isoform 4 of Collagen alpha1(XXVII) chain precursor |
| **COL28A1** COL28A1 isoform 4 of Collagen alpha1(XXVIII) chain precursor |
| **COL29A1** COL29A1 isoform 4 of Collagen alpha1(XXIX) chain precursor |
| **COL30A1** COL30A1 isoform 4 of Collagen alpha1(XXX) chain precursor |
| **COL31A1** COL31A1 isoform 4 of Collagen alpha1(XXXI) chain precursor |
| **COL32A1** COL32A1 isoform 4 of Collagen alpha1(XXI) chain precursor |
| **COL33A1** COL33A1 isoform 4 of Collagen alpha1(XXII) chain precursor |
| **COL34A1** COL34A1 isoform 4 of Collagen alpha1(XXIII) chain precursor |
| **COL35A1** COL35A1 isoform 4 of Collagen alpha1(XXIV) chain precursor |
| **COL36A1** COL36A1 isoform 4 of Collagen alpha1(XXV) chain precursor |
| **COL37A1** COL37A1 isoform 4 of Collagen alpha1(XXVI) chain precursor |
| **COL38A1** COL38A1 isoform 4 of Collagen alpha1(XXVII) chain precursor |
| **COL39A1** COL39A1 isoform 4 of Collagen alpha1(XXVIII) chain precursor |
| **COL40A1** COL40A1 isoform 4 of Collagen alpha1(XXIX) chain precursor |
| **COL41A1** COL41A1 isoform 4 of Collagen alpha1(XXX) chain precursor |
| **COL42A1** COL42A1 isoform 4 of Collagen alpha1(XXXI) chain precursor |
| **COL43A1** COL43A1 isoform 4 of Collagen alpha1(XXXII) chain precursor |
| **COL44A1** COL44A1 isoform 4 of Collagen alpha1(XXXIII) chain precursor |
| **COL45A1** COL45A1 isoform 4 of Collagen alpha1(XXXIV) chain precursor |
| **COL46A1** COL46A1 isoform 4 of Collagen alpha1(XXXV) chain precursor |
| **COL47A1** COL47A1 isoform 4 of Collagen alpha1(XXXVI) chain precursor |
| **COL48A1** COL48A1 isoform 4 of Collagen alpha1(XXXVII) chain precursor |
| **COL49A1** COL49A1 isoform 4 of Collagen alpha1(XXXVIII) chain precursor |
| **COL50A1** COL50A1 isoform 4 of Collagen alpha1(XXXIX) chain precursor |
| **COL51A1** COL51A1 isoform 4 of Collagen alpha1(XXX) chain precursor |
| **COL52A1** COL52A1 isoform 4 of Collagen alpha1(XXXI) chain precursor |
| **COL53A1** COL53A1 isoform 4 of Collagen alpha1(XXXII) chain precursor |
| **COL54A1** COL54A1 isoform 4 of Collagen alpha1(XXXIII) chain precursor |
| **COL55A1** COL55A1 isoform 4 of Collagen alpha1(XXXIV) chain precursor |
| **COL56A1** COL56A1 isoform 4 of Collagen alpha1(XXXV) chain precursor |
| **COL57A1** COL57A1 isoform 4 of Collagen alpha1(XXXVI) chain precursor |
| **COL58A1** COL58A1 isoform 4 of Collagen alpha1(XXXVII) chain precursor |
| **COL59A1** COL59A1 isoform 4 of Collagen alpha1(XXXVIII) chain precursor |
| **COL60A1** COL60A1 isoform 4 of Collagen alpha1(XXXIX) chain precursor |
| **COL61A1** COL61A1 isoform 4 of Collagen alpha1(XXX) chain precursor |
| **COL62A1** COL62A1 isoform 4 of Collagen alpha1(XXXI) chain precursor |
| **COL63A1** COL63A1 isoform 4 of Collagen alpha1(XXXII) chain precursor |
| **COL64A1** COL64A1 isoform 4 of Collagen alpha1(XXXIII) chain precursor |
| **COL65A1** COL65A1 isoform 4 of Collagen alpha1(XXXIV) chain precursor |
| **COL66A1** COL66A1 isoform 4 of Collagen alpha1(XXXV) chain precursor |
| **COL67A1** COL67A1 isoform 4 of Collagen alpha1(XXXVI) chain precursor |
| **COL68A1** COL68A1 isoform 4 of Collagen alpha1(XXXVII) chain precursor |
TABLE IV
Regulated NM proteins that are nuclear annotated and/or NM enriched and/or with corresponding mRNA change in the comparison of adenoma versus CIN + CRC

| Differential | Gene Symbol | Nuclear Annotation | mRNA Change | p-Value | X (Protein Enriched as XN% Fraction) |
|--------------|-------------|--------------------|-------------|---------|-------------------------------------|
| Proteins     |             |                    |             |         |                                     |
| P140         |             |                    |             |         |                                     |
| P150         |             |                    |             |         |                                     |
| P160         |             |                    |             |         |                                     |
| P170         |             |                    |             |         |                                     |
| P200         |             |                    |             |         |                                     |
| P210         |             |                    |             |         |                                     |
| P220         |             |                    |             |         |                                     |
| P230         |             |                    |             |         |                                     |
| P240         |             |                    |             |         |                                     |
| P250         |             |                    |             |         |                                     |
| P260         |             |                    |             |         |                                     |
| P270         |             |                    |             |         |                                     |
| P280         |             |                    |             |         |                                     |
| P290         |             |                    |             |         |                                     |
| P300         |             |                    |             |         |                                     |
| P310         |             |                    |             |         |                                     |
| P320         |             |                    |             |         |                                     |
| P330         |             |                    |             |         |                                     |
| P340         |             |                    |             |         |                                     |
| P350         |             |                    |             |         |                                     |
| P360         |             |                    |             |         |                                     |
| P370         |             |                    |             |         |                                     |
| P380         |             |                    |             |         |                                     |
| P390         |             |                    |             |         |                                     |
| P400         |             |                    |             |         |                                     |
| P410         |             |                    |             |         |                                     |
| P420         |             |                    |             |         |                                     |
| P430         |             |                    |             |         |                                     |
| P440         |             |                    |             |         |                                     |
| P450         |             |                    |             |         |                                     |
| P460         |             |                    |             |         |                                     |
| P470         |             |                    |             |         |                                     |
| P480         |             |                    |             |         |                                     |
| P490         |             |                    |             |         |                                     |
| P500         |             |                    |             |         |                                     |
| P510         |             |                    |             |         |                                     |
| P520         |             |                    |             |         |                                     |
| P530         |             |                    |             |         |                                     |
| P540         |             |                    |             |         |                                     |
| P550         |             |                    |             |         |                                     |
| P560         |             |                    |             |         |                                     |
| P570         |             |                    |             |         |                                     |
| P580         |             |                    |             |         |                                     |
| P590         |             |                    |             |         |                                     |
| P600         |             |                    |             |         |                                     |
| P610         |             |                    |             |         |                                     |
| P620         |             |                    |             |         |                                     |
| P630         |             |                    |             |         |                                     |
| P640         |             |                    |             |         |                                     |
| P650         |             |                    |             |         |                                     |
| P660         |             |                    |             |         |                                     |
| P670         |             |                    |             |         |                                     |
| P680         |             |                    |             |         |                                     |
| P690         |             |                    |             |         |                                     |
| P700         |             |                    |             |         |                                     |
| P710         |             |                    |             |         |                                     |
| P720         |             |                    |             |         |                                     |
| P730         |             |                    |             |         |                                     |
| P740         |             |                    |             |         |                                     |
| P750         |             |                    |             |         |                                     |
| P760         |             |                    |             |         |                                     |
| P770         |             |                    |             |         |                                     |
| P780         |             |                    |             |         |                                     |
| P790         |             |                    |             |         |                                     |
| P800         |             |                    |             |         |                                     |
| P810         |             |                    |             |         |                                     |
| P820         |             |                    |             |         |                                     |
| P830         |             |                    |             |         |                                     |
| P840         |             |                    |             |         |                                     |
| P850         |             |                    |             |         |                                     |
| P860         |             |                    |             |         |                                     |
| P870         |             |                    |             |         |                                     |
| P880         |             |                    |             |         |                                     |
| P890         |             |                    |             |         |                                     |
| P900         |             |                    |             |         |                                     |
| P910         |             |                    |             |         |                                     |
| P920         |             |                    |             |         |                                     |
| P930         |             |                    |             |         |                                     |
| P940         |             |                    |             |         |                                     |
| P950         |             |                    |             |         |                                     |
| P960         |             |                    |             |         |                                     |
| P970         |             |                    |             |         |                                     |
| P980         |             |                    |             |         |                                     |
| P990         |             |                    |             |         |                                     |

**Note:** This table includes a detailed list of regulated NM proteins, their nuclear annotations, mRNA changes, p-values, and protein enrichments. The data are presented in a tabular format for easy reference and analysis.
and such a protein yield is comparable with another study relatively small tissue samples (g/H11011 size that the work flow presented here allows for profiling of and scaffold filaments) fractions are harvested. We emphasized the importance of chromatin structure and function. As a first step toward identifying the NM core components, we have generated a “snapshot” of more than the extractions developed by Albrethsen et al., offers a “snapshot” of more than the subnuclear compartment enrichment analysis based on six biological samples, we provide, to the best of our knowledge, the most comprehensive view of the NM fraction so far.

Differential Extraction of Nuclear Proteins and Identification of NM-enriched Proteins

Previous studies have used differential extraction protocols that are most often based on the procedure elaborated by Penman and co-workers (13, 14). For analysis of the NM fraction in diverse cells and tissues, most studies have used 2DGE (17), but these studies have yielded only limited protein identification information to date. To provide a more comprehensive view and to overcome some of the limitations with solubility inherent to 2DGE, we used an alternative method consisting of GeLC-MS/MS on an LTQ-FT MS platform. In brief, the majority of the protein mass (>90%) isolated from tissue was first removed by sequential treatment with detergent, high salt buffers, and DNase treatment after which the so-called CB (proteins bound to/associated with chromatin), IF (a mixture of cytoskeletal scaffold filaments), and NM (the non-intermediate filament part of an insoluble nuclear scaffold and scaffold filaments) fractions are harvested. We emphasize the high performance of our work flow was shown in our reproducibility study where we identified 889 reproducibly detected proteins (CV% < 30%) of a total of 1,047 proteins in the NM fraction from CRC tissues. In the CB, IF, and NM fractions derived from six clinical samples, we identified a total of 2,059 proteins, about half of them with nuclear annotation, a clear enrichment as compared with unfractionated CRC tissue. Combined with our subnuclear compartment enrichment analysis based on six biological samples, we provide, to the best of our knowledge, the most comprehensive view of the NM fraction so far.

The biochemically defined NM, a detergent-, salt-, and DNase-resistant nuclear remnant after a series of biochemical extractions developed by Albrethsen et al. Berezney, Coffey and Penman (13, 14), offers a “snapshot” of more than the “NM proper” per se: it includes many factors that are (also) involved in the generation, processing, and export of ribosome subunits and RNA molecules. Many of these are (also) nucleolar constituents. We are interested in the NM proper as a fibrogranular scaffold and organizing entity involved in (regulation of) chromatin structure and function. As a first step toward identifying the NM core components, we have generated a list of proteins enriched in the NM fraction relative to where a total of 300–500 μg of NM protein was extracted from ~1 g of bladder tissue (29). The high performance of our work flow was shown in our reproducibility study where we identified 889 reproducibly detected proteins (CV% < 30%) of a total of 1,047 proteins in the NM fraction from CRC tissues. In the CB, IF, and NM fractions derived from six clinical samples, we identified a total of 2,059 proteins, about half of them with nuclear annotation, a clear enrichment as compared with unfractionated CRC tissue. Combined with our subnuclear compartment enrichment analysis based on six biological samples, we provide, to the best of our knowledge, the most comprehensive view of the NM fraction so far.

The biochemically defined NM, a detergent-, salt-, and DNase-resistant nuclear remnant after a series of biochemical extractions developed by Albrethsen et al. Berezney, Coffey and Penman (13, 14), offers a “snapshot” of more than the “NM proper” per se: it includes many factors that are (also) involved in the generation, processing, and export of ribosome subunits and RNA molecules. Many of these are (also) nucleolar constituents. We are interested in the NM proper as a fibrogranular scaffold and organizing entity involved in (regulation of) chromatin structure and function. As a first step toward identifying the NM core components, we have generated a list of proteins enriched in the NM fraction relative to
the CB and IF fractions. Many efforts in the past have aimed to identify a set of common or “core” NM proteins as well as tissue- and/or cancer-specific NM proteins; so far, ~300 NM mammalian proteins have been collected in an electronic database (30). However, most studies only describe a limited number of proteins, and core NM components cannot be deduced (7). In the NM fraction of six colorectal tissues, we identified out of a total of 1,153 identified proteins 158 significantly enriched proteins as compared with the CB and IF fractions. By GO mining, we found that these proteins are significantly associated with “mRNA processing” and “RNA metabolic process.” In line with these findings, particularly prominent in the molecular interaction map of proteins enriched in the NM was a dense cluster of proteins involved in pre-mRNA processing and, directly connected to it, multiple proteins involved in RNA splicing. How these functions are linked to the core fibrogranular NM scaffold remains to be investigated.

Among the proteins commonly detected by 2DGE in the NM fractions from diverse tissues and cell lines are several hnRNP family members as well as NUMA and laminas as “hallmark” constituents of the NM. Indeed in our analysis, many members of the hnRNP protein family were truly enriched in the NM fraction as compared with the CB and IF fractions (hnRNPA0, -A1, -A2B1, -A3, -AB, -D, -DL, -F, -H1, -H2, -H3, -L, and -R). It has been concluded from a large variety of experiments that hnRNP family members (such as hnRNPA1 and -B1) are responsible for a minor proportion of the protein content of the NM filaments (15). The nuclear laminas are members of the intermediate filament protein family and, as such, part of the tissue skeleton and therefore obvious candidates as main structural proteins of the NM. Lamins are primarily found at the inner side of the nuclear envelope and are probably not the primary constituents of the interior NM filaments (15); nonetheless, laminas have also been identified in the interior nucleus (31). Here, we have identified lamin A/C as an NM-enriched protein, whereas lamins B1 and B2 are enriched in the IF fraction, supporting several studies that find that specifically lamin A/C play important nuclear roles (60).

Association of the NM with RNA-binding proteins and RNA metabolism is in line with the notion that the NM is an RNA-protein-based network. Besides “RNA functions,” the NM-enriched proteins were also associated significantly with cytoskeletal protein binding, cell adhesion, and “extracellular matrix structural constituent” functions. Although one would not anticipate extracellular matrix, cell adhesion, and cytoskeletal proteins as NM constituents a priori, a nuclear role for at least some of these proteins cannot be ruled out. Indeed, several studies now support the concept of an interconnected protein skeleton in tissue, extending from the innermost structures of the nucleus through the cytoskeleton to the extracellular matrix, and in this way the entire tissue skeleton could be involved in gene regulation (17, 32, 33). It goes without saying that further work is needed to distinguish contaminants from shuttling proteins and proteins with a (structural) function in the nucleus. As to one promising candidate of this group, there is now sufficient evidence that actin is implicated in several nuclear activities together with a growing number of actin-binding proteins (34). Actin, myosin, cyto-keratins, and spectrin have all been identified in NM filaments by immunodetection (35). Indeed, in this study, many actin-binding proteins were enriched in the NM fraction (e.g. PDLIM1 and PDLIM4), and “actin binding” was a significant molecular function term associated with NM-enriched proteins. As such, our results support that in particular actin is a likely candidate as a major NM protein.

Other novel candidate NM proteins enriched in the NM preparations include the three isoforms CHMP1A, CHMP2B, and CHMP4B. Especially the chromatin modifying properties of at least one member of this family (CHMP1), found to be a nuclear matrix protein, are intriguing (36). Other proteins highly enriched in the NM include PSPC1, the RNA-binding protein ELAV-like protein (ELAVL1/HUR), transcriptional activator protein PURA, isoform long of TAF15, and isoform 1 of FUBP1 (up in CIN-CRC; see below) and FUBP3. Interestingly, the cancer stem cell protein nestin was identified in this study as an NM-enriched protein. Nestin may play a role in the trafficking and distribution of intermediate filament proteins.
and potentially other cellular factors to daughter cells during progenitor cell division (37). Further experiments are needed to determine which of the differentially enriched NM proteins are important structural components of the NM, and which ones are core NM-associated proteins. A combined effort of proteomics, microscopy, and immunodetection will be needed to further unravel the molecular composition and architecture of the NM.

**Nuclear Matrix Proteins Regulated in Colorectal Adenoma and Carcinoma Tissues**

To show the potential of our work flow for investigation of the nuclear matrix in the context of chromosomal instability, we further analyzed the nuclear matrix proteome in an exploratory study of adenoma and MIN+ and CIN+ carcinoma tissues. Our analysis revealed 129 and 169 differential proteins in the comparisons adenoma versus all CRC samples and adenoma versus CIN+ CRC, respectively. For clarity, we discuss these proteins in three categories: (i) chromosomal instability; (ii) nuclear protein complexes, nuclear bodies, and structure; and (iii) oncogenes, tumor suppressors, and binding partners.

**Chromosomal Instability**

The most common form of genomic instability observed in CRC is chromosomal instability, the molecular basis of which remains to be determined. The HMG proteins are associated with chromosomes, albeit it not exclusively, and have been implicated, through different mechanisms, in both benign and malignant neoplasias. HMGB1 was previously identified in the NM fraction from colon cancer tissue by 2DGE, and these results suggested enhanced localization to the NM in colon tumors relative to healthy tissue (38). In this study, we confirmed the up-regulation of HMGB1 in the NM fraction of...
chromosomal instable CRC and show, that besides HMGB1, HMGB2 and HMGB3 were also specifically up-regulated in the CIN+ CRC samples with HMGB3 up-regulation also being reflected at the mRNA level. HMGB1/2 have been implicated in cellular response to chemotherapy-induced DNA damage (39) and can up-regulate topoisomerase II (40), providing a possible link to chromosomal instability. Indeed, overexpression of HMGB1 induced unbalanced chromosomal rearrangements in human prostate cancer cell lines (41). Furthermore, HMGB1 and topoisomerase II co-localize at AT-rich scaffold/matrix-attached regions that are located at the base of DNA loop domains and thus present juxtapositions for chromosomal rearrangements (42).

B23/nucleophosmin (NPM) is a protein that is often encountered in NM preparations and has been implicated, among others, in pre-rRNA processing and ribosomal subunit export from the nucleus (43, 44) as well as centrosome duplication (45). Increased NPM expression has been linked to neoplastic colorectal mucosa (both adenomas and carcinomas) (46). We did not find a statistically significant difference between the NPM levels in the NM fraction of adenoma versus chromosomal instable carcinoma, but several of the NPM-interacting factors identified by Maggi et al. (44) are present in our list of differential proteins, being either up-regulated (NCL and DHX9) or down-regulated (CPSF6, hnRNPH2, DDX5, and Nup62). A putative NMP interactor that is highly differential between adenomas and carcinomas in our data set is PIN4/PAR14/parvulin, a peptidyl-prolyl cis-trans isomerase also harboring a DNA binding domain. This protein has been implicated in nucleolar pre-rRNA processing (47). Both PIN4 and NPM can be detected in pre-rRNP complexes, and PIN4 co-localizes with NMP both in interphase nucleoli and in mitotic spindles (47). In view of the latter observation and as PIN4 seems to accumulate at chromosomes during mitosis, it might have a role in proper partitioning of material (pre-rRNA, prenucleolar bodies, or chromosomes) during cell

---

**Fig. 7.** Protein-protein interaction map of 93 nuclear and/or NM-enriched colorectal CIN cancer-associated proteins. Solid circles indicate up-regulated proteins.
division. PIN4 does not seem to be enriched in NM preparations with respect to chromatin-binding and intermediate filament fractions. In view of the DNA binding and subcellular shuttling activities of PIN4 (48), this may not come as a surprise.

**Nuclear Protein Complexes, Nuclear Bodies, and Structure**

*Splicesosome*—Many NM proteins identified before and many of the differential proteins identified here, especially hnRNPs, have been implicated in splicesosome structure and nucleosome function (49, 50); some of them are up-regulated, whereas others are down-regulated. An association with cancer (according to literature in the respective entries in GeneCards) exists for many hnRNPs, KHSRP, DHX9, DDX5, SFRS1, and SFRS6, leaving the other differential proteins (e.g. SYNCRIP and hnRNPA/B) as potential novel candidates. In a recent proteomics study (51), galectin-3 (Gal3) was shown to stabilize SYNCRIP to maintain proliferation of human colon cancer cells. Interestingly, we found, in addition to up-regulation of SYNCRIP, also up-regulation of Gal3, underscoring a potential interaction between these two proteins in clinical material.

*Nucleolus*—The nucleolus is a major component of biochemical NM preparations. There is evidence that the nucleolus is the mirror of a series of metabolic changes that characterize cancer cells. Cell entry into the cell cycle is always associated with up-regulation of the nucleolar function and increased nucleolar size (52). Of the nucleolus-associated proteins with a cancer link, we note the strong up-regulation of NCL, a well known, established NM component. In addition, we detected several other cancer-related NM proteins (Table IV) with a previously reported association to the nucleolus (53), among others NOP5/NOP58 (up in CIN), S100 A-8 and A-9 (up in CIN; also with a corresponding mRNA change), HP1BP3 (up in CIN), cysteine- and glycine-rich protein 1 (CSR1; NM-enriched and up in CIN), 40 S ribosomal protein S19 (RPS19; up in CIN) and 28 S ribosomal protein S12 (RPS12; up in CIN).

**Nuclear Pore Complex**—The nuclear pore complex (NPC) is closely linked via lamins to (the rest of) the NM. A striking observation of this study is that many NPC proteins were down-regulated in the carcinoma samples, including NUP1L (which is also an NM-enriched protein), NUP214, NUP98, NUP133, NUP205, NUP88 (also mRNA down), NUP85, NUP93, NUP62 (also mRNA down), NUP54, NUP160, and NUP50 (Tables III and IV). This massive down-regulation of multiple NPC proteins may indicate cellular and nuclear crisis and loss of nuclear integrity. Interestingly, NUP98 is also a modulator of key spindle checkpoint proteins such as Mad and Bub proteins and has been linked to chromosomal instability (for a review, see Ref. 54). Together with TMPO, a putative lamin receptor in the nuclear envelope (which is up-regulated in CIN+ CRC samples), the NPC may form essential anchor points for the nuclear lamina, the outer shell of the NM scaffold.

**Oncogenes, Tumor Suppressors, and Binding Partners**

Although their relationship to the NM is unclear as yet, catenins α-1, β-1, and δ-1 were all down-regulated in the NM fraction of CRC samples. The adenomatous polyposis coli/β-catenin signaling pathway is of high importance in colon cancer (55, 56). Catenin β-1 is involved in the regulation of cell adhesion and in signal transduction through the Wnt pathway. Catenin δ-1 has been implicated in both cell transformation by SRC and ligand-induced receptor signaling through various receptor tyrosine kinases. Together, the differential regulation of these proteins may result in an altered regulation of dynamic actin-based, cytoskeletal, and signaling activities in the CRC tissues. Another example of a protein up-regulated specifically in CIN+ CRC and highly enriched in the NM fraction is FUBP1 (or FUSE-binding protein 1), a DNA helicase that regulates MYC expression by binding to a single-stranded far upstream element (FUSE) upstream of the c-myc promoter (57). It has generally been accepted that c-myc plays an important role in CRC progression, but its exact activators remain poorly understood. Here we identify FUBP1 as a novel candidate. Finally, as another example of an important CIN+ cancer-associated protein found in our NM preparations, the TP53 interactor YBX1 may be mentioned. YBX1 expression is correlated with tumor stage, drug resistance, and patient prognosis for many human tumors, including malignancies of the colorectum (58). A recent study correlated increased expression of YBX1 together with reduced E-cadherin expression to poor patient survival in invasive breast cancer (59). Here we also found correlation of up-regulation of YBX1 with down-regulation of E-cadherin in the CIN+ carcinomas.

**Conclusion**

We propose that the protocol used here for differential extraction of subnuclear fractions can be used for comparative protein complex profiling in small tissue samples. Important changes in tumor cells are alterations of nuclear structure and chromatin texture. We are the first to provide an extensive analysis of the protein constituents that are reproducibly detected as well as enriched in the NM fraction, and in a proof-of-concept study, we have pinpointed components that exhibit quantitative differences in chromosomal instable colorectal carcinomas (85% of CRC) relative to adenomas and microsatellite instable carcinomas. We anticipate that more extensive nano-LC-MS/MS-based subnuclear proteomics of patient tumor tissue in various stages of progression from colorectal adenoma to chromosomal instable carcinoma, combined with microscopy follow-up, could in future studies establish the missing link between diagnostic investigation of tissue in the pathology laboratory and the underlying cancer genetic mechanisms.
REFERENCES

1. Liotta, L. A., Kohn, E. C., and Petricoin, E. F. (2001) Clinical proteomics: personalized molecular medicine. JAMA 286, 2211–2214

2. Jimenez, C. R., Piersma, S. R., and Pham, T. V. (2007) High-throughput and targeted in-depth mass spectrometry-based approaches for biofluid profiling and biomarker discovery. Biomark. Med. 4, 541–565

3. Fey, E. G., and Penman, S. (1988) Nuclear matrix proteins reflect cell type and cell-cycle progression. J. Cell Sci. 98, 1973–1984

4. Fey, E. G., and Pennman, S. (1988) Nuclear matrix proteins reflect cell type of origin in cultured human cells. Proc. Natl. Acad. Sci. U.S.A. 85, 121–125

5. Nickerson, J. A., Krockmalnic, G., Wan, K. M., and Pennman, S. (1997) The nuclear matrix revealed by eluting chromatin from a cross-linked nucleus. Proc. Natl. Acad. Sci. U.S.A. 94, 4446–4450

6. Fey, D. S. (2002) Nuclear matrix proteins as proteomic markers of neoplastic and cancer lesions. Clin. Cancer Res. 8, 3031–3033

7. Leman, E. S., and Getzenberg, R. H. (2008) Nuclear structure as a source of cancer specific biomarkers. J. Cell. Biochem. 104, 1988–1993

8. Sjäkte, N., Sjäkte, T., and Vikmanis, U. (2004) Role of the nuclear matrix proteins in malignant transformation and cancer diagnosis. Exp. Oncol. 26, 170–178

9. Schouten, J. P., McGlunn, C. J., Waaijer, R., Zwijnenburg, D., Diepvens, F., and Pals, G. (2002) Relative quantification of 40 nuclear acid sequences by multiple ligation-dependent probe amplification. Nucleic Acids Res. 30, e57

10. Buffart, T. E., van Grieken, N. C., Tijssen, M., Coffa, J., Ylstra, B., Grabosch, H. I., van de Velde, C. J., Carvalho, B., and Meijer, G. A. (2009) High resolution analysis of DNA copy-number aberrations of chromosomes 8, 13, and 20 in gastric cancers. Virchows Arch. 455, 213–223

11. Coffa, J., van de Wiel, M. A., Diosdado, B., Carvalho, B., Schouten, J., and Meijer, G. A. (2008) MLPAssay: data analysis tool for reliable automated normalization of MLPA fragment data. Cell. Oncol. 30, 323–335

12. Denver, C., Postma, C., Carvalho, B., van den Bosch, S. M., Moerkerk, P. T., Hoekstra, H. J., Weijenberg, M. P., de Bruiñe, A. P., Meijer, G. A., and van Engeland, M. (2008) Integrated analysis of chromosomal, microsatellite and epigenetic instability in colorectal cancer identifies specific associations between promotor methylation of pivotal tumour suppressor and DNA repair genes and specific chromosomal alterations. Carcinogenesis 29, 434–439

13. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal. Chem. 68, 662–666

14. Keller, A., Nesvizhskii, A. I., Koller, E., and Aebersold, R. (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal. Chem. 74, 5383–5392

15. Nesvizhskii, A. I., Keller, A., Koller, E., and Aebersold, R. (2003) A statistical model for identifying proteins by tandem mass spectrometry. Anal. Chem. 75, 4653–4664

16. Schirmer, E. C., Florens, L., Guan, T., Yates, J. R., 3rd, and Gerace, L. (2003) Nuclear membrane proteins with potential disease links found by directed proteomic analysis of the human nucleus. Curr. Biol. 12, 1–11

17. Albrethsen, J., Knol, J. C., and Jimenez, C. R. (2009) Unraveling the nuclear matrix proteome. J. Proteomics 72, 71–81

18. Lengauer, C., Kinzler, K. W., and Vogelstein, B. (1997) Genetic instability in colorectal cancers. Nature 386, 623–627

19. Albers, A. J., Langen, R., Hess, P., Schmutzler, R. K., and Pommier, Y. (2003) Identification of DNA repair genes and specific chromosomal alterations between promoter methylation of pivotal tumour suppressor and DNA repair genes and specific chromosomal alterations. Cancer Res. 63, 850–858

20. Schotman, B. J., van der Heijden, B., Pathak, A., de Boer, M. F., van der Burg, P. A., and de Vries, P. L. (2006) Alterated expression and localization of creatine kinase B, heterogeneous nuclear ribonucleoprotein F, and high mobility group box 1 protein in the nuclear matrix associated with colon cancer. Cancer Res. 66, 763–769

21. Vliegenthart, J. F. R., van der Flier, M. J., Joosten, A. E., van der Loo, E., van der Kogel, A. J., and Oosterom, A. T. (2003) Chromatin-associated proteins HMG1/2 and PDIa trigger cellular response to chemotherapy-induced DNA damage. Mol. Cancer Ther. 2, 864–872

22.Strøm, M., Polanska, E., Struncová, S., and Pospisilová, S. (2009) HMG1 and HMG2 proteins up-regulate cellular expression of human topoisomerase I[γ]: a candidate architectural protein for chromosomal rearrangements in prostate cancer cells. Cancer Res.
62, 647–651
42. Bode, J., Benham, C., Ernst, E., Knopp, A., Marschalek, R., Strick, R., and Strissel, P. (2000) Fatal connections: when DNA ends meet on the nuclear matrix. J. Cell. Biochem. Suppl. 35, 3–22
43. Savkur, R. S., and Olson, M. O. (1998) Preferential cleavage in pre-ribosomal RNA by protein B23 endoribonuclease. Nucleic Acids Res. 26, 4508–4515
44. Maggi, L. B., Jr., Kuchenruether, M., Dadey, D. Y., Schwope, R. M., Grisendi, S., Townsend, R. R., Pandolfi, P. P., and Weber, J. D. (2008) Nucleophosmin serves as a rate-limiting nuclear export chaperone for the mammalian ribosome. Mol. Cell. Biol. 28, 7050–7065
45. Okuda, M., Horn, H. F., Tarapore, P., Tokuyama, Y., Smulian, A. G., Chan, P. K., Knudsen, E. S., Hofmann, I. A., Snyder, J. D., Bove, K. E., and Fukasawa, K. (2000) Nucleophosmin/B23 is a target of CDK2/cyclin E in centrosome duplication. Cell 103, 127–140
46. Nozawa, Y., Van Belzen, N., Van der Made, A. C., Dinjens, W. N., and Bosman, F. (2006) Systematic genome-wide annotation of spliceosomal proteins reveals differential gene family expansion. Genome Res. 16, 66–77
47. Reimer, T., Weiwad, M., Schierhorn, A., Ruecknagel, P. K., Rahfeld, J. U., Bayer, P., and Fischer, G. (2003) Phosphorylation of the N-terminal domain regulates subcellular localization and DNA binding properties of the peptidyl-prolyl cis/trans isomerase hPar14. J. Mol. Biol. 330, 955–966
48. Barbosa-Morais, N. L., Carmona-Fonseca, M., and Aparicio, S. (2006) Large-scale proteomic analysis of the human spliceosome. Genome Res. 12, 1231–1245
49. Yoo, B. C., Hong, S. H., Ku, J. L., Kim, Y. H., Shin, Y. K., Jang, S. G., Kim, I. J., Jeong, S. Y., and Park, J. G. (2009) Galectin-3 stabilizes heterogeneous nuclear ribonucleoprotein Q to maintain proliferation of human colon cancer cells. Cell. Mol. Life Sci. 66, 350–364
50. Derenzini, M., Montanaro, L., and Treré, D. (2009) What the nucleolus says to a tumour pathologist. Histopathology 54, 753–762
51. Shibao, K., Takano, H., Nakayama, Y., Okazaki, K., Nagata, N., Izumi, H., Uchiumi, T., Kuwano, M., Kohno, K., and Itoh, H. (1999) Enhanced coexpression of YB-1 and DNA topoisomerase II alpha genes in human colorectal carcinomas. Int. J. Cancer 83, 732–737
52. Evtodikinova, V., Vagner, C., Ng, T., Ruzanov, P., Melnyk, N., Fink, D., Sorokin, A., Ovchinnikov, L. P., Davicioni, E., Triche, T. J., and Sorensen, P. H. (2009) Translational activation of Snail1 and other developmentally regulated translation factors by YB-1 promotes an epithelial-mesenchymal transition. Cancer Cell 15, 402–415
53. Lee, D. C., Welton, K. L., Smith, E. D., and Kennedy, B. K. (2009) A-type nuclear lamins act as transcriptional repressors when targeted to promoters. Exp. Cell Res. 315, 996–1007