Cellular and Molecular Life Sciences

RESEARCH ARTICLE

HDAC2 and TXNL1 distinguish aneuploid from diploid colorectal cancers

Timo Gemoll · Uwe J. Roblick · Silke Szymczak · Till Braunschweig ·
Susanne Becker · Bernd-Wolfgang Igl · Hans-Peter Bruch · Andreas Ziegler ·
Ulf Hellman · Michael J. Difilippantonio · Thomas Ried · Hans Jörnvall ·
Gert Auer · Jens K. Habermann

Abstract DNA aneuploidy has been identified as a prognostic factor for epithelial malignancies. Further understanding of the translation of DNA aneuploidy into protein expression will help to define novel biomarkers to improve therapies and prognosis. DNA ploidy was assessed by image cytometry. Comparison of gel-electrophoresis-based protein expression patterns of three diploid and four aneuploid colorectal cancer cell lines detected 64 ploidy-associated proteins. Proteins were identified by mass spectrometry and subjected to Ingenuity Pathway Analysis resulting in two overlapping high-ranked networks maintaining *Cellular Assembly and Organization*, *Cell Cycle*, and *Cellular Growth and Proliferation*. CAPZA1, TXNL1, and HDAC2 were significantly validated by Western blotting in cell lines and the latter two showed expression differences also in clinical samples using a tissue micro-array of normal mucosa ($n = 19$), diploid ($n = 31$), and aneuploid ($n = 47$) carcinomas. The results suggest that distinct protein expression patterns, affecting TXNL1 and HDAC2, distinguish aneuploid with poor prognosis from diploid colorectal cancers.

T. Gemoll · U. J. Roblick · H. Jörnvall · J. K. Habermann
Department of Medical Biochemistry and Biophysics,
Karolinska Institutet, 17177 Stockholm, Sweden

T. Gemoll · U. J. Roblick · H.-P. Bruch · J. K. Habermann (✉)
Laboratory for Surgical Research, Department of Surgery,
University of Lübeck, Ratzeburger Allee 160,
23538 Lübeck, Germany
e-mail: Jens.Habermann@gmail.com

T. Gemoll · S. Becker · G. Auer
Karolinska Biomic Center, Karolinska Institutet,
17176 Stockholm, Sweden

T. Braunschweig
Institute for Pathology,
University Clinic RWTH Aachen, 52074 Aachen, Germany

Abbreviations

2-DE	Two-dimensional gel electrophoresis
AB	Antibody
AEC	3-amino-9-ethylcarbazole
Ambic	Ammoniumbicarbonate
CIN	Chromosomal instability

S. Szymczak · B.-W. Igl · A. Ziegler
Institute for Medical Biometry and Statistics,
University of Lübeck, 23538 Lübeck, Germany

U. Hellman
Ludwig Institute for Cancer Research Ltd.,
75124 Uppsala, Sweden

M. J. Difilippantonio · T. Ried
Department of Genetics, Center for Cancer Research,
NCI/NIH, Bethesda, MD 20814, USA

ECL	Enhanced chemiluminescence
HNPCC	Hereditary-non-polyposis-colorectal-cancer
IHC	Immunohistochemistry
IPA	Ingenuity pathways analysis
IPKB	Ingenuity pathways knowledge base
KEGG	Kyoto encyclopedia of genes and genomes
MMR	Mismatch repair
MSI	Microsatellite instability
RF	Random forest
SSP	Individual identification number for protein spots in analysis software
TMA	Tissue microarray
TNM	Classification of malignant tumors regarding tumor stage (T), lymph node metastasis (N), and distant metastasis (M)

Introduction

Colorectal carcinomas (CRC) constitute the third most common malignancy in the Western world. Two major molecular subtypes of CRC have been determined that are associated with distinct clinical prognosis: chromosomal instability (CIN) and microsatellite instability (MSI). MSI CRCs account for 15–20% of sporadic cancers and result from a loss of the DNA mismatch repair function. By failing to repair spontaneous errors that occur during replication, these tumors accumulate frame-shift mutations affecting tumor suppressor genes [1]. MSI tumors are known to present diploid tumor cell populations with few, if any, karyotypic abnormalities and show a favorable clinical prognosis [2, 3]. In contrast, CIN cancers are associated with an inferior prognosis [4]. These tumors are characterized by aneuploidy, i.e., tumor cell populations showing abnormal and highly scattered DNA stem line values higher than 5c. At the chromosomal level, aneuploid carcinomas show numerical chromosomal aberrations that define tumor entity specific patterns [5, 6]. Interestingly, a gain of chromosome 7 can already be observed in pre-malignant lesions of aneuploid colorectal carcinomas, e.g., in adenomas [7]. Such aberrations are faithfully maintained during disease progression and alter the protein expression [8, 9]. Aneuploid carcinomas (CIN type) account for about 80% of all sporadic colorectal carcinomas. Aneuploidy is associated with poor prognosis, disease recurrence, and death from colorectal cancer [4, 10]. Based on ploidy measurements of 260 CRCs, we could describe aneuploidy to be the strongest independent prognostic marker for colorectal cancer patients that underwent curative surgery (R0-resected) and were devoid of metastasis [11]. Our analysis showed that in addition to aneuploidy, age was also of prognostic value for R0-resected patients. Despite

these findings, neither causes nor consequences of aneuploidy for colorectal carcinogenesis have been sufficiently elucidated. Furthermore, no distinct proteins are known that define aneuploidy properties and are applied in a clinical setting.

Against this background we considered it highly important to identify aneuploidy-associated protein expression changes that could unravel novel targets for improved therapy options and prognosis markers. This was primarily accomplished by identifying aneuploidy-associated proteins by means of 2-DE analysis of colorectal cancer cell lines of defined ploidy types. It was further complemented by validation of identified candidate proteins using a tissue microarray containing clinical samples from diploid and aneuploid colon carcinomas with an overall 5-year follow-up survey. For overall study design, please see Fig. 1.

Materials and methods

Material

The human colorectal cancer cell lines DLD-1, HCT116, LoVo, COLO 201, HT-29, T84, and Caco-2 were
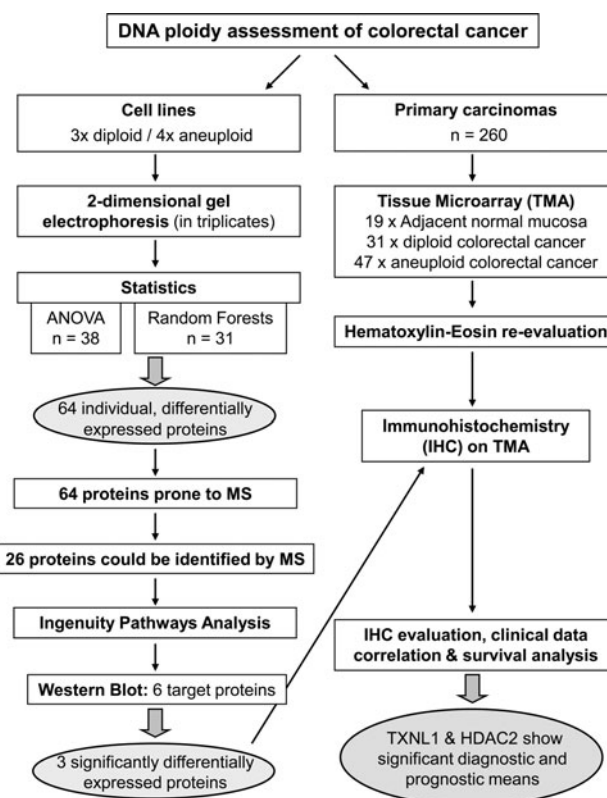
Fig. 1 Technical workflow of the study design [analysis of variance (ANOVA); mass spectrometry (MS)]

purchased from ATCC-Promochem® and cultured according to the manufacturers' recommendations (http://www.lgcstandards-atcc.org/, Supplemental Table 1) to a confluency of 80%. All mediums were supplemented with 1% (v/v) antibiotics (penicillin and streptomycin (#P11-010, PAA Laboratories GmbH, Pasching, Austria)) and 5–20% (v/v) unheated fetal bovine serum (#A15-151, PAA Laboratories GmbH, Pasching, Austria). Culture conditions for all cell lines were 37°C with 5% CO_2, except for DLD-1, LoVo, and COLO 201, that required no CO_2.

Patients

For clinical validation of differentially expressed proteins, colorectal carcinomas and adjacent normal mucosa of 78 patients were randomly selected from 260 patients that were part of a cohort of 1,840 patients undergoing surgery for colorectal malignancy at the Department of Surgery, University Clinic Schleswig-Holstein, Campus Lübeck, Germany, between 1994 and 2007. For the 260 patients, ploidy measurements were available as published by Gerling et al. [11]. Of those, 78 patients were randomly selected for tissue microarray validation. This subgroup, however, reflected the same prognostic means as the cohort of the 260 CRC patients. Clinical data comprised age, sex, post-operative TNM staging and grading (for an overview of patient cohorts see Table 1). Subsequent to operation, patients were seen on a regular basis in our clinic for post-operative surveillance (overall mean time of surveillance was $m = 4.74$ years; mean surveillance for patients included in survival analysis was $m = 4.08$ years). Clinical data are summarized in Table 2. Patients who underwent neoadjuvant radiotherapy for rectal tumors were excluded. Patients matching the Amsterdam II-criteria for the hereditary-non-polyposis-colorectal-cancer (HNPCC) syndrome were also excluded [12]. For survival analysis, patients with palliative treatment (R1 resection) and a survival of <30 days after surgery were excluded. Clinical material was collected from surgically removed tissue adhering to guidelines of the local ethical review board. Carcinoma tissue was used for touch preparation slides (imprints) for subsequent nuclear DNA content measurements by image cytometry. In addition, paraffin-embedded specimens of all 78 carcinomas were used for hematoxylin-eosin stained sections in routine histopathology assessment and for a tissue microarray in immunohistochemical evaluation of identified protein markers.

Methods

Genomic instability assessment

Genomic instability was assessed by nuclear DNA ploidy measurements by means of image cytometry using Feulgen stained cytospins (cell lines) and imprints (clinical specimens). The staining procedure, internal standardization, and cell selection criteria were based on methods described [13]. At least 500 nuclei per cell line and imprint were selected interactively and the DNA content was measured quantitatively using the ACAS imaging system (Ahrens ACAS, Hamburg, Germany). All DNA values were expressed in relation to the corresponding staining controls (lymphocytes), which were given the value 2c, denoting the normal diploid DNA content. The DNA profiles were classified according to Auer (Fig. 2a) [13].

Histograms characterized by a single peak in the diploid or near-diploid region (1.5–2.5c) were classified as type I. The total number of cells with DNA values exceeding the diploid region (>2.5c) was <10%. Type II histograms showed a single peak in the tetraploid region (3.5–4.5c) or peaks in both the diploid and tetraploid regions (>90% of the total cell population). The number of cells with DNA values between the diploid and tetraploid region and those exceeding the tetraploid region (>4.5c) was <10%. Type III histograms represented highly proliferating near-diploid cell populations and were characterized by DNA values ranging between the diploid and the tetraploid region. Only a few cells (<5%) showed more than 4.5c. DNA histograms of types I, II, and III thus characterize euploid cell populations. Type IV histograms showed increased (>5%) and/or distinctly scattered DNA values exceeding the tetraploid region (>4.5c) reflecting aneuploid populations of colon mucosa nuclei with decreased genomic stability. All DNA histograms were evaluated by three independent investigators who were unaware of the clinical and histopathological data of the patients.

Table 1 Patient cohorts

	Normal	Diploid SCC	Aneuploid SCC	Total
Total	19	31	47	97
Sex (male/female)	13/6	16/15	29/18	58/39
Average age at diagnosis (years)	68.89	69.58	71.43	69.97
Range (years)	55–92	54–92	54–88	54–92
Average observation time (months)		53.16	46.23	49.70
Maximum (months)		124.79	126.49	126.49

Table 2 Clinical parameters of 78 colorectal carcinomas

Value	Diploid SCC ($n = 31$)	Aneuploid SCC ($n = 47$)	Total ($n = 78$)
Tumor localization			
Right	12 (39%)	9 (19%)	21 (27%)
Left	9 (29%)	20 (43%)	29 (37%)
Rectum	10 (32%)	18 (38%)	28 (36%)
CEA-level at time of diagnosis			
Normal	14 (45%)	26 (55%)	40 (51%)*
High	5 (16%)	14 (30%)	19 (24%)*
ND	12 (39%)	7 (15%)	19 (24%)*
UICC stage			
1	7 (23%)	12 (26%)	19 (24%)*
2	11 (35%)	11 (23%)	22 (28%)*
3	13 (42%)	13 (28%)	26 (33%)*
4	0 (0%)	11 (23%)	11 (14%)*
T-status			
1	2 (6%)*	4 (9%)	6 (8%)*
2	10 (32%)*	10 (21%)	20 (26%)*
3	15 (48%)*	26 (55%)	41 (53%)*
4	4 (13%)*	7 (15%)	11 (14%)*
N-status			
0	18 (58%)	23 (49%)*	41 (53%)*
1	9 (29%)	12 (26%)*	21 (27%)*
2	4 (13%)	12 (26%)*	16 (21%)*
M-status			
0	31 (100%)	36 (77%)	67 (86%)
1	0 (0%)	11 (23%)	11 (14%)
Tumor grading			
1	0 (0%)	3 (6%)*	3 (4%)
2	21 (68%)	33 (70%)*	54 (69%)
3	10 (32%)	11 (23%)*	21 (27%)
Synchronous carcinoma	4 (13%)	7 (15%)	11 (14%)
Postoperative resection status			
0	30 (97%)	46 (98%)	76 (97%)
1	1 (3%)	1 (2%)	2 (3%)
Operative approach			
Curative	30 (97%)	35 (74%)	65 (83%)
Palliative	1 (3%)	12 (26%)	13 (17%)
Death within 30 days after operation	2 (6%)	2 (4%)	4 (5%)
Adjuvant therapy	17 (55%)	24 (51%)	41 (53%)

ND not determined

* Percentages were rounded up or down and deviate ±1 to 100%

Protein expression analysis

Protein extraction of colorectal cell lines

Colon cell lysates from diploid and aneuploid commercially available cell cultures were prepared in lysis buffer [5 mM sodium diphosphate ($Na_4P_2O_7$), 100 μM sodium orthovanadate (Na_3VO_4), 5 mM sodium fluoride (NaF), 830 μM benzamidine, 1 μg/ml aprotinin, 1 μg/ml leupeptin, and 1.74 μg/ml phenylmethylsulfonyl fluoride (PMSF)]. Briefly, after washing the cells twice with pre-chilled 1x PBS (4°C), the cells were scraped off into 2.5 ml of ice-cold lysis buffer

Fig. 2 a Example of a diploid and aneuploid DNA histogram type according to Auer; **b** tissue microarray (TMA) set-up of clinical colorectal carcinomas and adjacent normal mucosa samples for immunohistochemistry evaluation. *Blue* or *red* "N" highlights normal adjacent tissue of diploid (*blue*) and aneuploid (*red*) carcinomas. *Black crosses* mark necrotic, connective, or missing tissue that was not included in the evaluation
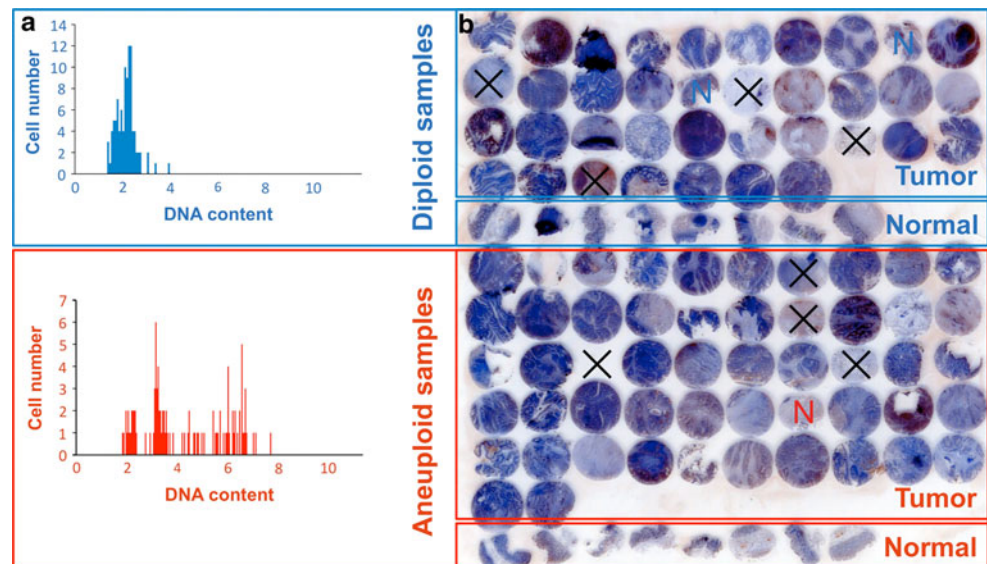
per T-175 flask. The cell solution was transferred to a pre-weighted Eppendorf tube and was centrifuged at 2,000 rpm for 3 min at 4°C. The pellet was dissolved in 1 ml of lysis buffer and re-centrifuged at 5,700 rpm for 5 min at 4°C. Prior to storage at −80°C and after discarding the supernatant, the wet weight (WW) of the pellet was determined. For protein-profiling, each pellet was thawed on ice and resuspended in 1.89 µl Milli-Q water per mg WW (1.89 × WW µl). The suspension was frozen and thawed four times. A volume of 0.089 × WW µl 10% SDS, including 33.3% mercaptoethanol, was mixed with the sample and incubated 5 min on ice with 0.329 × WW µl of a solution of DNase I (0.144 mg/ml 20 mM Tris-HCl with 2 mM $CaCl_2 \times 2\,H_2O$, pH 8.8) and RNase A (0.0718 mg/ml Tris-HCl). The sample was frozen and lyophilized. Sample buffer including 9 M urea, 0.15 mM PMSF, 0.5% Nonidet P-40, 0.36 mM 2, 6-di-butyl-4-methylphenol (BHT), 25 mM 3-(3-chloramido-propyl)-dimethylammonio-1-propanesulfonate (CHAPS), 1 M EDTA, 0.5 mM benzamidine and 65 mM dithiothreitol (DTT) was added carefully, mixed for 3 h, and centrifuged for 15 min at 10,000 rpm to remove insoluble material. Samples were stored at −80°C prior to protein determination.

Protein quantification according to Bradford

Protein concentrations of cell line extracts were determined in quadruplicates by addition of 25 µl concentrated Bio-Rad assay (Bio-Rad, CA, USA) to 1 µl solubilized sample diluted in 100 µl Milli-Q water using 96-well microplates [14]. A standard curve was constructed using different concentrations of bovine serum albumin. Plate reading was performed with a Multiscan reader (Labsystems, VA, USA).

Two-dimensional gel electrophoresis

After protein quantification, 2-DE was performed for all seven cell lines in triplicates as described [15, 16]. Briefly, all samples were diluted with rehydration buffer containing 7 M urea, 2 M thiourea, 1% CHAPS, 0.5% immobilized pH gradient (IPG) buffer, 0.3% (DTT) and a trace of bromophenol blue. In total, 75 µg of each sample was diluted in 300 µl rehydration buffer and applied to precast immobilized pH gradient strips (IPG), IPG 4–7 linear, 17 cm (Bio-Rad, CA, USA). The isoelectric focusing (IEF) was performed in a Protean IEF cell (Bio-Rad, CA, USA) at 20°C, reaching approximately 52,900 Vhs. The separation was performed stepwise, starting at 500 V (1 h), increasing from 1,000 V (gradient, 1 h) over 8,000 V (gradient, 3 h) to the final focusing step for 5.5 h at 8,000 V. After isoelectric focusing, the strips were immediately equilibrated for 2 × 15 min with 50 mM Tris-HCl, pH 8.8, in 6 M urea, 30% glycerol and 2% SDS. DTT (2%) was included in the first and iodoacetamide (2.5%) in the second equilibration step to reduce S–S bonds and alkylate-free thiols. 10–13% acrylamide gradient gels (1.5 × 200 × 230 mm) with 1% SDS were used for second-dimension gel electrophoresis. The IPG strip was placed on the surface of the second-dimension gel and sealed with 0.5% agarose in SDS-electrophoresis buffer (25 mM Tris base, 192 mM glycine, 0.1% SDS). The gels were run overnight at constant 100 V and 12°C to reach 2,000 Vhs.

Silver staining

After electrophoresis, the gels were fixed with 30% ethanol and 10% acetic acid for 2 × 30 min, incubated in

sensitization buffer [0.5 M potassium acetate (CH_3COOK), 20% ethanol, 8.3 M potassium tetrathionate ($K_2S_4O_6$)] and rinsed using six changes of Milli-Q water, 20 min each. After rinsing, the gels were incubated in silver nitrate [0.2% silver nitrate ($AgNO_3$), 5 mM HEPES, 0.026% formaldehyde ($HCHO$)]. The silver nitrate was discarded, and the gels were rinsed quickly twice with Milli-Q water, followed by development of the gels (0.22 M potassium carbonate (K_2CO_3), 0.05 mM sodium thiosulfate ($Na_2S_2O_3$), 0.009% formaldehyde) with intensive shaking for 10 min. The development was terminated with 5% acetic acid (CH_3COOH).

Image analysis

Silver-stained 2-DE gels were imaged using a flatbed scanner (GS 710, Bio-Rad, CA, USA), and were further analyzed by the PDQuest 2-DE analysis software (version 8.0.1, Bio-Rad, CA, USA). The analysis included protein spot detection, background subtraction, and quantification [17]. Each spot obtained an individual identification number (SSP). Matching results of the semi-automated gel-to-gel algorithm provided by the software were manually controlled in order to avoid and correct automated matching errors. The individual protein spot quantity was normalized to the total quantity of the valid spots present in the gel.

In-gel digestion and mass spectrometry

Digestion of the proteins in the excised spots was performed as described [18]. Gel pieces were de-stained, trypsin (80 ng in an appropriate volume of 50 mM ammonium bicarbonate/10% acetonitrile (ACN)) was added and incubation was carried out at 30°C over night. Digestion was stopped by acidification with trifluoroacetic acid (TFA) and resulting peptides were diluted with 20 μl 0.1% TFA. For matrix-assisted laser desorption/ionization (MALDI), the peptide extracts were concentrated and desalted with μC18 ZipTips (Millipore, MA, USA) by siphoning about ten times and washed twice with 15 μl 0.1% TFA. The tryptic fragments were eluted to the MALDI target with 75% ACN/0.1% TFA, containing half-saturated alpha-cyano-4-hydroxy cinnamic acid as matrix and analyzed by matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-ToF–MS) in an Ultraflex III TOF/TOF instrument (Bruker Daltonics, Bremen, Germany). MALDI-ToF mass spectrometry produced peptide mass fingerprints (up to 86 observed peptides per spot with a $S/N > 6$) were analyzed using the flexAnalysis v2.4 software (Bruker Daltonics, Bremen, Germany). Database searches were carried out with the software package *ProFound* (release January 2008,

http://prowl.rockefeller.edu) in the non-redundant protein database of the National Center for Biotechnology Information (NCBInr, v2008/01/01 with 5,869,058 sequences). "*Mammalia*" was chosen for the taxonomic category. All peptide masses were assumed to be monoisotopic and $[M + H]^+$ (protonated molecular ions). The search parameters included carbamidomethylation of cysteine as fixed modification, and oxidation of methionine as variable modification. One missed trypsin cleavage was allowed. Only protein hits with three or more matching peptide masses and a sequence coverage of at least 10% were considered. Internal calibration was achieved by analysis of autolytic trypsin cleavage products resulting in an accuracy of ±0.04 Da. Judgment of significance was based on the expectation value, number of matching peptide masses, and agreement between experimental and theoretical physical properties of the proteins.

Pathway analysis

As an exploratory approach, we used Ingenuity Pathways Analysis (IPA) software (www.ingenuity.com, Version 8.0-2602) to assess the involvement of differentially expressed proteins in pathways and networks. IPA is a widely utilized comprehensive database and software based on the Ingenuity Pathways Knowledge Base (IPKB) [19, 20]. The term "pathway" is used to refer to canonical pathways that annotate genes or proteins of interest to known metabolic or signaling cascades as described in the Kyoto Encyclopedia of Genes and Genomes (KEGG, www.genome.jp/kegg). Networks, however, are dynamically generated by IPA as groups of proteins that interact with a defined protein of interest, i.e., in our analysis differentially expressed proteins. IPA networks can include several genes or proteins and can also allow identification of indirect interactions between targets of interest. Thus, the uploaded proteins of interest are used to map molecular networks that indicate how these may influence each other above and beyond already known interaction maps as derived from, e.g., KEGG pathways. The IPA generated networks are ranked using a score that is calculated as the negative logarithm of the p value. This p value indicates the likelihood of the proteins of interest being found together in a common network due to random chance. A score ≥5 was considered to reflect a significant network.

Western blotting

Selected proteins based on the 2-DE/MALDI experiments were analyzed further using commercially available antibodies. A total of 15 μg of cell lysate was resolved by electrophoresis on a 7.5 or 16% SDS polyacrylamide gel (150 V) in a running gel buffer containing 25 mM Tris

base, 192 mM glycine, and 0.1% SDS. The samples were transferred to PVDF membranes for 80 min, at 150 mA. In order to control the protein transfer, each membrane was reversibly stained by Ponceau red. The membranes were then blocked for 60 min in 5% dry milk and were incubated overnight with the following antibodies (1 μg/ml): PRDX2 monoclonal antibody, clone 4E10-2D2 (Mouse monoclonal, Abnova, Taipei, Taiwan), HDAC2 monoclonal antibody, clone HDAC2-62 (Mouse monoclonal, GeneTex, TX, USA), TXNL1 monoclonal antibody, clone 1A10 (Mouse monoclonal, Abnova, Taipei, Taiwan), GNAS polyclonal antibody (Mouse polyclonal, Protein-Tech Group, MI, USA), CAPZA1 polyclonal antibody (Mouse polyclonal, Abnova, Taipei, Taiwan), YWHAQ monoclonal antibody, clone 3F7 (Mouse monoclonal, GeneTex, TX, USA), and beta-actin monoclonal antibody (Mouse monoclonal, Abnova, Taipei, Taiwan). After incubation of membranes with goat anti-mouse secondary antibody (1:2,000, Santa Cruz Biotech, CA, USA) final visualization was carried out with the ECL kit (Bio-Rad, CA, USA). Densitometric analysis was performed using QuantityOne software version 4.5.2 (Bio-Rad, CA, USA).

Tissue microarray-based immunohistochemistry

Based on ploidy status assessment of 260 colorectal cancer specimens described earlier [11], tissues of 31 diploid and 47 aneuploid colorectal carcinomas were selected as well as 19 adjacent normal mucosa specimens, for which sufficient paraffin-embedded tissue was available. These tissues were implemented into a tissue microarray as described [21, 22]. Briefly, small biopsies were retrieved from selected regions of donor paraffinized tissue blocks using hollow stainless-steel needles with an inner diameter of 1.5 mm. Cores were inserted into a standard-sized recipient array block. The TMA was constructed using a semiautomated arrayer (TMArrayer, Pathology Devices, MD, USA). For an overview of the TMA set-up, see Fig. 2b.

After construction of the array block, a hematoxylin-eosin stained 5-μm section was reviewed by one seasoned pathologist. For immunohistochemical analysis, 5-μm-thick sections were cut, heated to 60°C, deparaffinized, and rehydrated. The sections were blocked with peroxidase twice for 10 min (#S2001; company's replacement product is #S2003), peroxidase blocking reagent, DAKO, Glostrup, Denmark). Sections were incubated with primary antibodies against HDAC2 (monoclonal, GeneTex Inc., diluted 1:150), CAPZA1 (polyclonal, ProteinTech Group Inc., diluted 1:20), and TXNL1 (monoclonal, Abnova Corporation, diluted 1:125) overnight at 4°C. Staining was performed using the Strpt-ABComplex/HRP kit (#K0492; replacement product #K0675, DAKO, Glostrup, Denmark) and the peroxidase-specific substrate 3-amino-9-ethylcarbazole

(#K3464, DAKO, Glostrup, Denmark) according to the manufacturer's instructions. Immunopositivity of epithelial cells for each marker protein was assessed semiquantitatively as follows: 0, negative (no cells stained); 1, weakly positive (<20% of cells stained); 2, moderately positive (20–50% of cells stained); or 3, strongly positive (>50% cells stained). For CAPZA1 cytoplasmic staining was analyzed, for TXNL1 and HDAC2, nuclear staining. All stains were scored by three independent observers (T. G., T. B., J. K. H.) being unaware of sample assignments to ploidy group and clinical data. Inter-observer variability was measured using Kappa statistics.

Statistical analysis

2-DE protein expression data were pre-processed by log transformation and missing values were replaced with the aid of two different approaches. On the one hand, a missing value was substituted by the mean value of both remaining technical replications. On the other hand, a missing value was replaced with the minimum of all observed expression values. In order to measure agreement between replications, Bland–Altman plots were computed [23]. Based on the data pre-processing described above, two independent down-stream analyses were performed. First, an analysis of variance (ANOVA) with repeated measurements was calculated to determine the effect of ploidy, the replicate and their interaction on the observed expression value. Proteins with a p value for ploidy <0.05 and a $p > 0.05$ for the interaction term were selected for MS. Second, a random forest approach was used to classify cell lines as diploid or aneuploid [24]. This machine learning method ranks the protein spots according to their importance for classification. Variable importance was calculated using two methods based on the Gini index and permutation, respectively. In order to select the most important proteins, an iterative variable selection procedure was applied [25]. At each iteration, half of the variables with low importance scores were discarded and a new forest was built followed by re-estimation of variable importance. Only the most important variables based on the final forest with 25 variables were used for MS. Fold change was calculated as a ratio of the exponentiated weighted mean of median expression values in aneuploid over diploid cell lines. For analysis of Western-blot data, one-sided t tests were calculated with alternative hypotheses based on observed expression differences in 2-DE gel data.

Inter-individual observer agreement for TMA data was analyzed by overall agreement and kappa coefficients [26]. Samples remaining inconclusive after discussion were excluded from further analyses. One-sided Fisher's exact tests were used for comparison of aneuploid and diploid tumors. According to observed protein expression

differences in 2-DE gels, hypotheses were defined for Western-blot analysis. Cut-offs for grouping expression values into two categories were selected based on minimal p values. Survival curves were estimated by Kaplan–Meier and a log-rank test was calculated to compare survival in patients with aneuploid and diploid tumors. Patients that died within 30 days or received a palliative treatment were excluded. For each test, a significance level of 5% was used. Because of the explorative study design, p values were not adjusted for multiple comparisons. All statistical analyses were performed using the statistical software package R version 2.9.1.

Results

Genomic instability and ploidy-associated survival

DNA ploidy assessment was performed by means of image cytometry and in accordance with the Auer classification [13]. Ploidy assessment determined the colorectal cancer cell lines DLD-1, HCT116, and LoVo to be diploid. In contrast, Caco-2, HT-29, T84, and Colo 201 were classified as aneuploid (Table 3). Ploidy-type classification of cell lines was further supported by comparative genomic hybridization (CGH) and spectral karyotyping (SKY) analyses (Supplemental Fig. 1a–f and http://www.ncbi.nlm.nih.gov/sky/) [27, 28].

The clinical colorectal carcinoma samples were obtained from a larger patient cohort of 260 CRCs that were all assessed by image cytometry. In this cohort, aneuploidy proved to be the strongest independent prognostic marker for R0-resected colorectal cancer patients [11]. Out of this cohort, we selected 31 diploid and 47 aneuploid malignancies for which paraffin-embedded specimens were available (Table 1). These samples constituted the tissue microarray used for clinical validation of identified protein biomarkers (Fig. 2b). In this cohort, aneuploid colorectal carcinomas are presented with an inferior survival ($p = 0.04$) (Fig. 3). Thus, our tissue

Table 3 DNA image cytometry analysis of colorectal cell lines

Cell line	DNA content of the stem line	5c ER (%)	Analysis of the histogram
LoVo	2.07	1.1	Diploid
DLD-1	2.03	3.8	Diploid
HCT116	2.08	0.7	Diploid
HT-29	3.64	10.8	Aneuploid
COLO 201	3.29	90.8	Aneuploid
T84	2.99	27.6	Aneuploid
Caco-2	4.00	48.1	Aneuploid

5c ER 5c exceeding rate
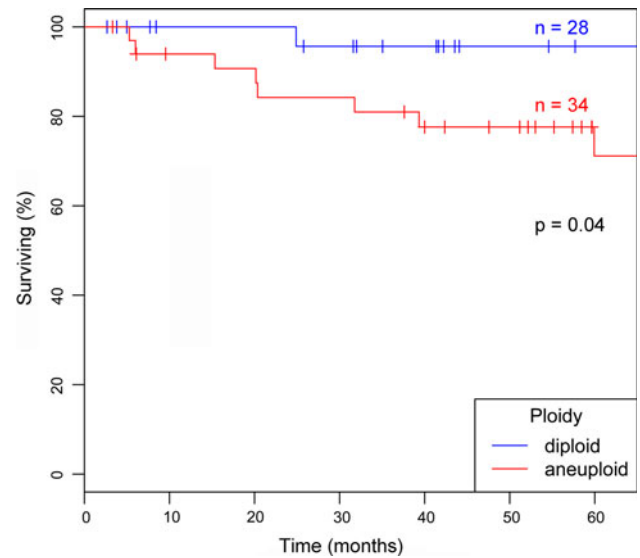
Fig. 3 Kaplan-Meyer 5-year survival curves of diploid and aneuploid colorectal carcinoma patients

microarray was sufficient for IHC validation of aneuploidy-associated protein expression changes of prognostic value.

Protein expression analysis

*Two-dimensional gel-electrophoresis (2-DE) and mass spectrometry*

PDQuest-based imaging of the gels detected on average 1,200 protein spots within a single gel (range 978–1,432) while achieving a match-rate of 73%. Bland–Altman plots were computed and demonstrated a low degree of heterogeneity between replicates (data not shown). Two independent statistical analyses were performed revealing 38 (ANOVA analysis) and 31 (random forest) protein spots that were differentially expressed between the diploid and aneuploid cell lines. A total of five protein spots were common to both statistical methods resulting in a total of 64 differentially expressed protein spots. All spots were subsequently subjected to MALDI-ToF MS. No conclusive identity could be obtained for 38 spots, with 20 of them showing only faint silver staining in the 2-DE gels. The remaining 26 spots were identified by peptide mass fingerprinting and scanning the latest version of NCBI's sequence database with the obtained mass data. Eight proteins were higher and 18 lower expressed in the aneuploid relative to the diploid cell lines (Fig. 4, Supplemental Figs. 2, 3).

Pathway analysis

In order to identify common networks and canonical pathways containing the 26 differentially expressed proteins, we used Ingenuity Pathway Analyses (IPA).

Fig. 4 **a** Representative gel image with highlighted protein spots of the 26 identified proteins separated by their isoelectric point (p*I*) and relative molecular mass (Mr) in kilodalton (kDa). **b** Identified proteins of the differentially expressed protein spots. The *p* value is given when ANOVA (analysis of variance) analysis had identified a differentially expressed protein spot. In all other cases, random forests analysis (RF) was performed. Identified proteins highlighted in *bold* were selected for subsequent Western blotting
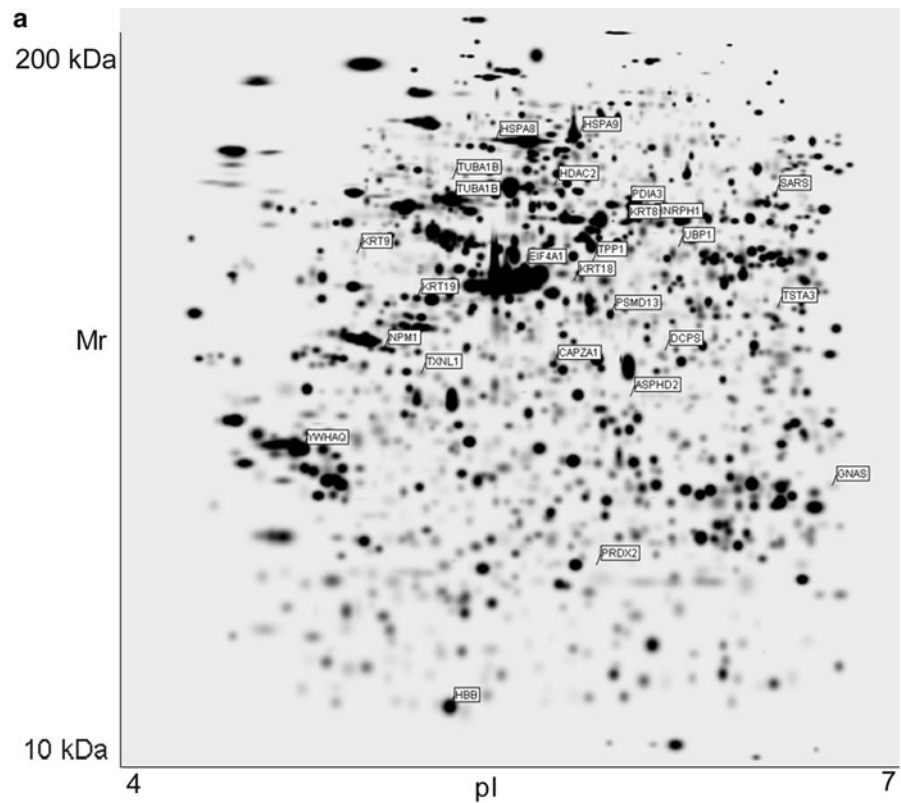
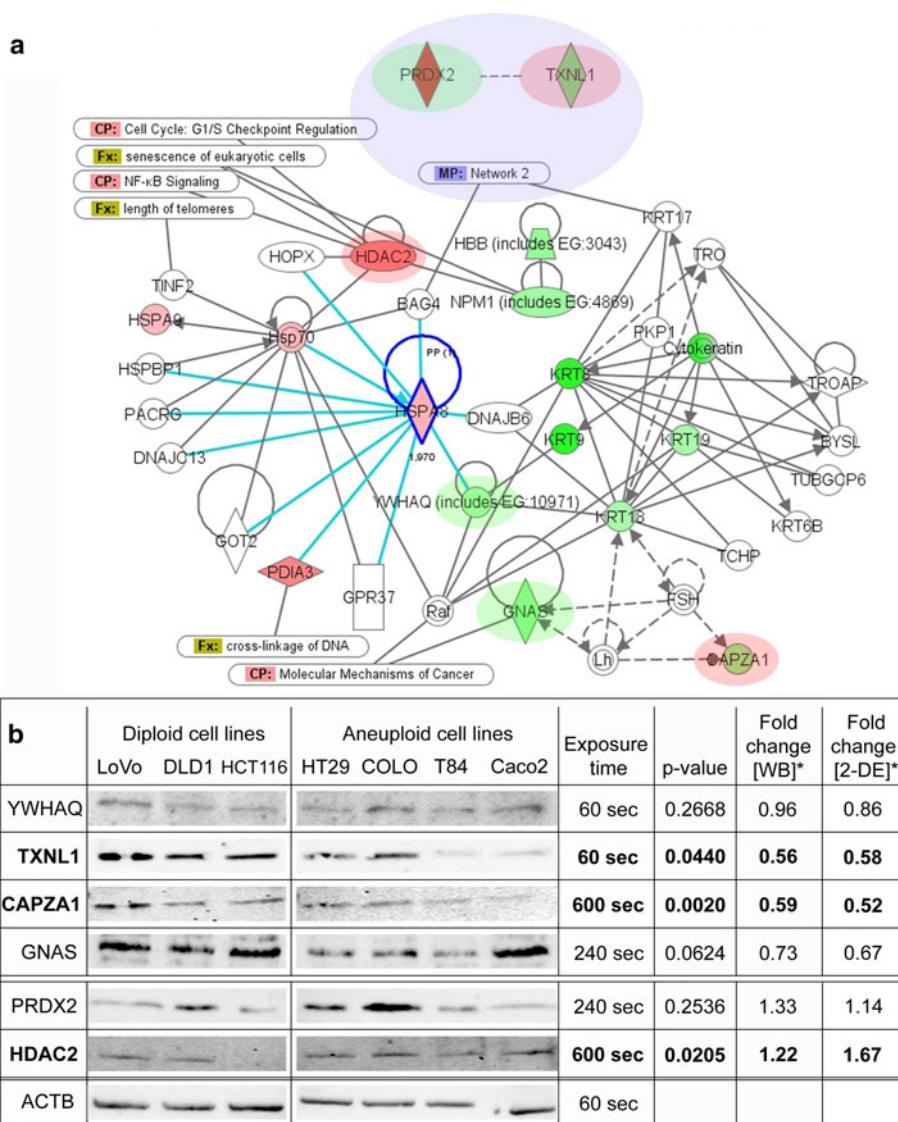| b Protein name	Gene symbol	Accession number	Chromosomal location	Sequence coverage (%)	Matched/ searched peptides	Score	Theoretical Mr (kDa)	Observed Mr (kDa)	Statistic
Lower expression in aneuploid samples									
14-3-3	YWHAQ	NP_006752	2p25.1	53	12/58	2.5×10^{-4}	29.33	30	p = 0.0046
Thioredoxin-like 1	**TXNL1**	**NP_004777**	**18q21.31**	**31**	**5/48**	**0.0028**	**32.63**	**40**	**p = 0.0195**
Capping protein muscle Z-line, alpha 1	CAPZA1	NP_006126	1p13.2	24	4/42	0.49	33.08	40	RF
Aspartate beta-hydroxylase domain containing 2	ASPHD2	NP_065170	22q12.1	10	4/38	0.49	38.86	40	p = 0.0081
Tissue specific transplantation antigen P35B	TSTA3	NP_003304	8q24.3	15	4/52	0.63	36.10	50	RF
Nucleophism 1	NPM1	NP_954654	5q35	29	7/52	8.7×10^{-3}	29.62	40	RF
Proteasome 26S subunit	PSMD13	NP_002808	11p15.5	27	9/25	3.5×10^{-6}	43.21	50	RF
Cytokeratin 8	KRT8	NP_002264	12q13	18	9/27	1.4×10^{-4}	53.54	65	p = 0.0051
Cytokeratin 9	KRT9	CAA52924	17q21.2	12	4/36	0.49	62.20	60	p = 0.0345
Cytokeratin 18	KRT18	CAA31377	12q13	18	5/37	0.060	47.32	55	p = 0.0458
Cytokeratin 19	KRT19	NP_002267	17q21.2	27	9/28	3.3×10^{-5}	44.09	55	RF
Hemoglobin, beta	HBB	NP_000509	11p15.5	30	3/43	0.47	11.54	10	RF
Tripeptidyl peptidase I	TPP1	AAQ72732	11p15	19	4/48	0.011	34.67	60	RF
Decapping enzyme, scavenger	DCPS	AAH14532	11q24.2	20	6/86	0.78	38.61	40	RF
Upstream binding protein 1	UBP1	NP_055332	3p22.3	17	7/86	0.20	60.46	60	RF
Seryl-tRNA-synthetase	SARS	AAH09390	1p13.3	17	8/47	7.0×10^{-3}	59.28	70	RF
Eukaryotic translation initiation factor 4A, isoform 1	EIF4A1	NP_001407	17p13	31	9/29	4.7×10^{-6}	46.36	60	RF
GNAS complex locus	**GNAS**	**NP_536351**	**20q13.3**	**19**	**4/53**	**0.28**	**44.70**	**30**	**RF**
Higher expression in aneuploid samples									
HSP70 protein 8	HSPA8	NP_006588	11q24.1	17	8/57	5.7×10^{-3}	71.11	90	p = 0.0318
Histone deacetylase 2	**HDAC2**	**NP_001518**	**6q21**	**15**	**6/39**	**0.054**	**55.92**	**70**	**p = 0.0496**
Protein disulfide isomerase family A, member 3	PDIA3	BAA11928	15q15	20	7/72	0.020	57.18	70	p = 0.0084
Peroxiredoxin 2	**PRDX2**	**NP_005800**	**19p13.2**	**15**	**3/44**	**0.38**	**22.05**	**25**	**p = 0.0467**
HSP70 protein 9	HSPA9	NP_004125	5q31.1	26	18/58	8.6×10^{-5}	74.12	90	RF
Tubulin, alpha (SSP # 2609)	TUBA1B	NP_006073	12q13.12	52	18/40	4.9×10^{-16}	50.82	70	p = 0.0278
Tubulin, alpha (SSP # 2719)	TUBA1B	NP_006073	12q13.12	29	8/50	2.1×10^{-4}	50.82	70	RF
Heterogeneous nuclear ribonucleoprotein H1	HNRPH1	NP_005511	5q35.3	19	6/73	0.72	49.40	65	p = 0.0401

Networks with a score ≥5 were considered significant. A total of 24 of the 26 identified proteins were recognized by the Ingenuity knowledge database as being eligible for network creation. Based on these, a total of three networks were projected by the software: the highest ranked network reached a score of 32 and included 13 of the 24 differentially expressed proteins, namely CAPZA1, GNAS, HBB, HDAC2, HSPA8, HSPA9, KRT8, KRT9, KRT18, KRT19, NPM1, PDIA3, YWHAQ (Fig. 5a). This network was significantly associated with functions regarding *Cellular Assembly and Organization* and *Cellular Function and Maintenance*. The differentially expressed protein HSPA8 was identified as a major central node of the first network. The second network reached a score of 23 and included ten

Fig. 5 a IPA network 1 (score 32) of 13 proteins differentially expressed between diploid and aneuploid colorectal cell lines. *Red* highlighted proteins are higher expressed in the aneuploid group and *green* ones are lower expressed in the aneuploid group. Please note the involvement of the differentially expressed protein HDAC2 in the canonical pathways (CP) of the cell cycle and NF-kB signaling and the overall association of this network with functions (FX) related to cell cycle checkpoints and senescence. This network overlaps with network 2 (MP) (*blue*; score 23 with ten differentially expressed proteins). Proteins accompanied by *red* and *green ovals* were prone to Western-blot evaluation with the *red* ones reaching significance.
b Western-blot analysis of selected target proteins confirmed TXNL1, CAPZA1, and HDAC2 as significantly differentially expressed proteins (highlighted in *bold*). TXNL1 and CAPZA1 are significantly lower expressed in aneuploid cell lines whereas HDAC2 is significantly higher expressed in aneuploid cell lines (*Fold change calculation for western blot (WB) and two-dimensional gel electrophoresis (2-DE): aneuploid over diploid)

b	Diploid cell lines			Aneuploid cell lines				Exposure time	p-value	Fold change [WB]*	Fold change [2-DE]*
	LoVo	DLD1	HCT116	HT29	COLO	T84	Caco2				
YWHAQ								60 sec	0.2668	0.96	0.86
TXNL1								**60 sec**	**0.0440**	**0.56**	**0.58**
CAPZA1								**600 sec**	**0.0020**	**0.59**	**0.52**
GNAS								240 sec	0.0624	0.73	0.67
PRDX2								240 sec	0.2536	1.33	1.14
HDAC2								**600 sec**	**0.0205**	**1.22**	**1.67**
ACTB								60 sec			

differentially expressed proteins (EIF4A1, HNRPH1, PRDX2, PSMD13, SARS, TPP1, TSTA3, TUBA1B, TXNL1, UBP1) (Supplemental Fig. 4). MYC, TGF-β, and NFκB were central nodes of this network, which is associated with functions of *Infection Mechanisms*, *Cell Cycle,* and *Cellular Growth and Proliferation*. The first and second networks are connected to each other via interactions with BAG4 and KRT17. Proteins of both networks showed cancer-associated functions of *Cellular Assembly and Organization*, *Cell-To-Cell Signaling,* and *Cell Death* ($p < 0.00001$ to $p = 0.0481$). The third network with a score of 3 did not reach significance and included only one differentially expressed protein (DCPS).

Western blotting

Based on IPA analysis, fold changes, molecular functions, and availability of antibodies, YWHAQ, CAPZA1, GNAS, PRDX2, HDAC2, and TXNL1 were selected for downstream analysis by Western blotting. Two of these six proteins had a higher and four a lower expression in the aneuploid cell lines relative to the diploid lines. Western-blot expression showed one specific band for YWHAQ, TXNL1, CAPZA1, and HDAC2, while multiple bands could be observed for GNAS (polyclonal AB) and PRDX2 (monoclonal AB) (Supplemental Fig. 5). Western-blot fold changes of all six proteins were in accordance with 2-DE data. However, only three of the six proteins reached a significance of $p < 0.05$ having either lower (TXNL1, CAPZA1) or higher expression (HDAC2) in the aneuploid group (Fig. 5b). For GNAS (polyclonal AB with multiple bands), the p value was marginally larger than 0.05 ($p = 0.062$). None of the additional bands of the polyclonal GNAS antibody showed significant differences between the diploid and aneuploid group. Densitometric analysis of Western-blotting expression for YWHAQ and

PRDX2 did not reach the significance level ($p > 0.05$). Of the six proteins evaluated, PRDX2 presented the lowest fold change of exploited proteins and multiple bands in Western blotting. None of the bands showed any significant differences between ploidy groups. YWHAQ showed the second lowest fold change in 2-DE and Western blotting with 0.89 and 0.96, respectively (Fig. 5b).

Validation by tissue microarray

For proof of clinical relevance by tissue microarray (TMA)-based immunohistochemistry, we chose only those three proteins whose significant differential expression was validated by Western blotting. Representative staining scores of HDAC2, TXNL1, and CAPZA1 are shown in Fig. 6a. The inter-individual observer agreement was measured by Kappa statistics showing good overall agreement for scoring of CAPZA1 (83%, $\kappa = 0.77$), HDAC2 (88%, $\kappa = 0.68$) and TXNL1 (87%, $\kappa = 0.63$). As shown in Fig. 6b, HDAC2 nuclear immunopositivity

(score 1, 2, and 3) was more frequently present in aneuploid (91.3%) than in diploid (70%) carcinomas ($p = 0.02$). Sensitivity was 91.3% and specificity 34.6%. For TXNL1, strong nuclear immunoreactivity (score 3) was more frequently observed in diploid (24%) than in aneuploid carcinomas (6.4%, $p = 0.04$). Sensitivity reached 93.4% and specificity 26.9%. While immunohistochemistry in clinical samples were in agreement with our cell line Western-blot results for HDAC2 and TXNL1, tissue microarray analysis for CAPZA1 did not and lacked statistical significance ($p = 0.19$). However, CAPZA1 immunohistochemistry showed a similar trend as in the Western-blot analysis with diploid carcinomas showing a higher grade of immunopositivity (score 2 and 3). Sensitivity amounted to 53.3% and specificity to 56.0%.

By comparing immunopositivity between normal adjacent mucosa and all carcinoma specimens irrespective of their ploidy status, we detected significantly stronger immunopositivity in the carcinomas for HDAC2 ($p < 0.001$)
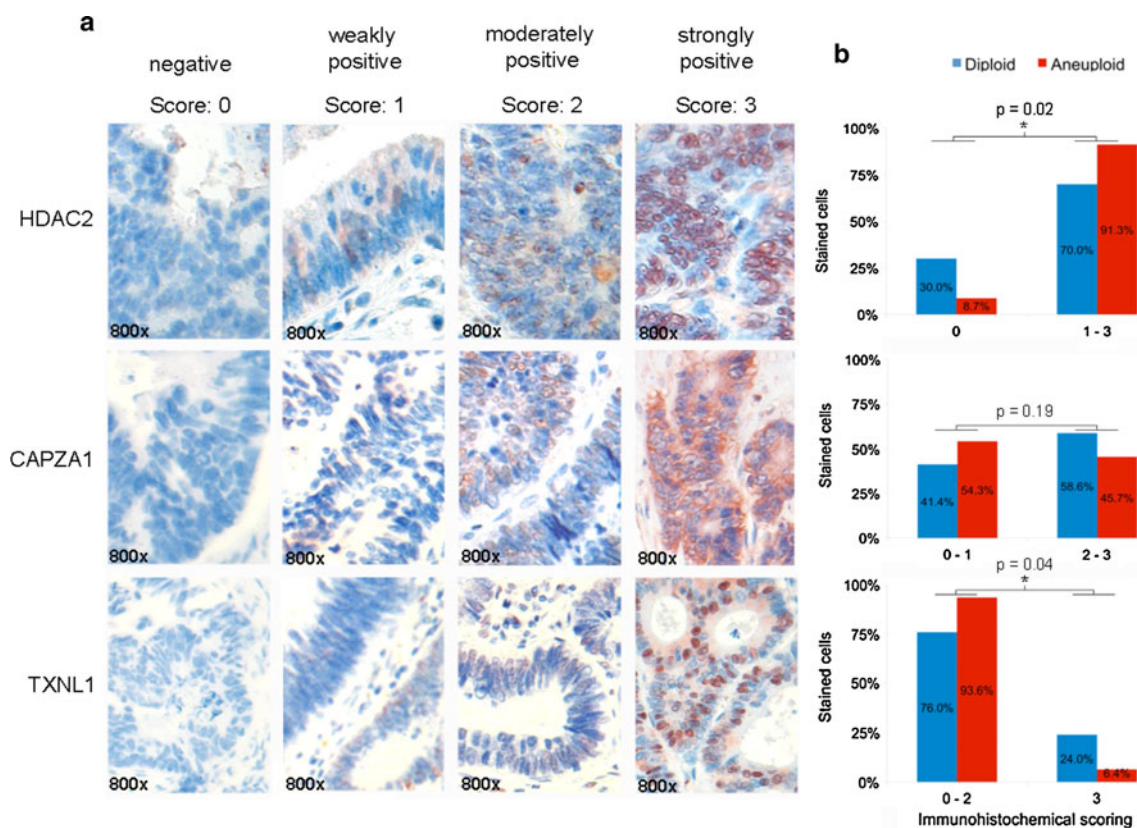
Fig. 6 a HDAC2, CAPZA1, and TXNL1 immunohistochemical detection in colorectal cancer specimens based on a tissue microarray. Image examples are given at a 800-fold magnification (800×). Immunopositivity for individual protein markers was assessed semiquantitatively as described in "Materials and methods". **b** Tissue-microarray-based immunohistochemical evaluation of HDAC2, CAPZA1, and TXNL1 comparing diploid versus aneuploid colorectal carcinoma specimens. Immunoreactivity was scored with "0" showing no positivity, "1" presenting up to 20% immunopositive cells, "2" up to 50%, and "3" above 50% stained cells. Barplots of the TMA analysis confirmed HDAC2 and TXNL1 as significantly (*asterisk*) differentially expressed proteins between diploid and aneuploid tumors

and TXNL1 ($p = 0.026$) (Supplemental Fig. 6). CAPZA1 did not present any significant expression differences.

Discussion

This study aimed at detecting aneuploidy-associated proteins with therapeutic and/or prognostic value for clinical application in colorectal cancer. Therefore, a total of seven colorectal cell lines were assessed for their DNA ploidy type. Comprehensive protein profiling detected 64 distinct spots that were differentially expressed between diploid and aneuploid colorectal cell lines. Twenty-five distinct proteins could be identified. Pathway analysis showed that most relevant functions associated with these proteins were related to *Cellular Assembly and Organization*, *Cellular Function and Maintenance*, *Infection Mechanism*, *Cell Cycle*, and *Cellular Growth and Proliferation*. Compromised cell cycle checkpoints favor malignant cells for accelerated proliferation thus leading to clonal expansion with uncompleted DNA repair [29].

Subsequent Western-blot analysis could not confirm significant differential expression for three of the six selected proteins: Expression levels of YWHAQ showed low fold changes in 2-DE gels (0.86) and Western blot (0.96). Therefore, expression differences might have been too subtle to reach significance for downstream analysis. While GNAS and PRDX2 presented promising fold changes in 2-DE gels (0.59 and 1.14) and Western blots (0.73 and 1.33), they also presented multiple bands. As the polyclonal antibody against GNAS most likely detected some of the seven isoforms of GNAS, the monoclonal antibody for PRDX2 should not have produced any additional bands. However, Manandhar et al. [30] also described additional bands with a PRDX2 antibody and concluded that its function was compromised. Low specificity of antibodies might therefore be attributable to the difficulties of detecting different isoforms and/or post-translational modifications of proteins in translational research. In addition, identification of 2-DE protein spots may be hampered by additional proteins that underlay the spot of interest. Such phenomena may also impact the success rate of Western-blot validation.

Against this background, we consider it valuable that differential expression by 2-DE could be confirmed by Western blotting for TXNL1, CAPZA1, and HDAC2. While the first two proteins were more highly expressed in the diploid than in the aneuploid group, the expression pattern for HDAC2 was reversed.

The TXNL1 gene encodes a protein that belongs to the thioredoxin family of small redox active proteins which have a conserved active site sequence. It is involved in cellular response to sugar starvation stress and regulates the redox states in higher eukaryotes [31, 32]. It has been reported that TXNL1 overexpression can increase the transcriptional repressor function through its binding to the transcription factor B-Myb [33]. Thus, TXNL1 overexpression specifically predisposes to G0/G1 phase arrest. Overexpression of thioredoxins has previously been shown for different tumor entities as compared to normal tissues [34]. In line with this report, strong immunopositivity (score 2 + 3) was significantly more often observed in our carcinoma specimens (39.2%) as compared to adjacent normal mucosa (10.5%) (Supplemental Figs. 6, 7). While maintaining genomic stability in normal tissue seems to require only low TXNL1 expression levels (89.5%, score 0 + 1), disturbances of the cellular equilibrium in malignancies might require increased expression levels to maintain or re-gain genomic stability. While this is achieved and reflected by strong immunopositivity (score 3) in a substantial number of diploid carcinomas (24%), aneuploid carcinomas seldom reached such high expression levels (6.4%). Overexpression in diploid malignancies could therefore help to maintain genomic stability.

CAPZA1 was also significantly less expressed in the aneuploid group of colorectal cancer cell lines, but did not reach the level of significance by immunohistochemistry validation. CAPZA1 was weak (44.4%) to moderately (55.6%) expressed in the normal tissue (Supplemental Fig. 7). In contrast, strong immunoreactivity could only be observed in malignant tissue (diploids 24.1%, aneuploids 15.2%). This finding supports our previous results where CAPZA1 was also higher expressed in primary colorectal carcinomas and associated metastases than in normal and adenoma tissue [9]. CAPZA1 is a member of the F-actin capping protein alpha subunit family and encodes the alpha subunit of the barbed-end actin binding protein. The protein regulates growth of the actin filament by capping the barbed end of growing actin filaments. While no definite link has been established between CAPZA1 and aneuploidy in cancer so far, it is intriguing to assume that CAPZA1 may also compromise the integrity of actin filaments and actin-binding proteins that are required for reliable chromosome segregation. Disturbed or insufficient CAPZA1 expression levels could thus contribute to genomic instability [35]. The reason why immunohistochemical validation of CAPZA1 expression lacked statistical significance requires additional evaluation.

In contrast to TXNL1 and CAPZA1, HDAC2 was expressed at high levels in the aneuploid colorectal cancer cell lines. TMA-based validation confirmed this finding: while HDAC2 immunoreactivity was absent (42.1%) or weak (52.6%) in normal mucosa, carcinomas showed moderate to strong immunoreactivity (51.3%) (Supplemental Figs. 6, 7). This is in line with the study by Ashktorab et al. [36] who described a correlation of HDAC2 expression with the progression from adenoma to carcinoma. Furthermore, we here present that a more frequent HDAC2 immunopositivity

characterizes aneuploid carcinomas (91.3%) that are known to show an inferior prognosis than diploid carcinomas [11]. HDAC2 positivity was only 70% in diploid carcinomas and 57.9% in normal tissue (Fig. 6b and Supplemental Fig. 7). HDAC2 plays an important role in transcriptional regulation and cell cycle progression [37]. It has been found to be overexpressed in different tumor entities including colon cancer [36, 38]. HDAC2 overexpression has also been associated with reduced survival of patients with oral and/or colorectal cancer [39, 40]. Poor prognosis, however, characterizes aneuploid malignancies overall including colorectal cancers as shown also in this work. Interestingly, HDAC2 overexpression seems to be induced by APC loss and to be sufficient on its own to prevent apoptosis, thus favoring the development of genomic instability and tumor growth [41]. The inhibition of HDAC in combination with chemotherapy (doxorubicin) is currently assessed in phase I clinical trials. Hereby, HDAC2 overexpression in the primary tumor serves as predictive marker for efficient HDAC inhibition [42]. Our data define a close correlation of HDAC2 overexpression, aneuploidy, and poor prognosis. It seems therefore reasonable to suggest that HDAC2 overexpression is a mere reflection of aneuploidy and that patients with aneuploid tumors in particular could benefit from HDACi treatment. In this respect, it seems worthy to acknowledge that about 70% of colorectal cancers are aneuploid. Furthermore, the proteins that were identified to be overexpressed and highly correlated with aneuploid carcinomas (e.g., HSPA8, HSPA9, and PDIA3) might serve as promising targets for chemotherapeutic intervention and are suggested for further analyses. In addition, IPA network genes need to be subjected to validation and functional evaluation studies.

In summary, HDAC2 and TXNL1 could be validated as aneuploidy-associated proteins in 78 primary colorectal cancers by immunohistochemistry on tissue microarrays. Further studies for their prognostic value are warranted.

References

1. Duval A, Hamelin R (2002) Mutations at coding repeat sequences in mismatch repair-deficient human cancers: toward a new concept of target genes for instability. Cancer Res 62:2447–2454

2. Bottger TC, Potratz D, Stockle M, Wellek S, Klupp J, Junginger T (1993) Prognostic value of DNA analysis in colorectal carcinoma. Cancer 72:3579–3587

3. Yamamoto T, Matsumoto K, Iriyama K (1998) Prognostic significance of the DNA index in a colorectal cancer. Surg Today 28:792–796

4. Araujo SE, Bernardo WM, Habr-Gama A, Kiss DR, Cecconello I (2007) DNA ploidy status and prognosis in colorectal cancer: a meta-analysis of published data. Dis Colon Rectum 50:1800–1810

5. Habermann JK, Upender MB, Roblick UJ, Kruger S, Freitag S, Blegen H, Bruch HP, Schimmelpenning H, Auer G, Ried T (2003) Pronounced chromosomal instability and multiple gene amplifications characterize ulcerative colitis-associated colorectal carcinomas. Cancer Genet Cytogenet 147:9–17

6. Ried T, Knutzen R, Steinbeck R, Blegen H, Schrock E, Heselmeyer K, du Manoir S, Auer G (1996) Comparative genomic hybridization reveals a specific pattern of chromosomal gains and losses during the genesis of colorectal tumors. Genes Chromosomes Cancer 15:234–245

7. Bomme L, Bardi G, Pandis N, Fenger C, Kronborg O, Heim S (1994) Clonal karyotypic abnormalities in colorectal adenomas: clues to the early genetic events in the adenoma–carcinoma sequence. Genes Chromosomes Cancer 10:190–196

8. Habermann JK, Paulsen U, Roblick UJ, Upender MB, McShane LM, Korn EL, Wangsa D, Kruger S, Duchrow M, Bruch HP, Auer G, Ried T (2007) Stage-specific alterations of the genome, transcriptome, and proteome during colorectal carcinogenesis. Genes Chromosomes Cancer 46:10–26

9. Roblick UJ, Hirschberg D, Habermann JK, Palmberg C, Becker S, Kruger S, Gustafsson M, Bruch HP, Franzen B, Ried T, Bergmann T, Auer G, Jornvall H (2004) Sequential proteome alterations during genesis and progression of colon cancer. Cell Mol Life Sci 61:1246–1255

10. Walther A, Houlston R, Tomlinson I (2008) Association between chromosomal instability and prognosis in colorectal cancer: a meta-analysis. Gut 57:941–950

11. Gerling M, Fuchs K, Igl BW, Fritzsche B, Ziegler A, Bader F, Kujath P, Schimmelpenning H, Bruch H-P, Roblick UJ, Habermann JK (2010) High frequency of aneuploidy defines ulcerative colitis-associated carcinomas—a comparative prognostic study to sporadic colorectal carcinomas. Ann Surg 2010 Jun 4. [Epub ahead of print]

12. Half EE, Bresalier RS (2004) Clinical management of hereditary colorectal cancer syndromes. Curr Opin Gastroenterol 20:32–42

13. Auer GU, Caspersson TO, Wallgren AS (1980) DNA content and survival in mammary carcinoma. Anal Quant Cytol 2:161–165

14. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. Anal Biochem 72:248–254

15. Franzen B, Linder S, Okuzawa K, Kato H, Auer G (1993) Nonenzymatic extraction of cells from clinical tumor material for analysis of gene expression by two-dimensional polyacrylamide gel electrophoresis. Electrophoresis 14:1045–1053

16. Carrette O, Burkhard PR, Sanchez JC, Hochstrasser DF (2006) State-of-the-art two-dimensional gel electrophoresis: a key tool of proteomics research. Nat Protoc 1:812–823

17. Garrels JI (1989) The QUEST system for quantitative analysis of two-dimensional gels. J Biol Chem 264:5269–5282

18. Hellman U (2000) Sample preparation by SDS/PAGE and in-gel digestion. EXS 88:43–54

19. Calvano SE, Xiao W, Richards DR, Felciano RM, Baker HV, Cho RJ, Chen RO, Brownstein BH, Cobb JP, Tschoeke SK, Miller-Graziano C, Moldawer LL, Mindrinos MN, Davis RW, Tompkins RG, Lowry SF (2005) A network-based analysis of systemic inflammation in humans. Nature 437:1032–1037

20. Ficenec D, Osborne M, Pradines J, Richards D, Felciano R, Cho RJ, Chen RO, Liefeld T, Owen J, Ruttenberg A, Reich C, Horvath J, Clark T (2003) Computational knowledge integration in bio-pharmaceutical research. Brief Bioinform 4:260–278

21. Braunschweig T, Chung JY, Hewitt SM (2005) Tissue micro-arrays: bridging the gap between research and the clinic. Expert Rev Proteomics 2:325–336

22. Kononen J, Bubendorf L, Kallioniemi A, Barlund M, Schraml P, Leighton S, Torhorst J, Mihatsch MJ, Sauter G, Kallioniemi OP (1998) Tissue microarrays for high-throughput molecular profil-ing of tumor specimens. Nat Med 4:844–847

23. Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1:307–310

24. Breiman L (2001) Random forests. Machine Learning 45:5–32

25. Diaz-Uriarte R, Alvarez de Andres S (2006) Gene selection and classification of microarray data using random forest. BMC Bioinform 7:3

26. Fleiss JL (1971) Measuring nominal scale agreement among many raters. Psychol Bull 76:378–382

27. Melcher R, Koehler S, Steinlein C, Schmid M, Mueller CR, Luehrs H, Menzel T, Scheppach W, Moerk H, Scheurlen M, Koehrle J, Al-Taie O (2002) Spectral karyotype analysis of colon cancer cell lines of the tumor suppressor and mutator pathway. Cytogenet Genome Res 98:22–28

28. Tsushimi T, Noshima S, Oga A, Esato K, Sasaki K (2001) DNA amplification and chromosomal translocations are accompanied by chromosomal instability: analysis of seven human colon cancer cell lines by comparative genomic hybridization and spectral karyotyping. Cancer Genet Cytogenet 126:34–38

29. Jefford CE, Irminger-Finger I (2006) Mechanisms of chromo-some instability in cancers. Crit Rev Oncol Hematol 59:1–14

30. Manandhar G, Miranda-Vizuete A, Pedrajas JR, Krause WJ, Zimmerman S, Sutovsky M, Sutovsky P (2009) Peroxiredoxin 2 and peroxidase enzymatic activity of mammalian spermatozoa. Biol Reprod 80:1168–1177

31. Jimenez A, Pelto-Huikko M, Gustafsson JA, Miranda-Vizuete A (2006) Characterization of human thioredoxin-like-1: potential involvement in the cellular response against glucose deprivation. FEBS Lett 580:960–967

32. Lee KK, Murakawa M, Takahashi S, Tsubuki S, Kawashima S, Sakamaki K, Yonehara S (1998) Purification, molecular cloning, and characterization of TRP32, a novel thioredoxin-related mammalian protein of 32 kDa. J Biol Chem 273:19160–19166

33. Kim KY, Lee JW, Park MS, Jung MH, Jeon GA, Nam MJ (2006) Expression of a thioredoxin-related protein-1 is induced by prostaglandin E(2). Int J Cancer 118:1670–1679

34. Powis G, Mustacich D, Coon A (2000) The role of the redox protein thioredoxin in cell growth and cancer. Free Radic Biol Med 29:312–322

35. Gerisch G, Faix J, Kohler J, Muller-Taubenberger A (2004) Actin-binding proteins required for reliable chromosome segre-gation in mitosis. Cell Motil Cytoskeleton 57:18–25

36. Ashktorab H, Belgrave K, Hosseinkhah F, Brim H, Nouraie M, Takkikto M, Hewitt S, Lee EL, Dashwood RH, Smoot D (2009) Global histone H4 acetylation and HDAC2 expression in colon adenoma and carcinoma. Dig Dis Sci 54:2109–2117

37. Harms KL, Chen X (2007) Histone deacetylase 2 modulates p53 transcriptional activities through regulation of p53-DNA binding activity. Cancer Res 67:3145–3152

38. Song J, Noh JH, Lee JH, Eun JW, Ahn YM, Kim SY, Lee SH, Park WS, Yoo NJ, Lee JY, Nam SW (2005) Increased expression of histone deacetylase 2 is found in human gastric cancer. APMIS 113:264–268

39. Chang HH, Chiang CP, Hung HC, Lin CY, Deng YT, Kuo MY (2008) Histone deacetylase 2 expression predicts poorer prog-nosis in oral cancer patients. Oral Oncol

40. Weichert W, Roske A, Niesporek S, Noske A, Buckendahl AC, Dietel M, Gekeler V, Boehm M, Beckers T, Denkert C (2008) Class I histone deacetylase expression has independent prognostic impact in human colorectal cancer: specific role of class I histone deacetylases in vitro and in vivo. Clin Cancer Res 14:1669–1677

41. Zhu P, Martin E, Mengwasser J, Schlag P, Janssen KP, Gottlicher M (2004) Induction of HDAC2 expression upon loss of APC in colorectal tumorigenesis. Cancer Cell 5:455–463

42. Munster PN, Marchion D, Thomas S, Egorin M, Minton S, Springett G, Lee JH, Simon G, Chiappori A, Sullivan D, Daud A (2009) Phase I trial of vorinostat and doxorubicin in solid tumours: histone deacetylase 2 expression as a predictive marker. Br J Cancer 101:1044–1050