Regulator of Chromosome Condensation 2 Identifies High-Risk Patients within Both Major Phenotypes of Colorectal Cancer

Jarle Bruun, Matthias Kolberg, Terje C. Ahlquist, Ellen C. Royrvik, Torfinn Nome, Edward Leithe, Guro E. Lind, Marianne A. Merok, Torleiv O. Rognum, Geir Bjorkoy, Terje Johansen, Annika Lindblom, Xiao-Feng Sun, Aud Svindland, Knut Liestol, Arild Nesbakken, Rolf I. Skotheim and Ragnhild A. Lothe

Linköping University Post Print

Tweet

N.B.: When citing this work, cite the original article.

Original Publication:

Jarle Bruun, Matthias Kolberg, Terje C. Ahlquist, Ellen C. Royrvik, Torfinn Nome, Edward Leithe, Guro E. Lind, Marianne A. Merok, Torleiv O. Rognum, Geir Bjorkoy, Terje Johansen, Annika Lindblom, Xiao-Feng Sun, Aud Svindland, Knut Liestol, Arild Nesbakken, Rolf I. Skotheim and Ragnhild A. Lothe, Regulator of Chromosome Condensation 2 Identifies High-Risk Patients within Both Major Phenotypes of Colorectal Cancer, 2015, Clinical Cancer Research, (21), 16, 3759-3770.
http://dx.doi.org/10.1158/1078-0432.CCR-14-3294
Copyright: American Association for Cancer Research
http://www.aacr.org/

Postprint available at: Linköping University Electronic Press
http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-122121

Regulator of chromosome condensation 2 identifies high-risk patients within both major phenotypes of colorectal cancer

Jarle Bruun[1,2,3], Matthias Kolberg[1,2,3], Terje Ahlquist[1,2], Ellen C. Røyrvik[1,4,5], Torfinn Nome[1,3], Edward Leithe[1,2,3], Guro E. Lind[1,2,3], Marianne A. Merok[1,3,6], Torleiv O. Rognum[7,8], Geir Bjørkøy[9], Terje Johansen[10], Annika Lindblom[11], Xiao-Feng Sun[12], Aud Svindland[7,13], Knut Liestøl[3,14], Arild Nesbakken[2,3,6,7], Rolf I. Skotheim[1,2,3,14], and Ragnhild A. Lothe[1,2,3,4*]

[1]Department for Cancer Prevention, Institute for Cancer Research, and [2]K.G.Jebsen Colorectal Cancer Research Centre - Oslo University Hospital, Oslo, Norway. [3]Centre for Cancer Biomedicine, Faculty of Medicine, University of Oslo, Oslo, Norway. [4]Department of Molecular Biosciences and [14]Department of Informatics, Faculty of Mathematics and Natural Sciences, University of Oslo, Oslo, Norway. [5]Department of Oncology, University of Oxford, ORCRB, Headington, Oxford, UK. [6]Department of Gastrointestinal Surgery, Aker Hospital - Oslo University Hospital, Oslo, Norway. [7]Faculty of Medicine, University of Oslo, Norway. [8]Division of Forensic Medicine, Department of Forensic Pathology and Clinical Forensic Medicine, the Norwegian Institute of Public Health, Oslo, Norway. [9]University College of Sør-Trøndelag, Trondheim, Norway. [10]Molecular Cancer Research Group, Institute of Medical Biology, University of Tromsø, Tromsø, Norway. [11]Department of Molecular Medicine and Surgery, Karolinska Institute, Stockholm, Sweden. [12]Division of Oncology, Department of Clinical and Experimental Medicine, Faculty of Health Sciences, County Council of Östergötland, University of Linköping, Linköping, Sweden. [13]Department of Pathology, Oslo University Hospital, Oslo, Norway.

***Corresponding author:**
Ragnhild A. Lothe
Street address: The Norwegian Radium Hospital, Institute for Cancer Research, Montebello, Ullernchausséen 70, NO-0379 Oslo, Norway
Phone (office): +47 2278 1728
Fax (office): +47 2278 1745
E-mail: Ragnhild.A.Lothe@rr-research.no

TRANSLATIONAL RELEVANCE

Colorectal cancer has a high incidence and mortality rate. Prognosis and postoperative treatments are still based on clinicopathological parameters alone. Patients with the MSI phenotype tumors generally have a good prognosis. Here we show that a mutation in the regulator of chromosome condensation 2, *RCC2*, gene can identify high-risk patients within the stage II MSI group. Furthermore, reduced RCC2 protein expression identifies female patients with poor outcome within the MSS group. Subgroup analyses suggest that RCC2 is able to stratify clinically important patient groups: patients above 75 years of age, patients with stage II and those with rectal cancer. Hence, RCC2 risk stratification can potentially guide clinical decision making for a large number of CRC patients. Importantly, both the mutation assay and the protein expression assessment can be done rapidly by cost-effective routine technologies.

ABSTRACT

Purpose: Colorectal cancer (CRC) has high incidence and mortality worldwide. Patients with microsatellite instable (MSI) tumors have significantly better prognosis than patients with microsatellite stable (MSS) tumors. Considerable variation in disease outcome remains a challenge within each subgroup, and our purpose was to identify biomarkers that improve prediction of CRC prognosis.

Experimental design: Mutation analyses of 42 MSI target genes were performed in two independent MSI tumor series (*n*=209). Markers that were significantly associated with prognosis in the Test series were assessed in the Validation series, followed by functional and genetic explorations. The clinical potential was further investigated by immunohistochemistry in a population-based CRC series (*n*=903).

Results: We identified the cell cycle gene Regulator of Chromosome Condensation 2 (*RCC2*) as a cancer biomarker. We found a mutation in the 5'UTR region of *RCC2* that in univariate and multivariate analyses was significantly associated with improved outcome in the MSI group. This mutation caused reduction of protein expression in dual luciferase gene reporter assays. SiRNA knockdown in MSI colon cancer cells (HCT15) caused reduced cell proliferation, cell cycle arrest and increased apoptosis. Massive parallel sequencing revealed few *RCC2* mutations in MSS tumors. However, weak RCC2 protein expression was significantly associated with poor prognosis, independently of clinical high-risk parameters, and stratifies clinically important patient subgroups with MSS tumors, including old patients (>75 years), stage II patients and those with rectal cancer.

Conclusions: Impaired *RCC2* affects functional and clinical endpoints of CRC. High-risk patients with either MSI or MSS tumors can be identified with cost-effective routine *RCC2* assays.

INTRODUCTION

Each year, approximately 1.4 million patients are diagnosed with colorectal cancer (CRC), and only half of them survive five years[1]. About 15% of the tumors have the microsatellite instability phenotype (MSI), which is caused by a malfunctioning DNA mismatch repair (MMR) system leading to accumulation of insertions and deletions (indels) in repetitive elements throughout the genome[2]. Patients with non-hereditary MSI tumors have better prognosis than those with microsatellite stable (MSS) tumors[3-5], and MSI is currently implemented in clinical guidelines as a prognostic biomarker[6] (www.nccn.org). However, considerable variation in disease outcome remains a challenge also within this subgroup. For the other 85% of CRC, staging remains the main crude prognostic factor, which underlines the need for robust biomarkers that can stratify these patient groups and improve treatment strategies.

A number of target genes with indels in microsatellites have been reported in MMR-deficient CRC[7], and many studies have evaluated their prognostic impact, but typically only few genes have been investigated at a time. Their ability to predict disease outcome remains inconclusive.

The cell cycle gene Regulator of chromosome condensation 2 (*RCC2*) was identified as a potential target gene in MMR-deficient CRC by a systematic database search in 2002[8]. *RCC2* is a member of the *RCC1* superfamily and is predicted to form a seven-bladed beta-propeller protein essential for the function of the chromosomal passenger complex (CPC)[9]. RCC2 has also been implicated as a negative dual regulator of RAC1 and ARF6, promoting directional cell migration[10]. However, reports on the relation between *RCC2* and cancer have been scarce, and to our knowledge, no direct links have been reported.

Here, we investigated the mutation status for 42 reported target genes of MMR-deficient CRC in two independent series to test the hypothesis that such mutations in cancer-critical genes provide a selective growth advantage and confer prognostic information. The functional, genetic and clinical relevance of the single gene, *RCC2*, that demonstrated prognostic relevance in the MSI group was further studied in cell line models as well as in a population-based consecutive series of CRC (MSI and MSS groups)

MATERIALS AND METHODS

Patient samples

DNA was extracted for mutation analyses from 81 MSI tumors (referred to as the Scandinavian series) from CRC patients admitted to Swedish (n=44)[11] and Norwegian (n=37)[3] hospitals. Of these, clinical data and long term patient follow-up data were available for the 37 Norwegian samples (referred to as the Test series, Table S1).

Validation and further experiments were based on a population-based consecutive series consisting of 1290 CRC patients admitted to Oslo University Hospital (OUS) – Aker, Norway in 1993-2003[12] (see Fig. S1). According to national health policies, all patients with a newly diagnosed colorectal cancer and living within defined catchment areas should be referred to the local hospital for treatment. OUS-Aker served a population of more than 200,000 inhabitants in five districts within or neighbouring the capital of Norway, city of Oslo, and close to 100% of the CRC patients who were treated in this hospital were included in the study. Of the 1290 patients, 929 patients underwent major resection, and the formalin-fixed paraffin-embedded (FFPE) tissue was collected in a tissue microarray (TMA) including 670 colonic, 233 rectal and 26 synchronous carcinomas. From the same series, DNA was extracted for mutation analyses from 128 MSI tumors (referred to as the Validation series, Table S2). More than 95% of patients were of Caucasian ethnicity (based on name-origin). All patient samples were analyzed retrospectively.

The study was approved by the Regional Committee for Medical and Health Research Ethics, South-Eastern Norway and the Norwegian Data Inspectorate. The research conformed to the Declaration of Helsinki and the research biobanks have been registered according to national legislation. The approved project and amendments (REK number 1.2005.1629) require that informed consent is obtained from patients being enrolled to the study.

Cell lines

HCT15, HT29 and HeLa CCl-2 cell lines were purchased from the American Type Culture Collection (ATCC, VA, USA) and maintained according to the ATCC recommendations. Identities were verified by fingerprinting according to the AmpFℓSTR Identifiler PCR Amplification Kit (Life Technologies by Thermo Fisher Scientific, MA, USA), and matched to the profiles reported by ATCC. Cell lines were regularly tested for mycoplasma contamination according to the MycoAlert Mycoplasma Detection Assay (Lonza Cologne AG, Germany).

Determination of MSI

MSI status was determined using the consensus markers provided by the National Cancer Institute as previously described[12].

Mutation analyses of gene targets downstream of MMR deficiency

Fragment analyses of the microsatellite containing regions of 42 genes were performed (details in Supplementary data Table S4). Electropherograms were visually examined for indels by two researchers independently (TA and ECR), against corresponding fragments from DNA from four different disease-free individuals (Fig. S6). Assays were duplicated in tandem runs using different PCR machines. Each gene was informative for at least 159 of the 209 samples.

Mutation verification by sequencing

In order to validate the fragment analysis results, 18 of the 42 genes (*ACVR2A, AIM2, ASTE1, AXIN2, BLM, EPHB2, GRK4, MBD4, PTHLH, RAD50, RBBP8, RCC2, SEMG1, SLC23A2, SYCP1, TAF1B, WISP3*, and *ZMYND8*) were sequenced in a random selection of 11 CRC tissue samples and 8 CRC MSI cell lines (details in Supplementary data).

Massive parallel sequencing

Whole-transcriptome paired-end sequencing was performed as previously described for seven cell lines (HCT15, HCT116, HT29, SW480, SW48, LS1034 and RKO)[13]. Whole-exome paired-end sequencing was carried out using the Illumina TruSeq DNA Sample Prep Kit, followed by the Illumina TruSeq Exome Enrichment Kit (Illumina Inc., CA, USA), and sequencing was subsequently performed on an Illumina Genome Analyzer IIx for nine samples (tumor and normal tissue). Whole-genome paired-end sequencing was performed by BGI (Hong Kong, China) for the four CRC cell lines: HCT15, HCT116, HT29 and SW480. The sequence reads were aligned by BWA version 0.6.1 (Li & Durbin, 2009). Mutation data from all sequenced tumors by The Cancer Genome Atlas (TCGA) was accessed via the cBioPortal (http://www.cbioportal.org/public-portal/)[14] on December 16, 2014.

Reporter gene assays

A dual luciferase reporter system (Promega, CA, USA) was used to determine reporter activity from vector constructs with mutated and wild-type *RCC2* (Fig. S3) in three cell lines: HCT15, HT29 and HeLa (details in Supplementary data).

Primary antibodies

For Western blotting, rabbit polyclonal anti-RCC2 antibodies were obtained from Novus Biologicals (Cat. No. NB110-40618, Cambridge, UK), recognizing an N-terminal epitope between residue 1 and residue 50. Mouse monoclonal anti-β-actin antibodies were

obtained from Sigma (Clone AC-47, Cat. No. A 5316, MO, USA), recognizing an N-terminal epitope of the β-isoform. Antibodies were employed at dilutions of 1:500 and 1:1000, respectively. For immunocytochemistry and immunohistochemistry, rabbit polyclonal anti-RCC2 antibodies were obtained from Novus Biologicals (Cat. No. NB110-40619, Cambridge, UK), recognizing a C-terminal epitope between residue 471 and residue 522, and employed at a 1:500 dilution. Alexa Fluor 594 phalloidin (Invitrogen, cat. No. A12381) was used to visualize actin for immunocytochemistry according to the manufacturer's instructions at a 1:40 dilution.

siRNA knockdown, Western blotting and immunocytochemistry

Two siRNA oligonucleotides targeted against *RCC2* were used to deplete endogenous *RCC2* protein in HCT15 cells: Invitrogen Stealth Select RNAi RCC2HSS125252 (I) and RCC2HSS125253 (II). The first *RCC2* siRNA (I) had the following sequence (sense strand): 5' - CCU GGU GAA GCU GUU UGA CUU CCC U - 3'. The sequence (sense strand) of the second RCC2 siRNA (II) was: 5' - GGG CUU CCC AGA UCU AUG CUG GUU A - 3'. Invitrogen Stealth RNAi Negative Control (Medium GC) served as negative control. The transfection reagent used was Lipofectamine 2000 (Invitrogen by Life Technologies, Paisley, UK) and the transfection procedure was carried out following the manufacturer's instructions, at a final concentration of 80 nM.

Western blotting and immunocytochemistry were employed to confirm knockdown of RCC2 protein expression and to visualize protein expression and cell morphology, respectively. Detailed protocols are reported elsewhere[15].

Flow cytometry

Flow cytometry analysis of cell cycle distributions and levels of apoptosis was carried out using a TUNEL assay where treatment with terminal transferase (TdT) followed by staining with Streptavidin-PE and Hoechst 33258 was employed. A Becton Dickinson LSR II flow cytometer was used for the analysis. Data analysis was carried out with BD FACSDiva Software version 5.0.3 and ModFit LT Version 3.2.1 (Verity Software House). Detailed protocol is reported elsewhere[15].

Immunohistochemistry

The immunohistochemical (IHC) analysis was performed on FFPE tissue as described[16] (details in Supplementary data).

Statistics

All statistics were performed using the SPSS 21.0 software (SPSS, IL, USA). Five-year overall survival (OS) and five-year time-to-recurrence (TTR) plots were generated using the Kaplan-Meier method. The logrank test was used to compare survival curves. TTR and OS were defined according to the guidelines given by Punt *et al.*[17] where time-to-

recurrence is defined as the time from surgery to the first event of either death from the same cancer, local recurrence or distant metastasis. Patients were censored at death from other cancer, non-cancer death and post-operative death (<3 months). Overall survival was defined as the time from surgery to death from any cause. No patients were lost to follow-up in the study period.

Multivariate Cox proportional hazards regression modeling (forward selection) was employed to determine all available parameters with significant independent impact on patient survival. Age and gender was included as bacground variables. Tumor location was also included in the protein expression analyses for informative purposes, and it did not affect the the multivariate models. Cases with missing data were excluded from analyses. Adjuvant treatment for patients with stage III colon cancer (<75 years of age) became standard treatment in 1997 and was considered in initial multivariate models. These patients were few and adjustment did not affect the models. Adjustment for pre- and post-operative radiotherapy for rectal cancer patients was also considered, but was pertinent to only a very limited number of patients and therefore not included in initial models. Tumor location was excluded as a covariate in the multivariate mutation analyses of patients with MSI tumors since there are very few cases in the distant colon and the rectum category as compared to the proximal category. Also, for the covariate Histopathological grade, the G1 and the G2 category were combined to make larger and more robust groups for comparison. We believe this is reasonable since G1 and G2 are similar with regard to prognosis, and there is in addition considerable inter-pathologist variation in this assessment. Interaction tests were integrated in the Cox models to assess whether effects were different between subgroups, but must be interpreted carefully due to the low power of such tests. The proportional hazards assumptions were verified by graphical evaluation of plots of log(-log survival time) *vs* log time. Bonferroni correction was used to correct for multiple testing in the mutation analyses for the Validation series.

For the analysis of *RCC2* mutation and protein expression in relation to other clinicopathological data, Wilcoxon rank-sum test (exact) was used to compare two groups with regard to an ordinal variable and Fisher's exact test was used to compare categorical variables. Student's *t*-test for paired samples was used to estimate *P*-values for comparisons in the dual luciferase experiments, siRNA knockdown experiments, cell proliferation experiments and the flow cytometry experiments. For the dual luciferase experiments, an analysis of variance was also carried out of the relative expression for the two vectors. Here, the cell lines were fixed factors and the experiments nested within each cell line as random factor.

All *P*-values were two-tailed, and considered statistically significant at $P \leq 0.05$ unless otherwise is stated in the text.

RESULTS

Mutation status of multiple cancer-critical target genes in CRCs with MSI

The mutation frequencies of 42 known target genes in MMR deficient CRC were assessed in two independent series, the Scandinavian series (*n*=81) and the Validation series (*n*=122) (Fig.1). Thirty-seven genes were mutated in more than 12% of the samples in both series (Fig. 2A). The mutation frequency for each gene varied from 6% to 91% in the Scandinavian series and from 4% to 92% in the Validation series, with a median of 17 (range 0-28) and 19 (range 0-29) mutated genes per sample, respectively.

Mutations in *RCC2* were located in the mononucleotide $(A)_{10}$-repeat within the 5'UTR and were identified in 62% of all MSI tumors, 51% (41/81) and 69% (84/122) in the two series, respectively. Deletion of one or two bases (103/125 and 19/125, respectively) were far more common than any other alterations in both series.

The *RCC2* 5'UTR mutation is an independent marker for improved outcome in patients with MSI

The mutation status of all 42 MSI target genes was analyzed for association with patient outcome in the Test series (*n*=37; 37 Norwegian patients from the 81 Scandinavian series) in order to identify the best candidate markers, which were then analyzed in the larger Validation series (*n*=122). From the six candidates identified by univariate five-year TTR analysis (*ACVR2A*, *AXIN2*, *EP300*, *MRE11A*, *OGT* and *RCC2*, Table S5), only the *RCC2* mutation was found to be significantly associated with outcome in both clinical series (Fig. 1B-C). Stratification by tumor stage in the Validation series showed that for stage II patients after five years follow-up, only 12% of the patients with mutated *RCC2* had experienced recurrence as compared with 39% for patients with wild-type *RCC2* (Fig 1E). The corresponding figures for stage III MSI patients where 44% and 55%, respectively. There were too few cases and/or events for robust analysis of stage I and IV patients with MSI.

Cox proportional hazards models were created (Table 1), and the *RCC2* mutation status demonstrated significant prognostic impact, independently of clinical high-risk parameters (*P*=0.0080). This result is significant according to standard Bonferroni correction for multiple testing (0.05/6 = 0.0083). Cox models for the Test series show a similar association (*P*=0.054), although this series is too small to give robust estimates for the various covariates (Table S6).

An additional more stringent analysis was performed *post hoc* in the Validation series, restricted to patients with right-sided tumors who were alive three months after a complete resection (R0, no evidence of residual tumor), and who were followed for at

least 36 months (n=60). The association between *RCC2* mutation status and TTR remained significant in univariate (P=0.038, Fig. 1D) and multivariate analysis (P=0.043).

For the Test series and Validation series, Wilcoxon and Fisher's exact tests showed that mutations in *RCC2* were not significantly associated (P<0.05) with other clinicopathological data, except surgical resection status (R-status, P=0.039) (Table S1-3).

Mutations in the complete gene sequence of *RCC2*

To search for mutations elsewhere in the *RCC2* gene, we used in-lab data sets from massive parallel sequencing of selected CRC samples and cell lines.

Three data sets were examined for *RCC2* mutations in 16 samples: whole-genome paired-end sequencing data from the four colon cancer cell lines HCT15, HCT116, HT29 and SW480 (performed at BGI Hong Kong, China), paired-end RNA sequencing data sets from the same four, as well as the three additional colon cancer cell lines SW48, LS1034 and RKO[13], and exome sequencing data of 9 CRCs (one MSI, 7 MSS and one unknown, unpublished data). Altogether, three deletions (one in the 5'UTR $(A)_{10}$ mononucleotide repeats (MNR) and two in the 3'UTR), one missense and one synonymous mutation were found (Fig.1F). Both the missense and the synonymous mutations were in the RKO cell line (MSI).

In addition, we used publically available data: The Cancer Genome Atlas (TCGA) accessed through the cBioPortal[14]. The provisional TCGA data set comprises somatic mutation data from exome sequencing of 223 CRCs. Here, *RCC2* was mutated in 6 cases (2.7%), of which 5 were missense mutations (4 in MSI and one in MSS) and one was a frameshift deletion (in an MSS tumor) (Fig. 1F). This frameshift mutation in codon 189 leads to a premature stop codon and truncation of the polypeptide sequence. The 5'UTR and the 3'UTR regions were not covered by the sequence capture protocol used by TCGA.

The 5'UTR *RCC2* mutation affects mRNA structure

Secondary structure models of a 184 nucleotide long RNA sequence comprising 122 nucleotides of the 5'UTR and 62 nucleotides of the coding sequence of wild-type and mutant *RCC2* were constructed using MFold (Fig. S2). The models revealed that a three-way RNA junction may be affected by the deletion of an A in the $(A)_{10}$-repeat. Three-way RNA junctions are important RNA folds which are often involved in distant tertiary contacts, protein binding and/or RNA binding[18]. In the structural model, the wild-type sequence has a single unpaired nucleotide spanning the junction while the mutant has two unpaired nucleotides. This extra nucleotide may affect the flexibility and projection

of the helices from this junction having adverse functional effects impinging on translation of the mRNA because either protein- or distant RNA contacts can be disturbed.

The 5'UTR *RCC2* mutation causes reduced protein expression

A dual luciferase gene reporter assay was carried out in three different cell lines (HCT15, HT29 and HeLa) to investigate the functional significance of the *RCC2* 5'UTR $(A)_9$ mutation as compared to the wild-type $(A)_{10}$. A wild-type or mutant 5'UTR *RCC2* fragment was inserted immediately upstream of the luciferase reporter gene in an SV40 promoter-based plasmid (Fig. S3). Luciferase activity was determined following transient transfection. The mutant construct was found to exhibit significantly reduced luciferase activity across the cell lines compared to the wild-type (P=0.0057, n=17, two-tailed paired Student's t-test, Fig. 2). Statistical significance was also evident from an analysis of variance of the relative expression for the two vectors, with the cell lines as fixed factors and the experiments nested within each cell line as random factor (P=0.024, n=17). A median 15% reduction of protein expression was observed.

Quantitative real-time PCR analysis (qPCR) of reporter transcript levels in the three cell lines showed no significant difference between wild type and mutant, indicating that the observed reduction in mutant reporter activity occurs at the translational level (P=0.62, n=15 , two-tailed paired Student's t-test).

RCC2 knockdown causes reduced cell growth and altered cell morphology

A colon cancer cell line with MSI, HCT15, was used to investigate the importance of *RCC2* in regulation of cell growth and proliferation. Two different siRNA constructs were employed to knock down *RCC2*, both of which reduced the RCC2 protein level by more than 90% compared to cells transfected with negative control siRNA sequences (mock) (Fig. 3). Depletion of RCC2 was associated with a significant 40-45% reduction in cell numbers (Fig. 3). Concurrently, *RCC2*-depleted cells exhibited a change in morphology, appearing larger and more distinctly epithelial-like (Fig. 3), indicating increased cell-spreading activity. Supporting evidence was provided by confocal microscopy analysis revealing that *RCC2*-depleted cells exhibited an altered, more elongated actin fiber pattern with predominant stress- and filopodia-like fibers (Fig. 3 and Fig. S4).

RCC2 knockdown causes G2-M arrest and increased apoptosis

Next, the effect of knockdown of *RCC2* in HCT15 cells on cell cycle progression and apoptosis was determined by flow cytometry. Depletion of *RCC2* by siRNA resulted in a 45% decrease of the G0-G1 population (P=0.012, n=3), a 62% increase of the G2-M population (P=0.011, n=3) and a 17% increase of the S population (P=0.016, n=3), as compared to mock transfected cells (Fig. 3). In addition, the level of aneuploid (>4N) cells was elevated for *RCC2* depleted cells (P=0.028, n=3). Interestingly, the level of

apoptosis in *RCC2* depleted cells was about three times higher than in control transfected cells (0.6% *versus* 1.7%, *P*=0.0077, *n*=3), as detected by a TUNEL-assay (Fig. 3).

Weak RCC2 protein expression *in situ* is associated with poor outcome in patients with MSS tumors

We investigated the prognostic value of the *in situ* protein expression of RCC2 by IHC in a consecutive single-hospital series on a tissue microarray of primary colorectal cancers (*n*=903; 781 and 797 evaluable tissue cores for cytosolic and nuclear staining respectively Fig. S1), including the two main subgroups of CRC, MSI (*n*=128, validation series) and MSS (*n*=714; 619 and 630 evaluable tissue cores for cytosolic and nuclear staining, respectively) tumors.

For cytosolic staining, both OS and TTR analyses showed significant associations with patient outcome ($P=5.2\text{x}10^{-7}$ and $P=8.3\text{x}10^{-5}$, respectively, Figure 4). The TTR analysis was restricted to R0 patients. Cox proportional hazards regression modeling including relevant background and clinical variables (Table 2) showed significant interaction between RCC2 cytosolic staining and gender, both for OS (*P*=0.0059 for the interaction test and *P*=0.0015 for RCC2) and for the TTR analyses (*P*=0.0037 for the interaction test and $P=1.4\text{x}10^{-5}$ for RCC2). Subsequent subgroup analysis (TTR) for gender revealed significant associations confined to female patients ($P=6.2\text{x}10^{-6}$, Fig. 4). The strong prognostic value of RCC2 cytosolic staining in females was confirmed in separate multivariate models for gender ($P=1.5\text{x}10^{-5}$, HR 0.37 , CI 0.24-0.58 for females and *P*=0.65, HR 0.90 , CI 0.56-1.43 for men). Interestingly, further stratification on MSS/MSI status showed that this relationship was confined to patients with MSS (Fig. 4). MSI tumors had stronger cytoplasmic expression compared to MSS tumors (*P*=0.012, Fisher's exact test, *n*=712), however, the number of patients with MSI and events were not sufficient to test for interaction with RCC2 cytosolic staining, nor carry out further subgroup analyses with regard to survival for the MSI group.

Interestingly, Wilcoxon rank-sum test (exact) revealed a significant association between RCC2 cytosolic staining and tumor stage ($P=2.1\text{x}10^{-4}$, *n*=777). Weak cytosolic staining was more predominant among stage III and IV tumors than among stage I and II tumors.

In multivariate Cox proportional hazards regression modeling (TTR, including adjustment for gender, age, MSI/MSS-status, and tumor location, differentiation and stage) nuclear expression of RCC2 showed significant interaction with gender (*P*=0.016, HR 3.0, CI 1.2-7.3 for the interaction test and *P*=0.001, HR 2.9, CI 1.6-5.3 for RCC2) (Fig. 4). Nuclear staining stratified by gender and MSS/MSI-status showed significant associations for female patients (*P*=0.045) and to MSS patients (*P*=0.013, Fig. 4).

Further exploratory subgroup analyses show interesting associations between reduced RCC2 expression and old age, rectal cancer and early stage (I/II) disease (Fig S5).

Wilcoxon rank-sum test for the MSI cancers did not reveal any significant differences in the level of RCC2 staining between tumors with *RCC2* 5'UTR mutation as compared with those having wild-type 5'UTR (cytosolic RCC2 staining: P=0.47, n=101; nuclear RCC2 staining: P=0.50, n=98).

DISCUSSION

This study connects *RCC2* directly to cancer showing that the mutation status of a MNR in the 5'UTR identifies high-risk patients in the MSI group with stage II CRC. We demonstrate that the mutation causes reduced protein expression *in vitro*, and that a reduced RCC2 protein level has functional impact on colon cancer cells. We also show that RCC2 protein expression carries prognostic information across all CRCs, but particularly for female patients with MSS tumors, and that mutations in the *RCC2* coding sequence are rare.

MNRs in UTRs are evolutionary conserved, indicating a functional role[19]. Repeat sequences are prone for alterations in cancers with defect mismatch repair (MSI) and some of these targets are involved in the regulation of gene and protein expression. UTRs are known to affect mRNA nuclear export, cytoplasmic localization, translational efficiency and stability[20]. The majority of translational control occurs at the level of initiation, thus implicating the 5'UTR region as an important site for translational regulation[21]. The $(A)_{10}$ repeat in *RCC2* is located in exon 1 in the 5'UTR of the gene, 77-86 bases upstream of the start codon. The luciferase reporters with mutant and wild-type *RCC2* promoters showed that a single base deletion of this repeat leads to reduced RCC2 protein expression. It should be noted that a recent study suggested *RCC2* to be a downstream target for the known cancer related micro-RNA miR-29c, through its 3'UTR miR-29c target sequence[22].

From global quantification studies of mammalian gene expression it is found that protein levels are predominantly controlled at the translational level[23]. RCC2 mRNA and protein levels display average half-lives of 9h and 48h, respectively. However, the translation rate constant in NIH 3T3 cells is about 230 proteins per mRNA per hour, much higher than the average of 140[23]. This clearly suggests that translational regulation of RCC2, and thereby the effect of a 5'UTR mutation reducing translation reported here, may have important consequences for the level of RCC2 protein.

Furthermore, we demonstrate that knockdown of *RCC2* in colon cancer cells with MSI (HCT15) results in G2-M arrest, in accordance with previous observations in HeLa cells[9]. As a protein fundamental to the integrity and proper function of the chromosomal passenger complex, RCC2 plays an important role in cell cycle regulation, especially during mitosis and cell cleavage. Whether the deletions seen in *RCC2* are sufficient to arrest the cells entering M-phase awaits further investigation, but the siRNA knockdown experiments, combined with the luciferase reporter analysis and the positive survival data for MSI tumors reported here, suggest that at least the cells are halted at the G2-M which is disadvantageous for the tumor cells and beneficial for the patient. Deregulation of the mitotic apparatus may cause some mitotic cells to enter apoptosis, but others to

halt during mitosis, not completing cytokinesis, and thus creating multinuclear cells. This hypothesis is supported by the TUNEL analysis (Fig. 3) that showed increased levels of apoptosis, as well as the cell cycle analysis that demonstrated increased levels of >4N cells. The occurrence of large multinuclear cells following *RCC2* downregulation was also apparent from the confocal microscopy analysis. Together, these results provide an explanation for why the 5'UTR mutation in *RCC2* is associated with improved survival in MSI patients; tumors with a slightly reduced level of RCC2 are likely to grow and develop less efficiently since more cells will reach mitotic arrest. However, it is important to keep in mind that the anticipated 15% protein reduction resulting from the 5'UTR *RCC2* mutation cannot be directly compared with the nearly complete knockdown of RCC2 protein in HCT15 cells, but it makes a functional connection probable. Deleterious mutations would generally not be selected for in a classical driver-oriented paradigm of cancer progression. However, recent evidence suggests that moderately deleterious mutations can evade negative selection and accumulate alongside tumor progression in a balance with "classical" driver mutations[24, 25]. Their model, created by combining evolutionary simulations of cancer development with analysis of cancer sequencing data, provide explanations for phenomena such as spontaneous regression, slow progression and heterogeneous growth rates.

Since only few prognostic markers have been reported for the MSI patient group, our finding has important clinical implications. The *RCC2* mutation test is a simple PCR-test and can be performed alongside assessment of MSI status which already is implemented in clinical guidelines. A simple assay determining the 5'UTR *RCC2* mutation status in tumors with MSI could further guide therapeutic decision-making for this subgroup of patients, particularly for those with stage II cancer, who are not routinely offered adjuvant treatment. Although MSI patients have generally better outcome than MSS patients, they have little benefit from conventional 5-FU treatment[26]. Therefore, high-risk MSI patients identified by the *RCC2* mutation test could be candidates for alternative adjuvant treatment.

The *RCC2* gene maps to chromosome band 1p36 from which allelic loss is an early and common event in colorectal adenomas and carcinomas[27, 28]. Several genes have been suggested to be targets of this loss, and the current data add *RCC2* as a candidate. The chromosome band 1p36 is most often lost in CRCs with MSS, which are typically in the triploid range and exhibit close to complete overlap with the CIN phenotype. The cause(s) underlying the CIN phenotype is presently unknown, but defects in components involved in the regulation of mitosis and cell cycle checkpoints are suggested as initiators[29]. Mutations found across the *RCC2* coding sequence in about 3% of all CRCs, and deletions of 1p sequences, may in part explain the variation of RCC2 staining among CRCs. It should be noted that none of the five missense mutations nor the single frameshift mutation recorded by TCGA were recurrent. The latter leading to a

premature stop codon disrupting the seven-bladed beta-propeller structure of RCC2 which would most likely produce a non-functional protein that is rapidly degraded, but as a consequence the total amount of wt RCC2 may be reduced. Due to the rareness of this mutation its clinical value is marginal.

IHC analysis of RCC2 in the consecutive CRC series revealed that patients with weak cytosolic RCC2 staining had a poor prognosis relative to patients with strong cytosolic RCC2 staining (Figure 4). Statistical tests for interaction demonstrated particular relevance for female patients post hoc, and subgroup analyses showed striking differences in the MSS group, for both cytosolic and nuclear staining. These findings seemingly contradict the previous results, which indicated that the 5'UTR *RCC2* mutation leads to reduced protein expression and a good prognosis. However, the mutation data relate to MSI tumors and a moderate reduction in protein amount. In MSS tumors, the ploidy stem line is, in contrast to MSI tumors, typically in the triploid range, and several pathway disturbances may lead to a decreased protein half-life.

An intriguing explanation for why reduced RCC2 protein expression in MSS tumors is associated with poor prognosis can be drawn from a study that identified RCC2 as a key player in the integrin $\alpha_5\beta_1$-fibronectin signaling network[10]. These experiments in MEF and B16-F10 cells suggested that RCC2 serves as a negative regulator of RAC1 and ARF6, and showed that *RCC2* knockdown resulted in accelerated cell spreading, cell adhesion and reduced directional cell migration. The authors suggested that RCC2 regulates and limits the required signaling by RAC1 and ARF6 to enable proper membrane protrusion and delivery. A recent study shows that RCC2 achieves this by binding to coronin-1C, which is crucial for RAC1 activation[30]. This is concordant with our observations of increased cell spreading and change in cytoskeletal organization following knockdown of *RCC2* in HCT15 cells. We speculate that strong cytosolic staining for the RCC2 protein in CRC indicates active RAC1 and ARF6 inhibition, and that weak RCC2 staining relates to lack of inhibition leading to an increased migratory and metastatic potential. The latter is in keeping with a significant correlation between cytosolic RCC2 expression and tumor stage, as stage I/II tumors more often have strong cytosolic staining, and stage III/IV tumors more often have weak cytosolic staining. In line with this scenario, another recent study demonstrated that RCC2 forms a complex with the well-established F-actin regulator cortactin[31], providing further support for a connection between RCC2 and cell motility and invasion. Cortactin, together with among others RAC1, is tightly involved in the generation of cell membrane protrusions, such as lamellipodias and invadopodias, and several studies have suggested that cortactin plays a role in cancer development[32]. As MSS tumors already suffer from a deranged mitotic apparatus, they might not experience severe negative effects following a reduced level of RCC2 protein. Instead, reduced levels of RCC2 protein might primarily lead to deregulation of RAC1, ARF6, coronin-1C and/or cortactin, followed by enhanced migratory and invasive capacities. A

tentative model of this network has been proposed[33], placing RCC2 between cortactin and RAC1, enabling a dynamic interplay among these partners with probable relevance to cancer development.

The present data adds cancer as an important dimension to the current RCC2 framework. We demonstrate functional consequences of “partly” as well as “total” protein absence in CRC models. RCC2 is the first prognostic marker validated for the MSI subgroup. By simple PCR or protein expression assay RCC2 may serve as a prognostic biomarker for patients with either MSI or MSS tumors.

AUTHOR CONTRIBUTIONS

J.B. participated in the study design, performed all functional studies and IHC experiments and scoring, participated in the genetic analyses, interpreted all results, performed all statistics and drafted the manuscript. M.K. performed IHC experiments and independent scoring, and participated in the statistical analyses. T.A., E.C.R. and G.E.L. performed and scored the mutation analyses of all repetitive tracts. J.B., T.A., T.N. and R.I.S. performed massive parallel sequencing of own samples and evaluated the TCGA dataset. E.L., G.B. and T.J. participated in design and interpretation of the functional assays. M.A.M., T.O.R., A.N., A.L and X-F.S collected the patient samples, provided the MSI data, collected and quality controlled the clinical data. A.N. participated in the survival analyses. A.S. performed morphological identification of qualified tumor areas for TMA construction. K.L. participated in the statistical analyses. R.A.L. conceived and coordinated the study, was responsible for the study design and research biobank, and participated in evaluation and discussion of results and in the drafting of the manuscript. All authors participated in manuscript writing and in scientific discussions.

ACKNOWLEDGEMENTS

We are grateful for all advice from Kirsti Solberg Landsverk at the Flow Cytometry Core Facility at The Cancer Research Institute, Oslo University Hospital HE – The Norwegian Radium Hospital. **Funding:** This study was supported by The Norwegian Cancer Society (RAL: grant no. PR-2006-0442, supporting JB as PhD student and TA as post doctoral research associate; RIS: grant no. PR-2007-0166) and by the South-Eastern Norway Regional Health Authority (RAL: grant. HSØ gruppe 2010. “Genome medicine of CRC”).

DECLARATION OF INTERESTS

We declare that we have no competing interests.

REFERENCES

1. International Agency for Research on Cancer.World Health Organization, "World Cancer Report 2014" (2014).

2. Sinicrope FA, Sargent DJ. Molecular pathways: microsatellite instability in colorectal cancer: prognostic, predictive, and therapeutic implications. *Clin Cancer Res 2012;* 18(6): 1506–1512.

3. Lothe RA, Peltomaki P, Meling GI, Aaltonen LA, Nystrom-Lahti M, Pylkkanen L, *et al.* Genomic instability in colorectal cancer: relationship to clinicopathological variables and family history. Cancer Res 1993; 53: 5849-5852.

4. Thibodeau SN, Bren G, Schaid D. Microsatellite instability in cancer of the proximal colon. Science 1993; 260: 816-819.

5. Popat S, Hubner R, Houlston RS. Systematic review of microsatellite instability and colorectal cancer prognosis. J. Clin. Oncol. 2005; 23: 609-618.

6. Duffy MJ, Lamerz R, Haglund C, Nicolini A, Kalousova M, Holubec L, *et al*. Tumor markers in colorectal cancer, gastric cancer and gastrointestinal stromal cancers: European group on tumor markers (EGTM) 2013 guidelines update. Int. J. Cancer 2013; 10

7. Royrvik EC, Ahlquist T, Rognes T, Lothe RA. Slip slidin' away: a duodecennial review of targeted genes in mismatch repair deficient colorectal cancer. *Crit Rev.* Oncog. 2007; 13: 229-257.

8. Kim NG, Rhee H, Li LS, Kim H, Lee JS, Kim JH *et al.* Identification of MARCKS, FLJ11383 and TAF1B as putative novel target genes in colorectal carcinomas with microsatellite instability. Oncogene 2002; 21: 5081-5087.

9. Mollinari C, Reynaud C, Martineau-Thuillier S, Monier S, Kieffer S, Garin J, *et al.* The mammalian passenger protein TD-60 is an RCC1 family member with an essential role in prometaphase to metaphase progression. Dev. Cell 2003; 5: 295-307.

10. Humphries JD, Byron A, Bass MD, Craig SE, Pinney JW, Knight D, *et al*. Proteomic analysis of integrin-associated complexes identifies RCC2 as a dual regulator of Rac1 and Arf6. Sci. Signal. 2009; 2: ra51.

11. Salahshor S, Kressner U, Fischer H, Lindmark G, Glimelius B, Pahlman L, *et al.* Microsatellite instability in sporadic colorectal cancer is not an independent prognostic factor. Br. J. Cancer 1999; 81(2): 190-193.

12. Merok MA, Ahlquist T, Royrvik EC, Tufteland KF, Hektoen M, Sjo OH, *et al.* Microsatellite instability has a positive prognostic impact on stage II colorectal cancer after complete resection: results from a large, consecutive Norwegian series. Ann. Oncol. 2012.

13. Nome T, Thomassen GO, Bruun J, Ahlquist T, Bakken AC, Hoff AM, *et al.* Common fusion transcripts identified in colorectal cancer cell lines by high-throughput RNA sequencing. Transl. Oncol. 2013; 6(5): 546-553.

14. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, *et al.* Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. Sci. Signal. 2013; 6(269): l1.

15. Sirnes S, Bruun J, Kolberg M, Kjenseth A, Lind GE, Svindland A, *et al.* Connexin43 acts as a colorectal cancer tumor suppressor and predicts disease outcome. Int. J. Cancer 2012; 131: 570-581.

16. Bruun J, Kolberg M, Nesland JM, Svindland A, Nesbakken A, Lothe RA. Prognostic significance of beta-catenin, E-cadherin, and SOX9 in colorectal cancer: results from a large population-representative series. Front. Oncol 2014.

17. Punt CJ, Buyse M, Kohne CH, Hohenberger P, Labianca R, Schmoll HJ, *et al.* Endpoints in adjuvant treatment trials: a systematic review of the literature in colon cancer and proposed definitions for future trials. J. Natl. Cancer Inst. 2007; 99: 998-1003.

18. de la Pena M, Dufour D, Gallego J. Three-way RNA junctions with remote tertiary contacts: a recurrent and highly versatile fold. RNA 2009; 15(11): 1949-1964.

19. Suraweera N, Iacopetta B, Duval A, Compoint A, Tubacher E, Hamelin R. Conservation of mononucleotide repeats within 3' and 5' untranslated regions and their instability in MSI-H colorectal cancer. Oncogene 2001; 20: 7472-7477.

20. Hughes TA. Regulation of gene expression by alternative untranslated regions. Trends Genet. 2006 22: 119-122.

21. Pickering BM,Willis AE. The implications of structured 5' untranslated regions on translation and disease. Semin. Cell Dev. Biol. 2005; 16: 39-47.

22. Matsuo M, Nakada C, Tsukamoto Y, Noguchi T, Uchida T, Hijiya N, *et al.* MiR-29c is downregulated in gastric carcinomas and regulates cell proliferation by targeting RCC2. Mol. Cancer 2013; 12(1): 15.

23. Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, *et al.* Global quantification of mammalian gene expression control. Nature 2011; 19;473(7347): 337-342.

24. McFarland CD, Korolev KS, Kryukov GV, Sunyaev SR, Mirny LA. Impact of deleterious passenger mutations on cancer progression. Proc. Natl. Acad. Sci. U. S. A. 2013; 19;110(8): 2910-2915.

25. McFarland CD, Mirny LA, Korolev KS. Tug-of-war between driver and passenger mutations in cancer and other adaptive processes. Proc. Natl. Acad. Sci. U. S. A 2014; 111: 15138-15143.

26. Sinicrope FA, Foster NR, Thibodeau SN, Marsoni S, Monges G, Labianca R, *et al.* DNA mismatch repair status and colon cancer recurrence and survival in clinical trials of 5-fluorouracil-based adjuvant therapy. J. Natl. Cancer Inst. 2011; 103(11): 863-875.

27. Lothe RA, Andersen SN, Hofstad B, Meling GI, Peltomaki P, Heim S, *et al.* Deletion of 1p loci and microsatellite instability in colorectal polyps. Genes Chromosomes. Cancer 1995; 14: 182-188.

28. Thorstensen L, Qvist H, Heim S, Liefers GJ, Nesland JM, Giercksky KE, *et al.* Evaluation of 1p losses in primary carcinomas, local recurrences and peripheral metastases from colorectal cancer patients. Neoplasia 2000; 2: 514-522.

29. Grady WM,Carethers JM. Genomic and epigenetic instability in colorectal cancer pathogenesis. Gastroenterology 2008; 135: 1079-1099.

30. Williamson RC, Cowell CA, Hammond CL, Bergen D, Roper J, Feng Y, *et al.* Coronin-1C and RCC2 guide mesenchymal migration by trafficking Rac1 and controlling GEF exposure. J. Cell Sci. 2014.

31. Grigera PR, Ma L, Borgman CA, Pinto AF, Sherman NE, Parsons JT *et al.* Mass spectrometric analysis identifies a cortactin-RCC2/TD60 interaction in mitotic cells. J. Proteomics. 2012.

32. Kirkbride KC, Sung BH, Sinha S, Weaver AM. Cortactin: a multifunctional regulator of cellular invasiveness. Cell Adh. Migr. 2011; 5: 187-198.

33. Byron A, Humphries JD, Bass MD, Knight D, Humphries MJ. Proteomic analysis of integrin adhesion complexes. Sci. Signal. 2011; 4: t2.

FIGURE LEGENDS

Fig. 1. Microsatellite mutations in the 5'UTR of *RCC2* are associated with improved outcome for patients with colorectal cancers of the MSI phenotype. Microsatellite mutation frequencies for MSI target genes in the compiled Scandinavian series and Norwegian validation series (**A**). Kaplan-Meier TTR analysis (logrank test) for CRC patients with MSI tumors stratified by the *RCC2* 5'UTR mutation in the test series (**B**), the validation series **(C**), and patients from the validation series with complete resection (R0, no evidence of residual tumor), right-sided tumors and follow-up data of more than 36 months (**D**). Subgroup analysis for patients with stage II CRC (**E**). Nucleotide-level somatic mutation data in CRC from the complete length of *RCC2* (**F**). Red pins represent mutations identified from own samples and yellow pins represent mutations in the dataset from The Cancer Genome Atlas. Orange: UTRs; Blue: coding sequence; Green: protein domains. Abbreviations: RLD - RCC1-like domain; NLS - Nuclear localization signal (putative). Abbreviations: wt – wild type; mut – mutant.

Fig. 2. The *RCC2* 5'UTR mutation causes a reduced protein expression. Constructs where wild type or mutated 5'UTR from *RCC2* were placed upstream of the coding region of the firefly luciferase gene, were transiently transfected into HCT15, HT29 and HeLa cells, followed by determination of luciferase activity in cell extracts. *P*-value from Student's *t*-test for paired samples (two-tailed). The y-axis represents the relative firefly luciferase expression (normalized to the *Renilla* luciferase expression) from the mutated to the wild-type construct.

Fig. 3. siRNA knockdown of *RCC2* in HCT15 cells has functional consequences on cell growth and morphology. Depletion of RCC2 causes reduced proliferation and increased cell spreading of HCT15 cells (**A-B**). HCT15 cells were either transfected with negative control siRNA (mock) or with two different siRNA sequences against *RCC2* (siRNA1 and siRNA2). The cells were assayed after four days of transfection. Representative Western immunoblots showing knockdown of *RCC2* by the two different siRNAs with quantification bar charts (**A**). Verification of knockdown was obtained by three independent experiments for each of the two siRNAs. The protein expression was normalized relative to cells transfected with mock. (**B**) Quantification of cell proliferation. Cells were counted after four days of growth on a 35 mm Petri dish. Each bar represents the average cell number from three independent experiments. Error bars indicate the standard error of the mean. *P*-value from Student's *t*-test for paired samples (two-tailed). Representative images of the morphology of mock-transfected cells (left) and RCC2-depleted cells (right) (**B**). The morphological changes following RCC2 depletion are accompanied by a rearranged actin filament network (**C-E**). HCT15 cells were either transfected with negative control siRNA (mock) (**C**), or with siRNA sequences against *RCC2* (**E-F**). After four days of transfection, the cells were fixed, stained

with anti-RCC2 antibodies and phalloidin to label actin, and visualized using immunofluorescence confocal microscopy. An apparent increase in the number of large multinuclear cells was observed in cells with reduced RCC2 protein (asterisk). Scale bar represents 20 µm and applies to all images. Depletion of RCC2 causes cell cycle arrest in the G2-M phase and increased levels of apoptosis (**F-H**). HCT15 cells were either left untreated or transfected with negative control siRNA (mock) or with siRNA sequences against *RCC2* (siRNA1) (**F**) Visualizations of representative cell cycle phase distributions for each treatment, as seen by flow cytometry. (**G**) Quantification bar charts for cell cycle phase distributions. (**H**) Quantification bar charts for levels of apoptosis, as seen by TUNEL-assay. Error bars indicate the standard error of the mean from at least three independent experiments with at least three biological replicates for each experimental variable. *P*-values indicate significant differences in cell cycle phase distributions between mock- and siRNA-transfected cells (two-tailed Student's *t*-test for paired samples).

Fig. 4. Weak staining for the RCC2 protein is associated with poor patient outcome, particularly for female CRC patients with MSS tumors. **(A-H)** Differential immunohistochemical staining of RCC2 in colorectal carcinomas. Staining of tumor sections using an RCC2 selective antibody demonstrated a specific staining in both the cytosol and in the nuclei of the cells. The staining was largely confined to epithelial cells, but some scattered staining was also observed in the tumor stroma. Representative staining of histospots (core diameter 0.6 mm) is illustrated at 400x magnification. (**A-D**) Strong-, moderate-, weak- and negative nuclear staining, respectively. (**E**) Strong nuclear and cytosolic staining. (**F**) Predominant cytosolic staining. (**G**) Staining in undifferentiated tissue. (**H**) Staining in normal colon. **(I-L)** Survival analysis for RCC2 cytosolic (left) and nuclear (right) protein expression in a consecutive CRC series (*n*=903). Green curves represent strong staining and blue curves represent weak staining. (**I**) Five-year overall survival. (**J-L**) Five-year time-to-recurrence analysis for patients with complete resection (R0, no evidence of residual tumor). (**K**) Subgroup analysis for gender. (**L**) Subgroup analysis for gender and MSS. The Kaplan-Meier method (logrank test) was used to draw the plots and determine statistical differences between patients with strong and weak RCC2 staining.

Variable	Patients (%)	*n*	Univariate analysis[a]		Multivariate analysis[b]		
			5-year TTR (%)	*P*	HR	95% CI	*P*
Total	122(100)						
Gender							
Female	92 (75)		67.0		1		
Male	30 (25)		78.6	0.29	0.64	0.25-1.66	0.36
Age[c]				-	1.03	1.00-1.06	0.047
RCC2							
Wild type	38 (31)		54.1		1		
Mutated	84 (69)		77.4	0.0071	0.37	0.18-0.77	0.0080
Stage							
I+II	83(68)		83.3		1		
III+IV	39 (32)		38.0	$1.8x10^{-8}$	6.5	3.06-13.8	$1x10^{-6}$
Histopathologic Grade							
G1 + G2	69 (57)		78.7		1		
G3	47 (38)		60.5	0.063	1.8	0.88-3.59	0.11
Mucinous*	5 (4)						
ND*	1 (1)						

Table 1. Univariate and multivariate analyses of *RCC2* 5'UTR mutation status in the Validation series. [a]Kaplan-Meier estimate (logrank test). [b]Cox regression model (Wald test) - all included parameters are displayed in the table. [c]Hazard ratios are given per year of age. *Excluded from the statistical analyses. Abbreviations: G1 – High differentiation; G2 – Moderate differentiation; G3 – Poor differentiation.

Variable	Patients *n* (%)	Univariate analysis[a]		Multivariate analysis[b]		
		5-year OS/TTR (%)	*P*	HR	95% CI	*P*
Overall survival						
Total	903 (100)					
Gender						
Female	474 (52)	49.6		1		
Male	429 (48)	51.0	0.40	0.83	0.61-1.14	0.25
Age[c]			-	1.04	1.03-1.05	$3.4x10^{-13}$
RCC2						
cytosolic staining						
Weak	319 (35)	39.5		1		
Strong	462 (51)	58.4	$5.2x10^{-7}$	0.61	0.45-0.83	0.0015
ND*	122 (14)					
Stage						
I	133(15)	77.4		1		
II	363(40)	62.3		1.61	1.03-2.52	
III	237(26)	47.3		2.49	1.59-3.91	
IV	165(18)	6.7	$1.6x10^{-56}$	3.66	2.00-6.70	$9.5x10^{-6}$
ND*	5(1)					
Tumor location						
Proximal colon	367 (41)	48.2		1		
Distal colon	302 (33)	46.4		0.98	0.76-1.27	
Rectum	234 (26)	58.5	0.0061	0.86	0.63-1.17	0.59
Histopathologic grade						
G1	84 (9)	63.1		1		
G2	674 (75)	51.0		1.17	0.76-1.79	
G3	108 (12)	33.3	$1.6x10^{-7}$	2.64	1.60-4.37	$1.1x10^{-6}$
Mucinous*	12 (1)					

ND*	25 (3)					
MSI/MSS						
MSI	119 (13)	58.8		1		
MSS	700 (78)	48.1	0.039	1.99	1.33-2.98	0.0085
ND*	84 (9)					
Residual tumor						
R0	714 (79)	61.2		1		
R1	19 (2)	42.1		1.73	0.81-3.70	
R2	170 (19)	5.9	6.2×10^{-85}	3.48	2.23-5.42	1.3×10^{-7}
Interaction parameter						
Gender x RCC2				1.84	1.19-2.84	0.0059
Time-to-recurrence						
Total	714 (100)					
Gender						
Female	365(51)	69.0		1		
Male	349(49)	68.3	0.96	0.66	0.42-1.05	0.079
Age[c]			-	1.03	1.01-1.04	0.0066
RCC2						
cytosolic staining						
Weak	(238)	60.2		1		
Strong	(381)	75.2	8.3×10^{-5}	0.38	0.24-0.58	1.4×10^{-5}
Stage						
I	133(19)	87.6		1		
II	340(48)	71.3		2.08	2.17-3.70	
III	221(31)	55.6		3.98	2.25-7.05	
IV	15(2)	28.6	6.2×10^{-14}	7.30	3.05-17.5	5.4×10^{-8}

ND*	5(1)					
Tumor location						
Proximal colon	284 (40)	68.9		1		
Distal colon	226 (32)	67.8		1.00	0.67-1.49	
Rectum	204 (29)	69.2	0.82	1.14	0.75-1.75	0.78
Histopathologic grade						
G1	79 (11)	76.6		1		
G2	537 (75)	69.2		0.94	0.54-1.62	
G3	70 (10)	52.9	0.0016	2.06	1.05-4.07	0.0056
Mucinous*	9 (1)					
ND*	19 (3)					
MSI/MSS						
MSI	103 (14)	73.0		1		
MSS	543 (76)	66.4	0.20	1.82	1.05-3.18	0.034
ND	68 (10)					
Interaction parameter						
Gender x RCC2				2.58	1.36-4.90	0.0037

Table 2. Univariate and multivariate analyses of cytosolic RCC2 *in situ* protein expression in a consecutive CRC series. [a]Kaplan-Meier estimate (logrank test). [b]Cox regression model (Wald test) - all included parameters are displayed in the table. [c]Hazard ratios are given per year of age. *Excluded from the statistical analyses. Abbreviations: G1 – High differentiation; G2 – Moderate differentiation; G3 – Poor differentiation; R0 – Complete resection – no residual tumor; R1 – microscopic residual cancer at the resection margin; R2 – Macroscopic or radiological evidence of residual cancer, locally or distant.