Cancer Immunol Immunother (2002) 51: 574–582
DOI 10.1007/s00262-002-0322-2

ORIGINAL ARTICLE

Aija Line · Zane Slucka · Aivars Stengrevics
Karina Silina · Geng Li · Robert C. Rees

Characterisation of tumour-associated antigens in colon cancer

Received: 23 May 2002 / Accepted: 1 August 2002 / Published online: 19 September 2002

Abstract In order to search for clinically relevant cancer-associated genes and to define further the spectrum of immunogenic proteins, we applied SEREX (serological identification of antigens by recombinant expression cloning) to analyse genes expressed in colon adenocarcinoma. Eight different serum-reactive cDNA clones were isolated by immunoscreening from a colon cancer-derived cDNA expression library. mRNA expression studies showed that 2 of them, RHAMM and AD034, have a differential tissue distribution, and that 3 genes, NAP1L1, RHAMM and AD034, are overexpressed in tumours in comparison with the adjacent non-cancerous tissues. 5′ RLM-RACE analysis of AD034, a sequence with a tyrosine kinase motif, revealed a frameshifting insertion of 32 bp, most likely generated by use of cryptic splice site in tumour-derived cDNA. Analysis of full-length RHAMM cDNA sequence revealed the presence of two splice variants, which are known to have a different sub-cellular localisation; expression of these splice variants is altered in colon cancer tissues. Serological responses to three antigens (C21ORF2, EPRS and NAP1L1) were found mainly in cancer patients' sera.

Keywords AD034 · Autoantibody · NAP1L1 · RHAMM · SEREX

Data deposition cDNA sequence of the AD034 splice variant reported in this paper has been deposited in the GenBank data base, accession number AY094356.

A. Line (✉) · Z. Slucka · K. Silina
Biomedical Research and Study Centre,
University of Latvia, 1 Ratsupites St, LV-1067, Riga, Latvia
E-mail: aija@biomed.lu.lv
Tel.: +371-780-8208
Fax: +371-744-2407

A. Stengrevics
Latvian Oncology Center, Riga, Latvia

G. Li · R.C. Rees
Department of Life Sciences,
Nottingham Trent University, Nottingham, UK

Introduction

Identification of immunogenic proteins in cancer is essential for the development of immunotherapeutic strategies where adoptive immunity is directed towards major histocompatibility complex (MHC) class I and class II-associated peptides [15, 20]. Many antigens are implicated in the aetiology and progression of cancer, and are associated with genetic and epigenetic events. Preclinical and clinical studies indicate that vaccination, targeting MHC-associated peptide antigens promotes tumour rejection [2, 23]. Serological identification of antigens by recombinant expression cloning (SEREX) is an approach for the identification of antigenic proteins in cancer, and was first used by Sahin et al. [25] in 1995. Since then, it has been applied to a variety of human tumour types resulting in the detection of more than 1,500 potential serologically active tumour antigens. SEREX involves the immunoscreening of tumour-derived cDNA expression libraries with cancer patients' sera, thus allowing a systematic search for genes whose products have elicited antibody production in the tumour-bearing host [21, 25, 32]. SEREX-detected antigens can arise from normal differentiation antigens, overexpressed antigens, cancer-testis antigens, mutated gene or translocation products, aberrant pre-mRNA processing and also from cancer-independent autoantigens [18, 21, 24, 32]. Although the clinical significance for the majority of SEREX-defined antigens still remains obscure, several of SEREX-defined antigens have already provided attractive candidates for the construction of cancer vaccines. For instance, "cancer-testis" antigen NY-ESO-1 was identified by SEREX in oesophageal cancer [6] and is regarded as one of the most immunogenic and widely expressed tumour antigens. Antibody responses to NY-ESO-1 have been observed in 40% to 50% of patients with NY-ESO-1-expressing tumours [29], and antibody production strongly correlates with $CD8^+$T cell responses in these patients [13]. Clinical trials, investigating the

immunological effects of vaccination with NY-ESO-1 peptides, are ongoing [14]. The fact that several previously known cytotoxic T lymphocyte (CTL) targets, for example tyrosinase and MAGE-1, have been detected by SEREX, also supports the assumption that mounting an immune reaction to a tumour antigen requires coordinated $CD4^+$ T, $CD8^+$ T and B cell responses. Moreover, mutated p53 [26], putative tumour suppressor ING1 [12], adhesion molecule galectin-9 [33] etc., have been detected by SEREX, thus showing that analysis of the autoantibody repertoire can also serve as tool for discovery of genes that play a role in cancer aetiology and may serve as diagnostic markers or indicators of progression of the disease.

Previous SEREX analysis of colon cancer by Scanlan et al. [26] resulted in the detection of 48 distinct antigens. Twenty-eight of them reacted only with autologous sera, 14 had similar reactivity with sera from both healthy donors and cancer patients, but 6 reacted only with one or more cancer patients' sera. In the present study we applied SEREX to study colon carcinoma with the aim of defining further the spectrum of immunogenic proteins in colon cancer. Eight different antigens were recognised by cancer patients' sera and were further characterised by sequence analysis, mRNA expression pattern and reactivity with allogeneic sera.

Materials and methods

Tissue specimens and patient sera

Colorectal and gastric cancer tissue and the adjacent non-cancerous tissue specimens from 22 patients undergoing surgery at the Latvian Oncology Center were resected and frozen in liquid nitrogen immediately after surgery. In addition, serum samples were obtained from colon, stomach and breast cancer patients undergoing diagnostic procedures, and from healthy volunteers. The study was approved by the Committee of Medical Ethics of Latvia and the tissue samples and sera were collected after the patients' informed consent had been obtained.

Isolation of total RNA and construction of the cDNA library

Total RNA was isolated from tumour and normal tissue samples, using Trizol according to the manufacturer's instructions (Life Technologies). A cDNA expression library was constructed from a tumour specimen of a moderately differentiated colon adenocarcinoma. $Poly(A)^+$ RNA was purified from total RNA using Dynabeads mRNA purification kit (Dynal, Norway) and cDNA was ligated into the lambda Uni-ZAP XR vector using a Gigapack III Gold cloning kit (Stratagene). After in vitro packaging, a library containing 10^6 primary cDNA clones was obtained and amplified once prior to immunoscreening.

Immunoscreening

Immunoscreening of the cDNA library was performed as described by Sahin et al. [25]. Briefly, *Escherichia coli* XL1 blue MRF' cells were transfected with the recombinant phages, plated at a density of approximately 5,000 pfu/150-mm plate (NZCYM-IPTG agar) and transferred to nitrocellulose filters following 8 h incubation at 37°C. In order to eliminate cDNA clones encoding human immunoglobins, filters were pre-screened with AP-conjugated rabbit anti-human secondary antibody (Pierce, USA) prior to incubation with sera, and reactive plaques were detected with 5-bromo-4-chloro-3-indolyl-phosphate/nitroblue tetrazolium (BCIP/NBT) and marked. Then the filters were incubated with 1:250 diluted patient's serum, which had been previously preabsorbed with *E. coli*-phage lysate; serum-reactive clones were detected with AP-conjugated secondary antibody and visualised by incubating with BCIP/NBT. The reactive phage clones were subcloned to monoclonality and converted to pBluescript phagemids. To assess frequencies of antibody responses to the SEREX-defined antigens in allogeneic sera, *E. coli* were transfected directly on the gridded agar plate by spotting 1 μl of monoclonal-positive phage (20–30 pfu/μl) side by side with non-recombinant phages. "Phage arrays" were screened with 1:200 diluted allogeneic sera as described above, excluding the IgG pre-screening step.

DNA sequencing and sequence analysis

Phagemid DNA was purified using QIAprep Spin Miniprep kit (Qiagen), analysed by *Eco*RI/ *Xho*I restriction enzyme digestion, and clones representing different cDNA inserts were sequenced using a BigDye Terminator Cycle Sequencing Ready Reaction kit on an ABI PRISM 3100 genetic analyser (Applied Biosystems). Gene-specific primers were designated to obtain full insert sequences. Genes were identified by homology search through the GenBank data base (www.ncbi.nlm.nih.gov/BLAST). Chromosomal localisation and exon-intron organisation of the cDNA was determined by comparison to the working draft of the human genome. Putative protein domains were predicted by scanning the sequences against PROSITE (www.expasy.org) and by using tools for sequence analysis at the SEREX website (www-ludwig.unil.ch/SEREX).

Western blot analysis

Immunoreactivity to the recombinant proteins in serum-reactive clones was confirmed by western blot analysis. *E. coli* XL1-blue cells were transformed with the recombinant pBluescript phagemids excised from the Uni-ZAP XR vector. The cells were grown in LB medium with ampicillin to an optical density (OD) of 0.4 at 540 nm, and then transcription from the *lac*Z promoter was induced with 2 mM IPTG. Samples of the bacterial cultures were collected before induction and 3 and 5 h after protein expression was induced. The cells were lysed with 3×Laemli buffer, lysates were separated by SDS-PAGE and blotted to Hybond c-extra filters (Amersham Biosciences). The filters were blocked with fat-free milk, incubated with the autologous patient serum and antigen-antibody complexes were detected with HRP-conjugated rabbit anti-human antibody using an ECL detection system (Amersham Biosciences).

Comparative RT-PCR analysis

The mRNA expression pattern of SEREX-defined antigens was analysed by RT-PCR using total RNA from a panel of normal tissues (whole brain, liver, heart, kidney, lung, trachea; Clontech), (stomach, colon, spleen, testis, ovary; Ambion), peripheral blood lymphocytes (PBL) and a specimen of colon cancer of the autologous patient. Relative mRNA levels were compared between cancerous and adjacent non-cancerous tissues of 15 colorectal cancer and 7 gastric cancer patients by comparative RT-PCR. The first-strand cDNA was synthesised from 4 μg of total RNA primed with oligo-dT(18) and random hexamer primers using a First-Strand cDNA Synthesis kit (Fermentas, Lithuania). Gene-specific PCR primers located within different exons were designed to amplify cDNA fragments (250–350 bp in length) of 3 SEREX-defined genes and GAPDH and β-actin as internal standard genes. One fiftieth of the RT mixture was amplified by a GeneAmp PCR System 2400 thermal cycler (Perkin-Elmer) in a total reaction volume of 20 μl containing 10 pmol of each primer, 200 μM dNTP and 2 U of Taq

polymerase (Fermentas, Lithuania). Optimisation of cycling conditions (amount of input cDNA and number of cycles) was performed as described by Toh et al. [31]. Amplification of all target genes was performed simultaneously at the same cycling conditions (45 s at 94°C, 30 s at 58°C, 45 s at 72°C), except for the number of cycles that was different for the amplification of each target gene. The primer sequences, number of cycles used and length of PCR products are shown in Table 1. The quantity of RT-PCR products was determined densitometrically after scanning the ethidium bromide-stained gel on digital gel documentation and analysis system GDS8000 (Ultra-Violet Products, U.K.) and the intensities of bands were calculated using GelWorks software. Standard curves of the amplification of each target gene were constructed from a series of PCR with ten 1.5-fold dilutions of the colon cancer cDNA. Amounts of PCR products were linearly dependent from the input cDNA over 10-fold dilutions of cDNA. The relative amounts of target mRNA were normalised to GAPDH and β-actin. The obtained values in tumours (T) were compared to those in matched normal epithelium (N), and T/N ratios were calculated for each mRNA in each patient's tissue samples. Each reaction was performed in duplicate.

This approach, although frequently used to compare the relative abundance of mRNA in clinical specimens [19, 30, 31], has a major drawback, namely variations in the expression of commonly used standard genes (actins, tubulins, GAPDH). Significant differences in the expression of β-actin and GAPDH in the tissue samples from 3 colon cancer patients were observed.

5′ RLM-RACE analysis of Co23

The full-length 5′ end of the Co23 cDNA sequence was cloned from autologous colon cancer tissues from a patient patient using a FirstChoice RLM-RACE kit (Ambion) according to the manufacturer's instructions. Briefly, 10 μg of total RNA was treated with calf intestinal phosphatase to remove 5'-phosphate from uncapped RNA (degraded mRNA, rRNA, tRNA or DNA), then the cap structure was removed from the full-length mRNA by tobacco acid pyrophosphatase and RNA adapters were ligated to mRNA molecules containing 5′ phosphate. A random-primed RT-nested PCR with gene-specific and adapter-specific primers was performed, products were cloned using InsT/Aclone PCR product cloning kit (Fermentas, Lithuania) and multiple clones were sequenced.

RHAMM full-length cDNA sequence analysis

In order to analyse a full-length cDNA sequence of RHAMM, the entire coding region was amplified from autologous colon cancer and normal colon RNA (Ambion). PCR primers were designed according to the published sequence, (GB, AF032862) and are shown in Table 1. A single band of approximately 2.2 kb was obtained and directly sequenced on an ABI prism 3100 genetic analyser (Applied Biosystems). Sequencing showed that this band represented a mixture of two RHAMM splice variants. Primers flanking the alternative exon 5 were selected, and two RT-PCR products of 394 and 346 bp were obtained from cancerous and normal colon RNA, separated on a 2% agarose gel, extracted from agarose by QIA quick gel extraction kit (Qiagen), and directly sequenced. Expression of these splice variants in a panel of normal tissues and paired colon cancer and adjacent tissues was analysed as described above for comparative RT-PCR analysis. To analyse the cDNA sequence around the alternative splice sites, these amplification products were extracted from the agarose gel and directly sequenced.

Results

Immunoscreening and identification of immunoreactive cDNA clones

Fourteen serum-reactive cDNA clones were detected by immunoscreening of 8×10^5 pfu from a colon cancer cDNA expression library with autologous serum. The

Table 1. Primers used for expression analysis of SEREX-defined antigens

Gene		Primer sequences (5′-3′)	No. of cycles	Size of product (bp)
C21ORF2	F	GACAACCAGGCTGTGACGG	32	303
	R	CAGCAGGATGGCAGTCAGG		
RHAMM (ex. 12–15)	F	AGAGGAAGCAAAATCTAGAGC	29	363
	R	GTTTCCTTTAGTGCTGACTTGG		
RHAMM[a](ex. 1–19)	F	CTTCAGTTTCTGGAGCTGGC	35	2228
	R	TAATCTTCAACAGGTTTCTCAG		
RHAMM[b](ex. 1–6)	F	CTTCAGTTTCTGGAGCTGGC	31	394v346
	R	CAGCATTTAGCCTTGCTTCC		
AD034	F	CTTATCTCTCAAAGGCCATGG	28	276
	R	GATTTTCTAGGAGTGCAGGG		
AD034[b](ex. 2–3)	F	ATGATGATGACTGGGACTGG	32	176v144
	R	GTAAGACCTTGTCTGCTGG		
EPRS	F	AAAACCACTAAGGTCACTTGG	30	252
	R	ATATGAAGAATCCTCTTCTCTGG		
NAP1L1	F	AAGAAAGACTTGATGGTCTGG	25	246
	R	CCATTCACATTCTTCTTCCG		
TAX1BP1	F	AGACAGAACGATGGCAGACC	29	378
	R	GTAGAGGCAGAAGTGGCTGG		
β-actin	F	AGTGTGACGTGGACATCCG	20	351
	R	AATCTCATCTTGTTTTCTGCGC		
GAPDH	F	GTCATCCCTGAGCTAGACGG	25	356
	R	GGGTCTTACTCCTTGGAGGC		

[a]Primers located in exons 1 and 19 were used to amplify entire coding region of RHAMM for cDNA sequence analysis but a set of internal primers was used for direct sequencing of the RT-PCR product

[b]These sets of primers were used for analysis of the expression of RHAMM and AD034 splice variants, respectively; *F* forward primer; *R* reverse primer

clones were purified, full-length sequences of their cDNA inserts were obtained and the genes were identified by homology search through the GenBank data base. This showed that the 14 reactive clones were derived from 8 different genes (see Table 2). No novel sequence variations were detected in these cDNA when compared with the corresponding sequences in the GenBank data base.

To verify the immunoscreening results, serum-reactivity to 4 SEREX-detected antigens (C21ORF2, RHAMM, EPRS and NAP1L1) was confirmed by western blot analysis. These antigens were expressed in *E. coli* as fusion proteins with β-galactosidase under the *lac*Z promoter and probed with the serum used for immunoscreening of the library. Besides bacterial proteins, a single IPTG-inducible protein of predicted molecular weight was detected in each case (an example of analysis is shown in Fig. 1). *E. coli* transformed with non-recombinant pBluescript was used as a negative control.

Frequency of antibody in allogeneic sera

The frequency of antibodies reactive to the identified antigens in the sera of healthy donors ($n=35$) and patients diagnosed with colon cancer ($n=24$), stomach cancer ($n=22$), breast cancer ($n=11$) and prostate cancer ($n=3$) was determined (Table 3). Three antigens (C21ORF2, EPRS and NAP1L1) did not react with any of the sera from the healthy individuals but did react with one or more serum samples from cancer patients tested in this study. Other antigens had similar reactivity with sera from both cancer patients and controls.

mRNA expression of SEREX-defined antigens

mRNA expression of SEREX-defined antigens was analysed by RT-PCR in normal tissues (brain, liver, heart, kidney, lung, trachea, spleen, colon, stomach, testis, ovary and PBL) and in an autologous specimen of colon cancer tissue from the same patient. Cycling conditions and the optimal number of cycles were chosen so that the PCR products were at the linear phase of amplification. GAPDH and β-actin were used as controls for RNA integrity and quantity. This allowed the assessment of the abundance of each mRNA in normal tissues relative to the autologous colon cancer tissue. C21ORF2, EPRS, NAP1L1 and TAX1BP1 were expressed in most of the tissues analysed, while RHAMM and AD034 showed differential tissue distribution. RHAMM showed the strongest expression in testis, spleen and colon cancer tissue and AD034 was expressed in testis, spleen, colon, stomach and colon cancer tissues (Fig. 2A).

Comparison of mRNA levels in colon cancer and adjacent non-cancerous tissues

To evaluate whether the antigens with a relatively high expression in the autologous tumour tissue (RHAMM, AD034, NAP1L1) were overexpressed in other colorectal and gastric cancers, we compared their relative mRNA levels between cancerous and paired adjacent tissue specimens of 15 patients with colorectal cancer and 7 patients with gastric cancer by RT-PCR. The conditions for amplification of each target gene were

Table 2. Genes identified by SEREX analysis of colon cancer

Clone	No. of clones	Gene symbol[a]	NCBI reference sequence	Chr. location	Product, function, features, expression pattern
Co2	3	C21ORF2	NM_004928	21q22.3	Function unknown. Ubiquitously expressed
Co7	4	RHAMM	NM_012485	5q33.2–qter	Receptor for hyaluronan-mediated motility. Involved in cell motility and metastasis. Overexpressed in various types of cancer
Co23	1	AD034	NM_031480	6p24.2	Hypothetical protein AD034. Function unknown. Tyr kinase catalytic domain motif and Asp and Lys- rich regions
Co24	1	EPRS	NM_004446	1q41–q42	Glu-Pro tRNA synthetase. Aminoacylation of tRNA
Co25	1	NAP1L1	NM_004537	12q15	Nucleosome assembly protein 1-like 1. Possibly involved in modulating chromatin formation and regulation of cell proliferation
Co26	1	GOLGA4	NM_002078	3p22–p21.3	Golgi autoantigen subfamily a 4; Autoantibodies are detected in several autoimmune diseases.
Co29	1	GOLGB1	NM_004487	3q13	Golgi autoantigen subfamily b; Autoantibodies are detected in several autoimmune diseases
Co32	2	TAX1BP1	NM_006024	7p15	Tax1 (human T-cell leukemia virus type I) binding protein 1. Interacts with anti-apoptptic protein A20. Inhibits apoptosis induced by TNF when overexpressed

[a]HUGO-approved gene symbols were used when available

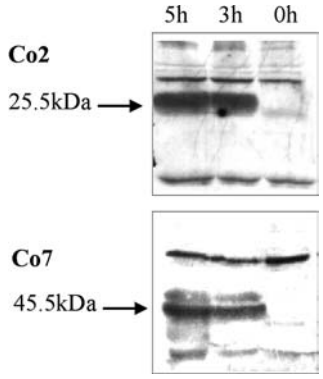

Fig. 1. Western blot analysis of reactivity of patient's serum to 2 SEREX-detected antigens, Co2 (C21ORF2) and Co7 (RHAMM). *E. coli* XL1-blue cells were transformed with the recombinant pBluescript phagemids excised from Co2 and Co7 phage clones, transcription was induced by 2 mM IPTG and the samples were collected before induction (0 h), and 3 and 5 h after induction. Besides *E. coli* proteins, a single IPTG-inducible protein of predicted molecular weight was detected in each case (predicted molecular weight of β-galactosidase-C21ORF2 fusion protein was 25.73 kDa; and of β-galactosidase-RHAMM, 45.66 kDa)

optimised so that the amount of PCR product was in linear relationship to the amount of cDNA input, at least over 10-fold dilution of cDNA input. GAPDH and β-actin were used as internal controls. An example of this analysis is shown in Fig. 3A. Relative quantities of RT-PCR products were determined by densitometric analysis, the amounts of target cDNA were normalised to that of β-actin or GAPDH and tumour/normal (T/N) ratios were calculated. Ratios ≥2 (the mean values of 2 independent experiments) were considered to represent significant overexpression. We observed a 2.1- to 9.3-fold increase of NAP1L1 in 7 colorectal and 1 gastric cancer specimen, a 2.2- to 7.5-fold increase of RHAMM in 6 colorectal and 2 gastric cancer specimens, and a 2.0- to 4.8-fold increase of AD034 in 4 colorectal cancer specimens when compared to the adjacent tissues (normalised to β-actin). In the remaining cases the amounts of PCR products were similar in cancer and non-cancerous tissues. In order to confirm the results obtained, further PCR analyses were performed (Fig. 3B), in which the amounts of cDNA input were varied according to the obtained T/N values (e.g. if 2.9-fold overexpression of NAP1L1 in tumour was observed, the amount of tumour cDNA used for PCR was decreased 2.9-fold in the next experiment, resulting in similar NAP1L1 PCR products in tumour and normal tissues and a respective decrease of the T/N ratio (approx. 3-fold) of GAPDH. These results indicate that this form of semi-quantitative analysis is capable of detecting differences in mRNA levels within a given range of variation, and that some genes are over-expressed in cancer compared to adjacent normal tissues.

Cloning of AD034 mRNA 5′ variants

Clone Co23 contains a partial cDNA sequence encoding hypothetical protein AD034. The longest ORF of AD034 encodes a 561-amino acid protein of approximately 64.6 kDa. Comparison of the predicted amino-acid sequence with Prosite and Pfam data bases revealed a similarity to the RIO1/ZK632.3/MJ0444 protein family (aa 186–380), the tyrosine kinase active-site signature (aa 313–325) and the aspartic acid and lysine-rich regions (aa 10–66 and 514–561, respectively). To determine the transcription start site and to search for possible sequence variations in the 5′ region of AD034 that was absent in clone Co23, 5′ RLM-RACE analysis was performed using total RNA from autologous tumour tissues. 5′ ends of the sequenced RLM-RACE clones differed by 165 bp, indicating that the transcription start site of AD034 was scattered within this region. The longest RACE clone extended the AD034 mRNA sequence by 37 bp; however, no additional translation initiation site was found. Of the 8 clones sequenced, 3 contained a 32 bp insertion (submitted to GenBank; AY094356). Alignment with the genomic sequence (NT_023412) showed that the inserted 32 bp were derived from the intronic sequence flanking exon 3, and presumably are included in the mRNA via the cryptic splice site. The insertion shifts the reading frame and introduces a stop codon, resulting in a truncated ORF of 91 aa. RT-PCR analyses of expression of the splice variants showed that the transcript containing the 32 bp sequence was just a minor mRNA variant and is detectable in all normal tissues where AD034 is expressed, and no significant differences in the ratios of either of the splice variants were observed between colorectal tumours and adjacent normal tissues.

Table 3. Reactivity of SEREX-defined antigens with sera from healthy donors and cancer patients

Antigen	Healthy donor	Colon cancer[a]	Gastric cancer	Breast cancer	Prostate cancer
C21ORF2	0/35	1/24	0/22	1/11	1/3
RHAMM	3/35	5/24	4/22	1/11	1/3
AD034	2/25	1/24	1/22	0/11	0/3
EPRS	0/25	4/24	1/22	2/11	0/3
NAP1L1	0/25	2/24	1/22	0/11	0/3
GOLGA4	5/13	2/5	nd	nd	2/3
GOLGB1	3/13	2/5	nd	nd	nd
TAX1BP1	9/35	8/24	7/22	3/11	3/3

[a]Includes the reactivity with autologous serum; *nd* not done

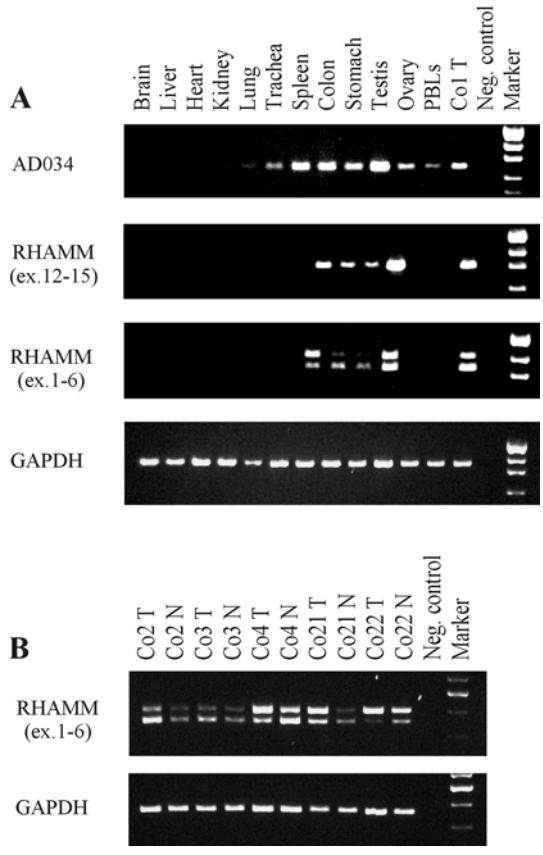

Fig. 2. **A** Expression of AD034 and RHAMM mRNA in normal tissues and autologous tumour (*Co1 T*) analysed by RT-PCR. RHAMM primers located in exons 1 and 6 are flanking the alternative exon, demonstrating the presence of two splice variants: $RHAMM^{FL}$ (394 bp) and $RHAMM^{-ex5}$ (346 bp). As no variations were observed in the 3′ region of RHAMM, amplification with primers located in exons 12 and 15 should represent total RHAMM cDNA. GAPDH was amplified as an internal control and demonstrates the equal amounts of mRNA used for RT-PCR. **B** Expression analysis of $RHAMM^{FL}$ and $RHAMM^{-ex5}$ mRNA in paired tumour (*T*) and adjacent non-cancerous tissues (*N*) of 5 colon cancer patients

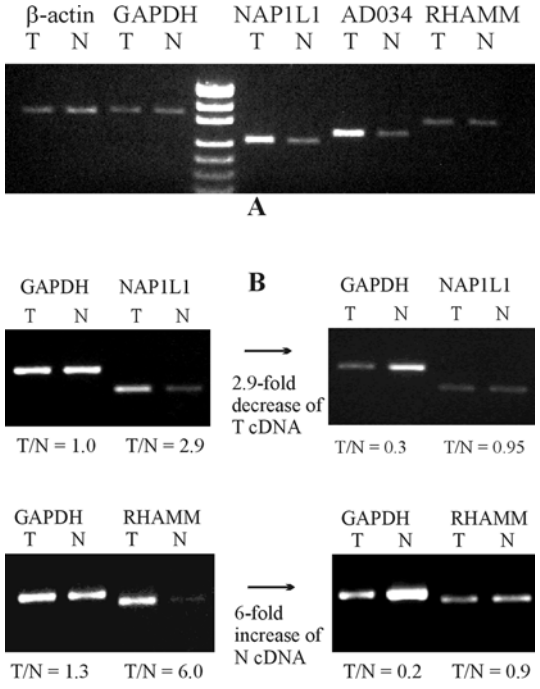

Fig. 3. **A** An example of a comparison between NAP1L1, AD034 and RHAMM mRNA levels in cancerous (*T*) and adjacent non-cancerous (*N*) tissues by RT-PCR. Cycling conditions were optimised so that the RT-PCR products were analysed when amplification was within the linear phase. Ethidium bromide-stained gels were scanned on a digital gel documentation system, band intensities were calculated, and relative expression coefficients were determined using standard amplification curves, and expression of each target gene was normalised to that of *β*-actin and GAPDH. In the example shown, a 3.3-fold increase in NAP1L1 and a 4.8-fold increase (the mean values of 2 independent experiments) in AD034 in cancerous tissues was observed (normalised to *β*-actin). **B** An example of a RT-PCR assay is shown in which the amounts of cDNA input were varied according to the obtained T/N ratios to demonstrate the accuracy of the densitomertric analysis. A 2.9-fold increase in NAP1L1 mRNA levels in a gastric cancer (*T*) specimen was observed as compared with the adjacent normal tissues (*N; upper left*), while the PCR products of GAPDH were similar. In the next experiment, 2.9-fold diluted T cDNA was used for amplification, resulting in a similar NAP1L1 amplification product and a respective decrease in the T/N ratio (approx. 3-fold) of GAPDH (*upper right*). Similarly, a 6-fold increase (4.6-fold when normalised to GAPDH) in RHAMM mRNA levels was observed in a gastric cancer specimen compared with matched normal tissue (*lower left*); when the amount of cDNA from normal tissue was increased 6-fold, similar RHAMM RT-PCR products and an approximately 6.5-fold decrease in the T/N ratio for GAPDH were observed

Expression of RHAMM splice variants

RHAMM cDNA contains several coding mononucleotide repeat sequences in the 5′ and 3′ sequence regions that are absent in clone Co7. Such sequences are particularly prone to mutations [8]. Therefore we made an attempt to obtain the full-length RHAMM cDNA sequence from the autologous colon cancer tissue specimen. Direct sequencing of the PCR product containing the entire coding region showed the presence of 2 splice variants differing in 48 bp encoded by exon 5 (here designated as $RHAMM^{FL}$ and $RHAMM^{-ex5}$), and revealed C-G polymorphism (Arg to Ser) in exon 6 (exons are numbered according to mRNA-genomic alignments at LocusLink on the NCBI server). RHAMM transcript variants differing in exon 5 have been described previously [5], but no evidence for the expression of RHAMM transcripts differing in exon 14, as previously described by Assmann et al. [4] in cancerous and normal colon tissues was found in this study. Sequence analysis showed that an additional AAG triplet (Lys) was inserted at the 5′ end of the exon 5 in comparison with the sequence reported by Assmann et al. [5]. Further analysis of this sequence in 10 paired samples of colon cancer and the adjacent tissues, 3 breast tumour samples and normal colon, spleen and testis showed that an AAG triplet was present in all cases, but no other sequence variations were found. Expression of the identified splice variants in normal tissues and paired cancer and adjacent tissues of 15 colon cancer patients was analysed by

RT-PCR with primers flanking exon 5. PCR products of 394 and 346 bp ($RHAMM^{FL}$ and $RHAMM^{-ex5}$, respectively) were obtained from spleen, colon, stomach, testis and all colon cancer, and the adjacent tissue specimens were analysed (Fig. 2A). Both isoforms were highly expressed at the same level in the testis; $RHAMM^{FL}$ predominated in the spleen, while the $RHAMM^{-ex5}$ signal was slightly stronger in the colon and stomach. The $RHAMM^{-ex5}$ isoform was dominant in the autologous cancerous tissue and it was overexpressed in comparison with that in tissue from normal colon. $RHAMM^{FL}$/$RHAM^{-ex5}$ ratios varied markedly among different patients and differed between paired cancerous and adjacent tissue samples (Fig. 2B).

Discussion

Our effort to define further the spectrum of immunogenic proteins in colon cancer has resulted in the identification of 8 distinct antigens. Four of them have been detected previously by SEREX analysis of various tumour types – C21ORF2 (Co2) in a colon cancer cell line, NAP1L1 (Co25) in breast and renal cell cancer, and golgins (Co26 and Co29) in gastric, breast and bronchial carcinomas and testis (SEREX database). However none of these antigens was detected by SEREX analysis of 4 colon cancer cases by Scanlan et al. [26], thus reflecting the heterogeneity of serological responses to colorectal tumours. To evaluate the possible reasons for the immunogenicity of the identified antigens, we characterised each antigen by cDNA sequence analysis, mRNA expression pattern and frequency of antibody responses in the sera from cancer patients and healthy donors.

With regard to serological responses, 3 antigens – C21ORF2, EPRS and NAP1L1 – showed cancer-associated reactivity and reacted preferentially with serum from cancer patients compared with normal human serum. Other antigens reacted with sera from both cancer patients and healthy individuals. Reactivity with the control sera could indicate that immune recognition of these antigens may not be cancer-dependent; however, it does not preclude the implication of the gene product in the development or progression of cancer, or its importance as a tumor antigen.

No sequence variations that could represent somatic mutations were detected by sequence analysis of cDNA inserts of the serum-reactive cDNA clones. However, it is also possible that sequence variations exist in the 5′ terminal regions that are absent in the clones, and that antibodies generated against altered protein cross-react with the wild-type counterpart. Analysis of full-length cDNA sequences showed that 2 genes – RHAMM and AD034 – are alternatively spliced in autologous tumour tissues. Deregulation of splice site selection may serve as an additional mechanism for the generation of protein diversity, thus contributing to the selection of more aggressive tumour cells [7, 22]. Moreover, ectopic expression of isoforms generated by altered splicing patterns can be recognised by the immune system and elicit antibody production.

Semi-quantitative RT-PCR was used to analyse the tissue distribution of the antigens and the optimal conditions for proportional production of PCR products with the number of PCR cycles was determined. Using the optimal PCR conditions, we observed that all of them were expressed in several normal human adult tissues, thus excluding them as differentiation or cancer-testis antigens. Comparison of mRNA levels between cancerous and adjacent non-cancerous tissues showed that at least 3 of the identified antigens appeared to be overexpressed gene products, and previous studies have suggested that overexpression of SEREX antigens can serve as immunogenic stimuli [21, 32]. The mechanism for induction of an immune response against overexpressed self-antigens remains largely unclear; it presumably involves breaking immunological tolerance by exceeding the thresholds of triggered T cell receptors required for T cell activation [16, 36].

We observed overexpression of NAP1L1 (nucleosome assembly protein 1-like 1) in 7 of 15 colorectal tumours and 1 of 7 gastric tumours. Relatively high expression of NAP1L1 has been previously shown in leukaemic cells and several cancer-derived cell lines. NAP1L1 is related to NAP-1 (nucleosome assembly protein-1) and has been proposed to be involved in nucleosome assembly and chromatin formation. Moreover, PMA/ionomycin stimulation causes a marked up-regulation of NAP1L1 in T cells, but treatment with NAP1L1 antisense oligonucleotides inhibits the induced T cell proliferation by about 50% [28]. Thus in theory, NAP1L1 could be considered as a proliferation marker of cancer cells and a possible therapeutic target.

Clone Co23 encodes a hypothetical protein AD034. Analysis of the predicted amino-acid sequence revealed a tyrosine kinase motif and a similarity to the RIO1/ZK632.3/MJ0444 family – evolutionary related uncharacterised proteins. We observed a relative up-regulation of AD034 mRNA expression in several colon cancer cases; however, the significance of AD034 expression in cancer development is not known. We also cloned a novel AD034 transcript variant, generated by use of a cryptic splice site. Translation of this transcript results in a truncated protein of 91 amino acids. However, RT-PCR analysis showed that the novel transcript variant represents less than 10% of the AD034 mRNA and is also detectable in several normal adult tissues including normal colon, thus showing that expression of this splice variant is not likely to be associated with immune recognition of AD034 in cancer patients. However, the biological role of this isoform remains to be investigated.

Another gene that we found to be up-regulated in cancerous tissues was RHAMM, a receptor for hyaluronic acid (HA) mediated motility. In a study by Yamada et al. [38], 1.05- to 13.6-fold increased RHAMM mRNA levels were observed in 29 out of 30 (97%) colon cancer specimens as compared with the adjacent tissues using real-time RT-PCR; this data supports the results

of the present study. Overexpression of RHAMM has been also observed in pancreatic cancer cell lines [1], primary gastric cancer [17] and various breast tumours, where its expression was elevated in particular tumour cell sub-populations and metastases, thus suggesting its role in motility and invasiveness of tumour cells [3]. RHAMM has been implicated in various physiological and pathological processes: HA-dependent cell motility, wound healing, cellular transformation and metastasis [9, 27]. However, its function still remains controversial. Initially RHAMM was described as a part of a cell surface-bound protein complex which upon binding of HA, triggers a signal transduction cascade leading to dissolution of focal adhesions and motility, and is also involved in cell cycle control [10, 34, 35, 37]. Later studies showed that RHAMM is an intracellular protein [5, 11] that interacts with microtubules and mitotic spindles [4]. Identification of several mRNA splice variants raised the possibility that various RHAMM isoforms may differ in their subcellular localisation and functional behaviour. In fact, both RHAMM isoforms differing in exon 5 have been shown to interact with mitotic spindles in dividing cells, but differ in their ability to bind to interphase microtubules. $RHAMM^{FL}$ co-localises with microtubules and is absent in the nucleus, while $RHAMM^{-ex5}$ does not bind to microtubules and is localised predominantly in the nucleus during interphase [4]. We observed tissue-specific differences in the expression of these splice variants, presumably showing that the efficiency of RHAMM splicing is regulated under physiological conditions or in a tissue-specific manner. $RHAMM^{FL}$ /$RHAMM^{-ex5}$ ratios varied significantly among different patients and were remarkably altered in the majority of tumour specimens as compared with the adjacent normal tissues. However, expression of the isoforms also differed in normal colon tissue specimens from different patients, and we do not know whether this reflects inter-individual differences, cell type-specific expression pattern differences, treatment or tumour-induced alterations in the regulation of the alternative splicing. It seems likely that alterations in the expression of RHAMM isoforms may influence cell morphology, motility and correlate with the metastatic behaviour of the tumour. However, this would necessitate confirmation by immunohistochemistry using RHAMM isoform-specific antibodies or by analysis of the RHAMM isoform expression in microdissected tumour and normal cell populations.

In conclusion, we have identified several new SEREX cancer genes, at least some of which are likely to have a functional role in cancer aetiology and could be considered as markers for colon cancer progression or as therapeutic targets. Nevertheless, the immunogenic nature of these antigens and consequently their potential as vaccine targets remains unclear, and requires further study. The relevance of cancer-associated autoantibodies to C21ORF2, EPRS and NAP1L1 in sero-diagnostics and as predictive or prognostic markers has yet to be established.

Acknowledgements This work was supported by grants from the Latvian Council of Science (2–2-011.1 and 01.0226), an UNESCO UVO-ROSTE grant, the John and Lucille van Geest Foundation and the Cancer and Polio Research Fund.

References

1. Abetamann V, Kern HF, Elsasser HP (1996) Differential expression of the hyaluronan receptors CD44 and RHAMM in human pancreatic cancer cells. Clin Cancer Res 2:1607
2. Ali SA, Lynam J, McLean CS, Entwise C, Loudon P, Rojas JM, McArdle SE, Li G, Mian S, Rees RC (2002) Tumor regression induced by intratumor therapy with a disabled infectious single cycle (DISC) herpes simplex virus (HSV) vector, DISC/HSV/murine granulocyte-macrophage colony-stimulating factor, correlates with antigen-specific adaptive immunity. J Immunol 168:3512
3. Assmann V, Gillett CE, Poulsom R, Ryder K, Hart IR, Hanby AM (2001) The pattern of expression of the microtubule-binding protein RHAMM/IHABP in mammary carcinoma suggests a role in the invasive behaviour of tumour cells. J Pathol 195:191
4. Assmann V, Jenkinson D, Marshall JF, Hart IR (1999) The intracellular hyaluronan receptor RHAMM/IHABP interacts with microtubules and actin filaments. J Cell Sci 112(22):3943
5. Assmann V, Marshall JF, Fieber C, Hofmann M, Hart IR (1998) The human hyaluronan receptor RHAMM is expressed as an intracellular protein in breast cancer cells. J Cell Sci 111(12):1685
6. Chen YT, Scanlan MJ, Sahin U, Tureci O, Gure AO, Tsang S, Williamson B, Stockert E, Pfreundschuh M, Old LJ (1997) A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening. Proc Natl Acad Sci USA 94:1914
7. Cooper TA, Mattox W (1997) The regulation of splice-site selection, and its role in human disease. Am J Hum Genet 61:259
8. Duval A, Rolland S, Compoint A, Tubacher E, Iacopetta B, Thomas G, Hamelin R (2001) Evolution of instability at coding and non-coding repeat sequences in human MSI-H colorectal cancers. Hum Mol Genet 10:513
9. Hall CL, Turley EA (1995) Hyaluronan: RHAMM mediated cell locomotion and signaling in tumorigenesis. J Neurooncol 26:221
10. Hardwick C, Hoare K, Owens R, Hohn HP, Hook M, Moore D, Cripps V, Austen L, Nance DM, Turley EA (1992) Molecular cloning of a novel hyaluronan receptor that mediates tumor cell motility. J Cell Biol 117:1343
11. Hofmann M, Fieber C, Assmann V, Gottlicher M, Sleeman J, Plug R, Howells N, von Stein O, Ponta H, Herrlich P (1998) Identification of IHABP, a 95-kDa intracellular hyaluronate binding protein. J Cell Sci 111(12):1673
12. Jager D, Stockert E, Scanlan MJ, Gure AO, Jager E, Knuth A, Old LJ, Chen YT (1999) Cancer-testis antigens and ING1 tumor suppressor gene product are breast cancer antigens: characterization of tissue-specific ING1 transcripts and a homologue gene. Cancer Res 59:6197
13. Jager E, Gnjatic S, Nagata Y, Stockert E, Jager D, Karbach J, Neumann A, Rieckenberg J, Chen YT, Ritter G, Hoffman E, Arand M, Old LJ, Knuth A (2000) Induction of primary NY-ESO-1 immunity: $CD8^+$ T lymphocyte and antibody responses in peptide-vaccinated patients with NY-ESO-1^+ cancers. Proc Natl Acad Sci USA 97:12198
14. Jager E, Nagata Y, Gnjatic S, Wada H, Stockert E, Karbach J, Dunbar PR, Lee SY, Jungbluth A, Jager D, Arand M, Ritter G, Cerundolo V, Dupont B, Chen YT, Old LJ, Knuth A (2000) Monitoring CD8 T cell responses to NY-ESO-1: correlation of humoral and cellular immune responses. Proc Natl Acad Sci USA 97:4760

15. Knights AJ, Zaniou A, Rees RC, Pawelec G, Müller L (2002) Prediction of an HLA-DR-binding peptide derived from Wilms' tumour 1 protein and demonstration of in vitro immunogenicity of WT1 (124–138)-pulsed dendritic cells generated according to an optimised protocol. Cancer Immunol Immunother 51:271
16. Lanzavecchia A, Lezzi G, Viola A (1999) From TCR engagement to T cell activation: a kinetic view of T cell behavior. Cell 96:1
17. Li H, Guo L, Li JW, Liu N, Qi R, Liu J (2000) Expression of hyaluronan receptors CD44 and RHAMM in stomach cancers: relevance with tumor progression. Int J Oncol 17:927
18. Line A, Stengrevics A, Slucka Z, Li G, Jankevics E, Rees RC (2002) Serological identification and expression analysis of gastric cancer-associated genes. Br J Cancer 86:1824
19. Mashino K, Sadanaga N, Tanaka F, Yamaguchi H, Nagashima H, Inoue H, Sugimachi K, Mori M (2001) Expression of multiple cancer-testis antigen genes in gastrointestinal and breast carcinomas. Br J Cancer 85:713
20. Mian S, Robins RA, Rees RC, Fox B (2001) Immunogenicity of tumour associated antigens: In: Robins RA, Rees RC (eds) Cancer immunology. Kluwer Academic Publishers, Dordrecht/Boston/London, p 1
21. Pfreundschuh M (2000) Exploitation of the B cell repertoire for the identification of human tumor antigens. Cancer Chemother Pharmacol 46:S3
22. Philips AV, Cooper TA (2000) RNA processing and human disease. Cell Mol Life Sci 57:235
23. Rees RC, Pawelec G, Muller L, Robins A (2002) Progress in vaccination against cancer (PIVAC) 2001, Robinson College, University of Cambridge, UK. Cancer Immunol Immunother 51:58
24. Sahin U, Li G, Tureci O, Pfreundschuh M (2001) Recognition of human tumors: SEREX expression cloning to identify tumour antigens: In: Robins RA, Rees RC (eds) Cancer immunology. Kluwer Academic Publishers, Dordrecht/Boston/London, p 45
25. Sahin U, Tureci O, Schmitt H, Cochlovius B, Johannes T, Schmits R, Stenner F, Luo G, Schobert I, Pfreundschuh M (1995) Human neoplasms elicit multiple specific immune responses in the autologous host. Proc Natl Acad Sci USA 92:11810
26. Scanlan MJ, Chen YT, Williamson B, Gure AO, Stockert E, Gordan JD, Tureci O, Sahin U, Pfreundschuh M, Old LJ (1998) Characterization of human colon cancer antigens recognized by autologous antibodies. Int J Cancer 76:652
27. Sherman L, Sleeman J, Herrlich P, Ponta H (1994) Hyaluronate receptors: key players in growth, differentiation, migration and tumor progression. Curr Opin Cell Biol 6:726
28. Simon HU, Mills GB, Kozlowski M, Hogg D, Branch D, Ishimi Y, Siminovitch KA (1994) Molecular characterization of hNRP, a cDNA encoding a human nucleosome-assembly-protein I-related gene product involved in the induction of cell proliferation. Biochem J 297(2):389
29. Stockert E, Jager E, Chen YT, Scanlan MJ, Gout I, Karbach J, Arand M, Knuth A, Old LJ (1998) A survey of the humoral immune response of cancer patients to a panel of human tumor antigens. J Exp Med 187:1349
30. Toh Y, Kuwano H, Mori M, Nicolson GL, Sugimachi K (1999) Overexpression of metastasis-associated MTA1 mRNA in invasive oesophageal carcinomas. Br J Cancer 79:1723
31. Toh Y, Oki E, Oda S, Tokunaga E, Ohno S, Maehara Y, Nicolson GL, Sugimachi K (1997) Overexpression of the MTA1 gene in gastrointestinal carcinomas: correlation with invasion and metastasis. Int J Cancer 74:459
32. Tureci O, Sahin U, Zwick C, Neumann F, Pfreundschuh M (1999) Exploitation of the antibody repertoire of cancer patients for the identification of human tumor antigens. Hybridoma 18:23
33. Tureci O, Schmitt H, Fadle N, Pfreundschuh M, Sahin U (1997) Molecular definition of a novel human galectin which is immunogenic in patients with Hodgkin's disease. J Biol Chem 272:6416
34. Turley EA, Austen L, Moore D, Hoare K (1993) *Ras*-transformed cells express both CD44 and RHAMM hyaluronan receptors: only RHAMM is essential for hyaluronan-promoted locomotion. Exp Cell Res 207:277
35. Turley EA, Austen L, Vandeligt K, Clary C (1991) Hyaluronan and a cell-associated hyaluronan binding protein regulate the locomotion of *ras*-transformed cells. J Cell Biol 112:1041
36. Viola A, Lanzavecchia A (1996) T cell activation determined by T cell receptor number and tunable thresholds. Science 273:104
37. Wang C, Thor AD, Moore DH, Zhao Y, Kerschmann R, Stern R, Watson PH, Turley EA (1998) The overexpression of RHAMM, a hyaluronan-binding protein that regulates *ras* signaling, correlates with overexpression of mitogen- activated protein kinase and is a significant parameter in breast cancer progression. Clin Cancer Res 4:567
38. Yamada Y, Itano N, Narimatsu H, Kudo T, Hirohashi S, Ochiai A, Niimi A, Ueda M, Kimata K (1999) Receptor for hyaluronan-mediated motility and CD44 expressions in colon cancer assessed by quantitative analysis using real-time reverse transcriptase-polymerase chain reaction. Jpn J Cancer Res 90:987