

# Serological identification and expression analysis of gastric cancer-associated genes

A Linē\*,<sup>1</sup>, A Stengrēvics<sup>2</sup>, Z Slucka<sup>1</sup>, G Li<sup>3</sup>, Ē Jankevics<sup>1</sup> and RC Rees<sup>3</sup>

<sup>1</sup>Biomedical Research and Study Centre, University of Latvia, 1 Ratsupites St, LV-1067, Riga, Latvia; <sup>2</sup>Latvian Oncology Center, 4 Hipokrata St, LV-1079, Riga, Latvia; <sup>3</sup>Department of Life Sciences, Nottingham Trent University, Clifton Lane, Nottingham NG11 8NS, UK

Serological identification of tumour antigens by recombinant expression cloning has proved to be an effective strategy for the identification of cancer-associated genes having a relevance to cancer aetiology and progression, and for defining possible targets for immunotherapeutic intervention. In the present study we applied this technique to identify immunogenic proteins for gastric cancer that resulted in isolation of 14 distinct serum-reactive antigens. In order to evaluate their role in tumorigenesis and assess the immunogenicity of the identified antigens, we characterised each cDNA clone by DNA sequence analysis, mRNA tissue distribution, comparison of mRNA levels in cancerous and adjacent non-cancerous tissues and the frequency of antibody responses in allogeneic patient and control sera. Previously unknown splice variants of TACCI and an uncharacterised gene Ga50 were identified. The expression of a newly identified TACCI isoform is restricted to brain and gastric cancer tissues. Comparison of mRNA levels by semi-quantitative RT-PCR revealed a relative overexpression of three genes in cancer tissues, including growth factor granulin and Tbdn-I – an orthologue of the mouse acetyltransferase gene which is associated with blood vessel development. An unusual DNA polymorphism – a three-nucleotide deletion was found in NUCB2 cDNA but its mRNA level was consistently decreased in gastric tumours compared with that in the adjacent non-cancerous tissues. This study has revealed several new gastric cancer candidate genes; additional studies are required to gain a deeper insight into their role in the tumorigenesis and their potential as therapeutic targets.

*British Journal of Cancer* (2002) **86**, 1824–1830. doi:10.1038/sj.bjc.6600321 www.bjcancer.com

© 2002 Cancer Research UK

**Keywords:** SEREX; tumour antigens; autoantibodies; TACCI; Tbdn-I; NUCB2

Cancer development is a multistep process, during which the cell acquires new phenotypic traits (overriding growth controls, induction of angiogenesis, evasion from host anti-tumour responses, extravasation and growth at metastatic sites etc.) as a result of successive genetic alterations (Hanahan and Weinberg, 2000). Identification of genetic/epigenetic alterations that contribute to the malignant phenotype capabilities is of major importance for the understanding the molecular basis of cancer, defining possible targets for therapeutic intervention, prognosis and diagnosis of human malignancies. Sahin *et al* (1995) introduced a novel approach – serological identification of tumour antigens by recombinant expression cloning – called SEREX, for the identification of tumour antigens recognised by cancer patient autoantibodies. This strategy is based on the construction of cDNA expression libraries from tumour specimens and immunoscreening of the libraries with cancer patient sera, thus allowing a systematic search for immunoreactive proteins (Sahin *et al*, 1995). SEREX has been applied to multiple human tumours, including melanoma (Jager *et al*, 2000a), renal cell cancer (Scanlan *et al*, 1999), astrocytoma (Sahin *et al*, 1995), breast (Jager *et al*, 1999) and colon (Scanlan *et al*, 1998), identifying tumour antigens for each tumour type, indicating that a humoral response is elicited in the majority of cancer patients.

SEREX-defined antigens represent a broad spectrum of structurally and functionally diverse proteins – transcription factors, adhesion molecules, signalling molecules, metabolic enzymes and cDNA sequences of 1549 antigens have been deposited in the SEREX database. Several categories of antigens have been detected by SEREX i.e. shared cancer-testis (C-T) antigens, differentiation antigens, mutated genes and translocation products, splice variant products, overexpressed antigens, viral antigens, cancer-related and cancer-independent autoantigens (Tureci *et al*, 1997; Sahin *et al*, 2001). Although the majority of the SEREX-identified genes have not been characterised beyond the preliminary sequence analysis, several of them have been proposed as attractive candidates for the construction of anti-cancer vaccines. For instance, ‘cancer-testis’ antigen NY-ESO-1 was identified by SEREX in oesophageal cancer and is regarded as one of the most immunogenic tumour antigens; antibody responses to NY-ESO-1 have been observed in 40–50% of patients with NY-ESO-1 expressing tumours and antibody production strongly correlates with CD8+ T cell responses in these patients (Jager *et al*, 2000c). Clinical trials, investigating the immunological effects of vaccination with NY-ESO-1 peptides, are ongoing (Jager *et al*, 2000b). Several previously known cytotoxic T cell targets, for example tyrosinase and MAGE antigens, have been detected by SEREX, supporting the premise of integrated CTL and B cell responses to tumour antigens. Moreover, antibody responses to mutated p53 (Scanlan *et al*, 1998), putative tumour suppressor ING1 (Jager *et al*, 1999) and amplified translation factor eIF-γ4 (Brass *et al*, 1997) have been detected by SEREX, thus demonstrating the potential

\*Correspondence: A Linē, E-mail: aija@biomed.lu.lv  
Received 23 January 2002; revised 21 March 2002; accepted 21 March 2002

of this technique for identification of genes that play a role in cancer aetiology and may serve as diagnostic markers or indicators of progression of the disease.

In the present study we applied SEREX to identify clinically relevant cancer-associated genes in human gastric carcinoma, and to define further the spectrum of immunogenic proteins in cancer. Fourteen 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

Gastric cancer and the adjacent non-cancerous tissue specimens were resected and snap frozen immediately after surgery. Tissue specimens and sera were obtained from 20 gastric cancer patients who had undergone surgical resection at the Latvian Oncology Center after the informed written consent was obtained (the study has been approved by the local ethical review board). In addition, serum samples were obtained from stomach, colon, breast and prostate cancer patients undergoing diagnostic procedures and from healthy volunteers.

### Construction of cDNA expression library

A cDNA expression library was constructed from a tumour specimen of a moderately differentiated, ulcerated gastric adenocarcinoma. Total RNA was isolated using Trizol reagent according to manufacturer's protocol (Life Technologies, Inc.). Poly(A)<sup>+</sup> RNA was purified from total RNA using Dynabeads mRNA Purification kit (DynaL AS, Norway) and cDNA was ligated into the lambda Uni-ZAP XR vector using Gigapack III Gold cloning kit (Stratagene GmbH). After *in vitro* packaging, library containing  $2 \times 10^6$  primary cDNA clones was obtained.

### Immunoscreening

Immunoscreening of cDNA library was performed as described by Sahin *et al* (1995). Briefly, after one round of amplification, the cDNA library was screened with 1:250 diluted autologous patient's serum and allogeneic sera, which had been previously preabsorbed with *E coli*-phage lysate. In order to eliminate cDNA clones encoding human IgG, nitrocellulose membranes containing phage plaques were pre-screened with AP-conjugated rabbit anti-human IgG secondary antibody (Pierce, USA) and reactive plaques were marked to exclude them from further study. Membranes were then incubated with patients' sera, and serum-reactive clones were detected with AP-conjugated secondary antibody and visualised by incubating with 5-bromo-4-chloro-3-indolyl-phosphate and nitroblue tetrazolium. The reactive phage clones were subcloned to monoclonality and converted to pBluescript phagemids.

### DNA sequencing and sequence analysis

Plasmid DNA was purified using QIAprep Spin Miniprep kit (QIAGEN GmbH), analysed by *EcoRI/XhoI* restriction enzyme digestion and clones representing different cDNA inserts were sequenced using BigDye Terminator Cycle Sequencing Ready Reaction kit on ABI PRISM 310 automatic sequencer (Applied Biosystems). Gene-specific primers were designed to obtain full insert sequences. cDNAs were identified by homology search through GenBank ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)). Multiple sequence alignments were performed with DNASIS (Hitachi Software Engineering Co Ltd) and MACAW (NCBI) software. Chromosomal localisation and exon-intron organisation for uncharacterised cDNAs was determined by comparison to the working draft of

the human genome. Putative protein domains were predicted by scanning the sequences against PROSITE and Pfam databases ([www.expasy.org](http://www.expasy.org) and [www.ludwig.unil.ch/SEREX](http://www.ludwig.unil.ch/SEREX)).

### Detection of antibodies in allogeneic sera

To assess frequencies of antibody responses to the SEREX-defined antigens in allogeneic sera, slightly modified immunoscreening procedure was used. *E coli* were transfected directly on gridded agar plate, by spotting 1  $\mu$ l of monoclonal positive phage ( $20-30$  pfu  $\mu$ l<sup>-1</sup>) 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.

### Comparative RT-PCR analysis

The mRNA expression pattern of SEREX-defined antigens was analysed by RT-PCR in a panel of normal tissue RNA from whole brain, liver, heart, kidney, lung, trachea (Clontech), spleen, colon, stomach, testes, ovary (Ambion), PBLs and a specimen of gastric cancer and adjacent tissues. Relative mRNA levels were compared between cancerous and adjacent non-cancerous tissues by semi-quantitative RT-PCR. Total RNA was isolated from paired tissue specimens using Trizol reagent according to manufacturer's protocol (Life Technologies, Inc.). The first-strand cDNA was synthesised from 4  $\mu$ g of total RNA primed with oligo-dT(18) and random hexamer primers using First-Strand cDNA Synthesis Kit (Fermentas, Lithuania). Gene specific PCR primers located within different exons were designed to amplify cDNA fragments (300-400 bp in length) of nine SEREX-defined genes and GAPDH,  $\beta$ -actin and histone H4 were used as internal standard genes. One fiftieth of RT mixture was amplified in GeneAmp PCR System 2400 thermal cycler (Perkin-Elmer Corp.) in a total reaction volume of 20  $\mu$ l containing 10 pmole of each primer, 200  $\mu$ M of dNTPs 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* (1997). Amplification of all target genes was performed simultaneously, at the same cycling conditions (1 min at 94°C, 30 s at 58°C, 45 s at 72°C), except for the number of cycles that was determined for each target gene (20 for  $\beta$ -actin, 25 for GAPDH, 27 for Zg4, 28 for Ga55 and Ga19, 29 for Ga27, Ga50, Zg2, histone H4 and 30 for Ga34, Zg14 and Zg15). Quantitative analysis of RT-PCR products was performed densitometrically after scanning the ethidium bromide stained gel on digital gel documentation and analysis system GDS8000 (Ultra-Violet Products Ltd, UK) and the intensities of bands were calculated using GelWorks software. Standard curves of amplification of each target gene were constructed from series of PCRs with ten 1.5-fold dilutions of normal stomach cDNA. Amounts of PCR products were linearly dependent on input cDNA over 10-fold dilutions of cDNA. The relative amounts of each mRNA were normalised to GAPDH and  $\beta$ -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 from each patient's tissue samples. Each reaction was performed in duplicate. Ratios  $\geq 2$  or  $\leq 0.5$  (the mean values of two independent experiments) were considered to represent significant alteration of expression levels.

### 5' and 3' RACE analysis

5' and 3' ends of full-length cDNAs were cloned from gastric cancer and adjacent tissues using FirstChoice™ RLM-RACE kit (Ambion) according to manufacturer's protocol. Briefly, for cloning of 5'-ends, 10  $\mu$ g of total RNA was treated with Calf Intestinal Phosphatase to remove 5'-phosphates from un-capped

RNAs, 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 reverse transcription and nested PCR with gene-specific and adapter-specific primers were performed and products were cloned using InsT/Aclone™ PCR Product Cloning Kit (Fermentas, Lithuania). For cloning of 3'ends, 1 µg of total RNA was reverse transcribed using an oligo (dT)-anchored 3'RACE adapter and nested PCR was performed.

## RESULTS

### Identification of serum-reactive cDNA clones

Eleven serum-reactive clones were identified by immunoscreening of approximately  $8 \times 10^5$  clones from gastric cancer cDNA expression library with autologous patient's serum. In addition  $\sim 3 \times 10^5$  phage plaques were screened using allogeneic serum from a patient with poorly differentiated infiltrative gastric adenocarcinoma, resulting in the isolation of nine positive clones. Full insert sequences of serum-reactive cDNA clones were obtained and identified by homology search through GenBank and SEREX databases. This revealed that serum-reactive clones are derived from 14 distinct genes, including eight genes whose function is unknown (listed in Table 1). Only two genes were detected by both autologous and allogeneic screening.

The frequency of antibodies from cancer patients and controls reacting with SEREX antigens was determined. Seven of the antigens had similar reactivity with serum from cancer patients and control individuals, three antigens reacted exclusively with autologous patient serum, while four antigens reacted with several cancer patient sera but not with sera from healthy individuals (Table 2).

### Analysis of sequence variations

Several sequence variations were found in three cancer-derived cDNAs in comparison with GenBank entries. An alteration of codon GAC (nt394) to GGC (Asp→Gly) in clone Ga34 (deposited in GenBank, AY039241) encoding Endomucin2 (NM\_016242) and three scattered nucleotide insertions within the 136 bp region of Zg14 (deposited in GenBank, AY039240) encoding for CGI-96 (AF151854), resulting in an altered 45 amino acid region in the predicted protein sequence were also observed. However sequen-

**Table 2** Reactivity of SEREX-defined antigens with sera from healthy donors and cancer patients

Clone	Healthy donor	Gastric Ca <sup>a</sup>	Colon Ca	Breast Ca	Prostrate Ca
Ga50	1/35	3/22	0/23	2/11	0/3
Ga52/Zg15	20/35	15/22	18/23	5/11	3/3
Ga55	0/35	1/22	0/23	0/11	0/3
Ga19	0/35	1/22	1/23	0/11	1/3
Ga27	0/35	1/22	0/23	0/11	0/3
Ga28/Zg3	4/35	3/22	1/23	3/11	2/3
Ga34	2/35	4/22	2/23	0/11	0/3
Ga44	0/35	3/22	0/23	1/11	0/3
Ga71	0/25	1/15	0/13	0/11	0/3
Zg2	1/25	2/22	0/23	0/11	0/3
Zg14	0/25	1/22	0/23	1/11	0/3
Zg4	0/25	2/22	0/23	0/11	0/3
Zg5	9/25	4/22	8/23	2/11	1/3
Zg10	3/25	4/22	2/23	2/11	0/3

<sup>a</sup>Includes the reactivity with sera used for immunoscreening of the cDNA expression library.

**Table 1** Genes identified by SEREX analysis of gastric cancer

Clone	Chr. location	Gene symbol <sup>a</sup>	NCBI reference sequence	Product, function, features, expression pattern
<i>Genes isolated by autologous screening</i>				
Ga19 <sup>b</sup>	4	<i>Tbdn-1</i>	NM_025085	Orthologous to mouse <i>tbdn-1</i> acetyltransferase which is involved in regulation of angiogenesis.
Ga27	17q21.32	<i>GRN</i>	NM_002087	Epithelial cell growth factor granulin. Upregulated in brain tumours and highly tumorigenic cell lines.
Ga28/Zg3	12	<i>KIAA0373</i>	NM_014684	Function unknown. Ubiquitously expressed <sup>c</sup> .
Ga34 <sup>b</sup>	4p15.33–q24	<i>Endomucin-2</i>	NM_016242	Function unknown. Trans-membrane helix.
Ga44	5.5pter–q11	<i>RARS</i>	NM_002887	Arginyl-tRNA synthetase. Aminoacylation of tRNA.
Ga50 <sup>b</sup>	12q12–12q14.3	<i>HSPC232</i>	NM_016488	Hypothetical protein. Function unknown. Multiple alternatively spliced isoforms.
Ga52/Zg15	15q22–q24	<i>KIAA1561</i>	AB046781	Uveal auto-antigen; similar to dog's C3VS protein, which is upregulated in thyroid by TSH stimulation.
Ga55 <sup>b</sup>	8p11	<i>TACCI</i>	NM_006283	Transforming, acidic coiled–coil containing protein 1. Involved in regulation of interaction between centrosomes and microtubules.
Ga71	3q28	<i>FXR1</i>	NM_005087	Fragile X mental retardation-related protein 1. Auto-antibodies are found in several auto-immune diseases.
<i>Genes isolated by allogeneic screening</i>				
Zg2	15	<i>KIAA0295</i>	AB002293	Function unknown. Zinc finger C2H2 motif. Ubiquitously expressed.
Zg4 <sup>b</sup>	11p15.1–p14	<i>NUCB2</i>	NM_005013	Ca <sup>2+</sup> binding protein. Contains DNA binding domain, helix–loop–helix motif, acidic amino acid region, leucine zipper and EF-hands.
Zg5	12	<i>KIAA1164</i>	AB032990	Partially similar to hypothetical protein KIAA1164. Function unknown. Ubiquitously expressed <sup>c</sup> .
Zg10	11p15	<i>SWAP-70</i>	AB014540	Involved in Ig heavy chain class switching. Expressed predominantly in B lymphocytes.
Zg14 <sup>b</sup>	22q13.2–13.31	<i>CGI-96</i>	NM_015703	Function unknown. RNA recognition motif. Ubiquitously expressed.

<sup>a</sup>HUGO approved gene symbols were used when available. <sup>b</sup>These cDNA sequences have been deposited in the GenBank (see text for accession numbers). <sup>c</sup>According to RT–PCR data at the Kazusa DNA Research Institute web-site ([www.kazusa.or.jp](http://www.kazusa.or.jp)).

cing of RT-PCR products derived from autologous normal gastric epithelium and tissue samples of two other patients revealed the same sequence variant for both genes, thus ruling out the possibility of somatic mutation in cancer cells.

Direct sequencing of NUCB2 cDNA derived from cancerous tissue of the patient whose serum was used for allogeneic screening showed heterogeneity of RT-PCR products at the 3' end of the coding region which is not present in the clone Zg4. The RT-PCR product was cloned into pTZ57R/T vector and two of five clones sequenced had ACA deletion (submitted to GenBank, AF450266). The 3-nucleotide deletion was also demonstrated by sequencing of amplified genomic DNA fragment. As the normal tissue sample from this patient was not available we could not directly determine whether this deletion represent somatic mutation or allelic polymorphism. Next we analysed DNA sequences from tumour and adjacent tissue samples of five gastric cancer patients, tumour and blood samples of five breast cancer patients and from blood samples of 25 donors. All paired tissue specimens had the same genotype thus showing that the deletion represents a germline DNA polymorphism and likely is not an immunogenic stimuli for production of anti-NUCB2 antibodies in the patients. Nevertheless, all gastric cancer patients and two breast cancer patients were heterozygous, one breast cancer patient had ACA deletion on both alleles, however five of 25 donors were also heterozygous at this locus, thus showing very high frequency of this unusual polymorphism not only among cancer patients but also in a general population.

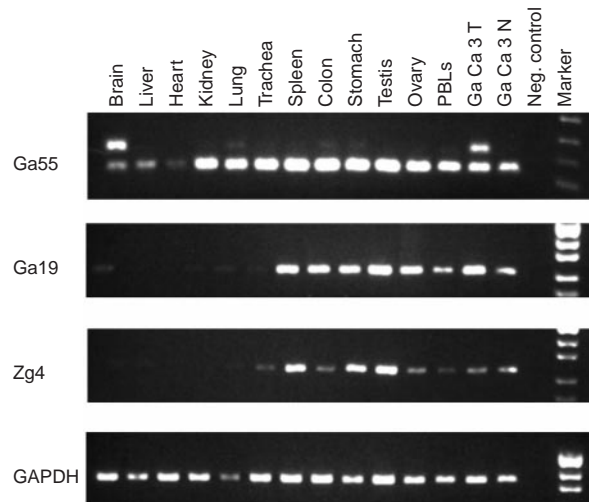
### Identification of TACC1 and HSPC232 splice variants

Comparison of Ga55 cDNA (submitted to GenBank, AY039239) with the published TACC1 sequence (GenBank, AF049910) showed that Ga55 represents a TACC1 isoform generated by inclusion of alternative 36 bp exon. The alternative exon is present in another GenBank sequence (KIAA1103, AB029026) derived from brain tissue, but is skipped in all other EST sequences, indicating that TACC1 pre-mRNA splicing may be regulated in tissue specific manner. The mRNA expression pattern of TACC1 isoforms was analysed by RT-PCR in a panel of RNAs derived from normal brain, liver, heart, kidney, lung, trachea, spleen, colon, stomach, testis, ovary, PBLs and in gastric cancer and adjacent non-cancerous tissue specimens using primers flanking the alternative exon. The transcript variant lacking the alternative exon was detected universally in all tissues tested. In contrast, the transcript containing the 36 bp exon was strongly expressed in brain and gastric cancer tissues; just trace amounts were detectable in lung and colon, but not in normal stomach and other normal tissues analysed (Figure 1). In subsequent analysis of paired gastric cancer and normal tissues, inclusion of the alternative exon was detected in three of 20 cancer tissue specimens but not in the adjacent tissues.

Clone Ga50 represents a novel splice variant of uncharacterised gene HSPC232. Alignment of GenBank cDNA entries and corresponding ESTs revealed at least four transcript variants differing in their 5' and 3' regions. RLM-RACE analysis showed that all splice variants are expressed in gastric cancer and adjacent tissues, and no additional variants were found. RT-PCR analysis showed that the splice variants are universally expressed in normal tissues but the efficiency of alternative splicing seems to be differentially regulated in different tissues. However no significant differences in expression of the splice variants in gastric cancer and adjacent tissues were found.

### mRNA expression pattern of SEREX-identified antigens

Multiple matching EST sequences were found in GenBank/EMBL databases for all of the identified antigens, however some genes shared sequence identity only to ESTs derived from cancerous

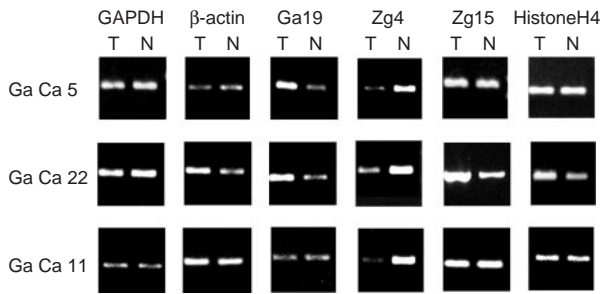


**Figure 1** Tissue distribution of Ga55 (*TACC1*), Ga19 (*Tbdn-1*), Zg4 (*NUCB2*) analysed by RT-PCR. Primers used were: Ga55 forward 5'-AC-CCAAGAAGGCAAAGTCGC-3' and Ga55 reverse 5'-AAATCTGGGATGATCACTGCC-3'. The primers are flanking the alternative exon – demonstrating the presence of the two splice variants, 119 and 155 bp. Ga19 forward primer 5'-CTAAAATCTATAAGCATGCTGG-3', Ga19 reverse 5'-ATTCATTGCTTTATAAGCCTGG-3', Zg4 forward 5'-TCAAGCAAGTGATTGATGTGC-3', Zg4 reverse 5'-TCAGGATTCAGGTGG-TTTAGG-3', Fragment of GAPDH cDNA was amplified as an internal control. 32, 29, 27, and 25 cycles of amplification were used for Ga55, Ga19, Zg4 and GAPDH, respectively. Ga Ca 3T and Ga Ca 3N – tumour and adjacent tissues from gastric cancer patient.

tissues. Therefore, we analysed the mRNA tissue distribution of selected antigens by RT-PCR in normal tissues (brain, liver, heart, trachea, lung, kidney, spleen, colon, stomach, testis, ovary and PBLs) and in gastric cancer and adjacent epithelium. Ga19, Zg15 and Zg4 mRNAs showed differential tissue distribution pattern. Relatively high expression of Ga19 and Zg15 was observed in normal testis, ovary, spleen, stomach, colon, and gastric cancer while it was weak or undetectable in other tissues tested. Zg4 was predominantly expressed in spleen, testis and normal stomach whereas RT-PCR signals were faint in all other tissues including cancerous and adjacent non-cancerous tissues from a gastric cancer patient (Figure 1).

### Comparison of mRNA levels in cancerous and adjacent normal tissues

In order to evaluate whether expression levels of genes encoding SEREX-antigens are altered in cancer tissues, we compared the relative mRNA levels in cancerous and paired non-cancerous tissues by semi-quantitative RT-PCR. Expression levels of nine genes isolated from gastric cancer cDNA library (Ga19, Ga27, Ga34, Ga50, Ga55, Zg2, Zg4, Zg14, and Zg15) were analysed in paired tissue samples from 20 gastric cancer patients. Relative mRNA levels of target genes were normalised to GAPDH and  $\beta$ -actin, and Tumour/Normal (T/N) ratios were calculated. Each T/N value represents the mean value of two independent experiments usually differing in less than 20%. T/N values 2 or  $\leq$ 0.5 were considered to represent significant difference. Histone H4 cDNA was amplified as a marker for the rate of proliferating cells in the tissue sample. Relative overexpression of three genes in various numbers of tumour specimens was observed. An example of the RT-PCR results is shown in Figure 2. We observed 2.0–6.5-fold increase in mRNA level of Ga19 in three of 20, 2.3–7.4-fold increase of Zg15 in three of 20 and 2.3- and 2.9-fold increase of Ga27 in two of nine gastric cancer specimens in comparison with



**Figure 2** Comparison of Ga19, Zg4 and Zg15 mRNA levels in three paired gastric cancer (T) and adjacent epithelium (N) samples by semi-quantitative RT–PCR. The same primers as shown in Figure 1 were used for amplification of Ga19 and Zg4. Primers used for Zg15 were: forward primer 5'-TCTGGAAGTATCTTTACCCAG-3', reverse 5'-GTTTGACGTGCTGAGCAAGC-3'. GAPDH and  $\beta$ -actin are used as internal standards. Histone H4 is used as a marker for the rate of cell proliferation.

matched normal tissues (when expression levels were normalised to  $\beta$ -actin). In the remaining cases, mRNA levels of these genes were similar in cancer and non-cancerous tissues. Interestingly, 2–11-fold downregulation of Zg4 was observed in 10 of 20 gastric cancer specimens. The major drawback for measurement of the relative abundance of cDNAs in clinical samples is the lack of an invariable internal standard. Expression levels of commonly used standards such as actins, tubulins and GAPDH may vary among different cell types (Weisinger *et al*, 1999) and in response to various factors such as hypoxia and, surgical resection (Thellin *et al*, 1999; Zhong and Simons, 1999) etc. For example, we observed apparently increased GAPDH mRNA levels in two tumour samples relative to  $\beta$ -actin and to target genes, resulting in down-estimation of the real expression levels of target genes when normalised to GAPDH.

### Sequence analysis of Ga19 cDNA

Sequence analysis of the Ga19 clone showed an uninterrupted ORF throughout the insert sequence indicating that the cDNA represented by this clone is truncated at both 5' and 3'-ends. To obtain the complete cDNA sequence RLM-RACE was performed. A single fragment of 930 bp was obtained in 5'RLM-RACE analysis and no variation of the transcription start site was found, but 3'RACE yielded in an 850 bp product. The assembled cDNA sequence contains 281 bp 5'UTR with several stop codons, 2595 bp coding sequence and 206 bp 3'UTR (deposited in GenBank, AY039242). Scanning of the predicted amino acid sequence against PROSITE and Pfam databases revealed N-terminal acetyltransferase domain (amino acids 250–602), bipartite nuclear localisation signal (aa 612–629) and TPR repeat region (aa 80–113; 46–181; 374–441).

### DISCUSSION

Definition of all immunogenic proteins in cancer ('cancer immunome') is regarded as the main long-term goal of SEREX (Jager *et al*, 1999; Pfreundschuh, 2000). Four of 14 antigens identified in the current study have been previously detected by SEREX in various tumour types (Table 3), and two of the identified genes – epithelial cell growth factor granulin (Lu and Serrero, 2000; Liau *et al*, 2000) and transforming coiled coil containing gene TACC1 (Still *et al*, 1999) have been previously shown to be associated with cancer. A further gene – FXR1 is autoantigen known to elicit antibody production in scleroderma patients (Bolivar *et al*, 1998). Another isolated gene is an Agr-tRNA synthetase. Autoantibodies against particular aminoacyl-tRNA synthetases are frequently found in myositis, interstitial lung

**Table 3** Antigens overlapping with previous SEREX analyses

TACC1	Stomach cancer (47)
SWAP70 (KIAA0640)	Renal cell carcinoma (699;338), Hodgkin's disease (765)
KIAA1561	Esophageal cancer (1558), breast carcinoma (1100), hepatocellular carcinoma
KIAA0373	Testis/CTCL sera (2019)

SEREX database ID is shown in brackets

disease, and arthritis (Hirakata *et al*, 1999) and various tRNA synthetases also have been detected by SEREX in renal (Scanlan *et al*, 1999), breast (Jager *et al*, 1999) and prostate cancer, and T cell leukemia (Itoh *et al*, 1999). No evidence for the implication of the genes in cancer or autoimmune states has been reported for the remaining antigens.

One possible reason for the immunogenicity of self-proteins is structural changes, resulting from mutations, translocations or the experiment of splice variants. Although we identified several sequence variations in three tumour-derived cDNAs (NUCB2, endomucin-2 and CGI-96), analysis of these sequences in tumour and normal tissue specimens indicates that these variations represent allelic polymorphisms and are not likely to be associated with the immunogenicity of these proteins. Nevertheless, three-nucleotide deletion within the coding region of NUCB2 that we detected with very high frequency among the cancer patients, is very interesting finding. As a result of the deletion Gln is omitted from the hydrophobic C-terminus of the protein. Moreover, comparison of mRNA levels between cancerous and adjacent non-cancerous tissues by RT–PCR showed that the expression of NUCB2 is downregulated in 50% of gastric tumours. NUCB2 is a  $Ca^{2+}$  binding protein, which also contains putative DNA binding domain, two helix–loop–helix motifs, EF-hands and leucine zipper (Barnikol-Watanabe *et al*, 1994; Kroll *et al*, 1999). It interacts with the postmitotic growth suppressor necdin and is proposed to be involved in the regulation of survival and death of postmitotic cells by controlling  $Ca^{2+}$  homeostasis in the cytoplasm (Taniguchi *et al*, 2000). Taken together, NUCB2 shows some characteristics of tumour suppressor gene but further studies investigating the functional significance of the 3-nucleotide deletion and downregulation of NUCB2 expression are required.

Alterations in the pattern and efficiency of alternative splicing of several pre-mRNAs (e.g. CD44, BRCA1, WT-1) have been implicated in tumorigenesis and correlate with tumour progression (Cooper and Dougherty, 1995; Liu *et al*, 2001). De-regulation of splice site selection may serve as an additional mechanism for the generation of protein diversity contributing to the selection of more aggressive tumour cells (Cooper and Mattox, 1997; Philips and Cooper, 2000). We identified several previously unknown splice variants of two genes – TACC1 represented by clone Ga55, and uncharacterised gene HSPC232 represented by clone Ga50. TACC1 is a recently identified gene, a member of TACC protein family that is involved in the regulation of interaction between centrosomes and microtubules (Gergely *et al*, 2000a). Overexpression of TACC1 in mouse fibroblasts results in cellular transformation and anchorage independent growth (Still *et al*, 1999). In *Drosophila* embryos, overexpression of TACC domain led to accumulation of microtubule asters and a complete failure in development of the embryos (Lee *et al*, 2001). In contrast, embryos carrying a mutation in *d-tacc* gene that decrease the level of D-TACC protein have shorter centrosomal microtubules and they also develop severe mitotic defects (Gergely *et al*, 2000b). These data show that perturbation of TACC gene expression may provoke various effects ranging from a genetic instability and missegregation of chromosomes to cell death. In the current study we identified a novel TACC1 isoform generated by inclusion of an alternative 36 bp exon. This isoform was strongly expressed in

three of 20 specimens of gastric cancer tissues while it was undetectable in the adjacent gastric epithelium and normal stomach epithelium. From normal adult tissues it was predominately expressed in brain. We propose that alteration in the pattern of TACC1 mRNA splicing in cancer cells might lead to perturbation of TACC1 function. Whether the ectopic expression of this isoform served as a stimulus for antibody production in the patient also remains to be established. Interestingly, TACC1 and TACC2 have also been detected by SEREX in gastric cancer by Yuichi Obata (SEREX database).

The largest category of SEREX antigens occur as overexpressed gene products. Gene amplification, elevated transcription and stability of mRNA or protein may lead to accumulation of protein and induction of immune response by exceeding the number of triggered TCRs required for T cell activation (Viola and Lanzavecchia, 1996; Lanzavecchia *et al*, 1999). In order to search for overexpressed antigens, we used semi-quantitative RT-PCR to examine the mRNA levels of the identified antigens in gastric cancer and adjacent non-cancerous tissues. Relative overexpression of Ga27 (granulin), Zg15 (orthologous of dog's C3VS gene) and Ga19 (orthologous to mouse Tbdn-1 gene) in gastric cancer was observed. Upregulation of the growth factor granulin has been shown in brain tumours (Liau *et al*, 2000) and in highly tumorigenic cell lines (Zhang and Serrero, 1998). Granulin has been shown to have mitogenic activity in fibroblasts, epithelial cells and glioblastoma cell lines (Xu *et al*, 1998; Liau *et al*, 2000; Lu and Serrero, 2000). Zg15 represents a human orthologue of dog's C3VS gene, its function is unknown but it is upregulated by mitogenic stimulation in thyroid cells (Wilkin *et al*, 1996). Thus, increased proliferation of tumour cells alone could account for the overexpression of Zg15 in cancerous tissues. In theory, it could be considered as a proliferation marker and its prognostic significance remains to be determined. Another gene that we found to be upregulated in three of 20 gastric cancer specimens is human orthologue of the mouse tbdn-1 gene encoding putative acetyltransferase. The levels and tissue distribution of tbdn-1 suggest that it may be involved in regulation of vascular and hematopoietic development and angiogenesis. Tbdn-1 is expressed at high levels at day 8 of gestation, when yolk sac vasculogenesis and blood island

formation is peaking but downregulated at day 10 when formation of larger coalesced vitelline vessels occurs. In adult mouse it is expressed in atrial endocardium, bone marrow and ovarian follicles only (Gendron *et al*, 2000). Although we observed relatively high expression of Tbdn-1 in spleen, colon, stomach, testis and ovary, it could be of interest to investigate which cell populations within a tumour overexpress it and whether this acetyltransferase has a role in tumour neovascularisation.

In conclusion, a number of the antigens that we have identified in this study may have a functional role in cancer aetiology or progression, and the altered expression pattern of these genes may provide the immunogenic stimuli in these patients. However other possibilities, for example, unidentified mutations or altered sub-cellular localisation which have been recently shown for  $\beta$ -actin in apoptotic breast cancer cells (Hansen *et al*, 2001), cannot be excluded. Further studies will be focused on exploration of the possible functional differences between TACC1 isoforms, definition of cell populations that express Tbdn-1 acetyltransferase and investigation of its possible role in tumour angiogenesis. It will also be of interest to elucidate the possible role of NUCB2 in tumour progression and the functional significance of the peculiar three-nucleotide polymorphism. In addition, serological responses to seven antigens (including TACC1, NUCB2, Tbdn-1 and granulin etc.) were restricted to cancer patients and investigation of their relevance for sero-diagnosis or prognosis is in progress.

## ACKNOWLEDGEMENTS

We thank Dr O Tureci and Dr U Sahin for helping us to establish the SEREX technique in our laboratories; Prof E Grens and Dr S Mian for motivating discussions. This work was supported by grants from Latvian Council of Science (No.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. Data deposition: cDNA sequences reported in this study have been deposited in the GenBank (accession numbers AY039238–AY039242 and AF450266).

## REFERENCES

- Barnikol-Watanabe S, Gross NA, Gotz H, Henkel T, Karabinos A, Kratzin H, Barnikol HU, Hilschmann N (1994) Human protein NEFA, a novel DNA binding/EF-hand/leucine zipper protein. Molecular cloning and sequence analysis of the cDNA, isolation and characterization of the protein. *Biol Chem Hoppe Seyler* **375**: 497–512
- Bolivar J, Guelman S, Iglesias C, Ortiz M, Valdivia MM (1998) The fragile-X-related gene FXR1 is a human autoantigen processed during apoptosis. *J Biol Chem* **273**: 17122–17127
- Brass N, Heckel D, Sahin U, Pfreundschuh M, Sybrecht GW, Meese E (1997) Translation initiation factor eIF-4gamma is encoded by an amplified gene and induces an immune response in squamous cell lung carcinoma. *Hum Mol Genet* **6**: 33–39
- Cooper DL, Dougherty GJ (1995) To metastasize or not? Selection of CD44 splice sites. *Nat Med* **1**: 635–637
- Cooper TA, Mattox W (1997) The regulation of splice-site selection, and its role in human disease. *Am J Hum Genet* **61**: 259–266
- Gendron RL, Adams LC, Paradis H (2000) Tubedown-1, a novel acetyltransferase associated with blood vessel development. *Dev Dyn* **218**: 300–315
- Gergely F, Karlsson C, Still I, Cowell J, Kilmartin J, Raff JW (2000a) The TACC domain identifies a family of centrosomal proteins that can interact with microtubules. *Proc Natl Acad Sci USA* **97**: 14352–14357
- Gergely F, Kidd D, Jeffers K, Wakefield JG, Raff JW (2000b) D-TACC: a novel centrosomal protein required for normal spindle function in the early *Drosophila* embryo. *EMBO J* **19**: 241–252
- Hanahan D, Weinberg RA (2000) The hallmarks of cancer. *Cell* **100**: 57–70
- Hansen MH, Nielsen H, Ditzel HJ (2001) The tumor-infiltrating B cell response in medullary breast cancer is oligoclonal and directed against the autoantigen actin exposed on the surface of apoptotic cancer cells. *Proc Natl Acad Sci USA* **98**: 12659–12664
- Hirakata M, Suwa A, Nagai S, Kron MA, Trieu EP, Mimori T, Akizuki M, Targoff IN (1999) Anti-KS: identification of autoantibodies to asparaginyl-transfer RNA synthetase associated with interstitial lung disease. *J Immunol* **162**: 2315–2320
- Itoh M, Watanabe M, Yamada Y, Furukawa K, Taniguchi M, Hata T, Schmitt M, Ikeda H, Yamaguchi M, Ohno T, Nakashima K, Shiku H (1999) HUB1 is an autoantigen frequently eliciting humoral immune response in patients with adult T cell leukemia. *Int J Oncol* **14**: 703–708
- Jager D, Stockert E, Jager E, Gure AO, Scanlan MJ, Knuth A, Old LJ, Chen YT (2000a) Serological cloning of a melanocyte rab guanosine 5'-triphosphate-binding protein and a chromosome condensation protein from a melanoma complementary DNA library. *Cancer Res* **60**: 3584–3591
- 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**: 6197–6204
- Jager E, Gnjatich 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 (2000b) 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**: 12198–12203

- 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 (2000c) Monitoring CD8 T cell responses to NY-ESO-1: correlation of humoral and cellular immune responses. *Proc Natl Acad Sci USA* **97**: 4760–4765
- Kroll KA, Otte S, Hirschfeld G, Barnikol-Watanabe S, Gotz H, Sternbach H, Kratzin HD, Barnikol HU, Hilschmann N (1999) Heterologous overexpression of human NEFA and studies on the two EF-hand calcium-binding sites. *Biochem Biophys Res Commun* **260**: 1–8
- Lanzavecchia A, Lezzi G, Viola A (1999) From TCR engagement to T cell activation: a kinetic view of T cell behavior. *Cell* **96**: 1–4
- Lee MJ, Gergely F, Jeffers K, Peak-Chew SY, Raff JW (2001) Msp/XMAP215 interacts with the centrosomal protein D-TACC to regulate microtubule behaviour. *Nat Cell Biol* **3**: 643–649
- Liau LM, Lallone RL, Seitz RS, Buznikov A, Gregg JP, Kornblum HI, Nelson SF, Bronstein JM (2000) Identification of a human glioma-associated growth factor gene, granulin, using differential immuno-absorption. *Cancer Res* **60**: 1353–1360
- Liu HX, Cartegni L, Zhang MQ, Krainer AR (2001) A mechanism for exon skipping caused by nonsense or missense mutations in BRCA1 and other genes. *Nat Genet* **27**: 55–58
- Lu R, Ferrero G (2000) Inhibition of PC cell-derived growth factor (PCDGF, epithelin/granulin precursor) expression by antisense PCDGF cDNA transfection inhibits tumorigenicity of the human breast carcinoma cell line MDA-MB-468. *Proc Natl Acad Sci USA* **97**: 3993–3998
- Pfreundschuh M (2000) Exploitation of the B cell repertoire for the identification of human tumor antigens. *Cancer Chemother Pharmacol* **46**(Suppl): S3–S7
- Philips AV, Cooper TA (2000) RNA processing and human disease. *Cell Mol Life Sci* **57**: 235–249
- Sahin U, Li G, Tureci O, Pfreundschuh M (2001) Recognition of human tumors: SEREX expression cloning to identify tumour antigens. In *Cancer Immunology*, Robins RA, Rees RC (eds) pp 45–57. Dordrecht/Boston/London: Kluwer Academic Publishers
- Sahin U, Tureci O, Schmitt H, Cochlovius B, Johannes T, Schmits R, Stenner F, Luo G, Schober I, Pfreundschuh M (1995) Human neoplasms elicit multiple specific immune responses in the autologous host. *Proc Natl Acad Sci USA* **92**: 11810–11813
- 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**: 652–658
- Scanlan MJ, Gordan JD, Williamson B, Stockert E, Bander NH, Jongeneel V, Gure AO, Jager D, Jager E, Knuth A, Chen YT, Old LJ (1999) Antigens recognized by autologous antibody in patients with renal-cell carcinoma. *Int J Cancer* **83**: 456–464
- Still IH, Hamilton M, Vince P, Wolfman A, Cowell JK (1999) Cloning of TACC1, an embryonically expressed, potentially transforming coiled coil containing gene, from the 8p11 breast cancer amplicon. *Oncogene* **18**: 4032–4038
- Taniguchi N, Taniura H, Niinobe M, Takayama C, Tomimaga-Yoshino K, Ogura A, Yoshikawa K (2000) The postmitotic growth suppressor necdin interacts with a calcium-binding protein (NEFA) in neuronal cytoplasm. *J Biol Chem* **275**: 31674–31681
- Theillin O, Zorzi W, Lakaye B, De Borman B, Coumans B, Hennen G, Grisar T, Igout A, Heinen E (1999) Housekeeping genes as internal standards: use and limits. *J Biotechnol* **75**: 291–295
- 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**: 459–463
- Tureci O, Sahin U, Pfreundschuh M (1997) Serological analysis of human tumor antigens: molecular definition and implications. *Mol Med Today* **3**: 342–349
- Viola A, Lanzavecchia A (1996) T cell activation determined by T cell receptor number and tunable thresholds. *Science* **273**: 104–106
- Weisinger G, Gavish M, Mazurika C, Zinder O (1999) Transcription of actin, cyclophilin and glyceraldehyde phosphate dehydrogenase genes: tissue- and treatment-specificity. *Biochim Biophys Acta* **1446**: 225–232
- Wilkin F, Savonet V, Radulescu A, Petermans J, Dumont JE, Maenhaut C (1996) Identification and characterization of novel genes modulated in the thyroid of dogs treated with methimazole and propylthiouracil. *J Biol Chem* **271**: 28451–28457
- Xu SQ, Tang D, Chamberlain S, Pronk G, Masiarz FR, Kaur S, Prisco M, Zanocco-Marani T, Baserga R (1998) The granulin/epithelin precursor abrogates the requirement for the insulin-like growth factor 1 receptor for growth in vitro. *J Biol Chem* **273**: 20078–20083
- Zhang H, Ferrero G (1998) Inhibition of tumorigenicity of the teratoma PC cell line by transfection with antisense cDNA for PC cell-derived growth factor (PCDGF, epithelin/granulin precursor). *Proc Natl Acad Sci USA* **95**: 14202–14207
- Zhong H, Simons JW (1999) Direct comparison of GAPDH, beta-actin, cyclophilin, and 28S rRNA as internal standards for quantifying RNA levels under hypoxia. *Biochem Biophys Res Commun* **259**: 523–526