A comparative analysis by SAGE of gene expression profiles of esophageal adenocarcinoma and esophageal squamous cell carcinoma

Jantine W.P.M. van Baal [a,*], Francesco Milana [a], Agnieszka M. Rygiel [a], Carine M.T. Sondermeijer [b], C. Arnold Spek [a], Jacques J.G.H.M. Bergman [b], Maikel P. Peppelenbosch [c] and Kausilia K. Krishnadath [b,*]

[a] Center for Experimental and Molecular Medicine, Academic Medical Center, Amsterdam, The Netherlands
[b] Dept. of Gastroenterology and Hepatology, Academic Medical Center, Amsterdam, The Netherlands
[c] Dept. of Cell Biology, University of Groningen, The Netherlands

Abstract. Esophageal adenocarcinoma (EA) and esophageal squamous cell carcinoma (ESCC) are the two main types of esophageal cancer. Despite extensive research the exact molecular basis of these cancers is unclear. Therefore we evaluated the transcriptome of EA in comparison to non-dysplastic Barrett's esophagus (BE), the metaplastic epithelium that predisposes for EA, and compared the transcriptome of ESCC to normal esophageal squamous epithelium. For obtaining the transcriptomes tissue biopsies were used and serial analysis of gene expression (SAGE) was applied. Validation of results by RT-PCR and immunoblotting was performed using tissues of an additional 23 EA and ESCC patients. Over 58,000 tags were sequenced. Between EA and BE 1013, and between ESCC and normal squamous epithelium 1235 tags were significantly differentially expressed ($p < 0.05$). The most up-regulated genes in EA compared to BE were SRY-box 4 and Lipocalin2, whereas the most down-regulated genes in EA were Trefoil factors and Annexin A10. The most up-regulated genes in ESCC compared to normal squamous epithelium were BMP4, E-Cadherin and TFF3. The results could suggest that the BE expression profile is closer related to normal squamous esophagus then to EA. In addition, several uniquely expressed genes are identified.

Keywords: Barrett's esophagus, SAGE, gene expression profile, esophageal squamous cell carcinoma, esophageal adenocarcinoma

Abbreviations

SAGE	serial analysis of gene expression;
EA	esophageal adenocarcinoma;
ESCC	esophageal squamous cell carcinoma;
BE	Barrett's esophagus;
RT-PCR	reverse-transcription polymerase chain reaction;
ESTs	expressed sequence tags;
SOX4	SRY-box 4;
TFF	trefoil factor;
FABP	fatty acid binding protein;
SBP	selenium binding protein.

*Corresponding authors: J.W.P.M. van Baal, Center for Experimental and Molecular Medicine, Academic Medical Center, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands. E-mail: jvanbaal@hotmail.com; K.K. Krishnadath, Dept. of Gastroenterology and Hepatology, Academic Medical Center, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands. Tel.: +31 20 5663632; Fax: +31 20 6977192; E-mail: K.K.Krishnadath@amc.uva.nl.

1. Introduction

Two main types of esophageal cancer are Esophageal Adenocarcinoma (EA) and Esophageal Squamous Cell Carcinoma (ESCC). EA is associated with Barrett's esophagus (BE), a metaplastic condition of the distal esophagus, in which through longstanding gastro-esophageal reflux disease, the normal squamous epithelium is replaced by columnar epithelium [1,2]. Malignant degeneration of BE is thought to be a multi-

step process in which metaplasia progresses through low grade and high grade dysplasia into an invasive adenocarcinoma [3]. BE patients have an increased risk of developing EA, with an estimated annual incidence varying from 0.4% to 1.8% [4–7]. Over the last 3 decades, the incidence of BE associated adenocarcinoma has increased in Western countries at a rate that exceeds that of any other malignancy [8–10]. Abnormalities in oncogenes, tumor suppressor genes and growth factors play an important role in the development of EA. These factors have an influence on cell cycle progression and are critical in malignant transformation. For instance mutation of p16 and p53 can be found in EAs [11–13].

ESCC develops in a multi-step, progressive process, as the result of a sequence of histopathological changes that typically involves esophagitis, atrophy, mild to severe dysplasia, carcinoma in situ and finally, invasive cancer. Worldwide ESCC is the predominant esophageal cancer and has a high mortality rate [14]. Several genetic changes are associated with the development of ESCC including mutations of the p53 gene, activation of oncogenes like EGFR and c-MYC, inactivation of several tumor suppressor genes and disruption of cell-cycle control in G1. The G1 phase of the cell cycle is controlled by several mechanisms that are disrupted in ESCC, like inactivation of p16, amplification of Cyclin D1 and alterations of the Retinoblastoma gene [15].

Although both types of esophageal cancers have comparable clinical outcomes characterized by early metastasis and poor patient prognosis, the pathophysiology of these cancers seems to be different. For understanding and optimizing future treatments it is important to understand the specific biology of these cancers. We hypothesized that by generating large molecular data sets of EA and ESCC and by quantitatively comparing these with non-dysplastic BE and normal squamous epithelium, respectively, we will accurately classify the different phenotypes of these carcinomas and improve our insight in the biological pathways and mechanisms involved in these malignancies. In this study gene expression profiles were obtained using the technique serial analysis of gene expression (SAGE). The method as first described by Velculescu et al. allows rapid, quantitative and simultaneous analysis of thousands of genetic transcripts from tissue samples [16]. It is based on two principles: a nucleotide sequence of 10 bp, a tag, is produced. The location of the tags within the transcripts is precisely defined; therefore these tags contain adequate infor-

mation to identify transcripts using public databases (SAGEgenie, http://cgap.nci.nih.gov). Additionally, by sequencing, a large amount of transcripts can be identified efficiently, because the tags are serially cloned along with a restriction enzyme recognition sequence that serves as an anchor. This reveals the identity of thousands of tags and at the same time it quantifies their level of expression. In this study transcriptomes were made of RNA isolated from an EA and ESCC biopsy with totally over a 58,000 tags. These transcriptomes were analyzed and compared with already known transcriptomes of BE and normal squamous epithelium [17]. A panel of biopsies obtained from another 10 EA patients and 13 ESCC patients was used for validation by Reverse-Transcription Polymerase Chain Reaction (RT-PCR) and immunoblotting.

2. Materials and methods

2.1. Patients and biopsy specimens

The study was approved by the hospital's medical ethical committee and all patients signed informed consent for the use of their biopsy material. Tissue samples were obtained during endoscopy of 11 patients with EA associated to BE (8 were male; mean age was 66 years, range 49–83 years). Four patients had T1N0M0 stage of cancer, 2 patients T2N0M0, 4 patients T3N0M0 and 1 patient T3N1M0. Additionally, endoscopic biopsies were obtained from 14 patients with known ESCC (8 were male; mean age 62 years, range 50–80 years). Seven patients had T3N1M0 stage of cancer, 2 patients T3N1M1b, 3 patients T4N1M0 and 2 patients T4N1M1b. None of the patients had received radiotherapy or chemotherapy prior to the endoscopy. Paired biopsies, taken immediately adjacent to each other, were obtained from BE and the cancerous lesion. Histological examination of the pair wise taken control biopsies confirmed presence of BE, EA or ESCC.

For SAGE analysis, RNA from an EA patient with T2N0M0 stage of cancer and from an ESCC patient with T3N1M0 stage of cancer, respectively, was used.

Tissue samples were obtained during routine surveillance endoscopy of another 16 patients known with intestinal type of metaplasia (BE; 13 were male; mean age was 62 years, range 41–83 years; average length of BE segment 3.6 cm, range 2–9 cm). All patients were

on long term proton pump inhibition of 40 to 80 mg daily to prevent active inflammation due to acid reflux. BE was defined as histologically recognized incompletely differentiated intestinal type of metaplasia in the distal esophagus. Paired biopsies, taken immediately adjacent to each other, were obtained from the Barrett's segment and normal squamous esophagus. The Barrett's segment was biopsied at least 2 cm above the gastroesophageal junction yet within the Barrett's segment, recognized endoscopically as typically pink colored columnar type of metaplasia. Normal squamous epithelium was biopsied at least 2 cm above the Barrett's segment. Endoscopically, none of the patients showed signs of erosive reflux. All patients had proven incompletely differentiated intestinal type of columnar epithelium without dysplasia in the histological control biopsies with no signs of active inflammation. Normal squamous esophagus epithelium was also confirmed histologically, in all the pair wise taken control biopsies.

2.2. RNA isolation

Total RNA was isolated from biopsies using Trizol Reagent (Life Technologies Inc, Invitrogen, Breda, The Netherlands) according to manufacturer's instructions. In brief: tissues were lysed by adding 200 µl Trizol. After phenol/chloroform extraction, RNA was precipitated with isopropanol, washed with 70% ethanol and air-dried. The RNA was then dissolved in RNase-free H_2O and stored at $-80°C$ until required. Spectrophotometry was performed with 1 µl of total RNA to determine the concentration on the Nanodrop® (type ND-1000, Wilmington, USA).

2.3. SAGE procedure

Two SAGE libraries were obtained, essentially following the SAGE protocol as described by Velculescu et al. using the Life Technologies I-SAGE kit and following manufacturer's instructions [16,17]. 5 µg of total RNA was used per SAGE analysis. The isolated concatemers, consisting of serially ligated tags, were ligated into the pZErO-1 vector (Life Technologies) and transformed in TOP10 Electrocompetent *Escherichia coli* cells (Life Technologies) by means of electroporation, following manufacturer's protocol (Biorad, Hercules, CA). PCR was performed on obtained colonies with specific primers Sp6-F (5′-GATTTAGGTGACACTATAG-3′) and T7A-R (5′-

TAATACGACTCACTATAGGG-3′) and PCR products were analyzed by agarose gel electrophoresis. A total of 1920 clones were selected for DNA sequencing using the Big Dye Terminator Kit (Applied Biosystems, Foster City, CA) and the T7A-R primers. Samples were run on an ABI3730 DNA Analyzer (Applied Biosystems) and analyzed with Sequence Analysis 5.1 software.

2.4. SAGE and statistical analysis

Statistical analysis of SAGE data was performed according to Van Baal et al. [17]. Briefly, for analysis of the SAGE data the program USAGE V2 (Academic Medical Center, bioinformatics department) and the public databases of the NCBI-site and SAGE Genie (http://cgap.nci.nih.gov) were used [18,19]. Statistical analyses and comparison of the SAGE libraries was done using a comparative Z-test (Pair-wise comparison, binominal approach) of the USAGE V2 program [20,21].

2.5. Reverse-transcription polymerase chain reaction

cDNAs from biopsies were synthesized from 1 µg of total RNA using an oligo dT primer and Superscript II MMLV-reverse transcriptase according to manufacturer's instructions (Life Technologies). Primers for selected genes (Table 1) were derived from mRNA sequences as deposited in GenBank (NCBI-site). Subsequent PCR analyses were carried out in 25 µl reactions containing 1 µl cDNA, 23 µl Reddy Mix PCR Master Mix (Applied Biosystems), 200 ng Forward primer and 200 ng Reverse primer. For each gene investigated, the number of cycles used, were confirmed to be in the linear region of the amplification curve. The products were electrophoresed on 1% agarose gel.

2.6. Immunoblotting

Immunoblotting was performed as described by Hardwick et al. [22]. Biopsies were lysed with 200 µl lysis buffer. Twenty µg of protein per lane was loaded onto SDS-PAGE. The blots were blocked with 2% BSA in Tris Buffered Saline supplemented with 0.1% Tween-20. The antibodies used and dilutions are summarized in Table 2.

Table 1

Primer sequences

Gene	Forward primer	Reverse primer	Annealing temperature (°C)	Fragment length (bp)
TFF1*	TTTGGAGCAGAGAGGAGG	TTGAGTAGTCAAAGTCAGAGCAG	60	438
TFF2	ATGGATGCTGTTTCGACTCC	GGCACTTCAAAGATGAAGTTG	55	247
TFF3	GTGCCAGCCAAGGACAG	CGTTAAGACATCAGGCTCCAG	58	303
Annexin A10	TTGTTCTCTGTGTTCGAGACAAACC	GTAGGCAAATTCAGGATAGTAGGC	52	609
SOX4**	CTTGACATGATTAGCTGGCATGATT	CCTGTGCAATATGCCGTGTAGA	64	100
FABP1***	TCATGAAGGCAATCGGTCTG	GTGATTATGTCGTCGCCGTTGAGT	55	277
BMP4	ACCTGAGACGGGGAAGAAAA	TTAAAGAGGAAACGAAAAGCA	62	348
Plakophilin 3	AGCCTGGAGGAGAAGGCTAAT	AGTGCTGGCTATCCCAAGATACT	60	234
Prosaposin	CCAGAGCTGGACATGACTGA	CAGTTCCCAACAAGGGCTTA	60	999
SBP1****	TCAGATGATCCAGCTCAGCCT	TCACAGCCTTCCCTGATGA	60	109
E-Cadherin	GACGCGGACGATGATGTGAAC	TTGTACGTGGTGGGATTGAAGA	56	280
Lipocalin 2	GGAGCTGACTTCGGAACTAAAGG	AGCCGTCGATACACTGGTCG	60	109
β-actin	GTCAGAAGGATTCCTATGTGG	GCTCATTGCCAATGGTGATG	52	628
β-2-microglobulin	CTCGCGCTACTCTCTCTTTCT	TGCTCCACTTTTTCAATTCTCT	60	185

Primer sequences used for RT-PCR with corresponding annealing temperatures and PCR fragment lengths. *TFF = Trefoil Factor; **SOX4 = SRY box 4; ***FABP1 = Fatty Acid Binding Protein 1; ****SBP1 = Selenium Binding Protein 1.

Table 2

Antibodies as used for immunoblot analysis

Antibody	Species	Company	Country	Dilution
Cytokeratin 5/6	Mouse monoclonal	Chemicon	USA	1:500
Cytokeratin 8	Mouse monoclonal	Chemicon	USA	1:500
Cytokeratin 10/13	Mouse monoclonal	Dako	Denmark	1:500
Cytokeratin 20	Mouse monoclonal	Progen	Germany	1:500
PKC β1	Rabbit polyclonal	Santa Cruz	Germany	1:500
BMP 4	Mouse monoclonal	R&D	United Kingdom	1:500
ID 2	Rabbit polyclonal	Santa Cruz	Germany	1:1000
Cyclin D1	Mouse monoclonal	Neomarkers	USA	1:1000
TGF-β	Rabbit polyclonal	Santa Cruz	Germany	1:500
EGF-receptor	Rabbit polyclonal	Cell Signaling	USA	1:500
p 19	Mouse monoclonal	Neomarkers	USA	1:1000
p 27	Mouse monoclonal	Santa Cruz	Germany	1:1000
PCNA	Mouse monoclonal	Santa Cruz	Germany	1:2000
β-Actin (I-19)	Goat polyclonal	Santa Cruz	Germany	1:2000

3. Results

3.1. SAGE library characteristics

Of EA and ESCC two unique SAGE libraries were made, totally consisting of over 58,000 tags. The EA library consisted of a total of 33,666 tags, containing 10,794 unique tags, while the ESCC library consisted of a total of 24,922 tags, containing 8,636 unique tags. The EA and ESCC SAGE libraries were compared to a non dysplastic BE and to a normal esophageal squamous epithelium SAGE libraries, respectively, that were previously published by our group [17]. The BE library consisted of 46,269 tags, while the normal esophageal squamous epithelium library contained 50,508 tags [17]. The SAGE library characteristics of the EA, ESCC, BE and normal squamous epithelium are described in Table 3 [17]. The complete SAGE libraries can be found on the Gene Expression Omnibus website (http://www.ncbi.nlm.nih.gov/geo/; Table 3).

Table 3
SAGE library characteristics

	Esophageal squamous cell carcinoma	Esophageal adenocarcinoma	Normal squamous esophagus*	Barrett's esophagus*
Total tags	24,922	33,666	50,508	46,269
Unique tags	8,636	10,794	14,835	16,058
Singletons	5,994	7,188	4,168	4,430
Tags 5-times present	666	945	1,201	1,202
Tags 10-times present	298	351	538	545
Accession code	GSM 110381	GSM 110379	GSM 52501	GSM 52502

Number of total tags in the esophageal squamous cell carcinoma, esophageal adenocarcinoma, normal squamous esophagus and metaplastic Barrett's esophagus [17] together with the corresponding accession code in the Gene Expression Omnibus website (http://www.ncbi.nlm.nih.gov/geo/), the number of unique tags, the number of singletons and the number of tags at least 5 times and 10 times present in each of the libraries. *Van Baal et al. [17].

3.2. Comparison of the expression profiles of esophageal adenocarcinoma and Barrett's esophagus

Between the EA and BE SAGE library 1,013 tags were significantly differentially expressed ($P < 0.05$), 673 tags were significantly up-regulated and 340 tags were significantly down-regulated in EA. From these, 55 tags were more than 10 fold up-regulated and 42 were more than 10 fold down-regulated in EA (supplemental data, Table 1).

Examples of genes corresponding to tags that were more than 10 fold up-regulated in EA were Insulin-like growth factor binding protein 7, small nuclear ribonucleoprotein D3 polypeptide, Nucleolin and Trf-proximal homolog. Examples of genes corresponding to tags that were more than 10 fold down-regulated in EA are Gastric lipase, Intelectin 1, Fatty acid binding protein (FABP) 1, Galectin 7, Trefoil factor (TFF) 1 and Tumor rejection antigen 1. Genes corresponding to tags that were significantly differentially expressed comparing EA with BE, were also clustered in groups of biological processes according to the Gene Ontology of the National Cancer Institute (http://cgap.nci.nih.gov). A higher expression level of genes in the groups of nucleobase/nucleoside/nucleotide and nucleic acid metabolism, cell growth, response to stimulus and signal transduction was found in EA compared to BE (Fig. 1(A)). Comparing EA with BE these biological processes were more than 2 fold higher in EA (Table 4).

3.3. Comparison of the expression profiles of esophageal squamous cell carcinoma and normal squamous esophagus

Between the ESCC and normal squamous esophagus SAGE library 1,235 tags were significantly dif-

ferentially expressed ($P < 0.05$), 1,022 tags were significantly up-regulated and 213 tags were significantly down-regulated in ESCC. From these, 129 tags were more than 10 fold up-regulated and 41 were more than 10 fold down-regulated in ESCC (supplemental data, Table 2).

Examples of genes corresponding to tags that were more than 10 fold up-regulated in ESCC were E-Cadherin, Tetraspanin 3, TFF1, Keratin 8, Claudin 18 and Galectin 4. Examples of genes corresponding to tags that were more than 10 fold down-regulated in ESCC are Epithelial membrane protein 1, Annexin A1, Calponin 2, Keratin 13 and S100 calcium binding protein A9. Genes, corresponding to tags that were significantly differentially expressed when comparing ESCC with normal squamous epithelium, were clustered in groups of different biological processes according to the Gene Ontology of the National Cancer Institute (http://cgap.nci.nih.gov). Genes in the biological clusters cell cycle, nucleobase/nucleoside/nucleotide and nucleic acid metabolism, metabolism, cell division, cell communication were more abundantly expressed in ESCC compared to normal squamous epithelium (Fig. 1(B)). These clusters were 10 fold or more increased in ESCC (Table 5). Genes in the cluster cell–cell signaling were predominantly expressed in normal squamous epithelium compared to ESCC (Fig. 1(B) and Table 5).

3.4. Validation of esophageal adenocarcinoma SAGE results by RT-PCR and immunoblotting

Expression levels of TFF1, TFF2, TFF3, Annexin A10, Selenium Binding Protein (SBP) 1, Lipocalin 2,
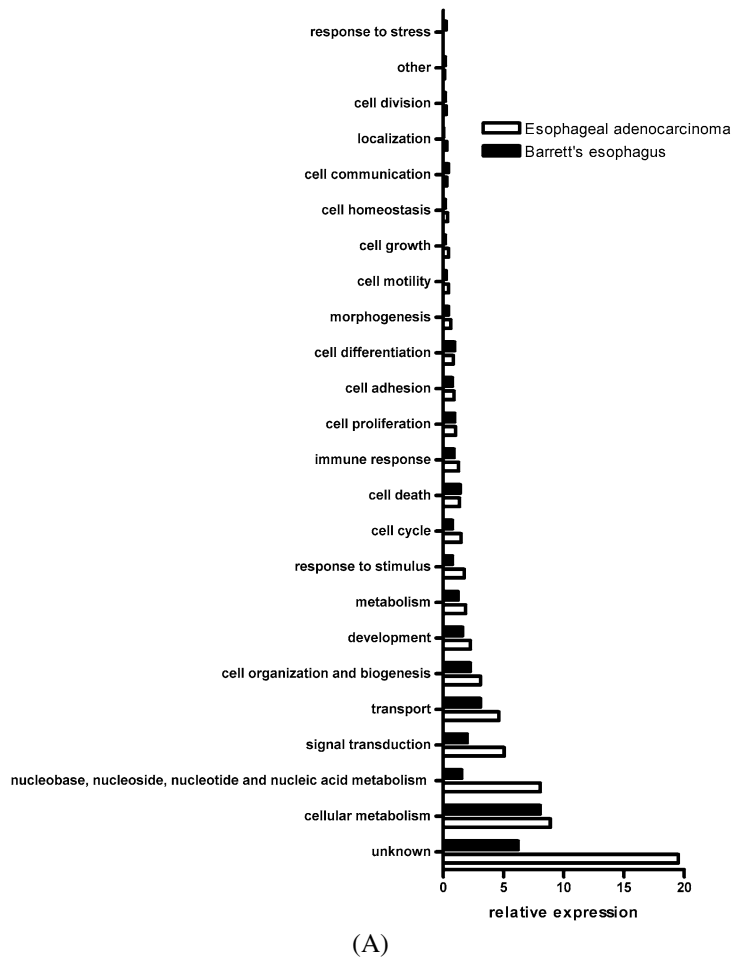
(A)

Fig. 1. Clustering of genes in biological processes. Clustering of genes corresponding to tags significantly differentially expressed comparing esophageal adenocarcinoma with Barrett's esophagus in groups of biological processes according to the Gene Ontology of the National Cancer Institute (http://cgap.nci.nih.gov) shows that genes in the clusters nucleobase/nucleoside/nucleotide and nucleic acid metabolism, cell growth, response to stimulus and signal transduction are predominantly expressed in esophageal adenocarcinoma (A). Genes in the clusters cell communication and cell differentiation are more expressed in Barrett's esophagus (A). Clustering of genes corresponding to tags significantly differentially expressed comparing esophageal squamous cell carcinoma with normal squamous esophagus shows that genes in the clusters cell cycle, metabolism, signal transduction, immune response and cell growth are more abundantly expressed in esophageal squamous cell carcinoma (B). Only genes in the cluster cell–cell signaling were predominantly expressed in normal squamous epithelium (B).

SOX4 and FABP1 were verified by RT-PCR to validate the EA SAGE results in 10 additional cases. In all cases examined, the expression of genes represented by tags in either SAGE library was confirmed. TFF1, TFF2, TFF3, Annexin A10 and FABP1 were lower expressed in all EA samples compared to the BE samples, whereas Lipocalin 2, SBP1 and SOX4 were higher expressed in all EA samples (Fig. 2).

Immunoblotting indicated that at protein level Cytokeratin (CK) 8, and BMP4 were lower expressed in EA compared to BE, whereas CK20 and EGF-receptor were higher expressed in EA compared to BE (Fig. 3).

3.5. Validation of esophageal squamous cell carcinoma SAGE results by RT-PCR and immunoblotting

To validate the ESCC SAGE results, expression levels of TFF3, Annexin A10, Prosaposin, BMP4, E-Cadherin and Plakophilin 3 were verified by RT-PCR in 13 additional cases. In all cases examined, the expression of genes represented by tags in each SAGE library was confirmed. TFF3, BMP4, Annexin A10, Prosaposin and E-Cadherin were higher expressed in all ESCC samples, whereas Plakophilin 3 was lower
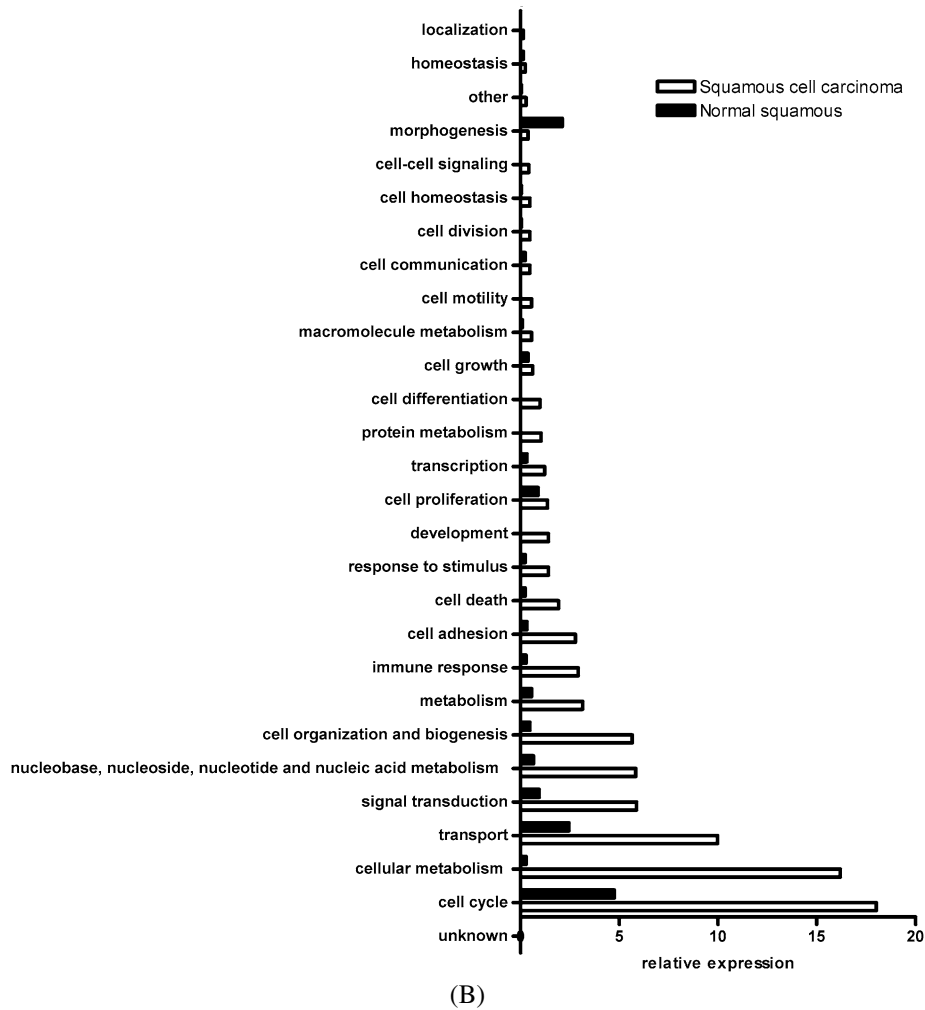
(B)

Fig. 1. (Continued).

expressed in ESCC compared to normal squamous esophagus epithelium (Fig. 4).

Immunoblotting performed for validation of certain genes on protein level showed that CK5/6, CK10/13 and EGF-receptor were lower expressed in ESCC compared to normal squamous epithelium (Fig. 5). CK8, PKC-β1, Cyclin D1, TGF-β, BMP4, ID2, p19 and p27 were higher expressed in ESCC compared to normal squamous epithelium (Fig. 5).

4. Discussion

In this study, SAGE technology was used to identify the gene expression profile of EA and ESCC, subsequently these expression profiles were compared with the gene expression profiles of BE and normal squamous epithelium. The specific information gained from this study may help to identify factors involved in esophageal carcinogenesis and to identify uniquely expressed tissue specific genes. Furthermore this information can be used to gain insight in the biology of these neoplastic lesions that may ultimately lead to a better disease management.

The major advantage of the SAGE technology compared to microarray is that SAGE generates a library of thousands of expressed genes without any previous knowledge of the tissue's repertoire. The transcriptome obtained using SAGE technology conveys not only the identity of each expressed gene but also quantifies its level of expression.

We preferred to use tissue samples of one male individual known with EA and one male individual

Table 4

Clustering of genes in biological processes comparing esophageal adenocarcinoma with Barrett's esophagus

Biologisch process	p EA	p BE	Up in EA	Down in EA
Nucleobase, nucleoside, nucleotide and nucleic acid metabolism	8.04	1.56	5.17	
Localization	0.32	0.06	5.00	
Unknown	19.52	6.23	3.14	
Signal transduction	5.06	2.01	2.52	
Cell growth	0.45	0.19	2.33	
Response to stimulus	1.75	0.78	2.25	
Cell homeostasis	0.39	0.19	2.00	
Cell cycle	1.49	0.78	1.92	
Cell motility	0.45	0.26	1.75	
Metabolism	1.88	1.23	1.53	
Transport	4.60	3.11	1.48	
Immune response	1.30	0.91	1.43	
Morphogenesis	0.65	0.45	1.43	
Development	2.27	1.62	1.40	
Cell organization and biogenesis	3.11	2.27	1.37	
Cell division	0.26	0.19	1.33	
Cell adhesion	0.91	0.78	1.17	
Cellular metabolism	8.88	8.04	1.10	
Cell proliferation	1.04	0.97	1.07	
Other	0.13	0.19		1.50
Cell communication	0.32	0.45		1.40
Cell differentiation	0.84	0.97		1.15
Cell death	1.36	1.43		1.05
Response to stress	0.00	0.26		

Clusters of biological processes with the corresponding proportion (p) in esophageal adenocarcinoma (EA) and Barrett's esophagus (BE). The fold up and down regulation of each cluster is presented.

known with ESCC, for making the SAGE libraries, and to compare EA with its precursor lesion BE and ESCC with normal squamous epithelium. We have chosen to use full biopsies, providing more accurate insight into the gene expression profiles in these important malignant conditions. A BE biopsy is a heterogeneous cell population, containing not only epithelial cells but also stromal tissue and inflammatory cells, whereas biopsies of the malignancies may contain a larger fraction of tumor than stromal cells. This may have confounded the results. Future studies aiming on specific analysis of the epithelial cells would be of interest to more specifically analyze the attribution of the epithelial cells using SAGE analysis. Another confounder is that only one EA and one ESCC SAGE library were generated and compared to a BE and normal squamous esophagus SAGE library, respectively. Since, these cancers may be considerably heterogeneous; the interpretation of these expression libraries should be taken with care. It is mandatory to verify RNA expression levels of genes of interest on a larger panel of samples for confirmation of the SAGE results. On the other hand, one should also realize that the expression of certain genes may be simply absent in this particular case. Hereto, we obtained biopsy samples of 10 patients with EA and 13 patients with ESCC to validate the differential expression of several genes by RT-PCR and proteins by immunoblotting. The results confirmed the SAGE findings and thus indicate that the SAGE data as presented here is representative.

In this study over 58,000 tags were analyzed. Further mapping the SAGE tags to known genes and mRNAs in the SAGE Genie database revealed a large number of genes known to be expressed in EA and ESCC, as well as many genes not previously recognized in EA and ESCC. For instance, the EA SAGE library confirmed high expression of Mucin 5 (TGCACAATAT), Mucin 1 (CCTGGGAAGT), Chemokine (C-X-C motif) ligand 3 (ATAATAAAAG) and Insulin-like growth factor binding protein 7 (CATATCATTA) [23–26]. In the ESCC SAGE library high number of tags were found for instance for TFF1 (CTGGCCCTCG) and

Table 5

Clustering of genes in biological processes comparing esophageal squamous cell carcinoma with normal squamous epithelium

Biologisch process	p ESCC	p SQ	Up in ESCC	Down in ESCC
Cell cycle	16.19	0.28	57.33	
Nucleobase, nucleoside, nucleotide and nucleic acid metabolism	5.65	0.47	12.00	
Metabolism	2.92	0.28	10.33	
Cell communication	0.47	0.05	10.00	
Cell division	0.47	0.05	10.00	
Signal transduction	5.84	0.66	8.86	
Immune response	2.78	0.33	8.43	
Cell adhesion	1.93	0.24	8.20	
Transport	5.88	0.94	6.25	
Cell death	1.41	0.24	6.00	
Cell growth	0.56	0.09	6.00	
Morphogenesis	0.28	0.05	6.00	
Cell organization and biogenesis	3.15	0.56	5.58	
Cellular metabolism	9.98	2.45	4.08	
Unknown	18.02	4.75	3.79	
Cell proliferation	1.22	0.33	3.71	
Cell motility	0.47	0.24	2.00	
Other	0.24	0.14	1.67	
Cell differentiation	0.61	0.38	1.63	
Development	1.36	0.89	1.53	
Cell–cell signaling	0.38	2.12		5.63
Response to stimulus	1.41	0.00		
Transcription	1.04	0.00		
Protein metabolism	0.99	0.00		
Macromolecule metabolism	0.56	0.00		
Cell homeostasis	0.42	0.00		
Homeostasis	0.14	0.00		
Localization	0.09	0.00		

Clusters of biological processes with the corresponding proportion (p) in esophageal squamous cell carcinoma (ESCC) and normal squamous esophagus (SQ). The fold up and down regulation of each cluster is presented.

TFF3 (CTCCACCCGA), known to be highly expressed in ESCC [27]. Furthermore, in ESCC low numbers of tags were found for instance for Keratin 13 (GCAGAGAGGA) and Keratin 4 (GTGACAACCT), these keratins are known to be expressed on protein level in normal squamous epithelium [28,29]. The complete SAGE libraries are uploaded on the Gene Expression Omnibus website (http://www.ncbi.nlm.nih.gov/geo/; Table 3). Using this website, all genes expressed in the different tissues can be compared to other SAGE and microarray data available on this website.

Previous transcriptome analysis studies have been performed comparing BE and esophageal carcinomas, or comparing the gene expression profiles of BE and intestinal metaplasia of the cardia to get a better overview of genes involved in BE transition [30,31]. In a microarray analysis study, Selaru et al. reported that EA clustered more closely to ESCC than to BE and they therefore concluded that the global gene expression profile fundamentally changed during the neoplastic progression of BE to EA [30]. Dahlberg et al. showed that normal squamous esophagus and gastric cardia are clustered closer together than to EA [25]. Furthermore in previous studies, SAGE, kinome and microarray analysis indicated that BE has strong similarities with the surrounding normal epithelium [17, 32,33]. Additionally Fox et al. reported that BE has a unique expression profile distinct from normal and EA specimens, whereas Wang et al. reported that the expression pattern of BE was significantly more similar to EA than to normal squamous epithelium [26,34].

In this study, comparison of the SAGE-generated gene expression profiles of EA and BE identified 1,013
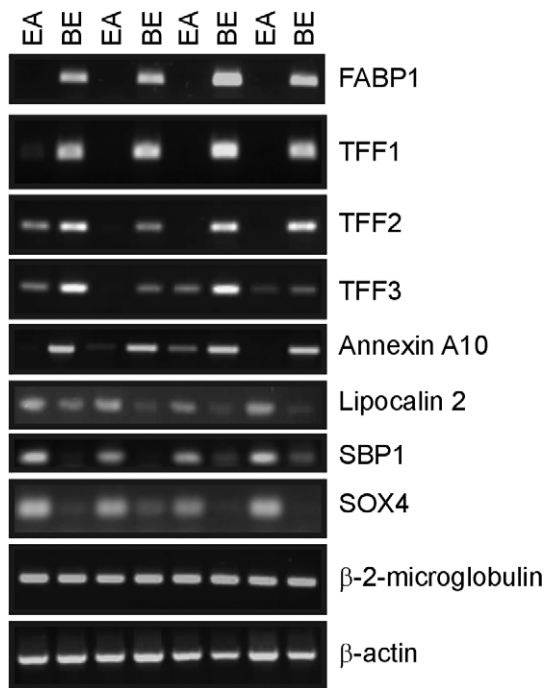
Fig. 2. RT-PCR to validate SAGE results of esophageal adenocarcinoma and Barrett's esophagus. RT-PCR on RNA from esophageal adenocarcinoma (EA) and Barrett's esophagus (BE) biopsies from different patients, demonstrates that Fatty Acid Binding Protein (FABP) 1, Trefoil Factor (TFF) 1, TFF2, TFF3 and Annexin A10 are highly expressed in the Barrett biopsies, but virtually absent in esophageal adenocarcinoma. Furthermore Lipocalin 2, Selenium Binding Protein (SBP) 1 and SRY box (SOX) 4 are highly expressed in esophageal adenocarcinoma biopsies and lower expressed in Barrett biopsies. β-actin and β-2-microglobulin were used as a control. Pictures are representative for results of 10 patients.
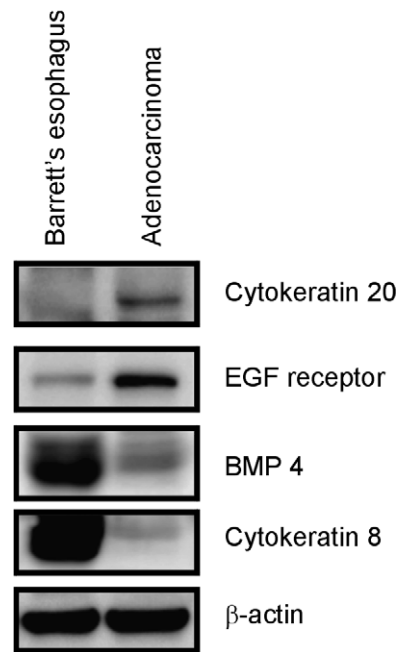

Fig. 3. Immunoblot validation of SAGE results of esophageal adenocarcinoma and Barrett's esophagus. Immunoblot analysis on protein samples from esophageal adenocarcinoma and Barrett's esophagus biopsies from different patients reveals that Cytokeratin 20 and EGF-receptor are highly expressed in esophageal adenocarcinoma (EA) and lower in Barrett's esophagus (BE). Furthermore BMP4 and Cytokeratin 8 are highly expressed in Barrett's esophagus and lower in esophageal adenocarcinoma. Pictures are representative for results of 10 patients.

significantly differentially expressed transcripts, while in a previous study we found 776 genes differentially expressed between BE and normal squamous esophageal epithelium [17,32]. From this we can conclude that EA has an individual, unique expression profile separate from BE. This is in agreement with the observations of Fox et al. [34]. We hypothesize that the BE expression profile is closer related to normal squamous esophagus then to EA, because the number of significantly differentially expressed tags between BE compared to normal squamous esophagus is 776 tags versus 1,013 tags, comparing BE to EA. A logical explanation for this observation is that metaplastic BE may be considered as a non malignant phenotype that has trans-differentiated from normal squamous epithelium.

In order to get a better overview of biological events occurring in EA and ESCC in comparison to respectively BE and normal squamous epithelium, genes corresponding to tags significantly differentially expressed between the two SAGE libraries were clustered into groups of biological processes using the website for Gene Ontology of the National Cancer Institute (http://cgap.nci.nih.gov). Interestingly, in EA an abundance of genes situated in the clusters nucleobase/nucleoside/nucleotide and nucleic acid metabolism, cell growth, response to stimulus and signal transduction were found (Fig. 1(A)), indicating that these processes play a major role in EA. Another remarkable finding is that the cluster analysis indicates that cell–cell signalling is five fold decreased in ESCC compared to normal squamous epithelium (Table 5). In contrast, the clusters cell cycle, nucleobase/nucleoside/nucleotide and nucleic acid metabolism, metabolism, cell division and cell communication were important in ESCC, all these clusters were more than 10 fold increased. In more detail, 13 main clusters of biological processes were more than 5 fold increased in ESCC in contrast to 1 cluster in normal
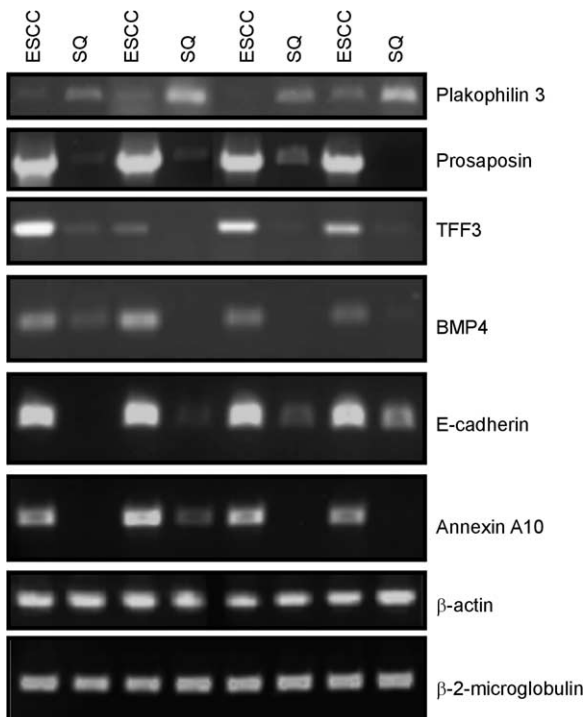
Fig. 4. RT-PCR to validate SAGE results of esophageal squamous cell carcinoma and normal squamous epithelium. RT-PCR on RNA from esophageal squamous cell carcinoma (ESCC) and normal squamous esophagus (SQ) biopsies from different patients, demonstrates that Prosaposin, Trefoil Factor (TFF) 3, Bone Morphogenetic Protein (BMP) 4, E-Cadherin and Annexin A10 are highly expressed in esophageal squamous cell carcinoma, whereas Plakophilin 3 is highly expressed in normal squamous esophagus biopsies. β-actin and β-2-microglobulin were used as a control. Pictures are representative for results of 13 patients.
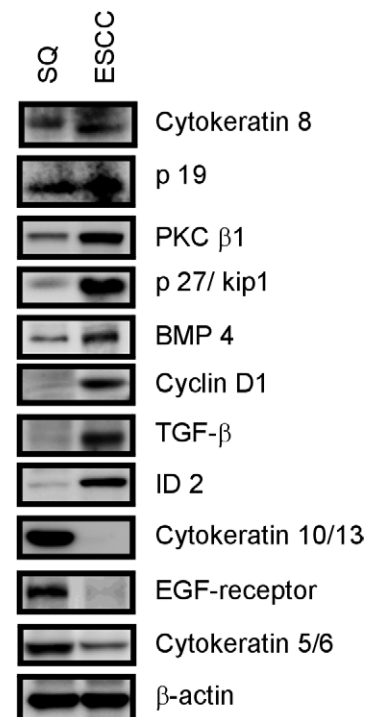

Fig. 5. Immunoblot validation of SAGE results of esophageal squamous cell carcinoma and normal squamous esophagus. Immunoblot analysis on protein samples from esophageal squamous cell carcinoma and normal squamous esophagus biopsies from different patients reveals that Cytokeratin 8, p19, PKC-β1, p27/kip, BMP4, Cyclin D1, TGF-β and ID2 are highly expressed in esophageal squamous cell carcinoma (ESCC) and lower in normal squamous esophagus (SQ). Furthermore Cytokeratin 10/13, EGF-receptor and Cytokeratin 5/6 are highly expressed in normal squamous esophagus and lower in esophageal squamous cell carcinoma. Pictures are representative for results of 13 patients.

squamous epithelium. This indicates that ESCC is a highly active epithelium in several types of processes, including cell adhesion, signal transduction, cell death, immune response and cell growth. Additionally, the comparison of the transcriptomes of ESCC and normal squamous epithelium revealed 1,235 differentially expressed genes and a total of 170 tags are more than 10 fold up- or down-regulated in ESCC compared to normal squamous esophagus. This together with the clustering analyses signifies that ESCC has a unique expression pattern distinct from normal squamous esophagus.

In summary, the comparison of the gene expression profiles of EA and ESCC with BE and normal squamous esophagus, respectively, shows that EA and ESCC have their own specific characteristics. The unique gene expression profiles described in this study harbors a wealth of information and provides us the identity of several genes involved in several cell sig-

naling pathways important in EA and ESCC. These profiles will contribute to a better understanding of the molecular alterations and elucidate important biological processes involved in cancer development of the esophagus. Therefore, this could improve tumor control and prevention and consequently could lead to a better disease management. Additionally, several unique genes that can be used as novel markers for distinguishing EA, ESCC, BE and normal squamous epithelium are identified.

Acknowledgements

References

[1] S.J. Spechler, J.M. Zeroogian, D.A. Antonioli, H.H. Wang and R.K. Goyal, Prevalence of metaplasia at the gastro-oesophageal junction, *Lancet* **344** (1994), 1533–1536.

[2] R.A. Ransford and J.A. Jankowski, Genetic versus environmental interactions in the oesophagitis-metaplasia-dysplasia-adenocarcinoma sequence (MCS) of Barrett's oesophagus, *Acta Gastroent. Belg.* **63** (2000), 18–21.

[3] C.P. Morales, R.F. Souza and S.J. Spechler, Hallmarks of cancer progression in Barrett's oesophagus, *Lancet* **360** (2002), 1587–1589.

[4] J.A. Jankowski, R.F. Harrison, I. Perry, F. Balkwill and C. Tselepis, Barrett's metaplasia, *Lancet* **356** (2000), 2079–2085.

[5] N.J. Shaheen, M.A. Crosby, E.M. Bozymski and R.S. Sandler, Is there publication bias in the reporting of cancer risk in Barrett's esophagus?, *Gastroenterology* **119** (2000), 333–338.

[6] W. Hameeteman, G.N. Tytgat, H.J. Houthoff and J.G. van den Tweel, Barrett's esophagus: development of dysplasia and adenocarcinoma, *Gastroenterology* **96** (1989), 1249–1256.

[7] D.J. Drewitz, R.E. Sampliner and H.S. Garewal, The incidence of adenocarcinoma in Barrett's esophagus: a prospective study of 170 patients followed 4.8 years, *Am. J. Gastroent.* **92** (1997), 212–215.

[8] S.S. Devesa, W.J. Blot and J.F. Fraumeni, Jr., Changing patterns in the incidence of esophageal and gastric carcinoma in the United States, *Cancer* **83** (1998), 2049–2053.

[9] A.T. Prach, T.A. MacDonald, D.A. Hopwood and D.A. Johnston, Increasing incidence of Barrett's oesophagus: education, enthusiasm, or epidemiology?, *Lancet* **350** (1997), 933.

[10] P. Bytzer, P.B. Christensen, P. Damkier, K. Vinding and N. Seersholm, Adenocarcinoma of the esophagus and Barrett's esophagus: a population-based study, *Am. J. Gastroent.* **94** (1999), 86–91.

[11] A.G. Casson, Molecular biology of Barrett's esophagus and esophageal cancer: role of p53, *World J. Gastroent.* **4** (1998), 277–279.

[12] K.K. Krishnadath, B.J. Reid and K.K. Wang, Biomarkers in Barrett esophagus, *Mayo Clin. Proc.* **76** (2001), 438–446.

[13] B.J. Reid, P.L. Blount and P.S. Rabinovitch, Biomarkers in Barrett's esophagus, *Gastrointest. Endosc. Clin. N. Am.* **13** (2003), 369–397.

[14] K.Y. Lam, S. Law, P.H. Tung and J. Wong, Esophageal small cell carcinomas: clinicopathologic parameters, p53 overexpression, proliferation marker, and their impact on pathogenesis, *Arch. Pathol. Lab. Med.* **124** (2000), 228–233.

[15] A.M. Mandard, P. Hainaut and M. Hollstein, Genetic steps in the development of squamous cell carcinoma of the esophagus, *Mutat. Res.* **462** (2000), 335–342.

[16] V.E. Velculescu, L. Zhang, B. Vogelstein and K.W. Kinzler, Serial analysis of gene expression, *Science* **270** (1995), 484–487.

[17] J.W. van Baal, F. Milano, A.M. Rygiel, J.J. Bergman, W.D. Rosmolen, S.J. van Deventer, K.K. Wang, M.P. Peppelenbosch and K.K. Krishnadath, A comparative analysis by SAGE of gene expression profiles of barrett's esophagus, normal squamous esophagus, and gastric cardia, *Gastroenterology* **129** (2005), 1274–1281.

[18] A.H. van Kampen, B.D. van Schaik, E. Pauws, E.M. Michiels, J.M. Ruijter, H.N. Caron, R. Versteeg, S.H. Heisterkamp, J.A. Leunissen, F. Baas and M. van der Mee, USAGE: a web-based approach towards the analysis of SAGE data. Serial analysis of gene expression, *Bioinformatics* **16** (2000), 899–905.

[19] A.E. Lash, C.M. Tolstoshev, L. Wagner, G.D. Schuler, R.L. Strausberg, G.J. Riggins and S.F. Altschul, SAGEmap: a public gene expression resource, *Genome Res.* **10** (2000), 1051–1060.

[20] A.J. Kal, A.J. van Zonneveld, V. Benes, M. van den Berg, M.G. Koerkamp, K. Albermann, N. Strack, J.M. Ruijter, A. Richter, B. Dujon, W. Ansorge and H.F. Tabak, Dynamics of gene expression revealed by comparison of serial analysis of gene expression transcript profiles from yeast grown on two different carbon sources, *Mol. Biol. Cell.* **10** (1999), 1859–1872.

[21] J.M. Ruijter, A.H. van Kampen and F. Baas, Statistical evaluation of SAGE libraries: consequences for experimental design, *Physiol. Genomics* **11** (2002), 37–44.

[22] J.C. Hardwick, G.R. van den Brink, S.A. Bleuming, I. Ballester, J.M. van den Brande, J.J. Keller, G.J. Offerhaus, S.J. van Deventer, and M.P. Peppelenbosch, Bone morphogenetic protein 2 is expressed by, and acts upon, mature epithelial cells in the colon, *Gastroenterology* **126** (2004), 111–121.

[23] P. Guillem, V. Billeret, M.P. Buisine, J.F. Flejou, M. Lecomte-Houcke, P. Degand, J.P. Aubert, J.P. Triboulet and N. Porchet, Mucin gene expression and cell differentiation in human normal, premalignant and malignant esophagus, *Int. J. Cancer* **88** (2000), 856–861.

[24] C.N. Chinyama, R.E. Marshall, W.J. Owen, R.C. Mason, D. Kothari, M.L. Wilkinson and J.D. Sanderson, Expression of MUC1 and MUC2 mucin gene products in Barrett's metaplasia, dysplasia and adenocarcinoma: an immunopathological study with clinical correlation, *Histopathology* **35** (1999), 517–524.

[25] P.S. Dahlberg, L.F. Ferrin, S.M. Grindle, C.M. Nelson, C.D. Hoang and B. Jacobson, Gene expression profiles in esophageal adenocarcinoma, *Ann. Thorac. Surg.* **77** (2004), 1008–1015.

[26] S. Wang, M. Zhan, J. Yin, J.M. Abraham, Y. Mori, F. Sato, Y. Xu, A. Olaru, A.T. Berki, H. Li, K. Schulmann, T. Kan, J.P. Hamilton, B. Paun, M.M. Yu, Z. Jin, Y. Cheng, T. Ito, C. Mantzur, B.D. Greenwald and S.J. Meltzer, Transcriptional profiling suggests that Barrett's metaplasia is an early intermediate stage in esophageal adenocarcinogenesis, *Oncogene* (2006).

[27] C. Labouvie, J.C. Machado, F. Carneiro, M. Sarbia, M. Vieth, R. Porschen, G. Seitz and N. Blin, Differential expression of mucins and trefoil peptides in native epithelium, Barrett's metaplasia and squamous cell carcinoma of the oesophagus, *J. Cancer Res. Clin. Oncol.* **125** (1999), 71–76.

[28] H.M. El-Zimaity and D.Y. Graham, Cytokeratin subsets for distinguishing Barrett's esophagus from intestinal metaplasia in the cardia using endoscopic biopsy specimens, *Am. J. Gastroent.* **96** (2001), 1378–1382.

[29] R., Moll, W.W., Franke, D.L., Schiller, B. Geiger and R. Krepler, The catalog of human cytokeratins: patterns of expression in normal epithelia, tumors and cultured cells, *Cell* **31** (1982), 11–24.

[30] F.M. Selaru, T. Zou, Y. Xu, V. Shustova, J. Yin, Y. Mori, F. Sato, S. Wang, A. Olaru, D. Shibata, B.D. Greenwald, M.J. Krasna,

J.M. Abraham and S.J. Meltzer, Global gene expression profiling in Barrett's esophagus and esophageal cancer: a comparative analysis using cDNA microarrays, *Oncogene* **21** (2002), 475–478.

[31] Y. Chang, J. Gong, B. Liu, J. Zhang and F. Dai, Gene expression profiling in Barrett's esophagus and cardia intestinal metaplasia: a comparative analysis using cDNA microarray, *World J. Gastroent.* **10** (2004), 3194–3196.

[32] M.T. Barrett, K.Y. Yeung, W.L. Ruzzo, L. Hsu, P.L. Blount, R. Sullivan, H. Zarbl, J. Delrow, P.S. Rabinovitch and B.J. Reid, Transcriptional analyses of Barrett's metaplasia and normal upper GI mucosae, *Neoplasia* **4** (2002), 121–128.

[33] J.W. van Baal, S.H. Diks, R.J. Wanders, A.M. Rygiel, F. Milano, J. Joore, J.J. Bergman, M.P. Peppelenbosch and K.K. Krishnadath, Comparison of kinome profiles of Barrett's esophagus with normal squamous esophagus and normal gastric cardia, *Cancer Res.* **66** (2006), 11605–11612.

[34] C.A. Fox, L.M. Sapinoso, H. Zhang, W. Zhang, H.L. McLeod, G.R. Petroni, T. Mullick, C.A. Moskaluk, H.F. Frierson, G.M. Hampton and S.M. Powell, Altered expression of TFF-1 and CES-2 in Barrett's esophagus and associated adenocarcinomas, *Neoplasia* **7** (2005), 407–416.