УДК 577.21:577.214.622+616–006.484.04

V.M. KAVSAN [1], V.V. DMITRENKO [1],
K.O. SHOSTAK [1], T.V. BUKREIEVA [1], N.Y. VITAK [1],
O.E. SIMIRENKO [1], T.A. MALISHEVA [2], M.I. SHAMAYEV [2]
V.D. ROZUMENKO [2], Y.A. ZOZULYA [2]

[1] Institute of Molecular Biology and Genetics, Zabolotnogo Str. 150, 03143 Kyiv, Ukraine
[2] A.P.Romodanov Institute of Neurosurgery, Manuilskogo Str. 32, 04050 Kyiv, Ukraine
E-mail: kavsan@imbg.org.ua

COMPARISON OF MICROARRAY AND SAGE TECHNIQUES IN GENE EXPRESSION ANALYSIS OF HUMAN GLIOBLASTOMA

To enhance glioblastoma (GB) marker discovery we compared gene expression in GB with human normal brain (NB) by accessing SAGE Genie web site and compared obtained results with published data. Nine GB and five NB SAGE-libraries were analyzed using the Digital Gene Expression Displayer (DGED), the results of DGED were tested by Northern blot analysis and RT-PCR of arbitrary selected genes. Review of available data from the articles on gene expression profiling by microarray-based hybridization showed as few as 35 overlapped genes with increased expression in GB. Some of them were identified in four articles, but most genes in three or even in two investigations. There was found also some differences between SAGE results of GB analysis. Digital Gene Expression Displayer approach revealed 676 genes differentially expressed in GB vs. NB with cut-off ratio: twofold change and $P \leqslant 0.05$. Differential expression of selected genes obtained by DGED was confirmed by Northern analysis and RT-PCR. Altogether, only 105 of 955 genes presented in published investigations were among the genes obtained by DGED. Comparison of the results obtained by microarrays and SAGE is very complicated because authors present only the most prominent differentially expressed genes. However, even available data give quite poor overlapping of genes revealed by microarrays. Some differences between results obtained by SAGE in different investigations can be explained by high dependence on the statistical methods used. As for now, the best solution to search for molecular tumor markers is to compare all available results and to select only those genes, which significant expression in tumor combined with very low expression in normal tissues was reproduced in several articles. 105 differentially expressed genes, common to both methods, can be included in the list of candidates for the molecular typing of GBs. Some genes, encoded cell surface or extra-cellular proteins may be useful for targeting gliomas with antibody-based therapy.

© V.M. KAVSAN, V.V. DMITRENKO, K.O. SHOSTAK, T.V. BUKREIEVA, N.Y. VITAK, O.E. SIMIRENKO, T.A. MALISHEVA, M.I. SHAMAYEV, V.D. ROZUMENKO, Y.A. ZOZULYA, 2007

Introduction. Two newly developed high-throughput gene expression profiling technologies, microarray-based hybridization [1—3] and Serial Analysis of Gene Expression (SAGE) [4—6] have dramatically reshaped the biomedical research on elucidating the role of genes in cancer. However, despite the power of microarray-based hybridisation and Serial Analysis of Gene Expression (SAGE) in the search for new molecular markers of human tumors, longstanding doubts persist as to the worth of the torrent of data that microarrays produce. Experiments have proved difficult to reproduce, and the lists of genes found in similar studies often have only limited overlap. The SAGE method also has its own problems: certain genes have no reliable tags; some tags have multiple matches against the EST databases. Also, sequencing errors could inflate the number of different tags. Alternative splicing of transcripts and sequence polymorphisms make the evaluation of results even more complicated, because multiple gene tags would correspond to a single gene.

To investigate the correlation and reproducibility between these currently often used approaches, we compared the available data obtained by microarray analysis and SAGE on changes of gene expression in glioblastoma, the most aggressive brain tumor that was the topic of our recent publications [7—9]. The practical aim of this study was to identify the changes of gene expression by comparison of results obtained by different approaches and described in different articles that might be helpful as molecular markers of glial tumors.

Materials and methods. PubMed search with different combinations of keywords was performed to find the publications cited oligonucleotide and cDNA microarrays analysis for gene expression profiles in glioblastomas.

Nine SAGE-libraries of glioblastoma (GB, WHO grade IV astrocytoma) and five normal adult human brain (NB) SAGE-libraries were analyzed to compare gene expression by accessing NCI CGAP web site (http://cgap.nci.nih.gov/SAGE) and using the search tool of Digital Gene Expression Displayer (DGED) provided by the SAGE Genie database. SAGE-library GBM1062, also available in this database, prepared from tissue of 4-month-old child, has a big difference in gene expression profile as compared to adult tissue. For example, tag counts for C1QA, CD74, CHI3L1, COL1A1, GFAP, IGFBP7, and IGHG1 genes were very low in this library (8 tags for IGFBP7 gene, 18

tags for CD74 gene and only one tag for other 5 genes) in comparison with several hundred tags for each gene in adult GB tumors. Only SAGE-libraries of normal adult cortex, adult cerebellum and adult thalamus have been chosen for the comparison. Available SAGE_Brain_fetal_normal_B_S1 and AGE_Brain_normal_peds_cortex_B_H1571 libraries, prepared from fetal brain and brain of 15-month-old child, correspondingly, differed also very much in gene expression profile as compared to adult NB libraries (data not shown). SAGE_Brain_ normal_leptomeninges_B_AL2 library was originated from the tissue of non-glial origin, and SAGE_Brain_normal_substantia_nigra_B_1 library was from substantia nigra, which is not the place of glioma arising. The analysis was carried out under two different cutoff ratio (5-fold change and twofold change of gene expression) and significance filter $P \leqslant 0.05$. The UniGene database of NCBI was searched to obtain the expressed sequence tags (ESTs) containing coding regions of corresponding mRNA. The selected cDNA clones were obtained from German Resource Center for Genome Research (RZPD).

Brain tumor tissue samples were collected from A.P. Romodanov Institute of Neurosurgery (Kyiv) under the approval of the Institute Review Board. Surgical specimens of histologically normal brain tissue adjacent to tumors were used as a source of normal adult human brain RNA. Northern analysis was performed as described in our previous works [8, 9] with following probes for hybridization: Annexin A1 (ANXA1) cDNA, clone IMAGp998L168452; Beta-2-microglobulin (B2M) cDNA, clone IMAGp998I211214; CD74 antigen, invariant polypeptide of major histocompatibility complex, class II antigen-associated (CD74) cDNA, clone IMAGp998P143584; Complement component 1, q subcomponent, alpha polypeptide (C1QA) cDNA, clone IMAGp958J19169; HC gp-39, human cartilage glycoprotein-39 (CHI3L1) cDNA, clone IMAGp998P09248; galectin 3 (LGALS3) cDNA, clone IMAGp998A089671; Sec61 gamma subunit (SEC61G) cDNA, clone IMAGp998G024700; Secreted protein, acidic, cysteine-rich (osteonectin) (SPARC) cDNA, clone IMAGp998E214660; Serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3 (SERPINA3) cDNA, clone IMAGp958K06246; and 500 bp long RT-PCR product of β-actin mRNA. Hybridization bands were normalized to β-actin and compared by densitometric analysis of hybridization signals using the Scion Image 1.62c program.

Semi-quantitative RT-PCR with gene-specific primers was performed on normal brain and brain tumor RNA samples as described by Rae et al [10]. Equal amounts of total cellular RNA (10 μg each) were transcribed into cDNA with an oligo(dT) primer. PCR was performed in 20 μl reaction mixture with cDNA synthesized from 0.5 μg of total RNA, 2 U Taq-polymerase, 1×PCR buffer, 0.2 mM dNTPs, and 1 μM gene-specific primers: ABCC3 (for TGTCCCACCTGCACACGTTTG, rev CGTGTCATTCACCACTTGGGG), COL1A1 (for GTGAACAAGGTCCCTCTGGA, rev CGCCATACTCGAACTGGAAT), COL3A1 (for TTGACCCTAACCAAGGATGC, rev GTGTGTTTCGTGCAACCATC), EGFR (for TGAAAACAGCTGCAAGGCCACA, rev ATGGCACAGGTGGACACATGG), FMOD (for CCTCCAAGGCAATAGGATCAAT, rev TGCCCATGCCACTTTTGAAGTT), MFAP2 (for GTCCAACAGGAAGTCATCCCAG, rev GGGGGACTGTCTGTCCTCAAAA), CHI3L2 (for GCAGGAACCAGGAAAATTCAC, rev AGGCTTCTCTTGACTGCTTGG). Thermal cycling parameters were: denaturation at 94 °C, 30 sec; annealing at appropriate temperature for each primer pair, 1 min; synthesis at 72 °C, 1 min for 30 cycles, followed by a further 7 min at 72 °C. The number of cycles was then decreased until the PCR product amplification rate was in the linear phase (27 cycles). Amplified products were electrophoresed in a 2 % agarose gel.

Results and discussion. *Comparison of the published results of glioblastoma analysis by cDNA and oligonucleotide microchips.* On May 1, 2006, we carried out a PubMed search and found 73 publications using the combination of words «microarray glioblastoma», 87 on «glioblastoma cDNA microarray», 54 on «glioblastoma gene expression microarray», and 82 on «glioblastoma oligonucleotide microarray». After removing redundancies and irrelevant publications, only 17 of 117 articles on gene expression profiling in glioblastoma by microarrays technique where authors compared gene expression profiles of native GBs and normal brain or lower grade gliomas and gave the names of differentially expressed genes, were remained for further analysis (Table 1).

Table 1

Results of gene expression analysis in glioblastoma by microarray method

Author [Reference]	Year of publication	Name of microchip[a]	Type of microchip (cDNA or oligo)	Q-ty of genes on chip	Q-ty of genes overexpressed in GB	Q-ty of described genes	Comparison of GB to which other tissue
Fuller et al. [11]	1999	Human Atlas cDNA Expression Array Blot I, Clontech Laboratories Inc., Palo Alto, CA, USA	cDNA	588	*	6	Normal brain
Sallinen et al. [12]	2000	Atlas Human Cancer cDNA Expression Array, Clontech Laboratories Inc., Palo Alto	cDNA	588	117	49	Normal brain
Kim et al. [13]	2002	Human Atlas cDNA Expression Array Blots, Clontech Laboratories Inc., Palo Alto	cDNA	597	*	10	Normal brain
Raza et al. [14]	2004	Human Atlas cDNA Expression Array Blots, Clontech Laboratories Inc., Palo Alto	cDNA	588	26	26	GBs with varying degrees of necrosis
Godard et al. [15]	2003	Atlas Human Cancer 1.2 Array, Clontech	cDNA	1176	239	34	Low grade astrocytoma
Somasundaram et al. [16]	2005	GeneMap Human Cancer Array, Genomic Solutions Inc.	cDNA	1152	*	9	Normal brain
Rickman et al. [17]	2001	Hu6800 GeneChip, Affymetrix, (Santa Clara, CA)	oligo	6800	167 703	49	Low grade astrocytoma Normal brain
Markert et al. [18]	2001	HUGeneFL Array, Affymetrix, (Santa Clara, CA)	oligo	6800	155	355	Normal brain
Van den Boom et al. [19]	2003	HuGeneFL, Affymetrix, (Santa Clara, CA)	oligo	6800	66 dif. expressed	7	Low grade astrocytoma
Ljubimova et al. [20]	2001	UniGEM™V gene microarray, Incyte Genomics, St. Louis, MO	cDNA	11004	2345	14	Normal brain
Tanwar et al. [21]	2002	Human V cDNA microarray Incyte Genomics, St. Louis, MO	cDNA	10000	*	36	Normal brain
Mischel et al. [22]	2003	U95Av2, Affymetrix, (Santa Clara, CA)	oligo	10000	90 dif. expressed	31	EGFR+ vs (EGFR-) GB
Nutt et al. [23]	2003	U95Av2 GeneChip, Affymetrix, (Santa Clara, CA)	oligo	12600	*	19	Anaplastic Oligodendroglioma
Nigro et al. [24]	2005	U95Av2, Affymetrix, (Santa Clara, CA)	oligo	10000	53	53	Normal brain
Freije et al. [25]	2004	HG U133 set, Affymetrix, (Santa Clara, CA)	oligo	30000	*	44	Between grade III and IV gliomas
Yokota et al. [26]	2006	Homemade array	cDNA	25344	54	99	Normal brain
Liau et al. [27]	2000	Homemade array	cDNA	26	*	8	Normal brain

Note. [a]All data are given as they are presented in the original articles. * Not available from the article.

In spite of different protocols and platforms used in microarray studies, we made an attempt to compare the described data and to reveal common genes with significantly changed expression in glioblastoma. Table 1 summarizes the characteristics of microchips, total quantity of overexpressed genes, quantity of described genes, and tissues, which were compared in these studies.

Unfortunately, authors present in their articles only the most prominent differentially expressed genes, however even comparison of available data shows quite poor overlapping of genes revealed by microarrays as it is possible to see in Table 2. Only limited number of 849 described differentially expressed in GB genes was identified at least in 4 of 17 investigations with microarray approach, other

ISSN 0564–3783. Цитология и генетика. 2007. № 1

Common differentially expressed genes revealed by cDNA-array or oligonucleotide microarray

Gene symbol	ReferencesFull name	Fold change of gene expression in GB as compared to NB a										
		[1]	[12]	[13]	[15]	[16]	[17]	[18]	[19]	[20]	[21]	[22
AQP1	Aquaporin 1/water transport, ion channel							>5.2				
CD44	CD44 antigen							>>10				
CD74	CD74 antigen				Inc							
CDK4	Cyclin-dependent kinase 4				>2.42		>6.2					>5
CDKN3	Cyclin-dependent kinase inhibitor 3 (CDK2-associated dual specificity phosphatase)						>1.6					
CENPF	Centromere protein F, 350/400ka (mitosin)								>5.7			
CH3L1 (HCgp-39)	Cartilage glycoprotein-39							>>10			Inc	
COL4A1	Collagen, type IV, alpha 1								>7.46	>3.2	Inc	
COL4A2	Collagen, type IV, alpha 2								>11.05			
COL6A3	Collagen, type VI, alpha 3										Inc	
EGFR	Epidermal growth factor receptor				>3.51					>7.8		In
FCGR3A	Fc fragment of IgG, low affinity IIIa, receptor (CD16a)				Inc							
FLNA	Filamin A, alpha (actin binding protein 280)						>9.4	>>10				
FN1	Fibronectin-1		>14.8		>5.16					>11.4	Inc	
GRN	Granulin							>>10				
HLA-DRA	Major histocompatibility Complex, class II, DR alpha				>2.95			>1.5				
HLA-DRB1	MHC Class II HLA DR- beta							>1.1				
IGF-II	Insulin-like growth factor 2							>1.0		>10.3		
IGFBP2	Insulin-like growth factor binding protein 2	Inc	>24.4	Inc	Inc		>12.7					
IGFBP3	Insulin-like growth factor binding protein 3		>9.6		Inc			>1.8		>13.4		
IGFBP4	Insulin-like growth factor binding protein 4							>2.7				
IGFBP5	Insulin-like growth factor binding protein 5		>4.6					>3.8		>5.3		
IGFBP6	Insulin-like growth factor binding protein 6		>2.9					>>10				
MGP	Matrix Gla protein /cartilage condensation, ossification, extracellular								>3.60			

Continue of table 2

Gene symbol	ReferencesFull name	Fold change of gene expression in GB as compared to NB and probability															
		[1]	[12]	[13]	[15]	[16]	[17]	[18]	[19]	[20]	[21]	[22]	[23]	[24]	[25]	[26]	[27]
MIF	Macrophage migration inhibitory factor (glycosylation-inhibiting factor)		>2.0					>>10									
NNMT	Nicotinamide N-methyltransferase							>>10						>4.5			
SPARC	Osteonectin		>4.8			Inc		<1.3									
TGFB1	Transforming growth factor, beta 1							<<10						>6.9			
TIMP-1	TIMP metallopeptidase inhibitor 1		>3.6		Inc									>5.9			
TNC	Tenascin C (hexabrachion)				>2.99		>9.2			>2.0				>2.4			
TOP2A	Topoisomerase II alpha (170 kD						>2.2		>8.8					>4.5	Inc		
VEGF	Vascular endothelial growth factor	Inc	>11.8		>2.94				>17.8	>2.3		Inc		>5.0	Inc		
VIM	Vimentin		>2.7		Inc	Inc				>5.4			>2.55	>4.9			
SRI	Sorcin											Inc				>4.3	
SNCA	Synuclein, alpha						<0.2								Dec		

N o t e. > — increase of expression, < — decrease of expression, >>10 — increase of expression more than 10-fold, << — decrease of expression more than 10-fold, Inc — increased expression in GB compared to NB, Dec — decreased expression in GB compared to NB.

overlapped genes were described only in three or even in two different articles. The poor overlapping of the results obtained by microarray approach can be partially explained by different number of genes analyzed in different works. In particular, the first Clontech cDNA arrays contained too small sets of genes. However, even the studies with the same type of microchips did not show significant reproducibility. For example, Hu6800 GeneChip was used in three works [17—19] but only one of described genes, TOP2A, was found as common between [17] and [19] and other gene, IGFBP4, was common for [18] and [19]. The explanation of such significant differences in obtained results was given in four recently appeared independent studies [28—31], which confirmed three persistent criticisms of the method: that the bewildering array of platforms and research protocols available can make results from different studies hard to compare; that, in the hands of less experienced labs, home-made arrays are less dependable than commercial chips; and that different labs doing the same study can often get very different results [32].

Determination of glioblastoma-associated genes by DGED analysis and comparison with published data obtained by microarray techniques. SAGE tag frequencies depend largely on the total number of tags counted [33, 34]. If more tags were counted, then the tag frequency for each gene would be higher. In our previous work [8], the comparison of five GB SAGE libraries with two NB SAGE libraries, which were available that time, has revealed 117 genes with more than 5-fold difference at the $P \leqslant 0.05$ level. Of these 117 genes, 24 increased their expression in GBs. Four new GB SAGE libraries have appeared recently in the SAGE Genie database. Comparing all nine GB tumor SAGE libraries with five NB SAGE libraries (cut-off ratio: 5-fold change, $P \leqslant 0.05$), the number of tags was 129, when gene tags that had no reliable matches in UniGene clusters, mitochondrial genes, ESTs, and the lower ranked tag (if more than one gene tag corresponded to the same UniGene cluster) were excluded from the list. 44 genes met the criteria for genes overexpressed and 85 genes met the criteria for genes down-regulated in tumors (Supplementary Table 1). In this table, the final column gives, for each gene (i.e., tag) the overall ratio for all of the GB samples, taken together as a group, to all of the NB samples, at a signifi-

Table 3

Common differentially expressed genes described by Markert et al. [18] and found by SAGE DGED

Genes found by microarray technique			Same genes revealed by SAGE DGED				
Identifier/ Accession	Gene name (as it was shown by Markert et al. [18])	Fold of gene expression change in GB vs NB (I) — increase (D) — decrease	SAGE tag sequence	Unigene cluster	Gene symbol and official name	Tag Odds*(GB/NB)	P
M95178	Actinin, alpha- (non-muscle)	1,5 (I) P=0,778	TCCTTCTCCA	Hs. 119000	ACTN1 Actinin, alpha 1	6.97	0.00
D14874	Adrenomedullin	1,7 (I) P=0,615	AAGAGAAAG	Hs. 441047	ADM adrenomedulin	12.30	0.00
U01691	Annexin V	4,9 (I) P=0,093	ATACTTTAAT	Hs. 145741	ANXA5 annexin A5	9.27	0.00
U41518	AQP-1	5,2 (I) P=0,339	AGCTTTGAAG	Hs. 76152	AQP1 aquaporin 1 (channel-forming integral protein, 28 kDa)	NaN**	0.00
J00105	Microglobulin, beta2-	2,2 (I) P=0,203	TTTGATGTAT	Hs.534255	B2M Beta 2 microglobulin	10.40	0.00
L21954	Peripheral benzodiazepine receptor	>10 (I) P=0,047	GGGGACGGCG	Hs.202	BZRP Benzodiazapine receptor (peripheral)	6.80	0.00
M94345	Macrophage capping factor	>10 (I) P=0,439	CTCCCCTGCC	Hs.516155	CAPG Capping protein (actin filament), gelsolin like	12.98	0.00
M83667	NF-IL6-beta	7,3 (I) P=0,335	GCCGCCGTGC	Hs.440829	CEBPD CCAAT/enhancer binding protein (C/EBP), delta	4.10	0.01
Y08374	GP-39 cartilage protein (CHI3L1)	>10 (I) P=0,038	GTATGGGCCC	Hs.382202	CHI3L1 Chitinase 3-like 1 (cartilage glycoprotein-39)	66.68	0.00
S80562	Acidic calponin	5,4 (I) P=0,09	ATCAGTGTGA	Hs.483454	CNN3 Calponin 3, acidic	7.33	0.00
L47738	Inducible protein	9,0 (D) P=0,077	CTGGTTTCTC	Hs.519702	CYFIP2 Cytoplasmic FMR1 interacting protein 2	0.13	0.00
X02761	Fibronectin (FN precursor)	1,2 (I) P=0,862	ATCTTGTTAC	Hs. 418138	FN1 fibronectin 1	21.53	0.00
V01512	c-fos	1,2 (I) P=0,846	TGGAAAGTGA	Hs.25647	FOS V fos FBJ murine osteosarcoma viral oncogene homolog	8.98	0.00
M37400	Cytosolic aspartate aminotransferase	>10 (D) P=0,552	CACGGACACG	Hs.500756	GOT1 Glutamic oxaloacetic transaminase 1, soluble (aspartate aminotransferase 1)	0.14	0.01
L13266	NMDA receptor 1 (NR1-1)	>10 (D) P=0,006	CCTCGGTCAG	Hs.495496	GRIN1 Glutamate receptor, ionotropic, N methyl D aspartate 1	0	0.00
L76224	NMDA receptor 2C (GRIN 2C)	1,2 (D) P=0,865	GTGAGGGCTG	Hs.436980	GRIN2C Glutamate receptor,	0.05	0.00

Continue of table 3

Genes found by microarray technique			Same genes revealed by SAGE DGED				
Identifier/ Accession	Gene name (as it was shown by Markert et al. [18])	Fold of gene expression change in GB vs NB (I) — increase (D) — decrease	SAGE tag sequence	Unigene cluster	Gene symbol and official name	Tag Odds*(GB/ NB)	P
					ionotropic, N methyl D aspartate 2C		
X62320	Epithelin 1 & 2 (granulin)	>10 (I) P=0,235	GGAGGTGGGG	Hs.514220	GRN Granulin	3.98	0.01
X00274	HLA-DR alpha heavy chain (MHC class II)	1,5 (I) P=0,533	GGGCATCTCT	Hs. 409805	HLA-DRA Major histocompatibity complex, class II DR alpha	12.40	0.00
M62403	IGFBP-4	2,7 (I) P=0,414	CTGGATTCAC	Hs. 1516	IGFBP4 Insulin-like growth factor binding protein 4	28.36	0.00
L27560	IGFBP-5	3,8 (I) P=0,151	GATAGCACAG	Hs. 369982	IGFBP5 Insulin like growth factorbinding protein	3.72	0.00
J04111	c-jun proto onco (JUN) clone hCJ-1	7,7 (I) P=0,002	TGCTGTGACC	Hs.525704	JUN V jun sarcoma virus 17 oncogene homolog (avian)	4.75	0.00
X53961	Lactoferrin	>10 (I) P=0,106	GCAAAACAAC	Hs.529517	LTF Lactotransferrin	NaN**	0.00
X14474	Microtubule-associated tau protein	>10 (D) P=0,322	GTAGACTCGC	Hs.101174	MAPT Microtubule-associated protein tau	0.07	0.00
M13577	Myelin basic protein (MBP)	>10 (D) P=0,011	TCTATTAATA	Hs.501262	MBP Myelin basic protein	0.04	0.00
M55131	CFTR (cystic fibrosis transmemb conduct reg)	1,6 (D) P=0,76	GCCACAAGCA	Hs.276808	MGAT3 Mannosyl (beta 1,4) glycoprotein beta 1,4 N acetylglucosaminyltransferase	0.11	0.01
HG3513-HT3707	Meromyosin, light	1,0 (I) P=0,99	GGTTTACAGA	Hs.190086	MRCL3 Myosin regulatory light chain MRCL3	14.01	0.00
X15306	NF-H 1 neurofilament protein	>10 (D) P=0,005	GTGGCGGTGG	Hs.198760	NEFH Neurofilament, heavy polypeptide 200kDa	0.03	0.01
U08021	Nicotinamide N-methyltransferase (NNMT)	>10 (I) P=0,061	GCCAACAACG	Hs.503911	NNMT Nicotinamide N methyltransferase	27.34	0.00
U61849	Neuronal pentraxin 1 (NPTX1)	2,4 (D) P=0,663	TGGTTCACAT	Hs.514556	NPTX1 Neuronal pentraxin I	0.03	0.00
U02020	Pre-B cell enhancing factor (PBEF)	>10 (I) P=0,12	GCCTTAACAA	Hs.489615	PBEF1 Pre B cell colony enhancing factor 1	10.93	0.00
X06318	PKC beta-I	>10 (I) P=0,104	TCTGTTATGT	Hs.460355	PRKCB1 Protein kinase C, beta 1	0.03	0.01
M16447	Dihydropteridine reductase (hDHPR)	6,4 (D) P=0,098	GATTGCTGGA	Hs.75438	QDPR Quinoid dihydropteridine reductase	0.27	0.00

Continue of table 3

Genes found by microarray technique			Same genes revealed by SAGE DGED				
Identifier/ Accession	Gene name (as it was shown by Markert et al. [18])	Fold of gene expression change in GB vs NB (I) — increase (D) — decrease	SAGE tag sequence	Unigene cluster	Gene symbol and official name	Tag Odds*(GB/ NB)	P
D87074	KIAA0237	>10 (D) P=0,059	ATTAAAGTCA	Hs.434924	RIMS3 Regulating synaptic membrane exocytosis 3	0.11	0.00
M38591	Cellular ligand of annexin II (p11)	>10 (I) P=0,083	AGGACACTTA	Hs. 143873	S100A10 S100 calcium binding protein A10 (annexin II ligand, calpactin I, light polypep-tide (p11))	12.25	0.00
J03474	Serum amyloid A (SAA)	>10 (I) P=0,001	CTCGGGGGAA	Hs.332053	SAA1 Serum amyloid A1	NaN*	0.00
L10338	Na+ channel beta-1 subunit (SCN1B)	1,7 (D) P=0,471	AAATAAAGAC	Hs.436646	SCN1B Sodium channel, voltage-gated, type I, beta	0.13	0.00
U76010	Zinc transporter ZnT-3	1,9 (D) P=0,233	TGTCTGTTTG	Hs.467981	SLC30A3 Solute carrier family 30 (zinc transporter), member 3	0.03	0.01
X54673	GAT1, GABA transporter	Very low level expression	GTCCAGCCCA	Hs.443874	SLC6A1 Solute carrier family 6 (neuro-transmitter transporter, GABA), member 1	0.07	0.00
D21267	SNAP25, highly expressed protein	>10 (D) P=0,052	TAATATTAAA	Hs.167317	SNAP25 Synaptosomal-associated protein, 25kDa	0.09	0.00
J03040	SPARC (osteonectin)	1,3 (D) P=0,762	ATGTGAAGAG	Hs. 111779	SPARC Secreted protein, acidic, cysteine-rich (osteonectin)	8.58	0.00
U20758	Osteopontin	3,4 (I) P=0,06	AATAGAAATT	Hs.313	SPP1 Secreted phosphoprotein 1 (osteopontin, bone sialoprotein I, early T-lymphocyte activation 1)	14.11	0.00
D87433	KIAA0246	>10 (I) P=0,001	CACAGGGAGG	Hs.301989	STAB1 Stabilin 1	29.22	0.00
S82024	SCG10 (neuron-specific growth-associated prot/ stathmin)	6,5 (D) P=0,067	CTGCGGAGGT	Hs.521651	STMN2 Stathmin-like 2	0.10	0.00
L37792	Syntaxin 1A	2,0 (D) P=0,351	CAGCAGGGGA	Hs.488683	STX1A Syntaxin 1A (brain)	0.11	0.00
D63851	UNC-18 homolog (sec1)	2,8 (D) P=0,167	CTTCAGGACC	Hs.288229	STXBP1 Syntaxin binding protein 1	0.07	0.00
X06389	Synaptophysin/p38	Very low level expression	GTGCAGTGAA	Hs.75667	SYP Synaptophysin	0.02	0.00
M55047	Synaptotagmin	4,5 (D) P=0,079	TTCAGTCTTC	Hs.310545	SYT1 Synaptotagmin I	0.09	0.00

Continue of table 3

Genes found by microarray technique			Same genes revealed by SAGE DGED				
Identifier/ Accession	Gene name (as it was shown by Markert et al. [18])	Fold of gene expression change in GB vs NB (I) — increase (D) — decrease	SAGE tag sequence	Unigene cluster	Gene symbol and official name	Tag Odds*(GB/ NB)	P
L10373	Clone CCG-B7 sequence	>10 (D) P=0,024	CCAACAAGAA	Hs.441664	TM4SF2 Transmembrane 4 superfamily member 2	0.31	0.00
U49869	Ubiquitin	1,8 (D) P=0,149	AAATTCTCTA	Hs.356190	UBB Ubiquitin B	14.86	0.00
M27281	Vascular permeability factor	1,9 (I) P=0,642	CAGATTATGC	Hs. 73793	VEGF vascular endothelial growth factor	8.63	0.00
D76435	Zic, zinc finger protein	>10 (D) P=0,096	TTACAGCTCT	Hs.41154	ZIC1 Zic family member 1 (odd paired homolog, Drosophila)	0.24	0.01

* The odds ratio uses a simple mathematical formula to provide a measure of the relative amount of a tag in pool of GB SAGE-libraries to pool of NB SAGE-libraries. **NaN — stands for «not a number» and occurs when the denominator of the equation is 0, i.e., there are no sequences of a gene in pool NB.

cance of $P \leqslant 0.05$. In the most cases, genes were overexpressed by more than 10-fold.

Two main questions are critical in such investigations: above which threshold has the overexpression of a particular gene to be considered as significant and relevant and what are the diagnostic, pathophysiological and therapeutical consequences of such overexpression? While there is no answer on the second question for any particular gene, every group of authors make their own decision. For example, Lal et al. [34] investigated genes that were differentially expressed by more than 5-fold with $P \leqslant 0.001$; Loging et al. [35] used the same approach and investigated genes that changed their expression 10-fold with $P \leqslant 0.001$. Markert et al. [18] discussed the possible participation in tumor initiation and progression even for those genes that differentially expressed less than 2-fold with $P \leqslant 0.05$, although they focused on genes with 5-fold change in expression. Ljubimova et al. [20] detected by GEM array a total of about 3,000 genes with changed expression in GBs. Of these genes, 14 were significantly (with ratios of $\geqslant 2$) up-regulated and 12 genes were down-regulated. Selecting a cutoff ratio of less than 5-fold change would lead to obtain more overexpressed genes, however, genes exhibiting the high differences in expression are likely more biologically relevant.

To compare SAGE results on genes, which changed their expression in GB with those obtained by microarray technique, the expression factor 2 and significance filter $P \leqslant 0.05$ were chosen because these parameters were used mostly in the microarray analyses. The comparison of the pools of 9 GB SAGE-libraries and 5-NB SAGE libraries has revealed 1,303 tags. Most part of the matching transcripts corresponded to characterized mRNA sequence entries, whereas 253 tags matched uncharacterised ESTs. When genes with no tags and tags that matched multiple genes were excluded, more than 2-fold differences of the expression were shown for 676 genes, of which 316 genes were determined as overexpressed.

The most detailed description of data, obtained by microarray analysis, was given by Markert et al. [18]. We compared their data with our SAGE results and revealed 51 common genes (Table 3). Conversion from gene names to tags, or from tags to genes was performed at the SAGEmap site (http://www.ncbi.nlm.nih.gov/SAGE/). To collate SAGE tags with GeneChip probe sets representing GenBank accession numbers, GenBank accession numbers were converted into UniGene clusters. Several genes had a good reproducibility by both methods. For example, expression of

human cartilage glicoprotein (HC gp-39) gene was increased more than 10-fold ($P = 0.038$) according microarray analysis and tag odds was 70.92 according SAGE. However, probability level $P \leqslant 0.05$ was obtained by Markert et al. [18] only for 10 of these 51 genes and there was no good correlation for several other genes with $P > 0.05$ between two methods. For example, tag odds for FN1 gene was 21.96, but in microarray analysis, its expression level was increased only in 1.2-folds ($P = 0.862$). Moreover, for several genes (PKC beta-1, SPARC and UBB) the results of microarray analysis are in contrary to SAGE: decreasing of the expression instead of increasing and vice versa.

Comparison of our SAGE results on overexpressed in GB genes shows only 3 common genes of 6 genes described by Fuller et al. [11], 8 of 49 by Sallinen et al. [12], 1 of 10 by Kim et al. [13], 12 of 34 by Godard et al. [15], 3 of 6 by Somasundaram et al. [16], 10 of 49 by Rickman et al. [17], 51 of 355 by Markert et al. [18], 6 of 7 by Van den Boom et al. [19], 12 of 14 by Ljubimova et al. [20], 4 of 36 by Tanwar et al. [21], 12 of 31 by Mischel et al. [22], 2 of 19 by Nutt et al. [23], 27 of 53 by Nigro et al. [24], 19 of 44 by Freije et al. [25], 11 of 99 genes described by Yokota et al. [26], and 2 of 8 by Liau et al. [27]. Alltogether, 105 of 849 described genes were overlapping in comparison with our results obtained by SAGE (Table 4). As it is possible to see, the main problem in evaluation of results obtained by comparison of gene expression in glioblastomas and normal brain samples is the lack of available data from each paper. The reason of poor overlapping of genes revealed by microarrays apparently is due to methodological artefacts (e.g. different gene numbers placed onto chips, poor quality of synthesised total cDNA probes or high background of hybridization patterns, problems with house-keeping gene controls, etc.) as well as to biological reasons (e.g. heterogeneity of molecular mechanisms of glioblastoma formation). A very big problem is the obtaining of normal brain samples. Mostly often, surgical specimens of histologically normal brain, adjacent to the tumor, are used as the source of NB RNA, however they can be considered as a normal control only with some precautions: gliomas are infiltrating tumors and scattered tumor cells are present far away from the dense tumor area removed during surgery.

Table 4

Common differentially expressed genes found in this work and by microarray analysis or SAGE (published data)

Gene symbol	SAGE tag	Tag Odds (GB/NB)	References	
			Microarray	SAGE
ABLIM1	CTTGTGTGTA	0.21	19	
ANXA1*	AGAAAGATGT	21.43	24	34, 35
ANXA5	ATACTTTAAT	9.35	18, 24	
APOC1*	TGGCCCCAGG	9.60	24	
AQP1	ACTGTCGCCA	NaN	18, 24	
B2M*	GTTGTGGTTA	10.25	18, 26	
BASP1	TCCGTGGTTG	0.16	22	
BZRP	GAATTTTATA	6.37	18	
CA2	ATTTCAAGAT	NaN	17	
CAPG	CTCCCCTGCC	13.72	18	34
CCND1	AAAGTCTAGA	5.83	22	
CCND2	ATATAGTCAG	4.05	22	
CD44	AAGATTGGGG	11.49	24	
CD74*	GTTCACATTA	10.29	23	34
CD99*	GGATGTGAAA	9.78	22	
CDK4	GAAGGAAGAA	6.86	17, 22	
CEBPD	GCCGCCGTGC	4.12	18	34
CHI3L1*	GTATGGGCCC	70.92	18, 21, 24	34, 36
CHI3L2**	TGGGATTCCC	NaN	15	34, 36
CNN3	ATCAGTGTGA	7.42	18	
COL1A2	GTTCCACAGA	38.09	24	
COL3A1**	GATCAGGCCA	NaN	24	
COL4A1	GACCGCAGGA	4.35	19, 20, 21	37
COL6A1	TTGCTGACTT	7.26	15	
COL6A3	ACTTTAGATG	NaN	21, 25	
CTGF	TTTGCACCTT	9.95	20	
CYFIP2	CTGGTTTCTC	0.13	18	
DCN	ACTTATTATG	9.26	25	
EGFR**	AGTACCTTAT	55.36	15, 20, 22, 24	
EMP3	CGACGAGGAG	4.77	24	
FCGR2A	TAAGTCTATA	NaN	24	
FCGR3B	GTAATAAAAT	11.32	24	
FCGRT	CTGTGAGACC	6.52	15	
FN1	ATCTTGTTAC	21.96	12, 15, 18, 20, 21, 24, 25	34, 36
GBP1	GGCAGGAGTA	NaN	24	
GJA1	TGTTCTGGAG	3.68	19	
GOT1	CACGGACACG	0.14	18	34
GRIN1	GCCCCAGCTG	0.01	18	34
GRN	GGAGGTGGGG	4.09	18, 27	
HLA-DRA*	GGGCATCTCT	12.71	14, 16	
HMOX1	CGTGGGTGGG	8.53	27	
IGFBP2*	GCCTGTACAA	7.72	11, 12, 13, 15, 17, 24, 25	
IGFBP4	ATGTCTTTTC	29.16	18, 19, 25	
IGFBP5*	GATAGCACAG	4.08	12, 18, 20	
JUN	CCTTTGTAAG	4.80	15, 18	
LGALS3*	TTCACTGTGA	12.04	24, 25	
LOX	TATGTATTTC	NaN	25	
LTF*	GCAAAACAAC	NaN	18	34

Continue of table 4

Gene symbol	SAGE tag	Tag Odds (GB/NB)	References	
			Microarray	SAGE
MAG	AAATAAATGC	0.30	20, 22	
MAL2	ATGAAAAGAA	0.09	25	
MAP1A	TGGGCTTGCC	0.19	19	
MAPT	GTAGACTCGC	0.11	18	
MBP*	TCTATTAATA	0.05	18, 22, 26	
MDK	CCCTGCCTTG	25.73	17	
MEST	CTGAATGTAC	25.05	17	
MGP	GTTTATGGAT	12.87	19, 25	36
MOBP	CAAAAAGTTA	0.04	20	
NEFH	CCGAATGCCA	0.06	18	
NNMT	GCCAACAACG	26.25	18, 24	34, 35
NPTX1	AATTGTTGAG	0.04	18	34
NRGN*	TGACTGTGCT	0.04	17, 26	
NTRK2	GAAAGTCTCT	0.05	25	
PACSIN1	ATTGTGTAAT	0.01	25	
PBEF1	GCCTTAACAA	10.18	18, 24	
PLP1	CATACATACA	0.28	20, 22	
PSD	CCAAGGCCCC	0.04	27	
PTPRZ1	TAATTTTAAC	9.61	22, 24	
QDPR	GATTGCTGGA	0.26	18, 26	
RIMS3	ATTAAAGTCA	0.11	18	
RPL27A	GAGGGAGTTT	2.34	27	
RPL39	TTACCATATC	5.98	27	
S100A10*	AGCAGATCAG	12.54	18	
S100A4*	ATGTGTAACG	NaN	25	
SAA1	ATGTGTAACG	NaN	18	36
SAT	TTTGAAATGA	8.11	26	
SEC61G*	GCAACAGCAA	16.48	24	35, 36
SEPT4	CCGGCCCCTC	0.06	20	
SNAP25	TAATATTAAA	0.08	18	
SNCA	AAACTATGCA	0.14	25	
SNCB	AATAAAGCTA	0.06	25	
SOX4	TCAAGTTCAC	5.92	24	
SPARC*	ATGTGAAGAG	8.35	12, 18	34, 35
SPP1	AATAGAAATT	13.67	18, 27	34
STAB1	TCACCAAAAA	29.85	18	36
STMN2	GTTTAAAAAG	0.11	18	
STXBP1	CTTCAGGACC	0.08	18, 26	34
SYT1	TACCCTTCTG	0.09	18, 25	34
SYT5	CTGGCCAACC	0.03	25	
TGFBI	GTGTGTTTGT	5.70	17, 24	
TIMP2	TCTCTGATGC	2.80	12	
TMSB4X	TTGGTGAAAG	20.24	27	
TM4SF2	CCAACAAGAA	0.32	18, 26	
TNC	AAGCTGTATA	10.81	15, 20, 24	36
TNFRSF1A	TTACACTAAT	11.84	12	
TOP2A	GTGCGGAGGA	NaN	17, 25	36
TPT1	TAGGTTGTCT	2.31	27	
TYMS	ATGTAGAGTG	NaN	17	36
TYROBP	AAGCACAAAA	6.47	15	
UBB	GCATTCGCAG	14.93	18	
VCAM1	GTACGGAGAT	NaN	17	

Continue of table 4

Gene symbol	SAGE tag	Tag Odds (GB/NB)	References	
			Microarray	SAGE
VEGF	TTTCCAATCT	9.26	11, 12, 15, 18, 20, 22, 24, 25	
VIM*	TCCAAATCGA	5.93	12, 15, 16, 20, 23, 24	
VSNL1	TGAAAAGTAA	0.12	25	
WWTR1	ACGTAATTAG	NaN	24	
ZIC1	TTACAGCTCT	0.26	18	
ZYX	CTGCCAAGTT	3.35	12	

Note. Parameters of DGEG analysis: total tags in Pool GB: 791621, total tags in Pool NB:271584. F (expression factor): 2X. P (significance filter): ≤ 0.05. NaN stands for «not a number» and occurs when the denominator of the equation is 0, i.e., there are no sequences of a gene in pool NB. Genes, analyzed by Northern, marked by *, analyzed by RT-PCR — marked by **.

Comparison with published results obtained by SAGE. As of May 2006, using the combination of words «SAGE and glioblastoma» we found only 10 publications. Evidently, the application of SAGE to the field of gene expression profiles for glioblastoma tumor and glioblastoma cell-culture is just a beginning. The first SAGE analysis of GB compared two GB SAGE libraries vs. two NB SAGE libraries and showed that 1.0 % (471) of the transcripts had more than a 5-fold ($P \leq 0.001$) of the 47,174 unique transcripts expressed in these two tissues [34]. Soon after this work, SAGEmap database was chosen by Loging et al. [35], who revealed 76 genes (0.16 %) as overexpressed in glioblastomas to the order of 10-fold or more and with P values <0.001. Again, authors mention only selected genes with a distinct difference in transcript levels between NB and GB samples. Three of discussed genes were described previously by Lal et al. [34]. Using the search tool DGED provided by the SAGE Genie database and the same 5-fold cutoff ratio for the comparisons of each gene, with a concomitant statistical likelihood of $P \leq 0.05$, we found only 33 differentially expressed trancripts when comparing pools of the same GB and NB SAGE libraries used by Lal et al. [34] and Lodging et al. [35]. Such large discrepancy in the number of genes that are over- or underexpressed in one sample relative to another at a given significance level can be explained by a very high dependence on the statistical methods used.

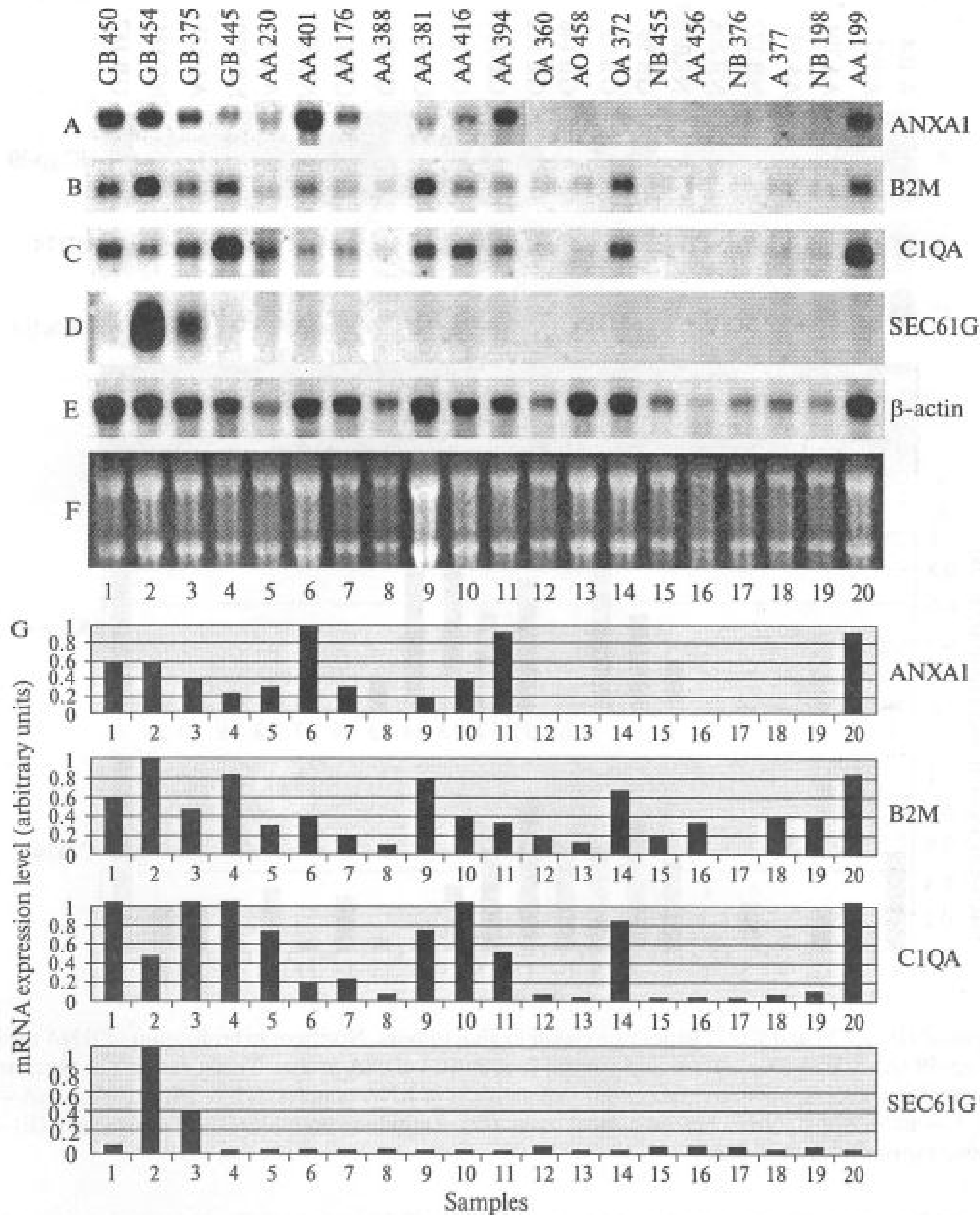

Fig. 1. Analysis of ANXA1, B2M, C1QA, and SEC61G genes expression in glial tumors. Northern hybridization of RNA panel with ^{32}P-labelled ANXA1 (A), B2M (B), C1QA (C), SEC61G (D), and control β-actin (E) cDNA probes. Tissue and tumor types are indicated above each lane of the blot, numbers are the conditional numbers of RNA samples. GB — glioblastoma, AA — anaplastic astrocytoma, A — astrocytoma, OA — oligoastrocytoma, AO — anaplastic astrocytoma, NB — human normal brain. (F) — ethidium bromide stained agarose gel. (G) — bar graph showing relative expression level of genes

Boon et al. [36] selected tags with at least a tenfold overexpression in gliomas. Despite the high heterogeneity among tumors, a small set of genes was consistently observed at high levels in more than a third of each grade of astrocytoma. Authors presented the list of 18 selected highly expressed genes in GB, only six of them were described previously by Lal et al. [34], and one gene (*SEC61G*) was described by Loging et al. [35]. Using Long SAGE libraries that contain 17 bp tags, Madden et al. [37] revealed 122 genes, which were induced (4-fold induction ratio) in the glioma endothelial

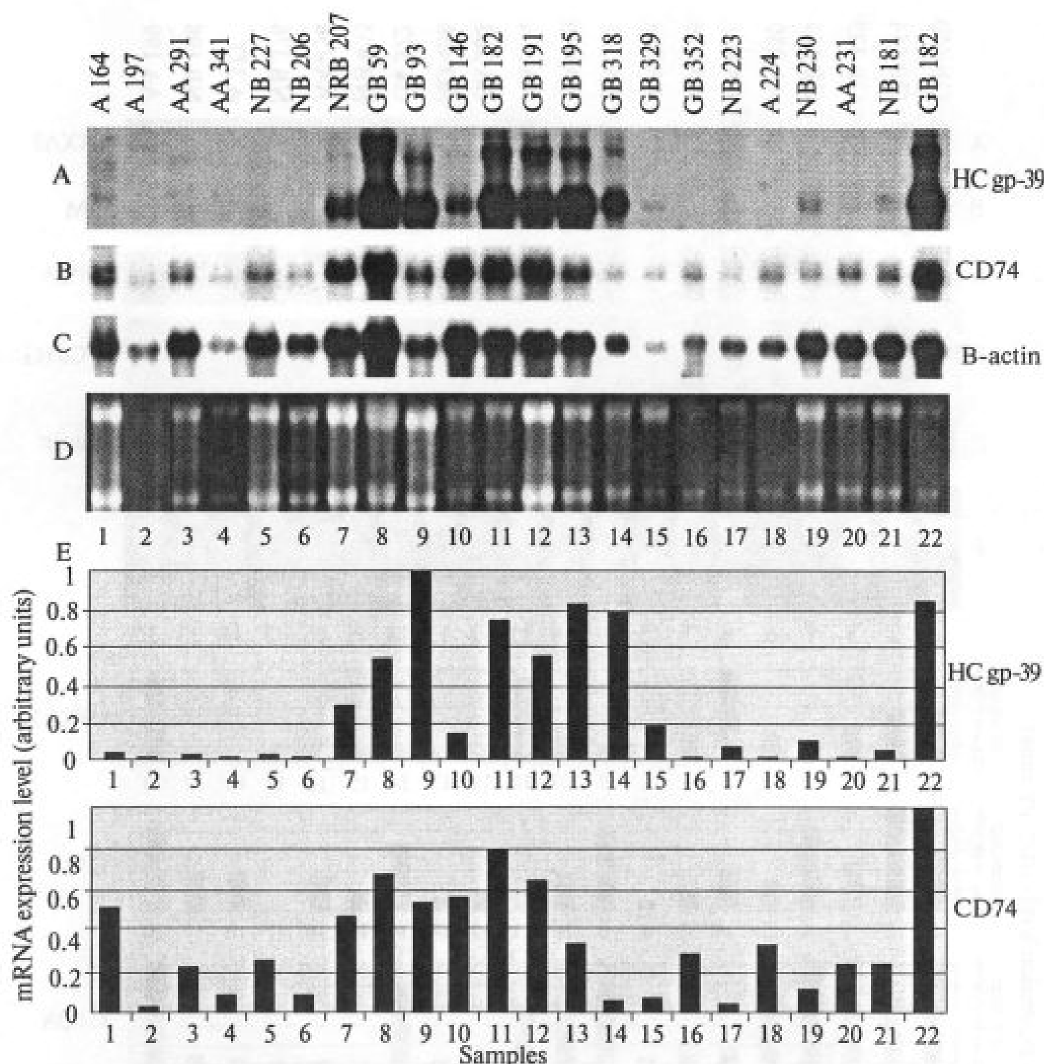

Fig. 2. Analysis of HC gp-39 and CD74 genes expression in glial tumors. Northern hybridization of RNA panel with ^{32}P-labelled HC gp-39 (A), CD74 (B), cDNA, and control β-actin (C) cDNA probes. Tissue and tumor types are indicated above each lane of the blot, numbers are the conditional numbers of RNA samples. GB — glioblastoma, AA — anaplastic astrocytoma, A — astrocytoma, NB — human normal brain. (D) — ethidium bromide stained agarose gel. (E) — bar graph showing relative expression level of genes

cells. These 122 genes were narrowed to 14 by applying additional statistical filters. Authors analysed gene expression in endothelial cells that may be an explanation, why only one gene (*COL6A2*) was common with SAGE data of other authors.

Comparison of 74 genes obtained by SAGE and described in four cited articles [34—37] with results of DGED analysis, which we performed on 9 GB and 5 NB SAGE libraries from SAGE Genie showed 32 common genes for F (expression factor): 5X and P (significance filter): 0.05, and 54 common genes for F: 2X and P: 0.05. The fact that not all 74 mentioned genes were revealed by DGED can be explained by different sets of libraries analysed in different works: Lal et al. [34] and Loging et al. [35] analyzed two GB and two NB SAGE libraries, Boon et al. [36] compared the pool of same GB SAGE libraries as we did, but to the pool of 5 NB SAGE libraries they added libraries of substantia nigra, pediatric cortex, and normal leukocytes. Loging et al. [35] and Boon et al. [36] gave genes only with increased expression in GB but did not discuss genes with down-regulated expression. In addition, the obtained differences

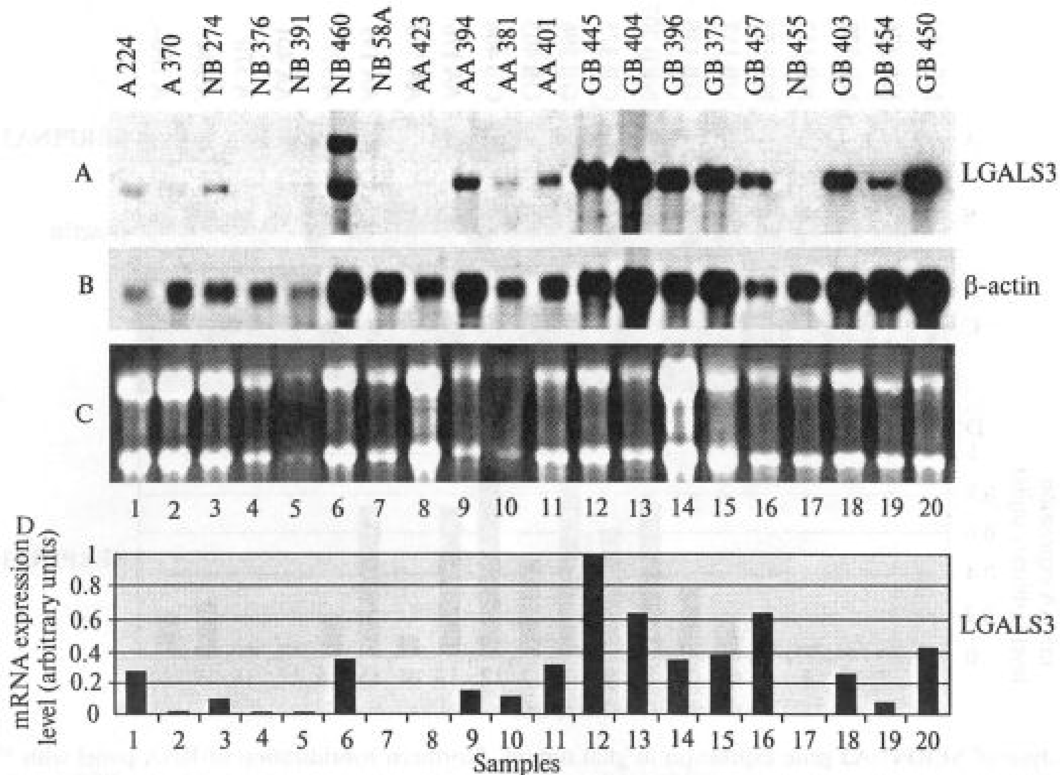

Fig. 3. Analysis of LGALS3 gene expression in glial tumors. Northern hybridization of RNA panel with ^{32}P-labelled LGALS3 cDNA (A) and control β-actin (B) cDNA probes. Tissue and tumor types are indicated above each lane of the blot, numbers are the conditional numbers of RNA samples. NB — human normal brain, A — astrocytoma, AA — anaplastic astrocytoma, GB — glioblastoma. (C) — ethidium bromide stained agarose gel. (D) — bar graph showing relative expression level of gene

may be explained by high dependence on the statistical methods used.

Apparently, the best solution of the problem is to compare all available data and to select only those genes, which significant expression in tumor combined with no detectable expression in normal tissues was reproduced in several articles. Because of the potential problems with normalization and other possible errors, it is best to base the decision to proceed with investigating a candidate tumor marker only on absolute differences in expression between tumor and normal tissues and not on small ratios of change [36]. After a gene expression profile has been obtained on a set of RNA samples the expression differences need to be confirmed and it is often useful to determine if the observation is repeatable in independent samples [38]. Northern blotting has been the gold standard for gene expression analysis for many years. To assess the reliability of expression patterns, we arbitrarily selected ten differentially expressed transcripts and evaluated them by Northern blot analysis, also allowing detection of alternative transcripts when their expression level was sufficient. Expression patterns were usually reproducible between different samples of independent set of tumors and normal tissue: genes with elevated expression in glioblastoma relative to normal brain as determined by SAGE, were detected in the most of GB samples and were expressed at considerably lower levels in the most samples of normal brain (Fig. 1—6). Differences in expression between individual tumors, exhibiting either high or low amounts of individual transcripts, undoubtedly contribute to the observed heterogeneity in the biological properties of glioblastomas.

When gene expression levels were too low to detect by Northern analysis, we used semiquantitative RT-PCR and confirmed increased expression of ABCC3, COL1A1, COL3A1, EGFR, CHI3L2, FMOD and MFAP2 genes in 30-80 % GBs (Fig. 7). In other investigations, real-time

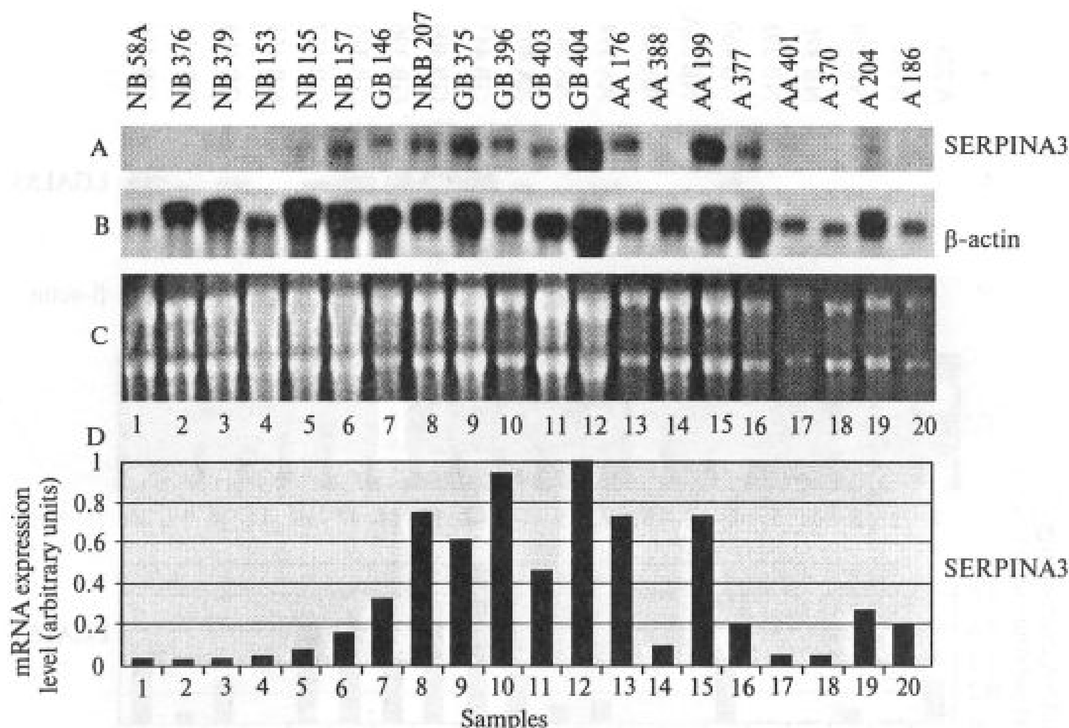

Fig. 4. Analysis of SERPINA3 gene expression in glial tumors. Northern hybridization of RNA panel with ^{32}P-labelled SERPINA3 cDNA (A) and control β-actin (B) cDNA probes. Tissue and tumor types are indicated above each lane of the blot, numbers are the conditional numbers of tumor or normal tissue samples. NB — human normal brain, A — astrocytoma, AA — anaplastic astrocytoma, GB — glioblastoma, NRB — neurolastoma. (C) — ethidium bromide stained agarose gel. (D) — bar graph showing relative expression level of gene

PCR was used to confirm the results obtained by microarray analysis for BMP2, DLL3, HDAC4, EDNRB, IP3K, RGS4, SYT1, VSNL1, MET, TOP2A, IGF2, CDC2, COL6A3, IGFBP4, LOX, THBS1genes [25] and SAGE results for ABCC3, ANXA1, GPNMB, NMB, NNMT, SEC61G [35], AQP1, TYMS, TOP2A, ABCC3, SAA1, CHI3L2, NMB, and MGP genes [36].

The confirmation of microarray or SAGE results on the protein level is advantageous, when the endpoint is knowledge of protein levels rather than mRNA levels. Unfortunately, expression changes on protein level were analyzed in reviewed publications with SAGE only for 3 genes. TOP2A and AQP1 were revealed in more than a half of GBs by immunohistochemistry [36]. Elevation of YKL-40 (CHI3L1) in 65 GBs was shown by Western blotting and as detected by ELISA analysis, YKL-40 was on substantially higher levels in serum of many GB patients [21]. We compared the production of YKL-40 protein in GB and NB by Western blotting and found much more higher level of the expression of YKL-40 in GB than in oligoastrocytoma, adult NB, fetal brain and fetal liver (Fig. 8).

Besides, we revealed larger sizes of YKL-40 bands in all three GBs analyzed in addition to 39 kDa protein described for the chondrocytes and synovial cells [39].

Conclusions. In this work we intended to find the overlapping genes with significantly enhanced expression in glioblastoma comparing the results of different scientific groups and obtained by two modern techniques — microarray-based hybridization and SAGE.

Unfortunately, authors present only the most prominent differentially expressed genes as compared to normal brain, but even comparison of available data shows quite poor overlapping of genes revealed by microarrays. The comparison of microarray analysis with SAGE showed more overlapping genes changing their expression in GB than between results obtained by microarrays due to the higher sensitivity of SAGE that permits to

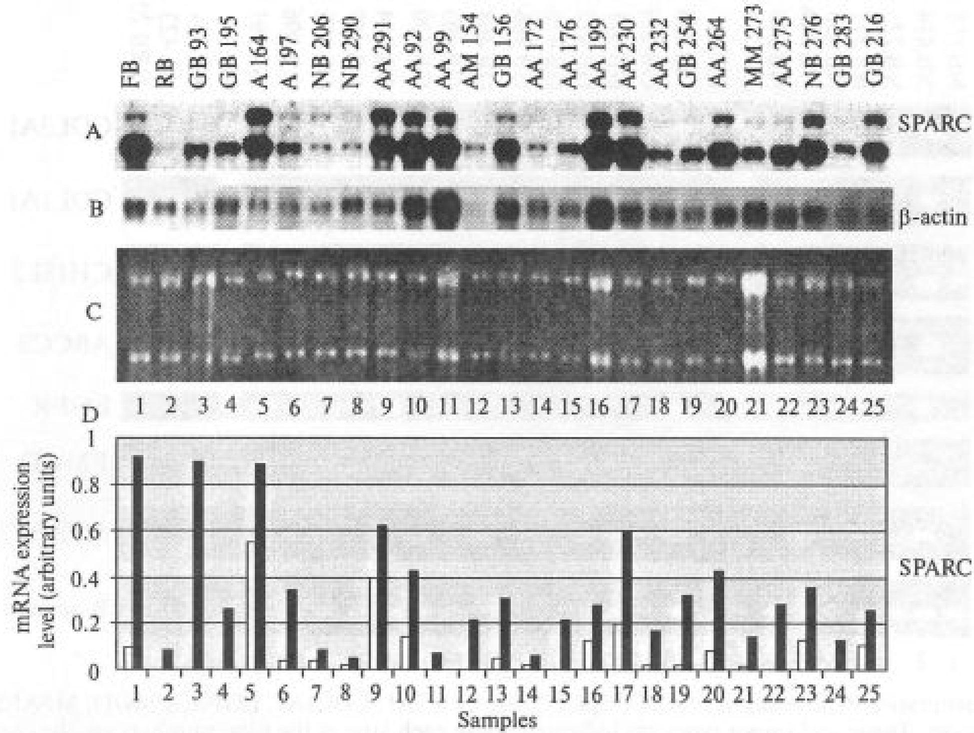

Fig. 5. Analysis of SPARC gene expression in glial tumors. Northern hybridization of RNA panel with ^{32}P-labelled SPARC cDNA (A) and control β-actin (B) cDNA probes. Tissue and tumor types are indicated above each lane of the blot, numbers are the conditional numbers of tumor or normal tissue samples. FB — human fetal brain, RB — rat brain, GB — glioblastoma, A — astrocytoma, NB — human normal brain, AA — anaplastic astrocytoma, AM — anaplastic meningioma, MM — meningioma meningothelial. (C) — ethidium bromide stained agarose gel. (D) — bar graph showing relative expression level of gene. Dark bars — small transcript, white bars — large transcript

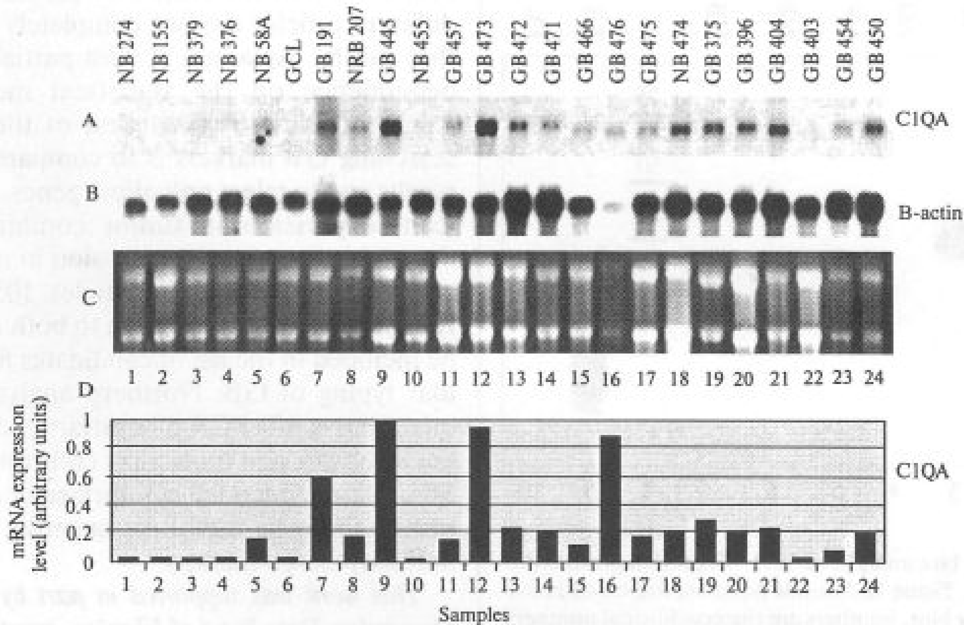

Fig. 6. Analysis of C1QA gene expression in glial tumors. Northern hybridization of RNA panel with ^{32}P-labelled C1QA cDNA (A) and control β-actin (B) cDNA probes. Tissue and tumor types are indicated above each lane of the blot, numbers are the conditional numbers of tumor or normal tissue samples. NB — human normal brain, GCL — rat glioma cell line C6, GB — glioblastoma, NRB — neuroblastoma. (C) — ethidium bromide stained agarose gel. (D) — bar graph showing relative expression level of gene

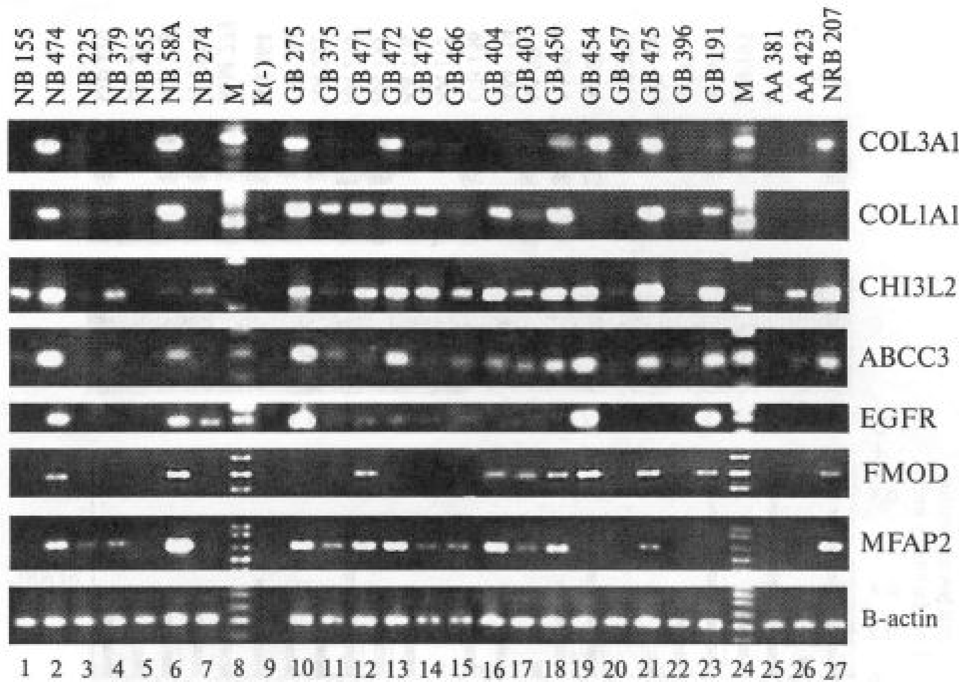

Fig. 7. Semi-quantitative PCR-analysis of ABCC3, CHI3L2, COL3A1, COL1A1, EGFR, FMOD, MFAP2 genes expression in glial tumors. Tissue and tumor types are indicated above each lane of the blot, numbers are the conditional numbers of tumor or normal tissue samples. NB — human normal brain, GB — glioblastoma, AA — anaplastic astrocytoma, NRB — neuroblastoma. K(-) — control reaction without DNA. M — marker of DNA molecular weight, GeneRuler 100bp DNA Ladder Plus (Fermentas)

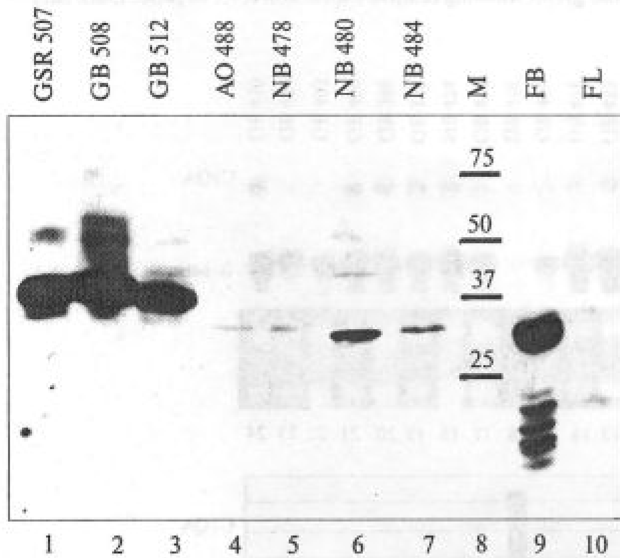

Fig. 8. Western blot analysis of YKL-40 protein production in glial tumors. Tissue and tumor types are indicated above each lane of the blot, numbers are the conditional numbers of tumor or normal tissue samples. GB — glioblastoma, AO — anaplastic oligodendroglioma, NB — human normal brain, FB — human fetal brain, 9 week old, FL — human fetal liver, 8 week old

reveal the differences even for rare transcripts. However, even the results of SAGE described in different articles are not completely reproducible that can be explained at least partially, by a high dependence on the statistical methods used. Apparently, the best solution of the problem in searching GB markers is to compare all available results and to select only those genes, which significant expression in tumor combined with no detectable or very low expression in normal tissues was reproduced in several articles. 105 differentially expressed genes, common to both methods, can be included in the list of candidates for the molecular typing of GB. Northern analysis and semi-quantitative RT-PCR of arbitrarily selected differentially expressed transcripts confirmed this result. Some genes, encoded cell surface or extra-cellular proteins may be useful for targeting gliomas with antibody-based therapy.

This work was supported in part by Fundamental Researches State Fund of Ukraine, grant F7/282-2001, and by National Academy of Sciences of Ukraine in frames of the program «Novel Problems of Medicine and Biology and Human Environment».

РЕЗЮМЕ. Для виявлення маркерів гліобластоми ми порівняли експресію генів в гліобластомі і нормальному головному мозку, використовуючи веб-сайт SAGE Genie, і співставили отримані результати з опублікованими даними. Дев'ять SAGE-бібліотек гліобластом и п'ять SAGE-бібліотек нормального головного мозку були проаналізовані з використанням програми Digital Gene Expression Displayer (DGED), результати DGED були перевірені нозерн-гібридизацією і ЗТ-ПЛР довільно відібраних генів. Обзор наявних даних із статей по профілюванню експресії генів гібридизацією микрочипів показав 35 загальних генів із підвищеною експресією в гліобластомі. Деякі з них були виявлені в чотирьох роботах, однак більшість генів знайдено тільки в трьох або навіть в двох дослідженнях. Є відмінності і в результатах досліджень з використанням техніки SAGE. Метод DGED виявив 676 диференційно експресуючих генів з понад двократною зміною експресії в гліобластомі та $P \leqslant 0.05$. Диференційна експресія відібраних генів, виявлених DGED, була підтверджена нозерн-гібридизацією і ЗТ-ПЛР. Всього тільки 105 з 955 генів, представлених в опублікованих роботах, були серед генів, знайдених DGED. Порівняння результатів аналізу мікрочипів і SAGE утруднене тим, що автори показують тільки найбільш репрезентативні гени, що диференційно експресуються, однак навіть наявні дані по аналізу мікрочипів погано перекриваються між собою. Деякі відмінності між результатами, отриманими SAGE в різних дослідженнях, можуть бути пояснені високою залежністю від статистичних розрахунків, що використовуються. Найліпшим рішенням при пошуку молекулярних пухлинних маркерів зараз може бути співставлення всіх наявних даних та відбір тільки тих генів, істотна експресія яких в пухлинах комбінується з дуже низьким рівнем експресії в нормальній тканині і репродукується в кількох роботах. Загальні для двох методів 105 генів можуть бути включені в список кандидатів для молекулярного типування гліобластом. Деякі гени, що кодують білки поверхні клітин або екстраклітинні білки, можуть бути мішенями при імунотерапії гліом.

РЕЗЮМЕ. Для выявления маркеров глиобластомы мы сравнили экспрессию генов в глиобластоме и нормальном головном мозге, используя вэб-сайт SAGE Genie, и сопоставили полученные результаты с опубликованными данными. Девять SAGE-библиотек глиобластом и пять SAGE-библиотек нормального головного мозга были проанализированы с использованием программы Digital Gene Expression Displayer (DGED), результаты DGED были проверены нозерн-гибридизацией и ОТ-ПЦР произвольно отобранных генов. Обзор имеющихся данных из статей по профилированию экспрессии генов гибридизацией микрочипов показал 35 общих генов с повышенной экспрессией в глиобластоме. Некоторые из них были выявлены в четырех работах, однако большинство генов обнаружено только в трех или даже в двух исследованиях. Имеются различия и в результатах исследований с использованием техники SAGE. Метод DGED выявил 676 дифференциально экспрессирующихся генов с более чем двукратным изменением экспрессии в глиобластоме и $P \leqslant 0.05$. Дифференциальная экспрессия отобранных генов, выявленных DGED, была подтверждена нозерн-гибридизацией и ОТ-ПЦР. Всего только 105 из 955 генов, представленных в опубликованных работах, были среди генов, обнаруженных DGED. Сравнение результатов анализа микрочипов и SAGE затруднено тем, что авторы показывают только наиболее представительные дифференциально экспрессирующиеся гены, однако даже имеющиеся данные по анализу микрочипов плохо перекрываются между собой. Некоторые различия между результатами, полученными SAGE в различных исследованиях, могут быть объяснены высокой зависимостью от используемых статистических расчетов. Наилучшим решением при поиске молекулярных опухолевых маркеров сейчас может быть сравнение всех имеющихся данных и отбор только тех генов, существенная экспрессия которых в опухолях комбинируется с очень низким уровнем экспрессии в нормальной ткани и репродуцируется в нескольких работах. Общие для двух методов 105 генов могут быть включены в список кандидатов для молекулярного типирования глиобластом. Некоторые гены, кодирующие белки поверхности клеток или экстраклеточные белки, могут быть мишенями при иммунотерапии глиом.

REFERENCES

1. *Schena M., Shalon D., Davis R.W., Brown P.O.* Quantitative monitoring of gene expression patterns with a complementary DNA microarray // Science. — 1995. — **270**. — P. 467—470.
2. *Lockhart D.J., Dong H., Byrne M.C., Follettie M.T., Gallo M.V., Chee M.S., Mittmann M., Wang C., Kobayashi M., Horton H., Brown E.L.* Expression monitoring by hybridization to high-density oligonucleotide arrays // Nat. Biotechnol. — 1996. — **14**. — P. 1675—1680.
3. *Lipshutz R.J., Fodor S.P., Gingeras T.R., Lockhart D.J.* High density synthetic oligonucleotide arrays // Nat. Genet. — 1999. — **21**(1 Suppl). — P. 20—24.
4. *Velculescu V.E., Zhang L., Vogelstein B., Kinzler K.W.* Serial analysis of gene expression // Science. — 1995. — **270**. — P. 484—487.
5. *Lash A.E., Tolstoshev C.M., Wagner L., Schuler G.D., Strausberg R.L., Riggins G.J., Altschul S.F.* SAGEmap: a public gene expression resource // Genome Res. — 2000. — **10**. — P. 1051—1060.

6. *Saha S., Sparks A.B., Rago C., Akmaev V., Wang C.J., Vogelstein B., Kinzler K.W., Velculescu V.E.* Using the transcriptome to annotate the genome // Nat. Biotechnol. — 2002. — **20**. — P. 508—512.
7. *Shostak K., Labunskyy V., Dmitrenko V., Malisheva T., Shamayev M., Rozumenko V., Zozulya Y., Zehetner G., Kavsan V.* HC gp-39 gene is upregulated in glioblastomas // Cancer Lett. — 2003. — **198**. — P. 203—210.
8. *Kavsan V., Shostak K., Dmitrenko V., Chausovskiy T., Zozulya Y., Demotes-Mainard J.* Peculiarities of molecular events in human glial tumors revealed by serial analysis of gene expression (SAGE) // Exp. Oncol. — 2004. — **26**. — P. 196—204.
9. *Dmitrenko V., Shostak K., Boyko O., Khomenko O., Rozumenko V., Malisheva T., Shamayev M., Zozulya Y., Kavsan V.* Reduction of the transcription level of the mitochondrial genome in human glioblastoma // Cancer Lett. — 2005. — **218**. — P. 99—107.
10. *Rae F.K., Stephenson S.A., Nicol D.L., Clements J.A.* Novel association of a diverse range of genes with renal cell carcinoma as identified by differential display // Int. J. Cancer. — 2000. — **88**. — P. 726—732.
11. *Fuller G.N., Rhee C.H., Hess K.R., Caskey L.S., Wang R., Bruner J.M., Yung W.K., Zhang W.* Reactivation of insulin-like growth factor binding protein 2 expression in glioblastoma multiforme: a revelation by parallel gene expression profiling // Cancer Res. — 1999. — **59**. — P. 4228—4232.
12. *Sallinen S.L., Sallinen P.K., Haapasalo H.K., Helin H.J., Helen P.T., Schraml P., Kallioniemi O.P., Kononen J.* Identification of differentially expressed genes in human gliomas by DNA microarray and tissue chip techniques // Cancer Res. — 2000. — **60**. — P. 6617—6622.
13. *Kim S., Dougherty E.R., Shmulevich I., Hess K.R., Hamilton S.R., Trent J.M., Fuller G.N., Zhang W.* Identification of combination gene sets for glioma classification // Mol. Cancer Ther. — 2002. — **1**. — P. 1229—1236.
14. *Raza S.M., Fuller G.N., Rhee C.H., Huang S., Hess K., Zhang W., Sawaya R.* Identification of necrosis-associated genes in glioblastoma by cDNA microarray analysis // Clin. Cancer Res. — 2004. — 10(Pt 1). — P. 212—221.
15. *Godard S., Getz G., Delorenzi M., Farmer P., Kobayashi H., Desbaillets I., Nozaki M., Diserens A.C., Hamou M.F., Dietrich P.Y., Regli L., Janzer R.C., Bucher P., Stupp R., De Tribolet N., Domany E., Hegi M.E.* Classification of human astrocytic gliomas on the basis of gene expression: a correlated group of genes with angiogenic activity emerges as a strong predictor of subtypes // Cancer Res. — 2003. — **63**. — P. 6613—6625.
16. *Somasundaram K., Reddy S.P., Vinnakota K., Britto R., Subbarayan M., Nambiar S., Hebbar A., Samuel C., Shetty M., Sreepathi H.K., Santosh V., Hegde A.S., Hegde S., Kondaiah P., Rao M.R.* Upregulation of ASCL1 and inhibition of Notch signaling pathway characterize progressive astrocytoma // Oncogene. — 2005. — **24**. — P. 7073—7083.
17. *Rickman D.S., Bobek M.P., Misek D.E., Kuick R., Blaivas M., Kurnit D.M., Taylor J., Hanash S.M.* Distinctive molecular profiles of high-grade and low-grade gliomas based on oligonucleotide microarray analysis // Cancer Res. — 2001. — **61**. — P. 6885—6891.
18. *Markert J.M., Fuller C.M., Gillespie G.Y., Bubien J.K., McLean L.A., Hong R.L., Lee K., Gullans S.R., Mapstone T.B., Benos D.J.* Differential gene expression profiling in human brain tumors // Phys. Genom. — 2001. — **5**. — P. 21—33.
19. *Boom J. van den, Wolter M., Kuick R., Misek D.E., Youkilis A.S., Wechsler D.S., Sommer C., Reifenberger G., Hanash S.M.* Characterization of gene expression profiles associated with glioma progression using oligonucleotide-based microarray analysis and real-time reverse transcription-polymerase chain reaction // Amer. J. Pathol. — 2003. — **163**. — P. 1033—1043.
20. *Ljubimova J.Y., Lakhter A.J., Loksh A., Yong W.H., Riedinger M.S., Miner J.H., Sorokin L.M., Ljubimov A.V., Black K.L.* Overexpression of alpha4 chain-containing laminins in human glial tumors identified by gene microarray analysis // Cancer Res. — 2001. — **61**. — P. 5601—5610.
21. *Tanwar M.K., Gilbert M.R., Holland E.C.* Gene expression microarray analysis reveals YKL-40 to be a potential serum marker for malignant character in human glioma // Cancer Res. — 2002. — **62**. — P. 4364—4368.
22. *Mischel P.S., Shai R., Shi T., Horvath S., Lu K.V., Choe G., Seligson D., Kremen T.J., Palotie A., Liau L.M., Cloughesy T.F., Nelson S.F.* Identification of molecular subtypes of glioblastoma by gene expression profiling // Oncogene. — 2003. — **22**. — P. 2361—2373.
23. *Nutt C.L., Mani D.R., Betensky R.A., Tamayo P., Cairncross J.G., Ladd C., Pohl U., Hartmann C., McLaughlin M.E., Batchelor T.T., Black P.M., von Deimling A., Pomeroy S.L., Golub T.R., Louis D.N.* Gene expression-based classification of malignant gliomas correlates better with survival than histological classification // Cancer Res. — 2003. — **63**. — P. 1602—1607.
24. *Nigro J.M., Misra A., Zhang L., Smirnov I., Colman H., Griffin C., Ozburn N., Chen M., Pan E., Koul D., Yung W.K., Feuerstein B.G., Aldape K.D.* Integrated array-comparative genomic hybridization and expression array profiles identify clinically relevant molecular subtypes of glioblastoma // Cancer Res. — 2005. — **65**. — P. 1678—1686.
25. *Freije W.A., Castro-Vargas F.E., Fang Z., Horvath S., Cloughesy T., Liau L.M., Mischel P.S., Nelson S.F.* Gene expression profiling of gliomas strongly predicts survival // Cancer Res. — 2004. — **64**. — P. 6503—6510.
26. *Yokota T., Kouno J., Adachi K., Takahashi H., Teramoto A., Matsumoto K., Sugisaki Y., Onda M., Tsunoda T.* Iden-

tification of histological markers for malignant glioma by genome-wide expression analysis: dynein, alpha-PIX and sorcin // Acta Neuropathol. — 2006. — **111**. — P. 29—38.

27. *Liau L.M., Lallone R.L., Seitz R.S., Buznikov A., Gregg J.P., Kornblum H.I., Nelson S.F., Bronstein J.M.* Identification of a human glioma-associated growth factor gene, granulin, using differential immuno-absorption // Cancer Res. — 2000. — **60**. — P. 1353—1360.
28. *Bammler T., Beyer R.P., Bhattacharya S., Boorman G.A., Boyles A., Bradford B.U., Bumgarner R.E., Bushel P.R., Chaturvedi K., Choi D., Cunningham M.L., Deng S., Dressman H.K., Fannin R.D., Farin F.M., Freedman J.H., Fry R.C., Harper A., Humble M.C., Hurban P., Kavanagh T.J., Kaufmann W.K., Kerr K.F., Jing L., Lapidus J.A., Lasarev M.R., Li J., Li Y.J., Lobenhofer E.K., Lu X., Malek R.L., Milton S., Nagalla S.R, O'malley J.P., Palmer V.S., Pattee P., Paules R.S., Perou C.M., Phillips K., Qin L.X., Qiu Y., Quigley S.D., Rodland M., Rusyn I., Samson L.D., Schwartz D.A., Shi Y., Shin J.L., Sieber S.O., Slifer S., Speer M.C., Spencer P.S., Sproles D.I., Swenberg J.A., Suk W.A., Sullivan R.C., Tian R., Tennant R.W., Todd S.A., Tucker C.J., Van Houten B., Weis B.K., Xuan S., Zarbl H.*; Members of the Toxicogenomics Research Consortium. Standardizing global gene expression analysis between laboratories and across platforms // Nat. Methods. — 2005. — **2**. — P. 351—356.
29. *Irizarry R.A., Warren D., Spencer F., Kim I.F., Biswal S., Frank B.C., Gabrielson E., Garcia J.G., Geoghegan J., Germino G., Griffin C., Hilmer S.C., Hoffman E., Jedlicka A.E., Kawasaki E., Martinez-Murillo F., Morsberger L., Lee H., Petersen D., Quackenbush J., Scott A., Wilson M., Yang Y., Ye S.Q., Yu W.* Multiple-laboratory comparison of microarray platforms // Nat. Methods. — 2005. — **2**. — P. 345—350.
30. *Larkin J.E., Frank B.C., Gavras H., Sultana R., Quackenbush J.* Independence and reproducibility across microarray platforms // Nat. Methods. — 2005. — **2**. — P. 337—344.
31. *Petersen D., Chandramouli G.V.R., Geoghegan J., Hilburn J., Paarlberg J., Kim C.H., Munroe D., Gangi L., Han J., Puri R., Staudt L., Weinstein J., Barrett J.C., Green J., Kawasaki E.S.* Three microarray platforms: an analysis of their concordance in profiling gene expression // BMC Genomics. — 2005. — **6**. — P. 63.
32. *Harris L.* Data stands up to scrutiny // The scientist. — 2005. — **18**. — P. 32—40.
33. *Boon K., Osorio E.C., Greenhut S.F., Schaefer C.F., Shoemaker J., Polyak K., Morin P.J., Buetow K.H., Strausberg R.L., De Souza S.J., Riggins G.J.* An anatomy of normal and malignant gene expression // Proc. Nat. Acad. Sci. USA. — 2002. — **99**. — P. 11287—11292.
34. *Lal A., Lash A.E., Altschul S.F., Velculescu V., Zhang L., McLendon R.E., Marra M.A., Prange C., Morin P.J., Polyak K., Papadopoulos N., Vogelstein B., Kinzler K.W., Strausberg R.L, Riggins G.J.* A public database for gene expression in human cancers // Cancer Res. — 1999. — **59**. — P. 5403—5407.
35. *Loging W.T., Lal A., Siu I.M., Loney T.L., Wikstrand C.J., Marra M.A., Prange C., Bigner D.D., Strausberg R.L., Riggins G.J.* Identifying potential tumor markers and antigens by database mining and rapid expression screening // Genome Res. — 2000. — 10. — P. 1393—1402.
36. *Boon K., Edwards J.B., Eberhart C.D., Riggins G.J.* Identification of astrocytoma associated genes including cell surface markers // BMC Cancer. — 2004. — **4**. — P. 39.
37. *Madden S.L., Cook B.P., Nacht M., Weber W.D., Callahan M.R., Jiang Y., Dufault M.R., Zhang X., Zhang W., Walter-Yohrling J., Rouleau C., Akmaev V.R., Wang C.J., Cao X., St Martin T.B., Roberts B.L., Teicher B.A., Klinger K.W., Stan R.V., Lucey B., Carson-Walter E.B., Laterra J., Walter K.A.* Vascular gene expression in non-neoplastic and malignant brain // Amer. J. Pathol. — 2004. — **165**. — P. 601—608.
38. *Cerutti J.M., Riggins G.J., de Souza S.J.* What can digital transcript profiling reveal about human cancers? // Braz. J. Med. Biol. Res. — 2003. — **36**. — P. 975—985.
39. *Hakala B.E., White C., Recklies A.D.* Human cartilage gp-39, a major secretory product of articular chondrocytes and synovial cells, is a mammalian member of a chitinase protein family // J. Biol. Chem. — 1993. — **268**. — P. 25803—25810.

Received 04.07.06