

# Using Serial Analysis of Gene Expression to identify tumor markers and antigens

Gregory J. Riggins

*Duke University Medical Center, Durham, NC 27710, USA*

*Tel: +1 919 684 3250; Fax: +1 919 681 2796;*

*E-mail: [greg.riggins@duke.edu](mailto:greg.riggins@duke.edu)*

Tumor markers and antigens are normally highly expressed in malignant tissue, but not in the surrounding normal tissue. Serial Analysis of Gene Expression (SAGE) is a technology that counts mRNA transcripts and can be used to find those genes most highly induced in malignant tissues. SAGE produces a comprehensive profile of gene expression and can be used to search for tumor biomarkers in a limited number of samples. Public sources of SAGE data, in particular through the Cancer Genome Anatomy Project, increase the value of this technology by making a large source of information on many tumors and normal tissues available for comparison. Although the perfect tumor-specific gene does not exist, the differences in gene expression between tumor and normal can be exploited for therapeutic or diagnostic purposes.

## 1. Introduction

During tumor growth, the pattern of expressed genes in the tumor diverges from that of the surrounding normal tissue. Simple knowledge of which tumor genes are induced compared with normal tissues can aid in locating clinically useful markers or antigens, even though the molecular basis of the altered expression is typically unknown.

Frequently, genes over-expressed in tumors have been sought as a marker for early detection. Tumor markers are used to indicate the presence of cancer, or to follow response to therapy. Typically tumor markers are assayed by the detection of protein from serum or other accessible body fluids or tissues. Tumor markers are clearly useful, but there is a lack of good markers for most cancers where early detection is warranted.

Proteins over-expressed relative to normal tissue have a second important practical use as Tumor Specific Antigens (not present in any normal tissue) or Tumor

Associated Antigens (expression in some normal cells). Tumor antigens may indeed be the same protein as a tumor marker, but their purpose is therapeutic rather than diagnostic. Toxic antibodies immunized to tumor antigens on the cell-surface or in the extra-cellular matrix may kill enough cancer cells to be therapeutic [21]. This approach ideally requires the cell surface protein to be uniquely expressed on all the tumor cells, but not expressed in any normal cells that would be in contact with the antibody during treatment. Also promising is a 'tumor vaccine' approach where the goal is to direct immune defenses toward the tumor by 'educating' host cells with tumor-derived material [4]. Expression of the marker on the cell surface is not a requirement of this system, but successful systemic administration of a tumor vaccine may require a much higher expression of the tumor antigen in the tumor compared to vital cells throughout the body. Either of these immune-based therapies would benefit from the discovery of new tumor antigens.

Another class of differentially expressed genes in cancer is prognostic markers. Recently, many groups have sought to classify tumors by gene expression pattern, in addition to histopathology. The introduction of large-scale gene expression analysis has been used successfully to classify tumors by RNA expression patterns, in particular using DNA microarrays. This has helped further separate aggressive from non-aggressive tumors and in some cases, help predict response to therapy.

Tumor markers, tumor antigens and prognostic markers (cancer biomarkers) have great potential for clinical application, but there has been a lack of high-quality markers for the various types of cancers [38]. Finding a candidate marker has frequently been the by-product of other studies and not the initial intent of the research. Furthermore, generating the expression profile for each suspect gene has often relied on time consuming techniques, such as Northern Blotting, *in situ* hybridization, or immunohistochemistry. Fortunately, the advent of large-scale gene expression analysis and information technology have accelerated

biomarker discovery. In particular, Serial Analysis of Gene Expression (SAGE) [33] can decipher complex expression patterns and be helpful for locating biomarkers.

This review will focus on the use of SAGE for locating clinically useful cancer-induced genes. SAGE is a technology that identifies and counts most all expressed transcripts from an RNA sample. Due to the quantitative and comprehensive nature of the data, it is particularly good for locating tumor markers. The transcripts found by SAGE to be induced to high level normally expresses the coded proteins at high levels, making this approach useful for locating candidate tumor antigens, as well. There are several examples that have been published already using SAGE for locating cancer markers and antigens, but the potential for this technology to discover disease biomarkers is just being realized.

## 2. What is SAGE?

Serial Analysis of Gene Expression (SAGE) is a sequence-based approach that produces and counts the transcripts expressed within a group of cells [35]. With SAGE a 10 base-pair 'tag' sequence plus a specific four base-pair restriction site is used to distinguish and identify transcripts. These tags are ligated and cloned into a sequencing vector, allowing the serial analysis of multiple transcripts using an automated sequencer. The number of times a particular tag is observed in a tag population made from one mRNA sample (SAGE library), is used to determine the relative abundance of each transcript in the mRNA sample. The counts of each transcript are stored on a computerized database and are used to make statistical comparisons between libraries.

Numerous reviews have been published describing the technology [13,17,24,34]. Many studies have been performed generating a comprehensive analysis from a diversity of tissues, in particular malignant tissues. A detailed protocol can be obtained through the SAGE Home Page from the Johns Hopkins Oncology Center (<http://www.sagenet.org>). The technology is patented by Johns Hopkins University and licensed to Genzyme Molecular Oncology (Framingham, MA) but freely available to academia and non-profit organizations for research purposes. Further information on the license agreements for commercial applications can be obtained directly from Genzyme ([www.genzyme.com/sage/welcome.htm](http://www.genzyme.com/sage/welcome.htm)).

An outline of SAGE is shown in Fig. 1. The construction of a SAGE library starts with the purification of polyadenylated mRNA which subsequently is converted into double stranded cDNA using a biotin-labeled oligo-dT primer during the first strand synthesis for the recovery of the cDNA. A frequent-cutting anchoring enzyme, usually *Nla* III, defines the position in the transcripts from which the sequence tags are derived. After digestion of the generated cDNA with the anchoring enzyme, the 3'-terminal cDNA fragments are bound to streptavidin-coated beads. Next, an oligonucleotide linker containing recognition sites for a type II restriction enzyme (tagging enzyme) is ligated to the bound cDNA fragments. Type II enzymes cut at a defined distance away from the recognition site. The SAGE tags are then released from the bound cDNA by cleavage with the tagging enzyme (usually *Bsm*FI) and dimerized by a tail-to-tail blunt-end ligation. The linked ditags (102 bp) are then amplified and digested with the same anchoring enzyme used for the initial digestion of the double stranded cDNA. The resulting ditag fragments of 26–28 bp lengths all have cohesive ends and can therefore be linked together to form concatemers and cloned into a plasmid for sequencing.

Modifications to the SAGE library construction have been made to allow libraries to be constructed with smaller starting samples. A 'microSAGE' procedure reported to assay small numbers of purified endothelial cells has worked robustly in our hands and other laboratories [30]. This procedure starts with total RNA or cell lysates and conjugates the mRNA in the sample to magnetic beads, where subsequent cDNA synthesis and digests are performed. This increase in efficiency allows libraries to be made from as little as 1  $\mu$ g of total RNA.

The final experimental step in the SAGE analysis is sequencing of the library. The data files generated by automated sequencing of the plasmids with ligated SAGE tags are analyzed by software that extracts and counts the occurrence of each tag. In a typical SAGE experiment, sequencing 2,000 inserts can produce over 50,000 tags, producing a sensitive level of detection. Cumulative databases can be formed, and bioinformatics applied to find transcripts with a particular pattern of expression. Therefore, SAGE data derived from a series of tumor and normal tissues can be queried directly for transcripts that are highly expressed in tumor, but not in normal tissue.

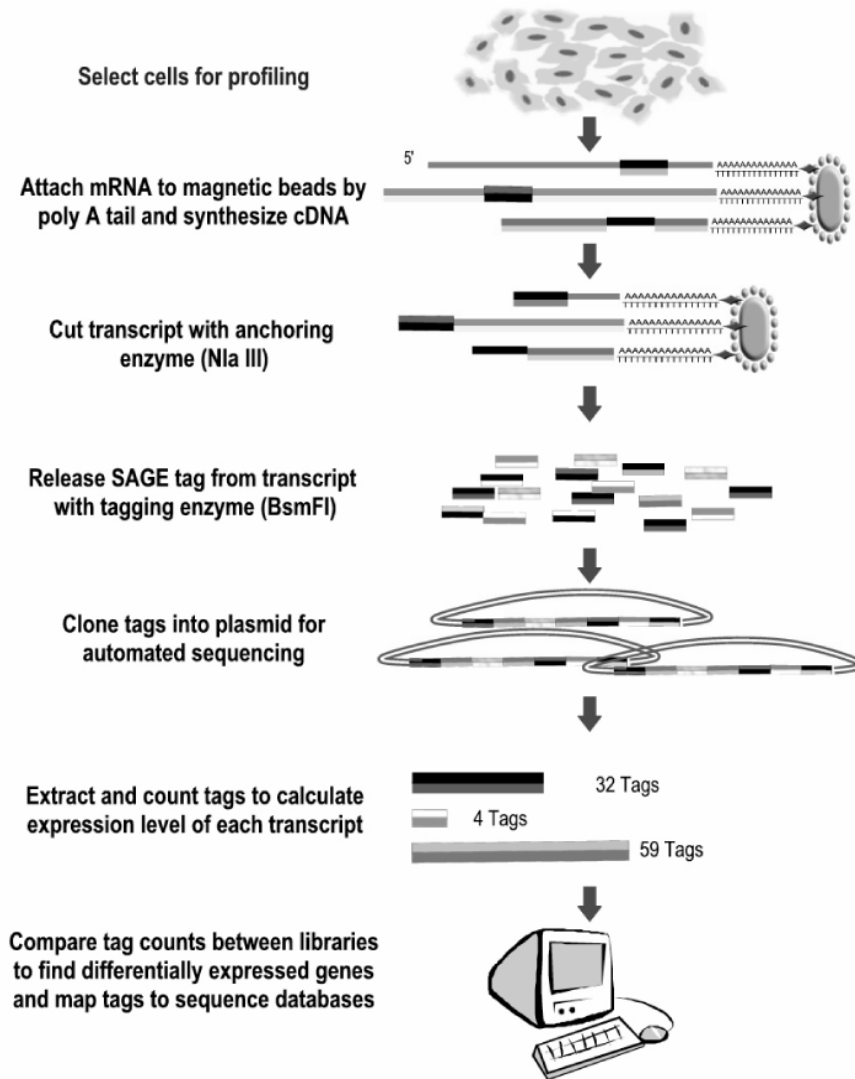


Fig. 1. Outline of SAGE approach for transcript profiling.

### 3. Assessment of SAGE technology

Since SAGE identifies and counts transcripts by nucleic acid sequence, it is frequently regarded as an accurate means for large-scale expression profiling. SAGE transcript levels are expressed as a fraction of the total transcripts counted, not relative to another experiment, standard, or a housekeeping gene, avoiding error prone normalization between experiments. The standardized nature of SAGE data makes cumulative data sets possible and historical comparisons valid. An additional strength of SAGE is that it determines expression levels directly from an RNA sample. It is not necessary to have a DNA probe arrayed to assay each gene, as with

chip technology. This allows SAGE to identify genes that are not included in an array [23] and avoids the infrastructure necessary to create and read large DNA arrays.

This flexibility of SAGE has some disadvantages. The number of samples that can be processed using SAGE is small compared to DNA arrays since it takes about two weeks of skilled labor to construct a SAGE library. Analysis of hundreds of samples by SAGE for a single laboratory is not a practical option for the technology in its present form. However, when an in-depth and quantitative profile is desired for a small number of samples, the extra work involved in creating a SAGE library can certainly be justified. To date, SAGE has

been successful for determining the differentially expressed transcripts in well-controlled experimental systems [1,6–8,10,16,23,39]. This type of data generated by SAGE is often complementary to a typical use of DNA arrays in cancer research for a wide survey of many patient tumor samples. However, the increasing amount of public SAGE data makes it possible to rapidly build upon the work of others to locate the genes of interest.

#### 4. Public sources of SAGE data

One advantage of SAGE technology is that public data that can be easily downloaded or queried online. Links to SAGE data, and SAGE resources are listed in Table 1. This data is a valuable resource for comparing internally generated expression data, or for mining of novel cancer biomarkers [15,28].

The Cancer Genome Anatomy Project (CGAP) specializes in creating databases and resources for cancer research [26,31,32] and has produced this large source of public SAGE information. CGAP adopted Serial Analysis of Gene Expression (SAGE) technology starting in 1998 with the introduction of the SAGEmap web site [12]. On-line tools built specifically to handle SAGE data [12,14] allow users to make statistical-based comparisons between libraries to find differentially expressed genes using the 'xProfiler', or by downloading data for local analysis. SAGE tags can be 'mapped' to UniGene clusters via SAGEmap, making the identification of a gene from a differentially expressed tag easier. However, the UniGene mapping is not always accurate and efforts are underway to produce a more accurate mapping of tags to transcript data. The SAGE data generated through this project is also used to create a 'Digital Northern' tool, where the expression level of a particular gene can be determined for each of the tissues used to make SAGE libraries. To date, over four million valid transcript tags have been processed from nearly 100 different malignant and normal cell types.

#### 5. Bioinformatics

For handling SAGE data most investigators rely on the SAGE software generated by Ken Kinzler and co-workers to process raw data. This software extracts tag sequences from raw sequence data and tabulating the counts in a database. The software also will

make comparisons between libraries of tags and calculate the statistical significance of differences based on Monte-Carlo simulations [40]. Additionally, the software helps create a relational database by extracting tags, gene name and gene information from sequence database. The program uses this information to match tags to known genes or ESTs. Additional 'tag to gene' mapping information can be downloaded from the NCBI from the SAGEmap website (Table 1). The SAGE software is freely available to non-commercial users of the technology and can be obtained via SAGEnet (Table 1). Investigators who plan a use of SAGE technology for commercial purposes should contact Genzyme Molecular Oncology for a license agreement.

#### 6. Verification of genes identified by SAGE

After a gene expression profile has been obtained on a set of RNA samples, it is desirable to experimentally confirm the expression differences and to extend the analysis to other samples. Normally a small set of interesting genes has been identified using DNA arrays or SAGE, but several different techniques are more efficient for assaying this smaller set of genes. In addition, each gene expression technique has inherent errors and an independent method is required for validating the original expression levels.

Although Northern Blotting is a time-consuming approach, it is still a useful and accurate way to confirm profiling data for a limited number of genes. When a good antibody is available for the gene of interest, a western blot or immunohistochemistry are reliable methods for confirming expression changes. This approach is advantageous; in particular when the endpoint is knowledge of protein levels rather than mRNA levels.

Real-Time PCR, sometimes called 'quantitative' or 'fluorescent' PCR, has gained popularity for rapid follow-up and confirmation of profiling data [15,25]. Expression determination by real-time PCR is based on continuous fluorescent monitoring of PCR products [20,36,37] from a reverse transcriptase-generated cDNA template. The number of cycles required to PCR-amplify a product to a certain level is proportional to the amount of starting template and can be used to accurately determine starting mRNA levels. Normally a serially diluted known sample is used for a standard curve to interpolate concentrations of unknown samples.

Table 1  
Links to SAGE data and information

Web site	URL	Description
SAGEmap	<a href="http://www.ncbi.nlm.nih.gov/SAGE">www.ncbi.nlm.nih.gov/SAGE</a>	Large RNA expression database from CGAP based on SAGE profiles of malignant and normal cells.
SAGEnet (Kinzler-Vogelstein Lab)	<a href="http://www.sagenet.org">www.sagenet.org</a>	SAGE database, protocols, references and additional links.
Genzyme Molecular Oncology	<a href="http://www.genzyme.com/sage/welcome.htm">www.genzyme.com/sage/welcome.htm</a>	SAGE information and applications for commercial users of the technology.
Cancer Genome Anatomy Project (CGAP)	<a href="http://cgap.nci.nih.gov">cgap.nci.nih.gov</a>	CGAP homepage with links to expression databases and cancer research resources.

To look at protein levels of many samples simultaneously, a tissue microarray system has been developed [11,19,29]. This system allows for up to one thousand small tissue samples, made from a narrow gauge biopsy needle, to be arrayed in a single block of tissue. This block of tissue can that be used to produce hundreds of slides that can be probed by immunohistochemistry or other means. In this way a standard set of the same samples can be probed for expression levels for many different genes. A digital imaging system is used to record and read the data. Although, robotics are now employed to array the tissues, many good quality samples must be collected and oriented for biopsy in the region of interest oriented by a pathologist. The results must also be scored in some fashion by signal intensity, done manually at this point in the technologies development. Finally, a good antibody is needed for each gene of interest that will work in the fixed tissue. This approach has the potential to make gene expression correlations with a vast archive of preserved tumor material.

## 7. Examples of cancer biomarker discovery

### 7.1. Colon cancer

The first application of SAGE to human tissues was to colon cancer [40]. Comparing colon tumors to normal colon epithelium showed that less than 1.5% of the transcripts were differentially expressed. Many genes elevated in cancer represented products known to be involved in growth and proliferation, while genes found in the normal colon were often related to differentiation. SAGE was also used more recently to locate candidate biomarkers for metastasis in colon cancer, using cell lines as a model [22].

### 7.2. Ovarian cancer

Ovarian cancer treatment would benefit from tumor markers capable of early detection, since most ovarian cancers have already metastasized at the time of diagnosis. In order to locate ovarian cancer markers, a total of 385,000 transcripts from ten different ovarian libraries were analyzed by SAGE [9]. From this data transcripts were identified that were high in all three primary ovarian cancers and low in all three non-malignant specimens. A total of 27 genes were identified that met these criteria and that were over-expressed more than 10-fold in ovarian tumors. Interestingly, a majority of those genes were predicted to encode membrane or secreted proteins, making them candidates for biomarkers or tumor targeting. Many of these secreted genes encoded protease inhibitors.

### 7.3. Tumor vascular endothelium

Endothelial cells provide the blood supply to solid tumors and are intimately involved in supporting their growth. The tumor antigens located on tumor endothelial cells could provide an excellent target for anti-tumor therapy. SAGE was used to identify genes differentially expressed between the endothelial cells from either normal colon or colon adenocarcinoma [30]. The study detected 79 different genes differentially expressed between these tissues, including 46 that were specifically elevated in tumor-associated endothelial cells. On the basis of these results, it was suggested that endothelium growing in a tumor is more like developing endothelium, and that these differences may be clinically relevant. Nine SAGE tags elevated in the tumor corresponded to novel, uncategorized genes. These genes were named tumor endothelial marker (TEM), and designated TEM-1 to TEM-9. Further experiments confirmed the tumor endothelium-specific expression of these genes, not only for colorectal tumors but also

for other major tumor types. These TEMs or other genes identified in this study may become targets of anti-angiogenic therapies.

#### 7.4. Brain cancer

SAGE has been used to study the most common adult malignant brain tumor, Glioblastoma Multiforme (GBM). The first SAGE analysis of GBM compared over 200,000 transcript tags from primary GBMs and normal brain cortex [12]. Approximately 1% of the genes detected were differentially expressed and included angiogenesis factors such as vascular endothelial growth factor, cell cycle regulators, and transcription factors. This data was also used by the Cancer Genome Anatomy Project to help start the public SAGEmap database and is available online at this site. Cancer-induced genes mined from this data were further tested using real-time PCR, western and Northern Blotting to see if candidate tumor marker could be located [15]. Most of the tumor over-expressed genes predicted by SAGE could be confirmed in a subset of glioblastoma. In general, a particular antigen was only highly expressed at most in about one-third of the GBMs tested, likely due to the molecular heterogeneity of this cancer. However, in combination, 75% of the tumors had at least one antigen that was strongly expressed, and not present in a panel of normal neural tissues. Two antigens were located that coded for cell surface proteins, and may be useful for targeting gliomas with antibody based therapy.

Investigators have also used SAGE to study a rat C6 glioma cell model [5]. Over-expressed genes were found that were related to invasion and cell-surface interactions. The SAGE results were confirmed by Northern Blotting.

Brain tumors other than GBM have been studied by expression profiling. The major malignant pediatric brain tumor, medulloblastoma, has been studied by SAGE [18]. Detailed SAGE expression profiles are also available for medulloblastomas and a variety of gliomas at the CGAP SAGEmap database [12].

#### 7.5. Pancreatic cancer

Pancreatic cancers are another major tumor that would greatly benefit by having an effective means for early detection. SAGE was used early on to profile the genes expressed in pancreatic cancer, although it was not possible to perform a SAGE analysis on the corresponding normal pancreatic ductal epithelium [40].

Despite this limitation, an effective tumor marker was found for pancreatic cancer, TIMP-1, in particular when it was used in conjunction with CA19-9 and carcinoembryonic antigen [41]. Clustering algorithms, first developed for DNA array data, were applied to SAGE expression data of pancreatic cancer [27]. In this study a group of invasion and metastasis specific genes were identified that may be useful as diagnostic or therapeutic markers for pancreatic cancer.

#### 7.6. Breast cancer

SAGE was used to compare breast carcinoma cells and normal breast epithelium [2]. The gene, 14-3-3 final sigma was found to be reproducibly repressed in breast carcinoma. The response of breast cancer cell lines to the effects of estrogen has also been studied using SAGE [1,10]. Among the genes found were WISP-2 (a Wnt-1 inducible signaling protein), and five novel genes (E2IG1-5).

### 8. Discussion and future goals

SAGE is currently one of the most useful methods for profiling as many of the expressed transcripts in a population of cells as possible. It provides perhaps the best chance to obtain an accurate and comprehensive picture of expressed transcripts in a particular tissue, although the technique is time consuming and laborious for multiple samples. Fortunately, the growing amount of public data makes it possible to search for candidate tumor biomarkers directly, or to augment private datasets with public data.

The first challenge is to determine how to find from complex gene expression data the best candidates for a tumor markers or antigens. Improved bioinformatics and computational methods allow the data to be queried more easily, but there is still much progress necessary to be able to integrate SAGE and other sources of molecular information in a meaningful way. Validation of candidate biomarkers at the RNA level is now much quicker with the use of real-time PCR techniques. The application of *in situ* hybridization or immunohistochemistry can be used to determine if all cells within a tumor are expressing the marker or if there is some small population of normal cells that highly expresses the gene of interest. When it is necessary to screen large sample sets for protein levels, immunohistochemistry using tissue microarrays can provide a rapid means [11]. Various improvements in proteomic

technology may also eventually provide a means to assay proteins on a level as comprehensive as currently available for mRNA [3].

One overall conclusion that can be made from gene expression profiling of cancer is that tumors, even with identical histopathology, are very heterogeneous at the expression level. Although this makes it difficult to select a molecular-based therapy using just histology, application of gene expression measurements to clinical samples will make it possible to identify the tumor that expresses the antigens for which a therapy is available.

The rate-limiting step for tumor marker application or discovery is still the work required to show that the marker will be clinically useful. It is, therefore, important that the best candidate markers or antigens can be predicted with some degree of accuracy from gene expression data. It still remains to be seen if the candidate markers or antigens discovered initially by SAGE will produce useful clinical tests or therapies. Although this process will take several or many years, it seems appropriate to use the most comprehensive data sets possible and careful validation of the candidates prior to embarking on the laborious task of further developing a tumor specific gene for clinical use.

## References

- [1] A.H. Charpentier, A.K. Bednarek, R.L. Daniel et al., Effects of estrogen on global gene expression: identification of novel targets of estrogen action, *Cancer Res* **60** (2000), 5977–5983.
- [2] A.T. Ferguson, E. Evron, C.B. Umbricht et al., High frequency of hypermethylation at the 14-3-3 sigma locus leads to gene silencing in breast cancer, *Proc Natl Acad Sci USA* **97** (2000), 6049–6054.
- [3] D. Figeys and D. Pinto, Proteomics on a chip: promising developments, *Electrophoresis* **22** (2001), 208–216.
- [4] E. Gilboa, S.K. Nair and H.K. Lyerly, Immunotherapy of cancer with dendritic-cell-based vaccines, *Cancer Immunol Immunother* **46** (1998), 82–87.
- [5] J.M. Gunnarsen, V. Spirkoska, P.E. Smith, R.A. Danks and S.S. Tan, Growth and migration markers of rat C6 glioma cells identified by serial analysis of gene expression, *Glia* **32** (2000), 146–154.
- [6] T.C. He, C. Rago, H. Hermeking et al., Identification of c-MYC as a target of the APC pathway, *Science* **281** (1998), 1509–1512.
- [7] T.C. He, T.A. Chan, B. Vogelstein and K.W. Kinzler, PPARdelta is an APC-regulated target of nonsteroidal anti-inflammatory drugs, *Cell* **99** (1999), 335–345.
- [8] H. Hermeking, C. Lengauer, K. Polyak et al., 14-3-3 sigma is a p53-regulated inhibitor of G2/M progression, *Mol Cell* **13** (1997), 3–11.
- [9] C.D. Hough, C.A. Sherman-Baust, E.S. Pizer et al., Large-scale serial analysis of gene expression reveals genes differentially expressed in ovarian cancer, *Cancer Res* **60** (2000), 6281–6287.
- [10] H. Inadera, S. Hashimoto, H.Y. Dong et al., WISP-2 as a novel estrogen-responsive gene in human breast cancer cells, *Biochem Biophys Res Commun* **275** (2000), 108–114.
- [11] J. Kononen, L. Bubendorf, A. Kallioniemi et al., Tissue microarrays for high-throughput molecular profiling of tumor specimens, *Nat Med* **4** (1998), 844–847.
- [12] A. Lal, A.E. Lash, S.F. Altschul et al., A public database for gene expression in human cancers, *Cancer Res* **59** (1999), 5403–5407.
- [13] A. Lal, I.-M. Sui and G. Riggins, Serial Analysis of Gene Expression: Probing transcriptomes for molecular targets, *Current Opinion in Molecular Therapeutics* **1** (1999), 720–726.
- [14] A.E. Lash, C.M. Tolstoshev, L. Wagner et al., SAGEmap: A public gene expression resource, *Genome Res* **10** (2000), 1051–1060.
- [15] W.T. Loding, A. Lal, I.-M. Siu et al., Identifying potential tumor markers and antigens by database mining and rapid expression screening, *Genome Res* **10** (2000), 1393–1402.
- [16] S.L. Madden, E.A. Galella, J. Zhu, A.H. Bertelsen and G.A. Beaudry, SAGE transcript profiles for p53-dependent growth regulation, *Oncogene* **15** (1997), 1079–1085.
- [17] S.L. Madden, C.J. Wang and G. Landes, Serial analysis of gene expression: from gene discovery to target identification, *Drug Discov Today* **5** (2000), 415–425.
- [18] E.M.C. Michiels, E. Oussoren, M. Van Groenigen et al., Genes differentially expressed in medulloblastoma and fetal brain, *Physiol Genomics* **1** (1999), 83–91.
- [19] H. Moch, P. Schraml, L. Bubendorf et al., High-throughput tissue microarray analysis to evaluate genes uncovered by cDNA microarray screening in renal cell carcinoma, *Am J Pathol* **154** (1999), 981–986.
- [20] T.B. Morrison, J.J. Weis and C.T. Wittwer, Quantification of low-copy transcripts by continuous SYBR Green I monitoring during amplification, *Biotechniques* **24** (1998), 954–958, 960, 962.
- [21] R.G. Panchal, Novel therapeutic strategies to selectively kill cancer cells, *Biochem Pharmacol* **55** (1998), 247–252.
- [22] A. Parle-McDermott, P. McWilliam, O. Tighe, D. Dunican and D.T. Croke, Serial analysis of gene expression identifies putative metastasis-associated transcripts in colon tumour cell lines, *Br J Cancer* **83** (2000), 725–728.
- [23] K. Polyak, Y. Xia, J.L. Zweier, K.W. Kinzler and B. Vogelstein, A model for p53-induced apoptosis, *Nature* **389** (1997), 300–305.
- [24] J. Powell, SAGE. The serial analysis of gene expression, *Methods Mol Biol* **99** (2000), 297–319.
- [25] M.S. Rajeevan, S.D. Vernon, N. Taysavang and E.R. Unger, Validation of array-based gene expression profiles by real-time (kinetic) RT-PCR, *J Mol Diagn* **3** (2001), 26–31.
- [26] G.J. Riggins and R. Strausberg, Genome and genetic resources from the Cancer Genome Anatomy Project, *Hum Molecular Genet* (2001), in press.
- [27] B. Ryu, J. Jones, M.A. Hollingsworth, R.H. Hruban and S.E. Kern, Invasion-specific genes in malignancy: serial analysis of gene expression comparisons of primary and passaged cancers, *Cancer Res* **61** (2001), 1833–1838.
- [28] D. Scheurle, M.P. DeYoung, D.M. Binninger, H. Page, M. Jahanzeb and R. Narayanan, Cancer gene discovery using digital differential display, *Cancer Res* **60** (2000), 4037–4043.
- [29] P. Schraml, J. Kononen, L. Bubendorf et al., Tissue microarrays for gene amplification surveys in many different tumor types, *Clin Cancer Res* **5** (1999), 1966–1975.
- [30] B. St Croix, C. Rago, V. Velculescu et al., Genes expressed in human tumor endothelium, *Science* **289** (2000), 1197–1202.

- [31] R.L. Strausberg, K.H. Buetow, M.R. Emmert-Buck and R.D. Klausner, The cancer genome anatomy project: building an annotated gene index, *Trends Genet* **16** (2000), 103–106.
- [32] R.L. Strausberg, C.A. Dahl and R.D. Klausner, New opportunities for uncovering the molecular basis of cancer, *Nat Genet* **15** (1997), 415–416.
- [33] V.E. Velculescu, S.L. Madden, L. Zhang et al., Analysis of human transcriptomes, *Nat Genet* **23** (1999), 387–388.
- [34] V.E. Velculescu, B. Vogelstein and K.W. Kinzler, Analysing uncharted transcriptomes with SAGE, *Trends Genet* **16** (2000), 423–425.
- [35] V.E. Velculescu, L. Zhang, B. Vogelstein and K.W. Kinzler, Serial analysis of gene expression, *Science* **270** (1995), 484–487.
- [36] C.T. Wittwer, M.G. Herrmann, A.A. Moss and R.P. Rasmussen, Continuous fluorescence monitoring of rapid cycle DNA amplification, *Biotechniques* **22** (1997), 130–131, 134–138.
- [37] C.T. Wittwer, K.M. Ririe, R.V. Andrew, D.A. David, R.A. Gundry and U.J. Balis, The LightCycler: a microvolume multisample fluorimeter with rapid temperature control, *Biotechniques* **22** (1997), 176–181.
- [38] J.T. Wu, Review of circulating tumor markers: from enzyme, carcinoembryonic protein to oncogene and suppressor gene, *Ann Clin Lab Sci* **29** (1999), 106–111.
- [39] J. Yu, L. Zhang, P.M. Hwang, C. Rago, K.W. Kinzler and B. Vogelstein, Identification and classification of p53-regulated genes, *Proc Natl Acad Sci USA* **96** (1999), 14517–14522.
- [40] L. Zhang, W. Zhou, V.E. Velculescu et al., Gene expression profiles in normal and cancer cells, *Science* **276** (1997), 1268–1272.
- [41] W. Zhou, L.J. Sokoll, D.J. Bruzek et al., Identifying markers for pancreatic cancer by gene expression analysis, *Cancer Epidemiol Biomarkers Prev* **7** (1998), 109–112.





# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

