



## REVIEW PAPER

## OPEN ACCESS

## Serial analysis of gene expression (SAGE): its modification and application in human disease

S. Abdolhamid Angaji\*, Ghazal Aminian, Bahram Khosravi, Masoomeh Taran

*Department of Cell and Molecular Biology, Faculty Biological Sciences, Karazmi University, Tehran, Iran*

**Key words:** serial analysis of gene expression(SAGE), microarray, Differential gene expression, human diseases.

<http://dx.doi.org/10.12692/ijb/5.1.49-60>

Article published on July 02, 2014

### Abstract

Among experimental techniques as regards providing gene expression profiles, serial analysis of gene expression (SAGE) yields quantitative, comprehensive data in the cell being studied. Short fragments of genetics information corresponding to those expressed genes of interest form the bedrock of SAGE method, in which these short sequences consisting of 9-10 bp, called SAGE tags, are bound with one another in order for sequencing to be efficient. Unlike the other techniques, requiring no prior knowledge of genes in addition to preparing both the qualitative and quantitative levels of gene expression make SAGE become singled out in fact. Considering different conditions so as to improve the efficiency of this technique, various modifications have been introduced to, including long SAGE, micro SAGE, super SAGE, to name a few. In this review, we have strived to present some new findings as to SAGE technology, along with discussing its practical applications to common diseases, peculiarly genetic disorders.

\* **Corresponding Author:** Abdolhamid Angaji ✉ [Angaji@khu.ac.ir](mailto:Angaji@khu.ac.ir)

### Abbreviation

DGE, Differential gene expression; SAGE, serial analysis of gene expression; TOGA, total gene expression analysis; READS, restriction enzyme analysis of differentially expressed sequences; VDR, vitamin D receptor; CYPA, cyclophilin A.

## Introduction

In recent years, several methods of gene expression have been discovered. Differential gene expression (DGE) technologies have been utilized in great many biological challenges as of their development in the early 1990s. DGE applications have greatly promoted the landscape of basic biological research, together with the discovery and development of pharmaceuticals. DGE is broadly divided into two families of technologies: closed architecture systems and open architecture systems (Geenet *al.*, 2001). Open system methods analyze the expression of genes which are unknown or there is only limited previous knowledge of their genome and therefore these methods are useful for recognizing limited number of the expressed genes. With open system methods discovery of many unknown genes is probable with no need for a priori comprehensive knowledge of the transcriptome sprang into existence, where the field of discovery is 'Open'. GeneCalling, SAGE, TOGA, READS are four examples of open systems in fact (Geen *et al.*, 2001). In contrast to open system, the closed system analyses are quite limited owing to the fact that genes ought to be known since completeness of our knowledge concerning genomes defines their coverage and the more complete knowledge of genome, the greater would be its analysis of the transcriptome. Of closed systems, Oligonucleotide or cDNA array hybridization technologies, and quantitative polymerase chain reaction (qPCR, or TaqMan) are among the most commonly used, also beads array for the detection of gene expression and microarray are examples of closed system methods (Marshall and Hodgson, 1998; Heidet *al.*, 1996). Although closed systems are useful to identify genes in a good way and work well for particular light novel genes, they are limited by the defect of the transcriptome knowledge. For example in the known genome such as human genome, the transcriptome would be more complex because of RNA editing, splice variants and polymorphisms. But if there is complete information, closed systems would provide a desirable knowledge of genome. In open system if a given transcriptome is complex, it won't change the quality of expression analysis and

won't preclude the discovery, but technically, the capacity of open systems is limited. In situations where complete knowledge of a transcriptome exists, they cannot really compete with closed architecture systems in terms of time and labor intensity. Even an open system that theoretically permits analysis of a whole genome is limited by the sensitivity of the method in detecting rare transcripts and splice variants (Lorkowski and Cullen, 2004). Having been touched upon, open systems might not depend upon existing transcriptome information, however exploiting existing expressed genome databases makes considerably these systems effective according to the both facilitating the analysis process as well as improving the efficiency of known and novel gene identification.

In identifying the new genes which require planned works, both closed and open systems are needed to achieve this aim. The unknown genes discover with open systems, then closed systems use the resulted information and can ended up in such discoveries as has been applied to recent genomes.

In this review we discuss the SAGE as an open system which requires no prior knowledge on expressed sequence. It in fact allows the identification of novel genes by using the SAGE tag as a specific polymerase chain reaction primer to amplify the unknown cDNA (van den Berg, 1999).

Indeed, Velculesco, Vogelstein, and Kinzler devised, serial analysis of gene expression (SAGE) method in 1995 at Johns Hopkins University (Velculescu, 1995). Higher organisms such as humans possess a long, massive genome constituting an enormous number of genes whose sequence determination appears to be an attainable objective these days. A highly efficacious technology, SAGE, provides a global expression profile of a specific cell's or tissue's genes. Furthermore, it is useful in discerning particular genes according to cellular conditions through the comparison of gene expression profiles generated for a pair of cells stored in different conditions (Bartlett, 2001; Polyak *et al.*, 1997; Yamamoto *et al.*, 2001).



microarray. These distinguishing attributes enables SAGE to be used as a primary discovery engine that can characterize human diseases at the molecular level while clarifying potential targets and markers for therapeutic and diagnostic purposes, respectively.

Gathering cumulative data, which can be continued at any time, is the other merit in the application of SAGE enabling a more detailed picture of the expression pattern (Hermeking, 2003; Madden *et al.*, 2000).

#### Disadvantages

Requiring a great deal of money and time are two major drawbacks of SAGE since thousands of polymerase chain reactions should be performed, for example virtually 10,000 reactions are needed for a comparison of two different states, in which each reaction is represented by 50,000 SAGE tags (Hermeking, 2003).

Both SAGE and microarray analysis have limits when applied to the identification of tumor markers. One obvious limit is the fact that differences in mRNA levels detected by these methods may not always be reflected at the protein level. Because proteins represent the targets used for either diagnostic or therapeutic approaches in most cases, confirmation of differential expression at the protein level is required for each candidate transcript. Automated proteome analysis using mass spectrometry has significantly progressed in recent years, suggesting that it may soon be used widely to identify proteins that are

specifically expressed in tumors. However, at present, a comprehensive picture of human gene expression can only be achieved with RNA-based techniques (Hermeking, 2003).

To solve such problems, various SAGE modifications have thus far been applied.

#### MiniSAGE

This method was devised by Ye *et al.* in 2000 followed by an additional PCR amplification, requires merely 1 µg of total RNA, which means hundreds of times less than the amount needed for the original SAGE protocol (Lorkowski and Cullen, 2004).

#### MicroSAGE and SAGE-lite

The other modified methods requiring far much less starting input of RNA are MicroSAGE, together with SAGE-lite with 500-5000fold less input for the former (Table 1), and less than 100 ng for the latter in fact (Peters *et al.*, 1999). As regards microSAGE, a 'single tube' procedure is utilized for all the steps through SAGE until the tags are released by digestion with the tagging enzyme. MicroSAGE, furthermore, requires fewer cycles of PCR to generate sufficient ditags, denoting the less input to profile gene expression. Therefore, concerning heterogeneous tissues, microdissection should be applied with microSAGE in order to build region-specific expression profiles. Moreover, profiling in cases of tissue scarcity, biopsy, and tumor metastasis are of the other usages of this method.

**Table 1.** Differences between SAGE procedure and modified procedures.

Differences	SAGE	microSAGE	miniSAGE
Amount of incoming materials	2.5 to 5 µg of RNA	1-5 ng of RNA	1 µg of RNA
Capture of cDNA	streptavidin-coated beads	magnetic streptavidin-coated PCR tube	streptavidin-coated magnetic beads
The number of tube reaction	Successive performed in multiple tubes	reaction Single tube reaction	Single tube reaction
Number of PCR cycle	25-28 of PCR cycle	a limited number of re-PCR cycle followed by 28 cycles	does not require the additional PCR amplifications

#### LongSAGE

LongSAGE makes longer tags (21bp) since it utilizes type II restriction enzyme MmeI. The tags show their

equivalent gene expression; therefore the longer the tags are, the more dependable the statistical analysis of gene expression is (Fig 2). Having said that,

however, there are more single nucleotide polymorphisms (SNPs) present in longer tags, which make a huge decline in the uniformity of tags in SAGE set. In addition, using longer tags costs heavily more, thereby circumscribing its applications (Porter *et al*, 2006).

#### SuperSAGE

The major difference between original SAGE and SuperSAGE is the fact that restriction enzymes which are used are different, BsmF1 and EcoP15

respectively. EcoP15I generates 26bp tags from transcripts (Fig. 2), where these tags in fact make superSAGE become more accurate, conferring the ability to profile plant gene expression, along with identification and quantification of new and known genes (Meisel *et al*, 1992; Matsumura, 2006).

#### Studies made by the use of SAGE

Significant studies performed by SAGE were summarized in table 2.

**Table 2.** Summary of significant studies performed by SAGE.

Cell, tissue	Total tags	Reference
Retina of eye	267,332	Dror Sharon <i>et al</i> , 2001.
VGP and MGP Melanoma Cell Lines	29,150 and 37,710 in two VGP melanoma SAGE libraries (DB1; DB2) respectively. 19,908 and 961 in two MGP melanoma SAGE libraries (DB3; DB4) respectively.	Smith <i>et al</i> , 2004.
Lobular carcinoma	67,834	Dengfeng, <i>et al</i> , 2008.
Endometrial epithelial and stromal cells	19,493 tags in Normal endometrium 80,783 tags in endom1 77,846 tags in endom2 76,100 tags in FbEM-1	Honda <i>et al</i> , 2008.
normal and neoplastic plasma cells	29,918 SAGE tags from normal and 10,340 tags from multiple myeloma tumor	Felix <i>et al</i> , 2009.
Colorectal Tumor Cells	26,060 unique tags	Meeh <i>et al</i> , 2009.
human vein umbilical endothelial cells	31,213 of Normoxia and 30,756 of Hypoxia	Liang <i>et al</i> , 2009.
Esophageal squamous cell carcinoma sample	14,430	Sakamoto <i>et al</i> , 2010.
Human breast tissue	Different SAGE-Seq libraries with different total tags	Wu <i>et al</i> , 2010.
human LNCaP prostate cancer cells	2,931,124	Romanuik <i>et al</i> , 2010.
Striatum	32,216 Uniquetriatal tags	Mazarei <i>et al</i> , 2010.
IPF and normal lung tissue	A total of 954,932 transcript tags were sequenced of which 168,272 were unique	Boon <i>et al</i> , 2011.
Colon Cancer	A total of 100,666 tags were generated including 45,560 unique tags	Fan <i>et al</i> , 2011.
The analysis of normal and adenoma libraries	21 030 and 22 329 unique SAGE tags in normal and adenoma libraries identified respectively	De Lima <i>et al</i> , 2012.
Peripheral Blood Cells	11 million tags	Mastrokolias <i>et al</i> , 2012.
liver fluke-associated Cholangiocarcinoma (CCA) tissue	Three libraries: CCA-K1 (60,319 tags), CCA-K2D (46,853 tags) and normal liver (66,308 tags)	Sawanyawisuth <i>et al</i> , 2012.
Human Neural Tube Fragments	269,043	Krupp <i>et al</i> , 2012.
Breast Carcinoma	66,128 tags in the breast ductal carcinoma library and 50,512 tags in normal epithelium)	Soloviev <i>et al</i> , 2013.

#### SAGE in human studies

SAGE technology has been used to collect many transcripts from a tissue sample, to compare the profile of mRNA expression in different types of cancer cells and to identify genes that are under the new regulation in response to exposure to external stimuli. It also has been widely used in a number of

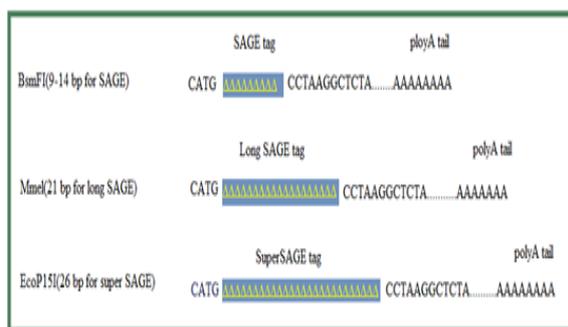
human studies. The following parts are some examples of these applications:

#### Ocular studies

A new study describes SAGE libraries constructed from the human peripheral retina, macula, and retinal pigment epithelium (RPE). They compared the

expression profile of the retinal periphery versus the macula of the same eye and the expression profile of the peripheral retina from two individuals (ages 44 and 88 years) then evaluated the potential value of the library as a resource for identifying candidate genes for hereditary retinal diseases (Sharon *et al*, 2002).

The cone photoreceptor contribution to all transcription in the retina was found to be similar in the macula versus the retinal periphery, whereas the rod contribution was greater in the periphery versus the macula. Genes encoding structural proteins for axons were found to be expressed at higher levels in the macula versus the retinal periphery, probably reflecting the large proportion of ganglion cells in the central retina. In comparison with the younger eye, the peripheral retina of the older eye had a substantially higher proportion of mRNAs from genes encoding proteins involved in iron metabolism or protection against oxidative damage and a substantially lower proportion of mRNAs from genes encoding proteins involved in rod photo transduction. The RPE library had numerous previously unencountered tags, suggesting that this cell type has a large, idiosyncratic repertoire of expressed genes (Sharon *et al*, 2002).



**Fig. 2.** Simplified scheme of SAGE, LongSAGE and SuperSAGE.

Comparison of these libraries with 100 reported nonocular SAGE libraries revealed 89 retina-specific or enriched genes expressed at substantial levels, of which 14 are known to cause a retinal disease and 53 are RPE-specific genes. We expect that these libraries will serve as a resource for understanding the relative expression levels of genes in the retina and the RPE

and for identifying additional disease genes (Sharon *et al*, 2002).

#### *Down syndrome (DS)*

Down syndrome is a congenital abnormality caused by the presence of an additional chromosome 21. The most important of Down syndrome abnormalities are mental retardation, heightened risk of Alzheimer's disease, increased occurrence of leukemia, immunity and heart defects, and muscle hypotonia (Epstein *et al*, 1991). It is assumed that a dosage imbalance of human chromosome 21 (HSA21) bring about the Down syndrome, also, Mao *et al*. in their work illustrated that there was an up regulation of HSA21 genes in DS fetal brain (Mao *et al*, 2003). By the means of SAGE technology, it is possible to create an expression profile of presumed genes involved with the generation and development of DS.

In an interesting study, Sommer and co-workers identified dysregulated genes in lymphocytes from children with Down syndrome (Sommer *et al*, 2008). Two SAGE libraries were constructed using pooled RNA of normal and Down syndrome children. Comparison between DS and normal profiles revealed that most of the transcripts were expressed at similar levels and functional classes of abundant genes were equally represented. Among the 242 significantly differentially expressed SAGE tags, several transcripts down-regulated in DS code for proteins involved in T-cell and B-cell receptor signaling. These results indicate that trisomy 21 induces a modest dysregulation of disomic genes that may be related to the immunological perturbations seen in DS (Malagó *et al*, 2005).

In the other work, gene expression profile of human Down syndrome leukocytes is prepared with SAGE to identify the differences in the gene expression patterns of Down syndrome and normal leukocytes. The generated data can be compared with that of other SAGE libraries constructed from different Down syndrome tissues in order to study the differences specific to temporal and spatial gene expression. The resulted library is currently being

used to analyze the expression profile of Down syndrome leukemic patients in an attempt to clarify the relationship between trisomy 21 and the increased risk of leukemia observed in Down syndrome and is useful in understanding the molecular mechanism of Down syndrome pathology (Malagó *et al*, 2005).

#### *Parkinson disease (PD)*

*Parkinson disease (PD)* is a common neurodegenerative disease that leads to shaking, difficulty with walking and movement and muscle rigidity and weakness. Some mutations like  $\alpha$ -synuclein and GIGYF2 have been elucidated in PD families which resulted that this disease is a polygenic disorder (Polymeropoulos *et al*, 1997; Lautier *et al*, 2008).

To investigation about the polygenic property of PD, by combining genomic convergence with the data generating capacity of SAGE, it demonstrated that a total number of 50 transcripts were found to be directly associated with genomic convergence peaks, thus establishing these genes as high priority candidates for PD risk (Noureddine *et al*, 2005). In recent studies several SAGE transcripts have been discovered which have association with PD and include the ring Wnger protein-11 (RNF11) gene, the glial Wbrillary acidic protein (GFAP) gene and the chemokine (C C motif) ligand- 2 (CCL2) gene .The study by Noureddine *et al*. also found that with the SAGE technique, all UNIGENE sets in linkage regions are detectable and also SAGE was useful in detecting a novel missense polymorphism (A5390G, Ile304Met) in the NADH dehydrogenase subunit 3 (ND3) gene which are seen in the substantianigra of Parkinsonianbrains. This polymorphism make an NlaIII site at the 3' end of the ND3 gene, then it is easy to utilize in SAGE tag isolation(Anderson *et al*,2007; Aponso *et al*,2008; Reale *et al*,2009).

The recent findings that RNF11, GFAP and CCL2 are associated with PD, illustrate the capacity of SAGE to create candidate genes for finding disease-association elements and use in functional surveys (Sommer *et al*, 2008; Malagó *et al*, 2005).

#### *Cancer study*

Cancer is a board range of diseases that has unregulated cell growth. In cancer cells grow with no control and divide for several times and invading to the other parts of body and forming malignant tumors. There are more than 200 different types of cancer which affect humans.

Cholangiocarcinoma (CCA) is the most common liver cancer where it's occurrence rate is so high(Shinet *al*, 2010). The progression of cancer cells toward malignancy and metastasis, including different events such as changes in gene expression can be studied by SAGE. In CCA tumor different alteration of gene expression happened. These include of TMSB10, GAL3, VDR, CYPA and some others. Among these genes, vitamin D receptor (VDR) is important because vitamin D3 [1,25(OH)<sub>2</sub>D<sub>3</sub>], has recently been used for prevention of human tumors and it has antitumor activity (Lamprecht and Lipkin, 2003; Kumagai *et al*, 2003). Sawanyawisuth and his co-workers constructed SAGE libraries of CCA and then compared to the publicly available SAGE library of normal liver tissue by using the SAGE DGED online tool on the CGAP website they identified 509 gene by SAGE that have different expression in CCA, 142 of these were up-regulated and 367 down-regulatedcompare to normal liver tissues (Sawanyawisuth *et al*, 2012). In vitro, vitamin D3 induces cell cycle arrest, differentiation, and apoptosis of cancer cells (Liu *et al*, 1996; Danielsson *et al*, 1999; Diazet *al*, 2000; Pálmer *et al*, 2001). The VDR belongs to the steroid/thyroid hormone nuclear receptor superfamily. VDR up-regulation has been shown in primary tumor tissues of breast, colon and pancreatic cancers (Friedrich *et al*, 1998).Accuracy of the VDR expression pattern corresponds to SAGE analysis. VDR was rarely expressed in normal bile duct epithelia but highly expressed in 74% of CCA tissues. The survival rate of CCA patients with positive VDR expression in tumor tissue was significantly better than that of patients with negative expression of VDR. Indeed, 1,25 (OH)<sub>2</sub>D<sub>3</sub>therapy, in the CCA cell lines with high expression of VDR significantly decrease cell proliferation in a dose-

dependent manner. The effect was not shown in lower VDR expressing CCA cell lines (Seubwai *et al*, 2007). These conclusion are according to those seen in carcinomas of gastric, breast and colon (Albrechtsson *et al*, 2003, Banerjee and Chatterjee, 2003; Pelczynska *et al*, 2005). These findings suggest that supplementation of 1,25 (OH) 2D3 or its analogs may be a potential strategy for long-term control of tumor development and progression in CCA patients. These data encourage further investigation of 1.25 (OH) 2D3 or its analogues as therapeutic agents in the treatment of CCA patients (Seubwai *et al*, 2010).

### References

- Albrechtsson E, Jonsson T, Möller S, Höglund M, Ohlsson B, Axelson J.** 2003. Vitamin D receptor is expressed in pancreatic cancer cells and a vitamin D3 analogue decreases cell number. *Pancreatology* **3**, 41-46.  
<http://dx.doi.org/10.1159/000069149>.
- Anderson LR, Betarbet R, Gearing M, Gulcher J, Hicks AA, Stefánsson K, Lah JJ, Levey AI.** 2007. PARK10 candidate RNF11 is expressed by vulnerable neurons and localizes to Lewy bodies in Parkinson disease brain. *Journal of Neuro pathology & Experimental Neurology* **66**, 955-964.  
<http://dx.doi.org/10.1097/nen.0b013e3181567f17>.
- Aponso PM, Faull RL, Connor B.** 2008. Increased progenitor cell proliferation and astrogenesis in the partial progressive 6-hydroxydopamine model of Parkinson's disease. *Neuroscience* **151**, 1142-1153.  
<http://dx.doi.org/10.1016/j.neuroscience.2007.11.036>.
- Banerjee P, Chatterjee M.** 2003. Antiproliferative role of vitamin D and its analogs--a brief overview. *Molecular and Cellular Biochemistry* **253**, 247-254.  
<http://dx.doi.org/10.1023/A:1026072118217>.
- Bartlett J.** 2001. Technology evaluation: SAGE, Genzyme molecular oncology. *Current Opinion in Molecular Therapeutics* **3**, 85-96.
- Boon K, Bailey NW, Yang J, Steel MP, Groshong S, Kervitsky D, Brown KK, Schwarz MI, Schwartz DA.** 2009. Molecular phenotypes distinguish patients with relatively stable from progressive idiopathic pulmonary fibrosis (IPF). *PLoS One* **4**, Epub 1-14.  
<http://dx.doi.org/10.1371/journal.pone.0005134>.
- Danielsson C, Torma H, Vahlquist A, Carlberg C.** 1999. Positive and negative interaction of 1, 25-dihydroxyvitamin D3 and the retinoid CD437 in the induction of human melanoma cell apoptosis. *International Journal of Cancer* **81**, 467-470.  
[http://dx.doi.org/10.1002/\(SICI\)10970215\(19990505\)81:3<467::AID-IJC22>3.0.CO;2-2](http://dx.doi.org/10.1002/(SICI)10970215(19990505)81:3<467::AID-IJC22>3.0.CO;2-2).
- De Lima DS, Martins CS, Paixao BM, Amaral FC, Colli LM, Saggiaro FP, Neder L, Machado HR, dos Santos AR, Pinheiro DG, Moreira AC, Silva WA Jr, Castro M.** 2012. SAGE analysis highlights the putative role of underexpression of ribosomal proteins in GH-secreting pituitary adenomas. *European Journal of Endocrinology* **167**, 759-768.  
<http://dx.doi.org/10.1530/EJE-12-0760>.
- Dengfeng C, Kornelia P, Marc KH, Hind N, Nina K, Christine I-D, Xinyan W, Saraswati S, Jessica H, Angelo DM, Pedram A.** 2008. Serial analysis of gene expression of lobular carcinoma *in situ* identifies down regulation of claudin 4 and overexpression of matrix metalloproteinase 9. *Breast Cancer Research* **10**, R91.  
<http://dx.doi.org/10.1186/bcr2189>.
- Diaz GD, Paraskeva C, Thomas MG, Binderup L, Hague A.** 2000. Apoptosis is induced by the active metabolite of vitamin D3 and its analogue EB1089 in colorectal adenoma and carcinoma cells: possible implications for prevention and therapy. *Cancer Research* **60**, 2304-2312.
- Dror SH, Seth B, Constance LC, Thaddeus PD.** 2001. Profile of the genes expressed in the human peripheral retina, macula, and retinal pigment

epithelium determined through serial analysis of gene expression (SAGE). Proceedings of the National Academy of Science of the United State of America **99**, 315-320.

<http://dx.doi.org/10.1073/pnas.012582799>.

**Epstein CJ, Korenberg JR, Annerén G, Antonarakis SE, Aymé S, Courchesne E, Epstein LB, Fowler A, Groner Y, Huret JL, Kemper TL, Lott IT, Lubin BH, Magenis E, Opitz JM, Patterson D, Priest JH, Puschel SM, Rapoport S1, Sinet PM, Tanzi RE, de la Cruz F.** 1991. Protocols to establish genotype-phenotype correlations in Down syndrome. The American Journal of Human Genetics **49**, 207-235.

**Fan J, Yan D, Teng M, Tang H, Zhou C, Wang X, Li D, Qiu G, Peng Z.** 2011. Digital transcript profile analysis with aRNA-LongSAGE validates FERMT1 as a potential novel prognostic marker for colon cancer. Clinical Cancer Research **17**, 2908-2918.

<http://dx.doi.org/10.1158/1078-0432.CCR-10-2552>.

**Felix RS, Colleoni GW, Caballero OL, Yamamoto M, Almeida MS, Andrade VC, Chauffaille M, de L, Silva Jr WA, Begnami MD, Soares FA, Simpson AJ, Zago MA, Vettore AL.** 2009. SAGE analysis highlights the importance of p53, ddx5, mapkapk2 and ranbp2 to multiple myeloma tumorigenesis. Cancer Letters **278**, 41-48.

<http://dx.doi.org/10.1016/j.canlet.2008.12.022>.

**Friedrich M, Rafi L, Tilgen W, Schmidt W, Reichrath J.** 1998. Expression of 1,25-dihydroxy vitamin D3 receptor in breast carcinoma. Journal of Histochemistry and Cytochemistry **46**, 1335-1337.

<http://dx.doi.org/10.1177/002215549804601114>.

**Geen CD, Simons JF, Taillon BE, Lewin DA.** 2001. Open systems: panoramic views of gene expression. Journal of Immunological Methods **250**, 67-79.

[http://dx.doi.org/10.1016/S0022-1759\(01\)00306-4](http://dx.doi.org/10.1016/S0022-1759(01)00306-4).

**Heid CA, Stevens J, Livak KJ, Williams PM.** 1996. Real time quantitative PCR. Genome Research **6**, 986-994.

<http://dx.doi.org/10.1101/gr.6.10.986>.

**Hermeking H.** 2003. Serial analysis of gene expression and cancer. Current Opinion in Oncology **15**, 44-49.

**Honda H, Barrieto FF, Gogusev J, Im DD, Morin PJ.** 2008. Serial analysis of gene expression reveals differential expression between endometriosis and normal endometrium. Possible roles for AXL and SHC1 in the pathogenesis of endometriosis. Reproductive Biology and Endocrinology **6**, Epub 1-13.

<http://dx.doi.org/10.1186/1477-7827-6-59>.

**Krupp DR, Xu PT, Thomas S, Dellinger A, Etchevers HC, Vekemans M, Gilbert JR, Speer MC, Ashley-Koch AE, Gregory SG.** 2012. Transcriptome profiling of genes involved in neural tube closure during human embryonic development using long serial analysis of gene expression (LongSAGE). Clinical and Molecular Teratology **94**, 683-692.

<http://dx.doi.org/10.1002/bdra.23040>.

**Kumagai T, O'Kelly J, Said JW, Koeffler HP.** 2003. Vitamin D2 analog 19-nor-1,25-dihydroxyvitamin D2: antitumor activity against leukemia, myeloma, and colon cancer cells. Journal of the National Cancer Institute **95**, 896-905.

<http://dx.doi.org/10.1093/jnci/95.12.896>.

**Lamprecht SA, Lipkin M.** 2003. Chemoprevention of colon cancer by calcium, vitamin D and folate: molecular mechanisms. Nature Reviews Cancer **3**, 601-614.

<http://dx.doi.org/10.1038/nrc1144>.

**Lautier C, Goldwurm S, Dürr A, Giovannone B, Tsiaras WG, Pezzoli G, Brice A, Smith RJ.** 2008. Mutations in the GIGYF2 (TNRC15) gene at the PARK11 locus in familial Parkinson disease. The

American Journal of Human Genetics **82**, 822–833.  
<http://dx.doi.org/10.1016/j.ajhg.2008.01.015>.

**Liang GP, Su YY, Chen J, Yang ZC, Liu YS, Luo XD.** 2009. Analysis of the early adaptive response of endothelial cells to hypoxia via a long serial analysis of gene expression. *Biochemical and Biophysical Research Communications* **384**, 415–419.  
<http://dx.doi.org/10.1016/j.bbrc.2009.04.160>.

**Liu M, Lee MH, Cohen M, Bommakanti M, Freedman LP.** 1996. Transcriptional activation of the Cdk inhibitor p21 by vitamin D3 leads to the induced differentiation of the myelomonocytic cell line U937. *Genes and Development* **10**, 142–153.  
<http://dx.doi.org/10.1101/gad.10.2.142>.

**Lorkowski S, Cullen P.** 2004. High-throughput analysis of mRNA expression: microarrays are not the whole story. *Expert Opinion on Therapeutic Patents* **14**, 377–403.  
<http://dx.doi.org/10.1517/13543776.14.3.377>.

**Madden SL, Wang CL, Landes G.** 2000. Serial analysis of gene expression: from gene discovery to target identification. *Drug Discovery Today* **5**, 415–425.  
[http://dx.doi.org/10.1016/S1359-6446\(00\)01544-0](http://dx.doi.org/10.1016/S1359-6446(00)01544-0).

**Malagó W Jr, Sommer CA, Del Cistia Andrade C, Soares-Costa A, AbraoPossik P, Cassago A, SantejoSilveira HC, Henrique-Silva F.** 2005. Gene expression profile of human Down syndrome leukocytes. *Croatian Medical Journal* **46**, 647–656.

**Mao R, Zielke CL, Zielke HR, Pevsner J.** 2003. Global up-regulation of chromosome 21 gene expression in the developing Down syndrome brain. *Genomics* **81**, 457–467.

**Marshall A, Hodgson J.** 1998. DNA chips: an array of possibilities. *Nature Biotechnology* **16**, 27–31.  
<http://dx.doi.org/10.1038/nbt0198-27>.

**Mastrokolas A, den Dunnen JT, van Ommen GB, 't Hoen PA, van Roon-Mom WM.** 2012. Increased sensitivity of next generation sequencing-based expression profiling after globin reduction in human blood RNA. *BMC Genomics* **13**, Epub 1-9.  
<http://dx.doi.org/10.1186/1471-2164-13-28>.

**Matsumura H.** 2006. SuperSAGE array: the direct use of 26-base pair transcript tags in oligonucleotide arrays. *Nature Methods* **3**, 469–474.  
<http://dx.doi.org/10.1038/nmeth882>.

**Mazarei G, Neal SJ, Becanovic K, Luthi-Carter R, Simpson EM, Leavitt BR.** 2010. Expression analysis of novel striatal-enriched genes in Huntington disease. *Human Molecular Genetics* **19**, 609–622.  
<http://dx.doi.org/10.1093/hmg/ddp527>.

**Meeh PF, Farrell CL, Croshaw R, Crimm H, Miller SK, Oroian D, Kowli S, Zhu J, Carver W, Wu W, Pena E, Buckhaults PJ.** 2009. A gene expression classifier of node-positive colorectal cancer. *Neoplasia* **11**, 1074–1083.  
<http://dx.doi.org/10.1593/neo.09808>.

**Meisel A, Bickle TA, KrügerDH, Schroeder C.** 1992. Type III restriction enzymes need two inversely oriented recognition sites for DNA cleavage. *Nature* **355**, 467–469.  
<http://dx.doi.org/10.1038/355467a>.

**Neilson L, Andalibi A, Kang D, Coutifaris C, Strauss JF 3rd, Stanton JA, Green DP.** 2000. Molecular phenotype of the human oocyte by PCR-SAGE. *Genomics* **63**, 13–24.  
<http://dx.doi.org/10.1006/geno.1999.6059>.

**Noureddine MA, Li YJ, van der Walt JM, Walters R, Jewett RM, Xu H, Wang T, Walter JW, Scott BL, Hulette C, Schmechel D, Stenger JE, Dietrich F, Vance JM, Hauser MA.** 2005. Genomic convergence to identify candidate genes for Parkinson disease: SAGE analysis of the substantia nigra. *Movement Disorders* **20**, 1299–1309.

<http://dx.doi.org/10.1002/mds.20573>.

**Pálmer HG, González-Sancho JM, Espada J, Berciano MT, Puig I, Baulida J, Quintanilla M, Cano A, de Herreros AG, Lafarga M, Muñoz A.** 2001. Vitamin D(3) promotes the differentiation of colon carcinoma cells by the induction of E-cadherin and the inhibition of beta-catenin signaling. *Journal of Cell Biology* **154**, 369-387.

**Patino WD, Mian OY, Hwang PM.** 2002. Serial analysis of gene expression technical considerations and applications to cardiovascular biology. *Circulation Research* **4**, 564-549.

<http://dx.doi.org/10.1161/01.RES.0000036018.76903.18>.

**Pelczynska M, Wietrzyk J, Jaroszewicz I, Nevozhay D, Switalska M, Kutner A, Zabel M, Opolski A.** 2005. Correlation between VDR expression and antiproliferative activity of vitamin D3 compounds in combination with cytostatics. *Anticancer Research* **25**, 2235-2240.

**Peters DG, Kassam AB, Yonas H, O'Hare EH, Ferrell RE, Brufsky AM.** 1999. Comprehensive transcript analysis in small quantities of mRNA by SAGE-Lite. *Nucleic Acids Research* **27**, e39-e44.

<http://dx.doi.org/10.1093/nar/27.24.e39>.

**Polyak K, Xia Y, Zweier JL, Kinzler KW, Vogelstein B.** 1997. A model for p53-induced apoptosis. *Nature* **389**, 300-305.

<http://dx.doi.org/10.1038/38525>.

**Polymeropoulos MH, Lavedan C, Leroy E, Ide SE, Dehejia A, Dutra A, Pike B, Root H, Rubenstein J, Boyer R, Stenroos ES, Chandrasekharappa S, Athanassiadou A, Papapetropoulos T, Johnson WG, Lazzarini AM, Duvoisin RC, Di Iorio G, Golbe LI, Nussbaum RL.** 1997. Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science* **276**, 2045-2047.

<http://dx.doi.org/10.1126/science.276.5321.2045>.

**Porter D, Yao J, Polyak K.** 2006. SAGE and related approaches for cancer target identification. *Drug Discovery Today* **11**, 110-118.

[http://dx.doi.org/10.1016/S1359-6446\(05\)03694-9](http://dx.doi.org/10.1016/S1359-6446(05)03694-9).

**Reale M, Iarlori C, Thomas A, Gambi D, Perfetti B, Di Nicola M, Onofri M.** 2009. Peripheral cytokines profile in Parkinson's disease. *Brain, Behavior, and Immunity* **23**, 55-63.

<http://dx.doi.org/10.1016/j.bbi.2008.07.003>.

**Romanuik TL, Wang G, Morozova O, Delaney A, Marra MA, Sadar MD.** 2010. LNCaP Atlas: Gene expression associated with in vivo progression to castration-recurrent prostate cancer. *BMC Medical Genomics* **3**, 43-62.

<http://dx.doi.org/10.1186/1755-8794-3-43>.

**Sakamoto N, Oue N, Noguchi T, Sentani K, Anami K, Sanada Y, Yoshida K, Yasui W.** 2010. Serial analysis of gene expression of esophageal squamous cell carcinoma: ADAMTS16 is upregulated in esophageal squamous cell carcinoma. *Cancer Science* **4**, 1038-1044.

<http://dx.doi.org/10.1111/j.1349-7006.2009.01477.x>.

**Sawanyawisuth K, Wongkham C, Araki N, Zhao Q, Riggins GJ, Wongkham S.** 2012. Serial analysis of gene expression reveals promising therapeutic targets for liver fluke-associated cholangiocarcinoma. *Asian Pacific Journal of Cancer Prevention* **13**, 89-93.

**Seubwai W, Wongkham C, Puapairoj A, Khuntikeo N, Wongkham S.** 2007. Overexpression of vitamin D receptor indicates a good prognosis for cholangiocarcinoma: implications for therapeutics. *Cancer* **109**, 2497-2505.

<http://dx.doi.org/10.1002/cncr.22716>.

**Seubwai W, Wongkham C, Puapairoj A, Okada S, Wongkham S.** 2010. 22-oxa-1,25 dihydroxyvitamin D3 efficiently inhibits tumor growth in inoculated mice and primary histoculture of cholangiocarcinoma. *Cancer* **116**, 5535-5543.

<http://dx.doi.org/10.1002/cncr.25478>.

**Sharma PC, Matsumura H, Terauchi R.** 2007. Use of serial analysis of gene expression (SAGE) for transcript profiling in plants. *Genomics-Assisted Crop Improvement Genomics Approaches and Platforms*, New York City,US: Springer **1**, 227-244 p.

**Sharon D, Blackshaw S, Cepko CL, Dryja TP.** 2002. Profile of the genes expressed in the human peripheral retina, macula, and retinal pigment epithelium determined through serial analysis of gene expression (SAGE). *Proceedings of the National Academy of Sciences* **99**,315– 320.

<http://dx.doi.org/10.1073/pnas.012582799>.

**Shin HR, Oh JK, Masuyer E, Curado MP, Bouvard V, Fang YY, Wiangnon S, Sripa B, Hong ST.** 2010. Comparison of incidence of intrahepatic and extrahepatic cholangiocarcinoma--focus on East and South-Eastern Asia. *Asian Pacific Journal of Cancer Prevention* **11**, 1159-1166.

**Smith AP, Weeraratna AT, Spears JR, Meltzer PS, Becker D.** 2004. SAGE identification and fluorescence imaging analysis of genes and transcripts in melanomas and precursor lesions. *Cancer Biology & Therapy* **3**, 104-109.

<http://dx.doi.org/10.4161/cbt.3.1.661>.

**Soloviev M, Esteves MP, Amiri F, Crompton MR, Rider CC.** 2013. Elevated transcription of the gene QSOX1 encoding quiescin Q6 sulfhydryl oxidase 1 in breast cancer. *PLoS One* **8**, Epub 1-11.

<http://dx.doi.org/10.1371/journal.pone.0057327>.

**Sommer CA, Pavarino-Bertelli EC, Goloni-Bertollo EM, Henrique-Silva F.** 2008. Identification of dysregulated genes in lymphocytes from children with Down syndrome. *Genome* **51**, 19–29.

<http://dx.doi.org/10.1139/g07-100>.

**Van den Berg A, van der Leij J, Poppema S.** 1999. Serial analysis of gene expression: rapid RT-PCR analysis of unknown SAGE tags. *Nucleic Acids Research* **27**, e17.

<http://dx.doi.org/10.1093/nar/27.17.e17>.

**Velculescu VE, Zhang L, Vogelstein B, Kinzler KW.**1995. Serial analysis of gene expression. *Science* **270**, 484–487.

<http://dx.doi.org/10.1126/science.270.5235.484>.

**Yamamoto M, Wakatsuki T, Hada A, Ryo A.** 2001. Use of serial analysis of gene expression (SAGE) technology. *Journal of Immunological Methods* **250**, 45–66.

**Wu ZJ, Meyer CA, Choudhury S, Shipitsin M, Maruyama R, Bessarabova M, Nikolskaya T, Sukumar S, Schwartzman A, Liu JS, Polyak K, Liu XS.** 2010. Gene expression profiling of human breast tissue samples using SAGE-Seq. *Genome research* **20**, 1730–1739.

<http://dx.doi.org/10.1101/gr.108217.110>.