

Combined laser capture microdissection and serial analysis of gene expression from human tissue samples

Jeong Hee Cho-Vega¹, Patricia Troncoso², Kim-Anh Do³, Carlo Rago⁴, Xuemei Wang³, Spiridon Tsavachidis³, L Jeffrey Medeiros⁵, Kevin Spurgers¹, Christopher Logothetis⁶ and Timothy J McDonnell¹

¹Department of Molecular Pathology, The University of Texas, MD Anderson Cancer Center, Houston, TX, USA; ²Department of Pathology, The University of Texas, MD Anderson Cancer Center, Houston, TX, USA; ³Department of Biostatistics, The University of Texas, MD Anderson Cancer Center, Houston, TX, USA; ⁴Department of Cellular and Molecular Medicine Program, Johns Hopkins School of Medicine, Baltimore, MD, USA; ⁵Department of Hematopathology, The University of Texas, MD Anderson Cancer Center, Houston, TX, USA and ⁶Department of GU Medical Oncology, The University of Texas, MD Anderson Cancer Center, Houston, TX, USA

Cell-specific gene expression profiling from heterogeneous human tissues is confounded by cell purification limitations. Here, we describe a technique to generate gene expression profiles of pure populations of prostate cancer cells obtained from fresh-frozen prostatectomy specimens and small initial quantities of RNA by combining laser capture microdissection and microserial analysis of gene expression (LCM–microSAGE). Two microSAGE libraries were obtained from approximately 100 000 laser pulses, estimated to contain fewer than 3×10^5 cells and 20–30 ng mRNA. Two libraries were sequenced to a depth of 10 111 and 10 463 unique tags from normal and cancer cells, representing 6453 and 6923 genes, respectively. Most transcripts were expressed at similar levels, but cancer cells compared with normal cells had increased expression of 385 tags and decreased expression of 389 tags. A total of 20 genes were differentially expressed ($P < 0.05$); five of these genes were upregulated and 15 were downregulated in cancer cells. Quantitative reverse transcriptase-polymerase chain reaction results from three selected genes corroborated the existence of cell-specific gene expression in LCM–microSAGE-derived libraries. In conclusion, the LCM–microSAGE approach demonstrates that large-scale expression profiles of known and unknown transcripts can be generated from pure populations of target cells obtained from human tissue samples comprised of heterogeneous mixtures of cell types.

Modern Pathology (2005) 18, 577–584, advance online publication, 29 October 2004; doi:10.1038/modpathol.3800327

Keywords: laser capture microdissection; serial analysis of gene expression; prostate cancer

Gene expression profiling has the potential to advance our understanding and treatment of cancer and other diseases. Using this approach, expression profiles of normal and cancer cell populations can be compared, allowing the identification of specific genes or groups of genes that are dysregulated.¹ Expression profiling studies of human prostate cancer have proven to be valuable and demonstrate

the promise of uncovering important molecular mechanisms of disease.^{2–6} A major challenge that impedes the genetic analysis of whole prostate cancer tissue has been the technical inability to collect pure populations of specific cells directly from complex heterogeneous prostate tissue. Therefore, the ability to assess specifically pure populations of prostate cancer cells isolated from non-neoplastic cell types is a significant technical achievement.

The serial analysis of gene expression (SAGE) method is one of several new technologies capable of assessing global gene expression. One advantage of SAGE is that it allows qualitative and quantitative analysis of thousands of transcripts simultaneously.^{7,8} In this approach, short sequence tags (~10 bp) are isolated from mRNA at a defined

Correspondence: Dr TJ McDonnell, MD, PhD, Department of Molecular Pathology, Box 89, The University of Texas, MD Anderson Cancer Center, 1515 Holcombe Blvd., Houston, TX 77030, USA.

E-mail: tmcdonne@mdanderson.org

Received 12 August 2004; revised and accepted 24 September 2004; published online 29 October 2004

position, ligated to form long concatemers, cloned and sequenced. The frequency of each tag in the cloned concatemer directly reflects transcript abundance within the messenger RNA (mRNA) population studied. Depending on the number of tags sequenced, changes in expression levels of rare transcripts can be detected. SAGE is a particularly advantageous technique for analyzing a cell population that is only a small subfraction of a particular tissue because it allows the unbiased detection of rare transcripts likely to be under-represented in pre-existing databases.^{1,9}

There are significant technical challenges associated with using SAGE to profile specific cell populations within heterogeneous human tissue.^{10–12} The limitation of relatively large amounts of input mRNA needed for typical SAGE applications must be overcome if the technique is to be adapted to routinely acquired human tissue samples. Laser capture microdissection (LCM) is a technique that allows the isolation of individual cell populations for genetic analysis, thereby circumventing the impact of tissue heterogeneity.^{13–15} We anticipated that combining an LCM technique with microSAGE would abrogate the limitations associated with a traditional whole tissue approach and produce cell-specific expression profiles from limited human tissue samples.

This study is the first to report the results of combining LCM and microSAGE technologies (LCM-microSAGE) to analyze global gene expression in routinely acquired prostate cancer tissue. Using this combined LCM-microSAGE technique enabled the generation of large-scale gene expression libraries of highly enriched cancer cells and normal cells from tissue sample.

Materials and methods

LCM of Human Prostate Tissue

Tissue samples from a single radical prostatectomy specimen (Gleason score 7(3+4), T₂N₀M₀) were used for this study. Samples of normal and cancer were procured and immediately snap-frozen, and stored at -70°C. Hematoxylin-eosin-stained frozen sections of these specimens were examined by two pathologists (JHC-V, PT).

Using a PixCell LCM with an infrared diode laser (Arcturus Engineering, Santa Clara, CA, USA), 5 µm thick serial sections of retrieved fresh frozen tissues were cut and immediately stained using a HistoGene™ LCM Frozen Section Staining Kit (Arcturus, Mountain View, CA, USA) and LCM was performed immediately thereafter. Cancer and normal epithelial cells were microdissected from different fresh frozen tissue blocks. The mRNAs of cancer cells from 12 serial frozen sections and normal cells from 10 serial frozen sections were isolated using 97 068 and 104 300 laser pulses, respectively, from a laser beam that was 30 µm in diameter.

MicroSAGE Analysis

Complete extraction and purification of total RNA from captured cells was performed by following the protocol of the PicoPure RNA Isolation Kit (Arcturus, Catalogue No. #KIT0202) and checked for RNA integrity and amount by spectrophotometry (DU^R7400 Spectrophotometer, Beckman). Cells were captured by approximately 10 000 pulses to produce total cellular RNA in a 10 µl volume. Then, 10 samples were collected for a total of 100 U (20–30 ng mRNA) for microSAGE analysis. Using the I-SAGE kit (Invitrogen, Carlsbad, CA, USA), microSAGE analysis with a modified ditag amplification procedure was performed as described previously.^{10,11}

The first series of ditag PCR amplifications were performed using 5 µl 10 × reaction buffer, 3 µl DMSO, 7.5 U 10 mM dNTPs, 2 µl each of ditag primer 1 and primer 2:

primer 1 5'GGATTTGCTGGTGCAGTACA3'
(175 ng/µl)

primer 2 5'CTGCTCGAATTCAAGCTTCT3'
(175 ng/µl)

29 µl dH₂O, and 0.5 µl *Taq* polymerase. PCR cycling conditions consisted of heating the reaction mixture to 95°C for 2 min followed by 28 cycles of denaturing at 95°C for 30 s, annealing at 55°C for 1 min and synthesis at 70°C for 1 min (5 min for final cycle). A 104 bp PCR product was generated consisting of linkers (44 + 42 bp) + ditag (18 bp). A negative control consisted of identical conditions with ligase omitted from the reaction mix. The products from 50 reactions were pooled. The 104 bp DNA fragments were purified using SDS-PAGE and the optimal dilution for ditags (1:10, 1:20 or 1:50) was determined. A preparative PCR (192 reactions) was then performed as described above, except that the amount of primers was reduced to 87.5 ng and the number of cycles was reduced to 12.

Three fractions of concatemer (0.5–0.8, 0.8–1.0 and 1.0–2.0 kb) were ligated into a pZERO/*Sphi* vector (Zero Background Cloning Kit, Invitrogen, Carlsbad, CA, USA). Negative controls consisted of a reaction without ligase and another without ligase or concatemer. Electroporation was used for bacterial transformation (50 µl TOP 10 One Shot Electrocompetent cells, 2 µl ligated vector). An amount of 20 µl of transformed bacteria was plated on low-salt LB/Zeocin/IPTG plates and clones were screened via PCR using M13 forward and reverse primers. PCR reactions containing concatemers comprised of at least 15 ditags were purified using a PCR purification column (PSI clone PCR 96 Kit, Princeton Separations, Adelphia, NJ, USA). Direct sequencing (ABI 3700, Applied Biosystems, Branchburg, NJ, USA) of PCR products was performed using a BigDye Primer Kit (Applied Biosystems) according to the manufacturer's instructions.

Sequence Analysis and Tag Extraction Strategies

The sequence and occurrence of each of the transcript tags was determined using the SAGE 2000 version 4.1 software (<http://www.sagenet.org>). SAGE 2000 produced a list of tags with their corresponding numerical values, providing a digital representation of global gene expression. The tags were then matched to the SAGE reliable map (<ftp://ftp.ncbi.nih.gov/pub/sage/map/Hs/NlaIII>). Monte Carlo simulations with adjustment for multiple comparisons were used to determine statistical significance and to calibrate the appropriate threshold values.¹⁶ A maximum value of 0.05 was chosen for *P*-chance. This yielded a false-positive rate that was no higher than 0.02 for the least significant *P*-chance value below the cutoff. The validity of this *P*-chance value was confirmed by using a second method that specifically calculates the probability of observing *y* copies of a particular tag in the cancer library given the observation of *x* copies of the same tag in the normal library.¹⁷ The *P*-chance was calculated by summing overall *y* values that were more extreme than the currently observed values.

Validation of LCM-MicroSAGE Data

To reveal the individual variation of gene expression, we assessed cancer and matched normal tissue samples using quantitative RT-PCR. Total RNA was isolated from 5 μ m-thick fresh frozen tissue sections using a PicoPure™ RNA Isolation Kit (Arcturus). The concentration and integrity of RNA was assessed using a DU^R7400 Spectrophotometer (Beckman). For this analysis, five well-characterized genes were selected including ATP synthase H+ transporting mitochondrial F0 complex subunit g (ATP5L),¹⁸ fatty acid synthase (FASN),¹⁹ kallikrein 3 or PSA

(KLK3),²⁰ calpain small subunit 1 (CAPNS1),²¹ prostate acid phosphatase (ACPP).²² Total RNA from nondissected whole tissue sections (300 ng/ μ l) or LCM samples (30 ng/ μ l) was used for the RT-PCR reaction. RT-PCR was performed using the ABI PRISM 7900HT Sequence Detection System using One-Step RT-PCR Master Mix Reagents (Applied Biosystems, Branchburg, NJ, USA) and Assays-on-Demand Gene Expression probes (Applied Biosystems). Serial dilutions of template isolated from the LNCaP prostate cancer cell line was used to generate standard curves. Statistical software produced the standard curve by measuring the crossing points of each standard and plotting them against the logarithmic value of concentrations. The concentrations of unknown samples were then calculated by setting their crossing points to the standard curve. The gene expression results were normalized to the expression of β -actin. The relative expression of each of the selected genes was calculated by dividing the value of expression for a gene by the value of β -actin.

Results

Generation of MicroSAGE Libraries using LCM

Laser capture microdissection was used to isolate selectively pure populations of prostate cancer cells and non-neoplastic epithelial cells from frozen sections of prostatectomy specimens (Figure 1). LCM-microSAGE libraries of cancer and non-neoplastic cells were generated using approximately 100 000 laser pulses from a PixCell Iie instrument (Arcturus) for each library. This yielded an estimated 3×10^5 cells using a 30 μ m diameter laser beam and approximately 20–30 ng mRNA used to generate each microSAGE library. The mRNA served as starting template for subsequent ditag generation

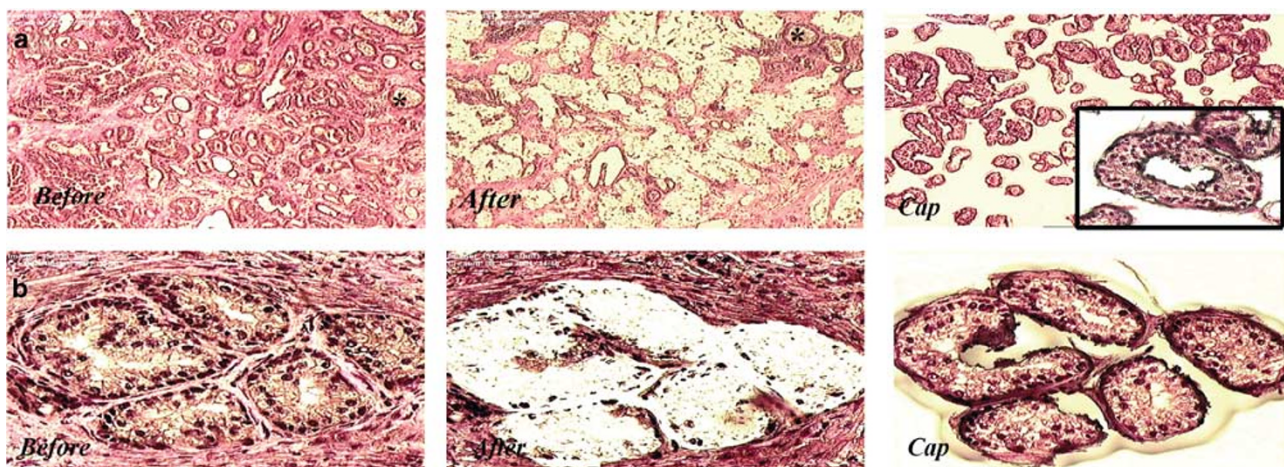


Figure 1 Using a PixCell LCM with an infrared diode, 5 μ m-thick serial sections of retrieved fresh frozen tissues were cut, and stained with a quick-acting hematoxylin–eosin (H&E) stain. Sections were air dried, and LCM was performed immediately afterward. Cancer cells (a, H&E, $\times 40$) and normal cells (b, H&E, $\times 200$) were microdissected from different fresh frozen tissue blocks. Most of cells were specific cancer cells except for a few myoepithelial cells (Cap). Inset: a high power view (H&E, $\times 400$) of a cap showing dissected individual cancer glands on the film. Before, before dissection; after, after dissection; Cap, cells picked up on the film.

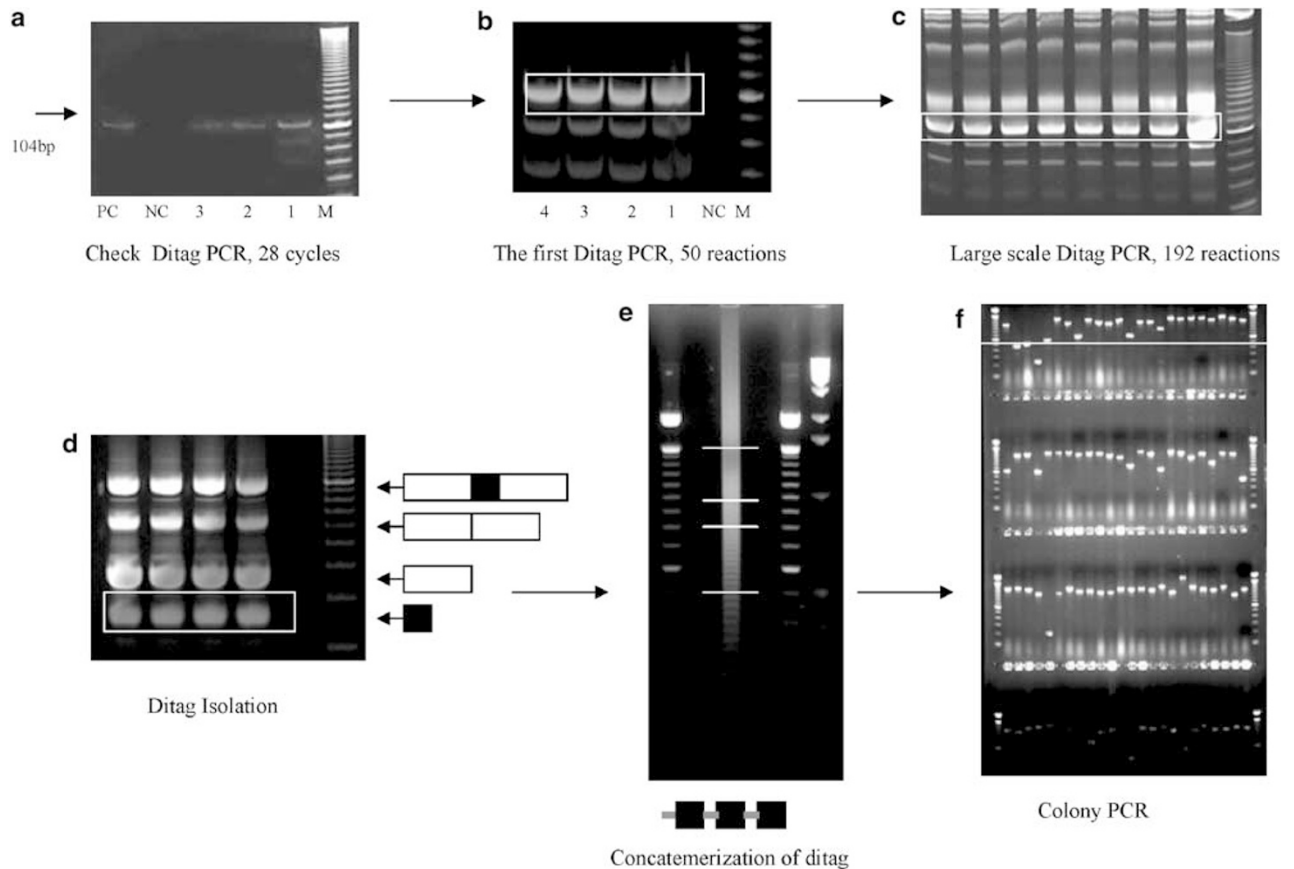


Figure 2 Polyacrylamide gels (8% in **a**, **b**, **c**, 12% in **d**) and agarose gels (1.5% in **e** and **f**) showing the serial steps of LCM-microSAGE in a library from normal cells. **(a)** Shown are 28 cycles of PCR of various dilutions (1, 1:10; 2, 1:20; 3, 1:50) of 1 μ l of the ligated ditag derived from LCM using normal cells to determine the optimal ditag concentration. The 104-bp band corresponding to the amplified ditags was sharply visible compared with the other faint background bands. **(b)** After establishing the optimal dilution, the ditag product was concentrated in 50 reactions. A box is used to delineate the collected ditag bands. The ditag bands were excised and DNA was extracted. **(c)** After large-scale rePCR (192 reactions) with 12 cycles, the PCR products were concentrated and run on a prepared gel from which the 100 bp ditag bands were excised. A box is used to delineate the ditag bands that were collected. **(d)** After digestion with *Nla*III to cleave off the linkers, the small ditag of 22–26 bp (encircled by a black box) was excised and purified. **(e)** The isolated ditags were ligated to concatemers that were separated by size on a 1.5% agarose gel. The regions of the gel containing concatemers ranging from 0.5–0.8 kb (lowest bar), 0.8–1.0 kb and 1.0–2.0 kb (upper most bar) were excised. **(f)** After the purified concatemers were cloned in pZero vector, 20 μ l of transformed bacteria were plated on low-salt LB/Zeocin/ IPTG plates and clones were screened by PCR using M13 forward/reverse primers. PCR reactions that contain concatemers comprised of at least 15 tags (>6161 bp (226 bp vector + 26 bp per ditagX15 ditags)) were purified. A bar indicated 600 bp. NC, negative control performed by H₂O for 35 cycles in **(a)** and **(b)**; PC, positive control RNA in I-SAGE kit; M, 10 bp ladder; ■, ditag; □, linker.

as described in detail in Materials and methods (Figure 2).

Expression Profiles in Normal and Matched Cancer Cell Libraries

Two libraries were generated from normal and matched cancer cells. Concatemerized ditags were sequenced from 923 and 1612 selected colonies from normal and cancer cell libraries, respectively. A total of 10 463 and 10 111 unique tags were identified from each library, respectively (Table 1). Of the tags that matched with known genes in the SAGE map, 5272 matched with a single gene and 3798 matched with more than one gene (Table 1). Genes known to be highly expressed in both cancer and normal

Table 1 Overall summary of two LCM-microSAGE libraries

	Cancer	Normal	Total
Total count of tags	10 463	10 111	20 574
No. of tags	4692	4377	9069
No. of genes	3097	3212	6309
No. of ESTs	847	827	1674
No. of tags matched to single gene	2717	2554	5271
No. of tags matched to multiple genes	1975	1823	3798
No. of tags unmatched to genes	3097	3212	6309
No. of tags matched to ribosomal protein	383	378	761

prostate cells including β -microseminoprotein (MSMB), sperm-associated antigen 7 (SPAG7), translationally controlled tumor protein (TPT1), lamin

Table 2 Genes highly expressed in both cancer and normal libraries

Tag		Cancer	Normal	UniGene	Gene	Description
ATAATTCTTT	*	125	104	539	RPS29	Ribosomal protein S29
ATTTGAGAAG		112	97	404283	RAD23B	RAD23 homolog B (<i>S. cerevisiae</i>)
ACTTTTCAA	*	111	76	133430	—	ESTs
GAGGGAGTTT	*	98	78	76064	RPL27A	Ribosomal protein L27a
CACAAACGGT	*	86	76	337307	—	ESTs, highly similar to hypothetical protein FLJ22662 [<i>Homo sapiens</i>] [<i>H. sapiens</i>]
TTCAATAAAA	*	85	89	2012	TCN1	Transcobalamin I (vitamin B12 binding protein, R binder family)
AAGACAGTGG	*	79	74	3352	HDAC2	Histone deacetylase 2
TTGGTCCTCT	*	73	89	356795	RPL41	Ribosomal protein L41
TAATAAAGGT	*	70	68	399720	RPS8	Ribosomal protein S8
TGGTGTGAG	*	67	63	275865	RPS18	Ribosomal protein S18
TAGGTTGTCT		64	55	401448	TPT1	Tumor protein, translationally controlled 1
GGATTTGGCC		61	65	297753	RPLP2	Ribosomal protein, large P2
AAGGAGATGG	*	59	40	164170	VRP	Vascular Rab-GAP/TBC-containing
GTGAAACCCC	*	57	37	268049	HSPC031	Hypothetical protein HSPC031
GGCAAGAAGA	*	53	41	83321	NMB	Neuromedin B
TGCACGTTTT		52	46	169793	RPL32	Ribosomal protein L32
GCATAAATAG	*	52	56	431927	RPL21	Ribosomal protein L21
CAATAAATGT	*	51	51	337445	RPL37	Ribosomal protein L37
ATCTCCAGT		51	55	234518	RPL23	Ribosomal protein L23
GCCGAGGAAG		48	43	434029	C1S	Complement component 1, s subcomponent
TGTGCTAAAT	*	46	48	250895	RPL34	Ribosomal protein L34
GGGCTGGGGT	*	44	38	90436	SPAG7	Sperm-associated antigen 7
CTGTGTTTAG		44	45	78825	MATR3	Matrin 3
CCTGTAATCC	*	43	38	181874	IFIT4	Interferon-induced protein with tetratricopeptide repeats 4
TCAGATCTTT	*	42	57	85112	IGF1	Insulin-like growth factor 1 (somatomedin C)
TGTGTTGAGA	*	41	38	356428	—	<i>Homo sapiens</i> mRNA expressed only in placental villi, clone SMAP83
GCCTGTATGA		40	36	180450	RPS24	Ribosomal protein S24
CCAGAACAGA		40	39	334807	RPL30	Ribosomal protein L30
CTGGGTTAAT	*	40	41	97457	—	ESTs
AGCTCTCCCT		39	37	82202	RPL17	Ribosomal protein L17
GAAAAATGGT	*	39	43	181357	LAMR1	Laminin receptor 1 (ribosomal protein SA, 67 kDa)
GGGAAGCAGA		38	31			
AATAGTCCA		35	33	409158	RPS25	Ribosomal protein S25
CCTATCAGTA		35	35	433392	MSMB	Microseminoprotein, beta-
TAAGGAGCTG	*	35	34	50651	JAK1	Janus kinase 1 (a protein tyrosine kinase)
CACTACTCAC		35	43			
CCACTGCACT	*	34	29	5338	CA12	Carbonic anhydrase XII
GTAAGTGTAC		34	28			
AGGCTACGGA	*	33	33	389335	RPL13A	Ribosomal protein L13a
CCCGTCCGGA		32	35	431392	RPL13	Ribosomal protein L13
ATGGCTGGTA	*	32	33	356364	—	ESTs, moderately similar to ribosomal protein S2; 40S ribosomal protein S2 [<i>Homo sapiens</i>] [<i>H. sapiens</i>]
AGGAAAGCTG		31	33	433411	RPL36	Ribosomal protein L36
AAGGTGGAGG	*	31	21	76171	CEBPA	CCAAT/enhancer binding protein (C/EBP), alpha
CCTCAGGATA		30	20	170009	TGFA	Transforming growth factor, alpha

The tags matching with more than one gene are marked with an asterisk. ID, identification number.

receptor 1 (LAMR1), and transforming growth factor alpha (TGFA) were expressed at high levels in the LCM-microSAGE libraries (Table 2). In all, 44% of the matching transcripts corresponded with mRNA sequence entries that have been characterized, whereas 46% matched either uncharacterized expressed sequence tags (EST) or complementary DNA (cDNA) entries (Table 1).

Most transcripts were expressed at similar levels in the cancer and non-neoplastic epithelial cell libraries (Figure 3). In cancer cells, 385 tags (4.2% of 9069 analyzed total unique tags) were increased more than fourfold and 389 tags (4.3%) were

decreased more than fourfold. Of these transcripts, 20 were exhibited significantly different ($P < 0.05$) expression between matched normal and cancer cells (Table 3). These differentially expressed transcripts consisted of five that were upregulated in prostate cancer cells and 15 that were downregulated.

Validation of LCM-MicroSAGE

To confirm the LCM-microSAGE results, fresh tissue sections of cancer and normal tissue samples that

were used for LCM-microSAGE analysis were also assessed by QRT-PCR for five selected genes (Figure 4). These genes were selected because they have been well studied in prostate cancer. The expression of ATP5L (tag ratio of normal vs cancer = 0.4),¹⁸ FASN (0.33)¹⁹ and CAPNS1 (0.75)²¹ were demonstrated to be expressed at higher levels in the cancer cell LCM-microSAGE libraries. These genes were

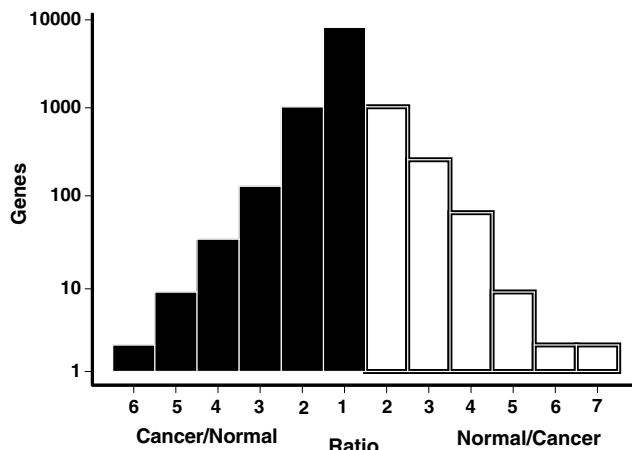


Figure 3 Semilogarithmic plot distribution of tags in prostate cancer and normal epithelial cell LCM-microSAGE libraries.

confirmed to be expressed at 2–3-fold higher levels in the cancer relative to normal cells with QRT-PCR. KLK3 (PSA) (1.0)²⁰ and ACPP (1.0)²² were present at equivalent levels both in the normal and cancer cell LCM-microSAGE libraries. However, these genes were shown to be expressed at approximately 10-fold higher levels in cancer tissue with QRT-PCR. ATP5L and FASN genes were selected and their expression validated against LCM paired normal, cancer, and cancer stromal cells (Figure 4b). FASN was expressed to a greater degree in cancer cells than in normal and cancer stromal cells.

Discussion

Several methodologies have recently been developed to generate genomewide expression profiles that are characteristic of specific cell types or states of differentiation. In general these techniques involve a hybridization step, usually either cDNAs or oligonucleotides fixed to a solid matrix,^{23–25} or the SAGE.^{1,7,8,12} As with any methodology, each of these techniques possesses advantages and corresponding disadvantages. The advantages of SAGE include the potential to quantify a large number of transcripts simultaneously in the absence of any *a priori* sequence information.^{1,7} Additionally,

Table 3 Genes showing differential expression in prostate cancer vs normal cells

Tag		Cancer	Normal	Unigene	GENE	Description	P-value
GTGAAACCCC	*	57	37	268049	HSPC031(CSF2RA)	Hypothetical protein HSPC031	0.026556
CCCATCATCC		5	0	306122	GPSN2	Glycoprotein, synaptic 2	0.031748
GGGTTTTTAT		5	0	74497	NSEP1	Nuclease sensitive element binding protein 1	0.031748
AAGGAGATGG	*	59	40	164170	VRP(RPL31)	Vascular Rab-GAP/TBC-containing(ribosomal protein L31)	0.037452
TGTGATCAGA		15	6	107476	ATP5L	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit g	0.040422
ACTTTTTTCAG		0	6	171495	RARB	Retinoic acid receptor, beta	0.007641
TCCATACACC		0	5	89497	LMNB1	Lamin B1	0.01533
TAAAAAAAAA	*	2	9	379466	UBE2A	Ubiquitin-conjugating enzyme E2A (RAD6 homolog)	0.018782
TTTAGGTAAA		1	7	76698	SERP1	Stress-associated endoplasmic reticulum protein 1	0.01909
GTGTATATTG		3	10	42251	MGC10067	Hypothetical protein MGC10067	0.027921
AAGGCAATTT	*	0	4	278479	TSPYL	TSPY-like	0.030758
AGCCCTACAG		0	4				0.030758
AGTTAAATTA		0	4				0.030758
GACACAGGCA	*	0	4	143641	LRP3	Low-density lipoprotein receptor-related protein 3	0.030758
GAGAAATGGT		0	4				0.030758
GATTCGTAAG		0	4				0.030758
GTGGCTGGTA		0	4				0.030758
TATTATCCA		0	4				0.030758
TGTATATCGT		0	4	79123	KIAA0084	KIAA0084 protein	0.030758
GGCCCTACAA		1	6	79058	SUPT4H1	Suppressor of Ty 4 homolog 1 (<i>S. cerevisiae</i>)	0.034469

According to Monte Carlo simulation, 20 transcripts revealed differential expression at significant different levels ($P < 0.05$); 15 of these were downregulated in prostate cancer cells, whereas five genes were upregulated in cancer cells. Six tags had no matched gene entries. The tags matching with more than one gene are marked with an asterisk. UID, Unigene identification number.

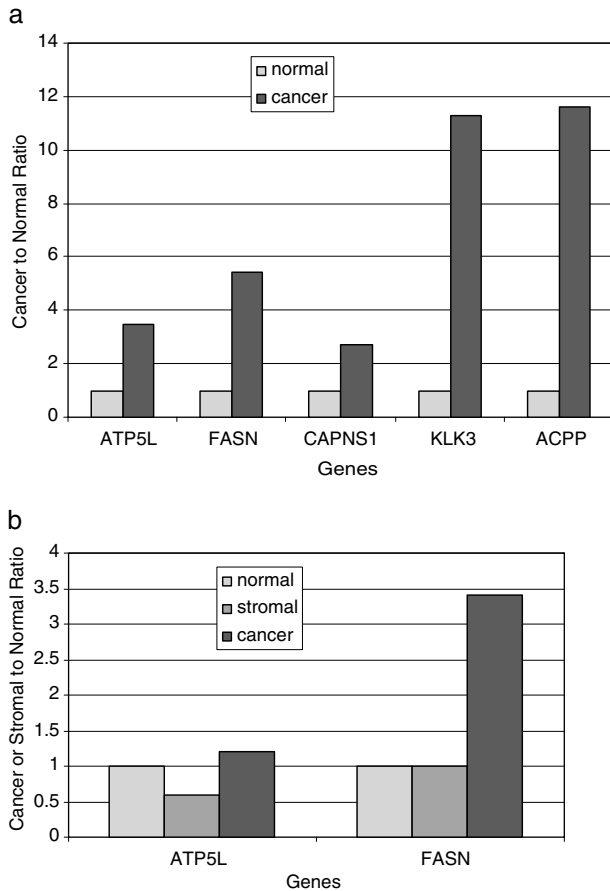


Figure 4 Results of quantitative RT-PCR of five selected human genes. (a) ATP5L, FASN, CAPSE1, KLK3, ACPP gene expression in non-dissected whole normal and paired cancer tissues from the same patient. (b) ATP5L and FASN gene expressions were analyzed with total RNAs derived from LCM paired normal, cancer, and cancer-stromal cells from the same patient. Values are normalized against β -actin.

previous studies have shown that there is not complete overlap in transcripts identified as differentially expressed using SAGE and array-based technologies.

However, the application of SAGE is technically more demanding and limited by the requirement for a relatively large amount of input mRNA.^{7,26} These disadvantages have necessarily limited the application of SAGE expression profiling to routinely acquired human tissue samples. Further, these tissue samples are comprised of heterogeneous cell populations, including the population of interest, thereby conveying the potential to complicate the interpretation of the expression profiling data.

Therefore, it was of interest to address these challenges in an attempt to adapt the SAGE technique to small human tissue samples that are obtained for diagnostic evaluation. LCM is now a standard technique to isolate cells of interest rapidly from tissue sections with estimates of purity of >95%.¹³ Recently, LCM performed on routinely stained frozen tissue sections have been used for

mRNA isolation amenable to RT-PCR and generation of cDNA array expression libraries.^{3,27-29} The use of LCM-derived mRNA for SAGE profiling has not previously been described. Owing to the limited quantity of template obtained from LCM isolated cell populations, an amplification procedure was used, similar to that of others,^{10,11} in order to generate sufficient amounts for the generation of SAGE libraries. This PCR step should be relatively free of bias because all ditags are of approximately equal length. However, the possibility that some ditag species are still preferentially amplified cannot be excluded. The advantage of using two rounds of PCR is that it enabled SAGE analysis to be performed from the limited amount template available from LCM-isolated cell populations. Furthermore, even though preferential amplification may occur, the SAGE software discards duplicate ditags generated by preferential amplification. The software, therefore, corrects this bias and provides more robust results. The high percentage of duplicate ditags excluded from subsequent analysis reduces the average number of analyzed tags obtained per clone compared to standard SAGE analyses.^{1,16}

The frequency of these potential artifacts can increase exponentially with the number of PCR cycles performed. In order to minimize this effect we performed multiple parallel PCR reactions of fewer cycles, rather than increase the number of PCR cycles using fewer replicates.¹¹ Using this procedure, we could obtain an expression profile from <30 ng of LCM-derived mRNA. A high percentage (>20%) of duplicate ditags were obtained using the LCM-microSAGE method. This may be a consequence of the relatively low complexity within the mRNA population, due to either the LCM protocol (a low amount and quality of starting material). Alternatively, it might be anticipated that variable preservation of mRNA in banked tissue samples would reduce complexity and result in an increase in the number of duplicate ditags.

Since a SAGE library generates candidate genes, independent methods, such as RT-PCR, Northern blot analysis^{10,11} *in situ* hybridization^{12,30} or immunohistochemistry⁵ are needed for validation. However, validation was considered a lower priority goal for this study, since it was thought that relatively few statistically significant differences in gene expression would be determined based on the amount of total tags sequenced for the study. Even with total 20 574 tags sequenced, the majority of epithelial and tumor-associated genes may not be expressed at sufficient levels to permit a stringent statistical assessment of differential expression.

QRT-PCR was used to confirm the expression of several well-characterized genes of matched cancer and normal prostate tissue. ATP5L, FASN, CAPNS1 revealed a good correlation with transcripts (tags) in both LCM-microSAGE libraries. In contrast, KLK3 and ACPP exhibited lower transcript levels in both the cancer and normal cell libraries, but higher

mRNA expression in whole cancer tissue. It is noteworthy that both the KLK3 (1464 bp) and ACPP (2239 bp) transcripts are relatively large. These transcripts may be preferentially degraded during the LCM and/or microSAGE procedures. In this regard, the β -microseminoprotein transcript (571 bp) was abundantly expressed in both LCM-microSAGE libraries. Ongoing studies will provide more detailed information regarding the variables that impact the sensitivity and reproducibility of the LCM-microSAGE technique.

In conclusion, we demonstrate the feasibility of combining LCM with microSAGE. This approach allows generation of cell-specific expression profiles and the assessment of the expression levels of known and unknown transcripts in specific cell types obtained from human tissue samples.

Acknowledgements

This work was supported by Grants NIH P50 CA40270 and EDRN-CA99-007.

References

- Zhang L, Zhou W, Velculescu VE, *et al*. Gene expression profiles in normal and cancer cells. *Science* 1997;276:1268–1272.
- Velculescu VE, Madden SL, Zhang L, *et al*. Analysis of human transcriptomes. *Nat Genet* 1999;23:387–388.
- Emmert-Buck MR, Strausberg RL, Krizman DB, *et al*. Molecular profiling of clinical tissues specimens: feasibility and applications. *J Mol Diagn* 2000;2:60–66.
- Krizman DB, Chuaqui RF, Meltzer PS, *et al*. Construction of a representative cDNA library from prostatic intraepithelial neoplasia. *Cancer Res* 1996;56:5380–5383.
- Waghray A, Schober M, Feroze F, *et al*. Identification of differentially expressed genes by serial analysis of gene expression in human prostate cancer. *Cancer Res* 2001;61:4283–4286.
- Xu LL, Su YP, Labiche R, *et al*. Quantitative expression profile of androgen-regulated genes in prostate cancer cells and identification of prostate-specific genes. *Int J Cancer* 2001;92:322–328.
- Velculescu VE, Zhang L, Vogelstein B, *et al*. Serial analysis of gene expression. *Science* 1995;270:484–487.
- Velculescu VE, Zhang L, Zhou W, *et al*. Characterization of the yeast transcriptome. *Cell* 1997;88:243–251.
- Chen J, Sun M, Lee S, *et al*. Identifying novel transcripts and novel genes in the human genome by using novel SAGE tags. *Proc Natl Acad Sci USA* 2002;99:12257–12262.
- Datson NA, van der Perk-de Jong J, van den Berg MP, *et al*. MicroSAGE: a modified procedure for serial analysis of gene expression in limited amounts of tissue. *Nucleic Acids Res* 1999;27:1300–1307.
- Virlon B, Cheval L, Buhler JM, *et al*. Serial microanalysis of renal transcriptomes. *Proc Natl Acad Sci USA* 1999;96:15286–15291.
- St Croix B, Rago C, Velculescu V, *et al*. Genes expressed in human tumor endothelium. *Science* 2000;289:1197–1202.
- Emmert-Buck MR, Bonner RF, Smith PD, *et al*. Laser capture microdissection. *Science* 1996;274:998–1001.
- Schutze K, Posl H, Lahr G. Laser micromanipulation systems as universal tools in cellular and molecular biology and in medicine. *Cell Mol Biol (Noisy-le-grand)* 1998;44:735–746.
- Webb T. Laser capture microdissection comes into mainstream use. *J Natl Cancer Inst* 2000;92:1710–1711.
- Lal A, Lash AE, Altschul SF, *et al*. A public database for gene expression in human cancers. *Cancer Res* 1999;59:5403–5407.
- Audic S, Claverie JM. The significance of digital gene expression profiles. *Genome Res* 1997;7:986–995.
- Itoh H, Takahashi A, Adachi K, *et al*. Mechanically driven ATP synthesis by F1-ATPase. *Nature* 2004;427:465–468.
- Baron A, Migita T, Tang D, *et al*. Fatty acid synthase: a metabolic oncogene in prostate cancer? *J Cell Biochem* 2004;91:47–53.
- Lilja H. Biology of prostate-specific antigen. *Urology* 2003;62:27–33.
- Carafoli E, Molinari M. Calpain: a protease in search of a function? *Biochem Biophys Res Commun* 1998;247:193–203.
- Zelivianski S, Glowacki R, Lin MF. Transcriptional activation of the human prostatic acid phosphatase gene by NF-kappaB via a novel hexanucleotide-binding site. *Nucleic Acids Res* 2004;32:3566–3580.
- Prashar Y, Weissman SM. READS: a method for display of 3'-end fragments of restriction enzyme-digested cDNAs for analysis of differential gene expression. *Methods Enzymol* 1999;303:258–272.
- Drmanac S, Stavropoulos NA, Labat I, *et al*. Gene-representing cDNA clusters defined by hybridization of 57 419 clones from infant brain libraries with short oligonucleotide probes. *Genomics* 1996;37:29–40.
- Lockhart DJ, Dong H, Byrne MC, *et al*. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotechnol* 1996;14:1675–1680.
- Polyak K, Xia Y, Zweier JL, *et al*. A model for p53-induced apoptosis. *Nature* 1997;389:300–305.
- Leethanakul C, Patel V, Gillespie J, *et al*. Distinct pattern of expression of differentiation and growth-related genes in squamous cell carcinomas of the head and neck revealed by the use of laser capture microdissection and cDNA arrays. *Oncogene* 2000;19:3220–3224.
- Ohyama H, Zhang X, Kohno Y, *et al*. Laser capture microdissection-generated target sample for high-density oligonucleotide array hybridization. *Biotechniques* 2000;29:530–536.
- Betsuyaku T, Griffin GL, Watson MA, *et al*. Laser capture microdissection and real-time reverse transcriptase/polymerase chain reaction of bronchiolar epithelium after bleomycin. *Am J Respir Cell Mol Biol* 2001;25:278–284.
- Hibi K, Liu Q, Beaudry GA, *et al*. Serial analysis of gene expression in non-small cell lung cancer. *Cancer Res* 1998;58:5690–5694.