

# Shotgun proteome expression profile of microdissected superficial transitional cell carcinoma and biomarker discovery from urine

HAI-TAO NIU<sup>1</sup>, SHU-JUAN LI<sup>2</sup>, TING XU<sup>3</sup>, QIN-CHAO YU<sup>1</sup>, JI-WU CHANG<sup>4</sup> and GUANG SUN<sup>4</sup>

<sup>1</sup>Department of Urology, Affiliated Hospital of the Medical College of Qingdao University, Qingdao 266003;

<sup>2</sup>Department of Pharmacology, Hebei North University, Zhangjiakou 075000; <sup>3</sup>Department of Gerontology, 401st Hospital of PLA, Qingdao 266000; <sup>4</sup>Department of Urology, Tianjin Institute of Urology, Second Hospital of Tianjin Medical University, Tianjin 300211, P.R. China

Received March 27, 2009; Accepted June 17, 2009

DOI: 10.3892/mmr\_00000162

**Abstract.** The aim of this study was to globally characterize the expression profile of superficial transitional cell carcinoma using shotgun proteome strategy and to discuss the biomarker panel identified from urine. We identified 440 commonly expressed proteins from four samples of superficial transitional cell carcinoma using laser capture microdissection coupled with two-dimensional liquid chromatography tandem mass spectrometry. The identified proteins were further analyzed using bioinformatics tools and compared with the published literature. Compared with the entire list of the International Protein Index, there were 41/22 Gene ontology (GO) terms found to be enriched/depleted within the biological process annotation. GO biological process enrichment/depletion analysis was consistent with the results of urine proteome analysis. Proteins classified under the terms cell adhesion, cell proliferation and cell differentiation are good candidate biomarkers for cancer detection from urine. In conclusion, the present study identified an extensive expression profile in superficial transitional cell carcinoma of the bladder, providing information for the understanding of cancer cell biology and the discovery of a biomarker panel from urine.

## Introduction

Transitional cell carcinoma of the bladder (BTCC) is the most common urologic tumor in China and the sixth most common malignancy in developed countries (1). The majority of bladder

tumors (70%) are diagnosed as superficial and are easily treated by transurethral dissection and chemotherapy (2). However, these tumors exhibit a high frequency of recurrence and can progress to muscle-invasive and metastatic tumors, posing a serious threat to patient survival (3). Currently, the molecular basis of these characteristics of superficial BTCC is unknown. Thus, there is a clear need to understand cancer cell biology and to identify effective biomarkers. Several studies involving two-dimensional electrophoresis (2-DE)-based proteomics have been performed, yet few differentially expressed proteins have been identified (4,5). Moreover, several classes of proteins, including poorly soluble proteins, proteins with extreme isoelectric points (pI) and molecular weights (MW), low abundance proteins and integral membrane proteins, tend to be poorly represented in 2-DE analyses (6). One promising alternative method, shotgun proteomic strategy, involves the use of multidimensional liquid chromatography (LC) in conjunction with tandem mass spectrometry (MS/MS) for the direct analysis of complex mixtures of peptides derived from an organism of interest (7).

In this study, we performed shotgun strategy as the initial step in our proteomic approach to understanding the biology of superficial BTCC and to the discovery of a biomarker panel. Gene ontology (GO) analysis, which has the advantage of wide proteome coverage, was possible. After analysis of the whole expression profile, we addressed the expression characteristics of superficial BTCC after laser capture microdissection (LCM). Following GO analysis, the protein categories that tended to be detected in urine were selected. The techniques employed in the present study, while effective, were associated with technical limitations, discussed below.

## Materials and methods

**Patients and cancerous tissue samples.** Four superficial BTCC specimens (confirmed by two individual pathological diagnoses) were obtained from patients treated at the Second Hospital of Tianjin Medical University immediately after radical cystectomy due to multiple superficial BTCC. None of the patients were previously treated with intravesical chemo/

---

*Correspondence to:* Dr Hai-Tao Niu, Department of Urology, Affiliated Hospital of the Medical College of Qingdao University, 16 Jiangsu Road, Qingdao 266003, P.R. China  
E-mail: niuhaitao1976@yahoo.com.cn

**Key words:** superficial bladder transitional cell carcinoma, proteome, gene ontology, biomarker panel

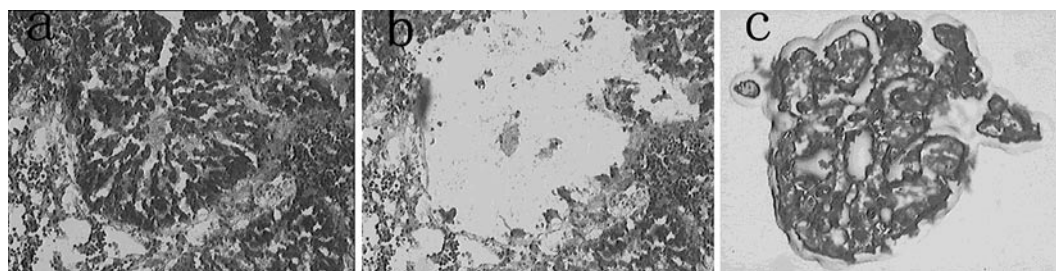


Figure 1. Cancer cells harvested by LCM. (a) Before LCM; (b) after LCM. (c) Microdissected cancer cells on the cap.

immunotherapy. The research protocol was approved by the Institutional Review Board, and informed consent was obtained from the patients. Tumor samples were excised and placed in liquid nitrogen until use.

**Laser capture microdissection.** Sections (8  $\mu$ m) of freshly prepared BTCC tissues were stained with toluidine blue according to the manufacturer's standard protocol with minor modifications. The sections were air-dried and microdissected with a Leica AS LMD Laser Capture Microdissection System. The conditions of LCM consisted of a laser beam diameter, 15  $\mu$ m; duration time, 2.0 msec; and energy, 80 mW. Approximately 250,000 cancer cells from each specimen were microdissected and stored on microdissection caps at  $-80^{\circ}\text{C}$  until lysed. Fig. 1 shows a representative LCM of a specimen.

**Digestion of protein mixture and 2D-LC-MS/MS.** Samples prepared by LCM technology were deposited in precipitation solution for at least 12 h at  $-20^{\circ}\text{C}$ . The pellets were washed with 100% acetone and 70% ethanol, then redissolved in 6 mM guanidine-HCl/100 mM Tris (pH 8.3), and the concentrations were measured using the Bio-Rad Protein Assay Kit. Next, 150  $\mu$ g of soluble proteins were reduced with dithiothreitol (DTT) and subsequently alkylated with iodoacetamide (IAA). After desalting and the removal of toluidine blue by ultrafiltration with Microcon-10, the protein mixture was incubated with trypsin at  $37^{\circ}\text{C}$  for 20 h.

Two-dimensional high-performance LC separations were performed on a Proteome X Workstation (Thermo Finnigan) equipped with two capillary LC pumps. A linear ion trap-Fourier transform (LTQ-FT) and linear ion trap-orbitrap (LTQ-Orbitrap) mass spectrometer (Thermo Finnigan) were used for peptide detection. Peptides and proteins were identified using TurboSequest software (Thermo Finnigan), which uses the MS and MS/MS spectra of peptide ions to search against the publicly available International Protein Index (IPI) database.

**Bioinformatics analysis.** The pI and MW of the proteins were analyzed using ExPASy proteomic tools accessed from [cn.expasy.org/tools/#proteome](http://cn.expasy.org/tools/#proteome). The grand average of hydropathicity (GRAVY) score was calculated as the arithmetic mean of the sum of the hydropathic indices of each amino acid. Transmembrane prediction was performed by the program TMHMM 2.0. To obtain an overview of the expression profile,

all identified proteins were categorized according to GO assignments ([www.geneontology.org](http://www.geneontology.org)), and GOfact software was used to identify statistically over- or under-represented GO categories in the biological data as the tool for enrichment analysis of our proteome dataset (8). The analysis was conducted using a hypergeometric test, and GO terms with a P-value  $<0.05$  or  $<0.01$  were categorized as enriched/depleted or significantly enriched/depleted, respectively. In order to facilitate the discovery of biomarkers from urine, the GO enrichment analysis was compared with the results of proteosome urine analysis.

## Results

**Identification of superficial BTCC proteins.** Using 2D-LC-MS/MS, we identified 489, 512, 493 and 479 proteins from the four cancerous tissue specimens, respectively. In order to compensate for the heterogeneous nature of humans, only the 440 proteins that commonly appeared in the four cancerous tissue specimens were analyzed. The smallest and largest MW values observed were 7.86 and 1005.20 kDa. Regarding pI distribution, the 440 proteins were distributed across a wide pI range (3.67-11.91). Seventy-six proteins (17.4%) had a MW  $<10$  kDa or  $>100$  kDa, and 84 (19.1%) had a pI  $>9$ . Of these, 23 proteins had a pI  $>9$  and a MW  $<10$  kDa or  $>100$  kDa (Table I). Twenty-four proteins (5.5%) were hydrophobic, and the remaining 416 proteins (94.5%) were hydrophilic. Using the program TMHMM 2.0, we found 38 proteins (8.6%) with at least one transmembrane helix.

**GO and GO enrichment/depletion analysis.** Of the 440 proteins, 312 (70.9%) had a biological process annotation. Sixty-three proteins (82.9%) with a MW  $<10$  kDa or  $>100$  kDa, and 61 proteins (72.6%) with a pI  $>9$  had a biological process annotation. Compared with the entire list of the IPI (IPI\_Human, version 3.13, 57,050 protein sequences), 41/22 GO terms were found to be enriched/depleted within the biological process annotation. Fig. 2 shows the enrichment/depletion analysis of the biological process annotation

**Comparison with GO analysis of urine.** Outcomes of the GO enrichment/depletion analysis of superficial BTCC were in good agreement with the results of urine proteome analysis. There were 16/7 GO terms found to be enriched/depleted in both the cancer and the urine proteome. Fig. 2 shows the outcome of the comparison of the GO analysis of urine.

Table I. Proteins with an isoelectric point (pI) &gt;9 and a molecular weight (MW) &lt;10 kDa or &gt;100 kDa.

IPI name	Swiss-Prot name	MW (kDa)	PI	GRAVY	GO biological process
IPI00215884	SFRS1	7.90	10.37	-1.16	Protein metabolism
IPI00012345	SFRS6	8.12	11.42	-1.55	mRNA processing
IPI00003377	SFRS7	8.34	11.83	-1.40	RNA metabolic process
IPI00301503	TRA2B	8.50	11.25	-1.62	RNA metabolic process
IPI00010204	SFRS3	8.89	11.64	-1.52	mRNA processing
IPI00009071	FUSIP	9.08	10.33	-1.76	None
IPI00477967		9.11	11.51	-1.34	None
IPI00013891	TRA2A	9.34	11.27	-1.59	Nuclear mRNA splicing
IPI00020546	ASH1L	132.33	9.46	-0.79	Cell-cell signaling
IPI00385783		133.57	9.62	-0.74	None
IPI00013043	P25A	133.61	9.48	-0.76	Protein metabolic process
IPI00386633		137.22	11.91	-0.85	Cell differentiation
IPI00026302	RL31	137.44	10.54	-0.81	Translation
IPI00175096		189.25	9.02	-0.89	Translation
IPI00303335	NEBU	191.10	9.10	-0.85	Organ development
IPI00289014	EVG1	205.60	9.83	-0.83	None
IPI00247583	RL21	220.57	10.49	-0.90	Translation
IPI00215965	ROA1	312.53	9.17	-0.91	None
IPI00011274	HNRDL	316.07	9.59	-0.95	Regulation of transcription
IPI00164755		331.78	9.08	-0.90	Signal transduction
IPI00002828	CENPU	332.79	9.18	-1.00	Regulation of transcription
IPI00249998		591.05	9.58	0.07	None
IPI00007188	ADT2	773.21	9.76	0.02	Transport

## Discussion

The LCM technology we employed has the advantage of accuracy; no cell lesions and no influence on the record before or after incision was observed. LCM is a revolutionary technique that deals with the problem of cell heterogeneity. However, there are a number of technical limitations to LCM, at least concerning the study of the proteome. First, LCM for use in proteomic research requires fresh frozen tissue, since formalin extensively crosslinks proteins. Unfortunately, most archival tissue specimens in pathology departments use formalin, requiring researchers to collect fresh tissue at the time of surgery. Second, LCM can be time consuming. Li *et al* reported that 50,000-100,000 cells were needed to extract more than 100  $\mu$ g proteins for analysis, while in our study it took approximately 50 h to microdissect 250,000 cells from each sample, which produced only 180  $\mu$ g of soluble proteins (9).

In our expression profile, 76 proteins (17.4%) had a MW <10 kDa or >100 kDa and 84 proteins (19.1%) had a pI >9. Of these, 23 proteins had a pI >9 and a MW <10 kDa or >100 kDa. These proteins were poorly represented in the 2-DE analyses. Compared with routine 2-DE, shotgun strategy increases the coverage of the expressed proteome. Several proteomic studies involving the identification of biomarkers for bladder cancer have been reported. However, these studies only focused on proteins with a moderate pI and MW. Sheng *et al* investigated 10 tumor-related proteins of transitional cell carcinoma by proteomic analysis. These proteins were concentrated

in a narrow pI range (4.9-6.6) and a narrow MW range (14.5-38.7 kDa) (4). A recent study revealed that proteins with extreme pI and MW can also serve as potential biomarkers for bladder cancer (10). In our study, most of the proteins with an extreme pI and MW had a biological process annotation, for example, the 23 proteins with a pI >9 and MW <10 kDa or MW >100 kDa. As listed in Table I, most of these proteins play important roles in cell biology, including cell-cell signaling, signal transduction and cell proliferation. This indicates that shotgun strategy can facilitate the identification of biomarkers, and that proteins with an extreme pI and MW should be highly valued as biomarkers.

Proteome analysis by the shotgun approach has reached a high level of sophistication with respect to sample processing, data acquisition and data analysis. However, a number of significant challenges remain. These are primarily related to the complexity of proteomes, which has so far precluded true proteomic analyses (that is, the analysis of all the components of a proteome) and has generated partially overlapping datasets from identical samples. There are also challenges to be overcome regarding the analysis of the information contained in proteomic datasets. Collectively, these problems have created the impression that published proteomic data are at times of dubious quality (11). Moreover, 2D-LC-MS/MS cannot offer quantitative information. Proteins that are identified from the tumor tissue itself and are known to be involved in the cancer process are reliable candidate biomarkers for cancer, and expressing quantity directly associated with the detection

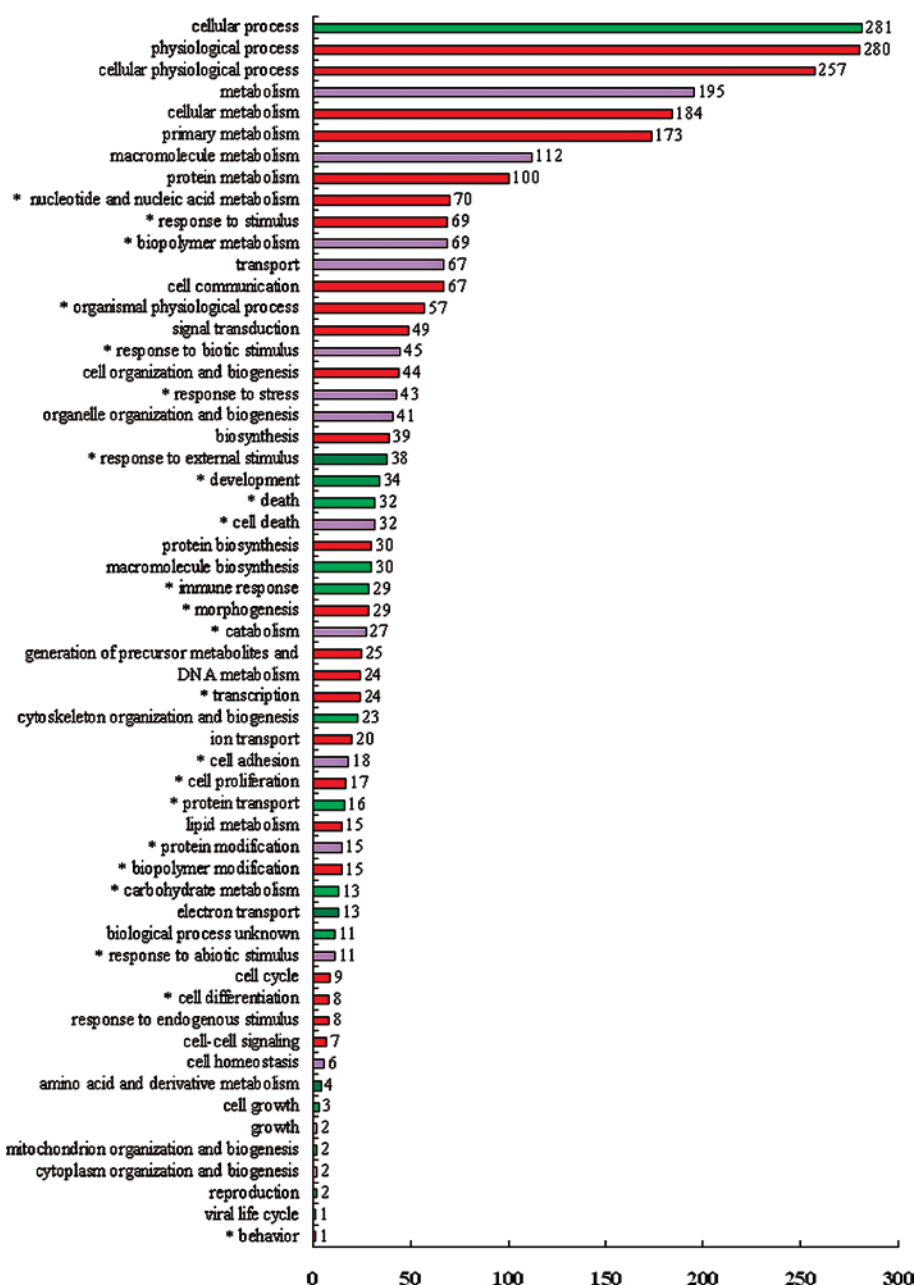


Figure 2. Enriched/depleted GO biological process terms for the set of identified proteins. Red indicates significantly enriched terms; purple, enriched terms; dark green, significantly depleted terms; and green, depleted terms. \*Terms overlapping with GO enrichment/depletion analysis of the urine proteome.

rate for abundant proteins tend to be detected in biofluid (12). Lack of quantitative information is another striking limitation of 2D-LC-MS/MS. Although innovations such as stable isotope labeling, spectral counting and spectral feature analysis appear to offer quantitative information, the effectiveness of these technologies is still disputed (13).

For a given expression profile, it is more important to identify the enrichment/depletion ontology category rather than just the functional distribution. Compared with all the human proteins available on the IPI, the enrichment/depletion of the functional classes means that the proteins in these categories are over/under-represented in this profile and reflect the specific biological categories of these data. The enriched proteins are more likely to be present in biofluid and

more likely to be identified as candidate biomarkers. In the biological process category, 41 GO terms were enriched and 22 GO terms were depleted. The GO terms related to cell proliferation, cellular metabolism, generation of precursor metabolites and energy, DNA metabolism, macromolecule metabolism, macromolecule biosynthesis and protein metabolism, which are associated with the biological characteristics of cancer cells, were significantly enriched. Additionally, the significant enrichment of GO terms related to immune and stimulus response represent the tumor-host immune interaction. The depletion of GO terms related to signal transduction and cell-cell signaling was unexpected. It is known that there is extensive crosstalk between cancer and stromal cells, and this crosstalk may be mediated through direct heterotypic cell-



cell contacts or through secreted molecules such as growth factors, cytokines, chemokines, extracellular matrix proteins, proteinases, proteinase inhibitors and lipid products (14,15). It is uncertain whether the removal of stromal cells by LCM brings about the depletion of these GO terms. However, a proteomic analysis of stromal cells is being carried out in our laboratory, and a reasonable explanation may be provided by comparing the expression profiles of cancer and stromal cells.

Urine is an ideal biofluid for the discovery of a BTCC biomarker panel (16,17). In order to confirm our experimental findings and to verify which type of protein was more likely to be detected in urine, we retrieved a recently published report on the human urinary proteome, in which urine samples were obtained from healthy volunteers and the cells in the urine were removed. After being clustered by BiNGO software, proteins in 128 GO biological process terms were significantly enriched, including cell adhesion, cell proliferation, cell differentiation, response to stimulus, response to stress and immune response (18). As shown in Fig. 2, this corroborates our findings. The molecular and genetic changes in cell adhesion, cell proliferation and cell differentiation were 'hot-spots' in the TCC research, and were shown to be enriched in our profile by GO analysis (19). Since proteins in these three interrelated processes were enriched in both cancer cells and urine, we conclude that the proteins involved in these processes are effective candidates for the comparison of independent complementary markers to detect cancer from urine. Meanwhile, the roles of these proteins in BTCC research require further investigation and validation.

GO enrichment analysis can easily generate a list of hundreds of proteins belonging to different biological processes. Examining the urine of cancer patients for these proteins requires reagents in the form of antibodies, aptamers, and/or isotopically labeled peptides. No single laboratory can be expected to produce these reagents for even a single process, and even if the reagents are readily available, no single laboratory can examine the proteins in the urine of cancer patients for more than one process. Another limitation of this procedure involves the complexity of the urine proteome, which includes proteins from the bladder, kidney and prostate. Therefore, in order to identify specific biomarkers for bladder cancer from urine, the proteins from these other genitourinary organs must be eliminated. This challenging task cannot be undertaken by a single laboratory.

## References

1. Landis SH, Murray T, Bolden S, *et al*: Cancer statistics, 1999. *CA Cancer J Clin* 49: 8-31, 1999.
2. Barocas DA and Clark PE: Bladder cancer. *Curr Opin Oncol* 20: 307-314, 2008.
3. Sylvester RJ: Natural history, recurrence, and progression in superficial bladder cancer. *Sci World Journal* 27: 2617-2625, 2006.
4. Sheng KH, Yao YC, Chuang SS, Wu H and Wu TF: Search for the tumor-related proteins of transition cell carcinoma in Taiwan by proteomic analysis. *Proteomics* 6: 1058-1065, 2006.
5. Memon AA, Chang JW, Oh BR and Yoo YJ: Identification of differentially expressed proteins during human urinary bladder cancer progression. *Cancer Detect Prev* 29: 249-255, 2005.
6. Gygi SP, Corthals GL, Zhang Y, Rochon Y and Aebersold R: Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc Natl Acad Sci USA* 97: 9390-9395, 2000.
7. Washburn MP, Wolters D and Yates JR III: Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 19: 242-247, 2001.
8. Li D, Li JQ and Yang SG: An integrated strategy for functional analysis in large-scale proteomic research by gene ontology. *Prog Biochem Biophys* 32: 1026-1029, 2005.
9. Li C, Hong Y, Tan YX, *et al*: Accurate qualitative and quantitative proteomic analysis of clinical hepatocellular carcinoma using laser capture microdissection coupled with isotope-coded affinity tag and two-dimensional liquid chromatography mass spectrometry. *Mol Cell Proteomics* 3: 399-409, 2004.
10. Barboro P, Rubagotti A, Orecchia P, *et al*: Differential proteomic analysis of nuclear matrix in muscle-invasive bladder cancer: Potential to improve diagnosis and prognosis. *Cell Oncol* 30: 13-26, 2008.
11. Nesvizhskii AI, Vitek O and Aebersold R: Analysis and validation of proteomic data generated by tandem mass spectrometry. *Nat Methods* 4: 787-797, 2007.
12. Smalley DM and Ley K: Plasma-derived microparticles for biomarker discovery. *Clin Lab* 54: 67-79, 2008.
13. Panchaud A, Affolter M, Moreillon P and Kussmann M: Experimental and computational approaches to quantitative proteomics: status quo and outlook. *J Proteomics* 71: 19-33, 2008.
14. Tlsty TD and Coussens LM: Tumor stroma and regulation of cancer development. *Annu Rev Pathol* 1: 119-150, 2006.
15. Li H, Fan X and Houghton J: Tumor microenvironment: the role of the tumor stroma in cancer. *J Cell Biochem* 101: 805-815, 2007.
16. Kim WJ and Bae SC: Molecular biomarkers in urothelial bladder cancer. *Cancer Sci* 99: 646-652, 2008.
17. Chiong E, Gaston Ke and Grossman HB: Urinary markers in screening patients with hematuria. *World J Urol* 26: 25-30, 2008.
18. Adachi J, Kumar C, Zhang Y, Olsen JV and Mann M: The human urinary proteome contains more than 1500 proteins including a large proportion of membrane proteins. *Genome Biol* 7: 186-202, 2006.
19. Williams SG and Stein JP: Molecular pathways in bladder cancer. *Urol Res* 32: 373-385, 2004.