Original Article

DOI: 10.5582/bst.2016.01055

iTRAQ-based quantitative proteomic analysis reveals potential early diagnostic markers of clear-cell Renal cell carcinoma

Limin Zhang[1,*], Haowen Jiang[1,*], Gang Xu[1,], Nan Chu[2,3,4], Ningxing Xu[5], Hui Wen[1], Bin Gu[1], Jun Liu[1], Shanghua Mao[1], Rong Na[1], Yan Jing[1], Qiang Ding[1], Yuanfang Zhang[1], Ling Wang[2,3,4,**]**

[1] *Department of Urology, Huashan Hospital, Fudan University, Shanghai, China;*
[2] *Laboratory for Reproductive Immunology, Hospital & Institute of Obstetrics and Gynecology, IBS, Fudan University Shanghai Medical College, Shanghai, China;*
[3] *Shanghai Key Laboratory of Female Reproductive Endocrine Related Diseases, Shanghai, China;*
[4] *The Academy of Integrative Medicine of Fudan University, Shanghai, China;*
[5] *Shanghai Medical College, Fudan University, Shanghai, China.*

Summary **Early detection is the key to improve the prognosis of kidney cancer. This study profiled and identified differentially expressed serum proteins in stage T1a renal cell carcinoma (RCC) using isobaric tags for relative and absolute quantification (iTRAQ)-based mass spectrometry. A total amount of 99 serum samples including 29 patients with ccRCC, 24 patients with a benign kidney mass, 28 patients with another type of urological tumor (20 cases of transitional cell carcinoma and 8 cases of prostate cancer or a male genital tumor), and 18 healthy controls were subjected to iTRAQ-based mass spectrometry. ProteinPilot software was used to identify the differentially expressed serum proteins in RCC compared to the other three populations. Hierarchical clustering analysis according to The Cancer Genome Atlas (TCGA) RCC database was then performed as the cross-platform validation. Immunohistochemistry was performed to verify the expression of selected proteins in tissue samples from these subjects. iTRAQ identified 27 differentially expressed serum proteins in the RCC patients, and 11 of these proteins were cross validated in RCC tissues from the TCGA database. The expression of C1QC, C1QB, S100A8, S100A9, ceruplasmin, and lumican was verified and associated with the tumor stage and/or grade. There were 27 differentially expressed proteins in early-stage RCC identified by iTRAQ; among them, the expression of C1QC, C1QB, S100A8, S100A9, ceruplasmin, and lumican were associated with the tumor stage and/or grade. Further studies are needed to confirm these data for their use as biomarkers for the early detection of RCC.**

Keywords: Renal cell carcinoma, iTRAQ, proteomics, tumor marker, bioinformatics

1. Introduction

Renal cell carcinoma (RCC) is a significant health problem in adults, accounting for more than 100,000 worldwide cancer-related deaths each year (*1*), and is the most lethal of the common urological cancers (*2*). Clear-cell type compromises about 80% of RCCs. With the improvement of surgical techniques in radical and partial nephrectomy, the postoperative overall 5-year survival of patients with organ-confined disease has increased to 97.9% in stage T1aN0M0, 94.9% in stage T1bN0M0, and 88.4% in stage T2N0M0 (*3*). Favorable prognosis relies on the acute early diagnosis of RCC (5-year survival rate-85%). Unfortunately, RCC often presents with few signs, symptoms, or laboratory abnormalities; thus, RCC is frequently only diagnosed at the advanced stage of the disease, when the prospects for a cure are dismal (a five-year survival rate of 9%) (*4*). Early diagnosis of RCC is usually based on traditional manifestations such as pain,

Released online in J-STAGE as advance publication June 17, 2016.

*These authors contributed equally to this works.
**Address correspondence to:
Dr. Gang Xu, Department of Urology, Huashan Hospital, Fudan University, 12 Wulumuqi Middle Road, Shanghai 200040, China.
E-mail: drxugang@163.com

Dr. Ling Wang, Obstetrics & Gynecology Hospital of Fudan University, 413 Zhaozhou Road, Shanghai 200011, China.
E-mail: Dr.wangling@fudan.edu.cn

mass and hematuria, which are not always effective. Diagnosis and subsequent resection of the RCC are also not accurate because it only relies on the basis of imaging findings. The absence of biomarkers in early detection of RCC in clinical practices has complicated the early diagnosis and treatment of RCC. A non-invasive test in serum or urine will have a significant impact on patient management.

To date, many researchers have been dedicated to the identification of potential biomarkers for the early detection of RCC using different proteome-based techniques, including the newly emergent mass spectrometry method using isobaric tags for relative and absolute quantification (iTRAQ) (*5,6*). iTRAQ-based mass spectrometry, first introduced by Ross *et al*. (*7*) in 2004, is a technique of multiplexed protein quantitative mass spectrometry using amine-reactive isobaric tagging reagents. The advantage of iTRAQ is that it enables quantitation of multiple samples simultaneously but requires only a small amount of sample. It is utilized frequently in research of many other malignant diseases to investigate their unique proteomic profile (*8-12*) and thus provide a new method to explore the biomarkers in malignant tumors.

In this study, we used quantitative proteomic analysis by iTRAQ-based proteomic identification technology to identify proteins that are dysregulated in the serum and LC-MS of T1a clear-cell RCC (ccRCC) and healthy people. Bioinformatics analysis and cross-platform validation according to the Cancer Genome Atlas (TCGA) RCC database, which catalogs genetic mutations as well as the gene/microRNA expression/regulation responsible for renal cancer, were also used in this study. To the best of our knowledge, this study is the first to highlight that T1a stage ccRCC accounts for 70-80% of all RCCs and always has a worse prognosis compared with other histological types of RCC (*13*). Profiling of the differentially expressed serum proteins in the mentioned patient populations will help us to understand the etiology of ccRCC and even provide a new strategy for the early detection of ccRCC.

2. Materials and Methods

2.1. *Study population and serum samples*

The subjects included in this study were from the Department of Urology at Huashan Hospital, Fudan University from January 2010 to December 2015. A total number of 99 serum samples were obtained including 29 patients with ccRCC, 24 patients with a benign kidney mass, 28 patients with another type of urological tumor (20 cases of transitional cell carcinoma and 8 cases of prostate cancer or a male genital tumor), and 18 healthy controls. Patients with accompanying kidney diseases, cardiovascular disorders, or other cancers were excluded. All ccRCC patients had stage T1a cancer as confirmed by radiological evaluation and pathological assessment of the surgical specimens. Sera from each of these four groups of subjects were pooled and abbreviated as R for ccRCC, C for benign kidney masses, M for other urological tumors, and H for noncancer controls. The serum samples were preserved at -76°C before iTRAQ analysis. This study was approved by the Institutional Review Board of Huashan Hospital, Fudan University, China (#KY2011-026, version 01.2011.1.12 and #2011-017), and written consent was obtained from each participant before enrollment into this study. The clinical information and data were obtained from industrial or hospital records. The clinical characteristics of the subjects in this study are shown in Table S1 (*http://biosciencetrends.com/docindex.php?year=2016&kanno=3*).

2.2. *iTRAQ labeling and nanoscale liquid chromatography coupled to tandem mass spectrometry (nano-LC-MS/MS)*

For iTRAQ analysis, 200 μL of each serum sample was dissolved and digested in a trypsin buffer (2 μg of trypsin in 40 μL of dissolution buffer) in a 37°C water bath for 16-18 h, then centrifuged, collected, and quantified by optical density (OD) measurements at an absorbance of 280 nm. Based on the OD_{280} values, an equivalent amount of peptide in each sample was subjected to iTRAQ labeling, according to the iTRAQ protocol (Applied Biosystems/MDS Sciex, Foster City, CA, USA). The samples were marked with iTRAQ tags as follows: iTRAQ117 for R, iTRAQ116 for C, iTRAQ115 for M, and iTRAQ114 for H. In order to obtain reliable results, the iTRAQ labeling experiment was replicated with the iTRAQ tags rearranged as follows: iTRAQ115 for R, iTRAQ117 for C, iTRAQ116 for M, and iTRAQ114 for H.

Next, the iTRAQ-labeled samples were pooled and fractionated by strong cation exchange chromatography on a polysulfoethyl 4.6 × 100 mm column (5 μm, 200 Å) (PolyLC Inc., Columbia, MD, USA) with a linear gradient of 0-500 mM KCl (10 mM KH_2PO_4, pH 3.0; 25% acetonitrile) over 75 min at a flow rate of 1 mL/min. According to the chromatography results, the collected fractions were recombined into 10 fractions and then freeze-dried. After that, each of the freeze-dried fractions from the SCX column was redissolved in 100 μL of 0.1% formic acid aqueous solution and then desalted using Vydac C18 supermicro-centrifugation chromatography. The sample was then extracted and analyzed by nano-LC-MS/MS. The nano-LC-MS/MS was completed by a system composed of an AB SCIEX Triple Time-of-Flight (TOF) 5600 mass spectrometer (Concord, Canada) and a liquid chromatograph with a cHiPLC nanoflex chip driven by an Eksigent nanoUltra 2D Plus nano-LC (Dublin, CA, USA). Each sample was run through a sampling course and subsequent separation with tandem MS analysis. The mode of

tandem MS was information-dependent acquisition (IDA). The resolving power of the screening performed by TOF MS was a full width at half maximum of 30,000, and the range was *m/z* 350-1250 in 250 ms. The top 30 abundant peptides with an ion peak greater than 120 counts/s and a range of charge from +2 to +5 were chosen to be analyzed by MS/MS with a range of *m/z* varying from 100 to 1800 in 100 ms for each TOF MS/MS screening time. The dynamic exclusion time was 18 s. When MS/MS was performed, the functions of enhanced iTRAQ splitting and auto-calculation of collision energy (AutoCE) were launched.

2.3. *Calculation of the ratios of differentially expressed proteins*

ProteinPilot (AB SCIEX, Framingham, MA, USA, 2012) software was then used to analyze the raw data. The signaling of the iTRAQ114 group served as the internal reference for the signal intensity, and all signals were normalized to it. The weighted average of the ratios of the respective peptides was calculated based on the protein quantitation results. False discovery rate analysis was conducted, and the detected protein threshold was set at less than 0.01. The proteomic database used in this study was the International Protein Index (IPI), Human v3.87. fasta (*http://www.ebi.ac.uk/*, Copyright© The European Bioinformatics Institute, 2013). Finally, the ratios of proteins of these four groups of samples with different iTRAQ tag labels were calculated as the averages of the ratios from two runs with different labeling sequences.

2.4. *Integration and hierarchical clustering analysis of iTRAQ data against the TCGA database on RCC*

The TCGA database catalogs genetic mutations as well as the gene/microRNA expression/regulation responsible for cancer risk and development using recently developed high-throughput genomic analysis techniques that were initiated in 2005. TCGA provides genomic characterization and sequence analysis of more than 20 different tumor types. In this study, the gene/microRNA expression pattern and clinical information of more than 500 patients with kidney renal clear-cell carcinoma (KIRC) were downloaded from the TCGA data portal (*https://tcga-data.nci.nih.gov/tcga/*).

For cluster profiling, the median expression value of each protein across the samples was set to zero. Cluster 3.0 and Tree View software (*http://rana.lbl.gov/EisenSoftware.htm*) were used for the cluster analysis and representation (*14*). The hierarchical clustering was performed on both genes and samples. Using a tree algorithm, these differentially expressed proteins were organized based on similarities in the expression profile. This allowed us to visualize and select genes based on individual expression profiles.

2.5. *Cross-platform validation vs. the TCGA database on RCC*

In the TCGA KIRC database, there are 31 normal control, 197 T1, 49 T2, 162 T3, and 6 T4 samples as well as 31 normal control, 5 G1, 173 G2, 169 G3, and 66 G4 samples. In this database, the expression levels of certain genes in these different grades or stages of KIRC were measured. The Student's *t*-test and analysis of variance were performed to statistically analyze the data using a cutoff of $p < 0.05$ and a fold change > 1.5 between different groups with different grades or T stages.

2.6. *Gene Ontology (GO) analysis*

The differentially expressed serum proteins in RCC were classified according to the GO category, including "biological process," "cell component," and "molecular function," respectively, by using the US National Institutes of Health (NIH) gene annotation software DAVID 6.7 (*15*) (*http://david.abcc.ncifcrf.gov/*; *http://david.abcc.ncifcrf.gov/gene2gene.jsp*). The differentially expressed protein list selected by ProteinPilot software was inputted into NIH DAVID 6.7 and thought to be significantly enriched if the *p* value was less than 0.05 using Fisher's exact test and the fold enrichment was more than 2-fold.

3. Results

3.1. *iTRAQ-based mass spectrometry profiling of the differentially expressed serum proteins from ccRCC patients*

In this study, we profiled differentially expressed serum proteins in 99 serum samples including 29 patients with stage T1a ccRCC, 24 patients with a benign renal mass, 28 patients with another type of urological tumor (18 for TCC), and 18 healthy controls using iTRAQ-based mass spectrometry. The basic clinical characteristics of subjects were collected as supporting information (Supplemental Data Table S1, *http://biosciencetrends.com/docindex.php?year=2016&kanno=3*)

Overall, 263 serum proteins were identified after two runs of iTRAQ and validated the ratios of proteins from both runs by a scatter diagram with ratios in the first run on the x-axis and ratios in the replicated run on the y-axis (Figure 1). The results from the two runs were consistent (Supplemental Data Table S2, *http://biosciencetrends.com/docindex.php?year=2016&kanno=3*). Using cutoff values of 1.5-fold for overexpression and 0.67-fold for underexpression of a protein, 74 differentially expressed proteins were identified in ccRCC vs. healthy controls (Table 1), with the identification of another 27 proteins (4 underexpressed and 23 overexpressed in ccRCC vs. the other three groups) according to the data from the TCGA database (Table 2; Figure 1).

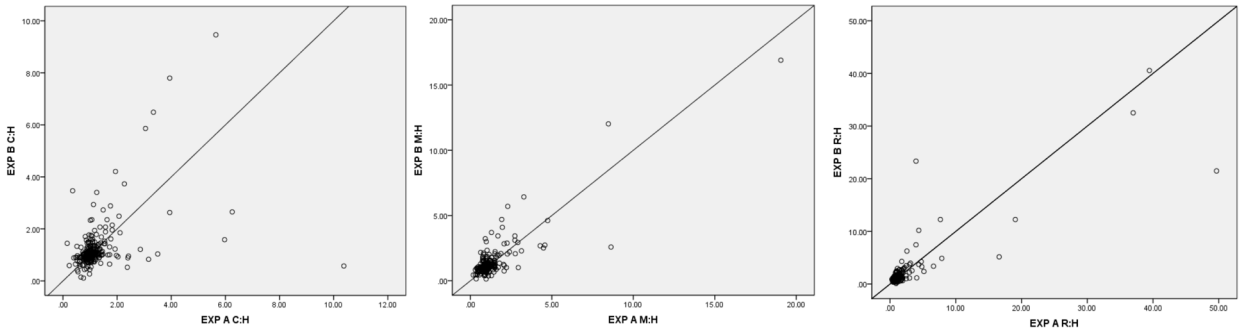


Figure 1. iTRAQ identification and validation of differentially expressed serum proteins in RCC. In this study, we first poured serum samples together from each of these four groups of subjects and performed the iTRAQ analyses two times. A total of 263 proteins after two iTRAQ runs were identified and validated, and the data were plotted on graphs. The x-axis indicates the value of the ratios in the first run, while the y-axis refers to those of the second run. Grouping of the differentially expressed proteins in this study. Note: R indicates patients with ccRCC, C indicates patients with a benign kidney mass; M indicates patients with another type of urological tumor; and H indicates noncancer controls.

Table 1. Differentially expressed serum proteins in ccRCC *vs.* healthy controls

IPI Serial No.	Gene symbol	Fold change for R:H	IPI Serial No.	Gene symbol	Fold change for R:H
IPI00290078.5	*KRT4 Cdna FLJ58275*	39.99825	IPI00329775.8	*Cpb2*	1.7606
IPI00009866.7	*Krt13*	35.56875	IPI00022392.1	*C1qa*	1.73835
IPI00022389.1	*Crp*	34.74575	IPI00010295.1	*Cpn1*	1.73765
IPI00909059.5	*KRT6A Cdna FLJ53910*	15.6504	IPI00021885.1	*Fga*	1.7024
IPI00643948.3	*C1qb*	13.6396	IPI00829636.2	*Ighd*	1.6924
IPI00218918.5	*Anxa1*	10.8741	IPI00879573.1	*Serpind1*	1.67605
IPI00783987.2	*C3*	9.9511	IPI00879709.3	*C6*	1.67525
IPI00022394.2	*C1qc*	7.27555	IPI00006114.5	*Serpinf1*	1.6597
IPI00025426.3	*Pzp*	6.3729	IPI00027235.1	*Atrn*	1.637
IPI01010737.1	*A2m*	5.69595	IPI00796990.4	*CFI Cdna FLJ58124*	1.62265
IPI00022974.1	*Pip*	4.9899	IPI00216065.3	*Proz*	1.6131
IPI00007047.1	*S100a8*	4.4943	IPI00914948.1	*Apol1*	1.5716
IPI01025667.1	*SERPINA3 Cdna FLJ35730 Fis*	4.40515	IPI00298828.3	*Apoh*	1.5514
IPI00641737.2	*Hp*	4.0558	IPI00023014.3	*Vwf*	1.53935
IPI00009867.3	*Krt5 Keratin*	4.03175	IPI00974055.1	*Crisp3*	1.52375
IPI00218407.6	*Aldob*	3.76455	IPI00032220.3	*Agt*	1.5207
IPI00022395.1	*C9*	3.5246	IPI00964994.1	*Habp2*	1.51505
IPI00027462.1	*S100a9*	3.33155	IPI00925621.1	*Il1rap*	0.65685
IPI00939824.1	*Cfb*	3.0278	IPI00479116.2	*CPN2 Carboxypeptidase N Subunit 2*	0.6502
IPI00029739.5	*Cfh*	2.9667	IPI00410333.2	*Treml1*	0.6474
IPI00553177.1	*Serpina1*	2.9385	IPI00296176.2	*F9*	0.64315
IPI00022417.4	*Lrg1*	2.6462	IPI00925635.1	*Igfals*	0.631
IPI00743766.2	*Fetub*	2.6123	IPI00010779.4	*Tpm4*	0.615
IPI00032291.2	*C5 Complement C5*	2.583	IPI00645849.1	*Ecm1*	0.6088
IPI00017601.1	*Cp*	2.52705	IPI00021856.3	*Apoc2*	0.60815
IPI00060800.5	*Zg16b*	2.3912	IPI00293925.2	*Fcn3*	0.58375
IPI00552768.1	*Txn*	2.2411	IPI00328609.3	*Serpina4*	0.57975
IPI00896380.1	*Ighm*	2.21655	IPI00021304.1	*Krt2 Keratin*	0.5755
IPI00019038.1	*Lyz*	2.1762	IPI00643348.4	*COMP Cdna FLJ60724*	0.56795
IPI00007240.2	*F13b*	2.0814	IPI00026314.1	*Gsn*	0.56025
IPI00023019.1	*Shbg*	2.07155	IPI00021854.1	*Apoa2*	0.559
IPI00019580.1	*Plg*	2.0611	IPI00647915.1	*Tagln2*	0.5075
IPI00021841.1	*Apoa1*	2.0512	IPI00940723.2	*Tnxb*	0.49205
IPI00953689.1	*Ahsg*	2.046	IPI00020986.2	*Lum*	0.48095
IPI00296608.6	*C7*	2.02405	IPI00163207.1	*Pglyrp2*	0.4501
IPI00218732.4	*Pon1*	1.9846	IPI00296099.6	*Thbs1*	0.43985
IPI00021727.1	*C4bpa*	1.87715	IPI00022446.1	*Pf4*	0.3987

R: patients with ccRCC; H: noncancer controls.

3.2. *Cross-platform analysis of the TCGA database vs. our differentially expressed proteins*

To compare these 27 differentially expressed serum proteins, we downloaded the kidney tissue gene expression patterns from 31 normal controls and 205 patients with T1 ccRCC from the TCGA data portal (*https://tcga-data.nci.nih.gov/tcga/*). We first analyzed the 27 gene expression patterns in the 30 normal controls and the 202 patients with T1 ccRCC, 1 normal control and 3 patients with T1a ccRCC (Figure 2A). Hierarchical and heat map analyses of 27 gene expression patterns showed that 9 (*C1qc, C1qb, Anxa1, Lyz, Cp, Agt, Zg16b, S100a8, And S100a9*) were upregulated and 2 (*Serpina4 And Lum*) were downregulated in both serum and tissue samples (Table 1). These 11 genes were selected for

Table 2. iTRAQ identification of differentially expressed serum proteins in RCC *vs.* the TCGA database

IPI serial No.	Proteins	Fold Ratios for differentially expressed proteins			Genes significantly increased/decreased in RCC tissues (TCGA)						Tissue genes	Serum proteins
					Normal Control		T1 RCC		T test			
		R:H	R:C	R:M	mean	standard deviation	mean	standard deviation	*p*	FC		
IPI00022394	C1QC	7.27555	2.213129	2.682626	15.24679	11.88908	134.6874	105.5283	1.59E-09	8.833825	up	up
IPI00643948	C1QB	13.6396	2.323038	2.81	20.60832	15.42208	196.4306	158.2369	2.96E-09	9.531614	up	up
IPI00218918	ANXA1	10.8741	9.902199	5.510198	42.0473	17.41306	103.3125	57.49486	1.39E-08	2.457055	up	up
IPI00019038	LYZ	2.1762	1.849881	1.518473	25.78253	22.49118	162.1104	131.2444	2.63E-08	6.287605	up	up
IPI00017601	CP	2.52705	1.936363	3.053099	4.599597	3.137652	165.0615	257.5653	0.000636	35.88608	up	up
IPI00032220	AGT	1.5207	2.268178	2.98821	14.15519	14.82322	34.54053	36.61214	0.002489	2.440132	up	up
IPI00060800	ZG16B	2.3912	2.046472	2.236962	0.160283	0.173884	0.410872	0.498079	0.00603	2.563412	up	up
IPI00007047	S100A8	4.4943	4.060075	2.269848	6.23674	6.064987	23.22495	34.21335	0.006364	3.723893	up	up
IPI00027462	S100A9	3.33155	3.610067	2.434098	13.1029	11.65109	36.49804	50.11201	0.010362	2.785493	up	up
IPI00328609	SERPINA4	0.57975	0.415815	0.534455	3.579333	3.009381	0.350075	1.142392	4.00E-23	0.097804	down	down
IPI00020986	LUM	0.48095	0.544061	0.632996	99.6874	75.92487	38.25532	64.23287	2.35E-06	0.383753	down	down
IPI00022389	CRP	34.74575	18.17294	1.93252	0.065947	0.127943	10.21652	91.41613	0.5377624	154.9205	up	up
IPI00009866	KRT13	35.56875	74.85006	64.85914	0.117634	0.344086	1.591729	18.09628	0.6511775	13.53121	up	up
IPI00290078	KRT4	39.99825	49.96658	54.30855	0.009406	0.015682	0.054731	0.167793	0.1348779	5.818911	up	up
IPI00909059	KRT6A	15.6504	2.858338	14.81905	0.048689	0.217767	0.248698	1.004503	0.2712461	5.107877	up	up
IPI00022417	LRG1	2.6462	2.094756	2.142672	2.409539	2.371319	7.391328	15.39787	0.0739449	3.067528	up	up
IPI00009867	KRT5	4.03175	2.778984	6.042791	0.314106	1.006641	0.520658	1.560287	0.4758168	1.657587	up	up
IPI00022974	PIP	4.9899	2.956102	3.536804	0.063886	0.0992	0.086039	0.292671	0.6769375	1.34676	up	up
IPI00552768	TXN	2.2411	3.140335	2.834862	197.7238	82.62611	207.598	107.3115	0.6242722	1.049939	up	up
IPI00021854	APOA2	0.559	0.611966	0.444427	1.883181	9.391258	1.823045	22.01121	0.988057	0.968066	down	down
IPI00218407	ALDOB	3.76455	4.097023	1.634345	1380.982	1424.919	64.59174	254.2692	1.27E-26	0.046772	down	up
IPI00743766	FETUB	0.4501	0.371523	0.524042	0.231238	0.197704	0.045323	0.237875	4.88E-05	0.196001	down	up
IPI00218732	PON1	1.9846	0.403858	0.424259	0.235652	0.507248	0.051987	0.168255	9.61E-05	0.220607	down	up
IPI00953689	AHSG	2.046	0.665561	0.51198	1.873805	9.962775	0.229014	2.373435	0.0432425	0.122219	down	up
IPI00021841	APOA1	2.0512	0.542983	0.36589	1.162947	4.175419	0.414824	2.427988	0.1541711	0.356701	down	up
IPI01025667	SERPINA3	4.40515	2.090275	1.718748	14.25598	22.14037	12.79282	27.03231	0.7743703	0.897365	down	up
IPI00163207	PGLYRP2	0.4501	0.371523	0.524042	0.010583	0.031642	0.210003	0.579237	0.0569706	19.84334	up	down

There were 4 downregulated and 23 upregulated serum proteins in T1a stage RCC, the top 11 of which were consistent with the tissue data from the TCGA database. R: T1a RCC; H: healthy subjects; C: patients with a benign mass lesion in the kidney; M: patients with another type of malignant urological tumor.

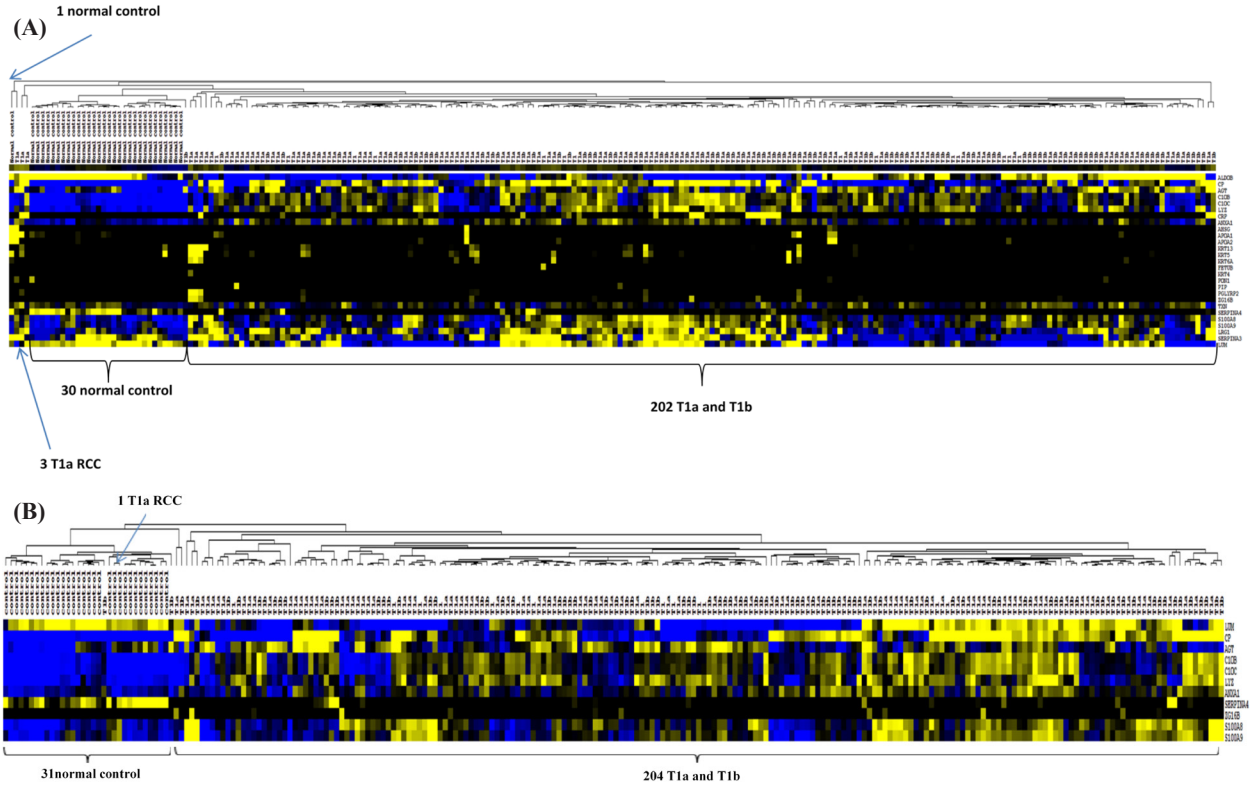


Figure 2. Analysis of differentially expressed serum proteins in RCC. (A), Hierarchical analysis of the 27 genes against data from the TCGA database including 31 normal controls and 205 patients with T1 RCC. The data indicate that the expression pattern of these 27 genes can distinguish normal controls from those with T1 ccRCC. **(B)**, Hierachical analysis of the selected 11 genes against data from the TCGA database including 31 normal controls and 205 patients with T1 RCC. The data indicate that the expression pattern of these 11 genes can distinguish the normal controls from patients with T1 RCC.

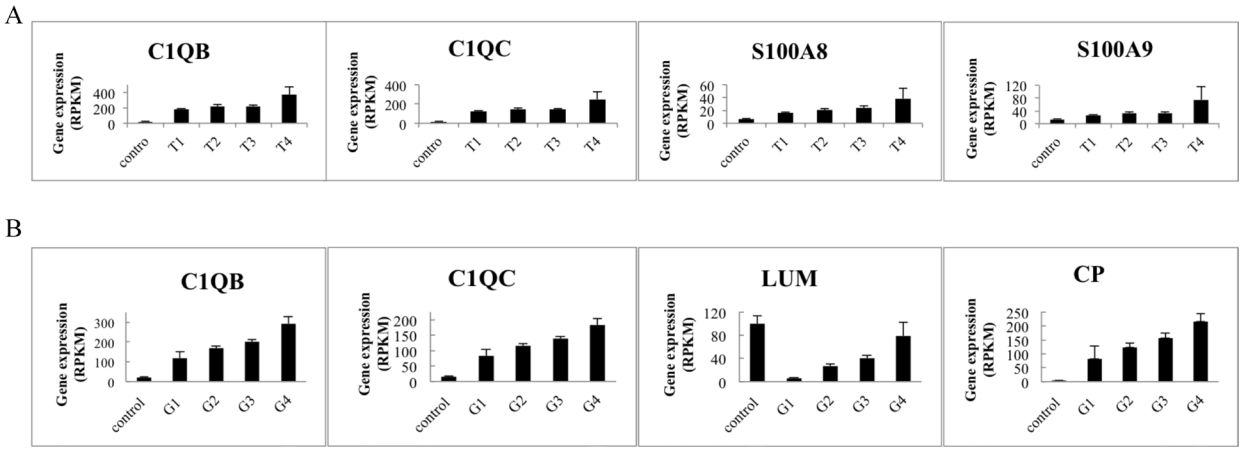
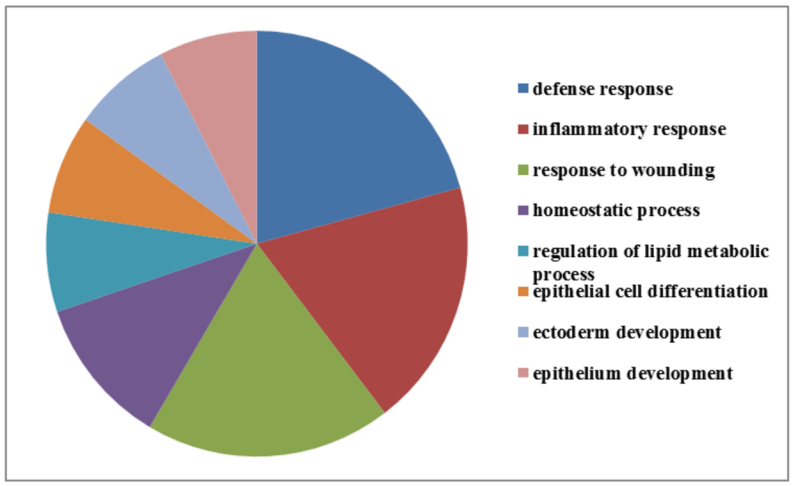


Figure 3. Association of different protein expression with the RCC stage and grade. (A), Association of C1QC, C1QB, S100A8, S100A9, and LUM expression with the RCC stage. We selected these differentially expressed serum proteins for association with the RCC stage using data from the TCGA database including 31 normal controls and 205 patients with T1 RCC. **(B)**, Association of C1QC, C1QB, CP, and LUM expression with the RCC grade. We selected these differentially expressed serum proteins for association with the RCC grade using data from the TCGA database including 31 normal controls and 205 patients with T1 RCC.

further analyses. Hierarchical analysis of 11 genes in the 31 normal controls and 205 patients with T1 RCC from the TCGA database indicates that the expression pattern of the 11 proteins can better distinguish normal controls from T1a RCC (Figure 2B).

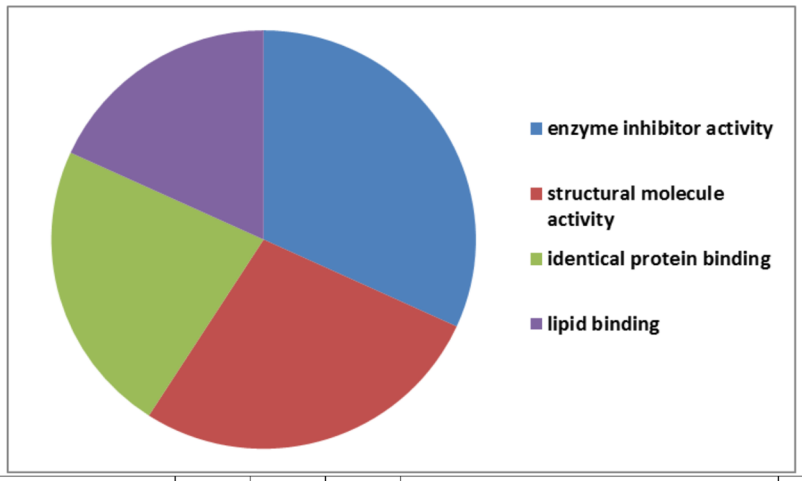
3.3. *Association of selected proteins with the tumor stage and grade in comparison to the TCGA database*

Next, we selected 11 proteins from these 27 differentially expressed serum proteins for confirmation



Category	Term	Count	%	PValue	Genes	Fold Enrichment	FDR
GOTERM_BP	defense response	11	40.74074	1.38E-08	C1QB, APOA2, S100A8, PGLYRP2, CRP, S100A9, ANXA1, SERPINA3, LYZ, C1QC, AHSG	10.52018381	2.07E-05
GOTERM_BP	inflammatory response	10	37.03704	9.02E-10	C1QB, APOA2, S100A8, CRP, S100A9, ANXA1, SERPINA3, LYZ, C1QC, AHSG	18.09765886	1.35E-06
GOTERM_BP	response to wounding	10	37.03704	6.41E-08	C1QB, APOA2, S100A8, CRP, S100A9, ANXA1, SERPINA3, LYZ, C1QC, AHSG	11.097621	9.6E-05

Figure 4. Gene ontology (biological process) analysis of the differentially expressed serum proteins in ccRCC *vs.* that of NIH DAVID. We performed gene ontology analysis to group the functions of these differentially expressed serum proteins in ccRCC using data from the NIH DAVID database.



Category	Term	Count	%	PValue	Genes	Fold Enrichment	FDR
GOTERM_MF	enzyme inhibitor activity	7	25.92593	5.74E-06	APOA2, FETUB, AGT, SERPINA4, ANXA1, SERPINA3, AHSG	14.02485	0.006241
GOTERM_MF	structural molecule activity	6	22.22222	0.004416	KRT6A, KRT5, LUM, ANXA1, KRT13, KRT4	5.119479	4.701003
GOTERM_MF	identical protein binding	5	18.51852	0.024534	C1QB, APOA2, APOA1, ALDOB, PON1	4.226237	23.67922
GOTERM_MF	lipid binding	4	14.81481	0.043804	APOA2, APOA1, ANXA1, PON1	4.808519	38.57092

Figure 5. The molecular functions of the differentially expressed serum proteins identified by the iTRAQ technique. They are mainly categorized into four classes: 1) enzyme inhibitor activity, 2) structural molecule activity, 3) identical protein binding, and 4) lipid binding.

and association with the clinicopathological data from the TCGA database. Our data showed that the expression of C1QC, C1QB, S100A8, and S100A9 proteins was significantly increased in RCC and associated with an advanced stage of disease ($p < 0.01$; Figure 3A; Supplemental Data Table S3, *http://biosciencetrends.com/docindex.php?year=2016&kanno=3*). The expression of ceruplasmin (CP) was also significantly increased, whereas lumican (LUM) was significantly decreased in the RCC samples vs. the controls ($p < 0.01$; Figure 3B; Supplemental Data Table S4, *http://biosciencetrends.*

com/docindex.php?year=2016&kanno=3).

3.4. *Identification of the functions of these differentially expressed serum proteins in ccRCC*

After that, we employed NIH DAVID 6.7 software (http://david.abcc.ncifcrf.gov/home.jsp) to assess the potential functional categories of these 27 differentially expressed serum proteins in early ccRCC. The cell component analysis in GO by NIH DAVID (Supplemental Data Figure S1, *http://biosciencetrends.com/docindex.php?year=2016&kanno=3*) showed that S100A8, S100A9, LUM, ZG16B, C1QC, C1QB, SERPINA4, and CP proteins are secreted proteins and that LUM, LYZ, AGT, ZG16B, C1QB, CP, and SERPINA4 proteins are found in the extracellular region. Biological process analysis presented that most of these 11 genes are involved in the host defense response, inflammatory response, and the response to wounding, indicating that the identified proteomic expression pattern might reflect the early change in the serum microenvironment of ccRCC (Figure 4). As for the molecular functions of the differentially expressed serum proteins, they might be involved in four kinds of biological processes including enzyme inhibitor activity, structural molecule activity, identical protein binding, and lipid binding (Figure 5).

4. Discussion

The treatment and prognosis of ccRCC largely depends on the tumor stage and the pre-existing conditions of the patients. In other words, surgery is still the best option for treating these patients, if diagnosed at an early stage. The lack of biomarkers to distinguish the malignant and benign lesions is the biggest challenge in ccRCC. Imaging has limited accuracy and might lead to unnecessarily surgery. Biopsy is a kind of accurate detection method for ccRCC but the invasive procedure and other side effects might limit its use. A non-invasive test such as urine or a blood-based test might provide a new idea for RCC diagnosis. In the current study, we performed proteomics analysis by using iTRAQ and several proteins have been identified as promising biomarkers for the accurate diagnosis of ccRCC although further investigation is still needed. In our study, we identified 27 differentially expressed serum proteins, 11 of which were cross-validated in RCC tissues against the TCGA database. Moreover, we found that expression of C1QC, C1QB, S100A8, S100A9, CP, and LUM proteins was associated with the RCC stage and/or grade. Of the dysexpressed proteins, 8 of them including S100A8, S100A9, LUM, ZG16B, C1QC, C1QB, SERPINA4, and CP proteins were secreted proteins and thus have the potential to be used as a serum biomarker in the diagnosis of ccRCC, especially in the early stage.

iTRAQ technology, the proteomic technology that we chose to use in this study, has been utilized frequently for the assessment of differentially expressed proteins in many other diseases (*8-12*), and previous publications have shown promising findings in the field of oncologic proteomic research. In RCC, to date, there have been two pioneering studies using the iTRAQ technique: one was related to the von Hippel-Lindau gene in RCC cells (*5*), and the other one investigated gene profiling in RCC tissues at different stages (*6*). To the best of our knowledge, our current study is the first to highlight differentially expressed serum proteins from patients with stage T1a disease, since these patients are usually diagnosed during a health check-up. Moreover, we also utilized the IPI human proteomic database to identify certain differentially expressed proteins and the NIH DAVID database to functionally analyze these differentially expressed serum proteins. The TCGA KIRC database was used for cross-platform validation of the selected serum proteins at the relevant genetic level in a totally different population. Thus, we identified 27 differentially expressed serum proteins in ccRCC patients vs. healthy controls, those with benign renal masses, and those with another type of urological tumor. Four proteins were underexpressed and 23 were overexpressed. Hierarchical analysis showed that these 27 genes could constitute a proteomic expression pattern that is able to distinguish T1 ccRCC from normal controls. However, the regulation of these 27 proteins in the tissue samples was not exactly consistent with the levels of serum proteins measured by the iTRAQ test. This might be attributed to the possibility of post-transcriptional modifications and the inherited inaccuracy of this protein quantification.

In addition, we found that the expression of the 11 differentially expressed serum proteins (C1QB, C1QC, ANXA1, LYZ, CP, ACT, ZG16B, S100A8, S100A9, SERPINA4, and LUM) was associated with the RCC tumor stage and grade. Functionally, C1QB and C1QC are complement subcomponent subunits (*16*) and are related to antibody-dependent and -independent immune responses in the human body (*17*). They are able to induce apoptosis of prostate cancer cells by targeting the tumor suppressor WWOX and hampering cell adhesion *via* a mechanism which is still unknown (*18*). As RCC has been demonstrated to be immunogenic, the association between C1q and the carcinoma is worthy of further study. Moreover, S100A8 and S100A9 are EF-hand Ca^{2+} binding proteins. They are abundant in the cytosol of phagocytes and play a critical role in numerous cellular processes, such as motility and danger signaling, by interacting and modulating the activity of target proteins. The expression levels of S100A8 and S100A9 are increased in many types of cancer, including gastric, colon, pancreatic, bladder, ovarian, thyroid, breast, skin, and prostate cancers, but not in ccRCC according to previous studies (*19-24*). All four of these proteins are secreted proteins and are involved in the

defense response, according to the GO analysis by NIH DAVID.

Our study suggested that ccRCCs might have 8 protein biomarkers that play important roles in the immune response, especially the defense response, although larger scale studies are needed to get more specific conclusions. The relationship between immune response including defense response and RCC needs further investigation and whether other defense response related proteins could also be biomarkers for early detection are also of great value. Dysregulated immune response is thought to be related to the rapid cell proliferation, metastasis and lower apoptosis of cancer cells.

Our findings are consistent with previous studies on functional analyses in ccRCC. The link between dysregulated immune response and ccRCC is not surprising. The reaction of the immune system plays a dual role in the development of carcinoma including ccRCC. The immune system could identify and control the proliferation of cancer cells to play the immunosurveilance function (*25*). However, on the other hand, it could also promote chronic inflammation, immunoselection of defective immunogenic variants and even suppress antitumor immunity and thus provide a beneficial microenvironment for progression of carcinoma. Defective immune-editing to balance activation and inhibition is responsible for the occurrence, angiogenesis, metastasis, apoptosis and inhibition of malignant lesions including in ccRCC (26).

Global analysis through protein analysis and pathway analysis may provide new significant applications in clinical practice. It might provide more than diagnostic markers. The pathway-derived metabolic products especially those proteins that are involved in the defense response might be predictive or even prognostic markers in patients and thus offer new methods for a deeper comprehension of malignant lesions (*27*).

In conclusion, we identified several serum proteins that can distinguish ccRCC and other controls in serum. Most of them are involved in the biological defense response. If the promising proteins could be confirmed as dysregulated in the serum of ccRC patients that would suggest their potential as noninvasive biomarkers for early detection of RCC. The functional analysis of these proteins in RCC may lead to a novel mechanism of RCC development and progression, possibly revealing a novel strategy to treat RCC patients in the future.

Acknowledgements

This work was supported in part by grants from the National Natural Science Foundation of China (#81001130) (to Gang Xu), the National Key Clinical Specialty Project of the Chinese Ministry of Health, and the Excellence of Research Project of Fudan University (to Gang Xu). This work was supported by the National Natural Science Foundation of China No. 31571196 (to Ling Wang), the Science and Technology Commission of Shanghai Municipality 2015 YIXUEYINGDAO project No. 15401932200 (to Ling Wang), the FY2008 JSPS Postdoctoral Fellowship for Foreign Researchers P08471 (to Ling Wang), the National Natural Science Foundation of China No. 30801502 (to Ling Wang), the Shanghai Pujiang Program No. 11PJ1401900 (to Ling Wang), Development Project of Shanghai Peak Disciplines-Integrated Chinese and Western Medicine No.20150407.

References

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA Cancer J Clin. 2011; 61:69-90.
2. Campbell S LB. Malignant renal tumors. Philadelphia: Elsevier Saunders, 2012, p. 1413-1474.
3. Kanao K, Mizuno R, Kikuchi E, Miyajima A, Nakagawa K, Ohigashi T, Nakashima J, Oya M. Preoperative prognostic nomogram (probability table) for renal cell carcinoma based on TNM classification. J Urol. 2009; 181:480-485; discussion 485.
4. Weiss RH, Lin PY. Kidney cancer: Identification of novel targets for therapy. Kidney Int. 2006; 69:224-232.
5. Haffey WD, Mikhaylova O, Meller J, Yi Y, Greis KD, Czyzyk-Krzeska MF. iTRAQ proteomic identification of pVHL-dependent and -independent targets of Egln1 prolyl hydroxylase knockdown in renal carcinoma cells. Adv Enzyme Regul. 2009; 49:121-132.
6. Siu KW, DeSouza LV, Scorilas A, Romaschin AD, Honey RJ, Stewart R, Pace K, Youssef Y, Chow TF, Yousef GM. Differential protein expressions in renal cell carcinoma: New biomarker discovery by mass spectrometry. J Proteome Res. 2009; 8:3797-3807.
7. Ross PL, Huang YN, Marchese JN, *et al*. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics. 2004; 3:1154-1169.
8. Muraoka S, Kume H, Watanabe S, Adachi J, Kuwano M, Sato M, Kawasaki N, Kodera Y, Ishitobi M, Inaji H, Miyamoto Y, Kato K, Tomonaga T. Strategy for SRM-based verification of biomarker candidates discovered by iTRAQ method in limited breast cancer tissue samples. J Proteome Res. 2012; 11:4201-4210.
9. Leong S, Nunez AC, Lin MZ, Crossett B, Christopherson RI, Baxter RC. iTRAQ-based proteomic profiling of breast cancer cell response to doxorubicin and TRAIL. J Proteome Res. 2012; 11:3561-3572.
10. Zeng GQ, Zhang PF, Li C, Peng F, Li MY, Xu Y, Yu FL, Chen MJ, Yi H, Li GQ, Chen ZC, Xiao ZQ. Comparative proteome analysis of human lung squamous carcinoma using two different methods: Two-dimensional gel electrophoresis and iTRAQ analysis. Technol Cancer Res Treat. 2012; 11:395-408.
11. Rehman I, Evans CA, Glen A, Cross SS, Eaton CL, Down J, Pesce G, Phillips JT, Yen OS, Thalmann GN, Wright PC, Hamdy FC. iTRAQ identification of candidate serum biomarkers associated with metastatic progression of human prostate cancer. PLoS One. 2012; 7:e30885.
12. Boylan KL, Andersen JD, Anderson LB, Higgins L, Skubitz AP. Quantitative proteomic analysis by

iTRAQ(R) for the identification of candidate biomarkers in ovarian cancer serum. Proteome Sci. 2010; 8:31.
13. Beck SD, Patel MI, Snyder ME, Kattan MW, Motzer RJ, Reuter VE, Russo P. Effect of papillary and chromophobe cell type on disease-free survival after nephrectomy for renal cell carcinoma. Ann Surg Oncol. 2004; 11:71-77.
14. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A. 1998; 95:14863-14868.
15. Dennis G, Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. 2003; 4:P3.
16. Nayak A, Pednekar L, Reid KB, Kishore U. Complement and non-complement activating functions of C1q: A prototypical innate immune molecule. Innate Immun. 2012; 18:350-363.
17. Hong Q, Sze CI, Lin SR, Lee MH, He RY, Schultz L, Chang JY, Chen SJ, Boackle RJ, Hsu LJ, Chang NS. Complement C1q activates tumor suppressor WWOX to induce apoptosis in prostate cancer cells. PLoS One. 2009; 4:e5755.
18. Healy J, Tipton K. Ceruloplasmin and what it might do. J Neural Transm (Vienna). 2007; 114:777-781.
19. Gebhardt C, Nemeth J, Angel P, Hess J. S100A8 and S100A9 in inflammation and cancer. Biochem Pharmacol. 2006; 72:1622-1631.
20. Emberley ED, Murphy LC, Watson PH. S100 proteins and their influence on pro-survival pathways in cancer. Biochem Cell Biol. 2004; 82:508-515.
21. Heizmann CW, Ackermann GE, Galichet A. Pathologies involving the S100 proteins and RAGE. Subcell Biochem. 2007; 45:93-138.
22. Hermani A, De Servi B, Medunjanin S, Tessier PA, Mayer D. S100A8 and S100A9 activate MAP kinase and NF-kappaB signaling pathways and trigger translocation of RAGE in human prostate cancer cells. Exp Cell Res. 2006; 312:184-197.
23. Ott HW, Lindner H, Sarg B, Mueller-Holzner E, Abendstein B, Bergant A, Fessler S, Schwaerzler P, Zeimet A, Marth C, Illmensee K. Calgranulins in cystic fluid and serum from patients with ovarian carcinomas. Cancer Res. 2003; 63:7507-7514.
24. Ghavami S, Rashedi I, Dattilo BM, Eshraghi M, Chazin WJ, Hashemi M, Wesselborg S, Kerkhoff C, Los M. S100A8/A9 at low concentration promotes tumor cell growth *via* RAGE ligation and MAP kinase-dependent pathway. J Leukoc Biol. 2008; 83:1484-1492.
25. Kaufman HL. Vaccines for melanoma and renal cell carcinoma. Semin Oncol. 2012; 39:263-275.
26. Bockorny B, Dasanu CA. Intrinsic immune alterations in renal cell carcinoma and emerging immunotherapeutic approaches. Expert Opin Biol Ther. 2013; 13:911-925.
27. Arsanious A, Bjarnason GA, Yousef GM. From bench to bedside: Current and future applications of molecular profiling in renal cell carcinoma. Mol Cancer. 2009; 8:20.

(Received March 31, 2016; Revised May 28, 2016; Accepted June 1, 2016)