Analyst

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/analyst

Preliminary Raman spectroscopic study of Urine: Diagnosis of breast cancer in animal models

T. Bhattacharjee[a#], A. Khan[a#], G. Maru[b], A. Ingle[c] and C. Murali Krishna*

[a]Chilakapati Lab, ACTREC, Kharghar, Navi-Mumbai, 410210

[b] Maru Lab, ACTREC, Kharghar, Navi-Mumbai, 410210

[c] Laboratory Animal Facility, ACTREC, Kharghar, Navi-Mumbai, 410210

[#]Authors have contributed equally to this work.

*Corresponding author's Address

Dr. C. Murali Krishna

Scientific Officer 'F' and Principal Investigator

Chilakapati Laboratory

Advanced Center for Treatment Research and Education in Cancer (ACTREC)

Tata Memorial Center (TMC)

Kharghar, Sector '22'

Navi Mumbai – 410210, INDIA

Telephone: +91-22-2740 5039

Email - mchilakapati@actrec.gov.in, pittu1043@gmail.com

Keywords: Breast cancer, Urine, Early Detection, Raman spectroscopy, PCA, PC-LDA

Abbreviations used: RS- Raman Spectroscopy, PC-LDA- Principal Component Linear Discriminant Analysis, LOOCV - Leave One Out cross validation

Abstract:

Prognosis of breast cancer, the most common cancer in females worldwide, has been shown to improve with early detection. Owing to disadvantages like low sensitivity, specificity, tedious sample preparation, long output times and inter-observer variance of currently available screening/ diagnostic tools; rapid, objective alternatives such as Raman spectroscopy (RS) are being extensively explored. Body fluid (serum, saliva) based RS assays have shown promising results in diagnosis of oral, lung and nasopharyngeal cancers. The current study aims to explore the feasibility of breast cancer diagnosis using urine based RS. In this study, spectra were acquired from unprocessed as well as concentrated urine of controls (C) and breast tumor bearing (T) rats and analyzed using Principal Component Analysis (PCA) and Principal Component - Linear Discriminant Analysis (PC-LDA). Classification efficiencies of 80% and 72% using unprocessed urine and 78% and 91% using concentrated urine for C and T rats respectively were achieved. Thus, results suggest possibility of breast cancer diagnosis using urine based RS. Further, spectra were also acquired from concentrated urine samples collected prior breast tumor development (TT) in rat and from rats that did not develop tumor despite carcinogen treatment (NTT). Concentrated urine of NTT rats could be classified as 'normal' (C or NTT) with ~ 83% efficiency whereas concentrated urine from visibly and palpably normal rats that eventually developed tumor (TT rats) could be classified as 'abnormal' (TT or T) with ~ 72.5% efficiency using PC-LDA. These results suggest possibility of detecting biochemical changes occurring prior tumor development using urine based RS.

1. Introduction:

Breast cancer is the most frequently diagnosed cancer and the leading cause of cancer death in females worldwide [1]. In 2012, 226,870 new invasive breast cancer cases and 39,510 breast cancer deaths were estimated to occur in US [2]. In developing countries like India, incidence rates as high as 39.5 per 1, 00,000 women have been reported [3]. Literature suggests improved prognosis with early detection of breast cancer [4]. In lieu of this, efforts have been directed towards development of screening techniques for early detection of this cancer. Clinical breast exam (CBE) and mammography are the two most widely used screening tools [5, 6]. However, the sensitivity of CBE is low whereas mammography suffers from disadvantages like low positive predictive value (only 25%) [6], unsuitability for women with dense breast, radiation exposure, etc. Alternatives like ultrasonography, thermography, Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET) have low sensitivity, cannot detect small tumors and are expensive. The gold standard for diagnosis of breast cancer – histopathology, also suffer from several disadvantages like tedious sample preparation, long output times and inter-observer variance [7, 8]. Rapid, objective and preferably non-invasive alternate screening /diagnostic techniques are hence being extensively explored.

Raman spectroscopy (RS), a rapid, objective tool with a potential for non-invasive/ minimally invasive applications, has shown promising results in the diagnosis of cervix, lung, gastrointestinal, brain, oral, skin, colon [9-13] and several other cancers [14] including breast cancers [15-22]. RS is based on an inelastic scattering process where the energy of photon scattered by the sample is different from the incident photon due to transfer of energy to or from vibrational modes of molecules in the sample. Since bands of Raman spectrum are characteristic of specific molecular vibrations unique to a molecule, RS can provide chemical fingerprint / biochemical profile of a sample. The ability of this technique to classify normal breast tissues from benign and malignant tissues [16, 17], metastatic from non metastatic cell lines [23], and invasive carcinoma from ductal *in situ* carcinoma using cryopreserved sections [24] have been reported. Detection of constituents deep inside breast tissue phantoms [20, 21] as well as transcutaneous detection of breast tumors in rats has been established [25, 26]. Tumor margin

assessment using *in vivo* RS during partial mastectomy surgery has also been demonstrated [22]. However, the requirement of invasive procedures for spectroscopy in human subjects remains a problem. Use of body fluid (serum, saliva, urine) based RS may circumvent this problem. Apart from minimal invasiveness, body fluids based tests have several advantages like accessibility, multiple sampling, easy handling, storage and transportation. Serum based RS to classify normal from cancer serum and monitor asthma [27-32] and saliva based Surface Enhanced RS assay for diagnosis of oral, nasopharyngeal and lung cancer [33-35] have been reported. Diagnosis of bladder cancer using cells voided into urine has also been demonstrated [36]. However, feasibility of using urine for breast cancer diagnosis is yet to be explored.

The current study aims to explore the feasibility of breast cancer diagnosis using urine based RS. In this study, spectra were acquired from unprocessed urine of controls and breast tumor bearing rats. To check the possibility of better classification between normal and tumor, spectra were also acquired after concentrating (by dehydration in vacuum) urine of control and tumor bearing rats. In order to further explore the sensitivity of urine based RS in breast cancer diagnosis and possibility of early detection; spectra were also acquired from urine samples collected prior breast tumor development in rat. The data were analyzed using Principal component analysis (PCA) and Principal Component-Linear Discriminant Analysis (PC-LDA). Results of the study are reported in the manuscript.

2. Materials and Methods:

2.1. Animals: A total of 42 Sprague-Dawley (SD) rats were used in this study. 25 fifty days old SD rats were administered 65 mg/kg 7, 12 dimethyl benzanthracene (DMBA, Sigma-Aldrich, purity 95%) dissolved in groundnut oil (Dhara, India) ingtragastrically by gavage. 20 rats developed breast tumors (histopathologically confirmed adenocarcinoma of breast) approximately six months post carcinogen treatment. 17 fifty days old SD rat were administered oil (control). None of the control rats developed breast tumors. The study was approved by Institutional Animal Ethics Committee, ACTREC endorsed by the Committee for the Purpose of Control and Supervision of Experiments on Animals (CPCSEA), Government of India. All animals were housed under standard laboratory conditions, fed a diet of in-house-prepared pellets and provided with water ad libitum.

2.2. Urine collection: The rats were restrained; airlifted and voided urine was collected in sterile pertidishes. The urine was then transferred to sterile eppendorf tubes using a micropipette. Separate petridish and eppendorf were used for each rat urine sample. Approximately 150-200ul urine was collected per rat. The urine samples were snap frozen immediately after collection and stored in -80^{o}C.Using this procedure, urine was from six groups:

a) Unprocessed control urine (n=9): urine samples of control rats were thawed and spectra were acquired from these samples.

b) Unprocessed tumor urine (n=9): urine samples of tumor bearing rats were thawed and spectra were acquired from these samples.

c) Concentrated control urine (n=8): urine samples of control were thawed, dehydrated in vacuum using Speed Vac and rehydrated with 40ul normal saline before spectra acquisition.

d) Concentrated tumor urine (n=7): urine samples of tumor bearing were thawed, dehydrated in vacuum using Speed Vac and rehydrated with 40ul normal saline before spectra acquisition.

e) Concentrated TT urine (n=4): urine samples were collected 5 months post carcinogen (DMBA) treatment from visibly and palpably normal rats. These rats were palpated every two weeks after urine collection. Approximately 1 month post urine collection (~ 6 months post carcinogen treatment), these rats developed breast tumors. Biopsy followed by histopathology confirmed the tumors to be adenocarcinoma. The urine samples collected from these rats were labeled 'Tumors Treated' and will hence forth be referred to as 'TT'. Before spectra acquisition, these samples were dehydrated and rehydrated as described above.

f) Concentrated NTT urine (n=5): urine samples were collected 5 months post carcinogen (DMBA) treatment from visibly and palpably normal rats. However, these rats failed to develop tumor even 8 months post carcinogen treatment. Urine samples from these rats were labeled as 'No Tumors Treated' and henceforth are referred as 'NTT'. The urine was processed in the same way before spectra acquisition.

Urine samples from control, TT, NTT and tumor bearing rats were collected at the same time. Thus, all samples were collected from age matched rats.

2.3. Spectra acquisition: After passive thawing/rehydrating, samples were subjected to Raman spectroscopy by placing 40 µl sample on calcium fluoride (CaF_2) window and spectra were recorded using Fiber Optic Raman microprobe (Horiba-Jobin-Yvon, France). This Raman system consists of laser (785 nm, Process Instruments) as an excitation source and HE 785 spectrograph (Horiba-Jobin-Yvon, France) coupled with CCD (Synapse, Horiba-Jobin-Yvon) as dispersion and detection elements respectively. Optical filtering of **laser line and** Rayleigh signals, is accomplished through 'Superhead', the other component of the system. Optical fibers were employed to carry the incident light from the excitation source to the sample and also to collect the Raman scattered light from the sample to the detection system. Raman microprobe was assembled by coupling a 40X microscopic objective (Nikon, Japan) to the superhead. Spectral acquisition details were: Excitation wavelength (λ_{ex}) = 785 nm, laser power = 40 mW. Spectra were integrated for 10 seconds and averaged over 6 accumulations. On an average, 8 spectra were recorded from each sample to generate a total of 355 spectra under 6 groups, 81 spectra from unprocessed urine of control rats, 82 from unprocessed urine of tumor bearing rats, 64 spectra from concentrated urine of control rats, 56 from concentrated urine of tumor bearing rats, 40 spectra from concentrated urine of TT rats and 32 spectra from concentrated urine of NTT rats.

2.4. Spectral pre-processing:

Raman spectra from all urine samples were corrected for instrument response with a National Institute of Science and Technology (NIST) certified Standard Reference Material 2241 (SRM 2241) followed by the subtraction of background signals from optical elements and substrate. To remove interference of the low frequency background, first derivatives of spectra (Savitzky-Golay method and window size 3) were computed [37-39]. Spectra were interpolated in 600-1800 cm^{-1} region, vector-normalized and used as input for multivariate analysis.

2.5. Multivariate analysis

First derivative, vector normalized spectra were subjected to multivariate unsupervised Principal Component Analysis (PCA) and supervised Principal Component-Linear Discriminant Analysis (PC-LDA). PCA is a routinely used method for data compression and visualization while LDA provides data classification based on an optimized criterion which is aimed for better class separation. LDA can be used in companion with PCA to increase efficiency of classification. For this, PCA scores obtained using a set of significant PCs with maximum variance amongst data are used as input data for LDA based classification. The advantage of doing this is to remove or minimize noise from the data and concentrate on variables important for classification. LDA models were validated by Leave-one-out cross-validation (LOOCV). LOOCV is a type of rotation estimation used mainly for smaller datasets i.e. a technique useful for assessing performance of a predictive model with a hypothetical validation set when an explicit validation set is not available. Algorithms for these analyses were implemented in MATLAB (Mathworks Inc., USA) based software using in-house codes [40].

Mean spectra were computed from the background subtracted spectra prior to derivatization for each class, by averaging Y-axis variations for each class, and baseline corrected by fitting a fifth order polynomial function. These baselines corrected spectra were vector normalized and then used for computing mean spectra. Standard deviation was computed using background subtracted baseline corrected spectra. Difference spectra were also calculated by subtracting mean spectra of control group from tumor group, TT and NTT group.

3. Results and Discussion:

3.1. Spectral analysis:

a) Unprocessed control and tumor urine: Vector-normalized average spectrum (Figure 1a i) of control rat urine exhibit urea peaks at 1004cm^{-1} (symmetrical C–N stretch) and 1161cm^{-1} (attributed to NH2 modes) and creatinine peaks at 680 cm^{-1} (C–NH2 and C═O stretching, ring vibrations) and 850 cm^{-1} (C–NH2 deformation and ring vibrations), as reported elsewhere [41]. Mean tumor bearing rat urine spectra (Figure 1a i) showed differences in the intensities of several peaks, indicating difference in the concentration of the urine's biochemical components. Differences were seen in the intensity of specific

peaks such as decreased intensity of the peaks of urea (1006 cm^{-1}) and creatinine (680 cm^{-1}) in the cancer group compared to control. To elucidate the spectral variations amongst groups, difference spectra were computed by subtracting mean control spectrum from mean tumor spectrum, respectively. Subtraction of mean spectra is one of the conventional ways of looking at spectral differences, it provides differences over a selected spectral range and thus understanding of the moieties that may have been modified is facilitated. The positive peaks of difference spectrum are from the mean tumor spectrum while negative peaks are from mean control spectrum. Tumor – control difference spectra (Figure 1a ii) also show a prominent positive urea peak at 1006 cm^{-1} suggesting increased urea concentration in urine during cancer.

b) Concentrated control and tumor urine: Mean concentrated control rat urine spectrum have features similar to unprocessed urine with additional peaks at 653, 756, 781, 885 and 925 cm^{-1} (Figure 1b i). Mean tumor spectrum (Figure 1b i) show difference in the intensity of urea and creatinine peaks with respect to control. Tumor – control difference spectra show a prominent positive urea peak at 1006 cm^{-1} suggesting increased urea concentration in urine during cancer (Figure 1b ii).

c) Concentrated NTT and TT urine: Mean concentrated NTT (Figure 3a i) and TT (Figure 3a ii) rat urine show difference in the intensity of urea peak. TT mean spectrum exhibit highest intensity compared to control and tumor while NTT mean spectum show lowest concentration compared to all. TT – control (Figure 2b i) and Control – NTT (Figure 2b iii) difference spectra suggest higher urea concentration in TT compared to control and higher urea concentration in control compared to NTT. TT - NTT difference spectra (Figure 2b ii) also suggest increased urea concentration in TT compared to NTT.

Standard deviations for each group, unprocessed urine control (Figure 3a), unprocessed urine tumor (Figure 3b), concentrated urine control (Figure 3c), concentrated urine tumor (Figure 3d), concentrated urine NTT (Figure 3e) and concentrated urine TT (Figure 3f) to assess reproducibility of data processing and sample-sample variation.

Although differences in the urine biochemistry of control and tumor rats are observed, the link between breast cancer and urine is not yet completely known. Studies have indicated excretion of androgen, corticoid and estrogen metabolites in urine to be associated with breast cancer [42, 43]. However, further studies are needed to establish an underlying connection between urine and breast cancer.

Further, it is unclear why TT and NTT have huge difference while control and tumor spectral difference is comparatively smaller. A probable explanation is as follows. TT is pre-tumor condition. Intense biochemical changes may be expected for establishing a tumor. In NTT group, successful immune response resulting in aborted carcinogenesis may alter metabolism. Altered metabolism is linked with changed excretion profiles. Since the metabolic and biochemical alterations in TT and NTT are intense and different, a large difference is probably observed. Compared to this, in control there is no tumor and hence additional biochemical changes are not expected. In tumor group, cancer is established. Hence, minor changes and secretions from tumor are expected. Hence, difference between these groups is probably less. Extensive studies in this direction are warranted.

3.2. Multivariate analysis:

a) Unprocessed control and tumor urine: Preprocessed spectra interpolated in 600-1800 cm^{-1} range were subjected to PCA for delineating trends in the data set. PCA variance plot and loadings are shown in Figure 4a and b. As can be seen in Figure 4a, cumulative variance covered by factor 2 and 3 are 81% and 84% respectively. Scatter plot of PCA factors (Figure 4c) shows a tendency towards classification of control and tumor bearing rat unprocessed urine. To explore the feasibility of classifying the above groups, PC-LDA was used. To avoid over fitting, 9 factors [44] contributing ~ 80 % percent of correct classification; were used (Figure 5a). The plot of PC-LDA factors 1, 2 and 3 (Figure 5b) show clusters of control and tumor unprocessed urine spectra. The confusion matrix for PC-LDA model building shows 69/81 control spectra correctly classify as control while 11/81 misclassify as tumor; whereas 63/82 spectra are correctly classified as tumor while 19/82 spectra misclassify as control. Leave-one-spectrum-out-cross validation (LOsOCV) was carried out to evaluate the results obtained by PC- LDA. LOsOCV builds a model

based on all observations but one, and tests the left out observation against the model built; this is repeated until all observations are left out once. The performance is estimated in terms of classification efficiency, which is percentage of spectra from each group that are correctly classified. In analysis of LOsOCV as shown in Table 1a; 65/81 control spectra correctly classify as control while 16/81 misclassify as tumor; whereas 59/82 spectra are correctly classified as tumor while 23/82 spectra misclassify as control.

Urine is a complex colloidal solution consisting mainly of urea, creatinine, salts and colloids made of glycoprotein, proteins and mucopolysaccharides [45]. Their concentration ranges from 9.3g/L (urea) to 0.67g/L (creatinine). The meager quantity present enhances the possibility of irregular distribution of the components mentioned. Further, as breast cancer progresses, minute concentration changes occur in limited number of urine components while the concentration of all other components of urine remain unchanged. These factors may contribute greatly to the misclassification observed amongst groups. Overall, the classification efficiency of control and tumor (using unprocessed urine samples) group was 80 % and 72 % respectively.

b) Concentrated control and tumor urine: Spectra of control and tumor bearing rat concentrated urine interpolated in 600-1800 cm^{-1} range were also subjected to PCA. PCA variance plot and loadings are shown in Figure 6a and b respectively. As can be seen in Figure 6a, cumulative variance covered by factor 2 and 3 are 82% and 84% respectively. Scatter plot of PCA factors (Figure 6c) shows clusters of concentrated control and tumor bearing rat urine. To explore the feasibility of classifying the above groups from control, PC-LDA was used. 4 factors contributing ~ 85 % percent of correct classification were applied (Figure 7a). The plot of PC-LDA factors 1, 2 and 3 (Figure 7b) shows well separated clusters of control and tumor spectra. The confusion matrix for PC-LDA model building shows that 50/64 control spectra correctly classify as control while 14/64 misclassify as tumor; whereas 51/56 spectra are correctly classified as tumor while 5/56 spectra misclassify as control. LOsOCV of results of PC-LDA model built (Table 1b); 50/64 control spectra correctly classify as control while 14/64 misclassify as tumor; whereas 51/56 spectra are correctly classified as tumor while 5/56 spectra misclassify as

control. As discussed earlier, limiting concentration of urine components and their irregular distribution may explain the observed misclassification.

Although, in this case, the samples are concentrated, the total amount of components present in the sample analyzed is still very low. Since 150-200ul samples were concentrated and used for spectra acquisition, the total quantity of major component urea expected in one urine sample will be ~2ug. Thus, concentrating samples have higher quantity and probably more regular distribution of components compared to unprocessed urine, but the quantities being analyzed are meager and possibly results in the misclassification observed. Overall, the classification efficiency of control and tumor (unprocessed urine samples) group was 78 % and 91 % respectively. While the classification efficiency of control group in case of both unprocessed and concentrated control urine is ~80%, classification efficiency of tumor group in case of concentrated urine is higher (91%) compared to unprocessed urine (72%). Therefore, further studies were conducted using concentrated urine.

c) Concentrated NTT and TT urine: To further explore the sensitivity of urine based RS in diagnosis of breast cancer and possibility of early detection, as mentioned earlier, urine samples were also collected prior tumor development Spectra acquired from concentrated urine of control, tumor bearing, NTT and TT rats were preprocessed, interpolated in 600-1800 cm^{-1} range were subjected to PCA and PC-LDA. The PCA variance plot and loading factors 1 and 3 are shown in Figure 8a and b respectively. The TT spectra in the PCA scatter plot (Figure 8c) shows a tendency of classification. The PC-LDA scatter plot (Figure 9b) of factors 1, 2 and 3 shows overlapping clusters of control, tumor, TT and NTT. It is however noteworthy, that control and NTT populate the left side of the plot whereas TT and tumor lie on the right side. The results of PC-LDA in the form of confusion matrix shows that 34/ 64 spectra are correctly classified as control, while 9/ 64 misclassified as NTT, 11/ 64 misclassified as TT and 10/ 64 misclassified as tumor. 17/40 NTT spectra were correctly classified as NTT, while 15/40, 2/40 and 4/40 misclassified as control, TT and tumor respectively. In case of TT, 23/32 were correctly classified whereas 7/32 and 2/32 misclassified with control and tumor respectively. 30/ 56 tumor

spectra classified correctly while 3/56, 11/56 and 12/56 misclassified with control, NTT and TT. The results of LOsOCV are shown in Table 2a. As can be seen, 34/ 64 spectra are correctly classified as control, while 9/ 64 misclassified as NTT, 11/ 64 misclassified as TT and 10/ 64 misclassified as tumor. 17/40 NTT spectra were correctly classified as NTT, while 16/40, 2/40 and 5/40 misclassified as control, TT and tumor respectively. In case of TT, 22/32 were correctly classified whereas 8/32 and 2/32 misclassified with control and tumor respectively. 30/ 56 tumor spectra classified correctly while 3/56, 11/56 and 12/56 misclassified with control, NTT and TT.

Despite misclassification amongst groups, 67.2% control spectra classify as either control or NTT, while 82.5% NTT spectra classify as either control or NTT (Table 2c). Control rats were not treated with carcinogen whereas NTT rats did not develop tumor in spite of carcinogen treatment. Therefore, control and NTT urine spectra represent 'normal' (non cancerous) condition. TT rats were rats that eventually developed tumor whereas tumor group rats had breast tumors at the time of urine collection. Thus, urine spectra of TT and tumor group rats represent 'abnormal' (cancerous) condition. As observed in Table 2a, 75% TT and 75% tumor spectra correctly classified as abnormal (TT/ Tumor).

Result of Leave-One-Rat-Out-cross-validation (LOrOCV) implemented using 2 PC-LDA factors is shown in Table 2b. As can be seen, 30/ 64 spectra are correctly classified as control, while 17/ 64 misclassified as NTT, 8/ 64 misclassified as TT and 9/ 64 misclassified as tumor. 14/40 NTT spectra were correctly classified as NTT, while 20/40, 0/40 and 6/40 misclassified as control, TT and tumor respectively. In case of TT, 19/32 were correctly classified whereas 8/32, 1/32 and 4/32 misclassified with control, NTT and tumor respectively. 22/ 56 tumor spectra classified correctly while 2/56, 15/56 and 17/56 misclassified with control, NTT and TT. Thus, 73.4% and 85% control and NTT spectra respectively classify as control/NTT while 72% and 70% spectra from TT and tumor respectively classify as TT/ tumor. The results corroborate the outcome of LOsOCV.

In a nutshell, results suggest that rats that did not develop tumor could be classified as 'normal' (with ~ 83% efficiency) even though these rats were treated with carcinogen and

had high probability of developing tumor, while rats that did develop tumor after carcinogen treatment were classified as 'abnormal' (with 72.5% efficiency) using urine collected prior any visible or palpable abnormality.

4. Conclusion: The aim of the study was to explore the feasibility of breast cancer diagnosis using urine based RS. Using unprocessed urine, control and tumor bearing rats could be classified with ~ 80% and ~72% efficiency respectively. Using concentrated urine, control and tumor groups could be classified with ~ 80% and 91% efficiency. These results suggest possibility of cancer diagnosis using urine based RS. Further, concentrated urine of rats that did not develop tumor even after carcinogen challenge could be classified as 'normal' with 83% efficiency whereas concentrated urine from visibly and palpably normal rats that eventually developed tumor could be classified as 'abnormal' with 72.5% efficiency. These results suggest possibility of detecting biochemical changes occurring prior tumor development using RS. Further studies in this direction may help development of urine based RS as early breast cancer detection tool.

References:

1. A. Jemal, F. Bray, M. Center, J. Ferlay, E. Ward and D. Forman, CA Cancer J Clin, 2011, 61(2), 69–90.
2. R. Siegel, D. Naishadham and Ahmedin Jemal, CA Cancer J Clin, 2012, 62, 10–29.
3. A. Nandakumar, T. Ramnath and M. Chaturvedi, Indian J Surg Oncol 2010, 1(1), 8-9
4. Howlader N, Noone AM, Krapcho M, Ney man N, Aminou R, Altekruse SF, Kosary CL, Ruhl J, T atalov ic h Z, Cho H, Mariotto A, Eisner MP, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds), SEER Cancer Statistic s Review, 197 5–2009, National Cancer Institute. Bethesda, MD, 2001.
5. G.I. Andreea, R. Pegza, L. Lascu, S. Bondari, Z. Stoica and A. Bondari, Current Health Sciences Journal, 2011, 37(2), 55-61.
6. D. Kennedy, T. Lee and D. Seely, Integrative Cancer Therapies, 2009, 8(1), 9-16.
7. S.M. Ismail, A.B. Colclough, J.S. Dinnen, D. Eakins, D.M. Evans, E. Gradwell, J.P. O'Sullivan, J.M. Summerell and R.G. Newcombe, , BMJ, 1989, 298, 707-710.
8. J.B. Sørensen, F.R. Hirsch, A. Gazdar and J.E. Olsen, Cancer, 1993, 71, 2971-2976.
9. V. Chang, P. Cartwright, S. Bean, G. Palmer, R. Bentley and N. Ramanujam, Neoplasia, 2009, 11(4), 325–332.
10. H. Yamazaki, S. Kaminaka, E. Kohda, M. Mukai and H. Hamaguchi, Radiat. Med., 2003, 21(1), 1–6.
11. S. Teh, W. Zheng, K. Ho, M. Teh, K. Yeoh and Z.Huang, Br. J. Cancer, 2008, 98(2), 457–465.
12. Y. Zhou, C.H. Liu, Y. Sun, Y. Pu, S. Boydston-White, Y. Liu and R.R. Alfano, J Biomed Opt., 2012, 17(11), 116021.

13. S. P. Singh, A. Deshmukh, P. Chaturvedi and C. M. Krishna, J. Biomed. Opt., 2012, 17(10), 105002.

14. D.J. Evers, B.H. Hendriks, G.W. Lucassen and T.J. Ruers, Future Oncol., 2012, 8(3), 307–320.

15. R. R. Alfano, C. H. Liu, W. L. Sha, H. R. Zhu, D. L. Akins, J. Cleary, R. Prudente and E. Celmer, Laser Life Sci., 1991, 4(1), 23-28.

16. S. Haka, K.E. Shafer-Peltier, M. Fitzmaurice, J. Crowe, R.R. Dasari, M.S. Feld, Proc Natl Acad Sci USA, 2005, 102 (35), 12371–12376.

17. M. V. Chowdary, K. K. Kumar, S. Mathew, L. Rao, C. Murali Krishna and J. Kurien, Biopolymers, 2009, 91(7), 539-546.

18. M.V. Chowdary, K.K. Kumar, J. Kurien, P. Mathew and C.M. Krishna, Biopolymers, 2006, 83 (5), 556-569.

19. M.D. Keller, E. Vargis, N. Granja, R.H. Wilson, M.A. Mycek, M.C. Kelley and A. Mahadevan-Jansen, J Biomed Opt. , 2011, 16(7), 077006.

20. N. Stone, R. Baker, K. Rogers, A.W. Parker and P. Matousek, 2007, Analyst 132, 899–905.

21. N. Stone and P. Matousek, Cancer Res, 2008, 68 (11), 4424-4430.

22. A. Haka, Z. Volynskaya, J. A. Gardecki,J. Nazemi, J. Lyons, D. Hicks, M.Fitzmaurice,R. R. Dasari, J. P. Crowe and M. S. Feld, Cancer Res, 2006, 66 (6), 3317-3322.

23. C. Nieva, M. Marro, N. Santana-Codina, S. Rao and D. Petrov, PLoS One, 2012, 7(10): e46456.

24. S. E. Clare, H. Neubauer, C. Schuetz, R. Kurek, E. Solomayer, Dr. Wallwiene, D. F. Dye, J. M. Zaleski, R. J. Goulet and T. Fehm, J Clin Oncol, 2006, 24 (18S), 10619.

25. L. Raniero, R.A. Canevari, L.N. Ramalho, F.S. Ramalho, E.A. dos Santos, R.A. Bitar, K.J. Jalkanen, H.S. Martinho and A.A. Martin, Theor Chem Acc., 2011, 130 (4), 1239-1247.

26. T Bhattacharjee, P. Kumar, G Maru, A Ingle and C Murali Krishna , Lasers in Medical Science, 2014, 29 (1), 325-333.

27. A. Sahu ,S. Sawant, H. Mamgain and C. M. Krishna, Analyst 138(14), 4161-74 (2013).

28. J. L. Pichardo-Molina, C. Frausto-Reyes, O. Barbosa-García, R. Huerta-Franco, J. L. González-Trujillo, C. A. Ramírez-Alvarado, G. Gutiérrez-Juárez and C. Medina-Gutiérrez, Lasers in Medical Science, 2007, 22(4), 229-236.

29. S. Feng, R. Chen, J. Lin, J. Pan, G. Chen, Y. Li, M. Cheng, Z. Huang, J. Chen and H. Zeng, Biosens. Bioelectron, 2010, 25, 2414–2419.

30. J. L. Gonz´alez-Sol´ıs, J. Rodr´ıguez-L´opez, J. C. Mart´ınez-Espinosa, C. Frausto-Reyes, L. F. Jave-Su´arez, A. C. Aguilar-Lemarroy, H. Vargas-Rodr´ıguez and E. Mart´ınez-Cano, AIP, Conf. Proc., 2010, 1226, 91–95.

31. X. Li, T. Yang and S. Li, Appl. Opt., 2012, 51, 5038–5043.

32. A. Sahu, K. Dalal, S. Naglot, P. Aggarwal and C. Murali Krishna, PLoS One, 2013, 8(11), e78921.

33. J. C. Kah, K. W. Kho, C. G. Lee, C. J. Sheppard, Z. X. Shen, K. C. Soo and M. C. Olivo, International J. Nanomed., 2007, 2(4), 785-798.

34. S. Feng, D. Lin, J. Lin, Z. Huang, G. Chen, Y. Li, S. Huang, J. Zhao, R. Chen and H. Zeng, 2014, Appl. Phys. Lett., 104, 073702.

35. X Li, T. Yang and J. Lin, J. Biomed. Opt, 2012, 17(3), 037003.

36. A. Shapiro, O. N. Gofrit, G. Pizov, J.K. Cohen and J. Maier, Eur Urol., 2011, 59(1), 106-12.

37. A. Nijssen, K. Maquelin, P. J. Caspers, T. C. B. Schut, M. H. A. Neumann, J. C. den Hollander, L. F. Santos and G. J. Puppels, J Biomed Opt, 2007, 12, 034004-1 –7.

38. S. Koljenovic, L. P. Choo-Smith, T. C. B. Schut, J. M. Kros, H. J. Van den Berge and G. J. Puppels, Lab Invest, 2002, 2, 1265-1277.

39. T. J. Harvey, E. Gazi, A. Henderson, R. D. Snook, N. W. Clarke, M. Brown and P. Gardner, Analyst, 2009, 134, 1083-1091.

40. A. D.Ghanate, S. Kothiwale, S. P. Singh, D. Bertrand and C. M. Krishna, J Biomed Opt, 2011, 16, 025003 1 -9.

41. J. A. Bispo, E. E. Vieira, L. Silveira and A. B. Fernandes, J. Biomed. Opt., 2013, 18 (8), 087004.

42. Bulbrook et.al., The Lancet, ,298 (7721), 395–398

43. Kabat et. al., Cancer Epidemiol Biomarkers Prev, 1997, 6, 505-509.

44. J. Kelly, J. Trevisan, A. Scott, P. Crmichael, H. Pollock, P. Martin-Hirsch and F. Martin, J. Proteome Res., 2011, 10, 1437–1448.

45. A. J. Butt, E. A. Hauser and J. Seifter, Calif. Med., 1952, 76(3), 123-129.

Figure legends:

Figure 1:a) i) Mean spectra of unprocessed urine interpolated in 600-1800cm^{-1} region from control and tumor bearing rats, ii) tumor – control difference spectrum; b) mean spectra of concentrated urine interpolated in 600-1800cm^{-1} region from control and tumor bearing rats, ii) tumor – control difference spectrum

Figure 2: a) Mean spectra of concentrated urine interpolated in 600-1800cm^{-1} region from i) NTT and ii) TT rats, b) i) TT – control, ii) TT - NTT, and iii) control - NTT difference concentrated urine spectra

Figure 2: Mean and standard deviations for each group, a) unprocessed urine control b) unprocessed urine tumor c) concentrated urine control d) concentrated urine tumor e) concentrated urine NTT and f) concentrated urine TT

Figure 4: PCA of unprocessed urine from control and tumor bearing rats a) variance plot, b) Loading factors 2and 3, and c) Scatter plot.

Figure 5: PC-LDA of unprocessed urine from control and tumor bearing rats a) Scree plot b) scatter plot

Figure 6: PCA of concentrated urine from control and tumor bearing rats a) variance plot, b) Loading factors 2and 3, and c) Scatter plot.

Figure 7: PC-LDA of concentrated urine from control and tumor bearing rats a) Scree plot b) scatter plot

Figure 8: PCA of concentrated urine from control, NTT, TT and tumor bearing rats a) variance plot, b) Loading factors 1and 3, and c) Scatter plot.

Figure 9: PC-LDA of concentrated urine from control, NTT, TT and tumor bearing rats a) Scree plot b) scatter plot

Table legend:

Table 1: PC-LDA Confusion matrix of a) Leave-one-out cross validation (LOOCV) of unprocessed control and tumor bearing rat urine, b) LOOCV of concentrated control and tumor bearing rat urine (Ex-diagonal elements are false positive predictions).

Table 2: PC-LDA Confusion matrix of C, NTT, TT and T concentrated urine for a) Leave-one-spectrum-out cross validation b) Leave-one-rat-out cross validation (Diagonal elements are true positive predictions and Ex-diagonal elements are false positive predictions)

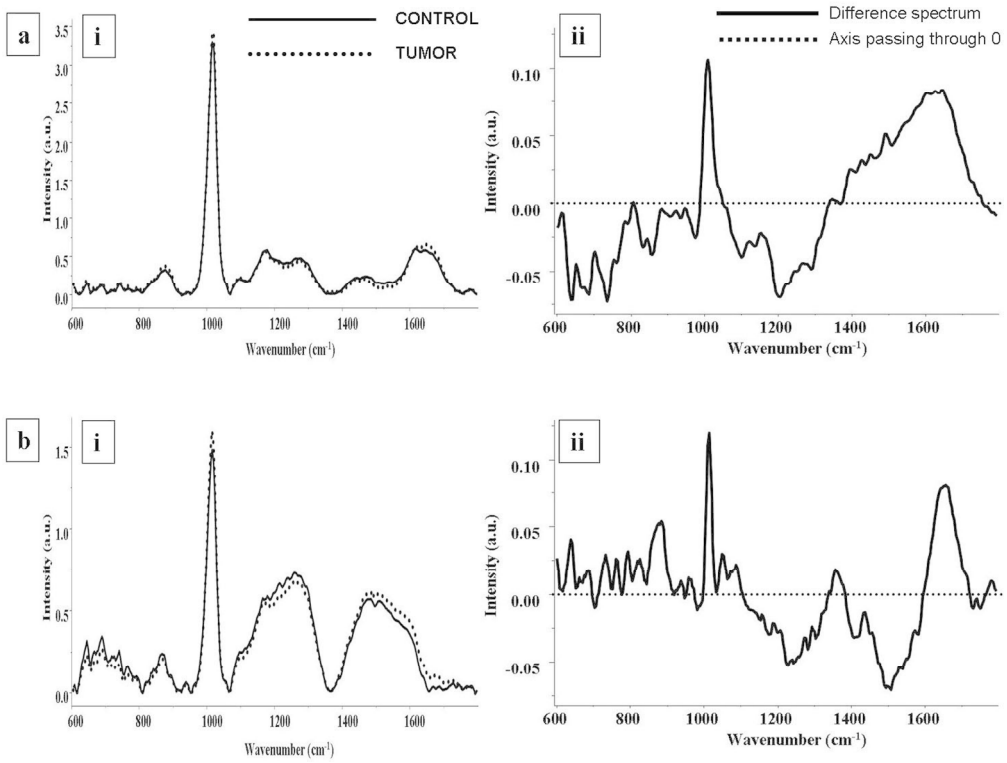
Table 1: PC-LDA Confusion matrix of a) Leave-one-out cross validation (LOOCV) of unprocessed control and tumor bearing rat urine, b) LOOCV of concentrated control and tumor bearing rat urine (Ex-diagonal elements are false positive predictions).

a)LOOCV (No. of animals, No. of spectra)	**Unprocessed urine control**	**Unprocessed urine tumor**
Unprocessed urine control (9, 81)	65 (**80.24 %**)	16
Unprocessed urine tumor (9, 82)	23	59 (**71.95 %**)
b)LOOCV (No. of animals, No. of spectra)	**Concentrated urine control**	**Concentrated urine tumor**
Concentrated urine control 8, 64)	50 (**78.12 %**)	14
Concentrated urine tumor (7, 56)	5	51 (**91.07%**)

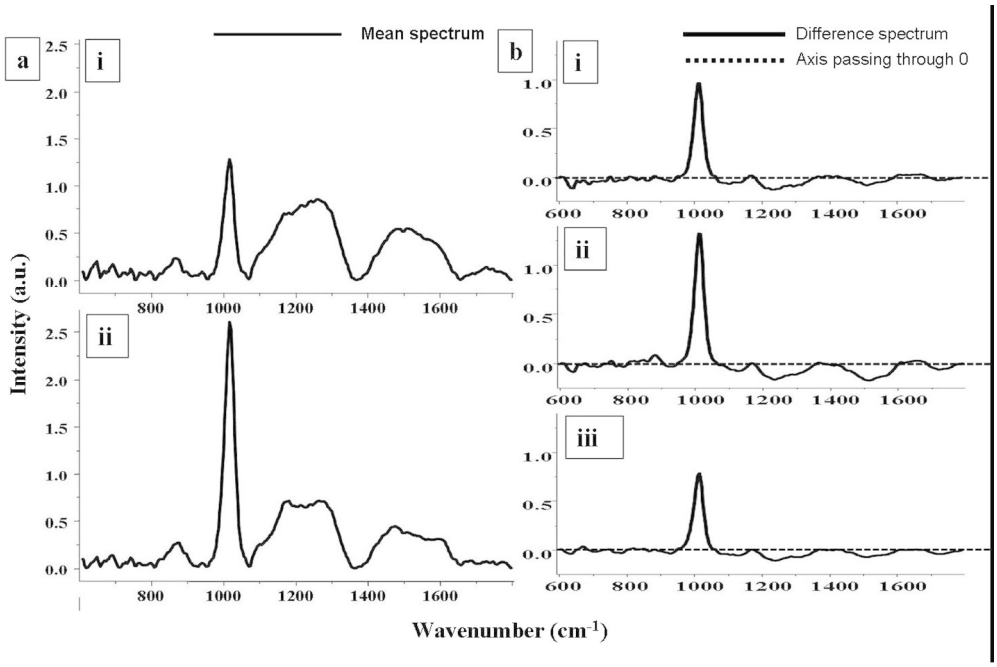
Table 2: PC-LDA Confusion matrix of C, NTT, TT and T concentrated urine for a) Leave-one-spectrum-out cross validation b) Leave-one-rat-out cross validation (Diagonal elements are true positive predictions and Ex-diagonal elements are false positive predictions)

a) **Group (No. of animals, no. of spectra)**	**CONCENTRATED URINE CONTROL**	**CONCENTRATED URINE NTT**	**CONCENTRATED URINE TT**	**CONCENTRATED URINE TUMOR**
CONCENTRATED URINE CONTROL (8, 64)	30 (46.8%)	17 (26.6%)	8 (12.5%)	9 (14.1%)
CONCENTRATED URINE NTT (5, 40)	20 (50%)	14 (35%)	0	6 (15%)
CONCENTRATED URINE TT (4, 32)	8 (25%)	1 (3.1%)	19 (59%)	4 (12.5%)
CONCENTRATED URINE TUMOR (7, 56)	2 (3.6%)	15 (26.8%)	17 (30.4%)	22 (39%)

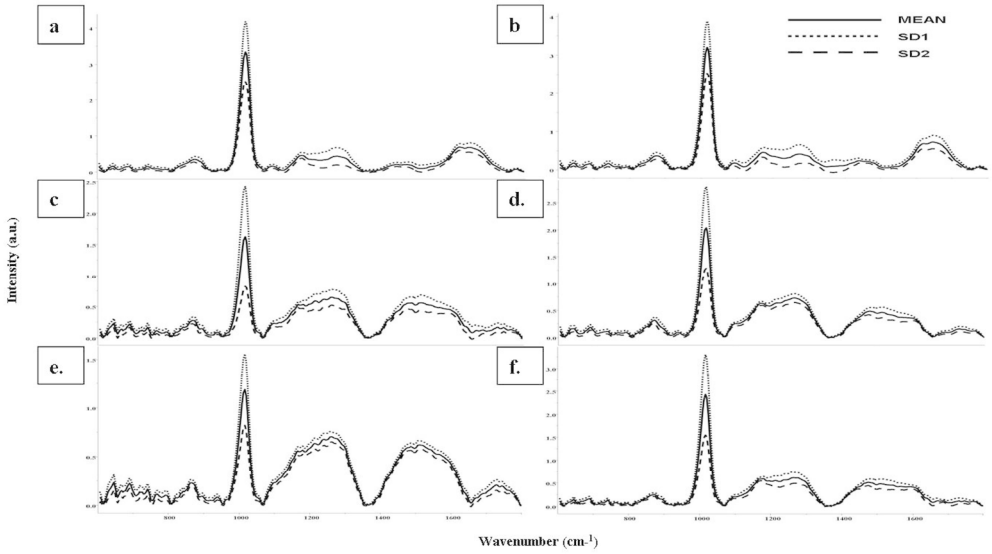
b) Group (No. of animals, no. of spectra)	CONCENTRATED URINE CONTROL	CONCENTRATED URINE NTT	CONCENTRATED URINE TT	CONCENTRATED URINE TUMOR
CONCENTRATED URINE CONTROL (8, 64)	34 (53.12 %)	9 (14.1%)	11 (17.2%)	10 (15.6%)
CONCENTRATED URINE NTT (5, 40)	16 (40%)	17 (42.5 %)	2 (5%)	5 (12.5%)
CONCENTRATED URINE TT (4, 32)	8 (25%)	0	22 (62.5 %)	2 (6.3%)
CONCENTRATED URINE TUMOR (7, 56)	3 (5.3%)	11 (19.6%)	12 (21.4%)	30 (51.78 %)



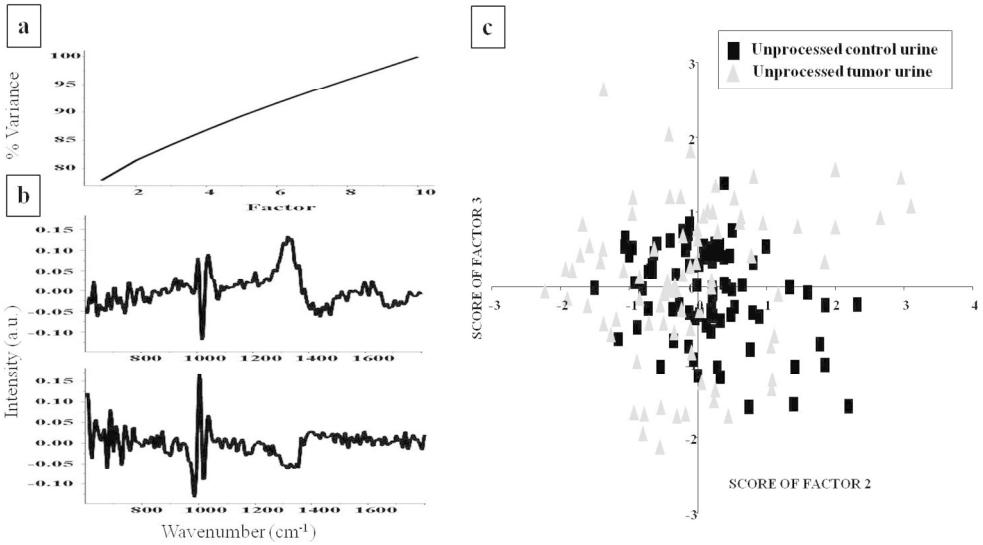
a) i) Mean spectra of unprocessed urine interpolated in 600-1800cm-1 region from control and tumor bearing rats, ii) tumor – control difference spectrum; b) mean spectra of concentrated urine interpolated in 600-1800cm-1 region from control and tumor bearing rats, ii) tumor – control difference spectrum
249x190mm (300 x 300 DPI)



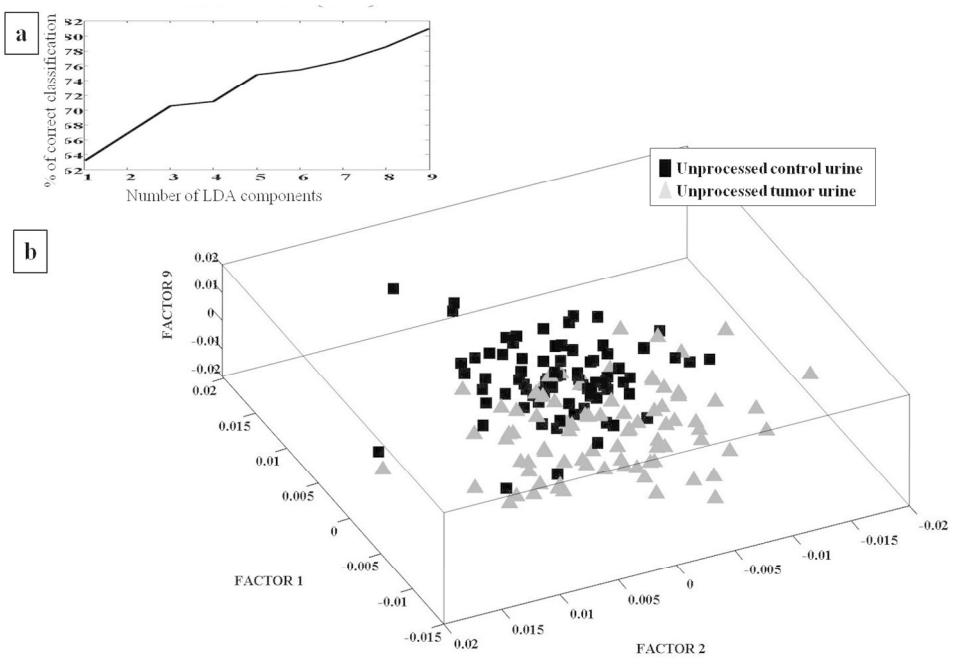
a) Mean spectra of concentrated urine interpolated in 600-1800cm-1 region from i) NTT and ii) TT rats, b) i) TT – control, ii) TT - NTT, and iii) control - NTT difference concentrated urine spectra
248x163mm (300 x 300 DPI)



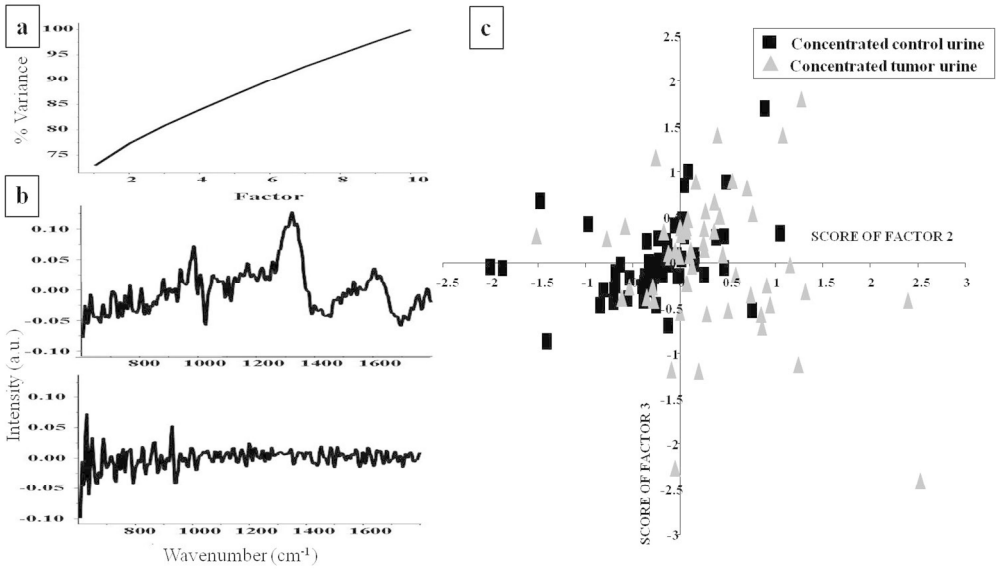
Mean and standard deviations for each group, a) unprocessed urine control b) unprocessed urine tumor c) concentrated urine control d) concentrated urine tumor e) concentrated urine NTT and f) concentrated urine TT
230x133mm (300 x 300 DPI)



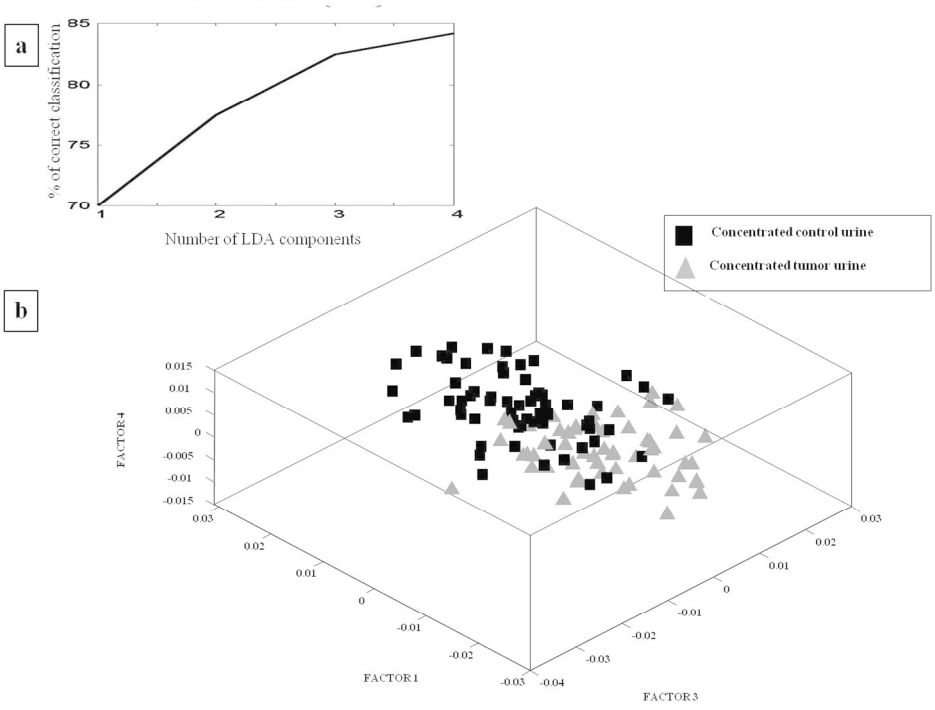
PCA of unprocessed urine from control and tumor bearing rats a) variance plot, b) Loading factors 2and 3, and c) Scatter plot.
254x144mm (300 x 300 DPI)



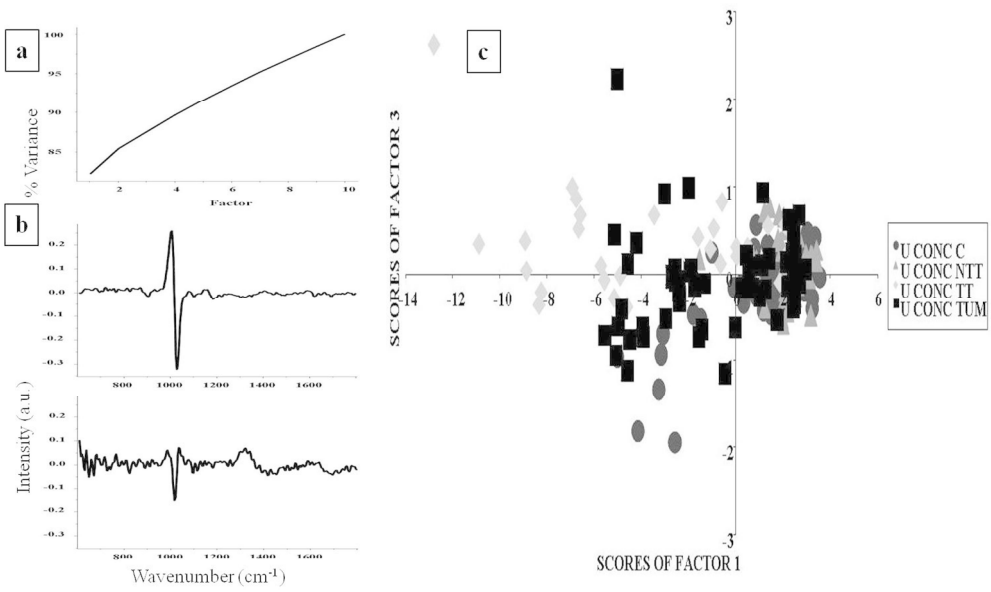
PC-LDA of unprocessed urine from control and tumor bearing rats a) Scree plot b) scatter plot
254x171mm (300 x 300 DPI)



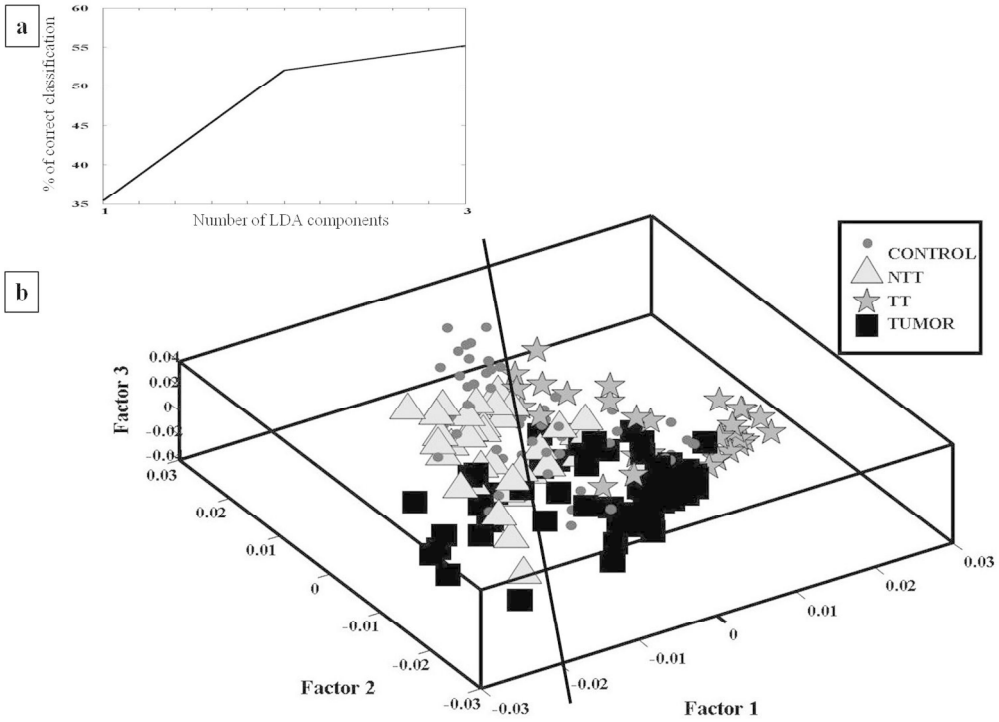
PCA of concentrated urine from control and tumor bearing rats a) variance plot, b) Loading factors 2and 3, and c) Scatter plot.
254x154mm (300 x 300 DPI)



PC-LDA of concentrated urine from control and tumor bearing rats a) Scree plot b) scatter plot
254x182mm (300 x 300 DPI)



: PCA of concentrated urine from control, NTT, TT and tumor bearing rats a) variance plot, b) Loading factors 1and 3, and c) Scatter plot.
255x150mm (300 x 300 DPI)



PC-LDA of concentrated urine from control, NTT, TT and tumor bearing rats a) Scree plot b) scatter plot
254x190mm (300 x 300 DPI)