### **Supplemental information**

# Low- and High-Grade Bladder Cancer Determination via Human Serum-Based Metabolomics Approach

Navneeta Bansal, Ashish Gupta, Nilay Mitash, Prashant Singh Shakya, Anil Mandhani, Abbas Ali Mahdi, SatyaNarainSankhwar, Sudhir Kumar Mandal

## <u>Table S1</u>

Summary of clinicopathological information of BC patients and healthy controls.

Characteristics	BC patients	Healthy controls	Significance level
Race	Indian	Indian	
No. of subjects	67	32	
Age (mean, range)	55, 40-70	52, 40-68	p > 0.05
Gender			
Male	67	32	
Female	0	0	
Cancer grade			
Log grade (LG)	36 (54%)	0	
High grade (HG)	31 (46%)	0	
BMI (median, range)	24.6, (17.3-27.6)	23.8 (16.3-28.2)	p > 0.05
Hematuria	6 (9%)	0	
Medications	0	0	
Smoking habit			
Nonsmokers	33 (49%)	20 (62.5%)	p = 0.37 by
Ex-smokers	10 (15%)	3 (9.4%)	Chi-square test
Smokers	20 (30%)	9 (28.1%)	
Information not available	42 (6%)	0	

#### Figure S1:

S-plot of multivariate analysis of different cohorts, (A) HC vs. LG+HG, (B) HC vs. LG, (C) HC vs. HG, (D) LG vs. HG. The corresponding OPLS-DA score plot explained in figure 3A-D. Dotted circle in S-plot symbolised the selected variable which have significant difference between their corresponding controls. Here HC: healthy control; LG: low grade; and HG: high grade of BC.



#### **Quantification of metabolites**

After careful baseline correction, the peak of identified metabolite and standard TSP were integrated and the following formula was used to estimate the concentrations of metabolites using a homemade computer program:

$$[C]_{x} = [C]_{TSP} \underbrace{N_{TSP} \cdot I_{x}}_{N_{x} \cdot I_{TSP}}$$

[*C*]<sub>X</sub> and [*C*]<sub>TSP</sub> are the concentration of metabolites and pre-calibrated known concentration of TSP, respectively.  $I_x$  and  $I_{TSP}$  are the NMR signal integration of metabolites and TSP, respectively.  $N_x$  is the number of protons per molecule giving rise to the integrated signal and  $N_{TSP}$ = 9.

#### Figure S2:

Comparisons of AUC of ROC showing the diagnostic accuracy based on discriminant predicted probability of various metabolites and validation of these biomarkers significance by doubleblind study (unknown samples). AUC of ROC curves of (A) HC vs. LG+ HG, (B) HC vs. LG, (C) HC vs. HG, (D) LG vs. HG using predicted probability of various combination of variables explained in table 3 and AUC of ROC curves of double blinded study of (E) HC vs. LG+ HG, (F) HC vs. LG, (G) HC vs. HG, (H) LG vs. HG using predicted probability of same variables as explained in know samples study, detailed explained in table 3 (here, HC: healthy control; LG: low grade; and HG: high grade of BC).

