Peer Review File

Article information: http://dx.doi.org/10.21037/tau-20-1057.

Reviewer A

Comment #1: Page 3; Line 26:

Wang et al. is cited for the quantitative determination of urine EVs. This author is not listed in the bibliography.

Answer: Refence #12 entitled "An integrated double-filtration microfluidic device for isolation, enrichment and quantification of urinary extracellular vesicles for detection of bladder cancer" had been added into the manuscript.

Comment #2: Page 3; Line 34:

Yes, in urine the EVs do not necessarily come from the bladder, often the urine has high leukocyte counts. Was it checked whether the detected mRNAs could also be related to leukocytes?

Answer: First, the number of urothelial cells exposed to urine outnumbers that of leukocytes. Besides, we had examined the expression levels of candidate genes in leukocyte (GTEx) and bladder cancer as well as para-carcinoma (TCGA). As given in the following table, the expression of candidate genes (median number) are far less then that in bladder cancer (house keeping genes SLC25A6, GAPDH and ACTB were listed as reference, see supporting document "SI-15_RNA.GTEx_and_TCGA"). Although the possibility of leukocytes derived mRNA cannot be entirely excluded, their impact on the diagnosis model should be very limited.

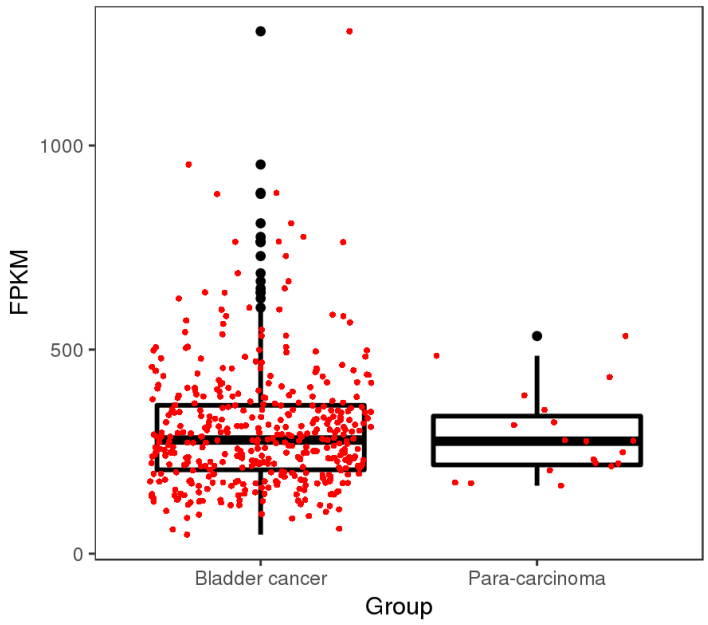
#ID	GTEx_Blood_rpkm	TCGA_Bladder-Cancer_fpkm	TCGA_Para-carcinoma_fpkm
SLC25A6	123.10	278.44	275.54
GAPDH	1087.00	1071.86	607.96
ACTB	2361.00	1636.14	2216.79
BIRC5	0.06	17.91	0.99
CCNB1	0.98	26.19	2.66
CDC20	1.38	37.85	2.82
CDK1	0.54	14.59	1.15
FOXM1	0.50	11.73	0.75
KRT7	0.14	489.46	66.88
MMP11	0.09	10.00	0.66
MYBL2	2.97	30.81	2.78
S100A2	0.11	171.61	124.96
TK1	2.98	50.09	3.82
TPX2	1.31	29.19	2.28
UBE2C	2.18	64.61	2.75

Comment #3: Page 5; Line 19:

The catalog number of the TSG101 antibody is incorrect. It is correctly "sc-13,6111; Santa Cruz; CA; USA".

Answer: The author would like to thank the reviewer for pointing out the typo. Correction has been made in the manuscript accordingly.

Comment #4: Page 5/7; lines 34/11:
How was the reference gene selected? In Figure 3, the expression of the reference gene is not shown. In the work of Murakami 2018 ALDOB is used as reference gene, because the other candidates showed a differential expression in the diagnostic groups. The proof, that SLC25A6 is not differentially expressed between the diagnostic groups must be provided.
Answer: The reference gene was selected on the basis of TCGA data analysis. On one hand, the expression level of SLC25A6 was stable in bladder cancer and para-carcinoma. As given in the box-plot as below, the FPKM of SLC25A6 showed no significant difference between bladder cancer and para-carcinoma. On the other hand, the expression level of SLC25A6 is closer to the candidate mRNAs comparing to GAPDH and β-actin (see table in the answer of comment #2). Therefore, SLC25A6 was chosen as the reference gene in this study.



Comment #5: Page 7; Line 7:
The TCGA database is based on 384 high and 21 low grade bladder tumors. The differentially expressed genes will therefore preferably characterize high grade tumors. What criteria were used to select KRT7 and S100A2 for low grade tumors? Murakami's work describes e.g. significant expression differences in pT1 tumors for the SLC2A1 gene.
Answer: The author agree with the reviewer. Due to the advantage of sample number, the differentially expressed gene will definitely favor the high grade tumors. In order to avoid missed detection of low grade bladder cancer, we performed an additional analysis. TCGA and GTEx data were downloaded from UCSC Xena browser platform.
After combining the data from GTEx (Genotype Tissue Expression, UCSC) and TCGA (21 low grade tumor and 9 health controls), Candidate genes with log2 fc>4 and TMP > 1000 in low grade tumor were selected as following table. KRT7 and S100A2, top 2 in the list (see supporting document “SI-DEGs+DEMs+LowGrade”), were selected to ensure the detection of low grade tumor.
Revision had been made in pg7, line 9-15.

#ID	NormalSamples	LowGradeSamples	log2FC	PValue
KRT7	70.63834203	2632.466491	5.219819875	1.68E-06
S100A2	81.52000883	1417.171834	4.119716679	1.36E-05
TSPAN6	27.27008131	66.0587681	1.276431064	0.056277595
TNMD	0.089999347	0.009999639	-3.169966598	0.009749871
DPM1	34.7407856	29.51039064	-0.23540736	0.304001286
SCYL3	10.54994269	4.110030285	-1.360014233	0.002994447
C1orf112	3.039903427	2.470008546	-0.299509459	0.160540226
FGR	11.46030964	3.600005992	-1.670574812	0.012004803
CFH	131.7801307	38.98002172	-1.757326063	0.006866357
FUCA2	25.94067987	31.51035747	0.28060983	0.929247404
GCLC	87.55292802	24.72059634	-1.824441881	0.004984501
NFYA	10.62996017	7.869743912	-0.433747597	0.533265395
STPG1	5.039965218	7.069930112	0.488282176	0.016517196
NIPAL3	6.949901014	8.230184186	0.243932288	0.177896926
LAS1L	28.88903407	21.73026119	-0.410816451	0.01248928
ENPP4	6.920093525	2.170060606	-1.673056201	5.02E-05
SEMA3F	30.35058599	75.6296965	1.317228458	0.008014734
CFTR	2.250052961	0	#NAME?	4.88E-05
ANKIB1	15.35999752	7.379916658	-1.057501554	0.000629755

Comment #6: Page 8; Line 20:
The training cohort includes 126 bladder tumor samples. Here 128 bladder tumor samples are erroneously listed.
Answer: Thank the reviewer for pointing out this typo. Correction has been made in the manuscript accordingly.

Comment #7: Page 8; Line 21:
S100A2 was examined as a marker for low grade tumors. What is the diagnostic performance in the training cohort for S100A2 for the differentiation between high and low grade tumors? The ExoPanel did not find 13/57 ~ 22% low grade tumors! Especially in the group of low grade tumors a reliable diagnosis is important, because here the course for an early successful treatment is set. Also in the validation cohort 3/13 ~ 23% low grade tumors were not identified.
Answer: The diagnostic performance in high and low grade tumors were given in the table below (also see supporting document "SI-S100A2"). All genes showed decreased performance in the diagnosis of low-grade tumors, while S100A2 had the best consistency among all candidate markers. Besides, we agree that the ExoPanel does miss 22-23% low grade tumor, which is a significant number of course. However, as we have mentioned in conclusion section, this assay can identify a significant amount of negative patients before carrying out cystoscopy test.Therefore, as supplementary detection, this ExoPanel can possibly reduce the unnecessary operation of cystoscopy and ensure that those who really need cystoscopy can be examined in a timely manner.

#ID	High_Train_AUC	Low_Train_AUC	High_Test_AUC	Low_Test_AUC
KRT7	0.614444844	0.530200457	0.556730092	0.639118457
MYBL2	0.827284705	0.545806225	0.868804322	0.432506887
S100A2	0.747754222	0.713952875	0.773666902	0.561524334
TK1	0.707599713	0.567698259	0.73255814	0.591827365
UBE2C	0.844202899	0.58638122	0.886774724	0.619834711

Comment #8: page 8; line 25:
The validation cohort includes 12 disease controls. These are characterized in more detail here. However, when added up, this enumeration yields 13 samples. What is correct?
Answer: It should be "4 patients with benign prostatic hyperplasia, 5 with cystitis, 2 with non-bladder urothelial carcinoma, 1 with urinary calculus" that is 12 disease control in total. Revision had been made.

Comment #9: Page 15; Figure 3:
The legend refers to 427 samples in which the 12 marker candidates were analyzed. How does this number arise? Please show the constant expression of SLC25A6 in all diagnostic groups.
Answer: Figure 3 is the heatmap of 12 marker candidates' expression in 427 samples in TCGA dataset. First, a total of 7053 differentially expressed gene (DEGs) were identified (see manuscript Page 6 line 1-7). We then applied two additional criteria (see manuscript Page 7 line 3-5) and lock on 10 candidate markers (i.e. TK1, CDK1, MYBL2, TPX2, FOXM1, UBE2C, BIRC5, CDC20, MMP11, and CCNB1, see table below). Because the number of high grade tumor is dominant position in TCGA dataset, we performed additional analysis to pick out 2 candidate marker (KRT7 and S100A2, see the answer to comment #7) . The expression of SLC25A6 in different group has been summarized as box-plot in the answer to comment #4.

#ID	Para-carcinoma	BladderCancer	log2FC	PValue
TPX2	2.276443718	29.18614609	3.680430025	2.84E-10
BIRC5	0.986471502	17.91115076	4.182436847	9.23E-09
MMP11	0.656143636	10.00078531	3.929957811	1.99E-11
MYBL2	2.779835027	30.81179993	3.470411791	1.06E-09
FOXM1	0.74652022	11.73063288	3.973955703	3.20E-09
CDC20	2.815183152	37.85362752	3.749130871	2.87E-10
CCNB1	2.659305722	26.18778331	3.299772398	1.70E-09
TK1	3.824401199	50.09051432	3.711231643	1.61E-09
CDK1	1.148811719	14.59470096	3.667230375	5.80E-11
UBE2C	2.745786026	64.60770097	4.55641503	5.11E-11
TSPAN6	20.61021295	22.70641052	0.139740248	0.08644411
TNMD	0.081267522	0.012122089	-2.745040611	3.91E-07
DPM1	25.49447647	36.38081398	0.51299311	1.07E-05
SCYL3	1.773239749	2.328276488	0.392874784	0.004453394
C1orf112	0.572302137	1.977005335	1.788467864	5.87E-09
FGR	3.134939439	2.259945258	-0.472149746	0.141797424
CFH	18.92927381	6.89012866	-1.458016237	2.75E-06
FUCA2	17.47833556	27.22119543	0.63916262	5.98E-06
GCLC	11.97871583	7.011298327	-0.772719726	0.049137279
NFYA	6.248114616	9.536692623	0.6100681	0.007102934
STPG1	1.162768022	1.923162725	0.725917538	0.004868692
NIPAL3	2.760814744	3.46568155	0.328045012	0.243312848
LAS1L	7.030619334	10.11731718	0.525103092	2.50E-06
ENPP4	4.571914213	2.543894388	-0.845759556	0.015357901
SEMA3F	24.3665247	31.72578439	0.380755334	0.012966626
CFTR	0.03552006	0.019806954	-0.842626967	0.024602024
ANKIB1	6.1556878	8.121525935	0.399830754	0.0005135
CYP51A1	0.675995774	0.815941616	0.271451698	0.359818807
KRIT1	3.877406876	4.565168416	0.235575952	0.036526744
RAD52	1.493132912	1.901795812	0.349019765	0.01139561
MYH16	0.009332801	0.049210851	2.398594501	7.36E-08
BAD	17.26779029	25.87275939	0.583350452	5.51E-05
LAP3	18.19671058	19.69659704	0.11426872	0.566381501
CD99	63.24087437	57.91673714	-0.126876987	0.665264318
HS3ST1	8.077741757	2.913816448	-1.471042025	0.058075093
AOC1	0.268700732	0.283056971	0.075092202	0.369882637
WNT16	0.019409482	0.062027949	1.676156821	0.038888633
HECW1	0.053738216	0.070129057	0.384063914	0.320376897
MAD1L1	3.140704158	5.410293646	0.784618846	3.51E-07
LASP1	24.8377279	42.57720435	0.777548019	6.68E-07
SNX11	5.689743662	7.719867972	0.440212517	8.23E-05
TMEM176A	8.645016883	5.446434598	-0.666556675	0.115550189
M6PR	13.31909042	18.84111381	0.500388692	0.041955516
KLHL13	2.035547553	0.326923939	-2.63839	1.25E-08
CYP26B1	0.814054564	0.433546407	-0.90893907	0.096671102
ICA1	4.845128766	7.963127093	0.716800076	0.123219878
DBNDD1	0.827942535	4.125135915	2.316839111	5.51E-06
ALS2	1.938239622	2.066246476	0.092265419	0.118213104
CASP10	2.349162528	2.530128461	0.107064106	0.331616892

Comment #10: Page 15; Figure 3

In the legend the cancer group is marked blue and the para-carcinoma group is marked red. The colored symbols in the figure are labeled the other way round.
Answer: Thank the reviewer to point it out. Such a mistake shouldn't have happened. We had made revision in the manuscript.

Comment #11: Page 20/22, Table 2/4:
How long was the time interval between suspicion of recurrence and the last tumor diagnosis in the group of non-recurrent patients?
Answer: In general, cystoscopy or imaging examination showed no recurrence of bladder tumor 2 years after surgery, were considered to be non-recurrent. So patients of non-recurrent group, the time interval between suspicion of recurrence and the last tumor diagnosis have to be more than 2 years.

Comment #12: Page 21; Table 3
In the column "Cystoscopy - negative" the ExoPanel is positive in 111 cases and negative in 131 cases. How many healthy controls, disease controls and non-recurrent samples are included in the respective numbers?
Answer: The number of healthy controls, disease controls and non-recurrent samples for cystoscopy negative patients were listed as below.

		cystoscopy	
		Positive	**Negative**
ExoPanel	**Positive**	112	111 (a)
	Negative	14	131 (b)

(a) 111 ExoPanel+/cystoscopy- individuals including 27 healthy controls, 29 disease controls, and 55 non-recurrent
(b) 131 ExoPanel-/cystoscopy- individuals including 75 healthy controls, 23 disease controls, and 33 non-recurrent

Reviewer B
Comment #1: You should include a supplementary table with all the mRNA and miRNA tested in the TCGA cohort, both for the main bladder cancer versus para-carcinoma comparision and the low-grade comparison used to select the two genes KRT7 and S100A2. The table should include the median expression in the two groups, FC and p-value.
Answer: We would like to thank the reviewer for his or her professional suggestion. We had added three additional supplementary tables.
1) DEGs Table, with all mRNA in the TCGA cohort
2) DEMs Table, with all miRNA in the TCGA cohort

3) Lowgrade analysis table

Comment #2: How did you selected the 5 mRNAs for the final model? What method/criteria did you use? How many and what models did you test? It would be interesting to see also models with less mRNAs, for instance 2 or 3 mRNAs. Can you report the AUC for each of the 12 mRNAs (you only report the selected 5 in the figures).

Answer: We had compared 2, 3 and 5 gene panels. The result were summarized as supporting document "SI-Performance of multiple gene panel". Besides, we also compare the NPV and AUC of different gene numbers for both training and validation cohort. As showed in Fig a and b, 5 gene panels has generally higher performance (NPV and AUC) and better consistency in both training and validation cohort. Therefore we decided to use 5 gene panel.

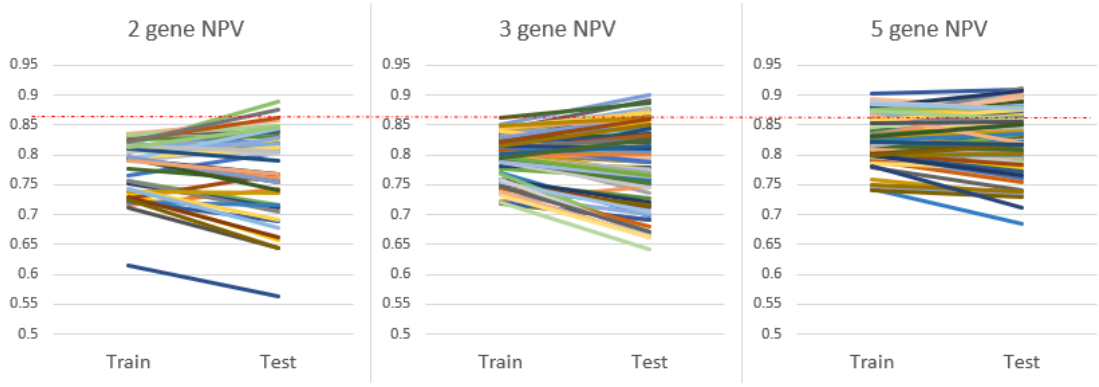


Fig. a Comparison of 2, 3 and 5 gene according to the their NPV for both training and validation cohort.

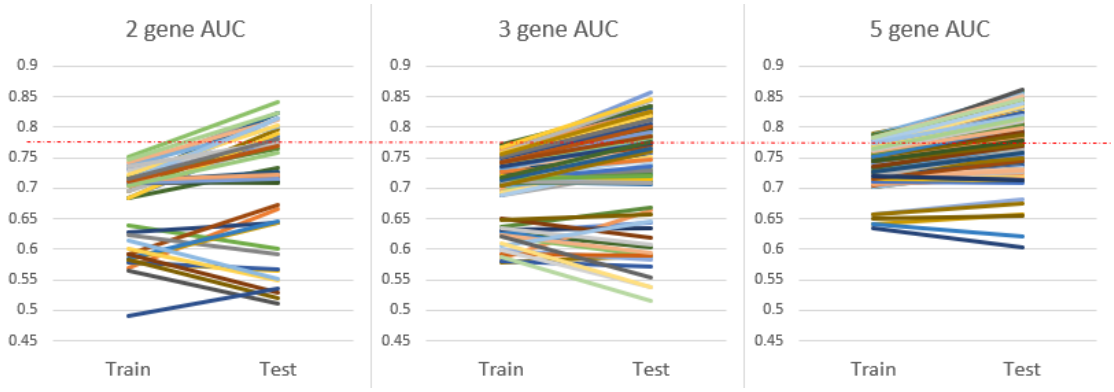
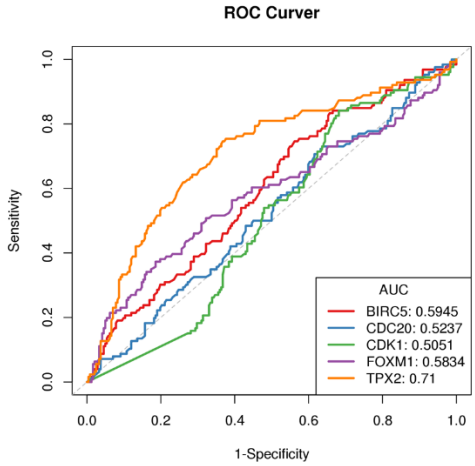


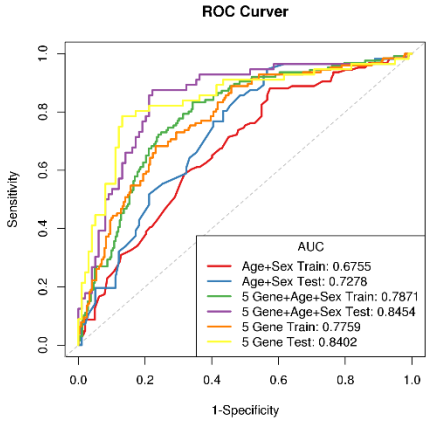
Fig. b Comparison of 2, 3 and 5 gene according to the their AUC for both training and validation cohort.

MMP11 and CCNB1 failed to be detected in a significant number of samples therefore they were excluded from further analysis. The performance of the rest 5 gene in training cohort were given as follow.



Comment #3: From table 2 and table 3, I can see that the bladder cancer affected individuals are older than the controls and with a higher proportion of males. In order to understand if these differences are relevant for the model, you should show two additional predictive models: one with only age and sex as a baseline and another one with age, sex and the 5 mRNAs, to see how much the mRNAs can add to the baseline.

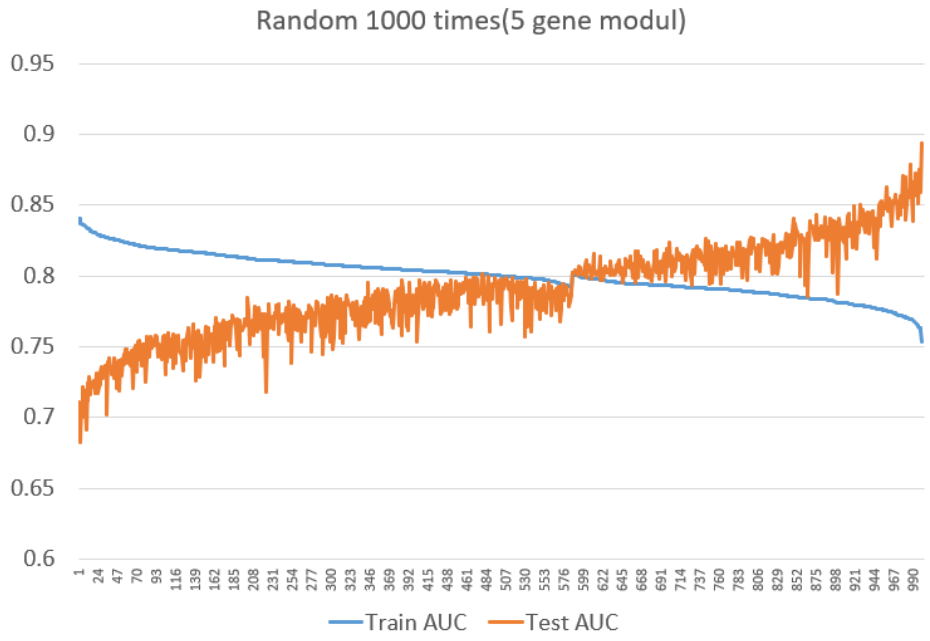
Answer: We had performed the analysis as suggested by the reviewer. Simply age and sex had a AUC 0.6755 for training cohort and 0.7278 for testing cohort, which were lower than 5 gene model. While combing age and sex with 5 gene model show a slight increase in AUC (0.7759 to 0.7871 in Training cohort, 0.8402 to 0.8454 in Testing cohort). Therefore, age and sex were not the main factors of the difference between two groups.



Comment #4:The much higher performance on the validation set with respect to the training set is strange: normally the model has some amount of overfitting and thus its performance is lower on the validation set. Can you explain why this does not happen in your case? Or can you try to refit the 5 mRNA model with a different random train/validation split?

Answer: To answer this question, we randomly divide the sample 1000 times to see if there is over fitting. Briefly, the samples were randomly divided according to 7:3 ratio. The 5 gene panel in the manuscript were used for training and testing. The AUC for each random division were summarized as following figure. It can be seen that in around 43% of all cases the AUC of training cohort is lower than the testing cohort. So the sample partition in our article may

just fall on the side of better performance of test set.



Minor points:

There are various small grammatical error (for instance missing articles).

pg4, line 17: "152 healthy control" should be "controls"
Answer: Correction has been made accordingly in the manuscript.

pg4, line 30: "as reported.(16)" should probably be "as reported in (16)"
Answer: Correction has been made accordingly in the manuscript.

pg6, lines 15-17: Please state here that the differential expression analysis is between the 408 cancer and the 19 para-carcinoma samples. It is written later, but it is not clear here.
Answer: Correction has been made accordingly in the manuscript.
"Through differential gene expression analysis of these 408 cancer and 19 para-carcinoma samples, a panel of 12…" (pg6, 17-18)

pg6, line 17: 12 candidates genes are mentioned here, but in the methods (at pg5, lines 26-27) only 10 genes are listed, you are missing two genes. Also the housekeeping gene should be mentioned?
Answer: Correction has been made accordingly in the manuscript.
The predictors used in the development of this multivariable model in this study were the expression level of RNA candidates (i,e, TK1, CDK1, MYBL2, TPX2, FOXM1, UBE2C, BIRC5, CDC20, MMP11, CCNB1, KRT7, S100A2) and house keeping gene SLC25A6.

pg7, line 4: here you report a FC threshold of 8, but previously (at pg6, line 6) you wrote "significance P value < 0.01, and FC >2", please fix the discordance.

Answer: Sorry for the confusing description.
P value < 0.01, and FC >2 at pg6, line 6 is the first step to identify differential expressed genes (DEGs) as highlight in yellow.
"In this step, two criteria were applied to identify candidate markers from existed DEGs:"
Then, to further identify valuable candidate genes from a total of 7053 DEGs, we set an additional criteria which is listed in pg7, line 4.

To eliminate such confusion, we had revised the manuscript as follow:
"In this step, two criteria were applied to identify candidate markers from existed 7053 DEGs identified in previous step: i) FC of gene expression between bladder cancer and para-carcinoma > 8, ii) FC of gene expression between bladder cancer and prostate/renal cancer > 5. " pg7, line 5-7

pg7, line 5: you speak of prostate/renal cancer samples, where do these samples come from and how many are they? I found not other information on the them in the paper.
Answer: Revision had been made accordingly in pg7 line 12-13, as below:
"The RNA-seq data of prostate (496 prostate cancer and para-carcinoma) and renal (885 renal cancer and 128 para-carcinoma) tissues were downloaded from TCGA."

pg7, line 8: which criteria (FC/p-value thresholds) did you use to select these two genes?
Answer:
Due to the advantage of sample number, the differentially expressed gene will definitely favor the high grade tumors. In order to avoid missed detection of low grade bladder cancer, we performed an additional analysis. TCGA and GTEx data were downloaded from UCSC Xena browser platform.
After combining the data from GTEx (Genotype Tissue Expression, UCSC) and TCGA (21 low grade tumor and 9 health controls), Candidate genes with log2 fc>4 and TMP > 1000 in low grade tumor were selected as following table. KRT7 and S100A2, top 2 in the list, were selected to ensure the detection of low grade tumor.

#ID	NormalSamples	LowGradeSamples	log2FC	PValue
KRT7	70.63834203	2632.466491	5.219819875	1.68E-06
S100A2	81.52000883	1417.171834	4.119716679	1.36E-05
TSPAN6	27.27008131	66.0587681	1.276431064	0.056277595
TNMD	0.089999347	0.009999639	-3.169966598	0.009749871
DPM1	34.7407856	29.51039064	-0.23540736	0.304001286
SCYL3	10.54994269	4.110030285	-1.360014233	0.002994447
C1orf112	3.039903427	2.470008546	-0.299509459	0.160540226
FGR	11.46030964	3.600005992	-1.670574812	0.012004803
CFH	131.7801307	38.98002172	-1.757326063	0.006866357
FUCA2	25.94067987	31.51035747	0.28060983	0.929247404
GCLC	87.55292802	24.72059634	-1.824441881	0.004984501
NFYA	10.62996017	7.869743912	-0.433747597	0.533265395
STPG1	5.039965218	7.069930112	0.488282176	0.016517196
NIPAL3	6.949901014	8.230184186	0.243932288	0.177896926
LAS1L	28.88903407	21.73026119	-0.410816451	0.01248928
ENPP4	6.920093525	2.170060606	-1.673056201	5.02E-05
SEMA3F	30.35058599	75.6296965	1.317228458	0.008014734
CFTR	2.250052961	0	#NAME?	4.88E-05
ANKIB1	15.35999752	7.379916658	-1.057501554	0.000629755

Revision had been made in pg7, line 9-15.

pg7, line 9: "low grade bladder cancer comparing to normal and para-carcinoma tissues", where do the normal samples come from? You only mentioned the cancer and para-carcinoma tissue samples previously.

Answer: To avoid missed detection of low grade bladder cancer, TCGA and GTEx data were downloaded from UCSC Xena browser platform, which included 9 normal samples.

pg7, line 34: How did you choose the 5 genes out of ten? Did you try multiple sets of genes? I think "relatively better performance and more evident differential expression" is too vague. This is an important step, it needs to be properly explained (did I miss it in the manuscript?). Both the method/criteria used for selecting the final model and which models were examined should be reported. See the major points.

Answer: On the basis of the qPCR result of ten mRNAs, we had examined a series of 5 genes panel (n=253), which were listed in descending order according to their negative prediction value (NPV). The final panel was decided basing on the NPV as well as sensitivity (SN).

Characteristics	Train_NPV	Train_SN	Train_SP	Train_AUC
MYBL2+TK1+UBE2C+KRT7+S100A2	0.9034483	0.888889	0.541322	0.77594123
FOXM1+UBE2C+S100A2+BIRC5+TPX2	0.8928571	0.857143	0.619835	0.768332677
FOXM1+UBE2C+KRT7+S100A2+TPX2	0.8922156	0.857143	0.615702	0.768627837
UBE2C+S100A2+BIRC5+CDC20+TPX2	0.8888889	0.857143	0.595041	0.763282172
FOXM1+UBE2C+S100A2+CDC20+TPX2	0.8850575	0.84127	0.636364	0.770431589
UBE2C+KRT7+S100A2+CDC20+TPX2	0.8848485	0.849206	0.603306	0.764331628
UBE2C+S100A2+CDC20+CDK1+TPX2	0.8846154	0.833333	0.665289	0.774367047
FOXM1+UBE2C+KRT7+S100A2+CDC20	0.8837209	0.84127	0.628099	0.760002624
MYBL2+TK1+UBE2C+CDC20+CDK1	0.8793103	0.833333	0.632231	0.775318116
FOXM1+UBE2C+S100A2+CDK1+TPX2	0.8781726	0.809524	0.714876	0.779024006
FOXM1+UBE2C+KRT7+S100A2+BIRC5	0.8777778	0.825397	0.652893	0.763314968
UBE2C+KRT7+S100A2+BIRC5+TPX2	0.875	0.825397	0.636364	0.758494031
MYBL2+UBE2C+BIRC5+CDK1+TPX2	0.8735632	0.825397	0.628099	0.76213433
UBE2C+KRT7+S100A2+BIRC5+CDC20	0.8735632	0.825397	0.628099	0.759379509
FOXM1+UBE2C+S100A2+BIRC5+CDC20	0.8729282	0.81746	0.652893	0.766299357
MYBL2+UBE2C+S100A2+CDC20+TPX2	0.8728324	0.825397	0.623967	0.763774105
UBE2C+BIRC5+CDC20+CDK1+TPX2	0.8728324	0.825397	0.623967	0.76613538
FOXM1+TK1+UBE2C+S100A2+TPX2	0.872449	0.801587	0.706612	0.785419126
FOXM1+UBE2C+S100A2+BIRC5+CDK1	0.8715084	0.81746	0.644628	0.769152565

pg8, lines 11-14: please round the numbers, there are way too many useless digits

Answer: Revision had been made.

(a= 0.2314, b= -0.0617, c= 0.4844, d= -0.3072, e= - 0.0268, f = – 0.2258).

pg8, lines 17: what do SN and NP stand for?

Answer: The full name of acronyms had been added. “the overall NPV, sensitivity (SN) and specificity (SP) of the final model” pg8, line 18-19.

pg8, lines 24-27: here you say that that the 155 patient validation cohort was enrolled, while previously (pg4, lines 22-24) you said they were randomly chosen with a 7:3 ratio, which one is it?

Answer: This validation cohort was randomly divided as describe in pg4. Therefore, this part had been revised as following to avoid confusion.

“To validate the performance of established model, a randomly assigned validation cohort including 56 bladder cancer patients, 50 healthy controls, 12 disease controls, (i.e. 4 patients with benign prostatic hyperplasia, 5 with cystitis, 2 with non-bladder urothelial carcinoma, 2 with urinary calculus), and 37 non-recurrent bladder cancer patients were used for analysis.”

pg4, line 17: what are the diseases of the 64 disease controls? Please list them like you do at pg8, lines 26-27

Answer: Revision had been made. Pg 4, line 17-19 “64 disease controls (i.e. 13 patients with benign prostatic hyperplasia, 30 with cystitis, 10 with non-bladder urothelial carcinoma, 11 with urinary calculus)”

pg8, lines 26-27: you say that there are 12 disease controls but number in the parenthesis sum to 13, please fix the inconsistency

Answer: It should be “4 patients with benign prostatic hyperplasia, 5 with cystitis, 2 with non-

bladder urothelial carcinoma, 1 with urinary calculus" that is 12 disease control in total. Revision had been made.