Integrative analysis of 1q23.3 copy number gain in metastatic urothelial carcinoma - Supplemental Information - Independent validation

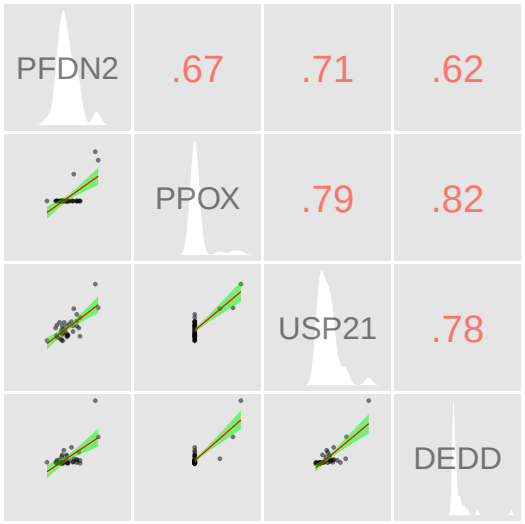
August 15, 2013

1 Expression analysis of 1q21.2 and 1q23.3 in an independent cohort

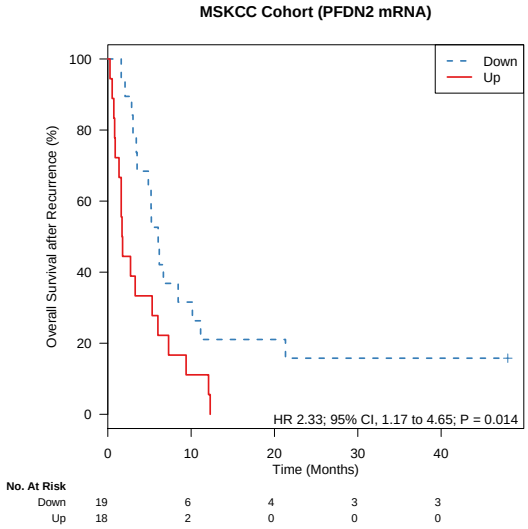
We next tested association of gene expression on 1q21.2 and 1q23.3 with survival after recurrence in an independent cohort of metastatic bladder cancer patients (MSKCC, [3]). Raw data was downloaded from GEO (GSE31684) and expression estimates were obtained with the Frozen Robust Multiarray Analysis (fRMA) method [2]. For genes with multiple probe sets, expression estimates were averaged after outlier removal with the Normalizer program from the Sleipnir library [1]. This normalization procedure further removed genes with expression indistinguishable from background noise. Including these genes would increase the number of test and thus artificially increase the false discovery rates.

In Table S8, we show the hazard ratios of genes available on the Affymetrix Hg U133 Plus 2.0 array in all three 1q23.3 GISTIC peaks. Consistent with our copy number results, over-expression of genes on peak 1 was associated with poor outcome after recurrence of metastatic disease. There was no significant association with OS after surgery (radical cystectomy). Table S9 shows that genes from the 1q21.2 genes were not associated with OS in the MSKCC cohort. Note that *MCL1* displayed expression patterns indistinguishable from background noise and was not associated with survival.

Figure S15 shows the high correlation of genes located in the 1q23.3 peak, indicating the presence of amplifications in this cohort. In our cohorts with matched expression and copy number data, these high correlations were only observed in patients with 1q23.3 amplification (Supplemental Figure S16).

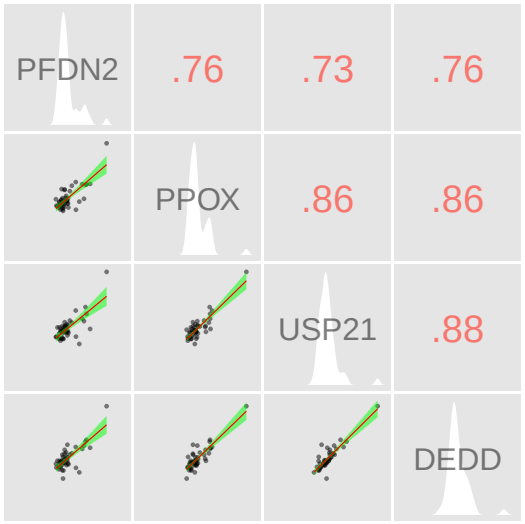


(a)

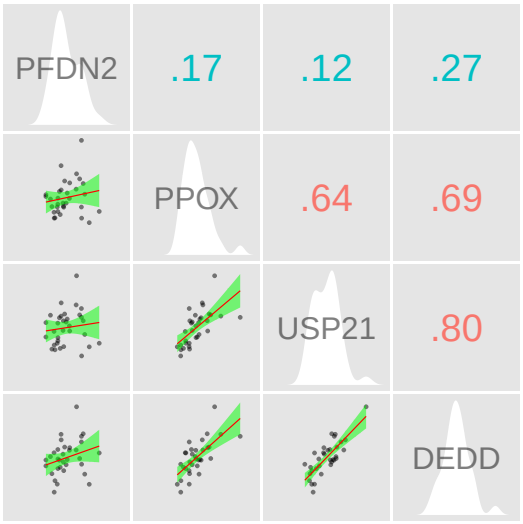


(b)

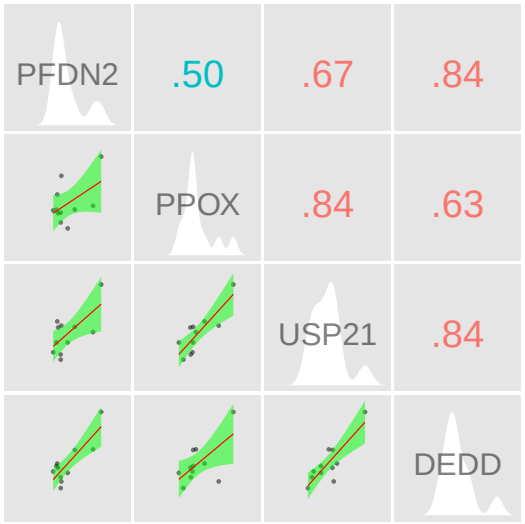
Figure S15: Expression of genes on 1q23.3 in an independent cohort [3]. (a) Pairwise correlation of the top hits (Table S8) of genes located in peak 1 and with hazard ratio > 1 (over-expression associated with shorter survival). Numbers in the upper-right, tringular half of the matrix are the Pearson correlation coefficients, which were all statistically significant ($P < 0.05$). Pairwise scatterplots of expression values are shown in the lower-left half and the expression histograms are shown on the matrix diagonal. (b) Association of *PDFN2* with survival after recurrence shown in a Kaplan-Meier plot in which patients where stratified by the cohort median of *PDFN2* expression.



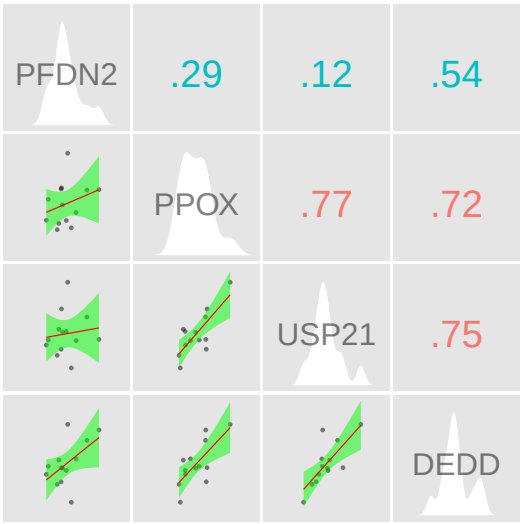
(a) Spanish 1q23 gain ($\log_2 > 0.15$, $n = 46$)



(b) Spanish no gain ($\log_2 \leq 0.15$, $n = 33$)



(c) DFCI 1q23.3 gain ($\log_2 > 0.25$, $n = 11$)



(d) DFCI no gain ($\log_2 \leq 0.25$, $n = 13$)

Figure S16: Expression of genes on 1q23.3 as in Supplemental Figure S15, but in the Spanish (a-b) and DFCI (c-d) cohorts. Panels (a) and (c) are for samples with 1q23.3 gain or amplification ($\log_2$ copy number ratio > 0.15), (b) and (d) for patients without amplification. Patients with amplification displayed significant expression correlation (colored in red) for all pairwise correlations. In patients without 1q23.3 amplification, some pairwise correlations were not significant ($P > 0.05$, colored in blue). *PFDN2* and *USP21* showed no correlation in absence of 1q23.3 amplification. Note that the threshold of > 0.15 is conservative and larger than the copy number ratios of 99% of normal segments.

Table S8: Overall survival (OS) Hazard Ratios (HRs) of genes in the three peaks, available on the Affymetrix HGU133 Plus 2.0 genechip and displaying patterns of gene expression different from background noise as identified by the Sleipnir library [1]. The LMX1A gene was flat, i.e., showed minimum expression intensity in all patients.

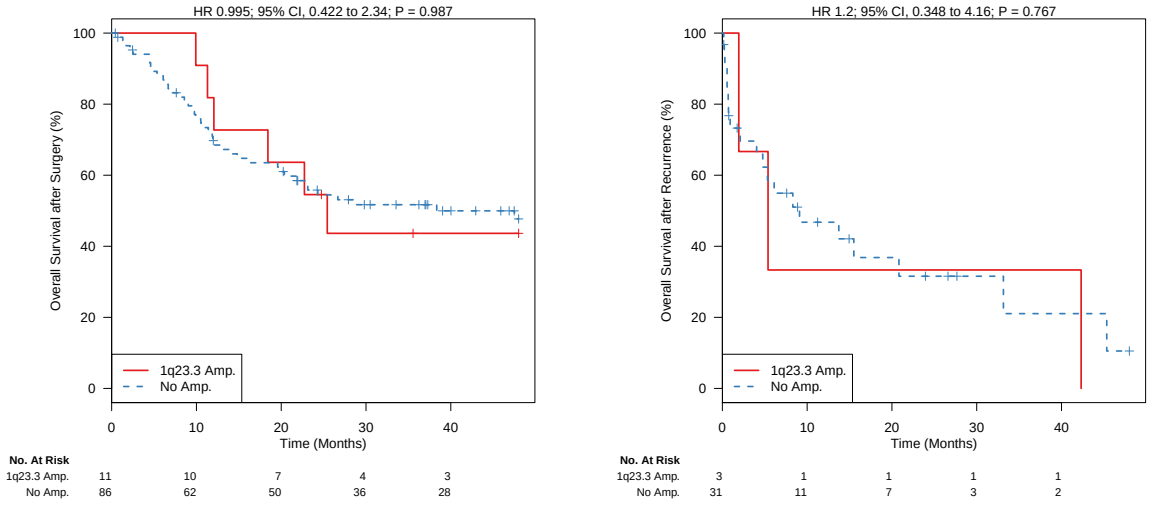
Genes	HR (95% CI) OS after Recurrence	FDR	HR (95% CI) OS after Cystectomy	FDR
PFDN2	1.91 (1.3-2.82)	0.01	1.22 (0.89-1.68)	0.53
PPOX	3.51 (1.45-8.52)	0.03	1.13 (0.53-2.39)	0.88
USP21	1.38 (1.05-1.81)	0.08	0.97 (0.77-1.23)	0.88
DEDD	1.34 (1.02-1.78)	0.12	0.92 (0.68-1.23)	0.88
UFC1	1.59 (0.87-2.9)	0.31	0.88 (0.54-1.45)	0.88
FCGR2B	0.91 (0.79-1.05)	0.40	1.01 (0.89-1.16)	0.88
PBX1	0.88 (0.71-1.09)	0.40	0.85 (0.66-1.08)	0.53
FCGR3B	0.92 (0.77-1.1)	0.46	0.99 (0.82-1.19)	0.88
USF1	5.02 (0.19-132.9)	0.46	30.63 (0.92-1020.81)	0.46
PVRL4	0.95 (0.8-1.12)	0.63	0.93 (0.8-1.08)	0.69
FCGR2C	1.07 (0.77-1.49)	0.67	1.26 (0.91-1.74)	0.53
KLHDC9	1.06 (0.81-1.4)	0.67	0.78 (0.59-1.03)	0.46
LMX1A	1 (1-1)		1 (1-1)	

Table S9: Overall survival (OS) Hazard Ratios (HRs) of genes in 1q21.2 as in Supplementary Table S8.

Genes	HR (95% CI) OS after Recurrence	FDR	HR (95% CI) OS after Cystectomy	FDR
CTSK	0.88 (0.76-1.03)	0.72	0.94 (0.81-1.1)	0.86
CTSS	0.9 (0.74-1.1)	0.72	0.88 (0.71-1.09)	0.77
ECM1	0.92 (0.72-1.18)	0.72	1 (0.78-1.29)	0.99
GOLPH3L	0.91 (0.64-1.28)	0.72	0.75 (0.51-1.1)	0.62
FAM63A	1.12 (0.7-1.8)	0.72	0.88 (0.55-1.39)	0.86
HORMAD1	1.07 (0.89-1.28)	0.72	1.33 (1.09-1.63)	0.05
HIST2H2BE	0.95 (0.77-1.18)	0.72	1.03 (0.85-1.24)	0.86
BOLA1	1.25 (0.87-1.81)	0.72	0.94 (0.67-1.31)	0.86
ANXA9	1.01 (0.86-1.18)	0.91	0.97 (0.83-1.14)	0.86

2 1q23.3 and survival after surgery

We finally tested whether 1q23.3 amplification was also associated with likelihood of recurrence. For that purpose, we tested 1q23.3 amplification with association of overall survival (OS) after initial surgery. For the Spanish cohort, surgery dates were not available. Only 2 patients with 1q23.3 amplification in the DFCI cohort had available radical cystectomy dates. As we have shown in Supplemental Table S8, mRNA expression of 1q23.3 genes was only associated with OS after recurrence of metastatic disease, not with OS after radical cystectomy. We tested for association with OS after surgery in an independet cohort from DFCI using multiplex ligation-dependent probe amplification. We found no association with OS after surgery in this cohort (Supplemental Figure S17). Three patients with 1q23.3 amplification recurred, prohibiting statistical confidence of association of this amplification with OS after recurrence in this cohort.



(a) (b)

Figure S17: Panel (a): No association of overall survival after surgery and 1q23.3 amplification in an independent cohort from DFCI. (b) No significant association of 1q23.3 amplification and OS after recurrence, possibly due to small patient numbers.

References

[1] C Huttenhower, M Schroeder, M D Chikina, and O G Troyanskaya. The sleipnir library for computational functional genomics. *Bioinformatics*, 24(13):1559–1561, Jul 2008.

[2] M N McCall, B M Bolstad, and R A Irizarry. Frozen robust multiarray analysis (frma). *Biostatistics*, 11(2):242–253, Apr 2010.

[3] M Riester, J M Taylor, A Feifer, T Koppie, J E Rosenberg, R J Downey, B H Bochner, and F Michor. Combination of a novel gene expression signature with a clinical nomogram improves the prediction of survival in high-risk bladder cancer. *Clin Cancer Res*, 18(5):1323–1333, Mar 2012.