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

August 15, 2013

1 Recurrent chromosomal gains and losses

As described in the main part of the paper, we used the GISTIC algorithm [1] to identify focal and broad recurrent copy number changes in the Spanish and the DFCI MIP cohorts. First we examined broad alterations, defined as those copy number alterations that spanned over 50% of the chromosome arm. Tables S1 and S2 list these broad events together with their frequencies, q-values and associations with overall survival after recurrence. No broad event was significantly associated with survival. In Figures S1 and S2, we examine the broad copy numbers at the arm level. These plots show whether the two arms of a chromosome are typically altered together or independently. As reported in the result section of the paper, both chromosome 9 arms are frequently lost together, whereas both arms of chromosome 20 is often gained. Figure S3 shows in another GISTIC output figure the negative correlation of the frequency of broad events and the number of genes on chromosome arm in both cohorts. Figures S4-S5 show the GISTIC plots for the focal amplifications and deletions (< 50% of the chromosome) in the Spanish and DFCI cohorts.

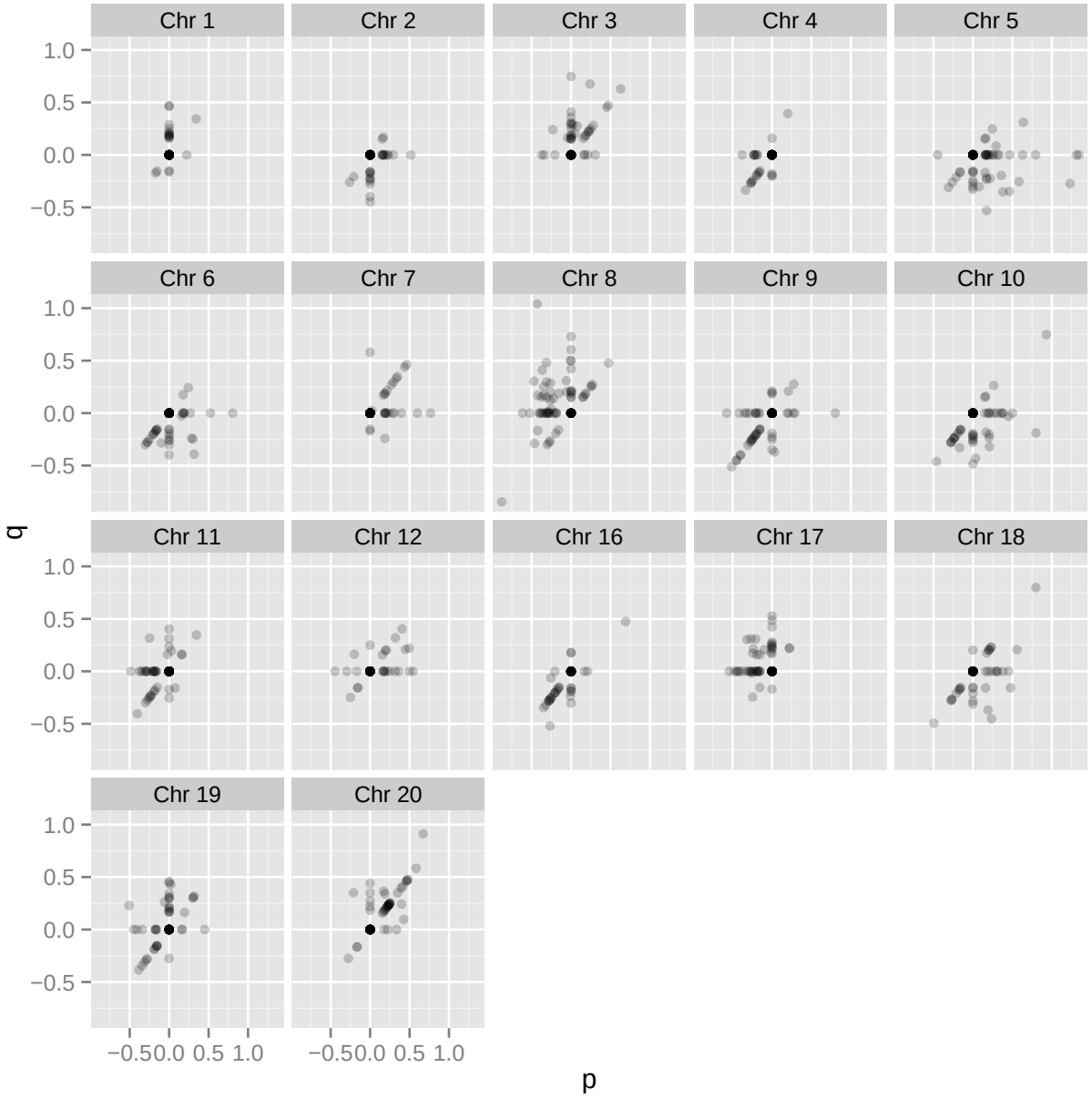


Figure S1: Broad copy numbers at the p- and q-arms in the Spanish cohort. Each point represents a patient and its $\log_2$ copy number ratios of the p- and q-arms. Points on the diagonal indicate whole chromosome gains or losses (chromosomes 9 and 20). Straight vertical lines (chromosome 1 for example) indicate gains or losses of the q-arm and horizontal lines frequent gains or losses of the p-arm (e.g., chromosome 11).

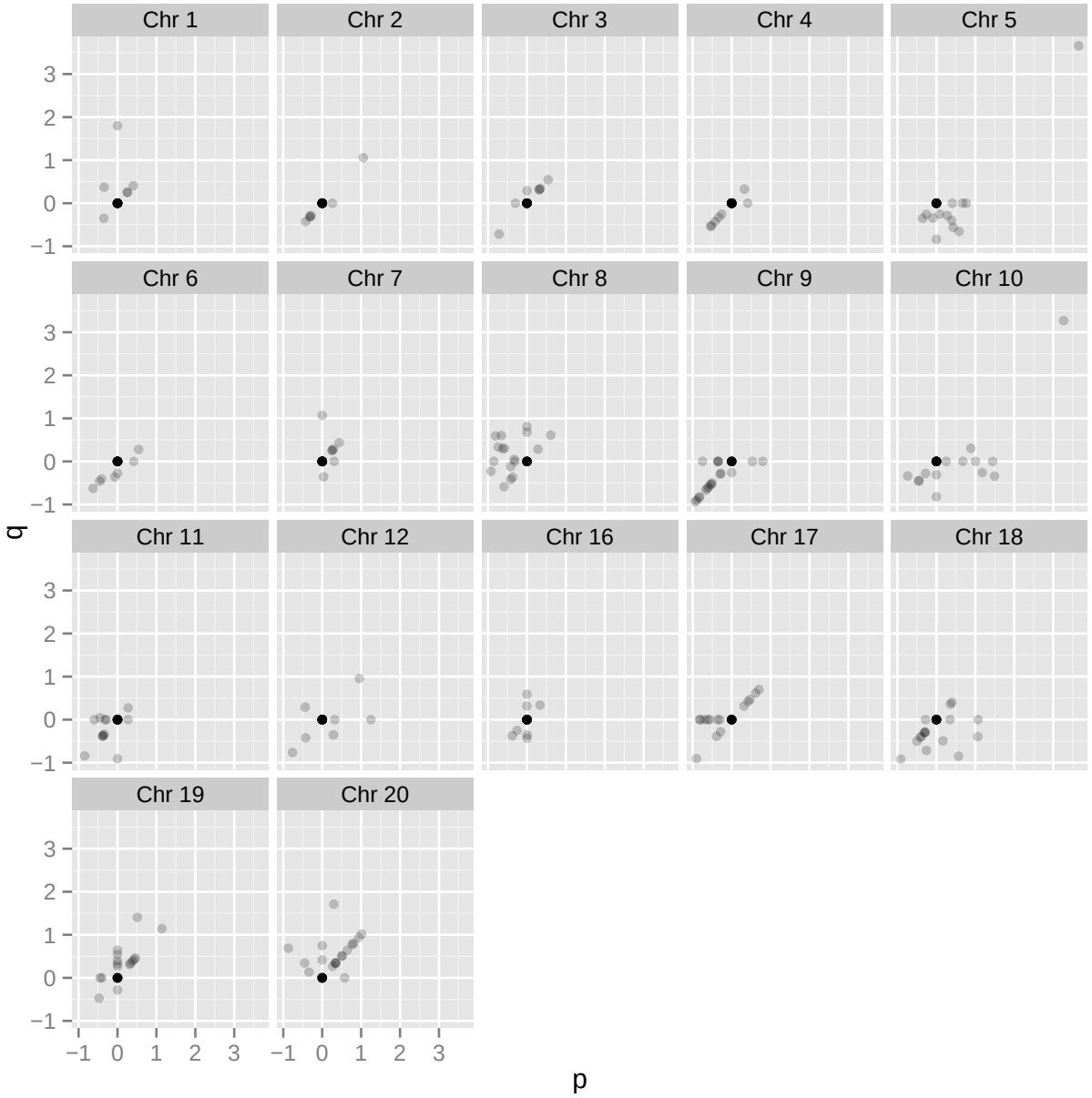
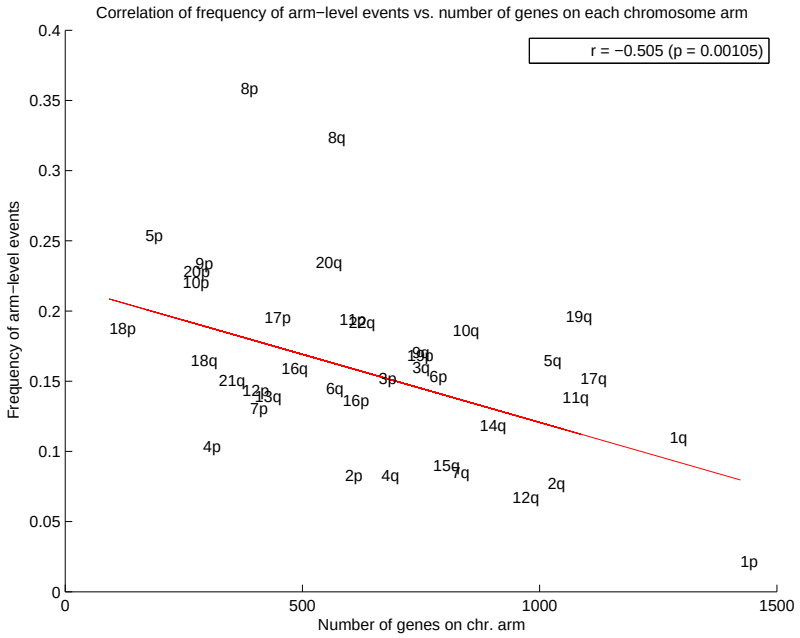
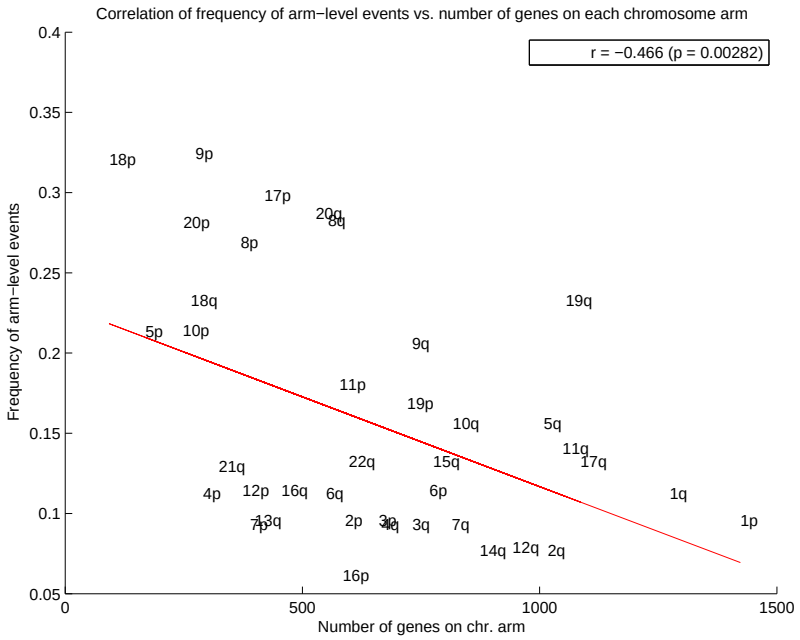


Figure S2: Broad copy numbers at the p- and q-arms in the DFCI cohort. (see Figure S1 for an explanation of this plot).



(a) Spanish



(b) DFCI

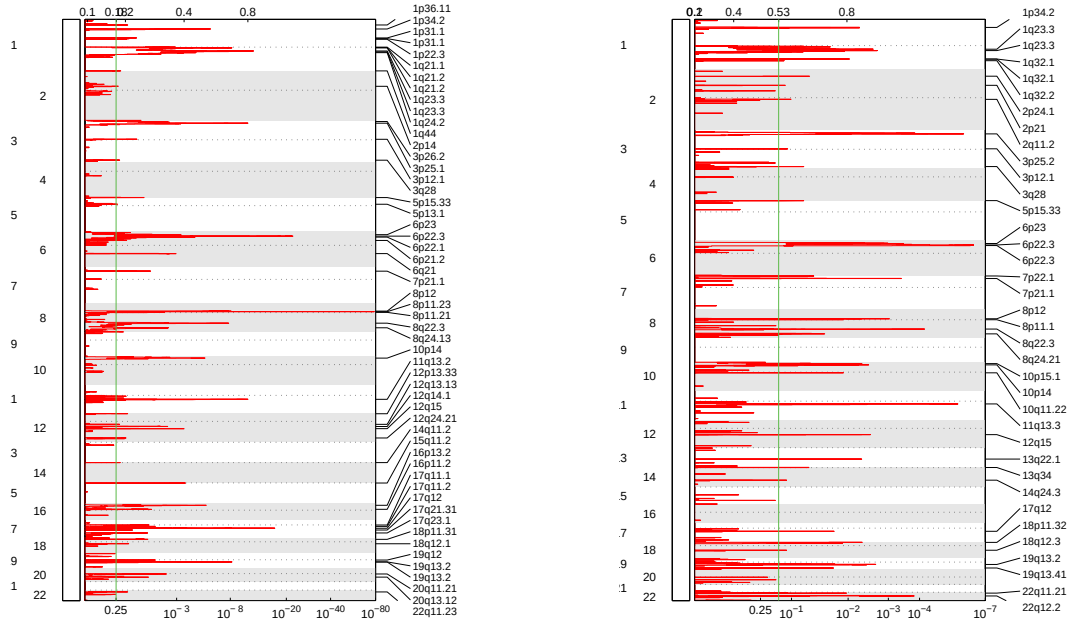
Figure S3: These GISTIC output files show the correlation between gene number and frequency of arm-level events in the (a) Spanish cohort and the (b) DFCI cohort. The red line represents the sum of least squares fit to the data. As gene number is a significant predictor of arm level events, this plot is important for the interpretation of Supplemental Tables S1-S2.

Table S1: Significant (q-value < 0.25) GISTIC broad aberrations and their association with overall survival after start of chemotherapy (Spanish cohort).

Arm	Number of Genes	Type	Frequency	z-score	q-value	Cox p-value	Cox FDR
8p	370	Del	0.55	8.77	< 0.001	0.20	0.64
20q	528	Amp	0.40	5.95	< 0.001	0.67	0.92
8q	554	Amp	0.41	5.79	< 0.001	0.11	0.46
5q	1009	Del	0.29	4.87	< 0.001	0.23	0.69
3q	732	Amp	0.32	4.73	< 0.001	0.06	0.43
20p	249	Amp	0.39	4.69	< 0.001	0.57	0.92
10q	817	Del	0.31	4.74	< 0.001	0.25	0.69
11p	578	Del	0.34	4.64	< 0.001	0.53	0.92
17p	420	Del	0.36	4.57	< 0.001	0.06	0.43
22q	598	Del	0.32	4.21	< 0.001	0.19	0.64
9p	275	Del	0.36	3.92	< 0.001	0.06	0.43
5p	169	Amp	0.36	3.76	< 0.001	0.06	0.43
17q	1086	Amp	0.24	3.69	< 0.001	0.97	0.98
9q	731	Del	0.27	3.28	0.00	0.10	0.46
19q	1055	Amp	0.23	3.30	0.00	0.65	0.92
16q	456	Del	0.27	2.55	0.02	0.60	0.92
6q	550	Del	0.26	2.50	0.02	0.65	0.92
19p	721	Del	0.24	2.46	0.02	0.98	0.98
1q	1274	Amp	0.16	2.04	0.10	0.84	0.97
16p	585	Del	0.23	1.81	0.11	0.07	0.43
7p	390	Amp	0.25	1.70	0.20	0.18	0.64
11q	1048	Del	0.16	1.40	0.22	0.39	0.92

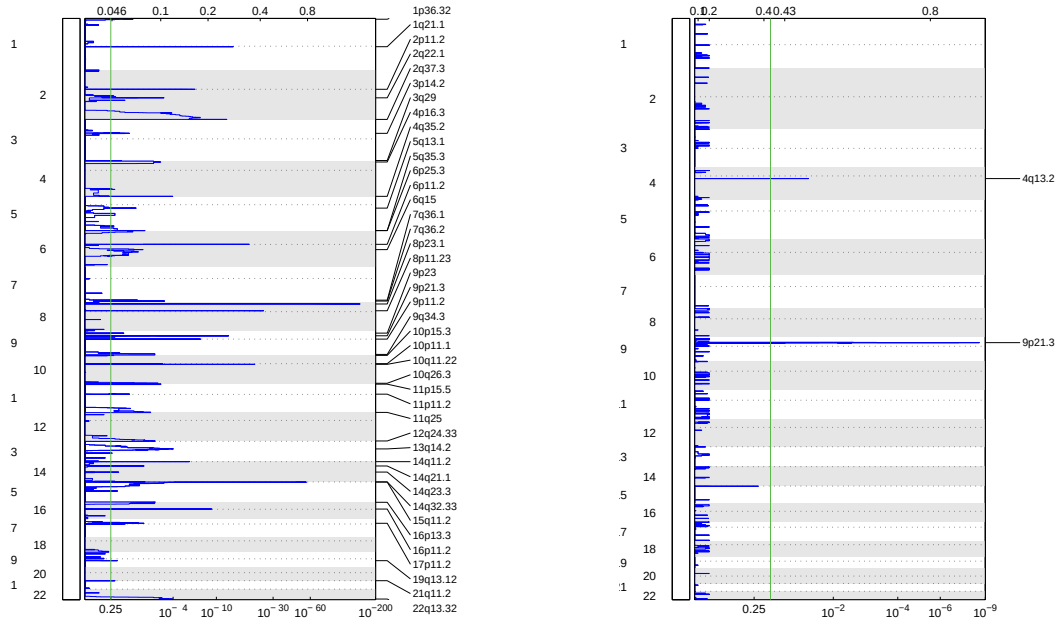
Table S2: Significant (q-value < 0.25) GISTIC broad aberrations and their association with overall survival after recurrence (DFCI cohort).

Arm	Number of Genes	Type	Frequency	z-score	q-value	Cox p-value	Cox FDR
19q	1055	Amp	0.38	4.77	< 0.001	0.58	1.00
9p	275	Del	0.53	4.73	< 0.001	0.77	1.00
20q	528	Amp	0.47	4.51	< 0.001	0.20	0.85
9q	731	Del	0.41	4.36	< 0.001	0.74	1.00
8p	370	Del	0.44	3.63	0.00	0.84	1.00
20p	249	Amp	0.42	3.04	0.02	0.12	0.85
5q	1009	Del	0.27	2.82	0.02	0.98	1.00
17p	420	Del	0.38	2.76	0.02	0.23	0.85
18p	93	Del	0.41	2.55	0.04	0.29	0.85
18q	265	Del	0.38	2.49	0.04	0.30	0.85
11q	1048	Del	0.24	2.39	0.04	0.26	0.85
10q	817	Del	0.27	2.26	0.05	0.15	0.85
11p	578	Del	0.28	1.79	0.14	0.66	1.00



(a) Spanish (b) DFCI

Figure S4: Recurrent copy number gains in the (a) Spanish and (b) DFCI cohort. False discovery rates (q-values; green line indicates the significance threshold of 0.25) for focal amplifications are plotted against the genome position. Dotted line show the centromers. GISTIC scores are plotted on the top x-axis, the corresponding q-values on the bottom x-axis.



(a) Spanish

(b) DFCI

Figure S5: Recurrent deletions in the (a) Spanish and (b) DFCI cohorts (see Figure S4 for an explanation of this plot). The smaller number of peaks in the DFCI cohort is mostly a result of the smaller sample size (the smaller the sample size, the more frequent an alteration must be to reach significance). Copy number differences between normal and altered loci are typically smaller in deletions, because rarely more than one copy is lost, whereas amplifcations can result in much higher copy number changes.

References

[1] C H Mermel, S E Schumacher, B Hill, M L Meyerson, R Beroukhim, and G Getz. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol*, 12(4), 2011.