Supplemental Note 1

The decision to combine data from two separate array platforms, the 250K SNP array and the higher density SNP6.0 genomic array, was made to enable GISTIC analysis across a larger sample set of gut adenocarcinomas. Here we evaluate the potential that combining data generated from two array platforms might introduce systematic bias.

First, we confirmed that the length distributions of copy-number alterations were similar between the two generations of genomic profiling arrays (Supplementary Figure 12). The distributions were similar between the two platforms for each tumor type. Second, we performed using unsupervised hierarchical clustering of the copy-number values at loci of GISTIC peaks. We saw no segregation of samples analyzed using SNP6.0 or 250K arrays. Rather, samples tended to separate more by tumor type with EA and CRC being the most different (Supplementary Figure 15).