

**Web Fig. B:** Selection of samples for training the classifier. For each tumour we calculated the number of genes for which the log fold-change was more than 1.5 above the median of the relevant group, and the number of genes for which the log fold-change was less than 1.5 below the median. Samples 1062-2, 1166-1, 1330-1, 320-7, and 956-2 show large differences from the group specific expression patterns and were not used to build the classifier.

