

### **Correlation of the cancer associated genes to tissue labels**

A two-fold change in gene expression levels between samples is often considered a lower limit for the detection of differential expression. To assess the association of the genes in  $M_{AD}$  with the tissue types, correlation coefficients of the expression profiles of the genes with a tissue vector were calculated. Figure 1 shows the histograms of the expression-tissue correlation coefficients for the genes. When compared with randomly assigned tissue labels, which show correlation coefficients close to and centered on zero, the experimental data shows two distinct peaks corresponding to genes down- (-ve values) and up- (+ve values) regulated in adenocarcinomas with few genes having values close to zero. The host cell factor binding transcription factor ZhangFei (ZF) has the highest correlation coefficient ( $r = 0.675$ ) to the tissue labels, while the gene RAGE was the most negatively correlated gene ( $r = -0.956$ ), consistent with its low expression levels noted above.