Supplemental Figure 4 (A) Principal component analysis (PCA) of 65 probe set signature defined in classifier 3 (FA and FEA merged versus FC). Each dot represents a sample, adenoma - orange, carcinoma - blue. The three dimensional PCA plot captures 74% of the variance described by these genes. The percentage of the total variance described by each of the three principal components is shown in parenthesis at each axis. The groups are well separated, although five samples are miss-classified during LOOCV training; three adenoma (FA-FEA) samples and two FC samples. The follicular carcinoma cluster is more heterogeneous than the adenoma samples showing the progressive nature of the cancer samples.

(B) Plot of the predictive Probability output obtained for classifier 3. The classifier was trained to distinguish between the merged group of fetal adenomas and follicular adenomas versus follicular carcinomas. If a sample has a predictive value above 0.5 (p(fa) > 0.5), it is classified as follicular adenoma, otherwise as a follicular carcinoma. Each dot represents a sample, where the color code shows the true class. Adenomas – orange and follicular carcinoma – blue. Five samples were misclassified, three adenoma (FA-FEA) samples and two FC samples.