New Composite Biomarkers for Colorectal Cancer Diagnosis

H Mansour, R Incitti, VB Bajic

Introduction

Colorectal cancer (CRC) is frequent in western countries and its incidence is increasing by 3%-4% per year since the 1970s. It is the second leading cause of death from cancer worldwide. Mass screening for CRC and pre-cancerous lesions (adenomatous polyps) using the biochemical faecal blood test has led to an 18%-25% reduction in mortality. However, the sensitivity of fecal occult-blood testing for colorectal cancer is low. Colorectal cancer neoplastic cells have proven to be an important paradigm due to the heterogeneity of molecular pathways involved in its development. Noninvasive tests that detects tumor-specific products with reasonable sensitivity and specificity are still lacking. Using a computational method we identified a new set of methylation-based biomarkers to enable potentially non-invasive diagnosis of CRC with high specificity and specificity.

Method

We developed a computational method that involves:

1/ Screening relevant CRC tissue methylated genes from Pubmed Bibliography (more than 2000 entries).

2/ Inferring serum methylation value of the genes from their tissue expression data available in GEO-NCBI (32 tissue cancer versus 32 autologous normal tissue).

3/ Selecting a panel of statistically significant hypermethylated genes in cancer versus normal.

4/ Testing the panel's combined predictive performance with a thresholding algorithm. We also tested predictive performance in the presence of resistance to methylation noise.

Results

We identified 95 relevant genes and inferred their serum methylation values. The classification of these genes according to their molecular functions and signaling pathways using the Panther Classification System (http://www.pantherdb.org), shows the heterogeneity of signaling pathways associated with CRC (Figure 1). 25 genes from the 95 selected are statistically hypermethylated in CRC versus normal (Mann-Whitney test, $p<10^{-5}$). Testing these 25 biomarkers by a thresholding algorithm including noise shows they are able to discriminate 100% (32) of the CRC patients from the normal, even with increasing the error margin from 5% to 33% (Figure 2).

Predictive performance tests identify 14 markers that, when used in combination, provide 100% of sensitivity and specificity..

Conclusion

Our finding proposes relevant non-invasive composite markers for the diagnosis of CRC and our strategy could be extended to identify other composites for the screening of other cancers.

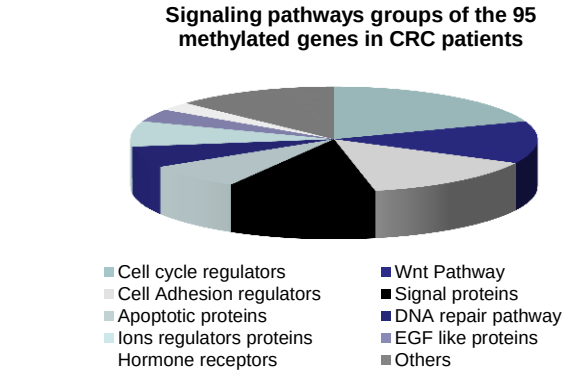

Figure 1: Signaling pathways of the selected methylated genes

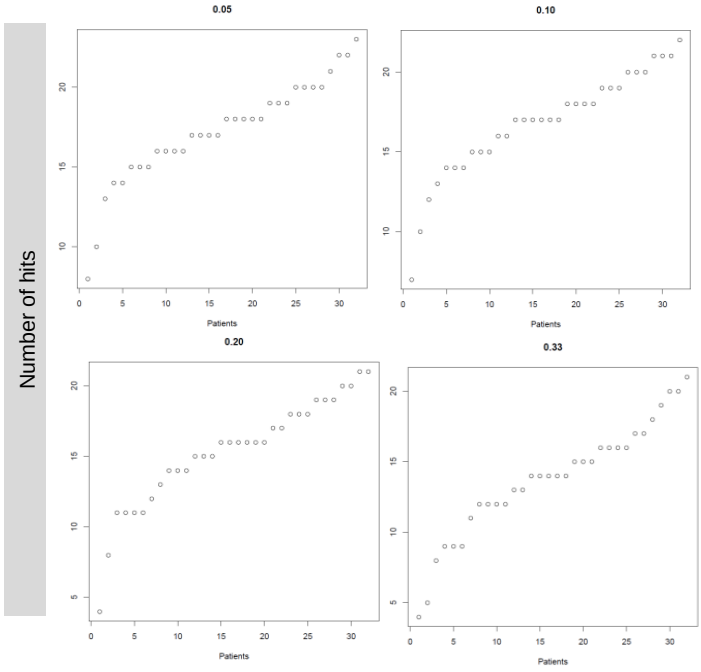

Figure 2: Hits Distribution at different cutoffs . For each patient (x-axis), we plot the number of hits (y-axis), i.e. methylation values above a gene's specific threshold plus an error margin of 5%, 10%, 25%, 33% as defined below. Patients are ordered by number of hits. For a gene g, the threshold Hg is the highest methylation value in normal samples. The error margin is defined as X times the difference between the full methylation value (100%) and Hg.