FACULTY OF **BIOSCIENCE ENGINEERING**

Discovery of methylation markers using a relaxation ranking algorithm

Maté Ongenaert[1], Wim Van Criekinge[1,2], Josef Straub[2], Bea Wisman[3], Ed Schuuring[3], Ate Van der Zee[3], Herman Spolders[2]

[1] Department of molecular biotechnology, Faculty of BioScience Engineering , Ghent University, Belgium
[2] Oncomethylome Sciences, Liege, Belgium
[3] Departments of Gynaecological Oncology and Pathology, University Medical Center Groningen, Groningen, The Netherlands.

Introduction

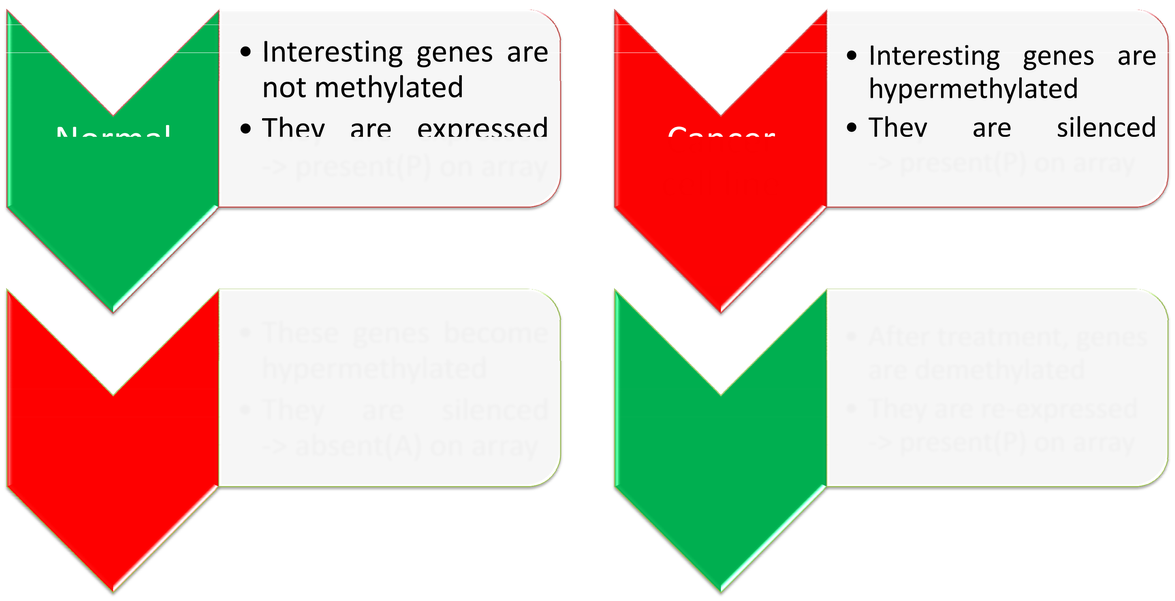
DNA methylation represents a modification of DNA by addition of a methyl group to a cytosine also referred to as the fifth base[1]. This epigenetic change does not alter the primary DNA sequence and might contribute to overall genetic stability and maintenance of chromosomal integrity and consequently facilitate organization of the genome into active and inactive regions with respect to gene transcription[2].

Genes with CpG islands in the promoter region are generally unmethylated in normal tissues. Upon DNA hypermethylation, transcription of the affected genes may be blocked, resulting in gene silencing. In neoplasia, abnormal patterns of DNA methylation have been recognized and hypermethylation is now considered one of the important mechanisms resulting in silencing expression of tumour suppressor genes, i.e. genes responsible for control of normal cell differentiation and/or inhibition of cell growth.

In the last few years, new hypermethylated biomarkers have been used in cancer research and diagnostics[3]. In many cancers various markers have been reported to be hypermethylated[4]. Most of these markers were studied because of their involvement in cancer development, cell growth and DNA damage[5]. Presently hypermethylation of only very few of these markers is of clinical relevance and/or cancer-type specific.

Methylation in cervical cancer

We want to investigate methylation in cervical cancer, and therefore use two technologies:
• micro-array technology: if gene is silenced by methylation, it is not visible on the array
• de-methylation strategy: by applying DAC and/or TSA we are able to remove methylation in cancer cell lines

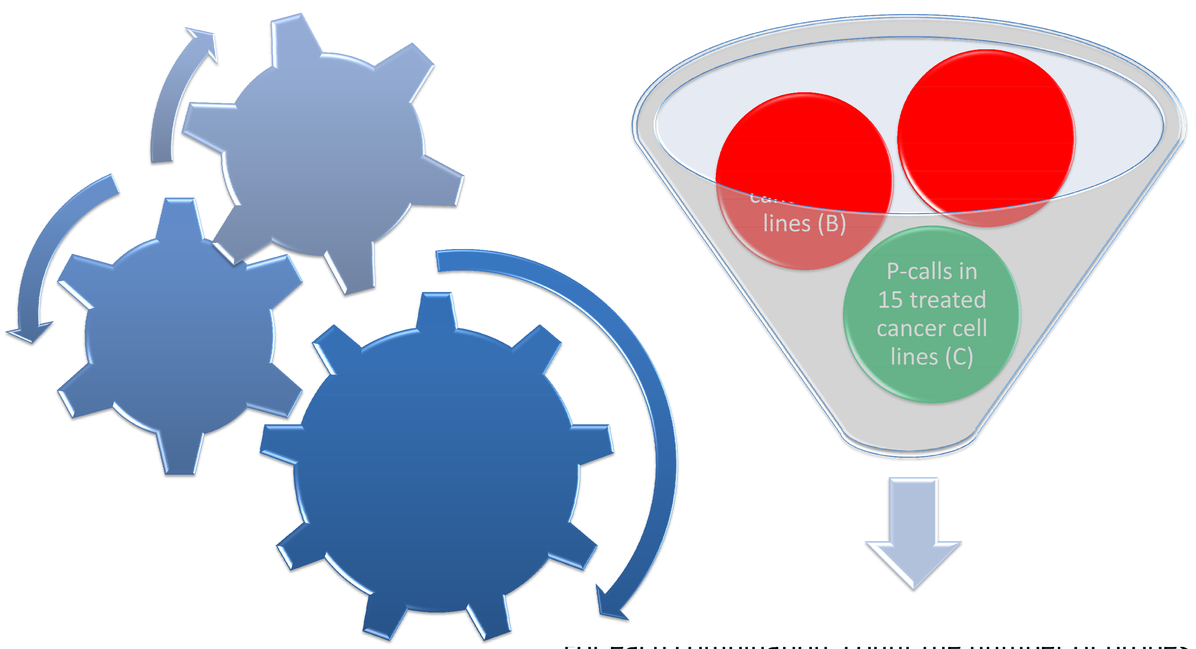


- 39 primary cancer samples
- 4 untreated cancer cell lines
- 15 treated (demethylated) cancer cell lines

Affymetrics micro-array with > 260000 human transcripts

Ideal geneprobe:
- absent in all primary cancers
- absent in cancer cell-lines
- present in demetylated cancer cell lines

How to select interesting probes in the array?



Generate all combinations of X, Y and Z

For each combination, count the number of probes (W)that pass this filter. Store X, Y, Z and W

Relaxation ranking

Now, we know the number of probes that would be selected with combinations of different selection criteria. In the ideal conditions (X=0: no expression in any primary cervical cancer sample; Y=0: no expression in any cervical cancer cell line and Z=15: re-expression of the gene in all demethylated cervical cancer cell lines) no any gene is selected.

So, we'll have to relax conditions. This is primary driven by W (the number of genes that pass the filter under conditions X, Y and Z). We sort, based on W (small to large) , followed by X (small to large), Y (small to large) and Z (large to small).

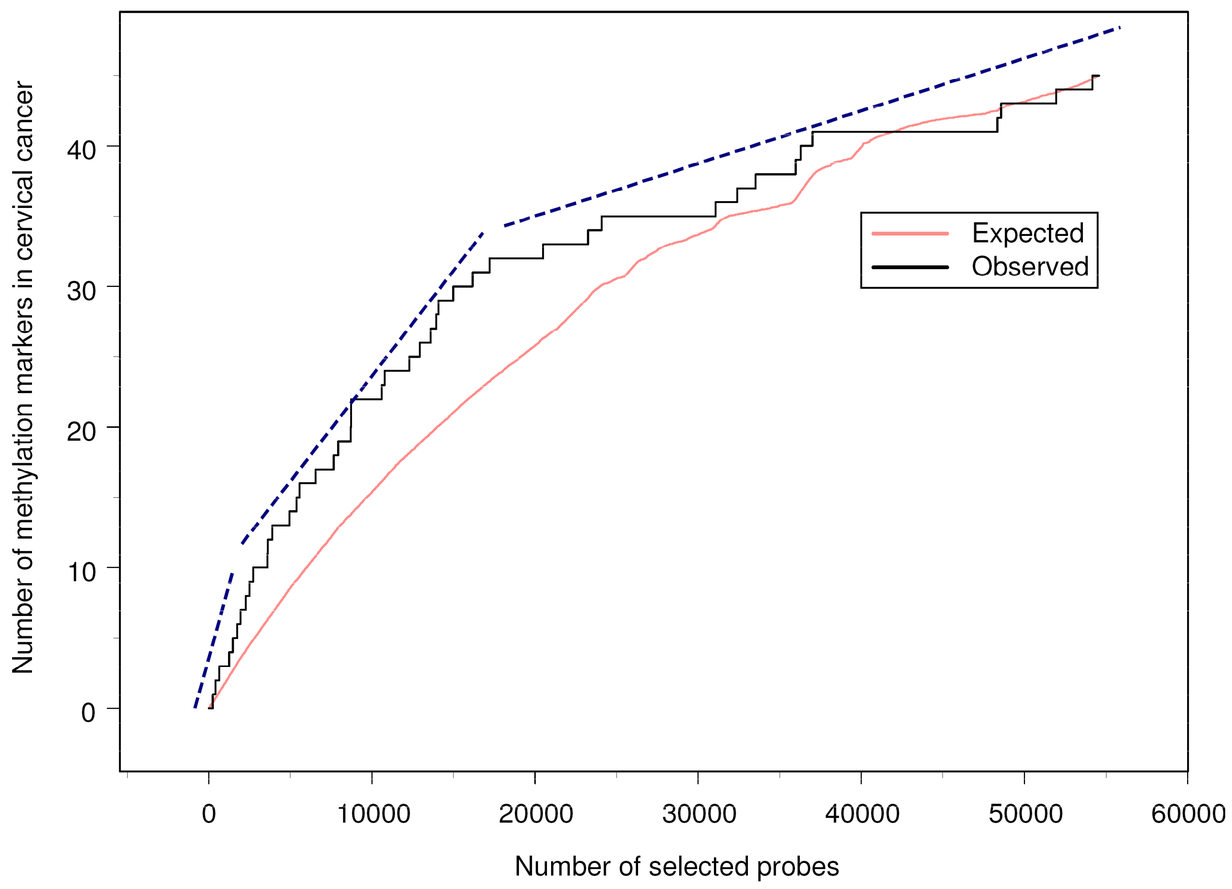
This way, we sort the entire dataset, and we start selecting interesting probes that would be found with the conditions in the first row; add this probe to a list. We continue to the next row, and if the genes that would be found with the conditions X, Y and Z are not already on the list, we add them. This process continues until there are as much genes on the list as we want.

Performance of the ranking algorithm

To test the performance of the relaxation ranking algorithm, we determine how many known methylation markers (described in literature) can be picked up and compare this result with random selection.

The algorithm is well suited to enrich towards methylation markers, as in the first 3000 probes, 10 (of 45 in total) known markers can be found, while if randomly selected, less than 3 would be in the top-3000.

Another striking example of the performance is the enrichment towards X-located genes. As in females, one copy of the X-chromosome is silenced and one of the main mechanisms is DNA-methylation, this confirms the potential of the ranking strategy.



Validation

After applying some additional criteria (such as not on the X-chromosome, not imprinted, a maximum of P-calls in primary cancer samples), 10 high-ranking genes are validated in 10 primary cervical cancer samples. 6 genes show DNA-methylation and 2 genes are methylated in cancer patients but not in normal persons. These and other genes are in validated in ongoing research to assess their diagnostic potential.

References

1. Doerfler,W. et al. Eukaryotic Dna Methylation - Facts and Problems. Febs Letters 268, 329-333 (1990).
2. Robertson,K.D. DNA methylation and chromatin - unraveling the tangled web. Oncogene 21, 5361-5379 (2002).
3. Esteller,M. Cancer epigenetics: DNA methylation and chromatin alterations in human cancer. New Trends in Cancer for the 21St Century 532, 39-49 (2003).
4. Paluszczak,J. & Baer-Dubowska,W. Epigenetic diagnostics of cancer - the application of DNA methylation markers. Journal of Applied Genetics 47, 365-375 (2006).
5. Rountree,M.R., Bachman,K.E., Herman,J.G., & Baylin,S.B. DNA methylation, chromatin inheritance, and cancer. Oncogene 20, 3156-3165 (2001).