DNA methylation-based classification of human central nervous system tumors

Martin Sill, German Cancer Research Center (DKFZ)
DNA methylation / Illumina methylation arrays

Cytosine  methylated Cytosine

Typical mammalian DNA methylation landscape

Sample Locus:

- Methylated locus
- Unmethylated locus

Methylation involves:

- Bisulfite conversion
- Whole genome amplification
- Enzyme fragmentation
- Hybridize

Allele-specific primer annealing
Single-base extension

- 450k probes / beta values between 0 and 1
Brain tumor reference set / TSNE

- 82 tumor classes
- 9 non-tumor classes
- 2801 samples

t-SNE dimensionality reduction

2,801 samples
Random Forest classifier

- variable selection by RF permutated importance measure
- subsampling to minority class
- ridge-penalized, multinominal GLM for score calibration
Performance evaluation

Figure 3

Cross-validation

<table>
<thead>
<tr>
<th>MCF MB, G3/4</th>
<th>MCF MB, SHH</th>
<th>MCF ATRT</th>
<th>MCF ENB</th>
<th>MCF GBM</th>
<th>MCF ATRT, MYC</th>
<th>ATRT, SHH</th>
<th>ATRT, TYR</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.4</td>
<td>100</td>
<td>3.4</td>
<td>94.6</td>
<td>86.7</td>
<td>96.6</td>
<td>99.0</td>
<td>99.5</td>
</tr>
</tbody>
</table>

Percent of samples in annotated class

<table>
<thead>
<tr>
<th>GBM, MES</th>
<th>GBM, MID</th>
<th>GBM, MYCN</th>
<th>GBM, RTK I</th>
<th>GBM, RTK II</th>
<th>GBM, RTK III</th>
</tr>
</thead>
<tbody>
<tr>
<td>96.6</td>
<td>99.0</td>
<td>99.5</td>
<td>99.0</td>
<td>99.5</td>
<td>99.5</td>
</tr>
</tbody>
</table>

Error rate | AUC | Brier score | Sensitivity | Specificity
---|-----|-------------|-------------|------------|
Raw score   | 4.89% | 0.9996 | 0.444 | 0.951 | 0.999 |
Calibrated score | 4.28% | 0.9998 | 0.070 | 0.957 | 0.999 |
Calibrated MCF score | 1.14% | 0.9999 | 0.024 | 0.989 | 0.999 |
library(minfi)  # basic methylation array analysis
library(conumee)  # copy-number variation analysis
library(mnp.v11b4)
RGset <- read.metharray("path2idats/")
MNPreport(RGset)
Depiction of chromosome 1 to 22 (and X/Y if automatic prediction was successful). Gains/amplifications represent positive, losses negative deviations from the baseline. 29 brain tumor relevant gene regions are highlighted for easier assessment. (see Hovestadt & Zapatka, http://www.bioconductor.org/packages/devel/bioc/html/conumee.html)
Welcome to MolecularNeuropathology.org - The platform for next generation neuropathology.

This website is your access point to a DNA methylation based molecular classifying algorithm for most types of tumors of the central nervous system and its coverings.

To implement the methylation profiling classifier you are required to generate and upload unprocessed IDAT-files of Illumina Human Methylation 450 BeadChip arrays or EPIC BeadChip arrays of your tumors of interest. This data is then automatically compared to methylation data of a reference cohort comprising over 2500 neuropathological tumors of almost all known entities (currently over 80 tumor classes or subclasses included). Within minutes you will receive an E-Mail report of the methylation profiling of your case, a low resolution copy number plot calculated from your array data (useful e.g. for 1p/19q analysis or the detection of all sorts of amplifications and deletions) and an estimation of MGMT promoter methylation status.

Classification using methylation profiling is a research tool under development, it is not verified and has not been clinically validated. Implementation of the results in a clinical setting is in the sole responsibility of the treating physician.

Upload statistics:
- IDAT files processed: 3269
- Samples approved for classifier development: 2547
- Statistics updated: 2017-06-28 11:52 CEST

Involved parties:
- University Hospital Heidelberg
  - Neurology
  - Pediatric Oncology
- Neurooncology
- Neurosurgery
- Radiation Oncology and Therapy
- German Cancer Research Center (DKFZ)
  - Pediatric Neurooncology
  - CCU Neuropathology
  - Biostatistics
  - Molecular Genetics
- German Consortium for Translational Cancer Research (DKTK)
Pediatric Neurooncology DKFZ
Stefan Pfister
David Jones
Lukas Chavez
Dominik Sturm

Biostatistics DKFZ
Axel Benner
Annette Kopp-Schneider

Neuropathology Heidelberg
Andreas von Deimling
David Capper
Daniel Schrimpf
Damian Stichel
Felix Sahm

Broad Institute
Volker Hovestadt

Genomics and Proteomics Core Facility DKFZ
Matthias Schick
Melanie Bewerunge-Hudler
Oliver Heil
Roger Fischer

molecularneuropathology.org