

MAID – Gene Filtering

In this tutorial we will pre-process human Affymetrix data and filter significantly regulated genes using MAID (MA-plot-based signal intensity-dependent fold-change criterion).

Data:

We analyse Affymetrix HG-U133A microarray data of three rheumatoid arthritis patients starting Etanercept therapy (MAID_sample_data.zip). For each patient gene expression levels were obtained at baseline (t0) as well as after three days post therapy initiation (t1). Here, we aim to identify genes consistently up- or down-regulated in this particular group of patients. The data are part of the data set published by Koczan et al. (2008). The full data set providing transcript levels of 19 patients is available from [ArrayExpress](#) (accession number [E-MTAB-11](#)).

Methods:

In the applied Affymetrix microarrays most probesets include probes matching transcripts of more than one gene and probes which do not match any transcribed sequence. Therefore, we utilize a [custom chip definition file](#) (CDF) that is based on the information contained in the [GeneAnnot](#) database. This allows us to avoid redundant probesets and to generate more reliable gene expression levels. Here we use [CDF package version 2.1.0](#) which is based on GeneAnnot version 1.9 and synchronized with [GeneCards](#) version 2.41.1. For convenience, GeneCards_2.41.1.RData provides selected information for each GeneCard in the database. The gene filtering is then based on two criteria: a t-statistic and so-called MAID-scores. By this means, we detect common transcriptional effects during the first three days of therapy. The respective methods are implemented in MAID_tutorial.r and the accompanying R code file MAID_tutorial_codes.r. MAID and the transcriptional regulation in response to Etanercept therapy are described in greater detail in Hecker et al. (2009).

Downloads:

-  [MAID_tutorial.r \(7,4 KiB\)](#)
-  [MAID_tutorial_codes.r \(8,9 KiB\)](#)
-  [MAID_sample_data.zip \(21,7 MiB\)](#)
-  [GeneCards_2.41.1.RData \(1,4 MiB\)](#)

Other GeneCards versions:

-  [GeneCards_2.40.1.RData \(1,4 MiB\)](#)
-  [GeneCards_2.40.2.RData \(1,4 MiB\)](#)
-  [GeneCards_2.41.RData \(1,4 MiB\)](#)
-  [GeneCards_2.42.RData \(1,7 MiB\)](#)

References:

- Hecker M, Goertsches RH, Engelmann R, Thiesen HJ, Guthke R (2009) Integrative modeling of transcriptional regulation in response to antirheumatic therapy. BMC Bioinformatics 10(1):262. [\[link\]](#)
- Koczan D, Drynda S, Hecker M, Drynda A, Guthke R, Kekow J, Thiesen HJ (2008) Molecular discrimination of responders and nonresponders to anti-TNFalpha therapy in rheumatoid arthritis by etanercept. Arthritis Res Ther 10(3): R50. [\[link\]](#)