

Readings for Module 2: Normalization and Missing Values

Affymetrix, Microarray Suite version 5.0 (2002)

Statistical Algorithms Reference Guide

(paper, comparison analysis. Normalization proposed in Affymetrix algorithms).

Schuchhardt, J., Beule, D., et al. (2000)

Normalization strategies for cDNA microarrays.

Nucleic Acids Research 28, No.10.

(paper, list of sources of unwanted variation).

Finkelstein D.B., Ewing R., Gollub J., Sterky F., Somerville S., Cherry J.M. (2002)

Iterative linear regression by sector

Methods of Microarray Data Analysis , eds. SM Lin, KF Johnson (Kluwer Academic), pp.

57-68. Stanford preprint available at:

http://afgc.stanford.edu/afgc_html/CAMDA_11_update.pdf

(paper, linear fits by sector, iterated with outliers elimination. Spotted arrays)

Yee Hwa Yang, Sandrine Dudoit, Percy Luu, Terry Speed (2001)

Normalization for cDNA microarray data

UC Berkeley Statistics tech. report available at:

<http://www.stat.Berkeley.EDU/users/terry/zarray/Html/normspie.html>

(use of non-parametric fits, and of alternative transformations of the signals).

Kepler T, Crosby L., and Morgan K.T. (2000)

Normalization and analysis of DNA Microarray Data by Self-Consistency and Local Regression

SFI tech. report available at:

<http://www.santafe.edu/sfi/publications/Abstracts/00-09-055abs.html>

Hartemink A.J., Gifford D.K., Jaakkola T.S., Young R.A. (2001)

Maximum likelihood estimation of optimal scaling factors for expression array normalization.

SPIE BIOS 2001, San Jose, California. MIT preprint 2001 available at:

<http://www.psrq.lcs.mit.edu/publications/Papers/normabs.htm>

(file, normalization using controls, and a model estimated with ML and Bayesian priors. Affymetrix arrays)

Kerr M.K., Martin M., Churchill G. (2000)

Analysis of variance for gene expression microarray data.

Journal of Computational Biology 7, 819-837. JAX preprint available at:

<http://www.jax.org/research/churchill/research/expression/anova.html>

(and many other papers by the same group, designed experiments)

Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D., and Altman, R. B. (2001).

Missing value estimation methods for DNA microarrays.

Bioinformatics 17, 520-525.

http://ismb00.sdsc.edu/9_8_pdfs/SunMethods/Troyanskaya.pdf

(using nearest neighbors or leading components for missing value imputation)