

**Poster L-35**  
**Graphical User Interfaces to**  
**Microarray Analysis Programs**



**Authors:**

Keith Satterley (*Bioinformatics Division, The Walter and Eliza Hall Institute of Medical Research*)

James Wettenhall (*Peter MacCallum Cancer Centre*)

Ken Simpson (*Bioinformatics Division, The Walter and Eliza Hall Institute of Medical Research*)

Gordon K Smyth (*Bioinformatics Division, The Walter and Eliza Hall Institute of Medical Research*)

**Short Abstract:** Graphical User Interfaces for Microarray Analysis. affyImGUI provides an integrated workflow for Affymetrix microarray data analysis. limmaGUI analyses two-colour spotted microarray experiments. Both packages interface to the limma package. affyImGUI also interfaces to various Affymetrix chip analysis packages. affyImGUI has recently been extended to include extra quality control tools.

**Long Abstract:**

Graphical User Interfaces to Microarray Analysis Programs.

James M. Wettenhall, Ken M. Simpson, Keith Satterley\* and Gordon K. Smyth

\* Presenter.

**Background**

While a wealth of advanced software for statistical analysis of microarrays is available in R packages on Bioconductor (<http://www.bioconductor.org>), these packages have sophisticated command-driven interfaces tuned to users from mathematical or computing backgrounds. Most biologists instead use commercial software packages, incorporating possibly less cutting edge statistical methods, because they are easier to use. One of the most commonly used Bioconductor packages is the limma (Linear Models for Microarray Analysis) package which provides data analysis and normalization for cDNA microarray data and analysis of differential expression for multi-factor designed experiments. The core of the limma package is an implementation of the Empirical Bayes linear modeling approach of Smyth. LimmaGUI provides a point and click interface to the core functionality of the limma package.

limmaGUI analyses two-colour spotted microarray experiments. affyImGUI provides an integrated workflow for Affymetrix microarray data analysis. Both packages interface to the limma package. affyImGUI also interfaces to various Affymetrix chip analysis packages. affyImGUI has recently been extended to include extra quality control tools.

**limmaGUI**

limmaGUI is a Graphical User Interface (GUI) used for the exploration and linear modeling of data from two-colour spotted microarray experiments, especially the assessment of differential expression in complex experiments. limmaGUI provides an interface to the

statistical methods of the limma package for R, and is itself implemented as an R package. The software provides point and click access to a range of methods for background correction, graphical display, normalization, and analysis of microarray data. Arbitrarily complex microarray experiments involving multiple RNA sources can be accommodated using linear models and contrasts. Empirical Bayes shrinkage of the gene-wise residual variances is provided to ensure stable results even when the number of arrays is small. Integrated support is provided for quantitative spot quality weights, control spots, within-array replicate spots and multiple testing. limmaGUI is available for most platforms on the which R runs including Windows, Mac and most flavours of Unix.

A particular area of need that limmaGUI has addressed is to facilitate the construction of appropriate design matrices for linear modeling and appropriate contrast matrices to extract comparisons of interest from those linear models. While the linear modeling approach allows for testing of differential expression in very complex experimental designs, users who are not professional statisticians may be discouraged by the difficulty of specifying appropriate design and contrast matrices using a command-line interface. limmaGUI automates the construction of these matrices, asking users only to specify the RNA sources they would like to compare.

limmaGUI reads text files containing raw intensity data exported by a variety of image analysis programs including GenePix, SPOT, ImaGene, QuantArray and ArrayVision.

Diagnostic plots include MA-plots and spatial image plots. Selections of standard and non-standard background correction and normalization options are provided, including those described by Smyth and Speed. Within-array normalization methods included print-tip loess and composite loess normalization. Multi-array or between-array methods include simple scale normalization.

The limmaGUI menus are intended to provide capabilities suitable for most users, but advanced users can use a command-line interface within limmaGUI to issue commands directly to the R interpreter.

## affyImGUI

affyImGUI is a graphical user interface (GUI) to an integrated workflow for Affymetrix microarray data. The user is able to proceed from raw data (CEL files) to QC and pre-processing, and eventually to analysis of differential expression using linear models with empirical Bayes smoothing. Output of the analysis (tables and figures) can be exported to an HTML report. The GUI provides user-friendly access to state-of-the-art methods embodied in the Bioconductor software repository.

affyImGUI provides an interface to the affy, gcrma, affyPLM and limma packages of Bioconductor. The software is itself implemented as an R package and uses the interface to Tcl/Tk provided by the R package tcltk. The package enables users to pre-process and visualize their data and generate lists of putatively differentially expressed genes. Users have a choice of several state-of-the-art pre-processing methods for Affymetrix CEL files and advanced statistical methods for assessing differential expression. The package provides powerful statistical methods for dealing with small sample sizes and with complex

experiments involving many different RNA sources.

The expression data are read from the CEL files using the affy package. At this point, the user may produce several diagnostic plots, including MA-plots of the perfect match probes and histograms of the raw intensities. Quality assessment is followed by background correction, normalization and summarization of the probe-level data into probe-set expression values.

After probe-set expression summaries are obtained, the user can proceed to differential expression. The approach taken by the limma package is to analyse the differential expression in terms of linear models. This approach has many advantages as it allows very general experiments to be analysed in a unified framework, including factorial, saturated or loop designs and time course experiments, but it requires some mathematical sophistication. It requires the user to specify two matrices, the design matrix, which provides a representation of the RNA targets that have been hybridized to the arrays, and the contrast matrix which defines which comparisons between the RNA targets are of interest to the experimenter. affylmGUI greatly eases this process by largely automating the formation of the two matrices. The design matrix is constructed without user intervention. A set of dialogs help the user to define a set of comparisons of interest from which the contrast matrix is constructed

Additional quality control tools are being added to affylmGUI. This includes Log density plots for all chips, chip images of probe weights for all chips (with user specifiable layouts available), Normalized Unscaled Standard Error (NUSE) box plots, Relative Log Expression (RLE) plots and RNA degradation plots.

Software is available from <http://bioinf.wehi.edu.au/limmaGUI> and <http://bioinf.wehi.edu.au/laffylmGUI>