rptDB: a prototype Affymetrix .rpt QC tool

James Hadfield[1], Marielle Vigouroux
[1]james.hadfield@bbsrc.ac.uk, John Innes Centre, Norwich, NR3 4DG, UK,

Affymetrix analysis results in the generation of large and complex data sets which need extensive downstream analysis to determine results of biological significance. Users can benefit from a careful QC of their data before beginning analysis to remove individual chips that may adversely affect the final interpretation.

rptDB is a prototype database for the analysis of Affymetrix expression report (.rpt) files. Report files are widely used to QC raw data after hyb, scan and .cel file generation. Only %Present calls, Scale Factors and 3':5' ratios of housekeeping genes are generally used, but there are many other metrics that may also be of importance for QC analysis.

This is a prototype and we would welcome community response and suggestions for development.

Why use rptDB?

• Currently all data QC is done internally, comparison to external data increases robustness

• Current methods are subjective

• Final goal is a Phred type score for Affymetrix arrays

• Analyse trends across techs/labs/countries

rptDB functions:

• Generate average .rpt file statistics for any chip type
• Upload data for QC analysis
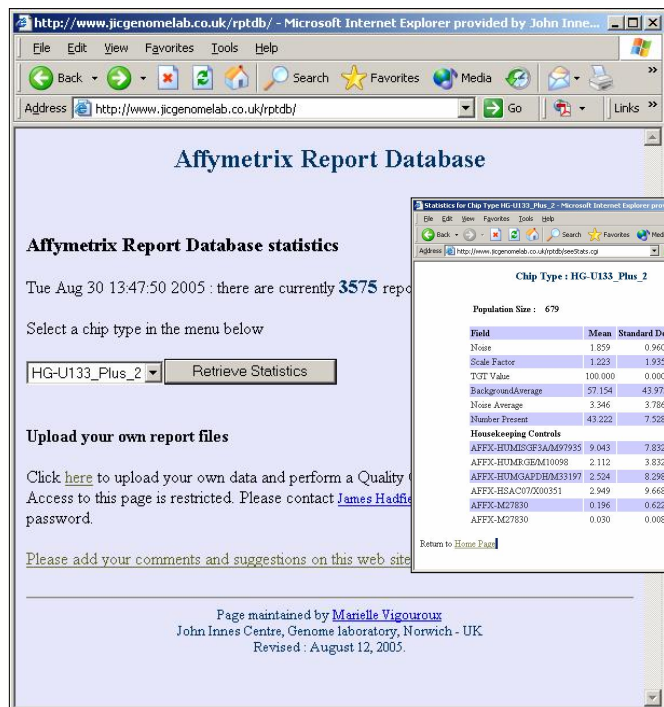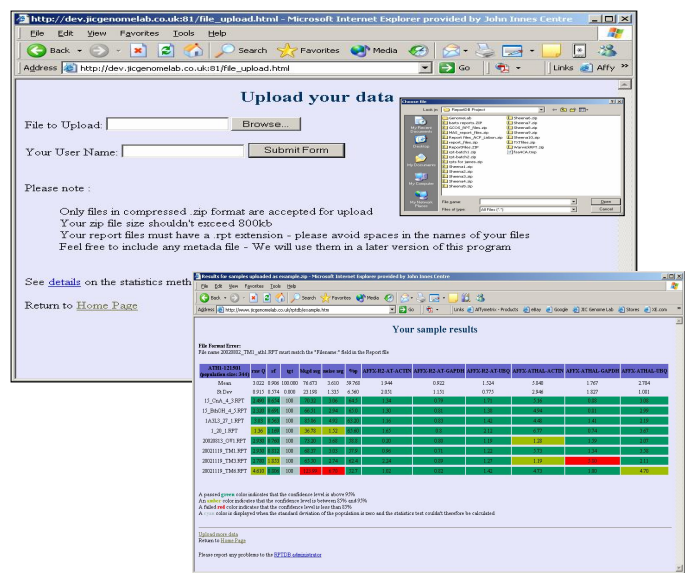
Generate average .rpt file statistics for any tissue type (demo)
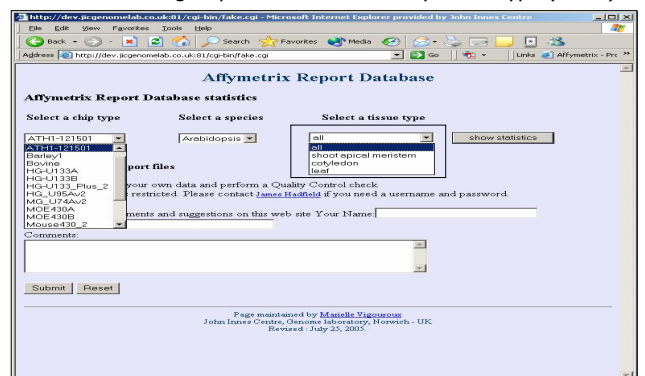
To Do List:

• Weighting .rpt metrics

• Removal of outliers (1.5*IQR or Median Absolute Deviation)

• Generate average .rpt file statistics for tissue type

• Variance of .rpts (how many .rpt do we need to have a 'good' population estimate?)

• MAIME input

• Arrayexpress/GEO data absorption

Development:

• Comparison of uploaded data to database and itself
• Non-expression reports
• What does the community want?

The Genome Laboratory offers a complete Affymetrix Authorized Expression Array Processing service. This service includes RNA sample QC, cDNA and cRNA synthesis and QC, frag, hyb, scan and data QC. **Prices from £580 per sample/chip including GeneChip array!**

We are also developing a two-colour custom microarray service.

Please contact us on +44 (0)1603 450842 via enquiries@jicgenomelab.co.uk or visit our website at www.jicgenomelab.co.uk.