

MeV Version 3.1 Release Notes

8 March 2005

Contents

Description
New Modules
New Features and Feature Enhancements
Existing Analysis Module Enhancements
Repairs
Developer Support Enhancements
Documentation
Contact Information
Acknowledgements

Description

TIGR MultiExperiment Viewer (MeV) is an application that generates informative and interrelated displays of expression and annotation data from single or multiple experiments and facilitates the identification of genes and expression patterns of interest. A variety of clustering analyses and data viewers allow the user flexibility in creating meaningful representations of the expression data.

New Modules

- ❖ **Correspondence Analysis (COA)**
Correspondence analysis is an explorative method to study associations between variables. Like principal components, it displays a low-dimensional projection of the data. However, in this case, both genes and samples can be projected onto the same space, revealing associations between them. (Manual Section 11.20)
- ❖ **Discriminant Analysis Module (DAM)**
This module is used for supervised multi-class prediction of samples or genes. Provided with elements that represent members of the known classes, unknown elements are partitioned. (Manual Section 11.19)

New Features and Feature Enhancements

- ❖ **Gene or Sample Search Utility**
The new feature allows you to track samples or genes throughout your analysis. Provide an ID and you will receive a list of shortcuts to result viewers that contain your gene or sample of interest. (Manual Section 9)
- ❖ **Import Gene or Sample ID Lists**
Supply keys to identify gene or sample lists and produce customized clusters in MeV. Lists of genes generated by an external criterion can be imported to produce clusters in MeV to permit easy color coded tracing of the genes of interest. (Manual Section 9)

- ❖ **Import Sample Annotation, Import Gene Annotation**
This utility allows you to enhance information about the loaded genes or samples by allowing the import of sample or gene annotation from formatted files. (Manual Section 9)

- ❖ **Distance Metric Selection Enhancement**
In addition to a global distance metric menu, all algorithms that make use of distance metrics also have metric selection controls on the initialization dialogs. This quickly identifies the current metric, indicates the default metric, and allows immediate selection for the current algorithm run.

- ❖ **Variance Filter and Enhance Data Filter Output**
For some experimental designs, genes that have relatively constant expression across loaded samples, while important to note, are not typically of great interest. The new variance filter provides three options to select the genes with the greatest variability. For version 3.1 all data filters have a viewer and gene count reported on the result tree for each filter operation. This is in addition to the report in the history log. (Manual Section 5.3)

- ❖ **Data Source Selection/Analysis Branching**
In earlier versions of MeV it has been possible to launch a new Multiple Array Viewer that contains a subset of genes from a result of interest. Now MeV supports the selection of data source from among the result viewers. After data source selection, subsequent analysis only involves the genes and samples in the selected data source. (Manual Section 5.4)

- ❖ **Enhanced Expression Image Options**
MeV now supports using a single color gradient for selecting colors to represent expression values. This can be a natural choice in cases where the input data is single channel or has no reference and the values are not \log_2 ratios. For people using the conventional two gradient scheme (e.g. Green -> Black (neutral) -> Red) the endpoints are not constrained such that the midpoint is selectable to be a value other than 0.0 and the endpoints of the color gradient scale need not be on opposite sides of zero. (Manual Sections 6.3-6.4)

- ❖ **Expression Color Range Selection**
The control of the color scale limits for Expression Images has been modified to report the number and percentage of elements that are off of the color scale (saturated) and permits multiple range updates without dismissing the dialog preview option). (Manual Section 6.4)

- ❖ **Agilent File Loader**
This feature, a contribution from University of Washington, adds the Agilent file format to list of file formats that can be loaded into MeV. (Manual Section 4.6)

- ❖ **Data Format: Text Tab Delimited, Multiple Sample Format (TDMS)**
This format description is consistent with and replaces what has been called the Stanford File Format in earlier versions of MeV. The format contains columns to hold values from multiple hybridizations but with the added advantage to allow multiple header rows to more descriptively annotate the loaded samples. (Manual Section 4.3, format description appendix 14.2)

- ❖ **Menu Reorganization**
The Multiple Array Viewer's main menu has undergone reorganization and has added the utilities menu (Manual Section 9) to support some of the features described above.

Existing Analysis Module Enhancements

- ❖ **T-test Enhancements**
 - Paired T-test option has been added.
 - False Discovery Rate (FDR) can now be used as a significance criterion.
 - T-test output now reports raw p-values as well as adjusted p-values when applying an adjustment (e.g. Bonferroni Adjustment)
 - One class T-test now generates a volcano plot.
- ❖ **One-Way ANOVA Enhancement**
 - False Discovery Rate can now be used as a significance criterion.
- ❖ **SAM Enhancement**
 - The observed and expected D scores are now both reported.
- ❖ **HCL Enhancements**
 - HCL trees can now be output to file in Newick format.
 - The description of average linkage clustering in the manual has been expanded to include a description of the weighting formula.
- ❖ **Statistical Module (TTEST, SAM, ANOVA, 2 Factor ANOVA, PTM) Enhancement**
 - The option now exists to construct hierarchical trees only on significant gene lists. This eliminates the cost of building trees on large lists of non-significant genes.
- ❖ **PCA Enhancements**
 - PCA now produces 2D viewers that correspond to combinations of the first three principal components.
 - 2D and 3D viewers can be generated for any set of selected components.
 - The control of element selection in 3D viewers has been improved.
 - Missing values within the expression matrix entering the analysis are now imputed using the K-nearest neighbors imputation method.
- ❖ **KMC Enhancement**
 - The K-Means mode of KMC has been optimized to permit efficient clustering of large data sets.
- ❖ **EASE Enhancements**
 - GO terms are now viewable from a new viewer that displays GO terms in their natural specificity hierarchy. The viewer allows sections of the hierarchy to be cloned for detailed inspection.
 - The algorithm has been optimized to run in approximately 25% of the previous run time.
- ❖ **ST Enhancements**
 - Support trees can now display support levels directly on the tree structure in addition to the typical color coded support legend.
 - Support trees can now be output to the Newick file format.

Repairs

- T-test, now uses n corresponding to available (non-missing data) for T calculation.
- EASE supports linking from table to viewer following scripting runs.
- Mac now supports implied theme reporting in EASE.
- Saving state with sample clusters fixed.

- QTC now supports construction of HCL trees on sample clusters.
- Fixed KNN imputation bug in SAM.
- Data loading for Linux and Windows XP does not stall in progress.
- Cluster operations (intersections, unions) enabled after analysis re-load.
- PCA 3D correctly maps gene names when using cutoff filters.
- Multiple Array Viewers launched from cluster viewers retain display settings of parent.

Developer Support Enhancements

- ❖ An ANT script has been developed for building MeV. The script supports partial compilation and jarring. The `ant_build_scripts` folder contains the script and a file to describe script modification for module development or feature development.
- ❖ The `build.bat` script has been modified to produce the partitioning of MeV classes into the same jar structure produced by the ANT script. A readme also describes script modification for module or feature development.
- ❖ Developer documentation has two new documents. One document describes the new jar partition scheme (`mev_jars_readme.txt`). Another new addition, `mev_package_overview.doc`, describes the basic MeV packages and the main classes within each package.

Documentation

Please refer to the MeV Manual for information on system requirements, program operations, screenshots and algorithm descriptions and references. The MeV Manual can be found in the *documentation* directory (*documentation/MeV_Manual_3_1.pdf*).

Contact Information

Please direct all questions, comments and bug reports to mev@tigr.org.

Acknowledgements

Please refer to the MeV Manual or the *Contributions* menu option (available on the main MeV menu bar, under the *References* menu) for a list of acknowledgements.