

Chapter 32

GEView (Gene Expression View) Tool for Intuitive and High Accessible Visualization of Expression Data for Non-Programmer Biologists

Libi Hertzberg

Tel Aviv University, Israel & Weizmann Institute of Science, Israel

Assif Yitzhaky

Weizmann Institute of Science, Israel

Metsada Pasmanik-Chor

Tel Aviv University, Israel

ABSTRACT

This article describes how the last decade has been characterized by the production of huge amounts of different types of biological data. Following that, a flood of bioinformatics tools have been published. However, many of these tools are commercial, or require computational skills. In addition, not all tools provide intuitive and highly accessible visualization of the results. The authors have developed GEView (Gene Expression View), which is a free, user-friendly tool harboring several existing algorithms and statistical methods for the analysis of high-throughput gene, microRNA or protein expression data. It can be used to perform basic analysis such as quality control, outlier detection, batch correction and differential expression analysis, through a single intuitive graphical user interface. GEView is unique in its simplicity and highly accessible visualization it provides. Together with its basic and intuitive functionality it allows Bio-Medical scientists with no computational skills to independently analyze and visualize high-throughput data produced in their own labs.

DOI: 10.4018/978-1-7998-1204-3.ch032

INTRODUCTION

Advanced high-throughput technologies have been developed and extensively used in the last decades (microarrays since the mid-1980s (Schena, Shalon, Davis, & Brown, 1995) and Next Generation Sequencing (NGS) (Goodwin, McPherson, & McCombie, 2016) since the mid-2000s). Mass spectrometry technologies were originally invented almost 100 years ago and further developed during the 1990s (Glish & Vachet, 2003). All these technologies and others are being routinely used by Bio-medical (Bio-Med) scientists for production of high-throughput data. However, currently, it is very difficult for Bio-Med researchers with no computational skills to analyze their own data. Various free, online and user-friendly tools have been developed and are routinely used for gene expression and proteomics data analysis. For example, Expander (Ulitsky et al., 2010), Chipster (Kallio et al., 2011), SAM (Tusher, Tibshirani, & Chu, 2001), Limma (Ritchie et al., 2015), DESeq (Love, Huber, & Anders, 2014), Morpheus (<https://software.broadinstitute.org/morpheus/>), MeV (Howe, Sinha, Schlauch, & Quackenbush, 2011) for gene expression analysis, and MaxQuant (Cox & Mann, 2008) for proteomics data analysis. However, these tools usually perform a large range of tasks, which are sometimes not straightforward and may necessitate some computational skills. In addition, they usually lack a single-gene trivial visualization (such as box plots), which is a basic request of lab researchers in order to study the expression pattern of a specific gene of interest. While the Morpheus tool (<https://software.broadinstitute.org/morpheus/>) does provide box plots for single genes, it has the limitation of not providing differential expression statistical analysis of more than two groups of samples. The result is that many Bio-Med researchers are reluctant to use such tools and often outsource for the analysis of their data. Outsourcing creates a gap in time and expenditure of financial resources, in addition to the fact that it is performed by a computational expert, who doesn't necessarily have a biological view of results. Thus, there is a need for a more basic and intuitive tool, with functionalities such as quality control measures, simple statistics, differential expression analysis and intuitive visualization of the results, including single-gene visualizations.

GEView is a free, easily downloadable tool suitable for Windows and Linux platforms, written in MATLAB (<https://www.mathworks.com>). It provides intuitive quality control analysis of the experiment, outlier exclusion and easy identification of differential expression of gene, microRNA or protein data. The great advantage of GEView is its usage simplicity and the output Excel table which provides immediate visualization of the analyzed data on a single-gene basis. It serves all available expression platforms, given that expression values for each gene/microRNA/protein are provided in a tab-delimited file as input, after normalization. GEView is freely available to non-commercial users and can be downloaded from <http://www.weizmann.ac.il/complex/compphys/software/geview/>. It includes a convenient "help" menu and an example test data (Teuffel et al., 2004).

We believe GEView may be highly useful for Bio-Med researchers who are not computational experts. Such researches, who were previously reluctant to use computational tools and were dependent on external bioinformatics services, have now an opportunity to analyze their own data. In addition, the readily available single-gene graphical representation enables laboratory scientists to simply return to the results table each time they become interested in a new gene or protein and explore its expression pattern in their data. Thus, by improving the access to basic expression analysis, GEView improves the ability of laboratory scientists to produce significant results from their various experiments.

11 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/geview-gene-expression-view-tool-for-intuitive-and-high-accessible-visualization-of-expression-data-for-non-programmer-biologists/243134

Related Content

A Machine Learning-Based Exploration of Relationship Between Security Vulnerabilities of IoT Devices and Manufacturers

Ritu Chauhan and Gatha Varma (2020). *International Journal of Data Analytics* (pp. 1-12).

www.irma-international.org/article/a-machine-learning-based-exploration-of-relationship-between-security-vulnerabilities-of-iot-devices-and-manufacturers/258917

Analysis of Heart Disease Using Parallel and Sequential Ensemble Methods With Feature Selection Techniques: Heart Disease Prediction

Dhyan Chandra Yadav and Saurabh Pal (2021). *International Journal of Big Data and Analytics in Healthcare* (pp. 40-56).

www.irma-international.org/article/analysis-of-heart-disease-using-parallel-and-sequential-ensemble-methods-with-feature-selection-techniques/268417

Hadoop Setup

(2019). *Big Data Processing With Hadoop* (pp. 45-62).

www.irma-international.org/chapter/hadoop-setup/216599

Conceptual View on Healthcare Digitalization: An Extended Thematic Analysis

Robert Furda and Michal Gregus (2017). *International Journal of Big Data and Analytics in Healthcare* (pp. 35-54).

www.irma-international.org/article/conceptual-view-on-healthcare-digitalization/197440

Big Data Analytics in Healthcare: Applications and Challenges

Jaimin Navinchandra Undavia and Atul Manubhai Patel (2020). *International Journal of Big Data and Analytics in Healthcare* (pp. 19-27).

www.irma-international.org/article/big-data-analytics-in-healthcare/253843