

Poster presentation

Open Access

## Techniques for the analysis of trichome *Pelagonium xhortorum* expressed sequence tags

Joseph Morris<sup>1</sup>, Eric Rouchka\*<sup>1</sup> and David Schultz<sup>2</sup>

Address: <sup>1</sup>Department of Computer Engineering and Computer Science, University of Louisville, Louisville, KY 40292, USA and <sup>2</sup>Department of Biology, University of Louisville, Louisville, KY 40292, USA

Email: Eric Rouchka\* - [eric.rouchka@louisville.edu](mailto:eric.rouchka@louisville.edu)

\* Corresponding author

from UT-ORNL-KBRIN Bioinformatics Summit 2008  
Cadiz, KY, USA. 28–30 March 2008

Published: 8 July 2008

BMC Bioinformatics 2008, 9(Suppl 7):P25 doi:10.1186/1471-2105-9-S7-P25

This abstract is available from: <http://www.biomedcentral.com/1471-2105/9/S7/P25>

© 2008 Morris et al; licensee BioMed Central Ltd.

### Background

An important aspect of bioinformatics is the infrastructure for storage and display of high throughput sequence data generated by biologists. The construction of databases and applications to display the results in a meaningful and useful way is not a trivial task. A well-designed and executed system can vastly improve the results obtained from an experiment. We present a system for storing and viewing EST sequences based on a set generated for a geranium genome.

### Results

A series of over 4,000 expressed-sequence tags (ESTs) originating from the trichome of the geranium, *Pelagonium xhortorum*, were generated in a wet lab. EST sequences were organized into clusters of overlapping sequences using CAP3. The resulting clusters, including single EST clusters, were analyzed and annotated by hand in order to classify the sequences into families and subfamilies of genes. Further refinement using Blast2GO classifies these sequences into annotated groups based on biological process, cellular component, and molecular function. Presentation of the results were accomplished through a MySQL database and java servlet web pages used to analyze, group, and display the resulting data. Queries into the data can be performed by using blast or by searching by known gene families.

### Acknowledgements

This work was supported in part by NIH-NCRR Grant P20RR16481 and NIH-NIEHS Grant P30ES014443. Its contents are solely the responsibility of the authors and do not represent the official views of NCRR, NIEHS, or NIH.