

## Visualization-Driven Analytical System for Comparative Microbial Genomics

Aurora Cain, Adam Price, Shatavia S. Morrison, Robert Kosara, Cynthia Gibas

Department of Bioinformatics and Genomics, University of North Carolina at Charlotte, NC, USA

Corresponding author:

Aurora Cain [aacain@uncc.edu](mailto:aacain@uncc.edu)

GenoSets is a visualization-driven support system for the analysis of multiple microbial genomes within a comparative context. The application combines data from several disparate data sources and bioinformatic tools unifying the data in a single analytical space. The prototype version of GenoSets currently supports data from several common annotative or comparative operations including assembly, annotation, functional categorization of genes, and homology detection. This brings together layers of information that can be explored simultaneously using the multiple charts, tables, and other multi-dimensional visualizations. Key features and relationships can be discovered using the interactive and connected visualizations. This enables the user to select a subset of data points in one visualization to highlight related elements in alternative visualizations. The visualizations support the aggregation of data into groups or sets based on user-defined properties. The user may define sets of genes based on annotation source, presence or absence of genes in groups of related organisms, or Gene Ontology (GO) functional classifications. To demonstrate the systems functionality, we present a case study for the comparative analysis of 9 *Vibrio* strains. In this study, we show how homology can be used to define differential genes that are selected at various levels of the taxonomic hierarchy. Once the differential genes have been defined, we show how the supportive visualizations can be used to interrogate the functional potential of these genes.