

Visual Analytics for Accelerating Discoveries in Translational Science

Suresh K. Bhavnani¹ PhD, Vickie McMicken², Rohit Divekar³, MD PhD

¹Inst. for Translational Sciences, Univ. of Texas Medical Branch, Galveston, TX; ²Management and Information Science, UHCL, Houston, TX; ³Div. of Allergy, Dept. of Med. Univ. of Texas Medical Branch, Galveston, TX

Abstract

Translational science increasingly involves the discovery of complex relationships among different granularities of information such as at the molecular, phenotype, and environmental levels. However, few visual analytical systems enable translational researchers to simultaneously represent, and intuitively interact with such complex data. Here we demonstrate a novel system that was designed and implemented by integrating (1) user needs for representing and interacting with complex biomedical data, (2) design heuristics from usability engineering and visual analytics, and (3) agile programming using HTML5 and scalar vector graphics (SVG). A formative evaluation of the system's functionality in two disease datasets demonstrated the efficacy of the system for discovering patterns in data, but required addressing design and performance limitations.

Introduction

Translational scientists such as biologists and clinicians need to discover relationships among different granularities of information such as patients and their molecular and phenotype information. Such analysis requires systems that (1) simultaneously represent molecular and phenotype information, and (2) enable intuitive interaction with the data to enable rapid discoveries. Unfortunately, current visual analytical systems have important limitations including constraints in the range of data that can be simultaneously represented (e.g., networks cannot simultaneously represent more than a few variables through the color and shape of nodes¹), and the lack of interactivity (e.g., the Circos system allows only a batch process to generate otherwise useful representations of complex data²).

Design

To address the above limitations, we developed a prototypical visual analytical system called *Interactive Circos* (iCircos) using a three-step design and evaluation methodology³: (1) Identified user needs by analyzing data using existing visual analytical systems such as networks and Circos ideograms; (2) Applied design heuristics from usability engineering and visual analytics to guide the selection of useful functionality, and the design of a usable interface; (3) Conducted a formative evaluation of the functionality to explore whether the interface and functionality could lead to discoveries in two domains.

Figure 1 shows the resulting interface for iCircos, which was developed using HTML5 and the SVG library Raphael. As shown, the iCircos enables a translational scientist to generate and interact with a *Circos ideogram*. This representation has at its center, a network consisting of nodes such as patient and cytokines, and edges that connect pairs of nodes representing relationships such as cytokine expression. The concentric circles around the patient nodes use a heatmap to represent phenotype variables such as lung function. Heeding usability guidelines, the interface has a minimalist design which enables a clear separation between the input panels on the left, and the visualization panels on the right. Users can click on the *demo button* (1) to understand the format of data input, make selections for the various visualization elements such as node color and edge thickness (2), directly interact with the visualization by, for example, double-clicking on nodes and edges to highlight them (3), and indirectly interact with the visualization through text boxes such as for zoom (4).

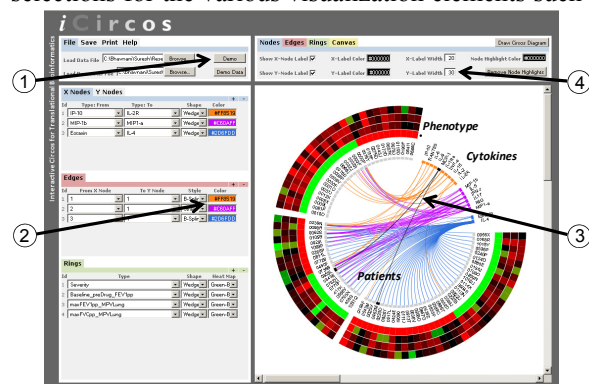


Figure 1. The iCircos interface enables the interactive generation and modification of a Circos ideogram which can simultaneously represent patients and their molecular information in a network in the center, and phenotype information of each patient in outer concentric rings.

A formative evaluation of the system was conducted by using iCircos to analyze two datasets. The results enabled the discovery of novel patterns in the data³, but revealed the need for mouse-over annotation, and better performance when interacting with the concentric rings.

Acknowledgements: CDC/NIOSH # R21OH009441-01A2.

References

- Bhavnani, S.K., et al. How Circos Ideograms Complement Networks. *Proceedings of AMIA TBI* (2011).
- Krzywinski, M., et al. Circos: an Information Aesthetic for Comparative Genomes. *Genome Res* (2009) 19:1639-1645.
- Bhavnani, S.K., et al., iCircos: Visual Analytics for Translational Bioinformatics. *Proceedings of IHI'2012* (2012).