How Complimentary Bipartite Visual Analytical Representations Reveal Patterns in Cytokine Expression: A Case Study in Rickettsioses

Suresh K. Bhavnani[1,2] **PhD, Bryant Dang**[1] **BS, Justin Drake**[1] **BS, Gowtham Bellela**[2] **PhD, Juan P. Olano**[3] **MD**
[1]Inst. for Translational Sciences, [3]Dept. of Pathology, UTMB; [2]Hewlett Packard Laboratories, Palo Alto, CA

Abstract

Although network visualizations use node proximity (e.g., through node clusters) to enable comprehension of node similarity, the resulting layouts often have many intersecting edges which can conceal important patterns among edges. We therefore explored the role of Circos ideograms (a node-edge representation that does not use proximity to show node similarity) for comprehending patterns related to edges in a bipartite network of patients with rickettsioses, and their cytokine expressions. The results suggest that while Circos ideograms are not designed to discover node clusters, they make salient how edge weight distributions differ among clusters identified through networks, and therefore elucidate how ideograms can complement network analysis.

Introduction

Networks have helped to reveal complex relationships in numerous biomedical datasets. For example, in a recent study[1] we used a bipartite network where nodes represented rickettsioses patients and candidate cytokines, and weighted edges connecting pairs of patients and cytokines represented normalized cytokine expression. The network layout revealed a small cluster of patients (Patient Cluster-1) with high cytokine expression of a small subset of cytokines (Cytokine Cluster-1), and a larger cluster of patients (Patient Cluster-2) with overall low cytokine expression of all cytokines (Cytokine Cluster-1 and Cytokine Cluster-2). However, due to the complex intersection of edges connecting the two patient and two cytokine clusters, there was less clarity related to patterns in the edges. Given this limitation in the network layout, we posed the question: *Can Circos Ideograms help to identify patterns related to cytokine expression across clusters of patients?*

Method

As shown in Figure 1, we generated Circos ideograms[2] where: (1) 85 patients and 26 cytokine nodes were laid out equidistant around a circle (hence distance between nodes had no meaning), organized in clusters identified through network analysis[1]; (2) links between the nodes represented cytokine expression, and their color and saturation represented cytokine type and expression value, respectively. For close inspection, the edge weights representing the cytokine expression were drawn for either Patient Cluster-1 (Figure 1A), or for Patient Custer-2 (Figure 1B).

Results and Conclusion

Two physicians inspected the ideograms, and identified a new pattern related to the edges that was not salient in the original network layout[1]. The variance of blue and red edges appeared higher for Cytokine Cluster-1 (red nodes) compared to Cytokine Cluster-2 (blue nodes) for both Patient Cluster-1 (purple nodes), and for Patient Cluster-2 (yellow nodes). Both differences in variances were highly significant using the Brown-Forsythe test (Patient Cluster-1 W = 9, p < .003) and (Patient Cluster-2 W=248, p < .00001, two-tailed), although the former appeared to be caused by a single patient (considered to be biologically feasible based on the cytokine profile). The results pinpointed the difference in expression profile between the two cytokine clusters suggesting that Cytokine Cluster-1 could contain a more reliable set of biomarkers for predicting the particular form of rickettsioses examined in this dataset. The study therefore suggested that because the Circos ideogram trade-offs node proximity with reduced edge crossings, it enables domain experts to quickly comprehend edge distributions between node clusters, leading to the rapid generation of hypotheses. Future research should explore similar trade-offs between different visual analytical representations to enable a more systematic understanding of how they can be used together for rapidly enabling translational insights.

References

1. Bhavnani S.K., et al. How Cytokines Co-occur across Rickettsioses Patients: From Bipartite Visual Analytics to Mechanistic Inferences of a Cytokine Storm. *AMIA Summit on Translational Bioinformatics* (in press).

2. Krzywinski, M. et al. Circos: an Information Aesthetic for Comparative Genomes. *Genome Res* (2009) 19:1639-1645.
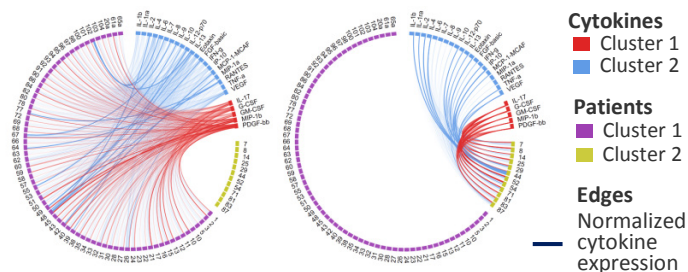
A. Patient Cluster-1 Profile   **B.** Patient Cluster-2 Profile

Figure 1. The Circos ideogram complements network analysis by revealing differences in the variance of edge weights (representing normalized cytokine expression) between patient and cytokine clusters.