ImmunoExplorer: A Web-based Multivariate Visualization System for Exploratory Analysis of Immunotherapy

Mai Elshehaly
Computer Science and Electrical Engineering
University of Maryland, Baltimore County
Baltimore, MD, USA
mai@umbc.edu

Gregory Szeto
Chemical, Biochemical and Environmental Engineering
University of Maryland, Baltimore County
Baltimore, MD, USA
gszeto@umbc.edu

Zhigeng Pan
School of Information Science and Engineering
Hangzhou Normal University
Hangzhou, China
zgpan@hznu.edu.cn

Jian Chen
Computer Science and Electrical Engineering
University of Maryland, Baltimore County
Baltimore, MD, USA
jichen@umbc.edu

***Abstract*—We present ImmunoExplorer, a web-based multivariate visualization environment that supports exploratory analysis of experimental datasets typical in immunotherapy research. Research advances in immuno-oncology have opened up new frontiers for experimental and clinical research that aims to understand the complex interactions between cancer and the immune systems. Immunotherapy focuses on the development of novel cancer therapies based on leveraging these interactions. Extensive analysis of experimental datasets is required to rigorously validate preclinical results and clinical outcomes to design new therapies and to identify who will benefit from them. Visualization is a crucial part of this analysis, given the complexity, multidimensionality, and heterogeneity of the data involved and the lack of automatic computational solutions to derive patterns. In this work we first characterize the data types and analysis tasks, and then present two visualization techniques for comparative studies of various therapies. Finally, we demonstrate the usefulness of the design through a case study using a real dataset.**

***Keywords*-Data visualization, multivariate analysis, design study, tumor therapy**

I. Introduction

The field of immuno-oncology aims to develop therapies that harness the immune system to provide enduring and adaptable cancer control. Combination immunotherapy approaches have been studied extensively in animal models to help scientists understand complex networks in the immune system, aiming to identify factors contributing to the progression or rejection of tumors [1]. The goal of the therapy-outcome relationship study is to identify the analytes (biological pathways, co-regulated proteins) and the multivariate relationships among them that yield more effective therapies or therapy failure.

In doing this, extensive analysis of experimental data is crucial to support hypothesis generation and testing of the complex phenomena that govern this interplay. Computational methods alone are not enough or do not exist for this kind of analysis, due to high dimensionality and heterogeneity of the data involved in the analysis. Hence visualization methods are crucial to leverage the experts? domain knowledge and to give them a more complete understanding of the data and the underlying phenomena.

Specifically, scientists seek to establish correlations among clusters of proteins working together as co-regulated entities that may contribute to the progression or regression of disease. The effect of these co-regulated proteins on disease outcomes, such as tumor size and survival rates is studied with different experimental parameters. For example, different combination therapeutic components are mixed in different ways and administered to the animals. The scientists then seek to conduct multivariate regression analysis to establish the correlation among the different therapeutic alternatives, the measured analytes, and the observed outcomes.

This work focuses on a web-based visualization environment, ImmunoExplorer (Figure 1), a multiview interactive exploration environment for ranking and comparison studies of immunotherapy data. The first contribution of this paper is a characterization of the data used in immunotherapy research. The second contribution is a characterization of the analysis tasks of interest in this research domain into three groups: association and differentiation pattern tasks, distribution pattern tasks, and connectional pattern tasks. This design of two main visual encoding methods constitutes our third contribution. The two methods are parallel coordinates and a sunburst view of therapy measurements. The goal is to address two two major challenges scientists face while conducting data analysis. To demonstrate how the system works and addresses scientists' tasks, we present a

978-1-5090-5188-5/16 $31.00 © 2016 IEEE
DOI 10.1109/ICVRV.2016.87

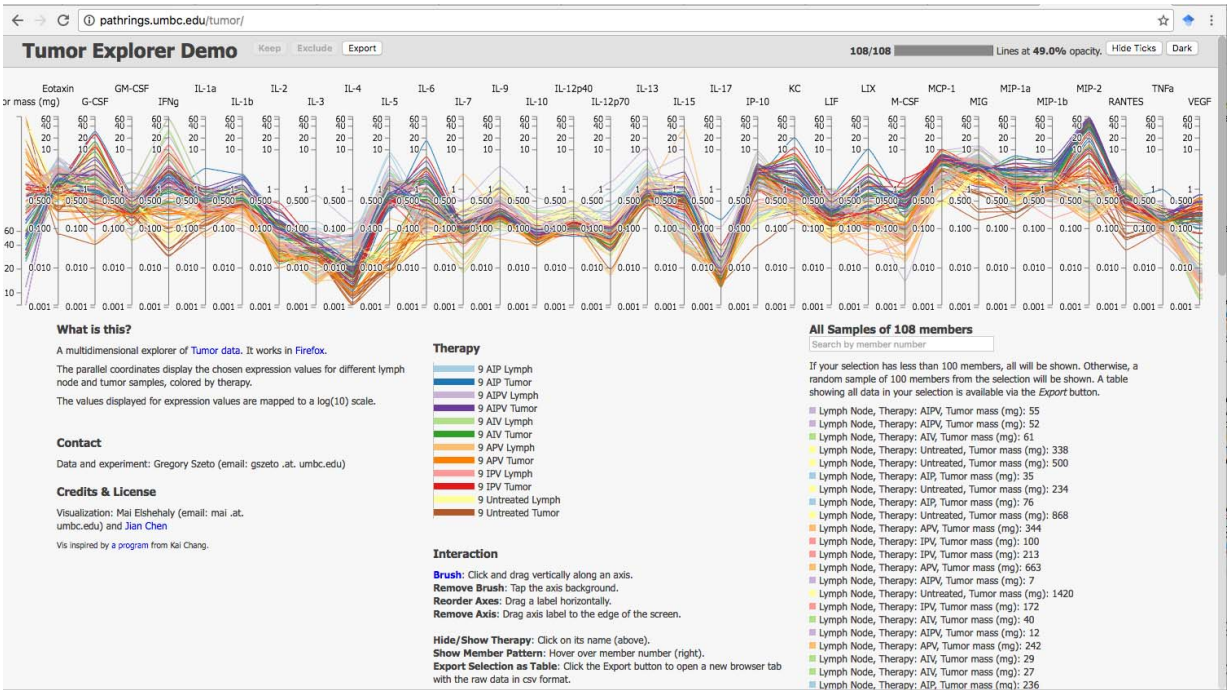

Figure 1: ImmunoExplorer: A web-based multi-view interface. From top to bottom and left to right: (1) parallel coordinates showing nine therapies measurements (those names listed in the axes, e.g., IL-3) and related outcomes (tumor size and survival), (2) links to the source data and help information, (3) interactive filtering by therapy names, and (4) interactive filtering by each measurements in all therapies. The program is accessible online at http://pathrings.umbc.edu/tumor.

case study using a real experimental dataset that describes how a domain expert performe analysis using the techniques supported in ImmunoExplorer.

II. Background and Related Work

A. Background and Data Used in Our Study

We worked with a real experimental dataset to measure the outcome from nine different therapies containing 110 samples from a diseased (tumor) and a normal control groups. The data contains three independent variables that are categorical: sample, therapy, and organ. Dependent variables observed through the experiment include 32 quantitative analytes that constitute *protein levels*, *tumor size* and *survival* (whether or not the mice die).

The structure of immunotherapy experimental data can be broken down into two main blocks. The first, which is the analytes block, contains measurements taken simultaneously for a number of cytokines at different stages of treatment. Cytokines are proteins released by both immune and cancer cells into either the tumor or lymph node microenvironment. They recruit and reprogram other types of cells to condition each microenvironment and are key factors in modulating the immune response either against or favoring the tumor. The interaction of cytokines, growth factors, and cells creates a network largely responsible for the overall progression or rejection of the tumor [2].

The second data block is the outcomes block. In an experimental setting, tumor cells are injected in a population of animals (here mice) and allowed to establish themselves for a few days before therapy is begun. Tumor size and other biomarkers are recorded at different times before and during the treatment administration period. These outcomes are at the heart of the analysis process. An overarching analysis question is whether outcomes can be predicted by the interaction of analytes, and whether a model can be constructed for this prediction. Scientists usually seek to collect as many outcomes as possible to learn about the different effects of the applied treatment.

B. Multidimensional Data Visualization

Several application domains have benefited from multidimensional data visualizations. Work in clinical cohort analysis, for instance, relies on techniques like contingency arrays [3] and parallel coordinates [4] [5] to explore interactions among multiple parameters. In biology, parallel coordinates [6] and circular glyphs [7] are commonly used. In addition to techniques addressing their high-dimensional nature, visual analysis of multivariate data over time has used creative glyph-based [8], [9]. We explore the design space and alternative visual encodings for multivariate visualization techniques in Section V-A. Our design choices combine strengths from existing approaches and

develop new ones to offer a web-based environment for effective exploration and analysis of experimental data in immunotherapy research. A data and task characterization have guided our design process.

C. Parallel Coordinates

Parallel coordinates are a widely adopted technique for the visualization of multidimensional data. First appeared in the literature in the context of nomography, they have been used to provide overview and support data summarization tasks [6]. They offer an effective visual encoding for the identification of sets of data items exhibiting similar characteristics, thus supporting visual clustering and pattern association tasks [10]. In addition to association, parallel coordinates make it possible to differentiate divergent behavior and outliers through density estimation of raw data [11]. Furthermore, regression tasks aiming to predict the values of dependent variables with respect to one or more independent variables are supported by parallel coordinates, in the form of visual regression [12], or through visualizing statistical properties of regression models [13].

III. Task Characterization

Our first contribution in this paper is a characterization of analysis tasks in immunotherapy research. We are interested in answering the high-level question: What is the most effective treatment combination with maximal antitumor efficacy? This high-level question translates into a number of lower-level analysis tasks that scientists perform to gain comprehension and insight on analytes of interest and their effect on outcomes. Specifically, scientists look for changes within tumors (e.g., in intratumoral cytokines and chemokines) that distinguish treatments that induce sustained tumor regression.

In the nutshell, scientists have two analysis challenges. The first is to detect interactions among a multitude of analytes in order to effectively identify clusters of co-regulated proteins. This cluster identification requires understanding the interactions among a relatively large number of parameters. Considering the effect of one analyte at a time would yield inaccurate conclusions about the outcomes because a one-factor-at-a-time approach cannot capture the complex connectivity in the biological processes involved. For example, when the scientist looks at different analytes, s/he can tell whether a number of proteins are highly expressed and correlate with increased tumor size. A desirable scenario can occur when the cluster of highly expressed genes have one common entity in the upstream network in the pathway diagram. The scientist can then conclude that this entity is a regulator that may be targeted for inhibition to decrease tumor size.

The second challenge is to establish cause-effect relationships between the detected clusters of analytes and disease outcomes. The goal here is to identify clusters that produce the greatest tumor size reduction, thereby yielding highest treatment efficacy. This challenge is complicated by the multidimensional nature of both analytes and outcomes, which renders computational methods insufficient for the analysis. Furthermore, these cause-effect relationships must be compared across different treatment groups and different microenvironments.

To answer questions about these types of patterns, scientists go through the following analysis stages: (1) overview the raw data to gain trust in their results before conducting further analysis, (2) build a qualitative understanding of data clusters and pattern distributions, (3) build an understanding of the interplay and probable cause-effect relationships among analytes and outcomes in the data, and (4) use domain knowledge and integrate external data sources to achieve comprehension and construct predictive models.

Based on these two analysis challenges, we have identified a set of visualization tasks summarized in Table I. These tasks involve different pieces of data used by the domain experts to generate and test hypotheses. Each task can be viewed as a function where the inputs and outputs are individual or a group of data components. To optimize experts? use of information, appropriate visual encodings must be provided for each type of pattern that is sought in the data.

Tasks T1 to T3 are the principal building blocks of the analysis. The main referential components in these tasks are the samples collected in an experimental or clinical setting, which are taken from multiple microenvironments from different animals or patients, at different time steps. In addition, each sample has a number of characteristics (dependent variables) that constitute the observed analytes and outcomes. Each sample is therefore regarded as a multivariate data vector. These vectors act as inputs to the analysis tasks. The output of these tasks takes the form of association or differentiation patterns. An association pattern unifies references into a whole that can be handled together (e.g., a cluster). This association is typically based on identical or close characteristics. A differentiation pattern, on the other hand, distinguishes those references that remain when other references have been united in an association pattern [14].

In contrast to these sample-level association and differentiation patterns, distribution patterns represent aggregate information about the data, such as extrema and summary statistics. These patterns are essential components of the target cognitive model that the visualization aims to construct in the expert's brain to achieve comprehension. Tasks T4 to T6 fall in this task category. Referential components involved in these distribution patterns are typically groupings of individual references (i.e., reference set). In immunotherapy data, these groupings are typically based on the sample site (microenvironment) and the type of therapy.

Finally, connectional patterns seek to establish cause-

Table I: Pattern Search Task Types and Example Tasks

Task type	Index	Task
Association	1	Identify clusters of co-regulated cytokines
	2	Identify co-regulated clusters that are inversely correlated with tumor size
Differentiation	3	Detect convergent, divergent, and outlier behavior in clinical or experimental samples
Distribution	4	Compare cytokine signatures across different microenvironments
	5	Compare cytokine clusters across different therapies
	6	Compare disease outcomes across different therapies
connectional pattern	7	Identify cytokine signatures that can predict efficacy of each therapy
	8	Identify tumor compartments that offer a predictive model of therapeutic efficacy
	9	Characterize temporal behavior of tumor progression/regression within/across compartments
	10	Identify upstream/downstream regulators that can be targeted for knock-down
	11	Identify treatment combinations that elicit durable cures and sustained resistance mechanisms

effect relationships at a higher level of cognition. These tasks require a high-level understanding of the data and the underlying phenomena and rely on the user's expertise and integration of multiple data sources. The target connective patterns constitute the final outcomes of the analysis. Tasks T 7 to T 11 belong to this group. Again, these high-level tasks do not typically deal with individual data references. Rather, aggregate information, clusters, and knowledge extracted from the other two task categories act as input to these connective tasks.

IV. Design Aims

In order to support these analysis tasks, we have identified the following encoding aims:

- Aim 1. Support the detection of association and differentiation patterns through data overview;
- Aim 2. Support the comprehension of distribution patterns through data aggregates;
- Aim 3. Support the visual separation of the two main data blocks involved in the analysis: analytes and outcomes; And
- Aim 4. Provide interaction capabilities that leverage the user's domain knowledge to support construction of more complete comprehension that consequently leads to the identification of connective patterns.

V. ImmunoExplorer

We have designed and implemented ImmunoExplorer, a web-based visualization system that leverages the powers of multiple coordinated views, brushing and linking, and interactive multivariate data visualization techniques to support the analysis of experimental immunotherapy data.

A. Visual Encoding

To address the aims listed in Section III, we designed two visual encoding approaches: parallel coordinates and Sunburst chart. Parallel coordinates aim to visually encode existing raw data in their original dimensions. ImmunoExplorer offers a simultaneous overview of multiple variables and their underlying patterns (Aim 1). Color is used to encode different groups of samples. In Figure 1, the top row where the parallel coordinates encoding is shown, samples are grouped by therapies and microenvironments from which they were obtained. By overlaying these approaches in a parallel coordinates display, scientists can discern patterns involving multiple analytes while also observing the effect of these analytes on a multitude of outcomes. We note here that in order for experts to trust a subsequent analysis, the first thing they look for is an informative data overview. This parallel coordinates help create a overview and establish a level of trust in the data and what to expect from it.

The legend in Figure 1 (bottom row middle column) lists the different sample categories and displays their distribution in the elements currently shown in the parallel coordinates display. Further, the legend acts as a filtering tool from which the user can toggle on and off specific categories. The user can select the categories of interest by selecting the names on the legend to turn on and off the visual display of the data in the parallel coordinates (Figure 2).

Contrast to *association* and *differentiation* patterns that require an overview of raw unprocessed data, *distribution* patterns are detected when aggregates and frequency measures are computed from the data. Among the most common visual encodings that serve distribution patterns are small multiples (of pie charts and bar charts). To support visual separation between the two main data blocks involved in the analysis (analytes and outcomes), we designed and implemented a sunburst chart in a multiview interface (Aim 3). Two such views are shown in Figure 3. An array of sunburst glyphs is created in which each glyph represents one record (i.e., sample) in the data. Bar lengths in the sunbursts are proportional to the values observed for the corresponding proteins' regulation (i.e., analytes). To enable comprehension of the different bars in each suburst and distinguish the analytes they correspond to, labels are provided only on the first element of the glyph array. Figure 4 shows a single view of the first glyph element and accompanying labels.

Given such multiples, it is intuitive for scientists to decide on the distribution of raw elements in the entire dataset as well as the enrichment of additional information in selected subsets. In ImmunoExplorer, a bar chart is linked to the main parallel coordinates overview, and provides distribution information about the entire dataset (Aim 2). As the expert

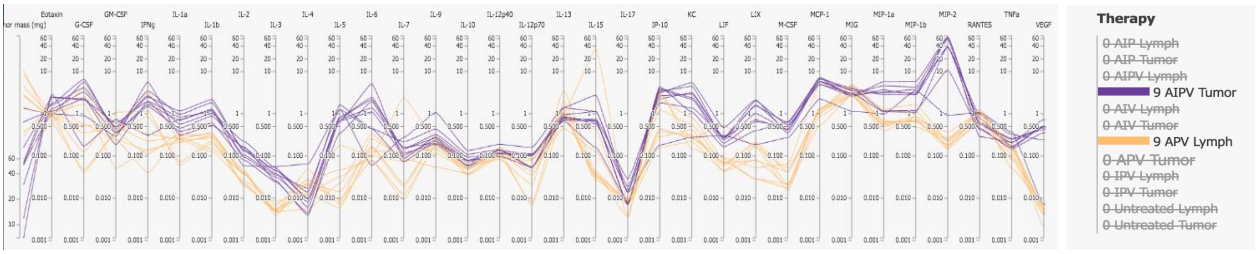

Figure 2: Parallel coordinates view showing filtered samples taken from two different microenvironments (tumor and lymph node) with two therapeutic combinations (AIPV and APV).

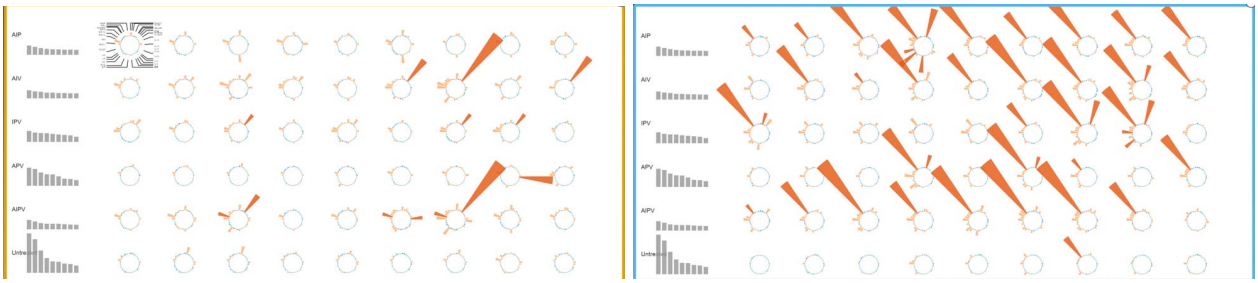

Figure 3: Sunburst array bubbles for lymph node samples (left) and tumor samples (right). Each glyph represents a sample (a record) in the data. The vertical bar charts on the left of each bubble summarize tumor sizes (as outcomes) of every row in the array.

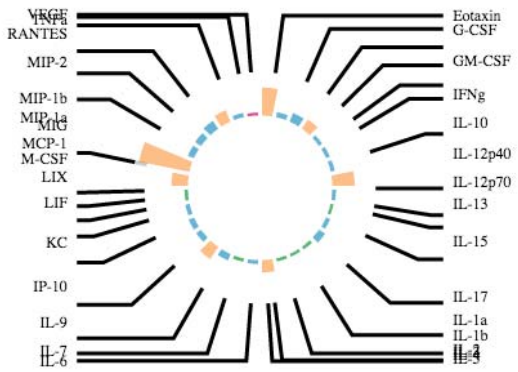

Figure 4: The sunburst visualization of multivariate attributes. The height of each burst is mapped to the scale of that analytes measurement.

explores different subsets and/or clusters of data, a brush is applied and linked to the bar chart to depict the enrichment of specific characteristics in the brushed references.

There are two main advantages to this visualization approach: (1) separable visual encodings between analytes and outcomes in the sunburst view make possible an overview of cause-effect relationships, and (2) separable visual encodings (glyphs) of individual data records make possible clutter-free exploration of these individual entities.

The interaction capabilities in ImmunoExplorer are one of its major strengths. The ability to reorder axes of parallel coordinates is particularly useful in leveraging domain expertise. For example, if the expert knows from theory that a group of proteins are co-regulated, s/he can place the axes for these proteins next to each other to generate a less cluttered, more consistent overview. Further, brushing and linking capabilities strongly support the expert?s understanding of the interactions among data components (Aim 4).To further support *association* and *differentiation* tasks, data can be sorted in the glyph array to reveal patterns.

B. Implementation

The ImmunoExplorer prototype is implemented in Data-Driven Documents (D3) [15]. The parallel coordinates view interface is derived from an online tool [16]. The sunburst view is implemented in a multiview bubbles interface that lets scientists construct incremental views [17].

VI. Case Study

We demonstrate the capabilities of ImmunoExplorer in supporting the analysis tasks described in Section 4 using the real experimental dataset that was generated and described in Section 3. The data was generated by injecting tumor cells in a population of mice and leaving them to establish for 8 days before initiation of therapy. Initial tumor sizes captured at that point in time fell in the range 40-60 mm2. Four different

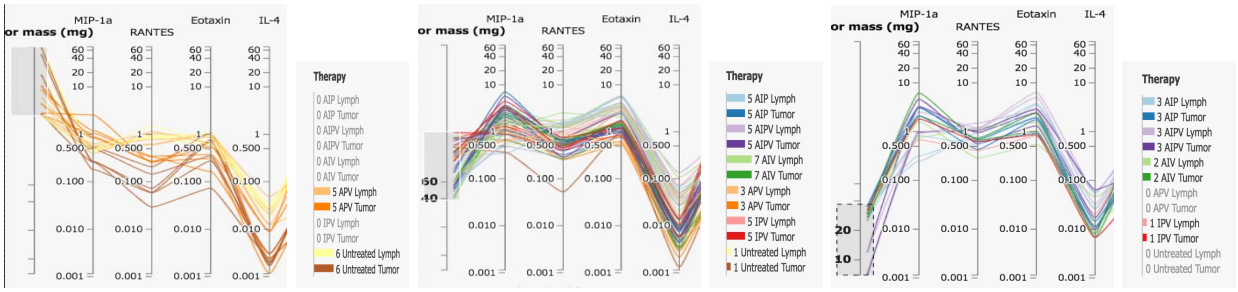

Figure 5: Parallel coordinates view showing three different tumor size ranges selected with a brush. The regulation of MIP-1a, RANTES, eotaxin, and IL-4 are inversely correlated with tumor size and can therefore be predictive of therapeutic efficacy. Legends to the right of every brushed range depict the enrichment of the selected samples in different therapies.

therapeutic components were combined and administered: A (anti-tumor antibody), I (MSA-IL-2), P (anti-PD-1), and V (amphiphile vaccine). Luminex quantification of intratumoral cytokines and chemokines was used to collect the data.

We began by inspecting the data overview provided by the sunburst array and parallel coordinates. We noticed in the sunburst array that two cytokines have consistently larger spikes in many samples. Using this information, we moved the axes of these cytokines next to one another in the parallel coordinates view. We were able to see that the data traces form an almost horizontal pattern between these axes. We concluded that these cytokines are co-regulated.

Next, we began filtering and brushing the data to inspect different patterns that emerged as different tumor size ranges were selected. Again, forming an initial guess was served by the sunburst array view, which sorts samples as rows in the array (see Figure 3). In this case study, each row in the glyph array consists of samples from a specific therapeutic combination. Within each row, the samples are sorted in descending order of tumor size, as can be seen in the vertical bar charts at the beginning of each row. This view made it easy for us to look at specific glyph columns to see which cytokines have consistently large spikes in the last column. This means that when these cytokines are upregulated, the tumor size is consistently smaller. We noticed such an inverse correlation between tumor size and the following cytokines: MIP-1a, RANTES, eotaxin, and IL-4. We suspected that these biomarkers are predictive of treatment efficacy. We moved the axes of these analytes next to the tumor size axis in parallel coordinates and created a brush on tumor size. As we moved the brush up and down the tumor size axis, we observed the behavior of these cytokines, as shown in Figure 5.

Next, we looked for convergent and divergent cytokine behavior in parallel coordinates, since data samples are intuitively displayed as trajectories in this view. However, detecting outlier behavior was not as intuitive as we had anticipated in parallel coordinates, since outliers constituted a relatively small number of samples. These few samples of interest can be missed due to clutter. We decided that the sunburst view can serve this purpose better, since it provides a clutter-free view of the raw data. Using this view, we immediately spotted the outliers in the lymph node samples (Figure 3 (left)) under AIV and AIPV therapies. Interestingly, we noticed that these outliers belong to the same cytokine and are significantly larger than the rest of the samples within similar tumor sizes. We then checked the legends on the first glyph of the array to learn that these spikes belong to the IFNg analyte. We took this observation to the parallel coordinates view for further investigation, brushed these two samples on the IFNg axis and observed their interaction with other analytes. Figure 6 shows this interaction.

In addition, we compared samples belonging to two specific therapy/microenvironment combinations: lymph-node samples from subjects treated with APV and tumor samples taken from subjects treated with AIPV. From this comparison, we determined cytokine signatures that distinguish the two clusters, and thus can be used in predictive models (Figure 2).

VII. Discussion

One limitation of parallel coordinates is the lack of visual distinction between the axes corresponding to the analytes block and those in the outcomes block. This distinction may be crucial for understanding cause-effect relationships in experimental settings where the number of outcomes is relatively large. For analysis scenarios involving only one or two outcomes, however, the user can drag and move the outcome axes and observe their interaction with the different analytes.

Another known limitation of parallel coordinates is clutter. Despite their strength in revealing clusters of convergent and divergent behaviors, thereby supporting tasks targeting association and differentiation patterns and satisfying Aim 1, the study of individual data elements or groups of elements (e.g., outliers) can be cumbersome, especially as the number

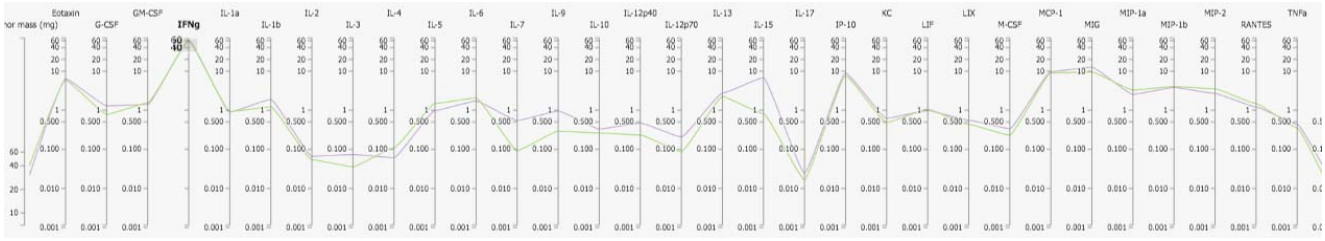

Figure 6: Parallel coordinates view showing a brush selecting the 2 samples observed as outliers for the IFNg analyte with AIV and AIPV therapies in the sunburst view shown in Figure3(left).

of records increases. The sunburst view of ImmunoExplorer complements the capabilities of the parallel coordinates view to accomplish this type of analysis.

The visual distinction between analytes and outcomes is one of the main advantages of the sunburst array view. This separation sets the stage for connectional pattern tasks to be performed based on observations made by the expert about the interplay of these two major data blocks. As the number of outcomes involved in the analysis increases, however, novel visual encoding alternatives are needed to establish these complex multivariate correlations.

Integration of additional data sources is crucial for this type of analysis. For example, the scientist may want to seed a pathway diagram, selected from an online curated database, around specific co-regulated genes of interest to explore the upstream network and detect potential knock-down genes. For example, the highly expressed genes can have one common entity in the upstream network in the pathway diagram: this entity is a regulator that can be targeted and knocked down or inhibited. Such inhibition can lead to down-regulation of the genes of interest. The scientist will also want to find out what the literature says about the regulatory inference layers in this network.

Another useful comparison is that performed over time. A goal here is to tell medical specialists where and when to perform a biopsy for a tumor. A low-level question can be: when are tumors and lymph nodes similar? When are they different? Further, temporal analysis leads to the identification of regulatory patterns (what genes are up/down at the same time). Some analytes may have very low baseline but fluctuate significantly over time. Such fluctuations are very important and ‘can lead to conclusions about profiles and predictive capability of tumor/lymph-node microenvironments. This type of temporal analysis is the focus of our future work.

VIII. Conclusion

Immunotherapy research is an exciting new area for visualization, and the hope of inventing new therapies for cancer by harnessing the immune system make this area of research especially rewarding. Our discussion with our collaborators suggested (1) the web-based system lets the community analyze their data in a common environment, (2) the parallel coordinates view and filtering can support global pattern discovery to observe trend and finding extreme values, and (3) the sunbursts view is a local technique which has limited use in global pattern search but can support inter-therapy comparisons of the quantitative measurements. In the future, we will design new methods to reduce the visual clutter of the parallel coordinates to reveal both global and local patterns.

Acknowledgment

This work was supported in part by NSF DBI-1260795, DBI-1147029, and EPS-0903234. The authors thank Katrina Avery for her editorial support. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

[1] A. Drake, N. S. Joshi, G. L. Szeto, E. F. Zhu, H. N. Eisen, and D. J. Irvine, “Koch institute symposium on cancer immunology and immunotherapy,” *Cancer Immunology Research*, vol. 1, no. 4, pp. 217–222, 2013.

[2] B.-C. Sheu, W.-C. Chang, C.-Y. Cheng, H.-H. Lin, D.-Y. Chang, and S.-C. Huang, “Cytokine regulation networks in the cancer microenvironment.” *Frontiers in bioscience: a journal and virtual library*, vol. 13, pp. 6255–6268, 2007.

[3] P. Klemm, S. Oeltze-Jafra, K. Lawonn, K. Hegenscheid, H. Völzke, and B. Preim, “Interactive visual analysis of image-centric cohort study data,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 1673–1682, 2014.

[4] X. Dai and M. Gahegan, “Visualization based approach for exploration of health data and risk factors,” in *Proceedings of the International Conference on GeoComputation*, vol. 31, 2005.

[5] M. D. Steenwijk, J. Milles, M. Buchem, J. Reiber, and C. P. Botha, “Integrated visual analysis for heterogeneous datasets in cohort studies,” in *IEEE VIS Workshop on Visual Analytics in Health Care*, vol. 3, 2010, p. 3.

[6] J. Heinrich and D. Weiskopf, "State of the art of parallel coordinates," *Proceedings of Eurographics, STAR*, pp. 95–116, 2013.

[7] M. Krzywinski, J. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, S. J. Jones, and M. A. Marra, "Circos: an information aesthetic for comparative genomics," *Genome research*, vol. 19, no. 9, pp. 1639–1645, 2009.

[8] J. Fuchs, F. Fischer, F. Mansmann, E. Bertini, and P. Isenberg, "Evaluation of alternative glyph designs for time series data in a small multiple setting," in *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, 2013, pp. 3237–3246.

[9] J. Fuchs, P. Isenberg, A. Bezerianos, F. Fischer, and E. Bertini, "The influence of contour on similarity perception of star glyphs," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 2251–2260, Dec 2014.

[10] J. Heinrich and D. Weiskopf, "Continuous parallel coordinates," *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 6, pp. 1531–1538, 2009.

[11] M. Novotny and H. Hauser, "Outlier-preserving focus+ context visualization in parallel coordinates," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 5, pp. 893–900, 2006.

[12] E. J. Wegman and Q. Luo, "High dimensional clustering using parallel coordinates and the grand tour," in *Classification and Knowledge Organization*. Springer, 1997, pp. 93–101.

[13] C. A. Steed, J. E. Swan, T. Jankun-Kelly, and P. J. Fitzpatrick, "Guided analysis of hurricane trends using statistical processes integrated with interactive parallel coordinates," in *IEEE Symposium on Visual Analytics Science and Technology*, 2009, pp. 19–26.

[14] N. Andrienko and G. Andrienko, *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*. Springer-Verlag New York, Inc., 2005.

[15] M. Bostock, V. Ogievetsky, and J. Heer, "D^3 data-driven documents," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2301–2309, 2011.

[16] K. Chang, *Nutrient parallel coordinates*. http://bl.ocks.org/syntagmatic/3150059, August 2016 (last accessed).

[17] A. Garbarino, L. Sun, Z. Garbarino, C. Schmidt, and J. Chen, "Visgumbo, vismirror, viscut: interactive narrative strategies for large biological pathway comparisons," *IEEE VIS Workshop on Exploring Graphs at Scale (EGAS)*, 2016.