POINTS OF VIEW

Visualizing biological data

Data visualization is increasingly important, but it requires clear objectives and improved implementation.

Researchers today have access to an unprecedented amount of data. The challenge is to benefit from this abundance without being overwhelmed. Data visualization for efficient exploration and effective communication is integral to scientific progress. For visualization to continue to be an important tool for discovery, its practitioners need to be present as members of research teams.

One of the goals of data visualization is to enable people to explore and explain data through interactive software that takes advantage of human beings' ability to recognize patterns. Its success depends on the development of methods and techniques to transform information into a visual form for comprehension. It is a process that synthesizes skills from engineering, statistics and graphic design along with a number of other disciplines.

In recent years we have witnessed a growing appetite for the visual display of information. The ease of generating sophisticated computer graphics has encouraged the use of visualization. However, the value and utility of this popular form of communication remains unclear. We can easily be fooled into believing that we absorb more than we do when looking at large, colorful displays of information.

Data visualization, when applied to scientific research, has to be more than just the graphical display of information. Clear objectives are needed to drive design so we can assess the utility and practicality of visualizations. What is it that the researchers want to and need to see in the data? Which computational approaches and visual encodings will best bring out the trends? It is essential for the visualization practitioners to work side by side with the researchers to ensure that design decisions are continually refined to meet research objectives.

Unfortunately, there are few models for highly integrative teams consisting of visualization experts and biological researchers. The existence of distinct professional meetings and publication venues may be partially responsible for the barrier to working together. For example, the major professional meetings in information visualization such as Visualizing Biological Data (VIZBI) and BioVis (part of the Institute of Electrical and Electronics Engineers' VisWeek) attract few biologists. As a consequence, advances in visualization are not adequately described and shared with the biological community.

Identifying shared funding sources will certainly help to unite the professions under a common set of responsibilities and deliverables. Unfortunately, it is uncommon to hear stories in which agencies supporting scientific research have been successfully convinced to fund visualization efforts. Without the tight working relationships of an integrated team, we give up the ability to rapidly turn sketches into software prototypes at a pace relevant to the research: as data types and research questions evolve, there is a constant need to refine and adapt visualizations. One funding mechanism that has enabled an integrated focus on visualization in the context of biological research is supplemental support associated with awarded grants. The primary objective of such work is to extend the utility and accessibility of the research findings.

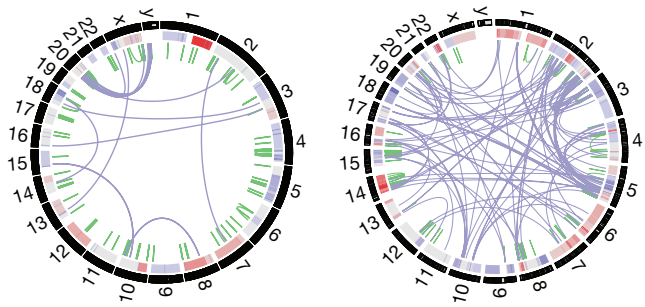

Figure 1 | Visualization of whole-genome rearrangement. Representative Circos plots[1] of whole-genome sequence data from two different tumors showing gene duplications and chromosome rearrangements. The outer ring depicts chromosomes arranged end to end. The inner ring displays copy-number data in green and interchromosomal translocations in purple. Reprinted from ref. 2 with permission from Elsevier.

It is useful to recognize that not every biological question will benefit from visualization, and graphical approaches should therefore be reserved for projects for which they will produce the greatest impact. Many data challenges can be addressed perfectly with computation alone. For a subset of research questions, however, visualizations can offer specific advantages over computation. In instances when we do not yet know the regularities in the data, visualization provides a powerful approach to explore the data for patterns. Visualization can also be useful for projects in which it complements algorithmic approaches. In genomics, for example, automated processes can reliably find sites where rearrangement occurs, but visualization is then needed to provide a mental image so that the detail of structural variation can be fully appreciated and understood (**Fig. 1**).

Data visualization represents a powerful aid to understanding data because well-designed graphical depictions of information can replace arduous cognitive assessment with simple perceptual inferences. For this reason, visualization can have a significant impact in biology, especially in the age of big data. For the last two and a half years, I have covered visual strategies for depicting scientific data. Although Points of View will go on hiatus after this month, these columns represent part of Nature Publishing Group's commitment to meeting visual communication challenges of scientific data.

Bang Wong

1. Krzywinski, M. *et al. Genome Res.* **19**, 1639–1645 (2009).
2. Imielinski, M. *et al. Cell* **150**, 1107–1120 (2012).

Bang Wong is the creative director of the Broad Institute of the Massachusetts Institute of Technology & Harvard and an adjunct assistant professor in the Department of Art as Applied to Medicine at The Johns Hopkins University School of Medicine.