POINTS OF VIEW

The overview figure

Our goal when writing research papers is to convey information as clearly as possible. In past columns I have suggested several graphic design techniques to improve the clarity of figures. In addition to refining data figures, including overview figures in a research paper provides a framework for readers to understand the experimental design and reported findings.

Illustrative schematics in overview figures can make publications accessible to a wider audience. They give context to the data presented. An example of such a figure is one I illustrated (**Fig. 1**)[1]. It depicts technology called Hi-C used to determine how cells organize the billions of DNA base pairs. This opening figure is effective because it constructs a mental model for understanding the technology and primes readers to expect DNA sequence information as the primary data type.

Typical overview figures illustrate a procedure (**Fig. 1**) or compare conditions such as 'control' and 'experimental' (**Fig. 2**)[2]. These figures portray a continuous process as discrete steps. As such, it is imperative that we create continuity through imagery and written descriptions. Each step in the progression is understood by relating it to the previous and subsequent step. For comparisons, differences in the corresponding steps between processes should also be highlighted (**Fig. 2**).

In the design of procedural schematics, it is useful to adopt an '*A* to *B*' structure in which *A* and *B* are states connected by an action. The states are often depicted graphically, and the action is text describing the transformation from *A* to *B* (for example, cut with restriction enzyme). To create good visual linkage between steps, redraw the elements from the previous step highlighting only the effective change. Because readers need to follow a series of events, it is helpful to account for all graphical elements introduced and removed from the figure. When the numbers of elements do not match from one step to the next, it can confuse readers and compromise the utility of overview figures.

With visual communications, it is essential that symbols have minimal overlapping meanings. For example, arrows can be used to point and to indicate motion. When the same graphical representation is used to mean different things, it impedes efficient and accurate decoding of information. In designing **Figure 2**, I used arrows to indicate progression and leaders—lines without arrowheads—for labeling. In **Figure 1**, I used arrows to represent and indicate the directionality of sequencing primers. Clear delineation in meaning enables readers to quickly learn the visual vocabulary and group information into hierarchy. Similarly, using language consistently makes it easier for readers to follow the word story. One sentence structure could be used to describe actions and another to label objects (that is, 'cut with restriction enzyme' and 'restriction fragments').

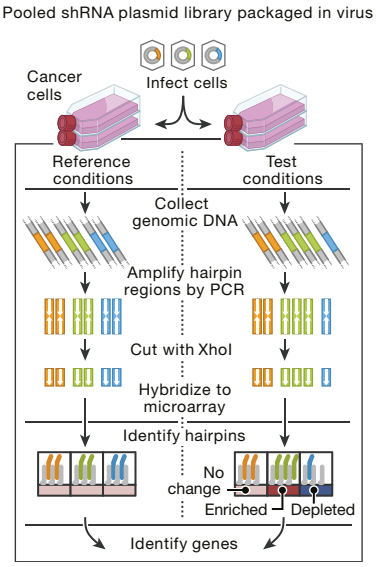

Figure 2 | Well-ordered compositions and clear visual encodings make schematics easy to follow. Schematic comparing experimental conditions in a pooled RNA interference screen. Reprinted from reference 2.

Fundamentally, overview figures are intended to convey general concepts and not to present data. When selecting graphics to represent each step, consider how a reader might interpret the imagery. In **Figure 2,** the authors initially selected a heatmap taken from elsewhere in the manuscript to illustrate the 'identify hairpins' step. Although the researchers did identify hairpins by analyzing heatmaps, a schematic representation (as shown) better demonstrates the experimental strategy. Research data in the context of an overview figure are disconcerting. Are we supposed to read them as graphs or see them as stand-ins for something else?

Despite their general usefulness, overview figures are usually the first to be eliminated when space becomes limited. One strategy to have them included in the final publication is to design the illustrations with an economy of marks and to make them as compact as possible. I designed the overview of Hi-C (**Fig. 1**) without intervening arrows and used the action labels as headers to save space. The horizontal layout provides a natural left-to-right ordering. Space-efficient designs can be achieved by fully using available whitespace[3] and organizing visual elements into groups according to the Gestalt principles [4,5].

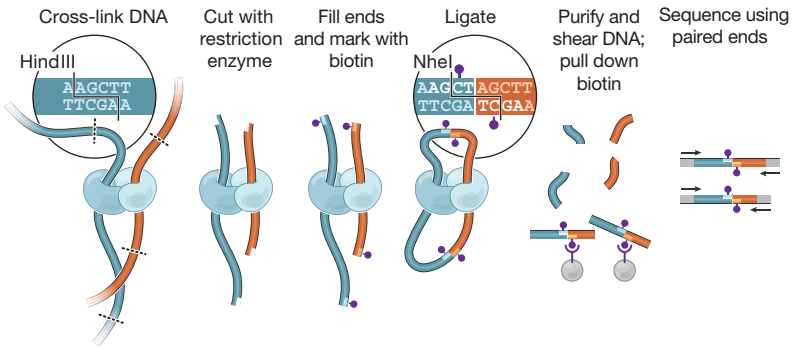

Figure 1 | Overview figures can clarify concepts. Outline of the Hi-C technique used to decipher the three-dimensional structure of the human genome. Reprinted from reference 1.

Bang Wong

1. Lieberman-Aiden, E. *et al. Science* **326**, 289–293 (2009).
2. Luo, B. *et al. Proc. Natl. Acad. Sci. USA* **105**, 20380–20385 (2008).
3. Wong, B. *Nat. Methods* **8**, 5 (2011).
4. Wong, B. *Nat. Methods* **7**, 863 (2010).
5. Wong, B. *Nat. Methods* **7**, 941 (2010).

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