

VIZBI 2012

VISUALIZATION PRINCIPLES

Practical Session – Proposed Solutions



Martin Krzywinski

martink@bcgsc.ca

mkweb.bcgsc.ca

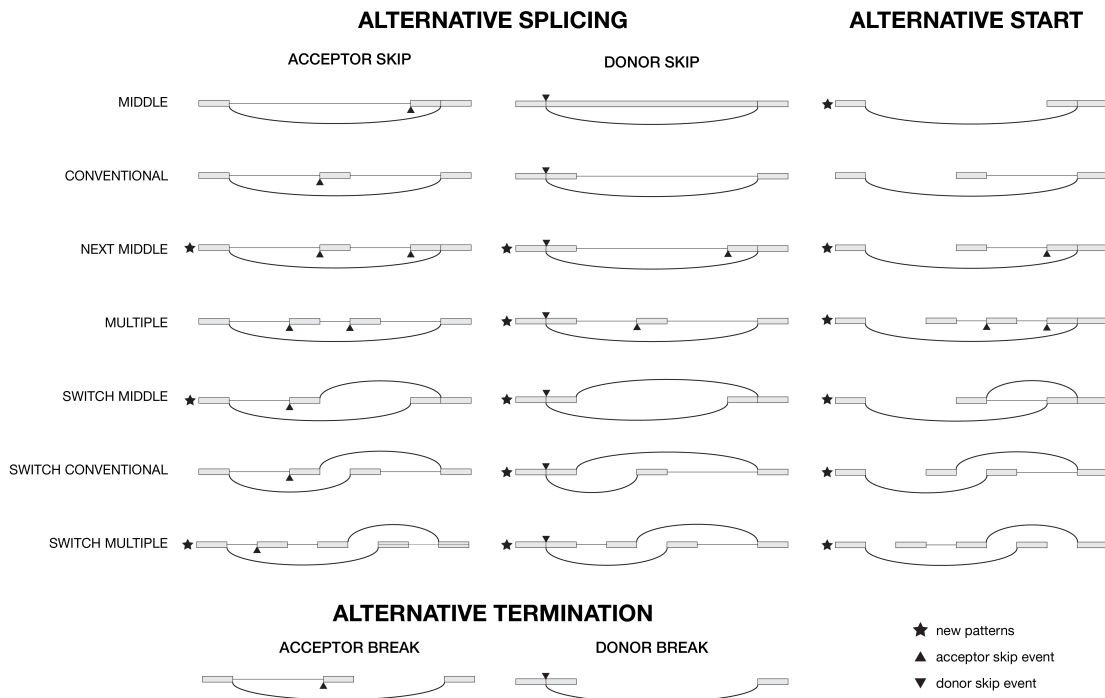
This document is available at

<http://mkweb.bcgsc.ca/vizbi/2012/practical-solutions.pdf>

All images can be downloaded from

<http://mkweb.bcgsc.ca/vizbi/2012/practical-solutions.zip>

IMAGE 1



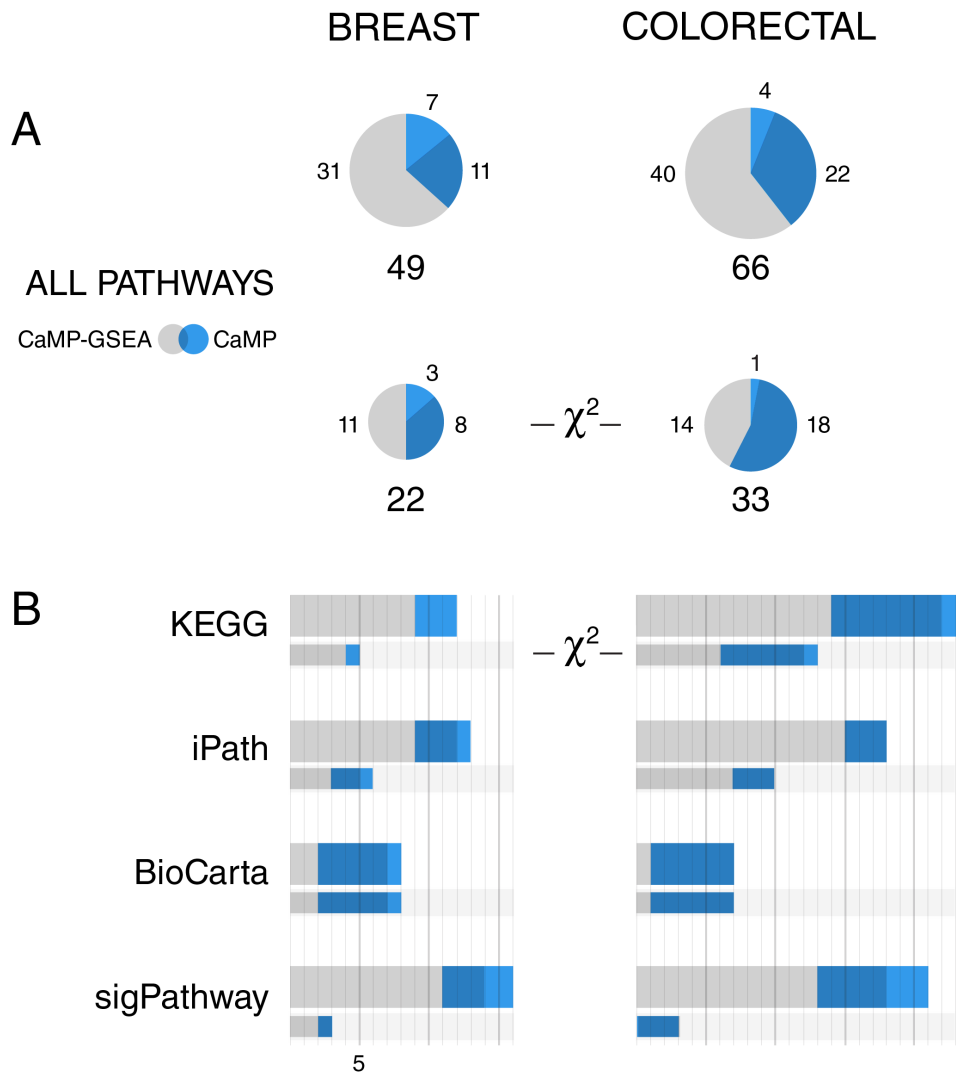
The original image suffers from three problems. First, patterns in splicing are difficult to distinguish because the exons, which are drawn as heavy boxes, steal your attention. Second, the variable spacing and size of exons adds clutter and further hides splicing patterns. Third, the data falls into a clear hierarchy which is lost in the repetitive labeling.

The redesigned figure emphasizes the salience of the splice lines, which are drawn as curves to provide better separation from the remaining components, which are linear. Exons are drawn subtly, evenly spaced and sized. Splicing is an adjacency relationship – spacing and size of exons is irrelevant.



<http://mkweb.bcgsc.ca/vizbi/2012/sharov-transcripts-redesign.pdf>

IMAGE 2 – PROPOSED SOLUTION

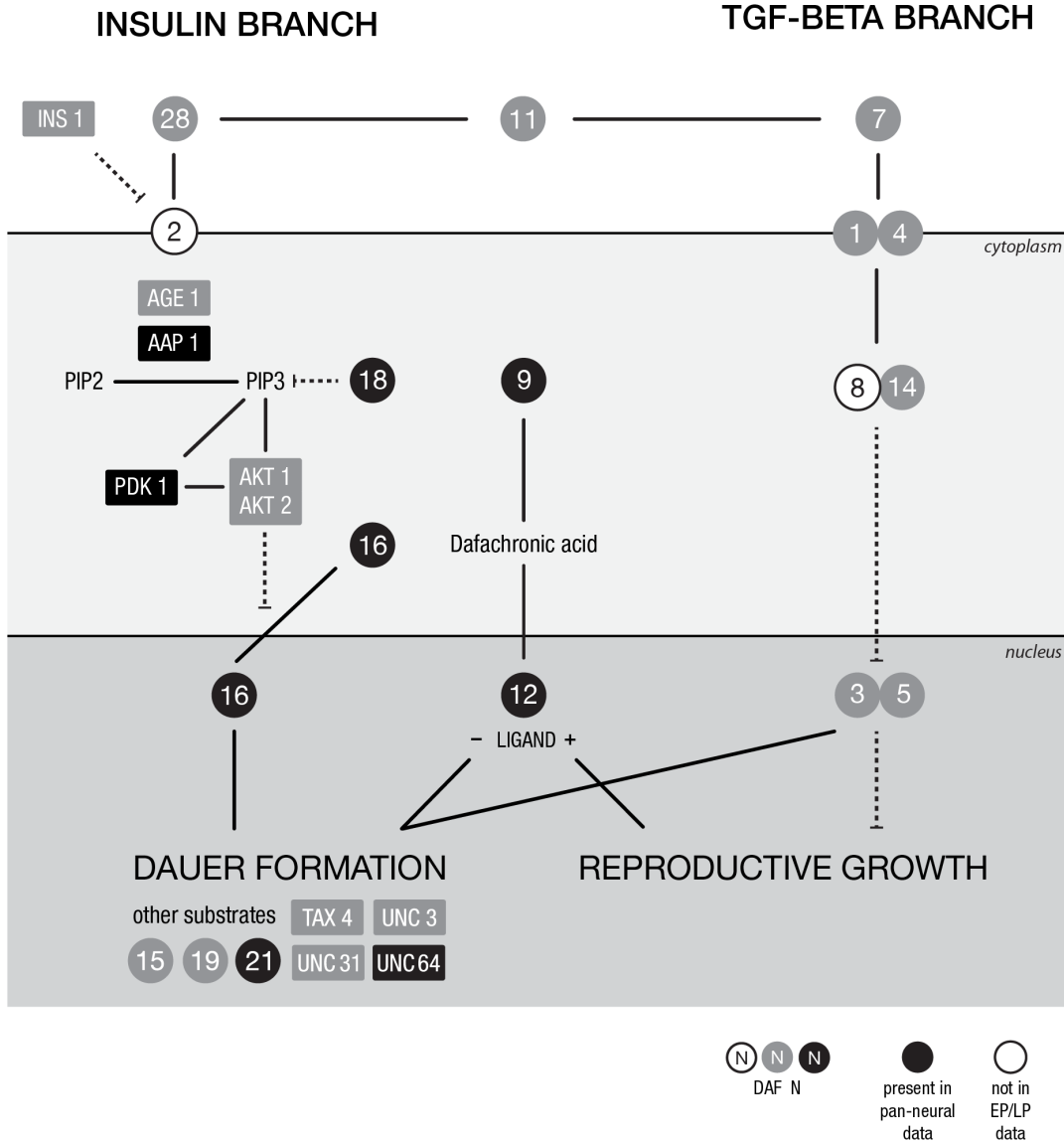


The original figure is unparseable. Multidimensional data should be communicated in steps, each introducing the reader to progressively more detail. The redesigned figure uses the first panel (A) to introduce you to data categories and proportions of values. The panel is meant to be a qualitative introduction to the data, with a more quantitative exposition in (B). There are patterns in the data that are quick to glean from this redesign and impossible to identify in the original figure. Can you find them?

.....

<http://mkweb.bcgsc.ca/vizbi/2012/lin-mutations-redesign.pdf>

IMAGE 3 – PROPOSED SOLUTION

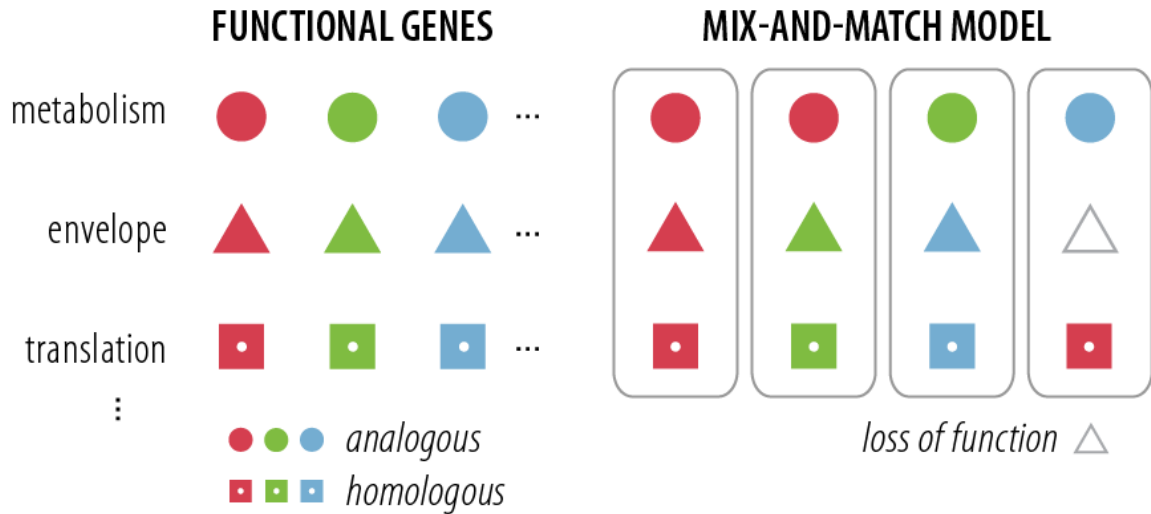


The original figure is too cluttered with chart junk. Too many colors and shapes are used to encode entities. With so much variation, all salience is lost. There is much redundancy in labels (e.g. DAF), which can be refactored. The redesign attempts to organize the pathway around the DAF components. The original figure uses bold, grey and italic text to encode encodings. This encoding is gratuitous because the classification is tripartite – every molecule falls into one of the categories. It is better to treat one condition as a reference (e.g. grey background for the enriched transcript shown originally in bold).

.....

<http://mkweb.bcgsc.ca/vizbi/2012/stetina-celegans-redesign.pdf>

IMAGE 4 – PROPOSED SOLUTION

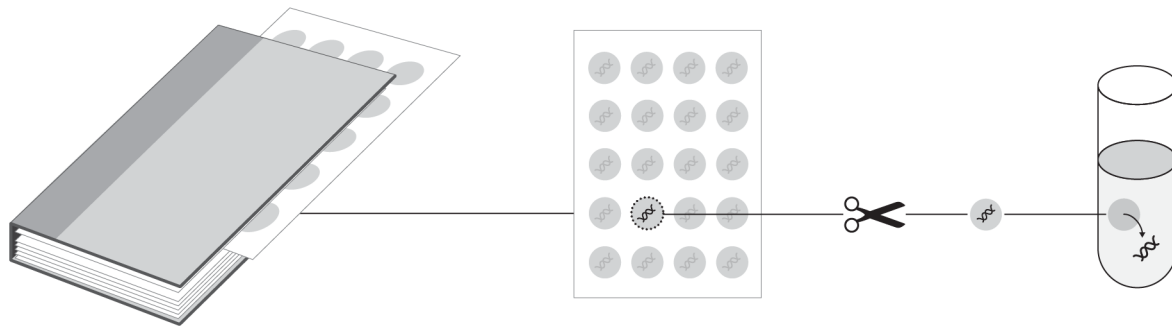


Use the fewest color and shapes required to communicate the message. The variety of shapes and colors in the original figure is unnecessary. Gene uniqueness can be encoded by a unique combination of color and shape. Additional relationships, such as homology which imply a specific type of similarity, can be encoded by a compatible encoding (shapes of homologous genes are made more similar by dot). In the mix-and-match model, loss of function is a feature but not shown in the original figure. This can be easily shown by using the same shape, but make it hollow, which is a natural encoding of diminished capacity (loss of function).

.....

<http://mkweb.bcgsc.ca/vizbi/2012/charlebois-evolution-redesign.pdf>

IMAGE 5 – PROPOSED SOLUTION

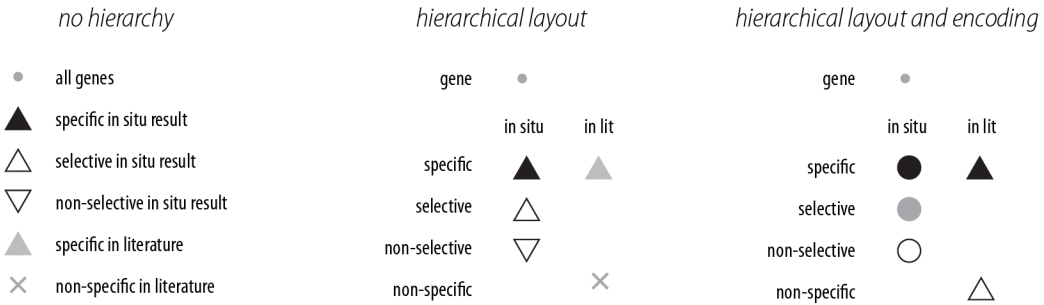


This is the core message of the DNA book. As much as $\frac{3}{4}$ of the original figure is preoccupied with the production and delivery of the book, aspects that are unimportant.

.....

<http://mkweb.bcgsc.ca/vizbi/2012/kawai-dna-book-redesign.pdf>

IMAGE 6 – PROPOSED SOLUTION



The data set falls into a natural hierarchy. Without making it explicit, would you have noticed that *non-selective in literature* was not an observed combination? Use traditional symbols, like circles, squares and triangles. Use fill and tone to emphasize importance. A hollow glyph will always be seen as subordinate to a solid one.



<http://mkweb.bcgsc.ca/vizbi/2012/nelander-genes-redesign.pdf>