Summarizing and Presenting Data

Do-s and Don’t-s for presenting genomic data

Ridiculogram: meaningless albeit visually impressive image of a network

Thursday, July 17, 2008

When visuals gone wrong ...

If you are in biosciences related work many times you have seen a fabulous but meaning-less image in the sci papers, typical error to represent data with wrong visuals perceptions, M. E. J. Newman named “Ridiculograms” (you can see the related video in Youtube).

A ridiculogram can be defined as:

• Visual stunning
• Scientifically worthless
• Published in Nature or Science
Data transformations
Consider data transformations
Consider data transformations
\[
\log(X/Y) = \log(X) - \log(Y) = - \left( \log(Y) - \log(X) \right) = - \log(Y/X)
\]

Ingo Ruczinski  | Asian Institute in Statistical Genetics and Genomics  | July 22-23, 2016
Figure 1  Plots of the female/male genetic-distance ratio against sex-averaged genetic location (in cM) along six selected chromosomes. Approximate locations of the centromeres are indicated by the triangles. The dashed lines correspond to equal female and male distances.

Comparisons
Don’t distort the quantities
value proportional to radius

Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 22-23, 2016

[ KB ]
Don’t use areas at all
value proportional to length

Should you include zero..?
Ease comparisons
things to be compared should be adjacent

Distribution of All TFBS Regions

- Pseudogene/ambiguous: 17%
- 5' to known gene: 22%
- Novel: 24%
- Within or 3' flanking to a known gene: 36%

866 Total TFBS Regions
Color and 3D

Optical illusions

en.wikipedia.org/wiki/Optical_illusion
Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 22-23, 2016

Nature Methods Point of View

Discrepancy in appearance relating to height

Occluding object

http://clearscience.info/wp/?p=546
R package dichromat

http://www2.warwick.ac.uk/fac/sci/moac/people/students/peter_cock/r/heatmap/
Nonetheless, there are a few small concerns. The book has beautiful graphics that brilliantly illustrate the points, provided that one is not red-green colorblind (about 12% of males, and a very small percentage of females). This oversight will plague many readers, and it is very unfortunate that a different choice of color schemes was not made early in the process of writing. We hope the second edition will correct this problem.

David Banks