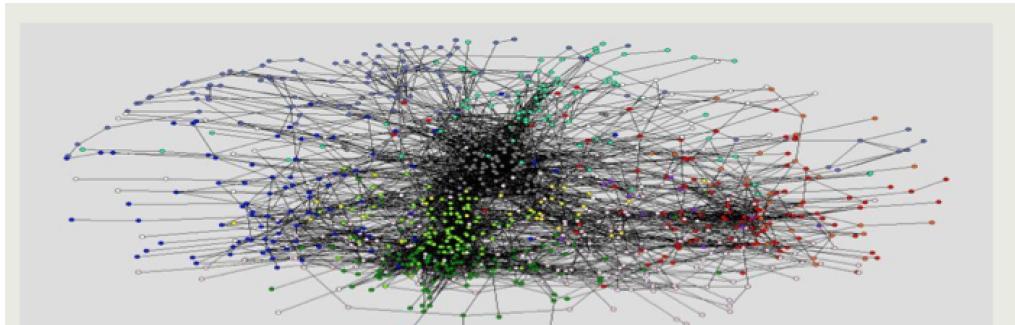


# Summarizing and Presenting Data

## Do-s and Don't-s for presenting genomic data

Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017



*Ridiculogram: meaningless albeit visually impressive image of a network*

Thursday, July 17, 2008

### When visuals gone wrong ...

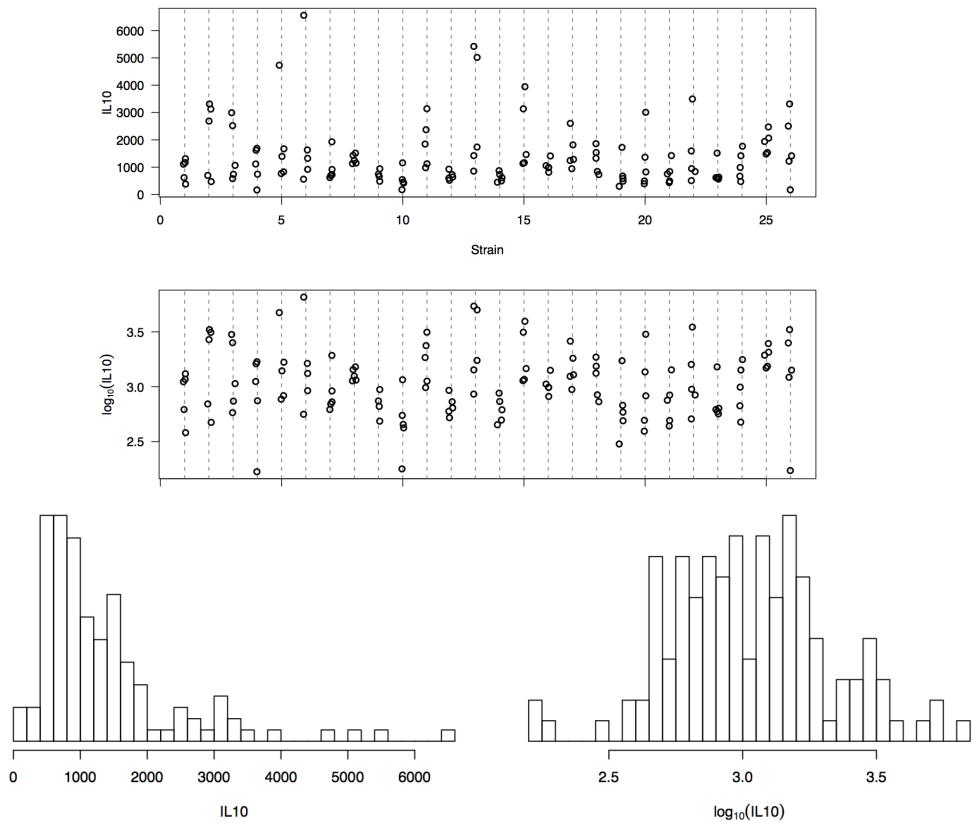
If you are in biosciences related work many times you have seen a fabulous but meaning-less image in the sci papers, typical error to represent data with wrong visual perceptions, M. E. J. Newman named "Ridiculograms" (you can see the [related video](#) in Youtube).

A ridiculogram can be defined as:

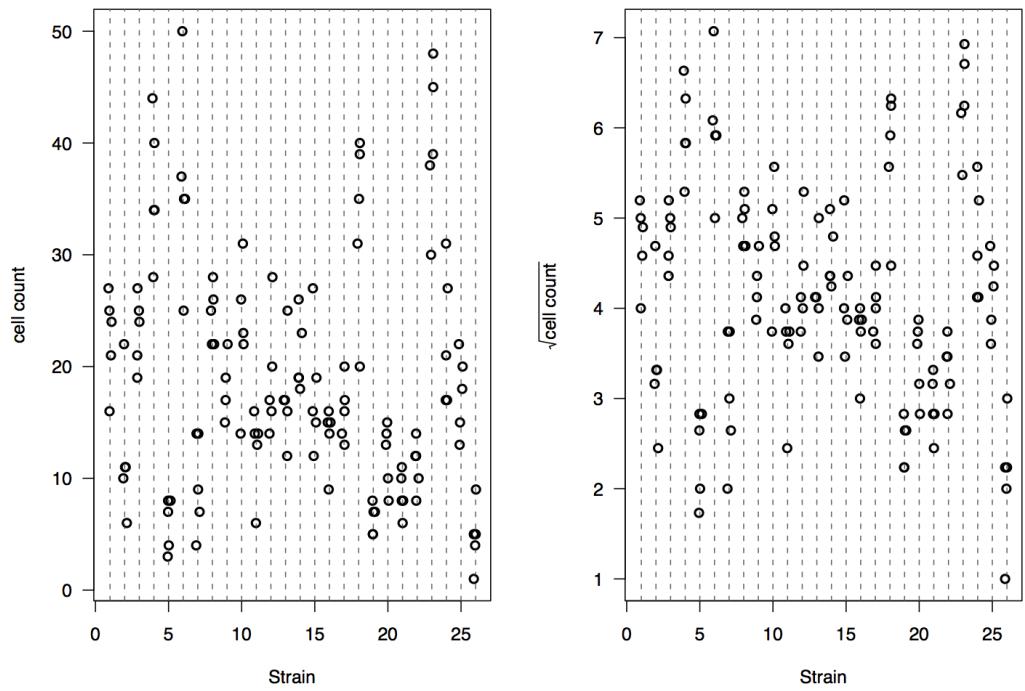
- Visual stunning
- Scientifically worthless
- Published in Nature or Science

## Data transformations

Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

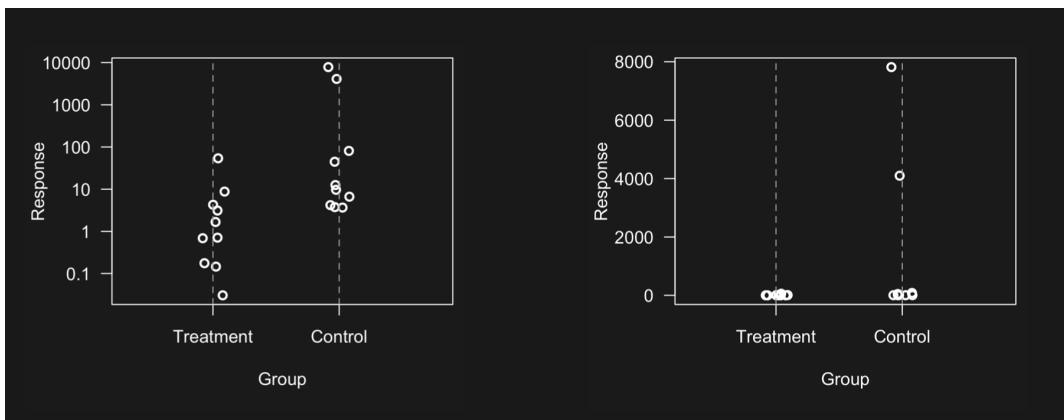


Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017



Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

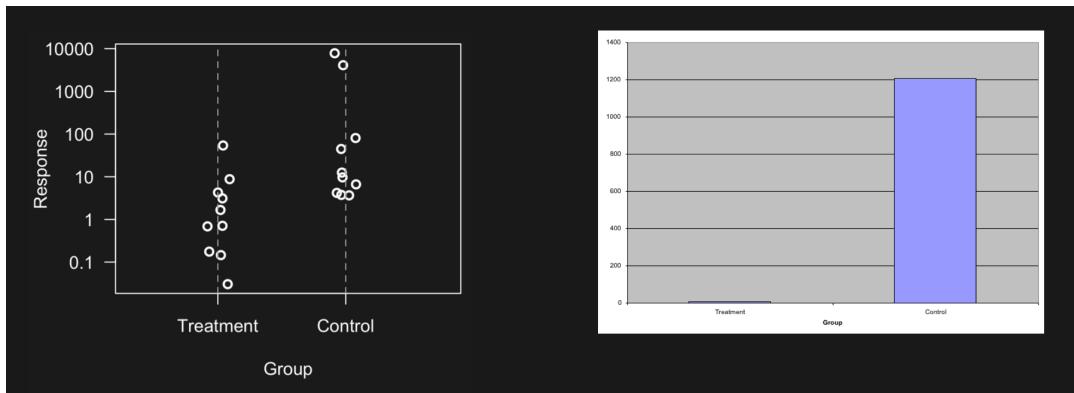
## Consider data transformations



Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

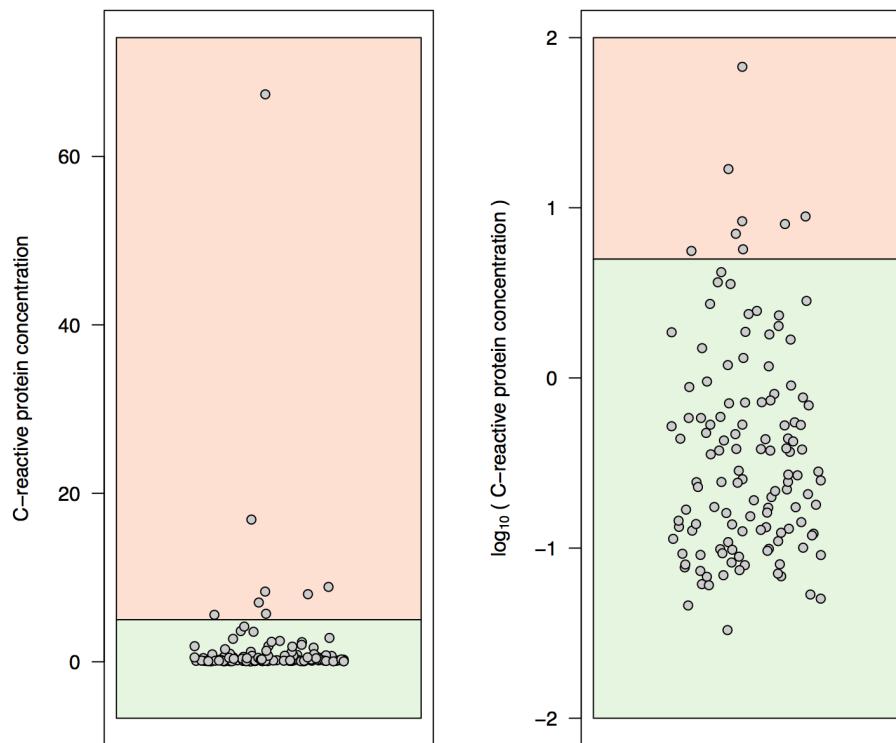
[ KB ]

## Consider data transformations

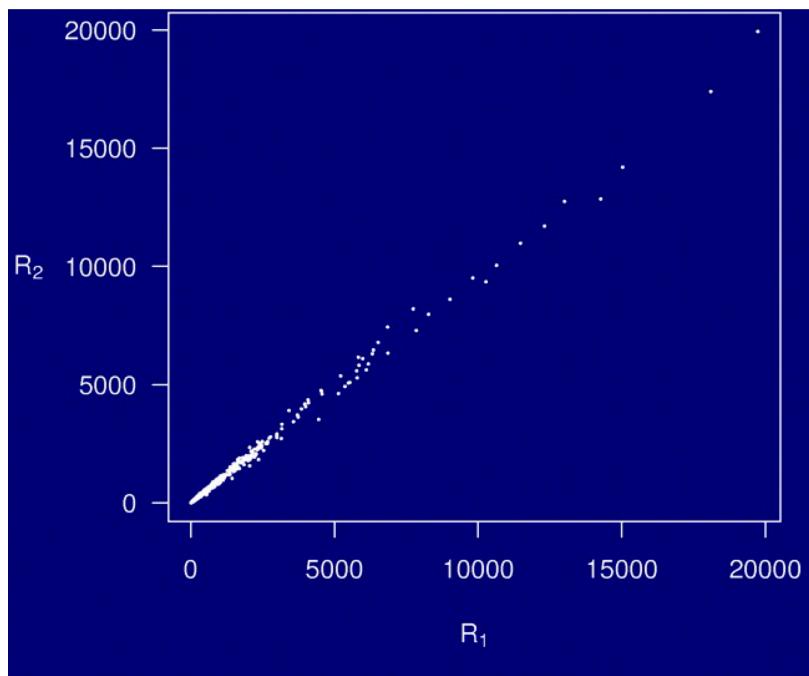


Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

[ KB ]

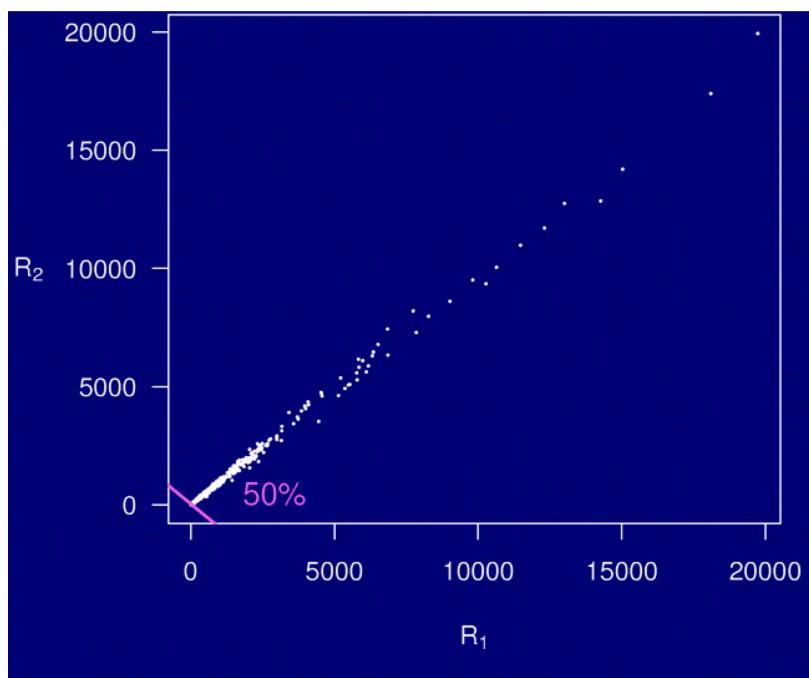


Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017



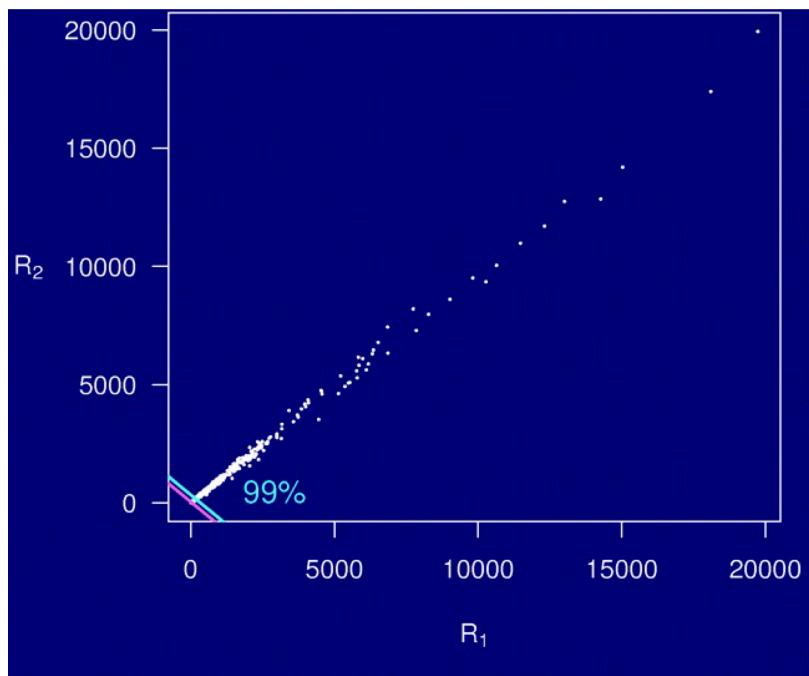
Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

[ KB ]



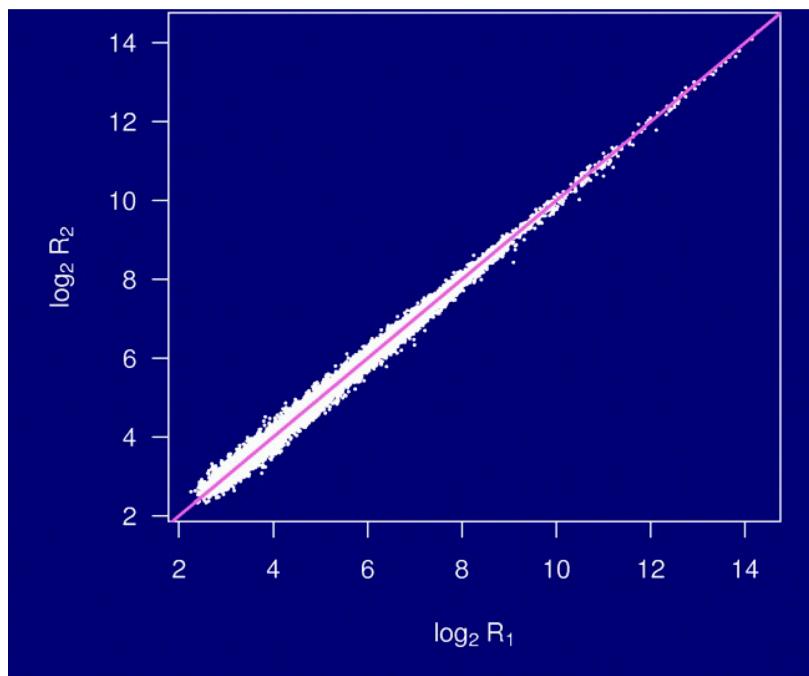
Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

[ KB ]



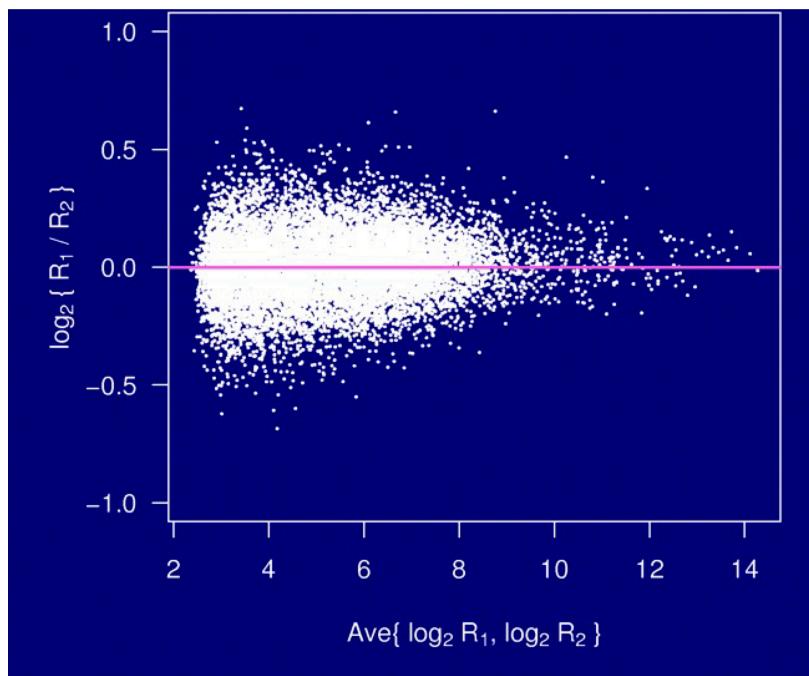
Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

[ KB ]



Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

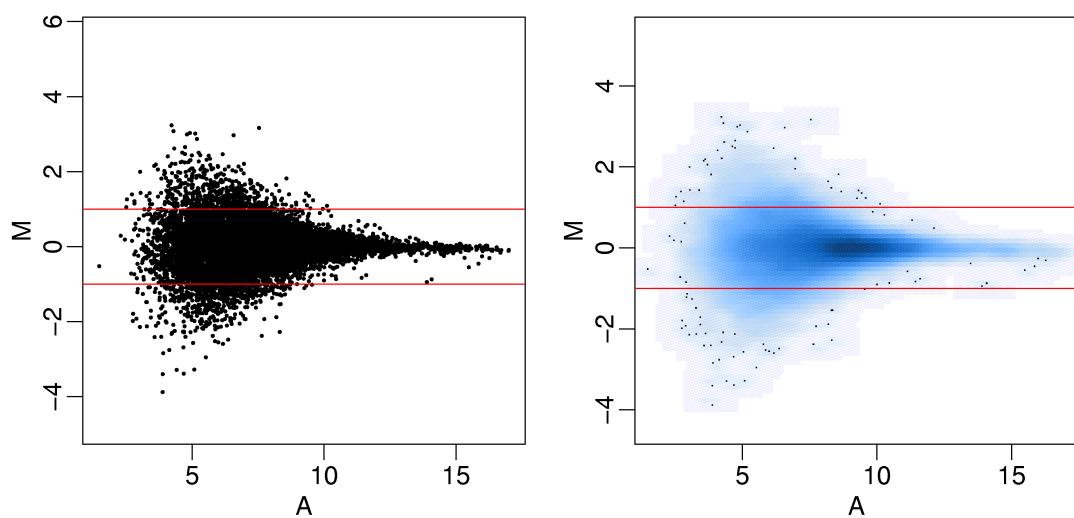
[ KB ]



Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

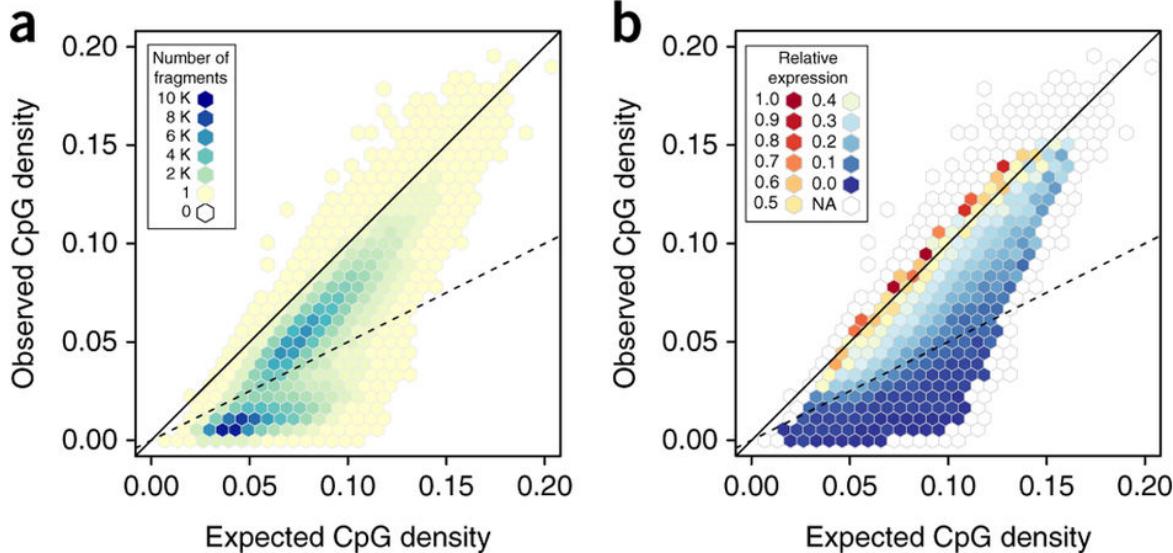
[ KB ]

## Scatter smooth



Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

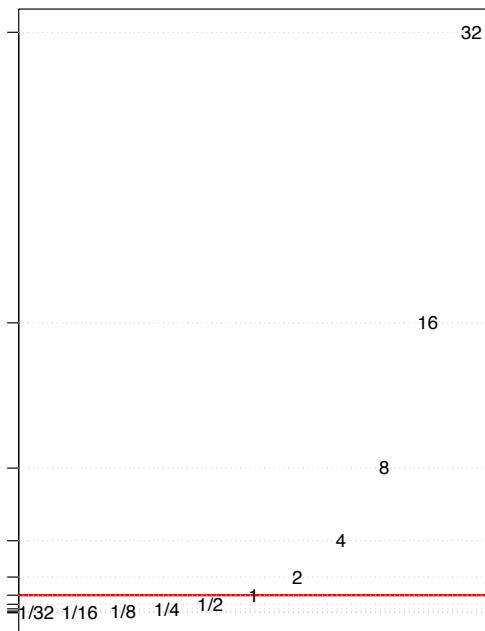
## hexbin



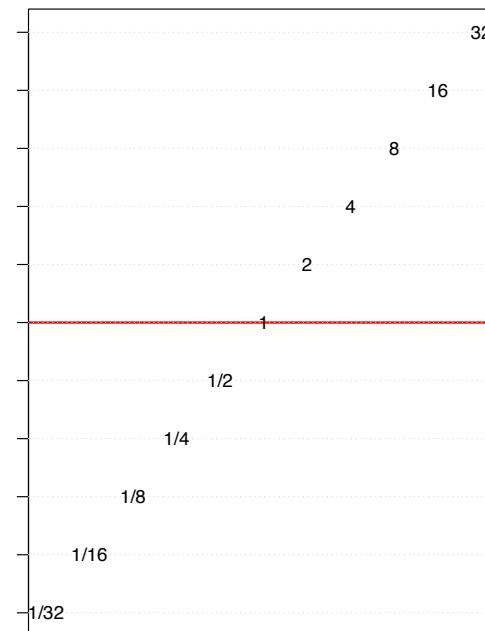
Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

PMID 28024146

Original scale

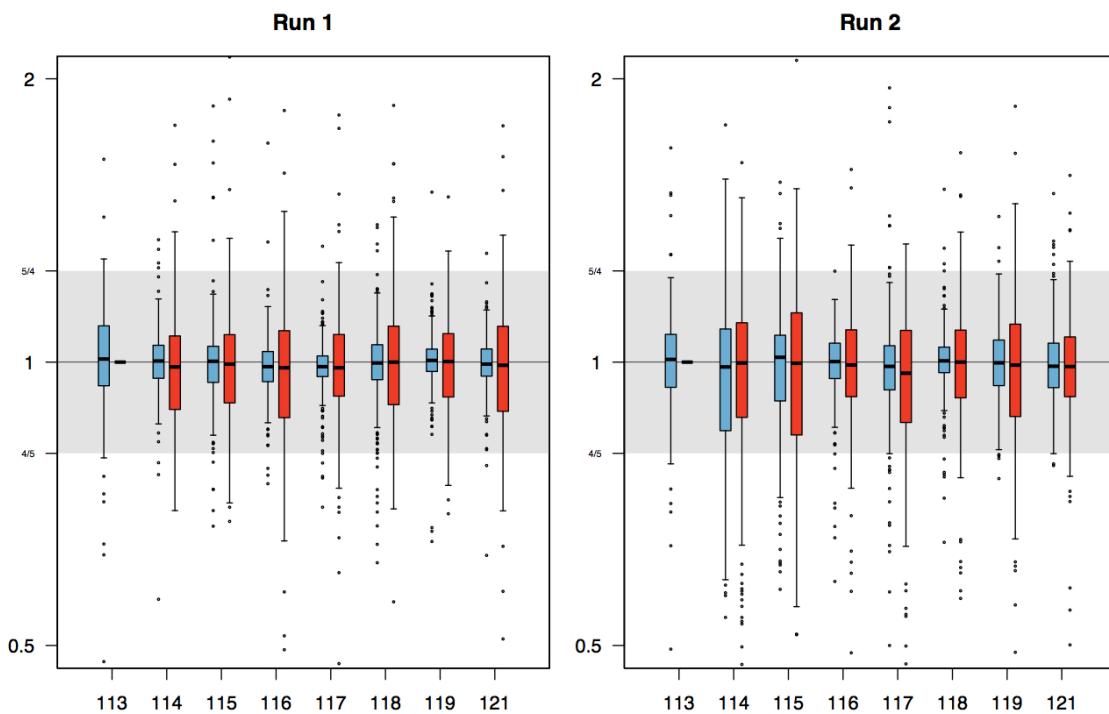


Logarithmic scale

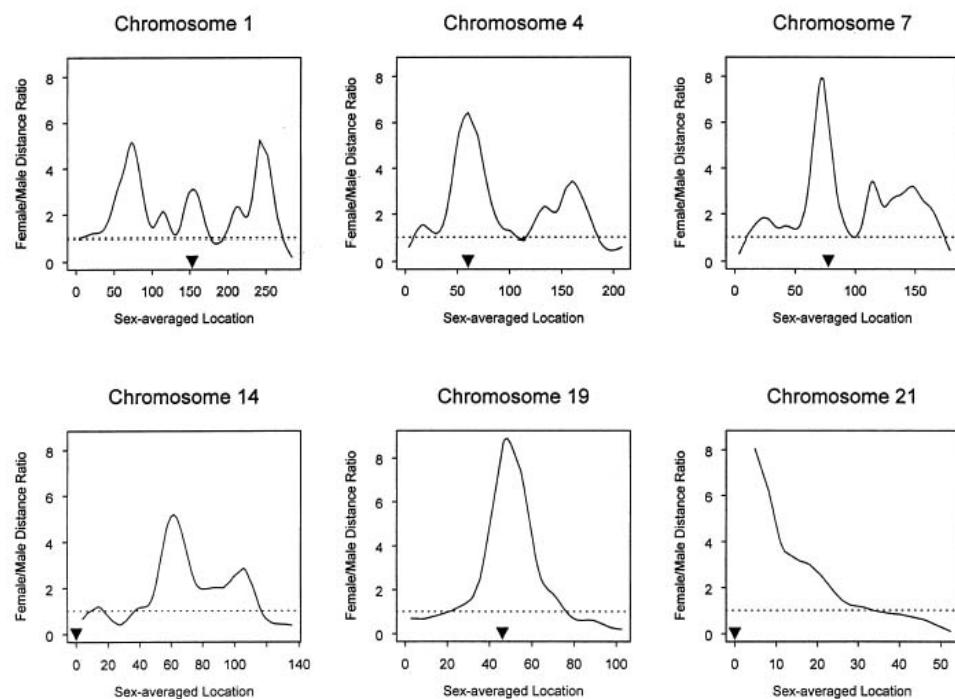


$$\log(X/Y) = \log(X) - \log(Y) = -(\log(Y) - \log(X)) = -\log(Y/X)$$

Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017



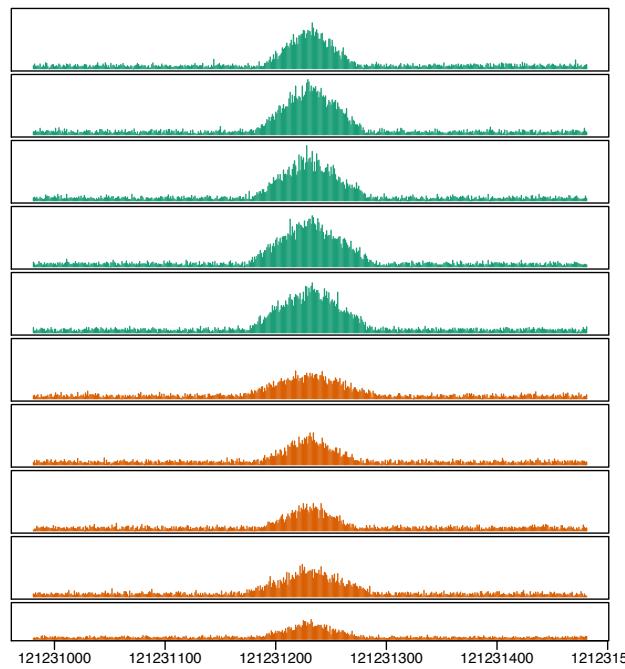
Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

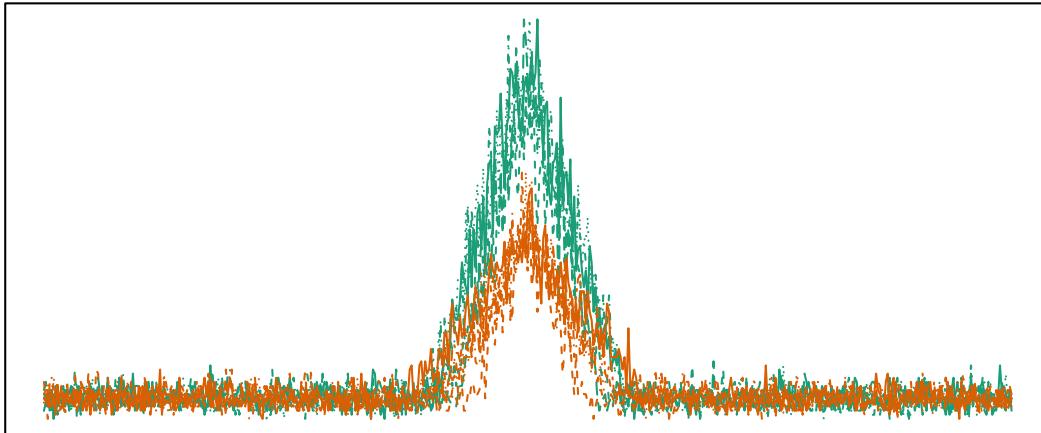


**Figure 1** Plots of the female:male genetic-distance ratio against sex-averaged genetic location (in cM) along six selected chromosomes. Approximate locations of the centromeres are indicated by the triangles. The dashed lines correspond to equal female and male distances.

# Comparisons

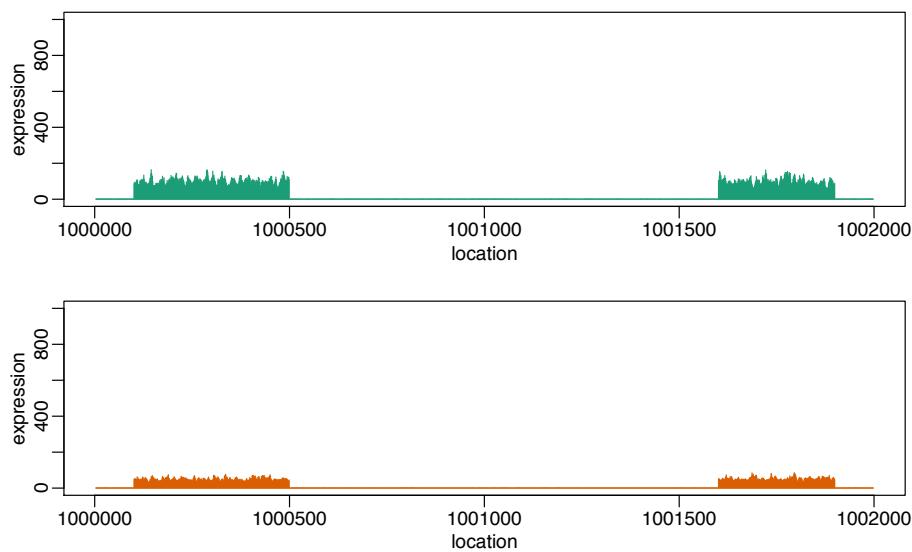
Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017





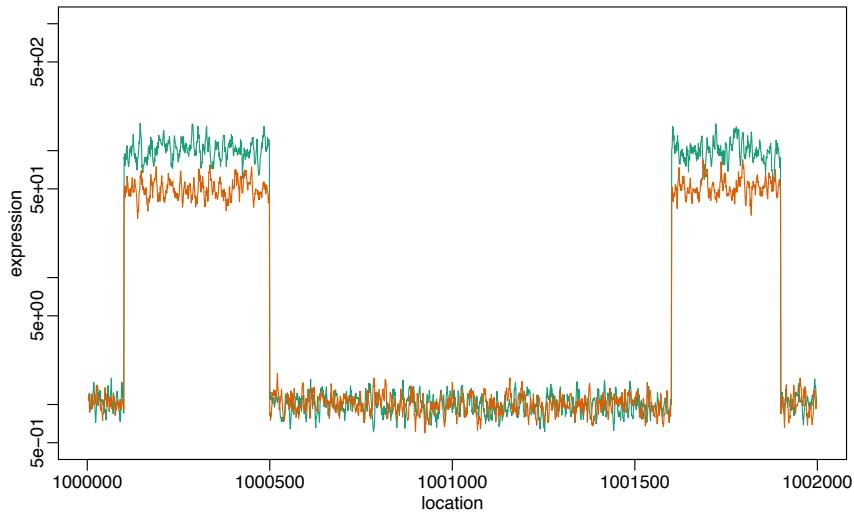
Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

[ RI ]



Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

[ RI ]



Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

[ RI ]

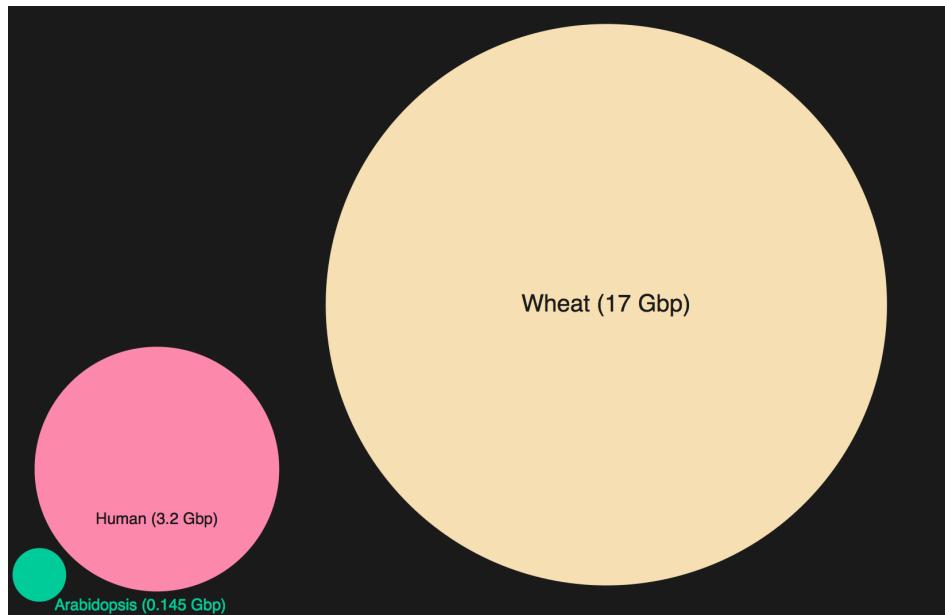
## Don't distort the quantities value proportional to radius



Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

[ KB ]

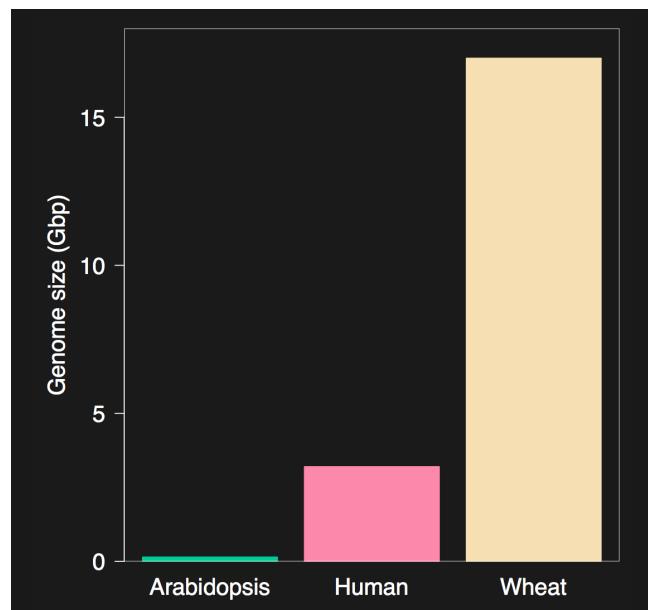
## Don't distort the quantities value proportional to area



Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

[ KB ]

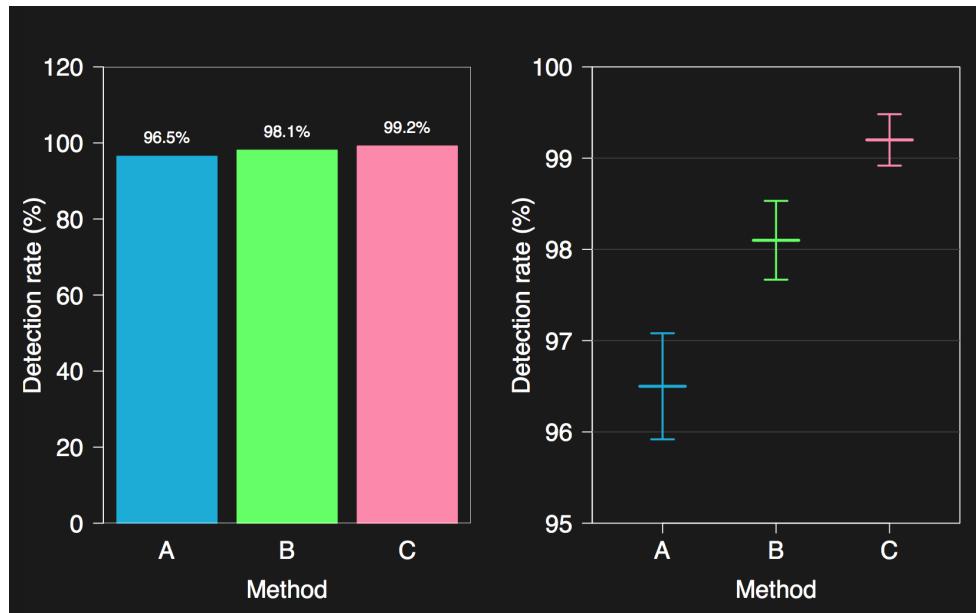
## Don't use areas at all value proportional to length



Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

[ KB ]

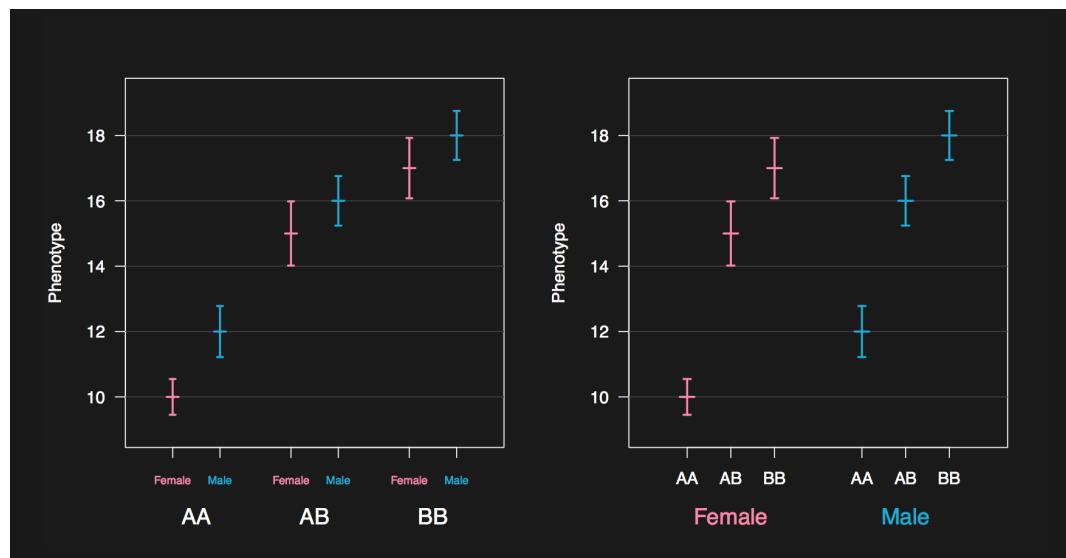
## Should you include zero..?



Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

[ KB ]

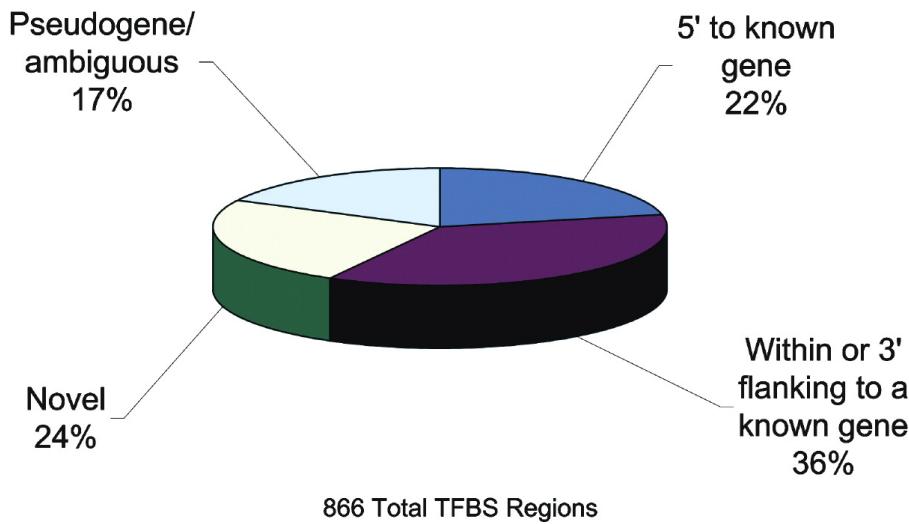
Ease comparisons  
things to be compared should be adjacent



Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

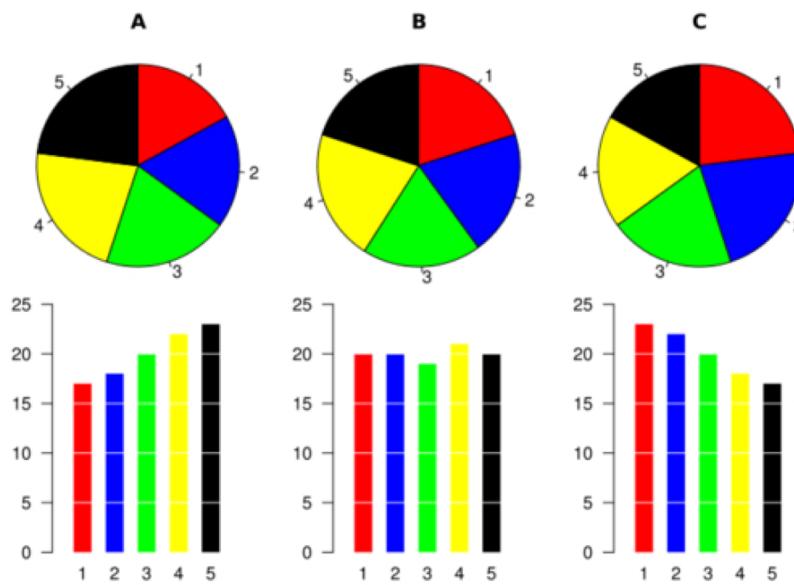
[ KB ]

# Distribution of All TFBS Regions



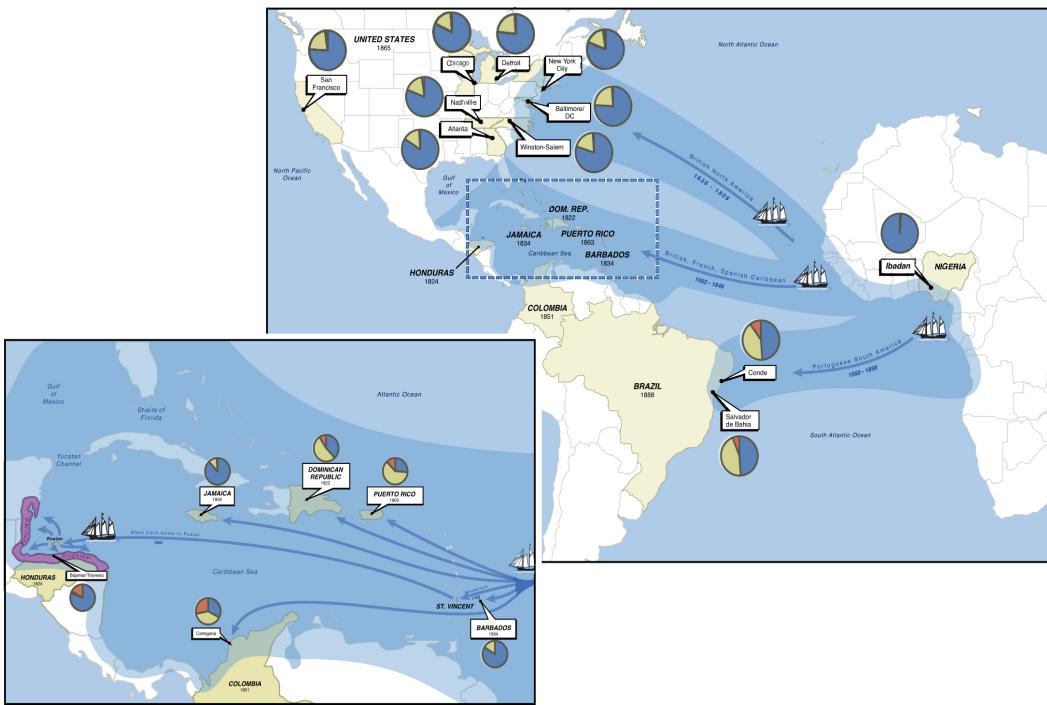
Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

PMID 14980218



Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

[www.businessinsider.com](http://www.businessinsider.com)

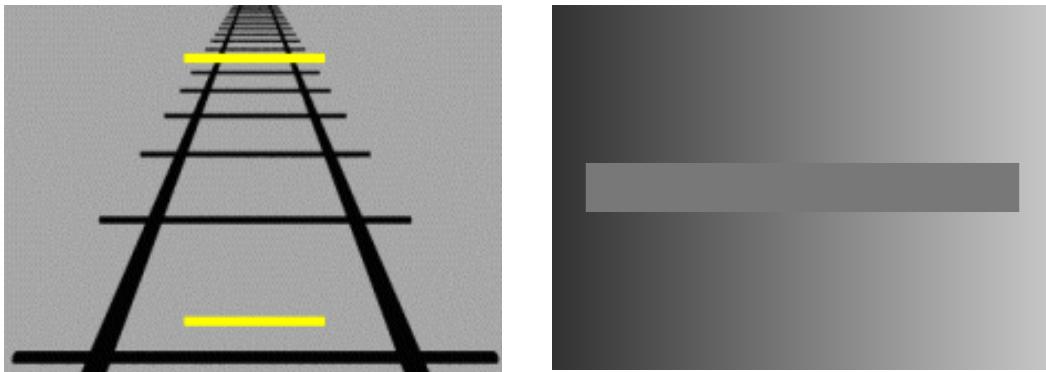


Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

[www.caapaproject.net](http://www.caapaproject.net)

## Color and 3D

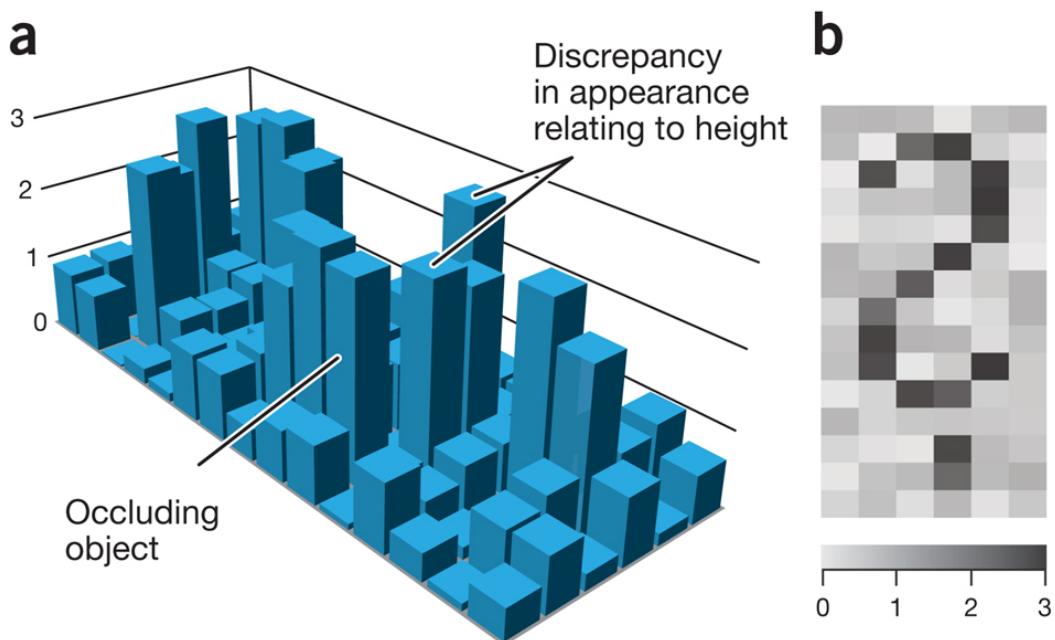
# Optical illusions



Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

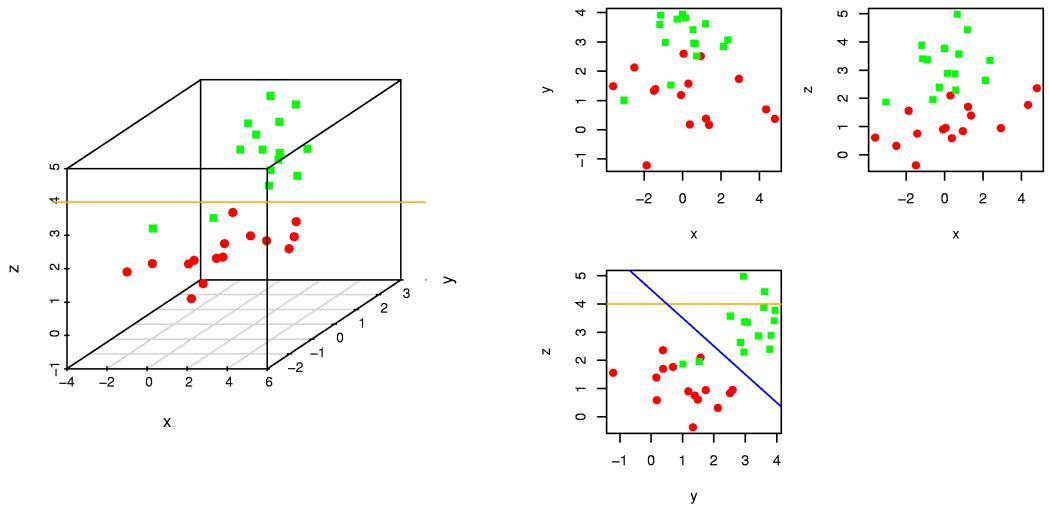
[en.wikipedia.org/wiki/Optical\\_illusion](http://en.wikipedia.org/wiki/Optical_illusion)

## Nature Methods Point of View



Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

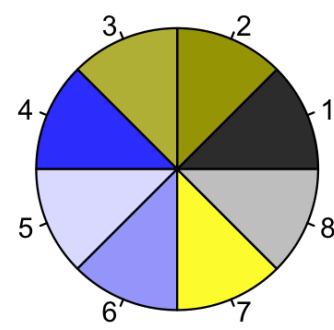
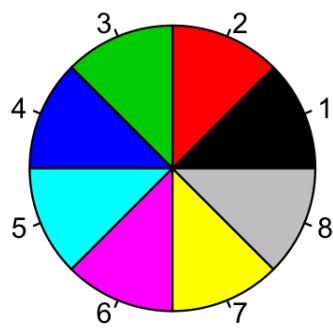
<http://clearscience.info/wp/?p=546>



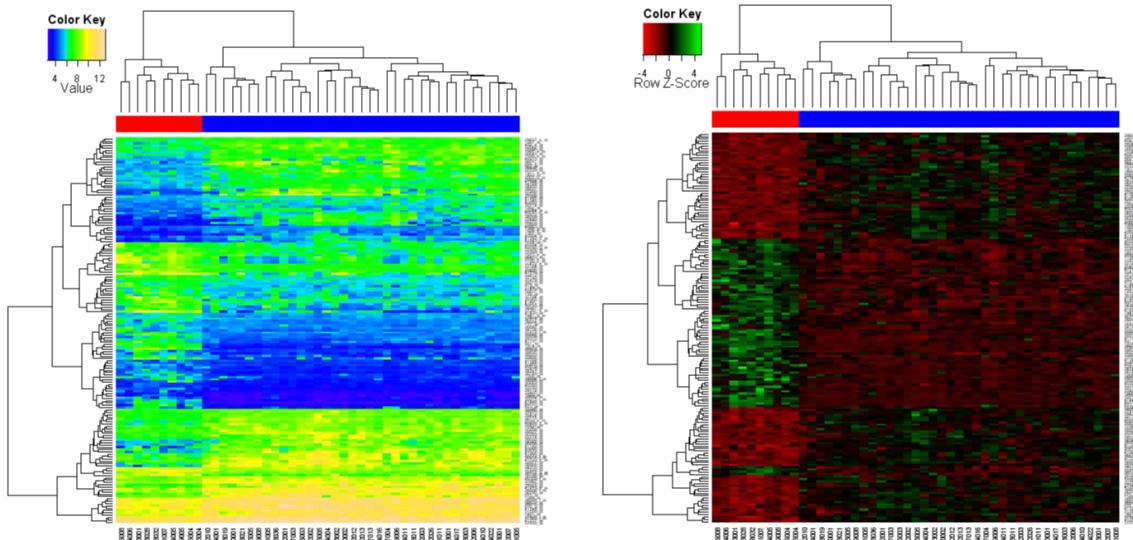
Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

[ RI ]

## R package dichromat

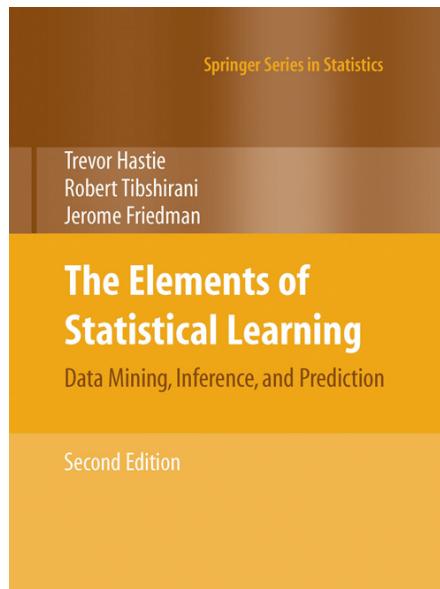


Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017



[http://www2.warwick.ac.uk/fac/sci/moac/people/students/peter\\_cock/r/heatmap/](http://www2.warwick.ac.uk/fac/sci/moac/people/students/peter_cock/r/heatmap/)

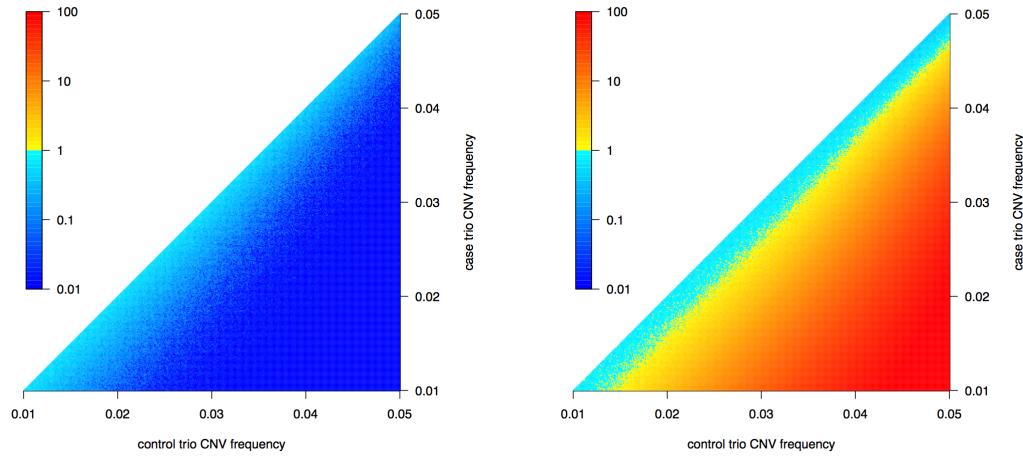
Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017



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

Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017



Ingo Ruczinski | Asian Institute in Statistical Genetics and Genomics | July 21-22, 2017

PMID 24528994

