Relaxation Techniques for the Upset Data Scientist

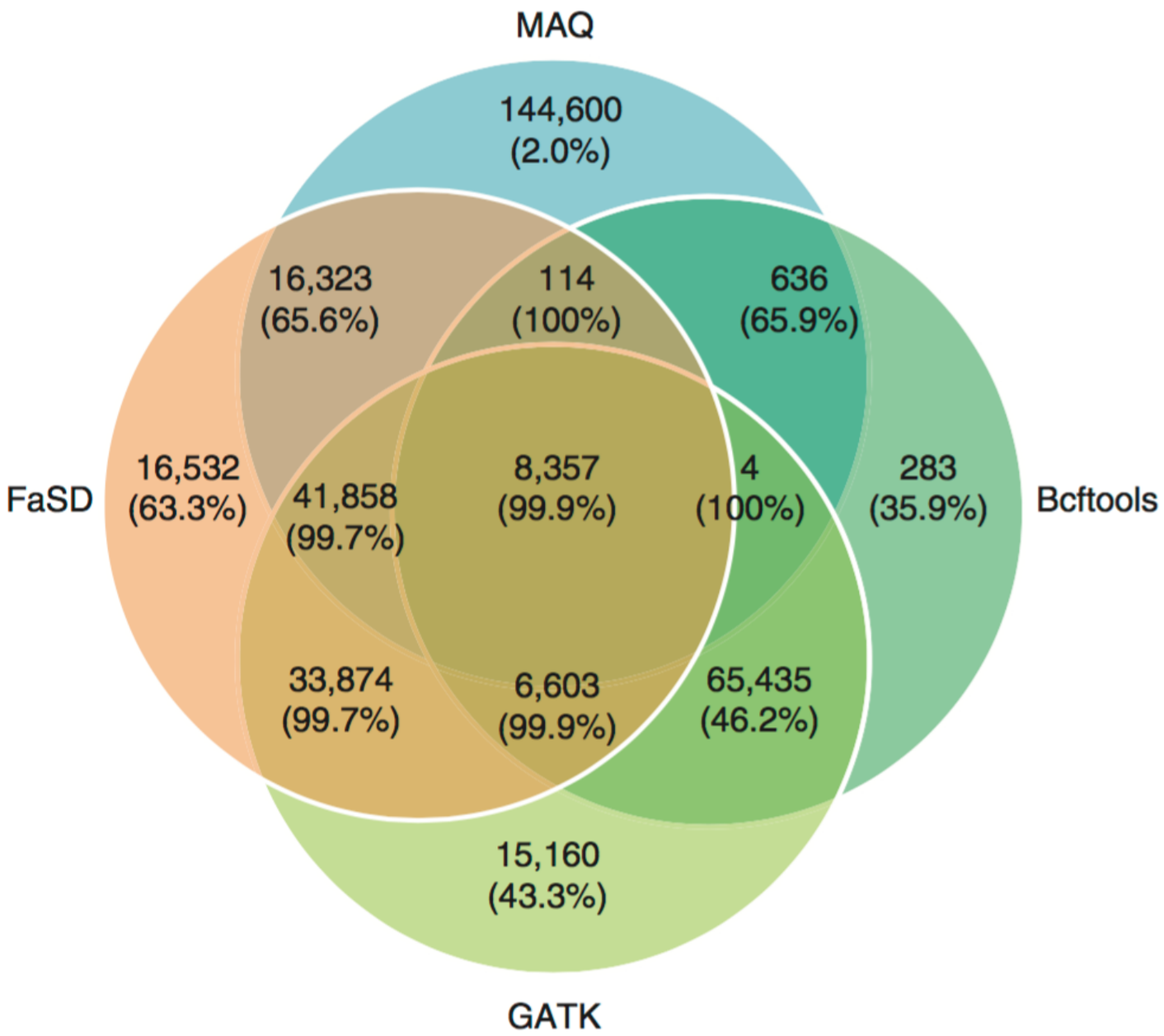
Nils Gehlenborg, PhD

http://gehlenborglab.org | @*nils_gehlenborg*

What are the sources of your tension and anger?

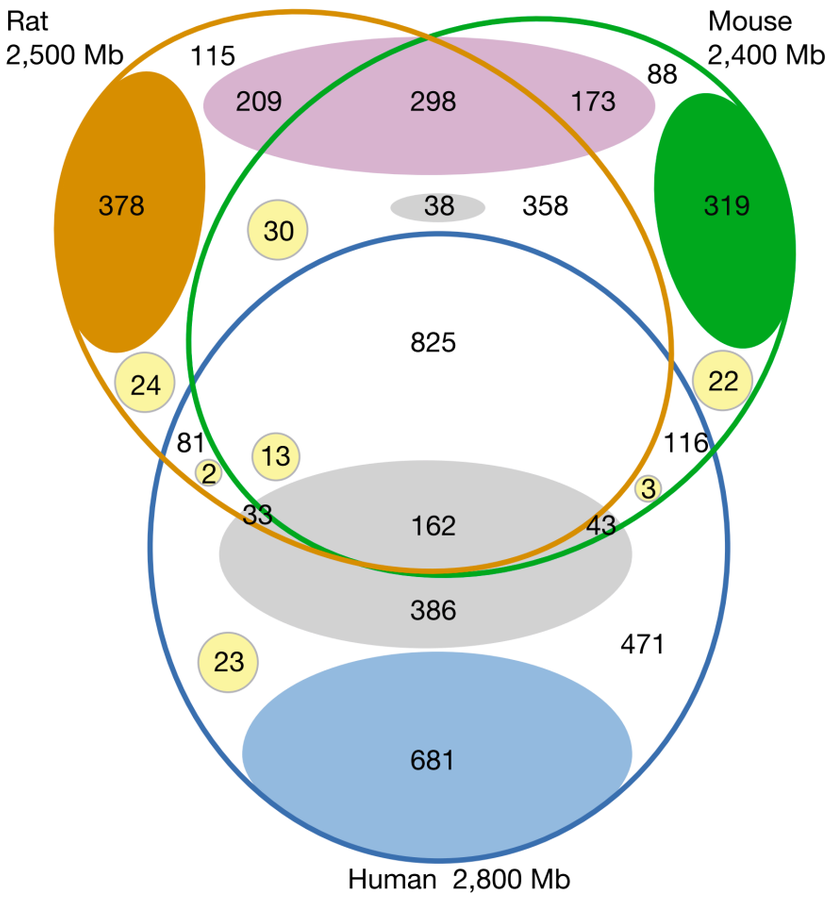
Xu et al., Nature Communications 3, 1258 (2014)

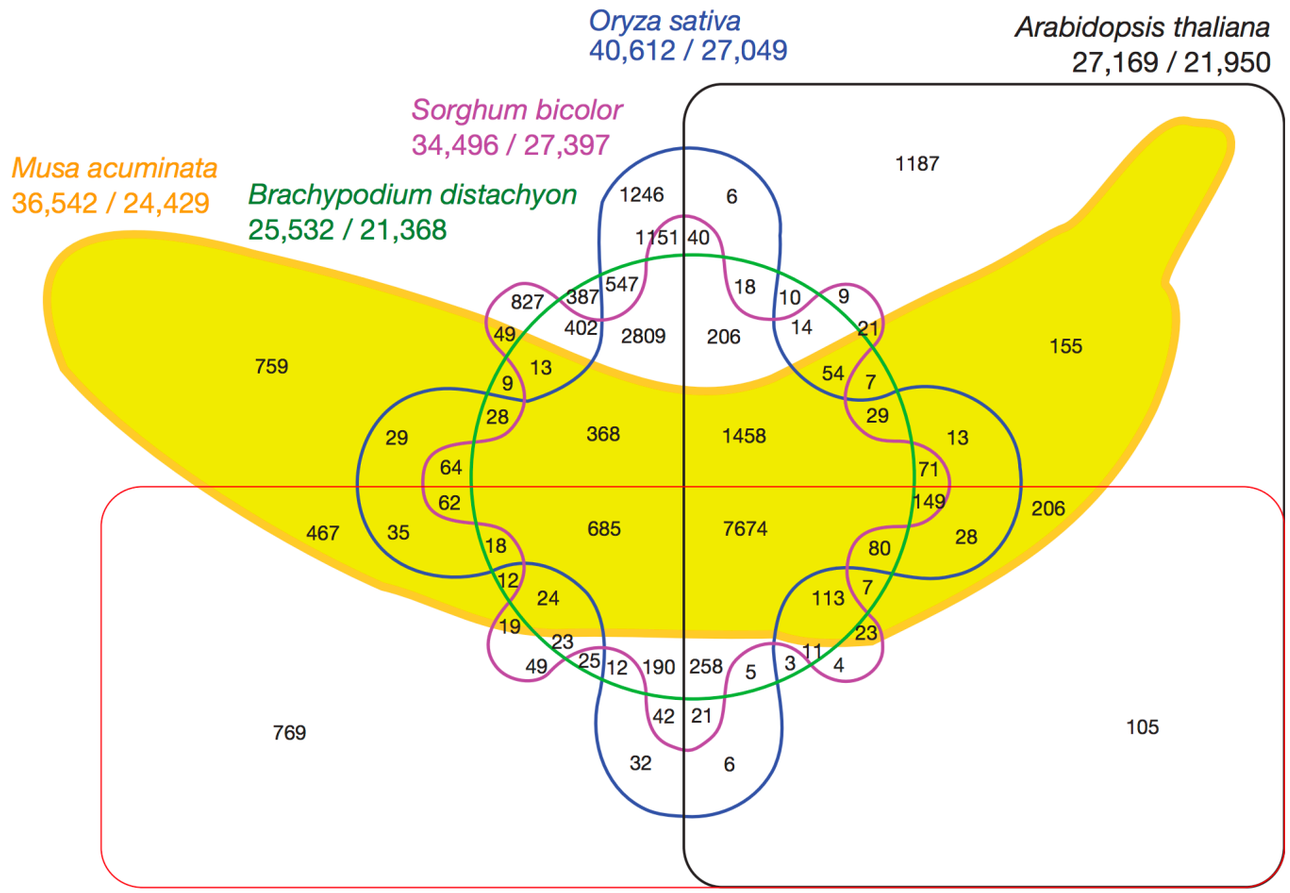
Where is the biggest intersection?



Rat Genome Sequencing Consortium, Nature 428, 493 (2004)
D'Hont et al., Nature 488, 213 (2012)

"Rats," you'll say, "I'll go bananas!"





Why is this stressful and exhausting?

This visualization makes you work!

This visualization makes you work!

[Good] visualization uses perception to free up cognition!

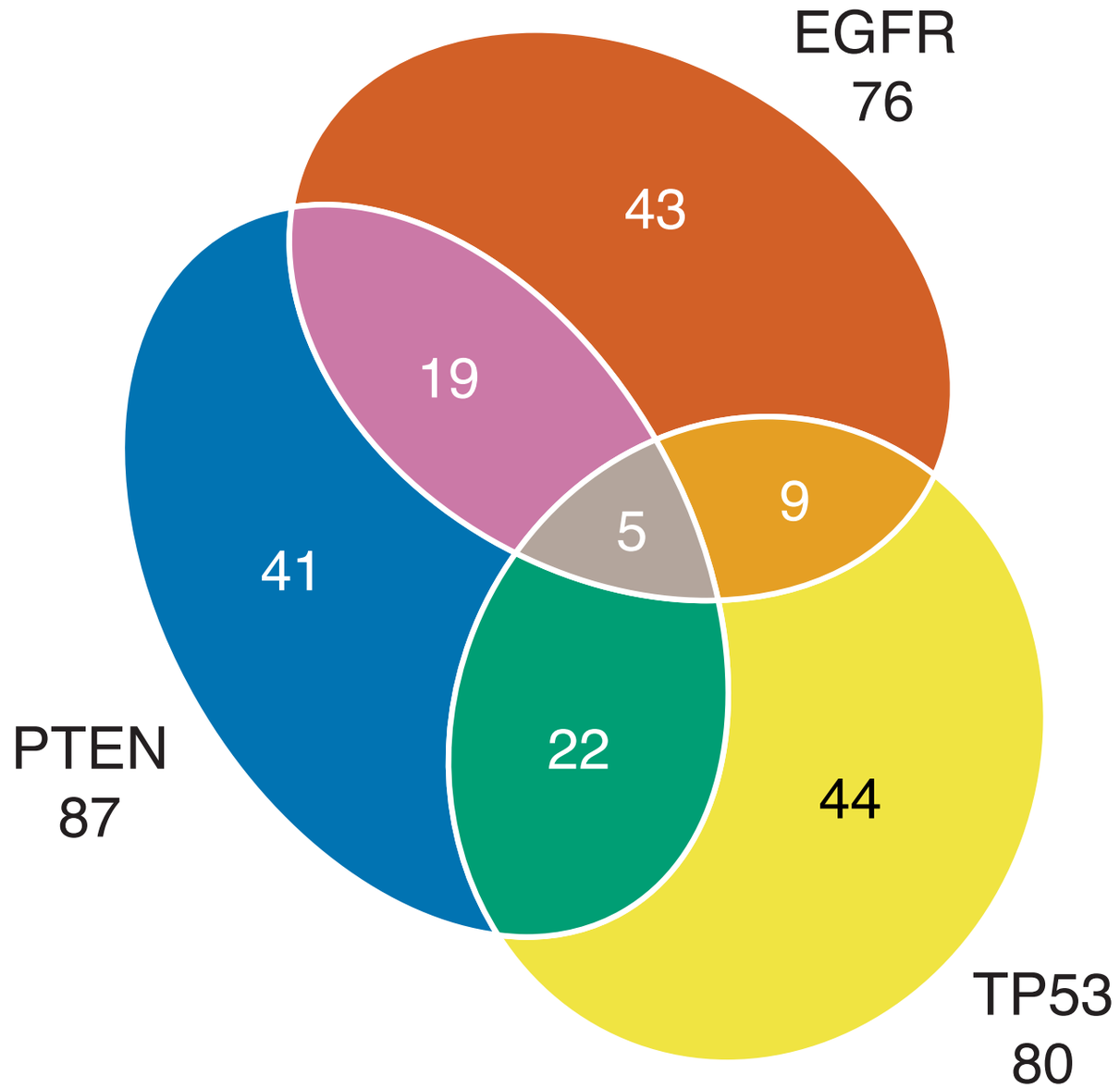
What does that mean?

What does that mean?

$$
\begin{array}{r}
453 \times 862 = \\
\hline
906 \\
+\ 27,180 \\
+\ 362,400 \\
\hline
390,486
\end{array}
$$

What are some proven relaxation techniques?

Lex & Gehlenborg, Nature Methods 11, 1 (2014)



Lex & Gehlenborg, Nature Methods 11, 1 (2014)

Lex & Gehlenborg, Nature Methods 11, 1 (2014)

1. Let Go!

Lex & Gehlenborg, Nature Methods 11, 1 (2014)

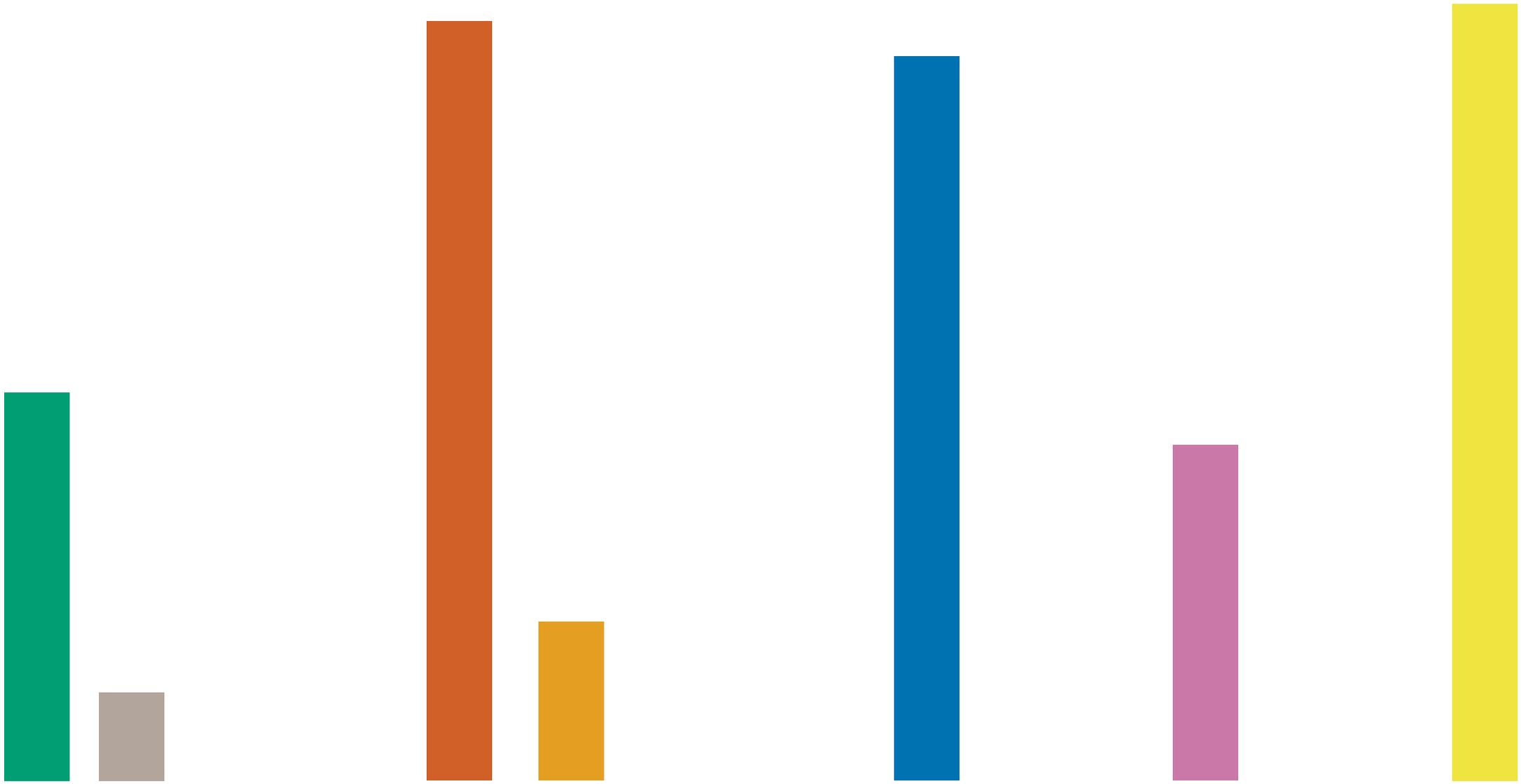
2. Stretch!

Lex & Gehlenborg, Nature Methods 11, 1 (2014)

Lex & Gehlenborg, Nature Methods 11, 1 (2014)

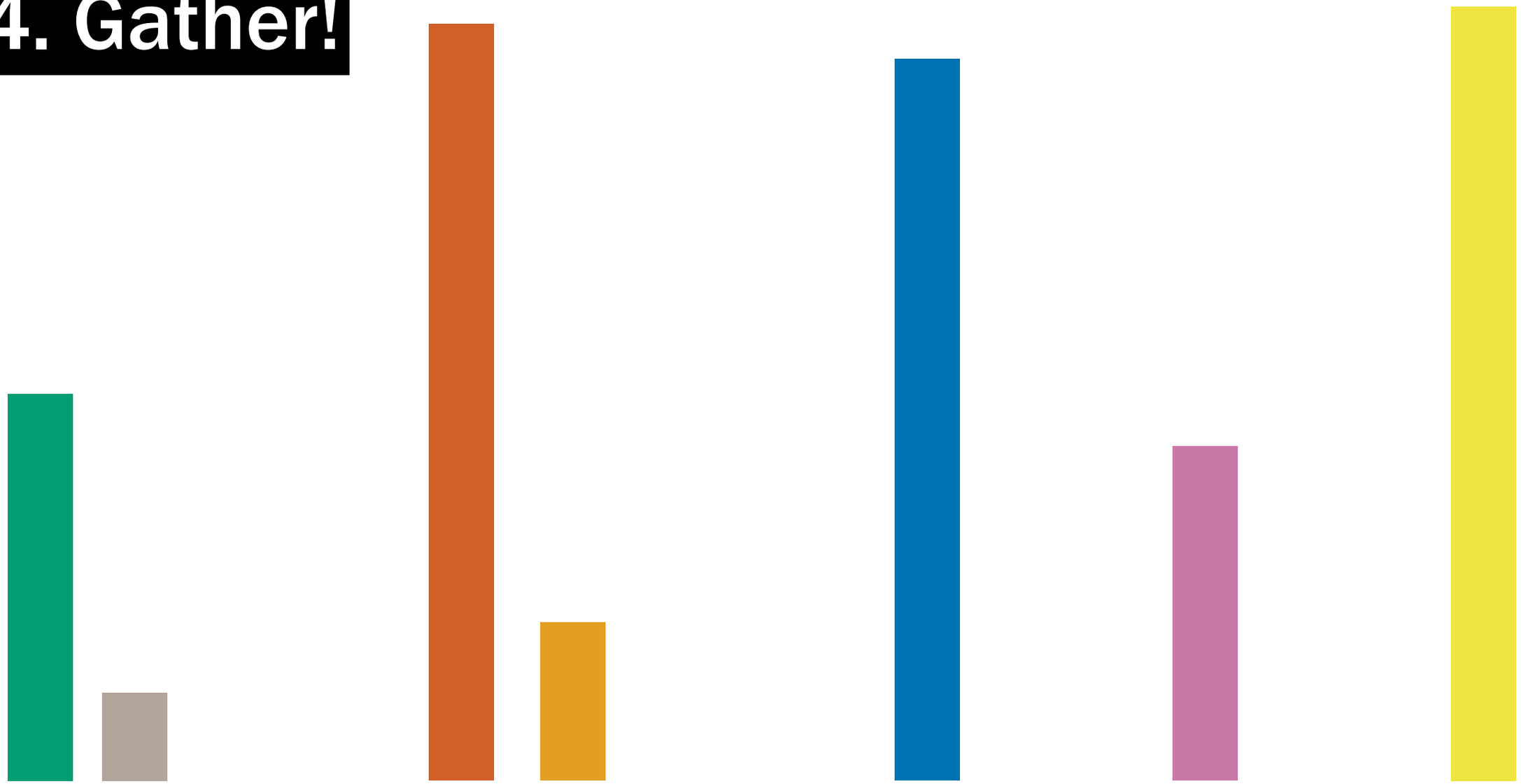
3. Align!

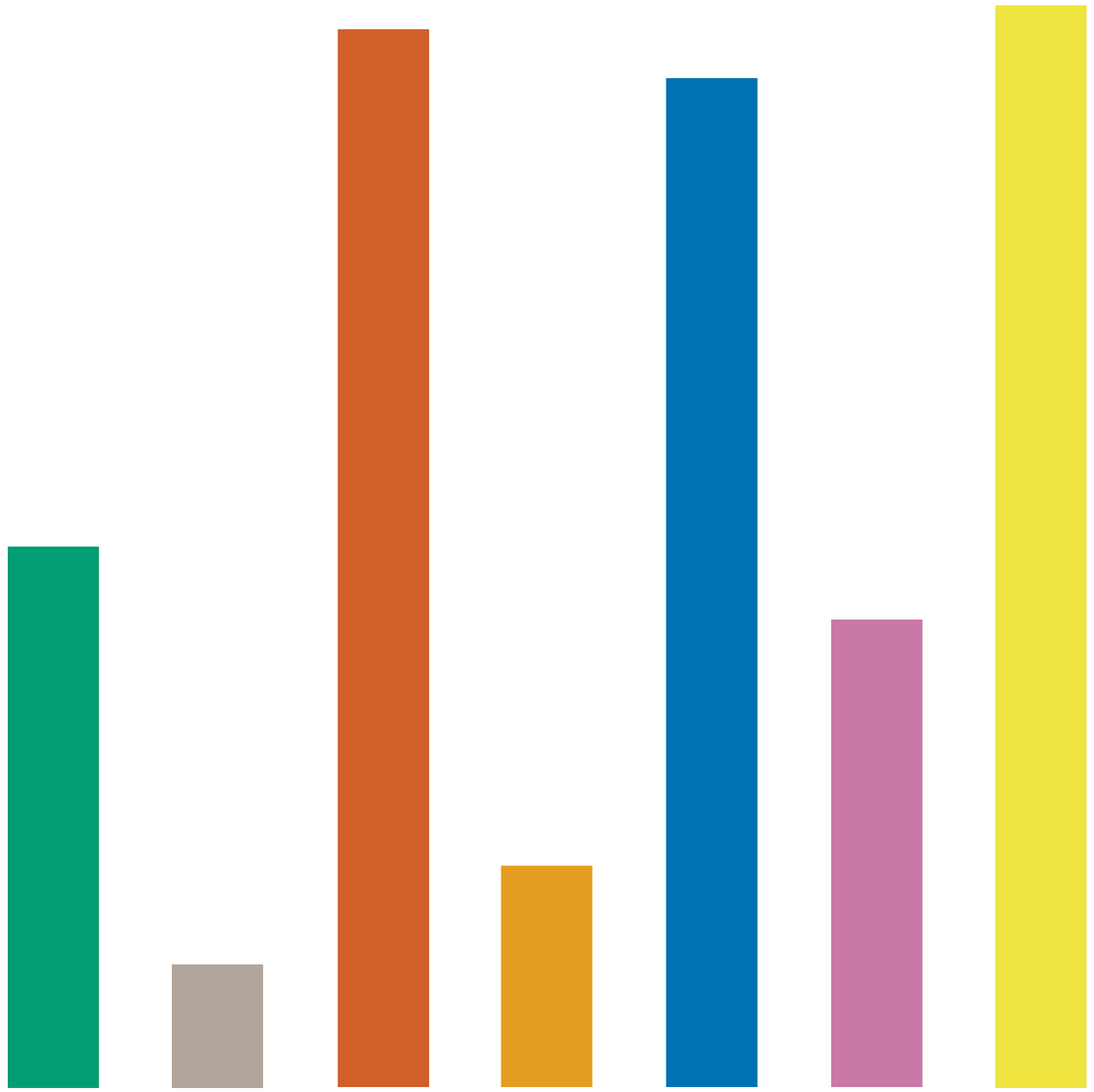
Lex & Gehlenborg, Nature Methods 11, 1 (2014)



Lex & Gehlenborg, Nature Methods 11, 1 (2014)

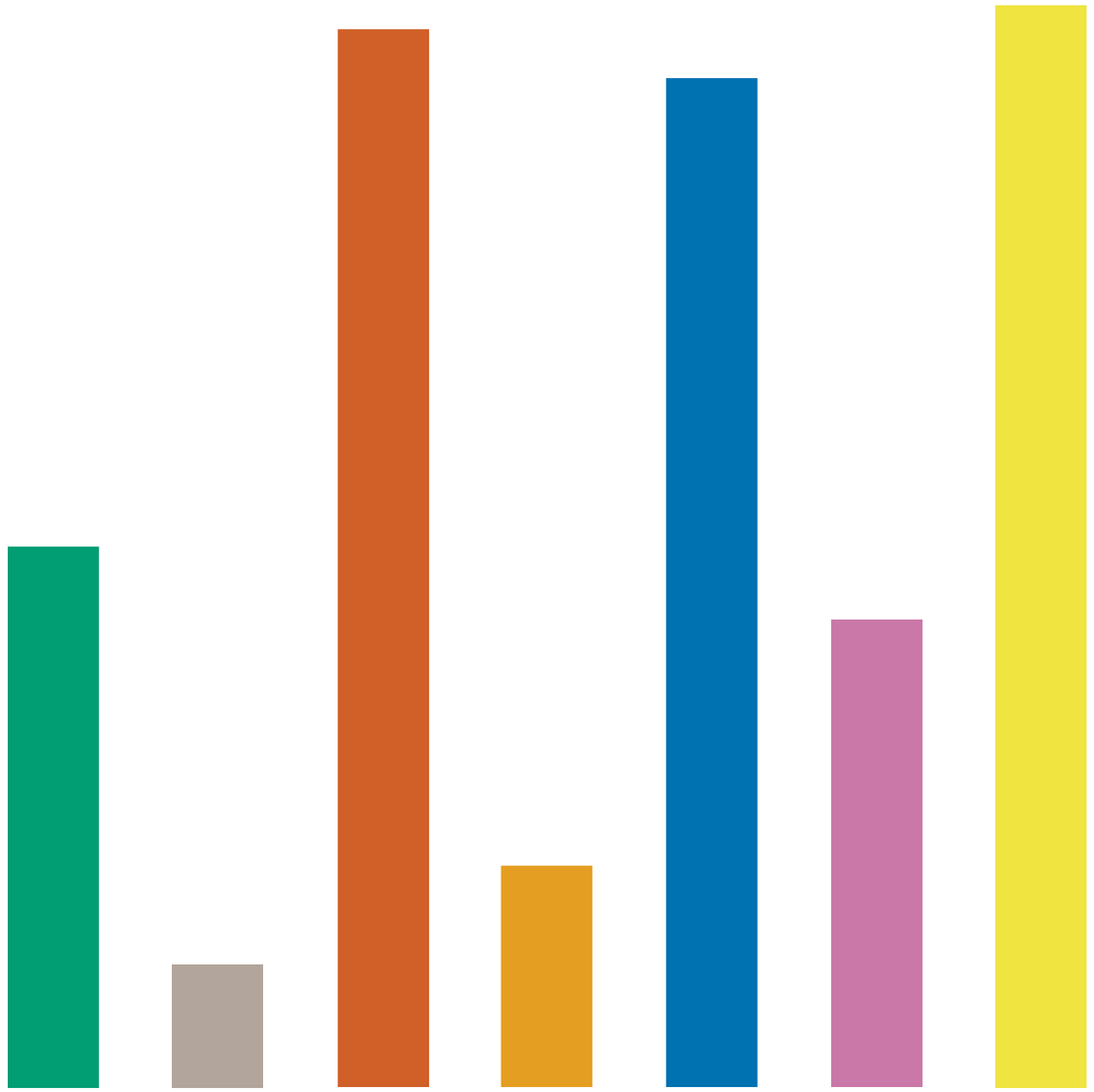
4. Gather!

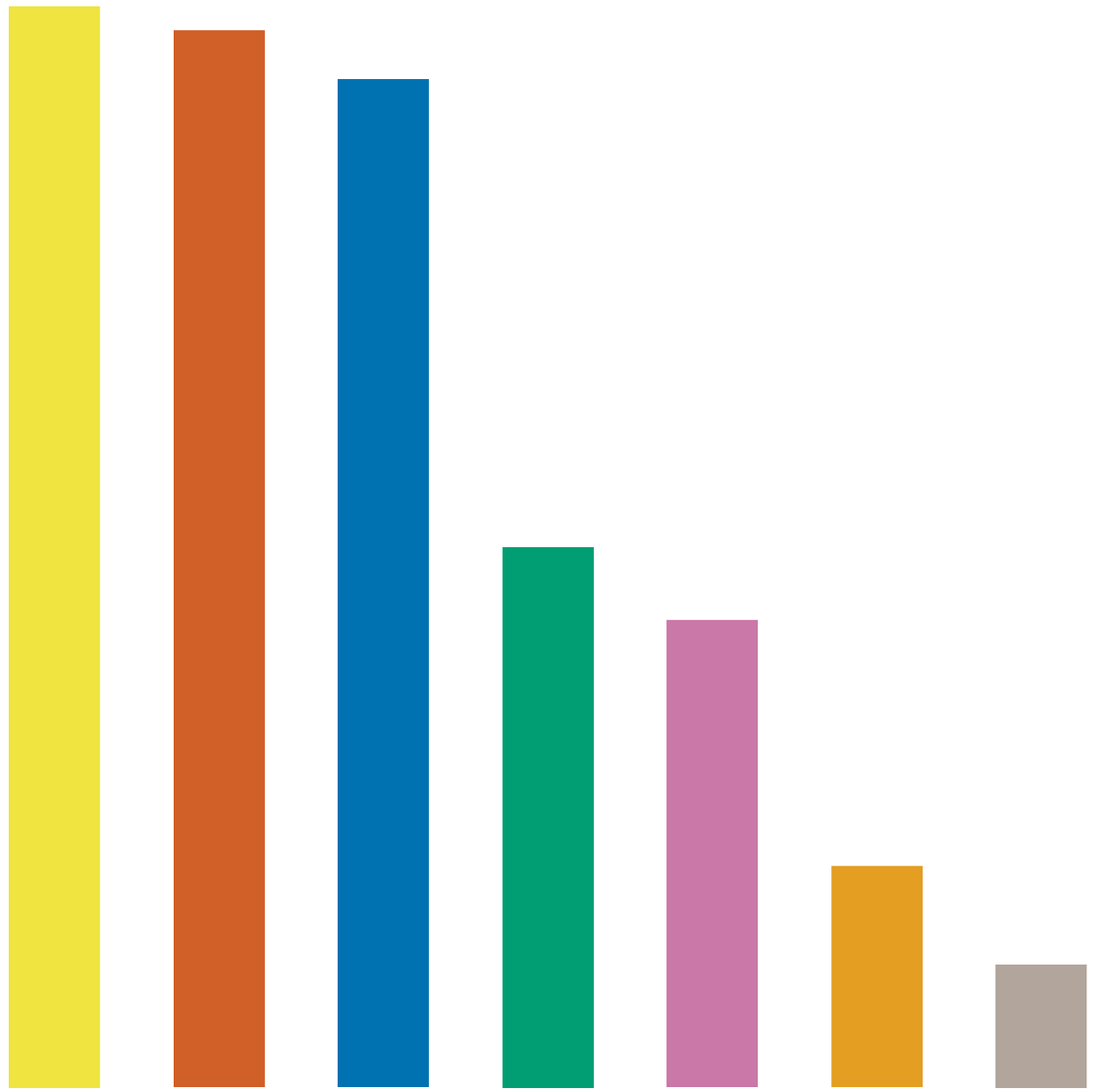




Lex & Gehlenborg, Nature Methods 11, 1 (2014)

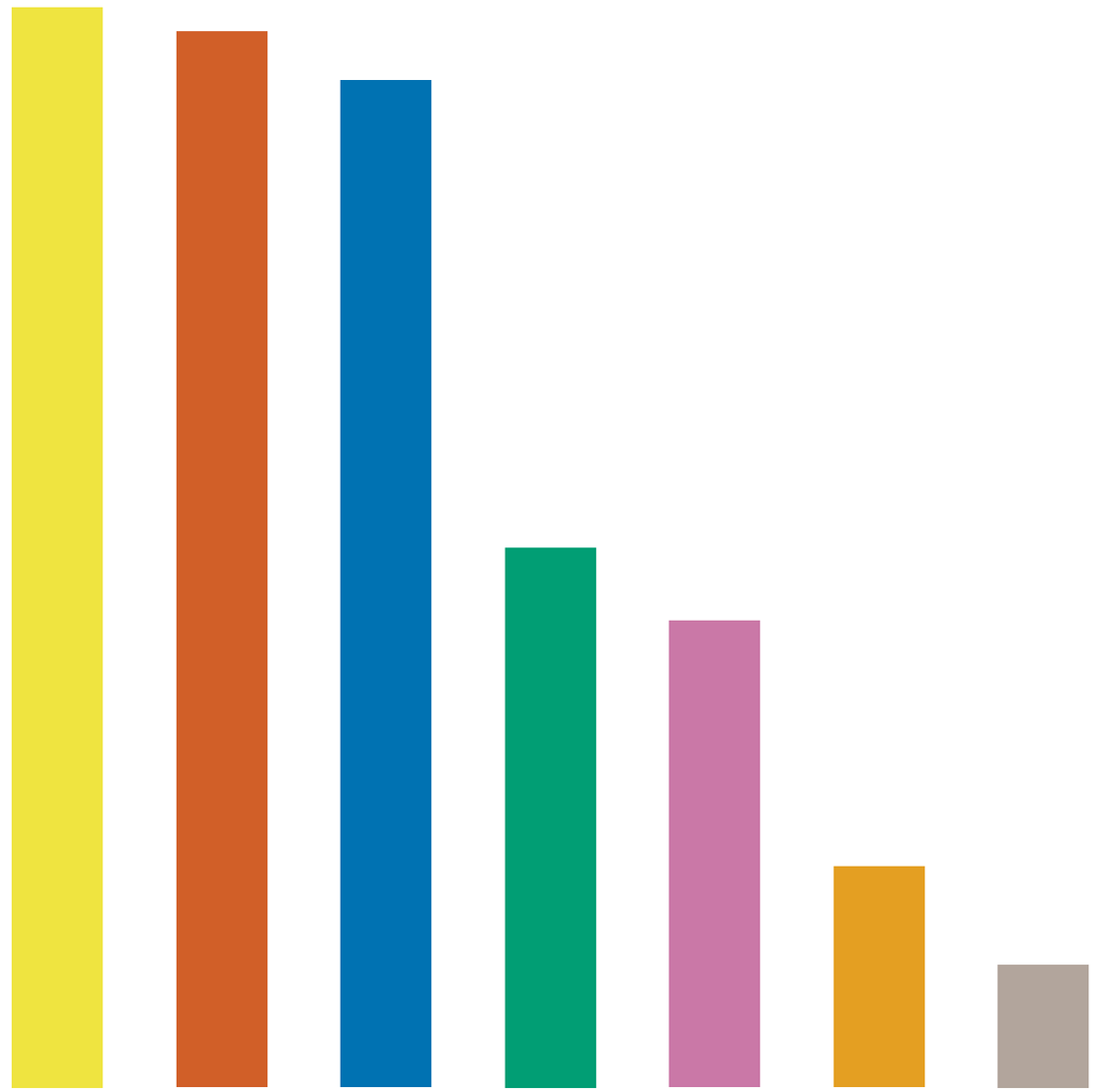
5. Sort!



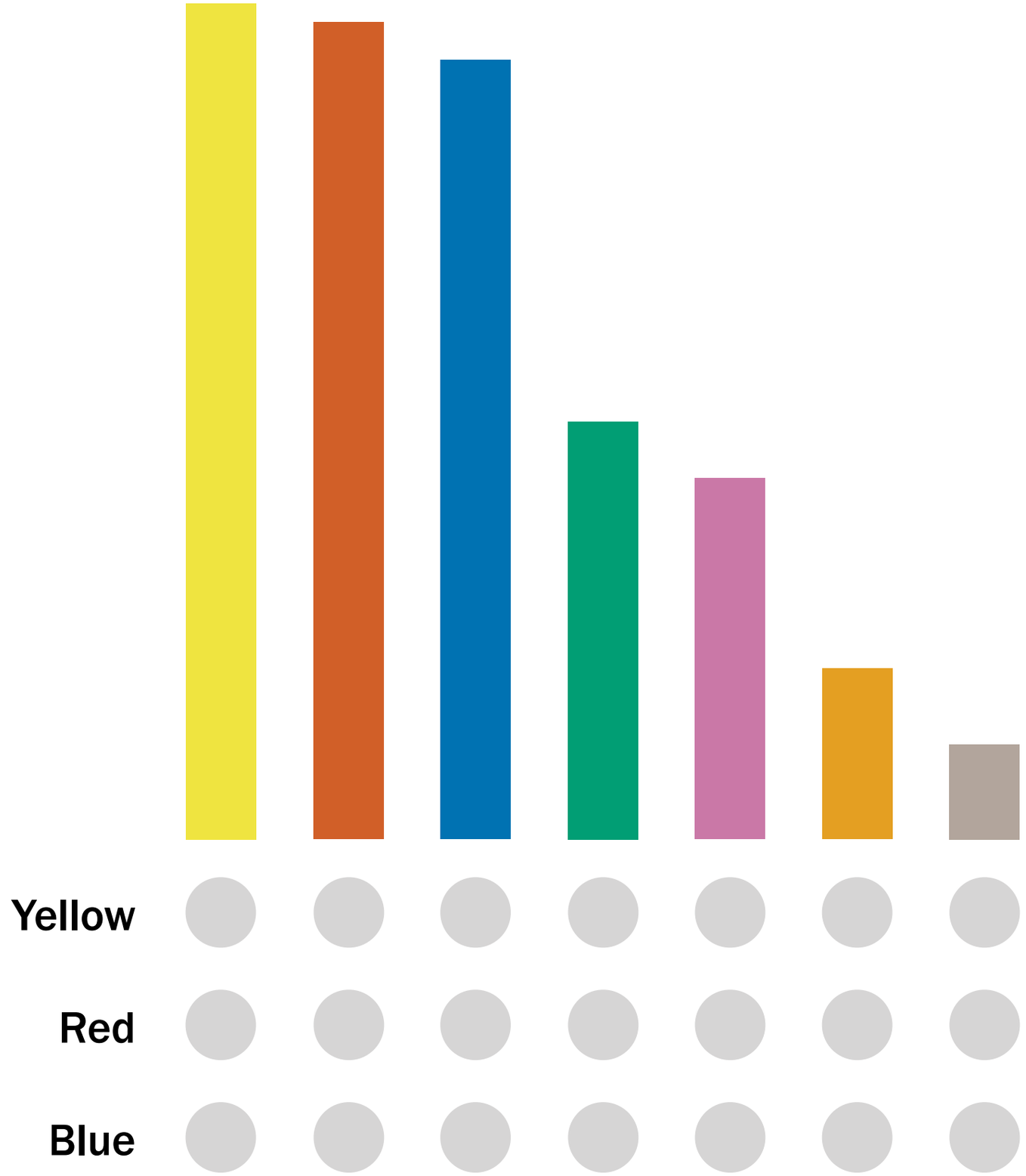


Lex & Gehlenborg, Nature Methods 11, 1 (2014)

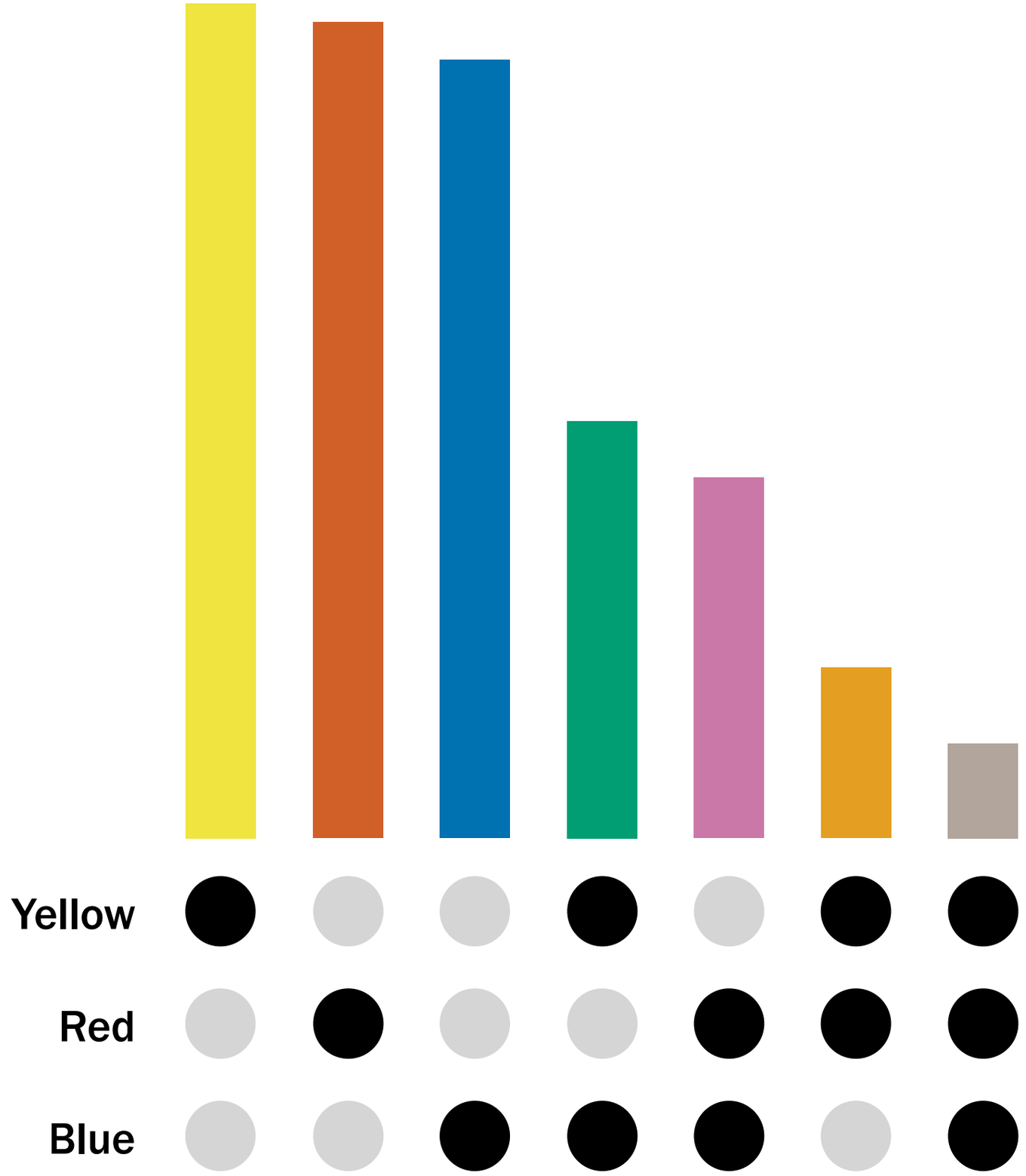
6. Let Go! (again)



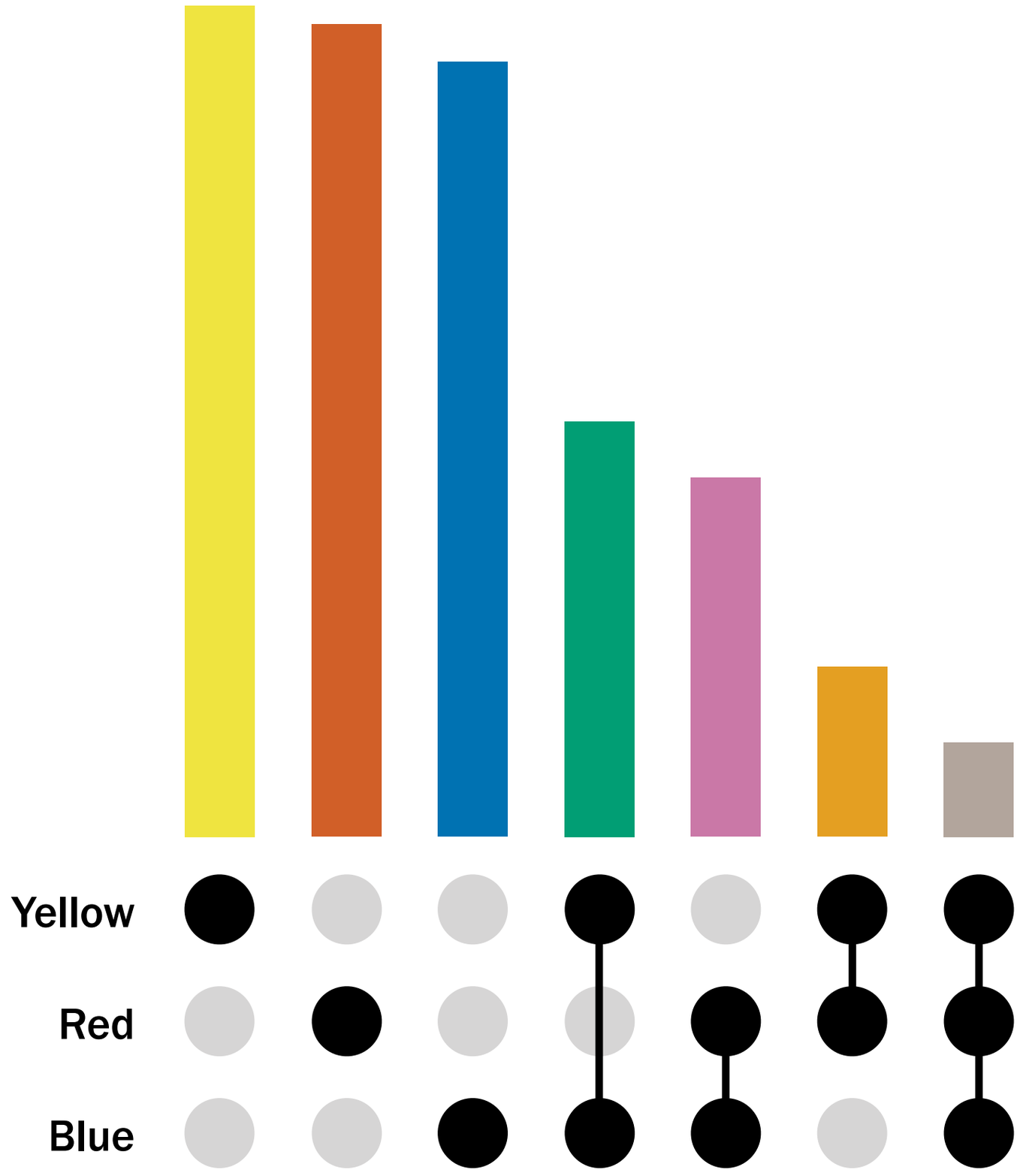
Lex & Gehlenborg, Nature Methods 11, 1 (2014)



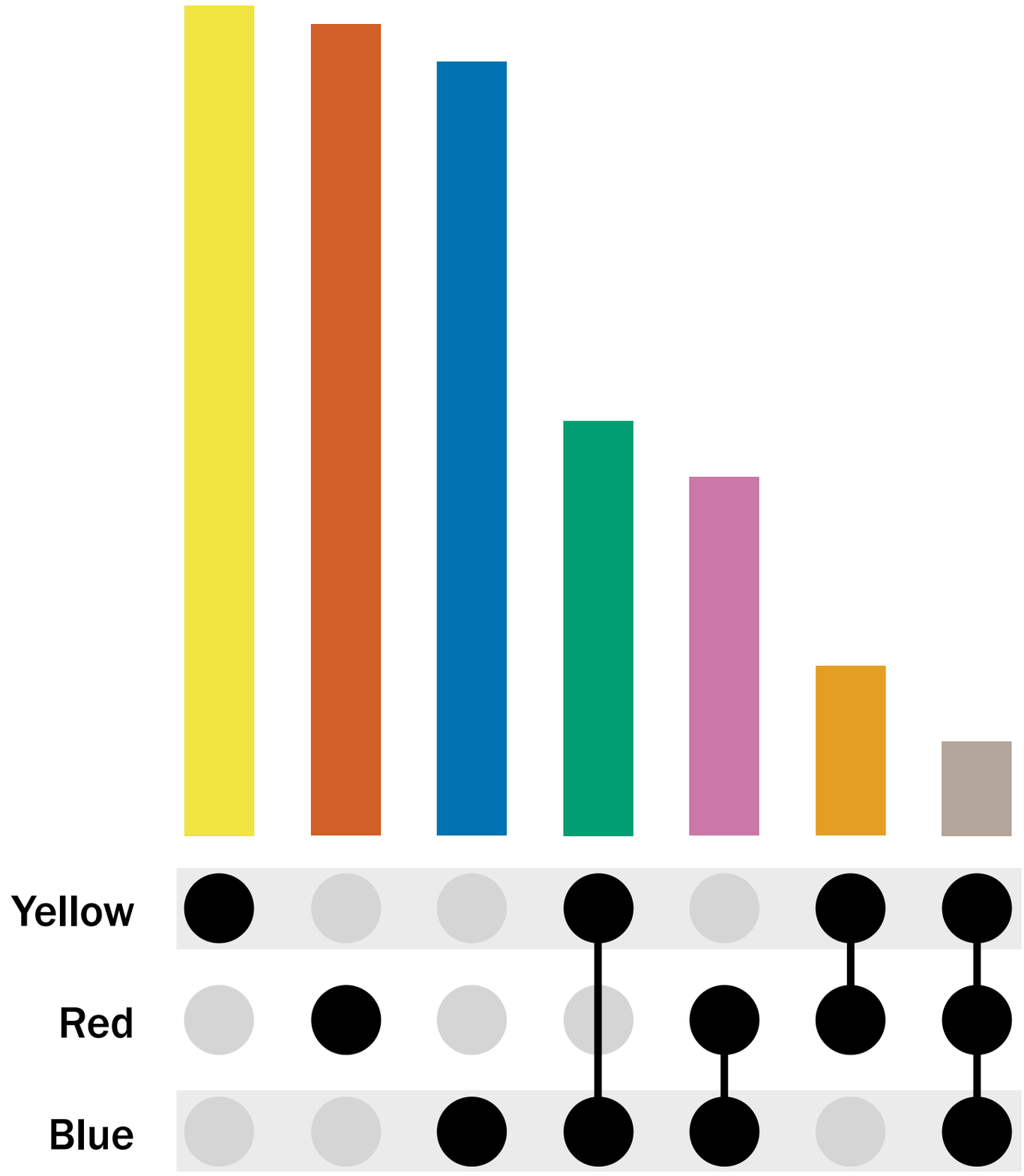
Lex & Gehlenborg, Nature Methods 11, 1 (2014)



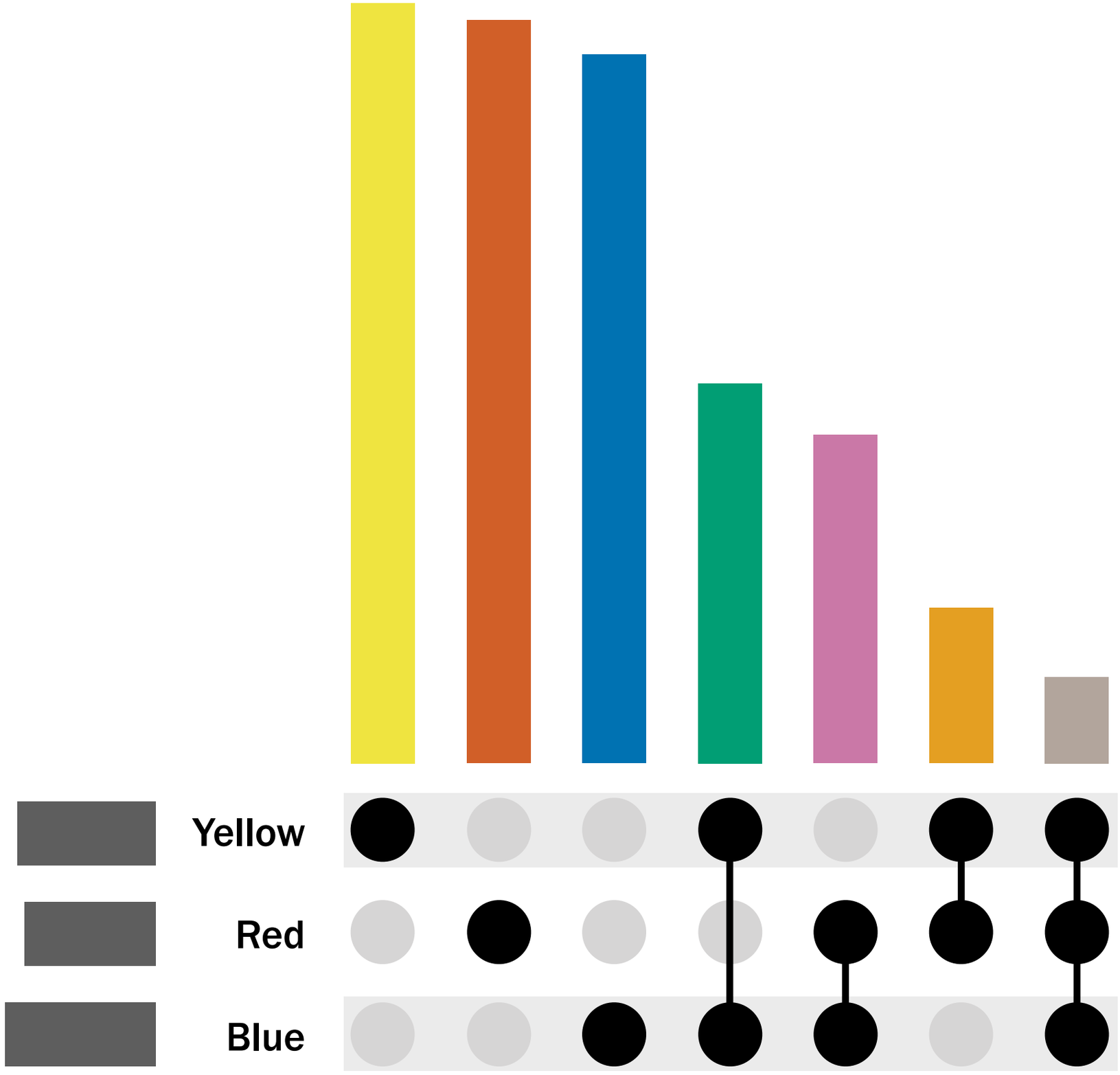
Lex & Gehlenborg, Nature Methods 11, 1 (2014)



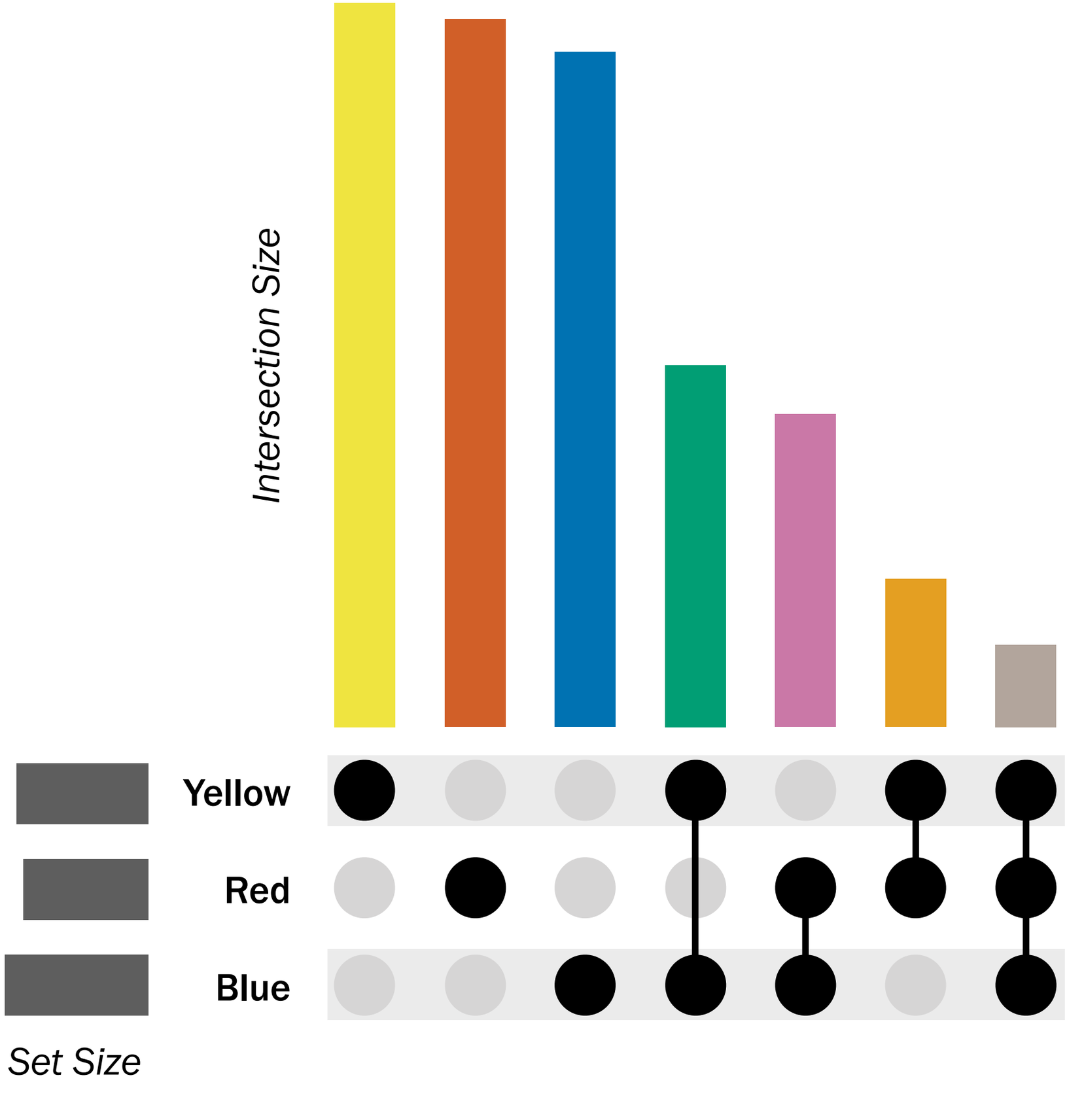
Lex & Gehlenborg, Nature Methods 11, 1 (2014)



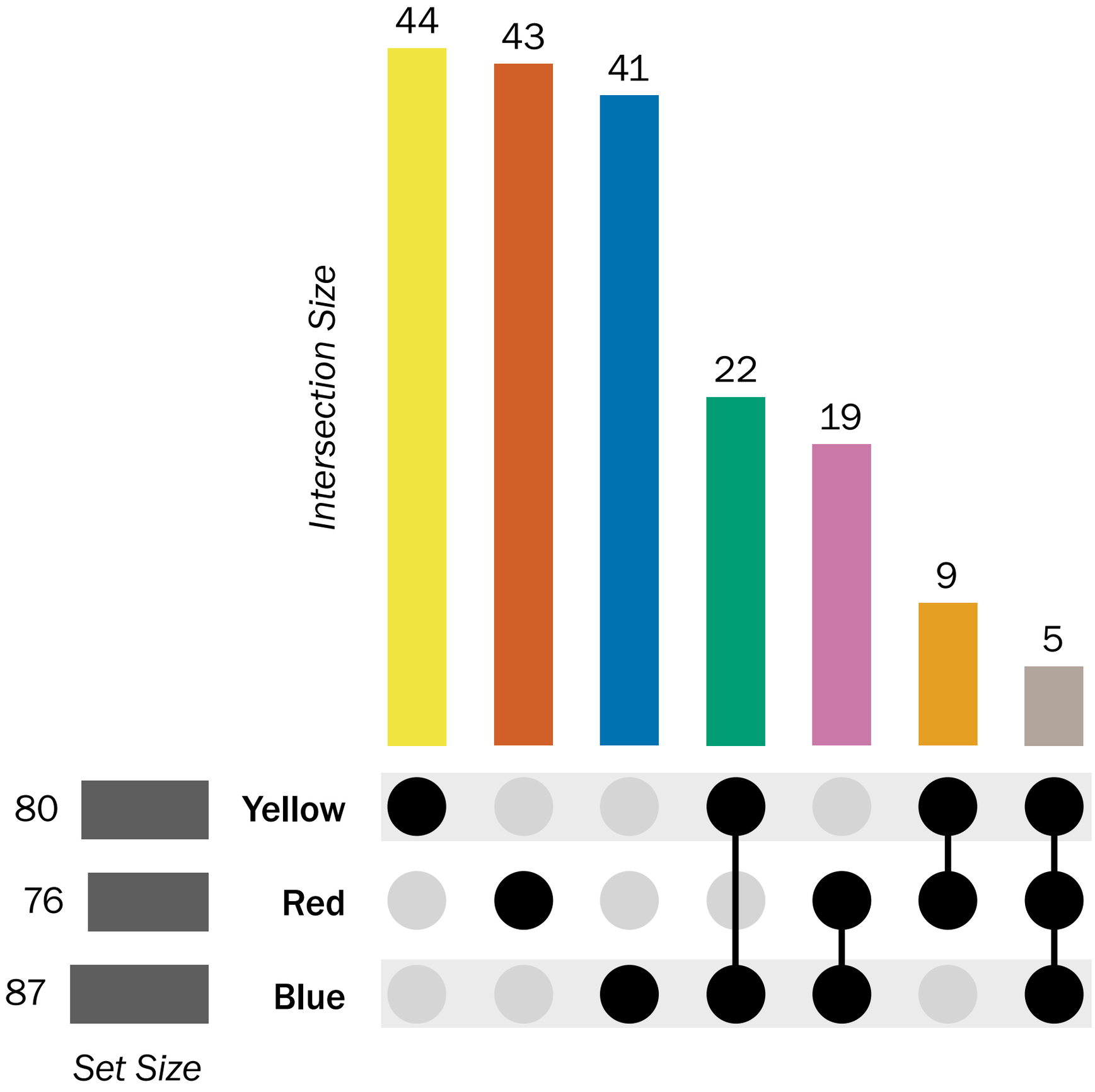
Lex & Gehlenborg, Nature Methods 11, 1 (2014)



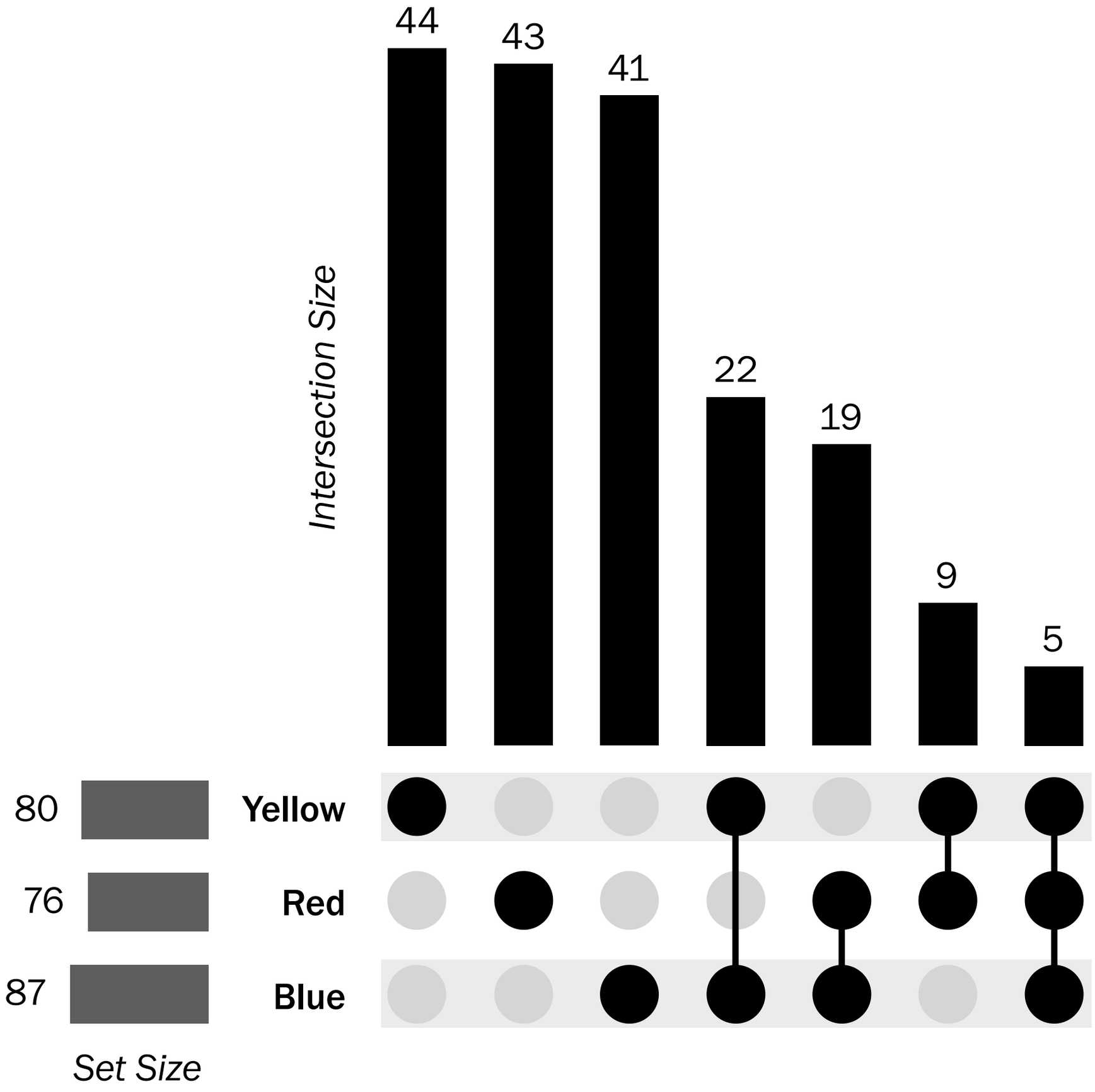
Lex & Gehlenborg, Nature Methods 11, 1 (2014)



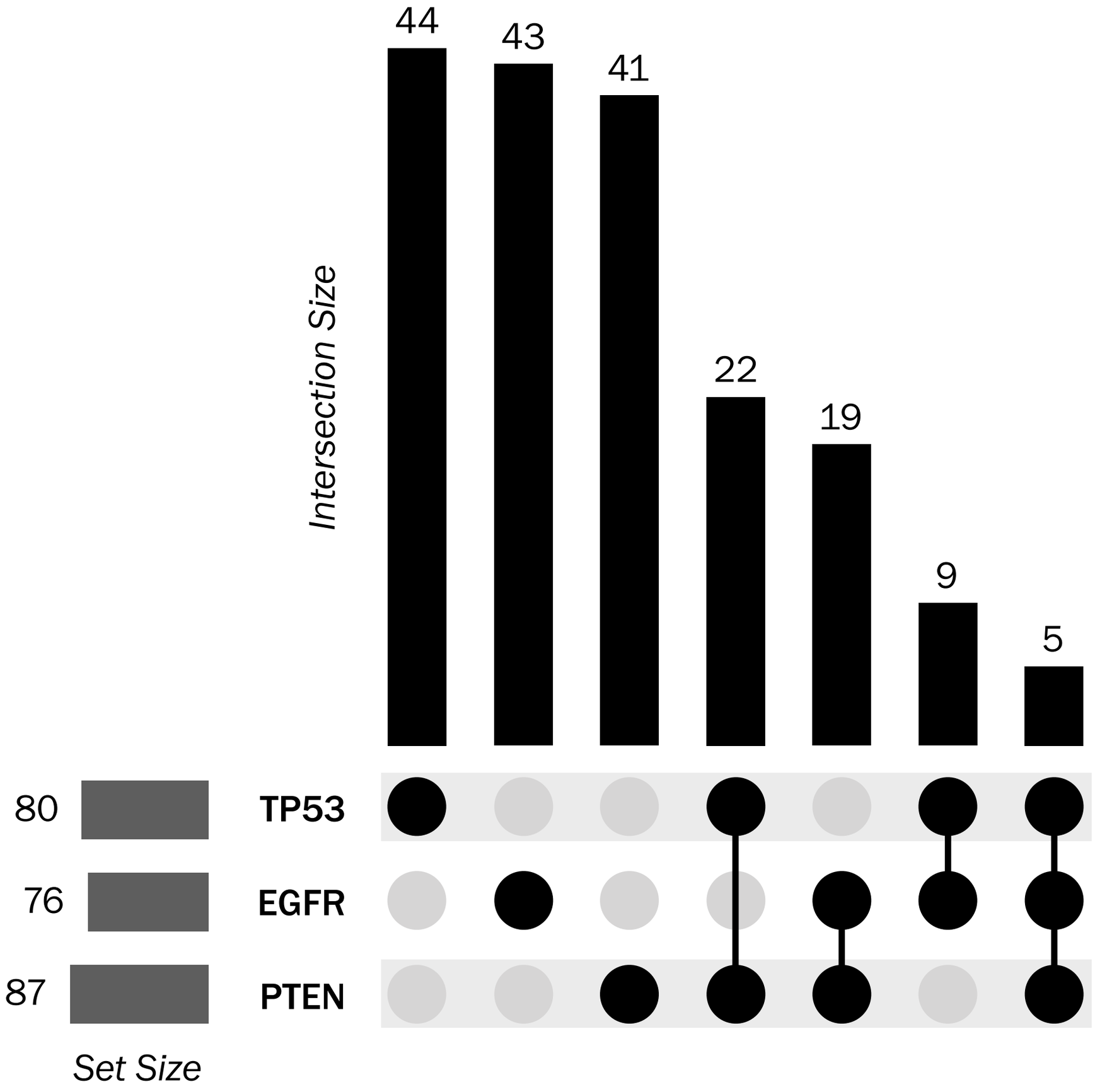
Lex & Gehlenborg, Nature Methods 11, 1 (2014)



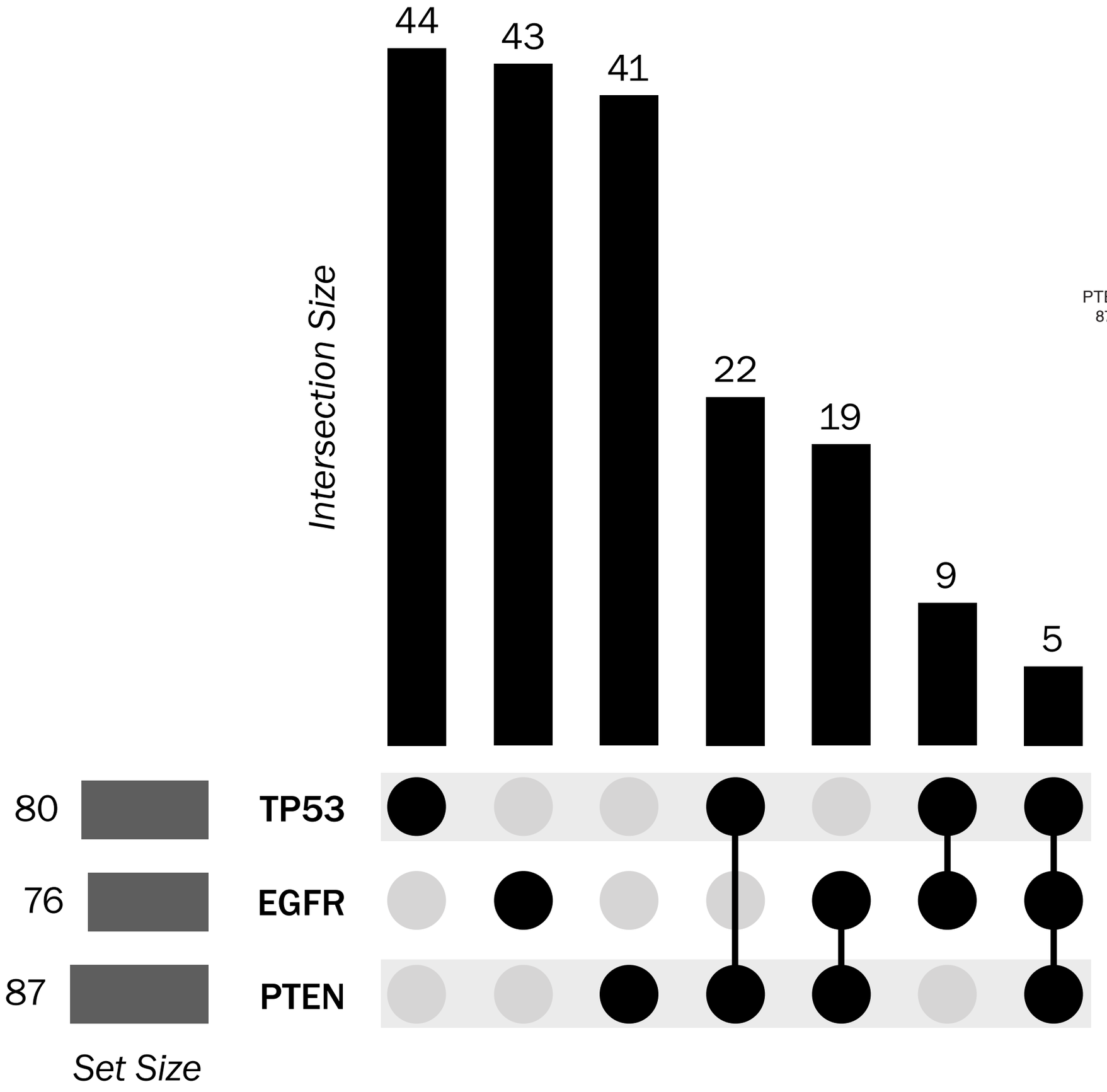
Lex & Gehlenborg, Nature Methods 11, 1 (2014)

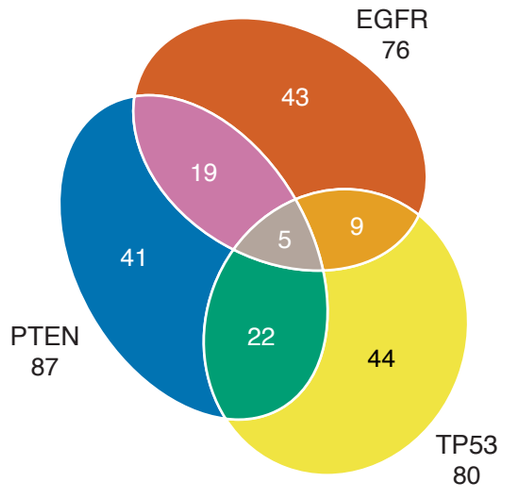


Lex & Gehlenborg, Nature Methods 11, 1 (2014)



Lex & Gehlenborg, Nature Methods 11, 1 (2014)





Hey! It's an UpSet Plot!

Bioinformatics, 2017, 1–3
doi: 10.1093/bioinformatics/btx364
Applications Note

Genome analysis

UpSetR: an R package for the visualization of intersecting sets and their properties

Jake R. Conway,[1] Alexander Lex[2] and Nils Gehlenborg[1,*]

[1]Department of Biomedical Informatics, Harvard Medical School, Boston, MA 02115, USA and [2]SCI Institute, School of Computing, University of Utah, Salt Lake City, UT 84112, USA

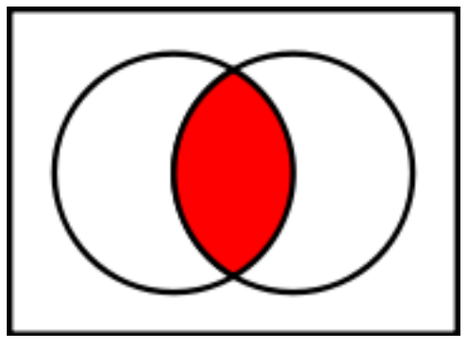
*To whom correspondence should be addressed.
Associate Editor: John Hancock

Received on March 25, 2017; revised on May 19, 2017; editorial decision on May 31, 2017; accepted on June 5, 2017

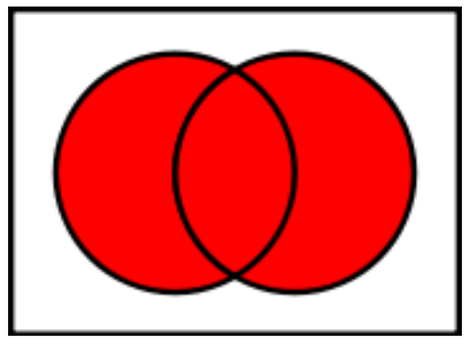
Disclaimer

http://en.wikipedia.org/wiki/Venn_diagram

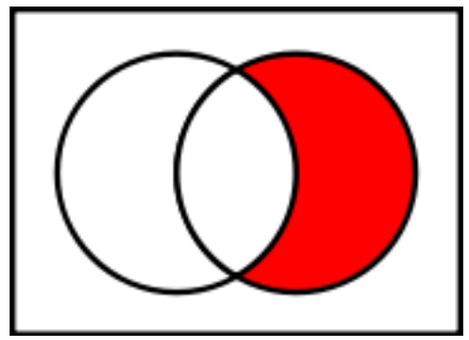
Venn Diagrams are actually very useful!



$A \cap B$



$A \cup B$



$B \setminus A$

HARVARD MEDICAL SCHOOL
DEPARTMENT OF BIOMEDICAL INFORMATICS

I am hiring postdocs!

Data visualization, analysis, and management for:

- genomic structural variants
- dynamics of the 3D genome
- cancer subtypes in patient cohorts
- exploration tools for data repositories
- provenance graphs

See http://gehlenborglab.org or http://dbmi.med.harvard.edu for details.