

# designing effective visualizations in the biological sciences

## BIOINFORMATICS SEMINAR

**MARTIN KRZYWINSKI**

Michael Smith Genome Sciences Center  
BC Cancer Research Center  
Vancouver, Canada

GENOME SCIENCES CENTER

BC Cancer Research Center  
26 Aug 2011



BC Cancer Agency  
CARE + RESEARCH



CANADA'S MICHAEL SMITH  
GENOME  
SCIENCES  
CENTRE

# design : effective visual communication

## OBJECTIVE

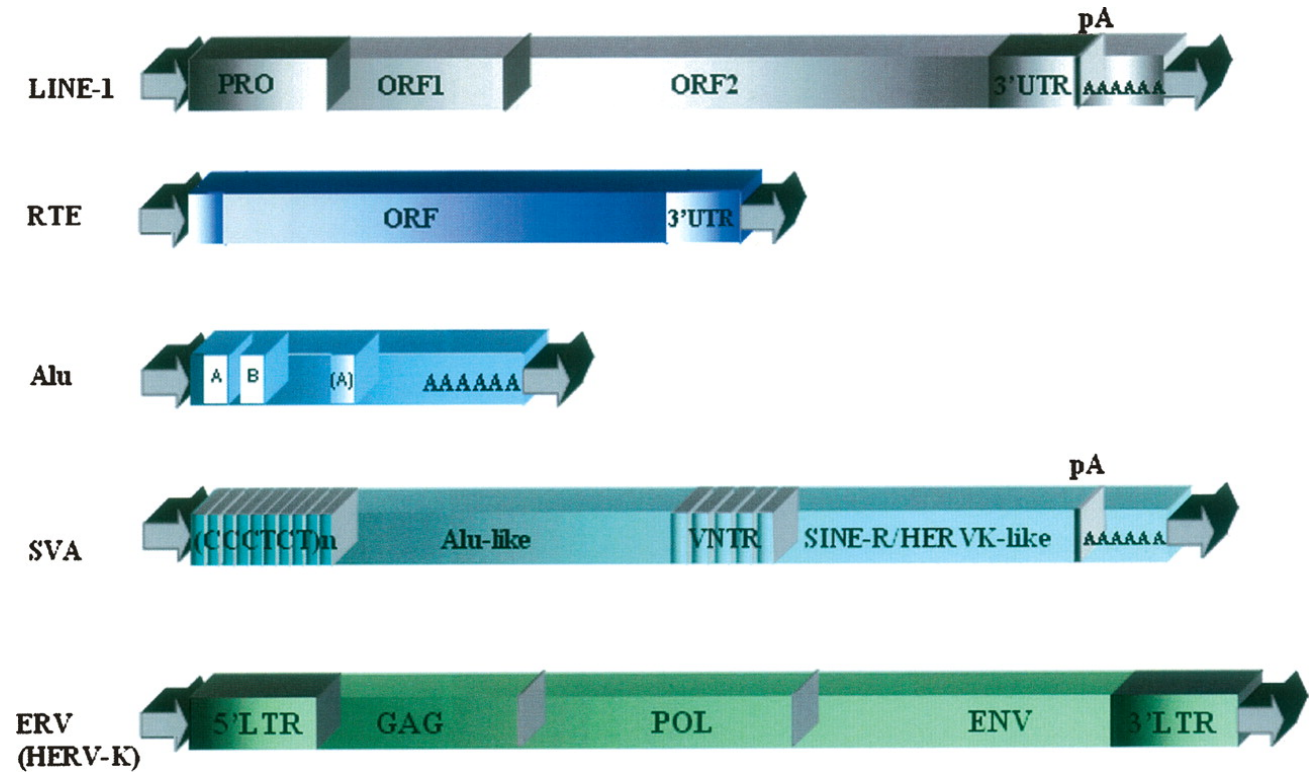
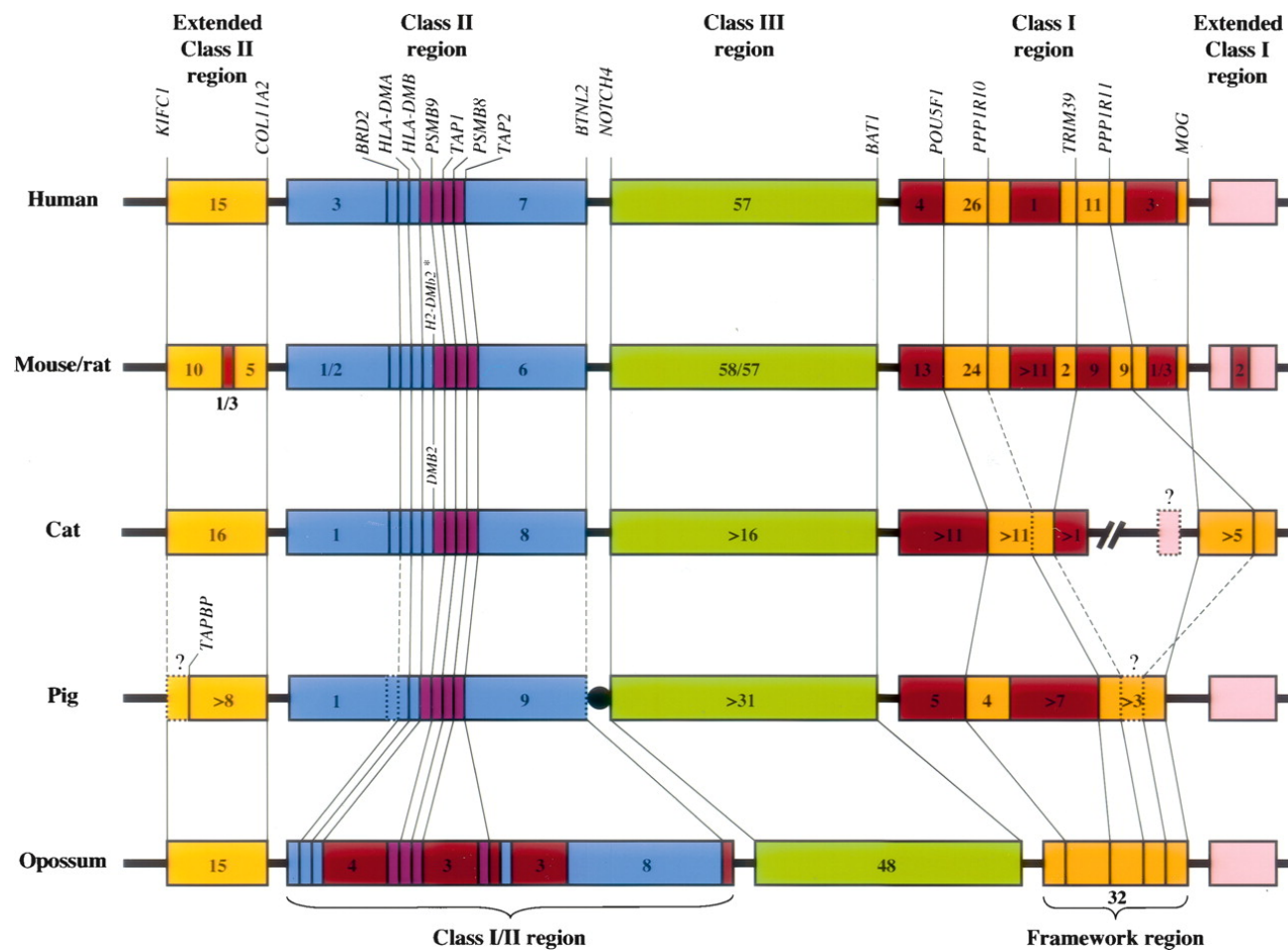
FOLLOWING GUIDELINES

## SUBJECTIVE

BREAKING GUIDELINES PURPOSEFULLY



# OBJECTIVE ASPECTS OF CLEAR COMMUNICATION



Excellent organization and consistency. Clear use of color. Vertical lines cue continuity.

Samollow, P.B., The opossum genome: insights and opportunities from an alternative mammal. *Genome Res*, 2008. 18(8): p. 1199-215.

Overly ornamental. Illegible and inconsistently formatted text. Redundant elements.

Gentles, A.J., et al., Evolutionary dynamics of transposable elements in the short-tailed opossum *Monodelphis domestica*. *Genome Res*, 2007. 17(7): p. 992-1004.



# OBJECTIVE ASPECTS OF ATTRACTION

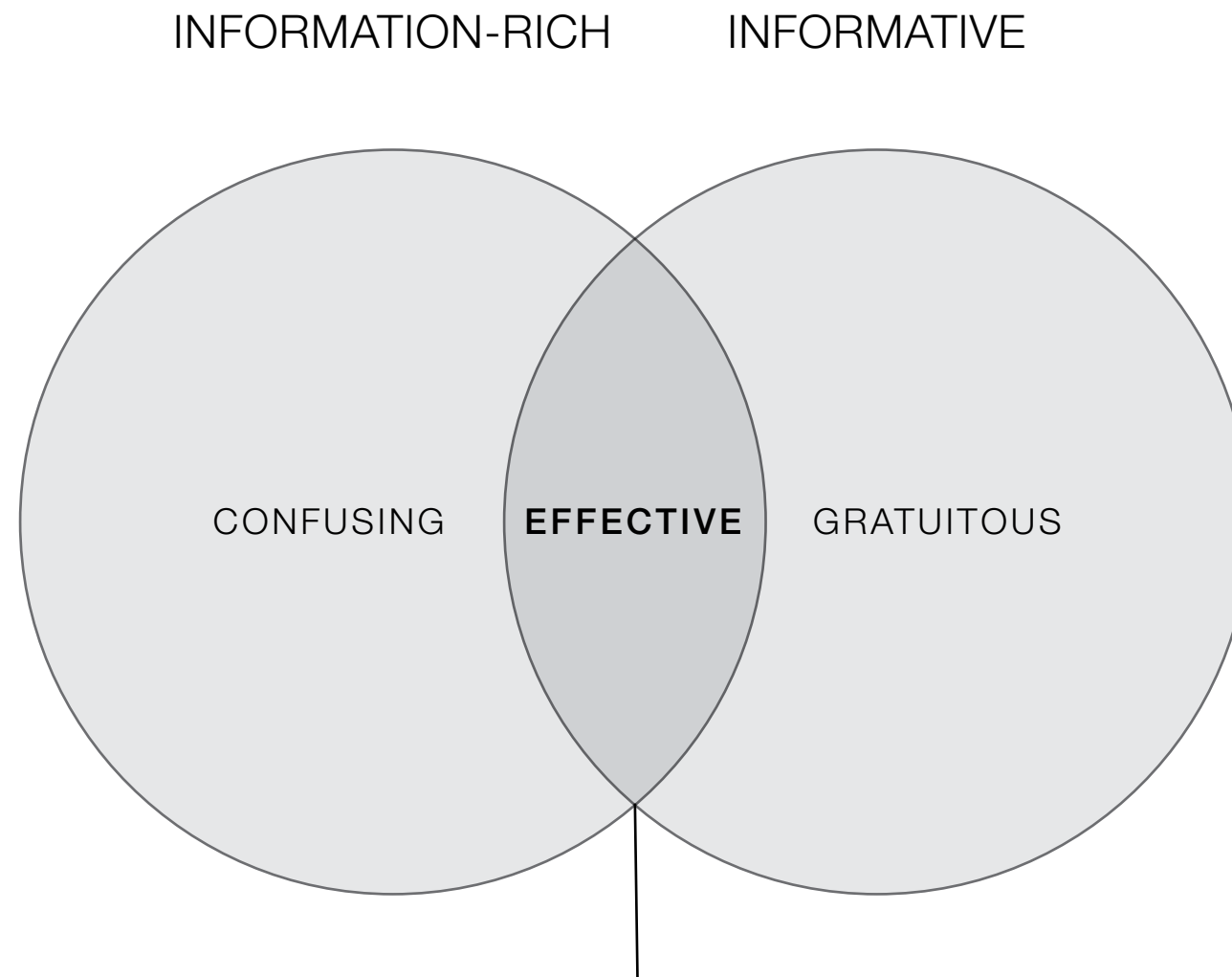


Healthy coat. Pleasant colour scheme. Vigorously adorable proportions. Attentive gaze. Sympathetic eyes.



Where do I start?

# WHAT IS AN EFFECTIVE VISUALIZATION?



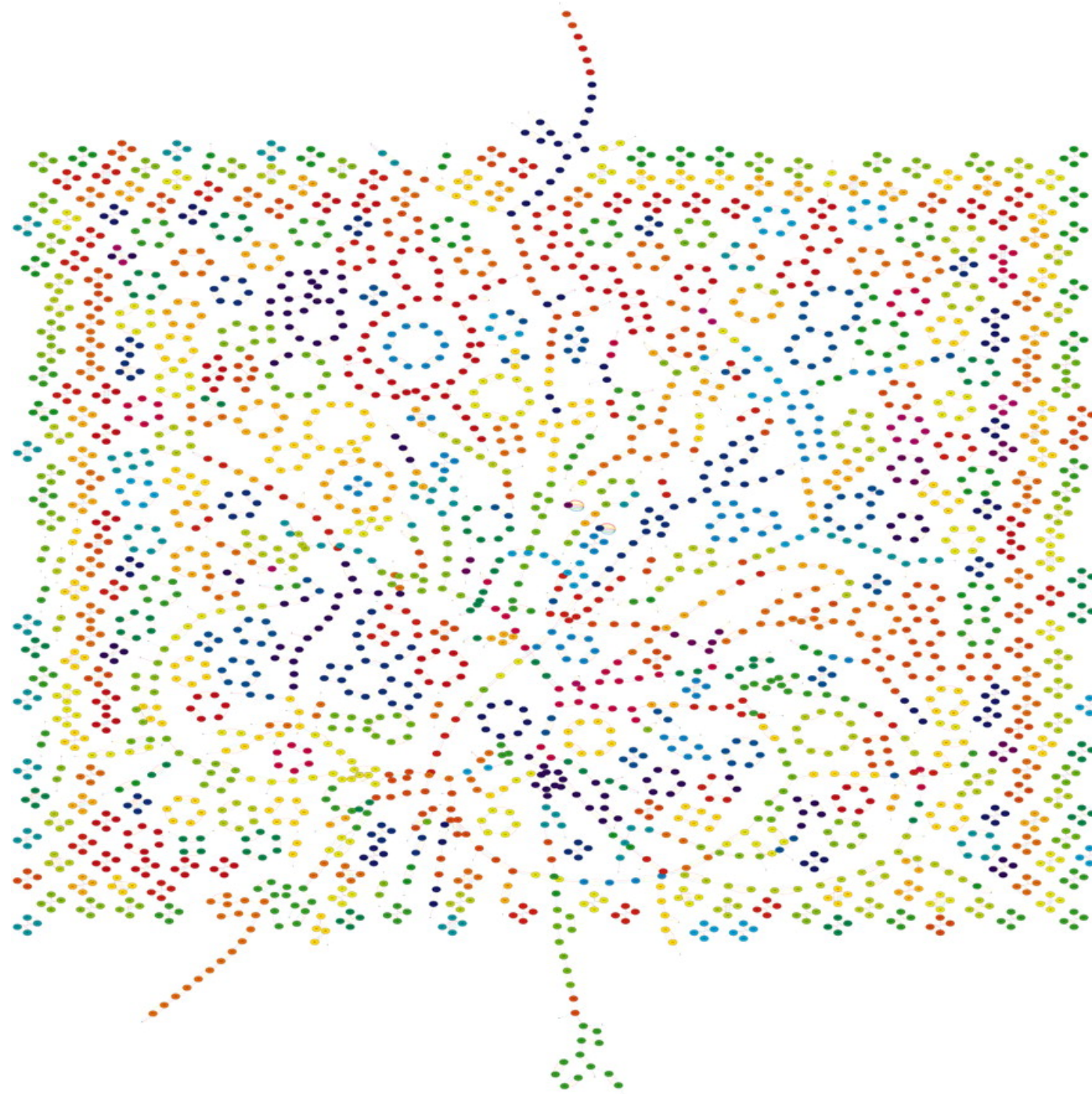
**CLEAR MESSAGE**

**HIGH DATA-TO-INK RATIO**

**ACCESSIBLE COMPLEXITY**



# INFORMATION-RICH, NOT INFORMATIVE



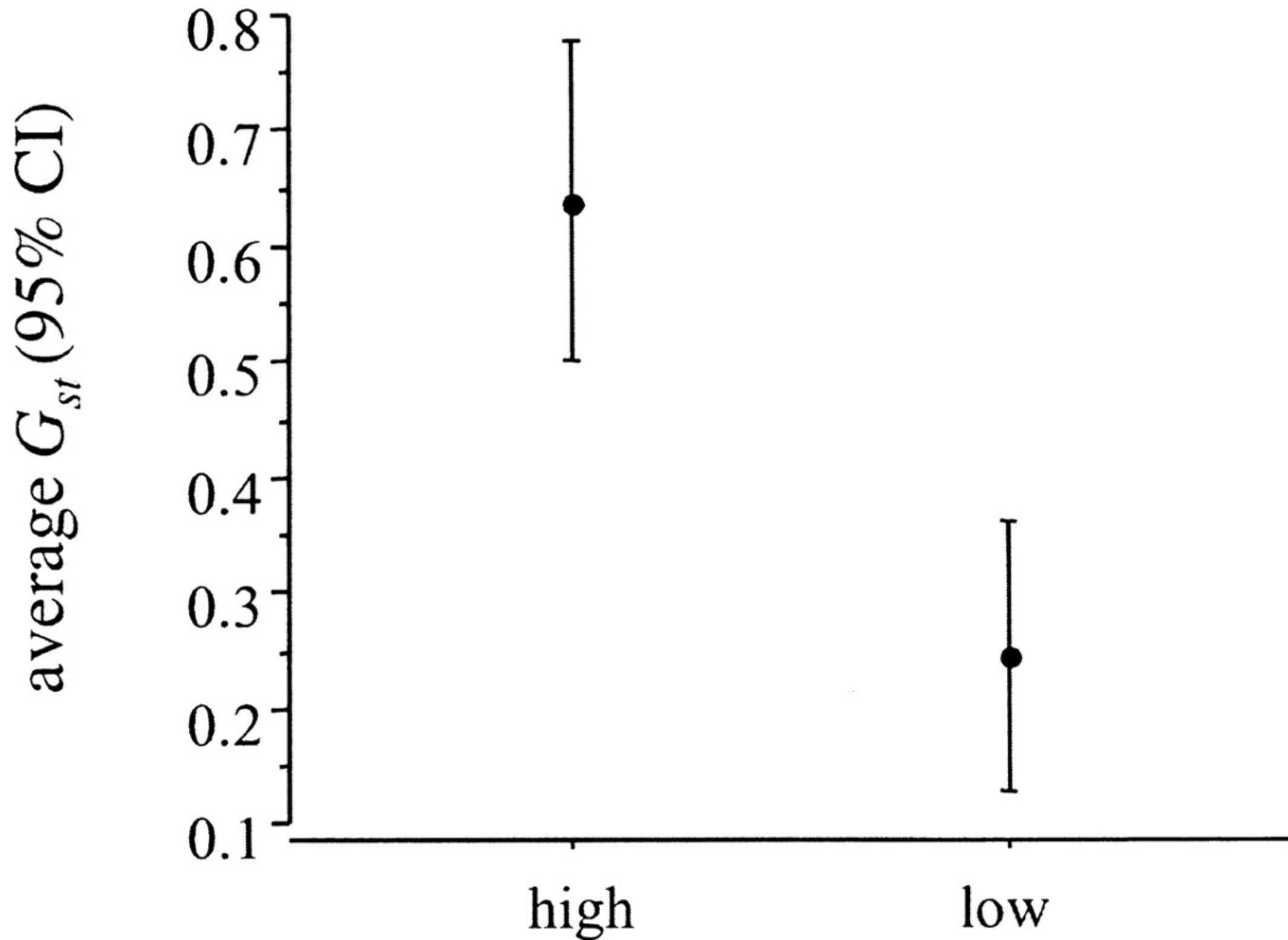
Chromosome colors:



The breakpoint graph  $G(M,R,D,Q,H,C)$  (obverse edges are not shown) of six mammalian genomes. Alekseyev, M.A. and P.A. Pevzner, Breakpoint graphs and ancestral genome reconstructions. *Genome Res*, 2009. 19(5): p. 943-57.

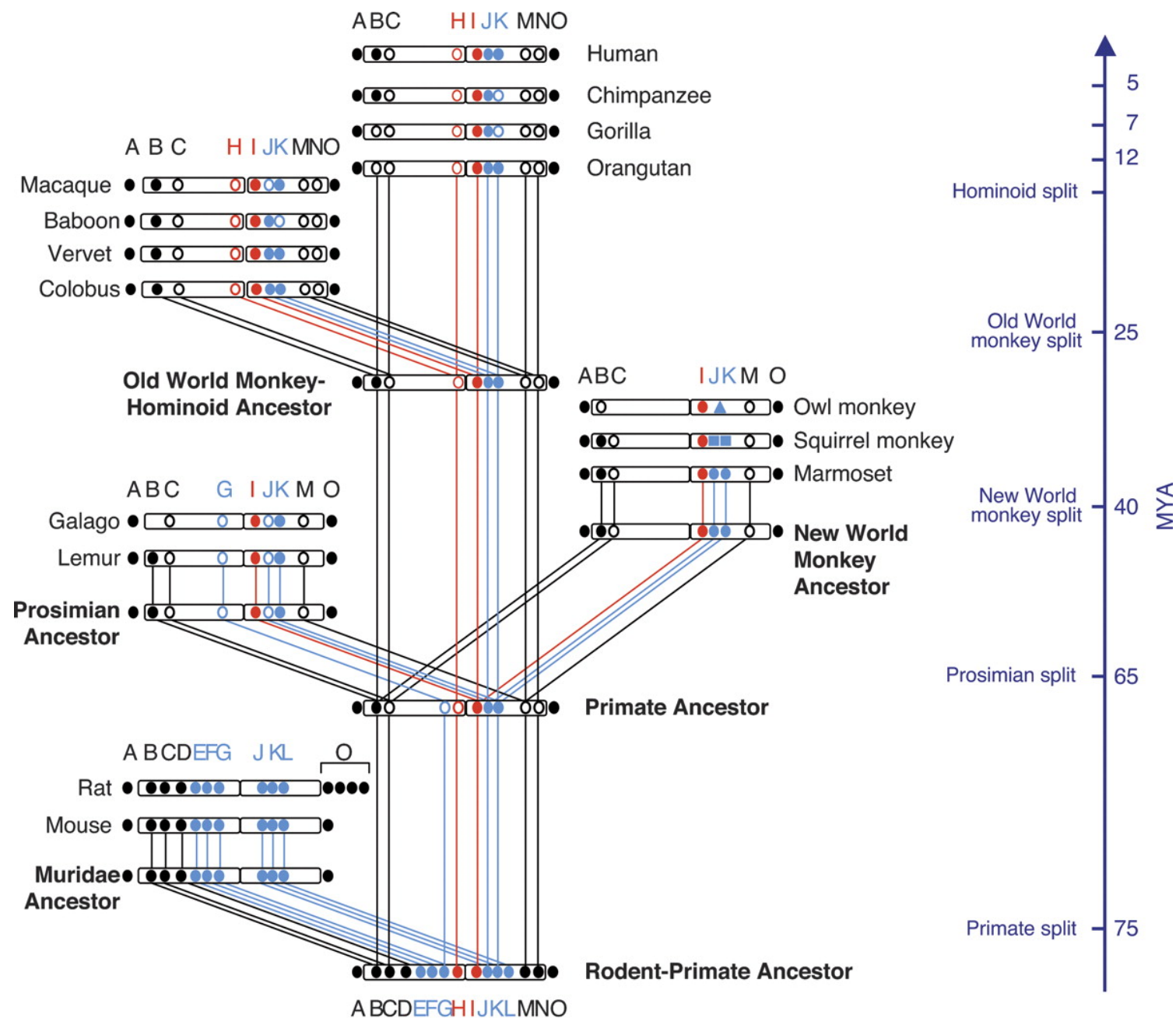


# INFORMATIVE, NOT INFORMATION-RICH



Mean and 95% confidence interval of average  $G_{st}$  based on direct sequencing of fragments within high (N = 8) and low (N = 6) differentiation regions in natural population of house mice. Harr, B., Genomic islands of differentiation between house mouse subspecies. *Genome Res*, 2006. 16(6): p. 730-7.

# INFORMATION-RICH AND INFORMATIVE



<b>A</b> <i>WFDC5/Wfdc5</i>	<b>D</b> <i>WFDC15a/Wfdc15a</i>	<b>M</b> <i>WFDC15c/Wfdc15c</i>
<b>B</b> <i>WFDC12/Wfdc12</i>	<b>E, F, G, J, K, L</b> <i>SEMG/Svs family</i>	<b>N</b> <i>WFDC15d/Wfdc15d</i>
<b>C</b> <i>WFDC15b/Wfdc15b</i>	<b>H, I</b> <i>Trappin (PI3) family</i>	<b>O</b> <i>SLPI/Slpi family</i>

Evolutionary history of the centromeric WFDC sublocus. Hurle, B., W. Swanson, and E.D. Green, Comparative sequence analyses reveal rapid and divergent evolutionary changes of the WFDC locus in the primate lineage. *Genome Res*, 2007. 17(3): p. 276-86.

# requirements for effective visual communication

**LEGIBILITY**

**CLARITY**

AND OPTIONALLY,

**ATTRACTIVENESS**





# LEGIBLE

are all elements  
discernable?

does text contrast well  
with background?

is there simultaneous  
contrast?

RESOLUTION  
PARSABILITY  
COLOR



# LEGIBLE

are all elements discernable?

does text contrast well with background?

is there simultaneous contrast?

RESOLUTION  
PARSABILITY  
COLOR

# CLEAR

does the reader see the noise or trends?

message delivered?

are there redundant or ornamental elements?

STRONG MESSAGE  
REDUNDANCY  
REPRESENTATION  
CONSISTENCY  
EXCESS INK  
GLYPHS





# LEGIBLE

are all elements discernable?

does text contrast well with background?

is there simultaneous contrast?

RESOLUTION  
PARSABILITY  
COLOR

# CLEAR

does the reader see the **noise** or **trends**?

message delivered?

are there redundant or ornamental elements?

STRONG MESSAGE  
REDUNDANCY  
REPRESENTATION  
CONSISTENCY  
EXCESS INK  
GLYPHS

# ATTRACTIVE

are spacing and orientation of visual cues consistent?

GRID LAYOUT

# LEGIBLE

are all elements discernable?

does text contrast well with background?

is there simultaneous contrast?

# CLEAR

does the reader see the noise or trends?

message delivered?

are there redundant or ornamental elements?

# ATTRACTIVE

are spacing and orientation of visual cues consistent?

DO I HAVE A MESSAGE ?    HAVE I COMMUNICATED IT ?

RESOLUTION  
PARSABILITY  
COLOR

STRONG MESSAGE  
REDUNDANCY  
REPRESENTATION  
CONSISTENCY  
EXCESS INK  
GLYPHS

GRID LAYOUT

**legibility** clarity attractiveness

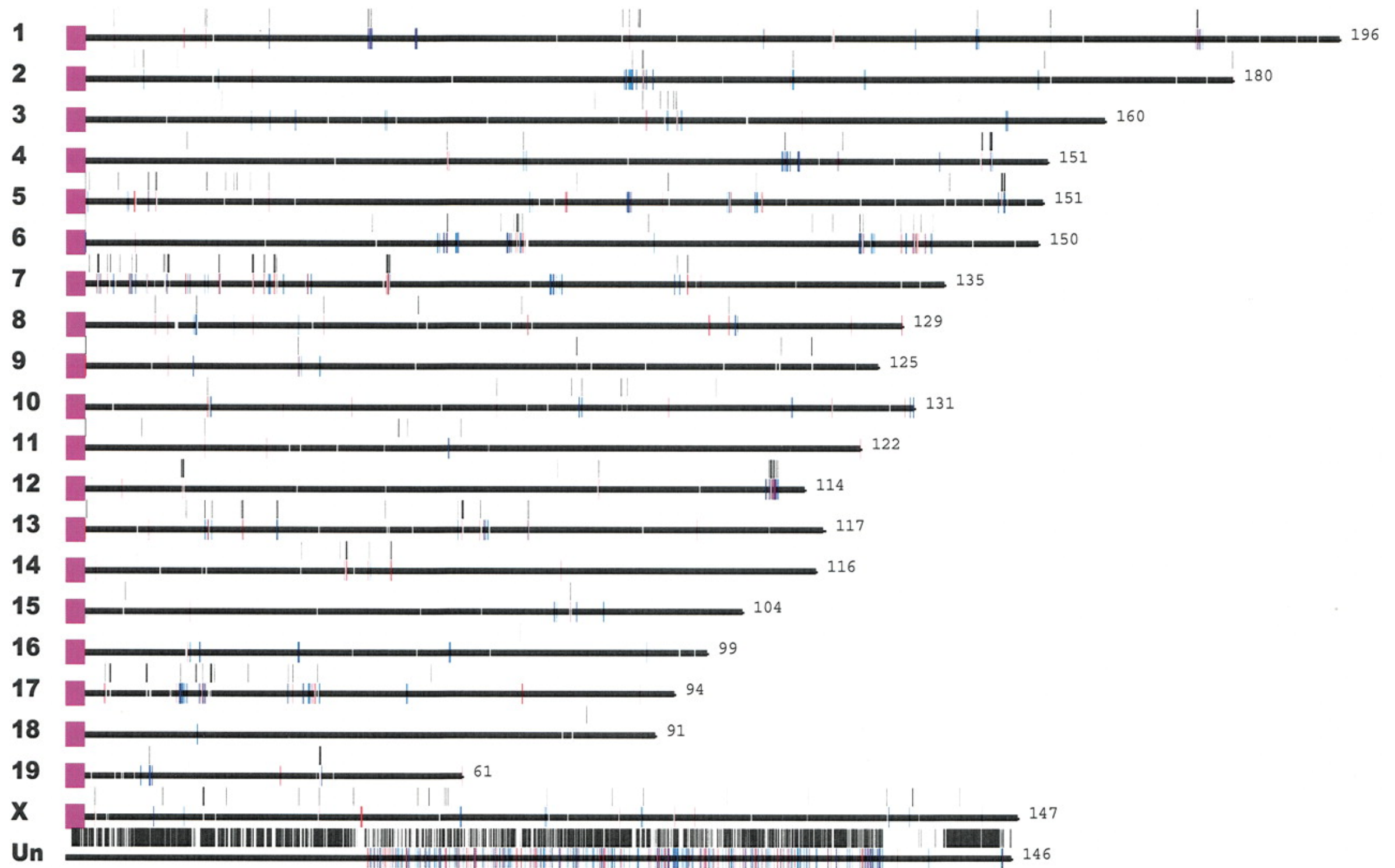
**RESOLUTION**

**PARSABILITY**

**COLOR**

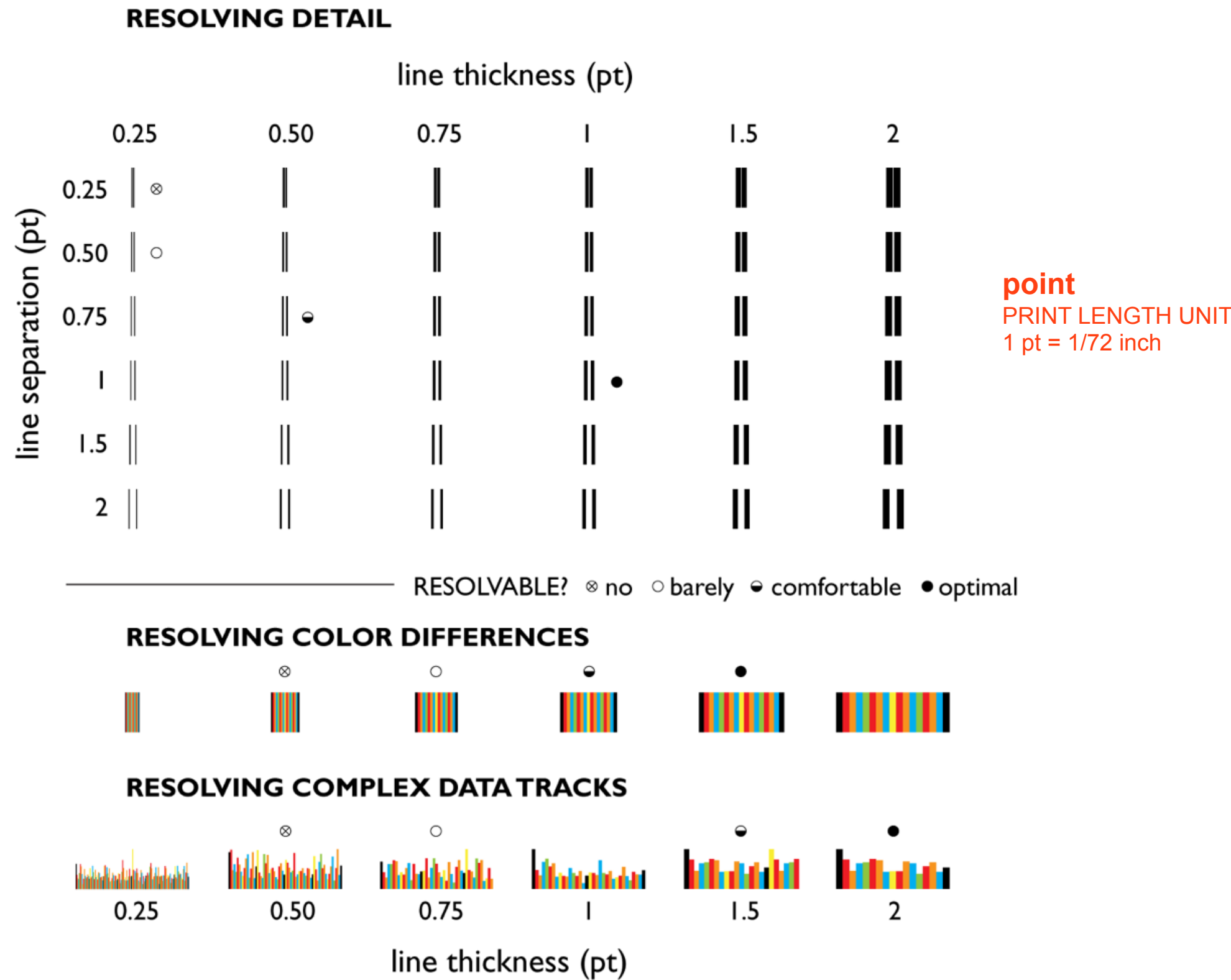


# RESOLUTION & ACUITY LIMITS



Mouse segmental duplications. J. A. Bailey, D. M. Church, M. Ventura, M. Rocchi, E. E. Eichler, Genome Res 14, 789 (May, 2004).

# RESOLUTION & ACUITY LIMITS

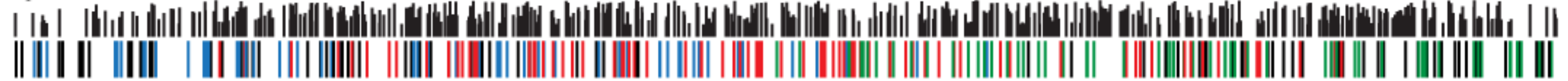


The resolving power of the eye is approximately 50 cycles per degree. This limits us from distinguishing features smaller than 0.1 mm at a reading distance of 30 cm. Larger features must be used to maintain legibility and comprehension. 1 point = 1/72 inch = 0.0353 cm

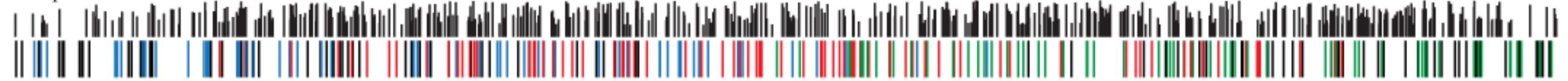
# RESOLUTION & ACUITY LIMITS

**DO NOT DIVIDE YOUR SCALE INTO MORE THAN 500 INTERVALS**

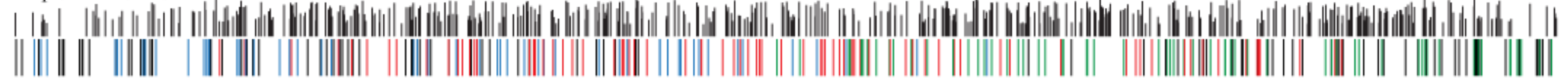
1 pt, 2.9 kb, ~ 500 divisions



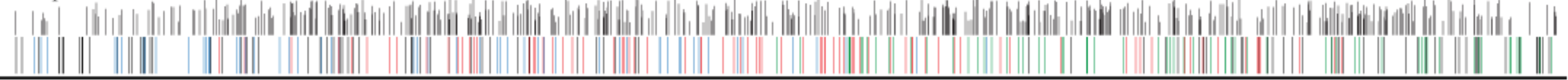
0.75 pt, 2.2 kb, ~ 700 divisions



0.5 pt, 1.5 kb, ~ 1,050 divisions



0.25 pt, 0.7 kb, ~2,100 divisions



*S. cerevisiae* chrIV, 1,531 kb   scale length 183 mm, 7.2 in, 519 pt   scale resolution 8.4 kb/mm, 213 kb/in, 2.9 kb/pt

As a rule of thumb, you should not divide your scale into more than 500 intervals per 8.5 inch (216mm, US letter size). This corresponds to 1 pt on a 183 mm figure, 4 pixels on a 1920 horizontal resolution display, or 2 pixels on a typical LCD projector.

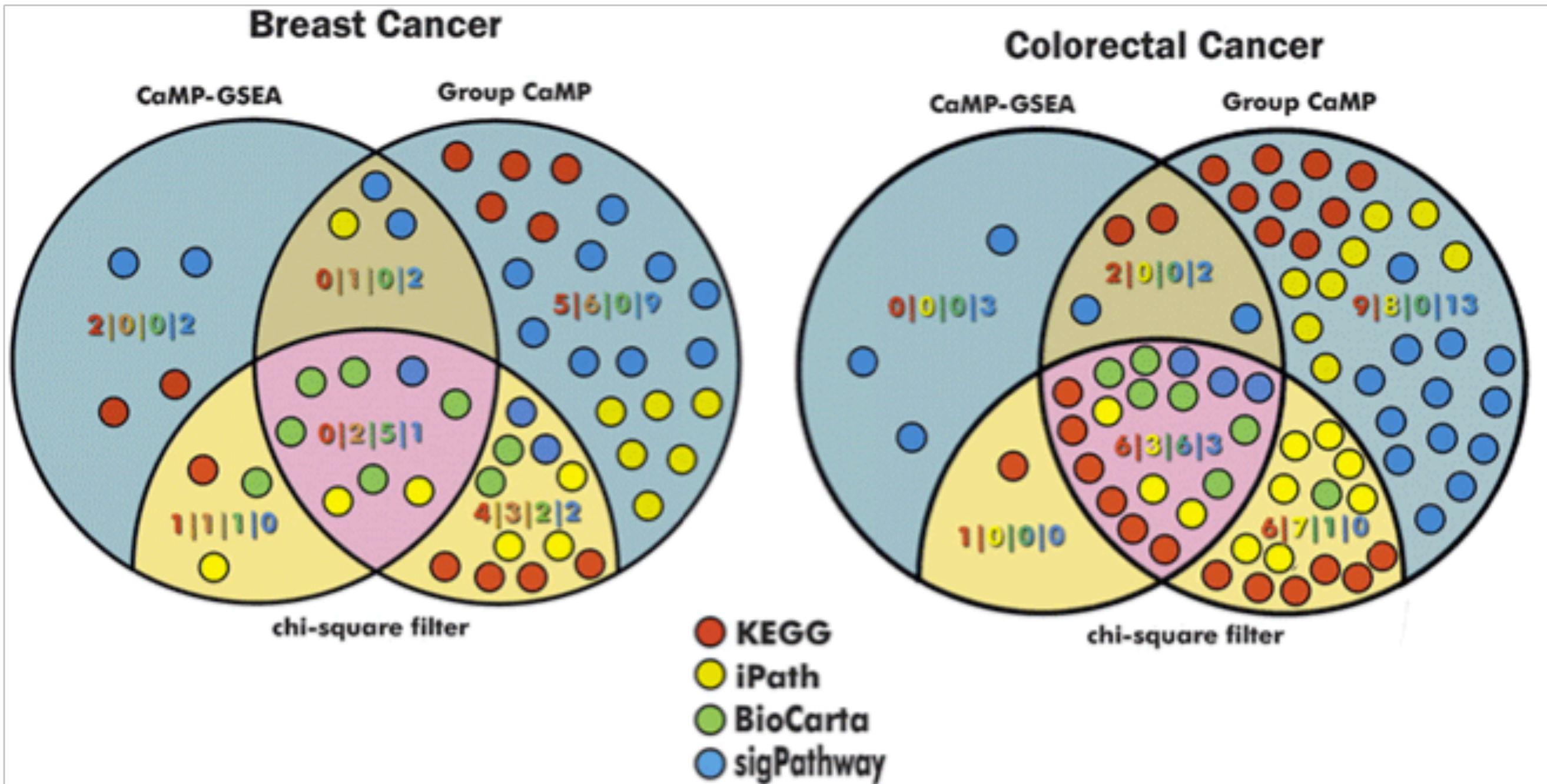


# PARSABILITY LIMITS



The specific domains and combinations in the domain graph. Ye, Y. and A. Godzik, Comparative analysis of protein domain organization. Genome Res, 2004. 14(3): p. 343-53.

# PARSABILITY LIMITS

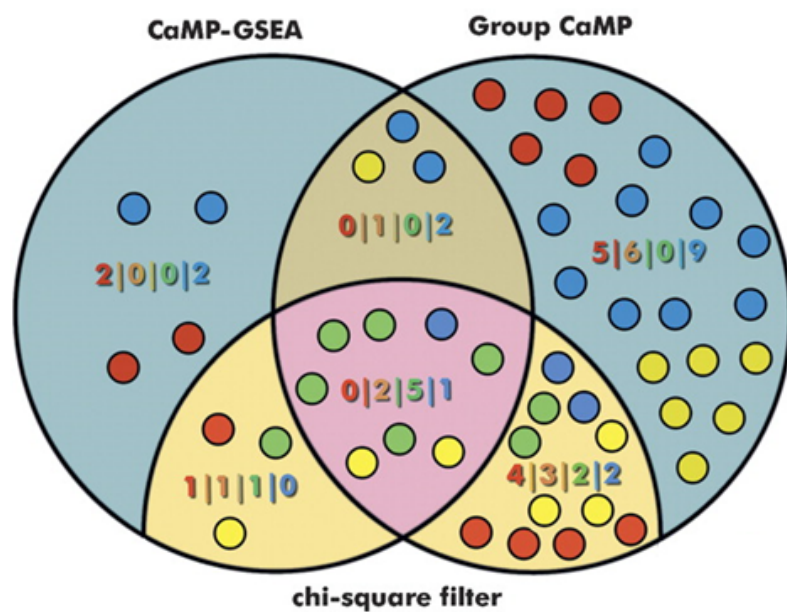


Comparison of mutation enrichment in cellular pathways using complementary statistical approaches. Lin, J., et al., A multidimensional analysis of genes mutated in breast and colorectal cancers. Genome Res, 2007. 17(9): p. 1304-18.

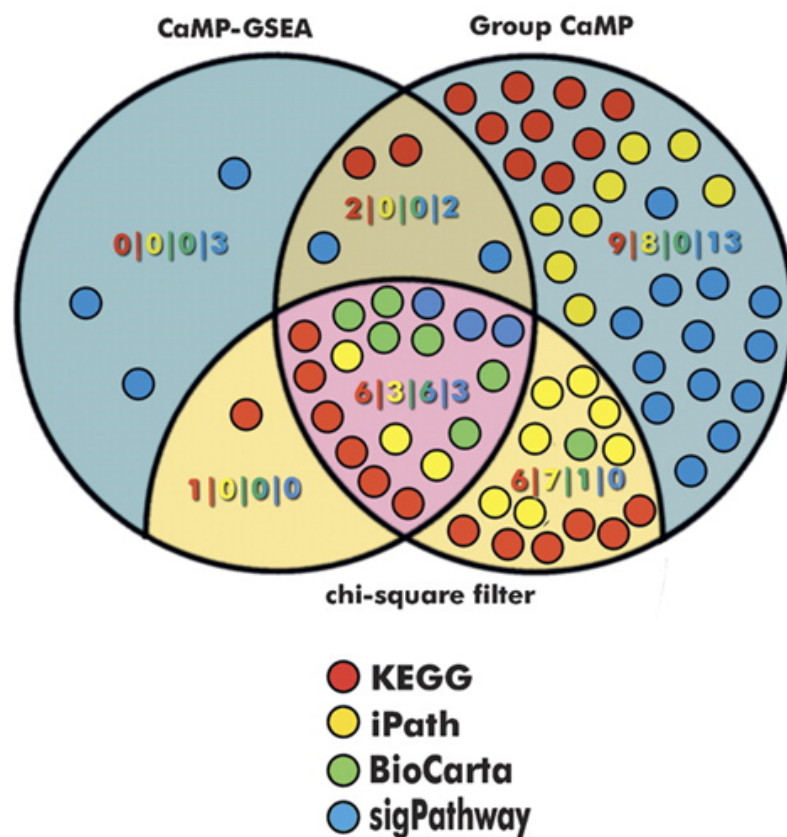


# REFACTORING COMPLEXITY

## Breast Cancer



## Colorectal Cancer

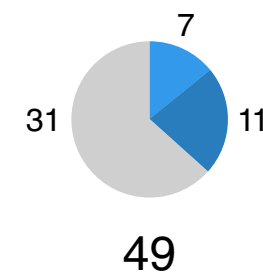


A

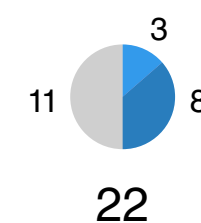
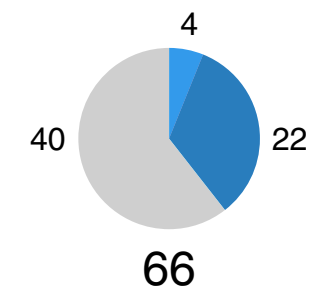
ALL PATHWAYS

CaMP-GSEA ● CaMP

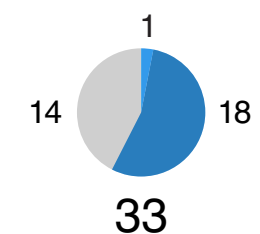
BREAST



COLORECTAL



—  $\chi^2$  —

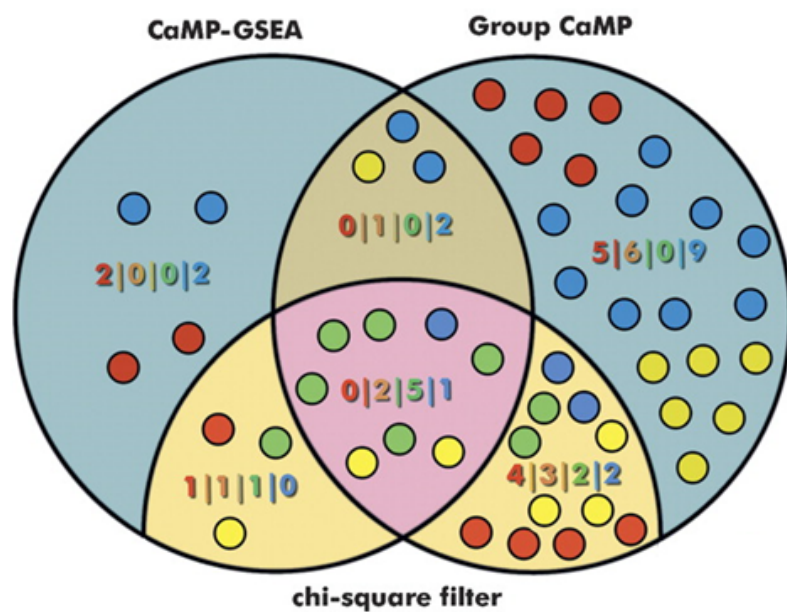


Comparison of mutation enrichment in cellular pathways using complementary statistical approaches. Lin, J., et al., A multidimensional analysis of genes mutated in breast and colorectal cancers. Genome Res, 2007. 17(9): p. 1304-18.

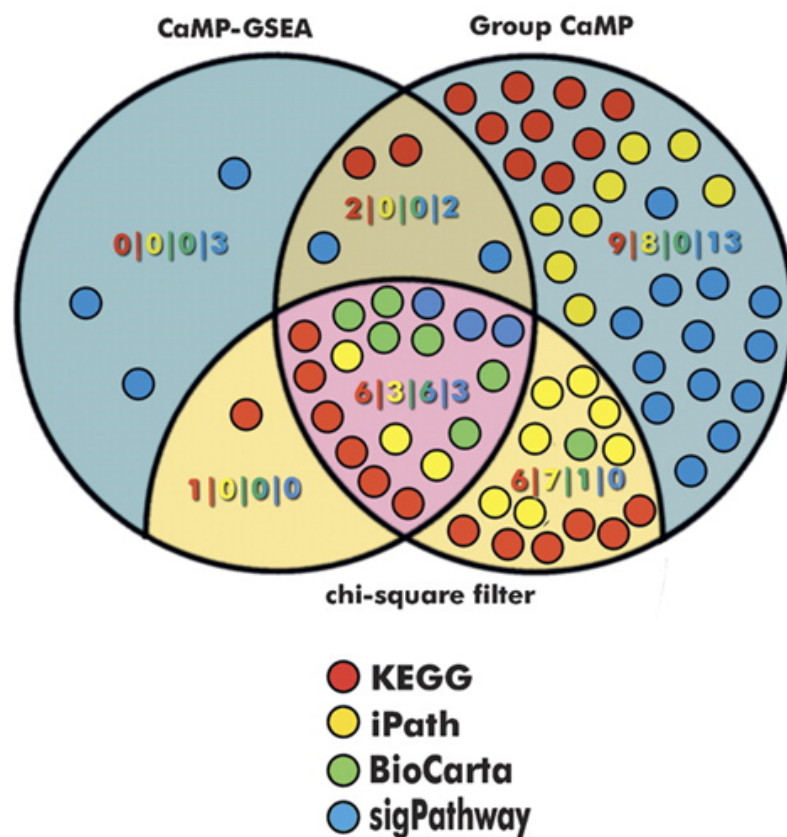


# REFACTORING COMPLEXITY

## Breast Cancer



## Colorectal Cancer

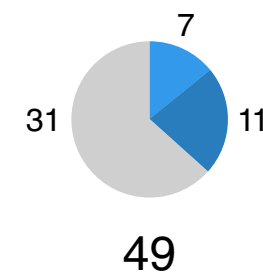


A

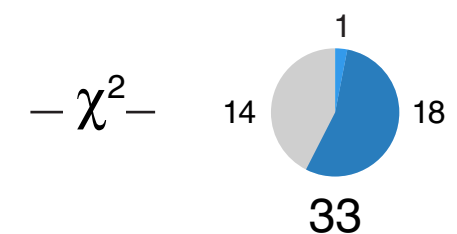
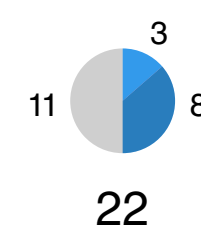
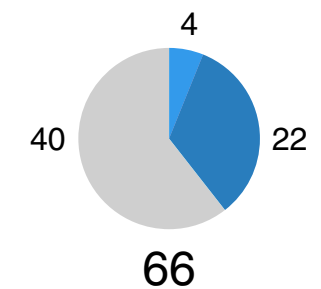
ALL PATHWAYS

CaMP-GSEA ● CaMP

BREAST



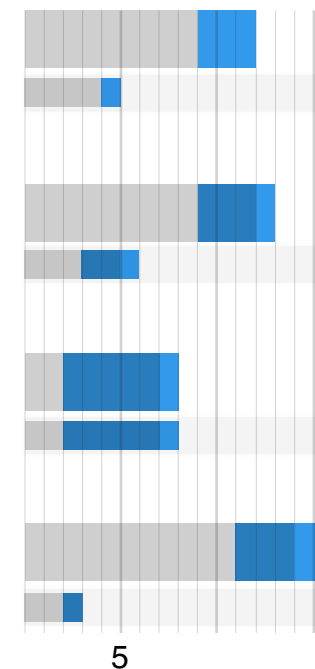
COLORECTAL



—  $\chi^2$  —

B

KEGG

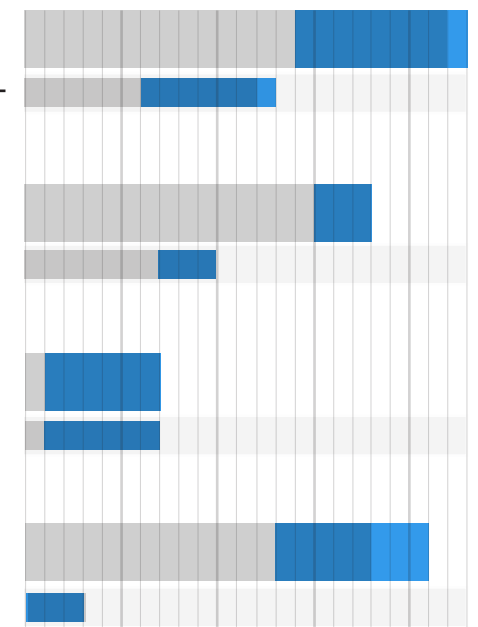


iPath

BioCarta

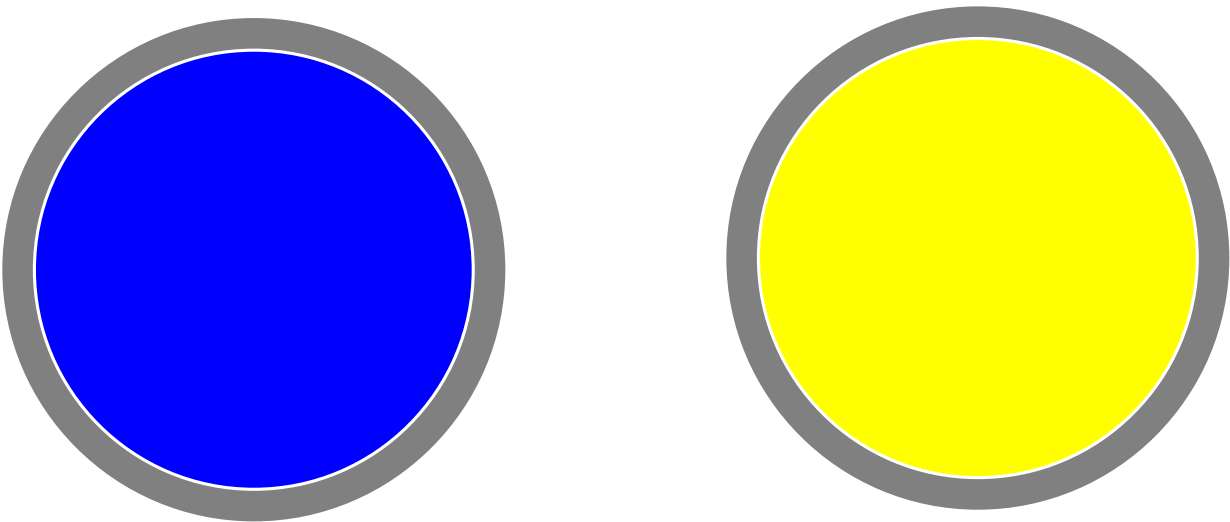
sigPathway

—  $\chi^2$  —



Comparison of mutation enrichment in cellular pathways using complementary statistical approaches. Lin, J., et al., A multidimensional analysis of genes mutated in breast and colorectal cancers. Genome Res, 2007. 17(9): p. 1304-18.

# COLOR PERCEPTION



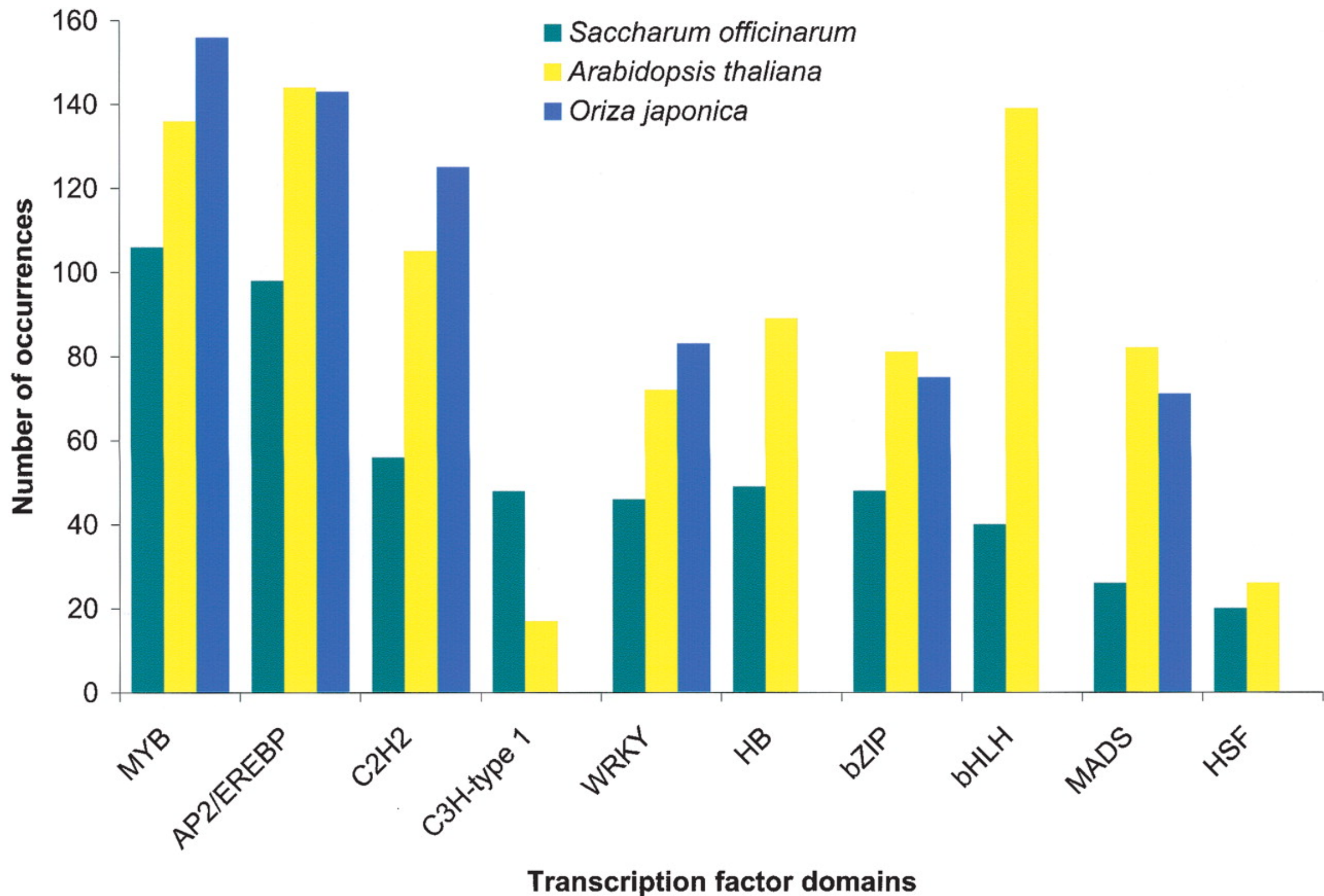
## HSB COLOR SPACE

HUE	240	60
SATURATION	1	1
BRIGHTNESS	1	1

## LCH COLOR SPACE

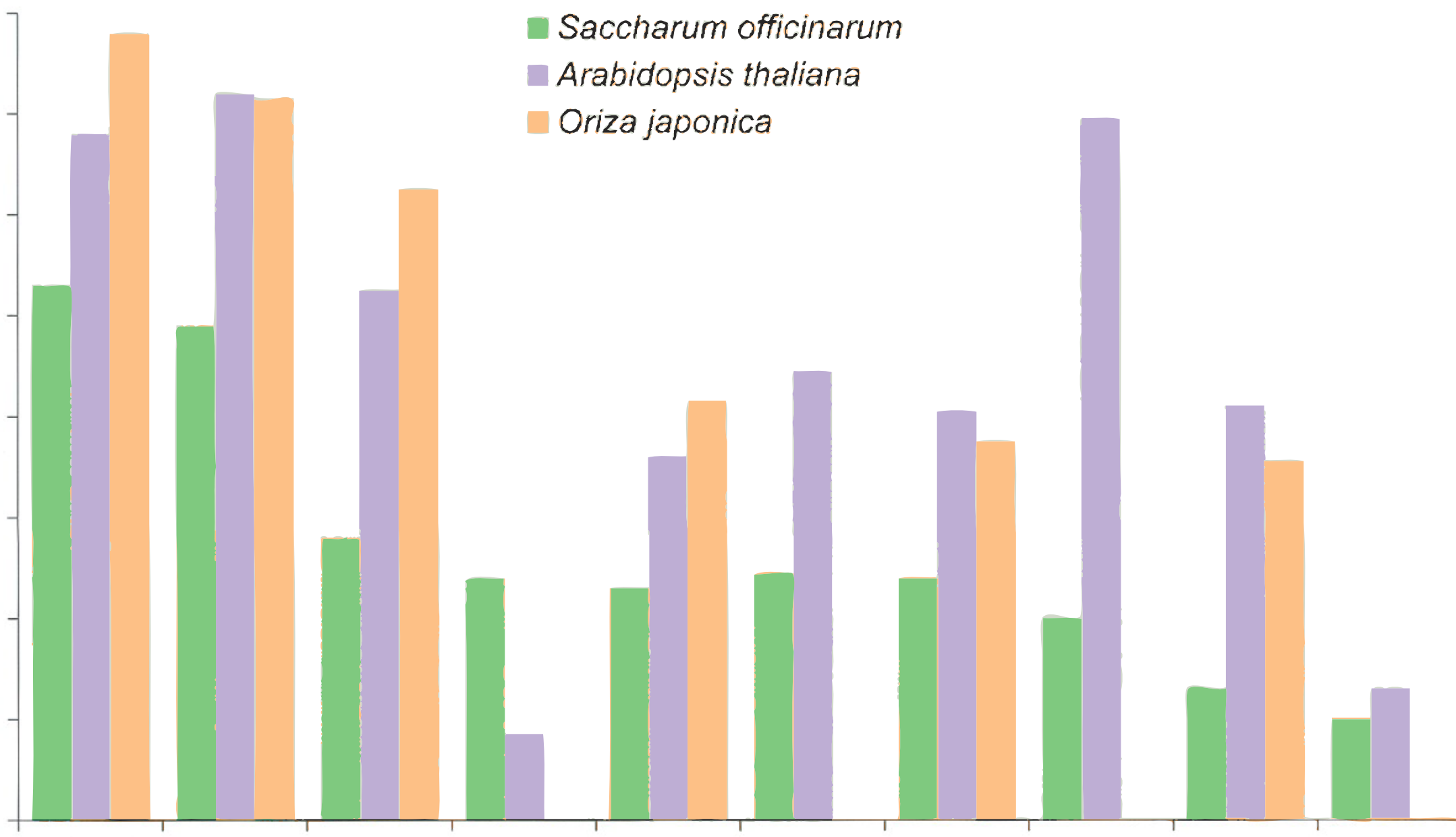
HUE	266	86
CHROMA	130	107
LIGHTNESS	0.32	0.97

# COLOR PERCEPTION



The 10 most common transcription factor Pfam domains in SAS proteins. Vettore, A.L., et al., Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane. Genome Res, 2003. 13(12): p. 2725-35.

# COLOR PERCEPTION



## BREWER QUALITATIVE 3-COLOR PALETTES

ACCENT ■ ■ ■

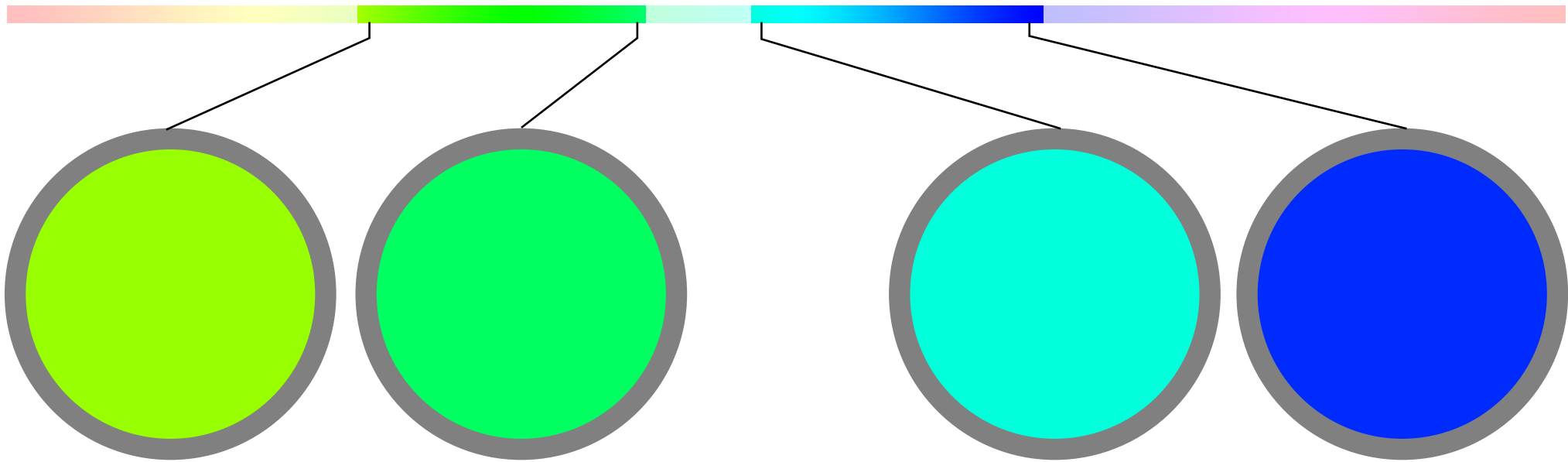
PASTEL 1 ■ ■ ■

SET 2 ■ ■ ■

WWW.COLORBREWER.ORG

In a qualitative palette, colors have no perceived order or importance.

# PERCEPTUAL UNIFORMITY



## HSB COLOR SPACE

	$\Delta H = 60$		$\Delta H = 60$	
HUE	83	143	171	231
SATURATION	1	1	1	1
BRIGHTNESS	1	1	1	1

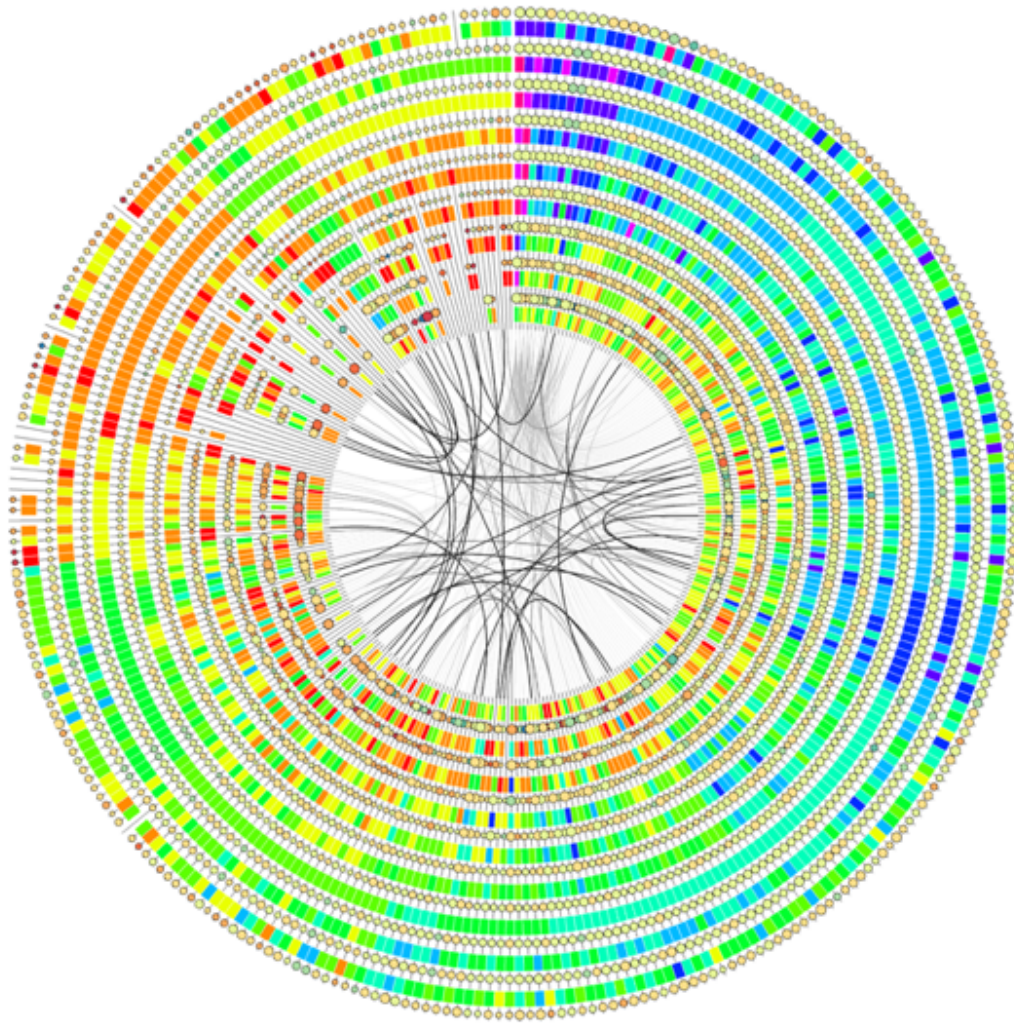
## Lab COLOR SPACE

	$\Delta E_{ab} = 35$		$\Delta E_{ab} = 176$	
L	91	88	90	35
a	-59	-81	-58	70
b	87	60	4	-102

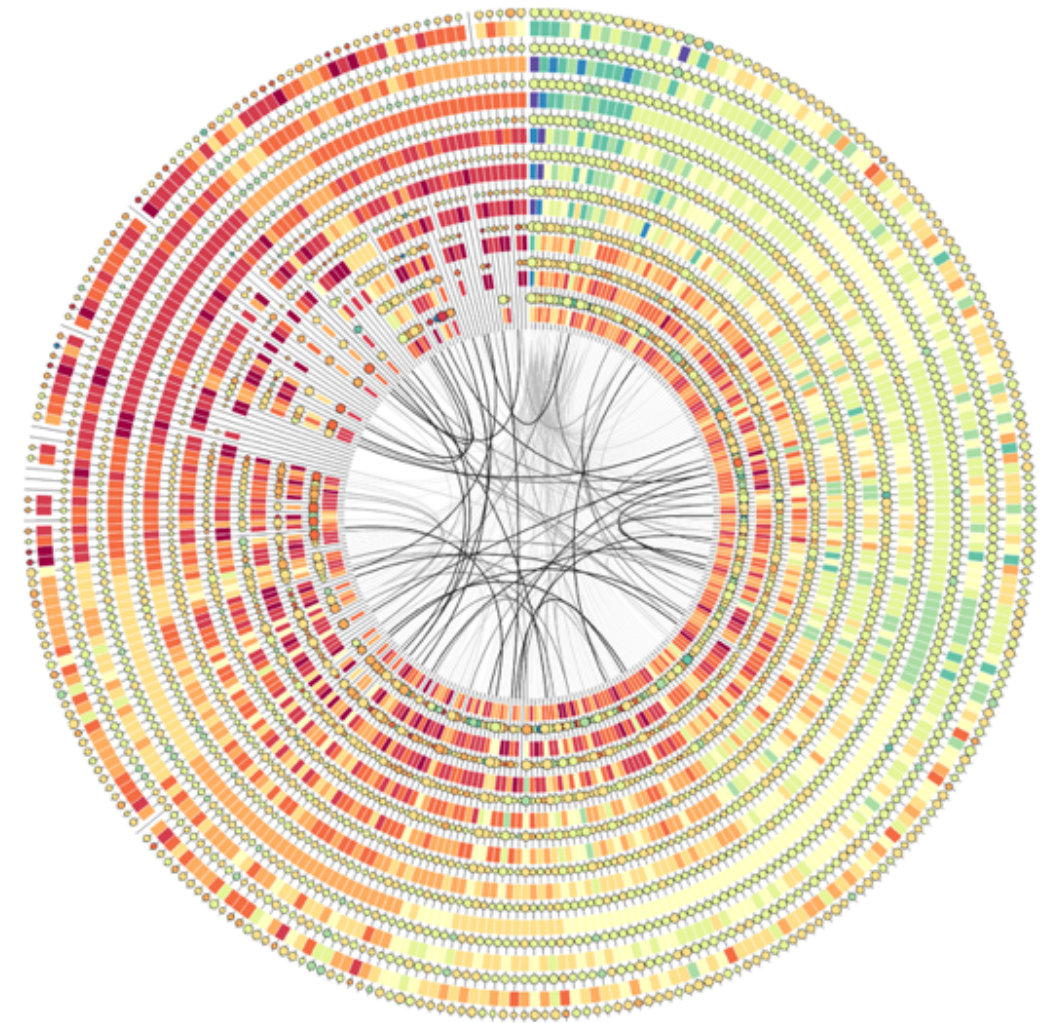


# DIVERGING BREWER PALETTES

UNIFORM HSB 11 COLOR PALETTE



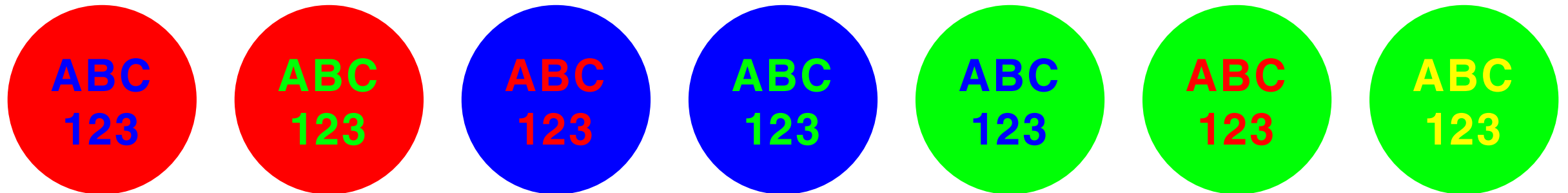
BREWER 11 COLOR SPECTRAL PALETTE



last.fm music listening habits by age.  
<http://mkweb.bcgsc.ca/spl/circos-lastfm-legend.png>

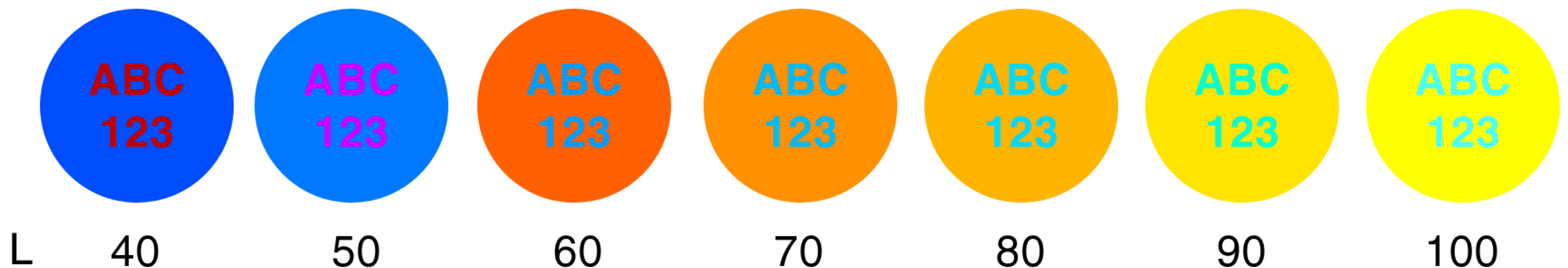
# CONTRAST

## AVOID ADJACENT PURE COLORS



## AVOID ADJACENT COLORS WITH SIMILAR LUMINANCE

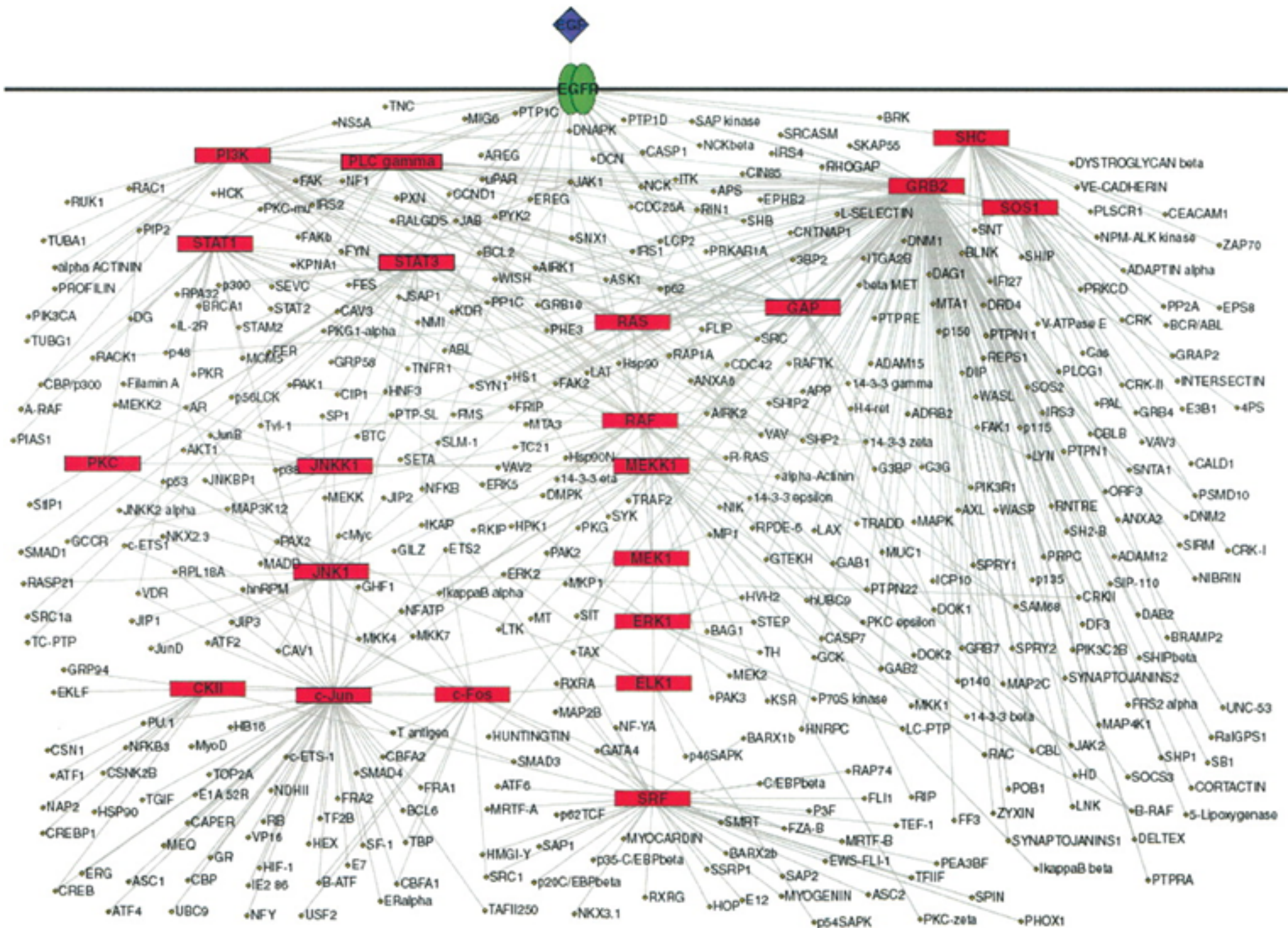
$$\Delta L = 0$$



Simultaneous contrast occurs when two pure colors are adjacent.  
Poor contrast occurs when two colors have similar luminance (perceived brightness).



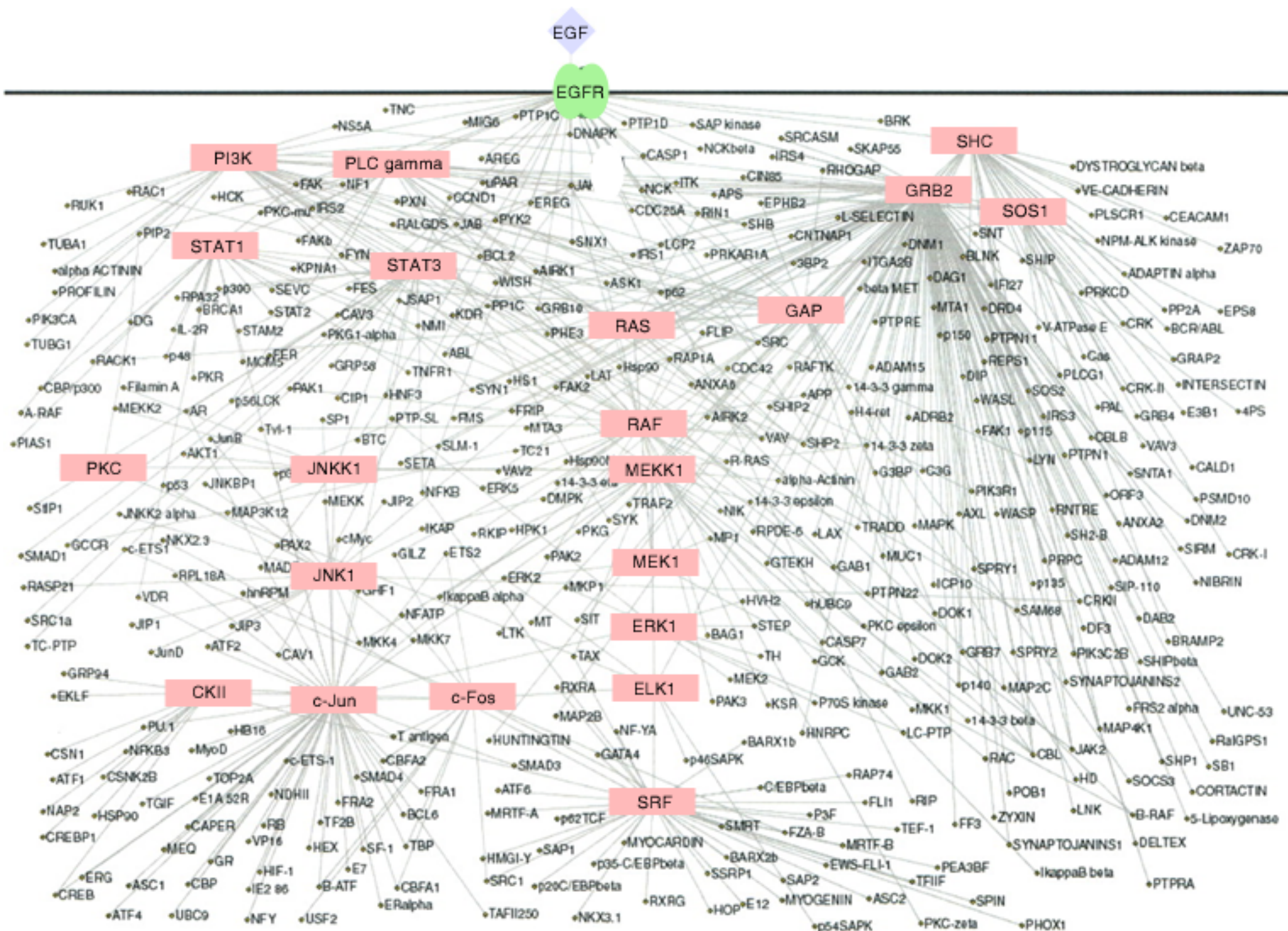
# CONTRAST



Black text on dark colors is illegible due to insufficient luminosity contrast.

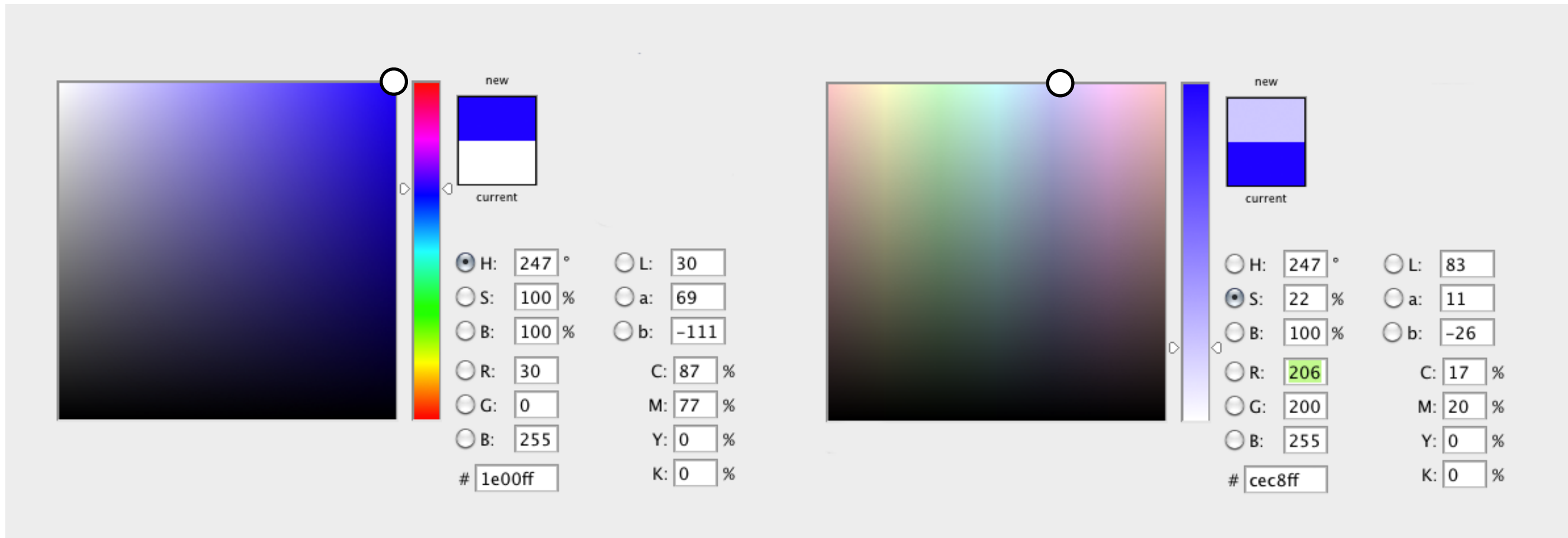


# GOOD CONTRAST



Decreasing the saturation of the background colors improves contrast.

# USING COLOR PICKERS



Photoshop shows a color's coordinates in several color spaces. Decrease saturation to make a color appear lighter. Use the "L" channel of the Lab space to determine a color's perceived brightness (luminance).

legibility **clarity** attractiveness

**STRONG MESSAGE**

**REDUNDANCY**

**REPRESENTATION**

**CONSISTENCY**

**EXCESS INK**

**GLYPHS**



# IMPORTANCE OF CLARITY



This is fairly clear. Either

- (a) do not poop near the tiny trees, or
- (b) no pooping for giants



This is less clear. Perhaps

- (a) no talking
- (b) did you remember your black glove?
- (c) stop and wait

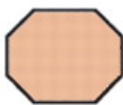







# IS YOUR MESSAGE NECESSARY?



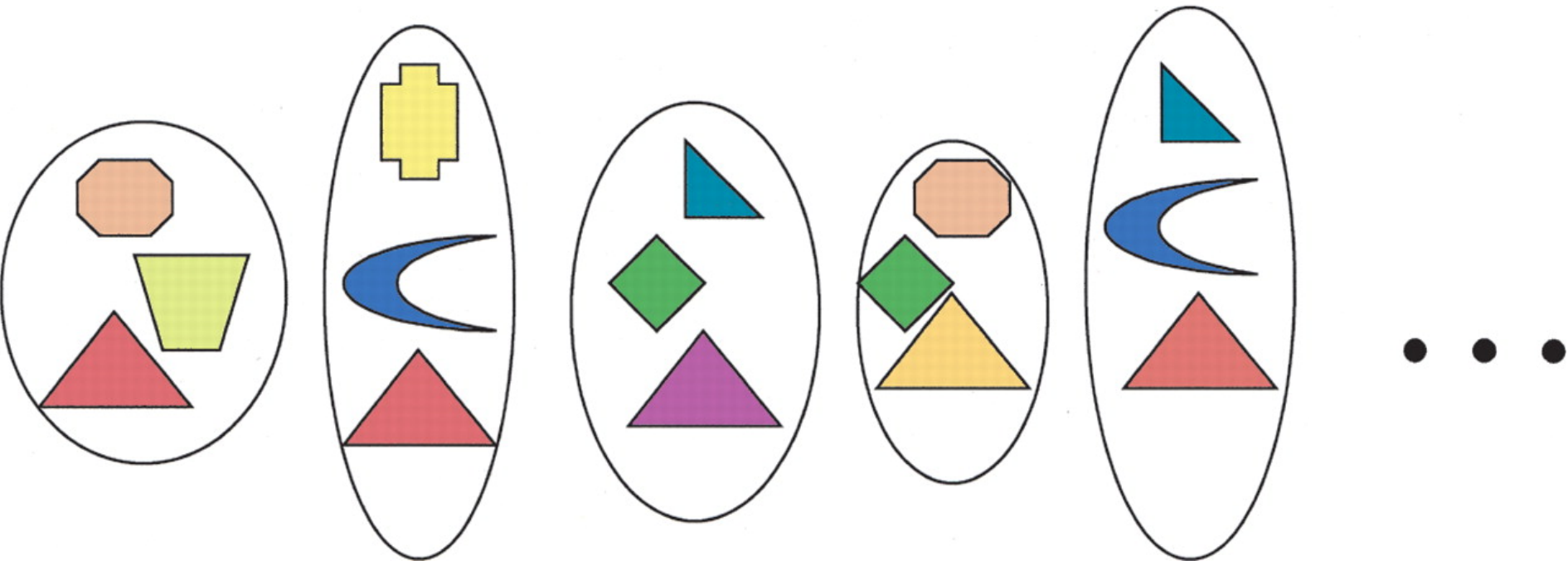
<http://lushlush.livejournal.com/190093.html>

# ACADEMIC FIGURE OR MENSA TEST?

 =  =  =  ENVELOPE

 =  =  =  METABOLISM

 =  =  =  TRANSLATION



A mix-and-match model for prokaryotic genome evolution. Charlebois, R.L. and W.F. Doolittle, Computing prokaryotic gene ubiquity: rescuing the core from extinction. Genome Res, 2004. 14(12): p. 2469-77.



# CRAFTING A MESSAGE

WHAT IS SHOWN?

WHAT IS COMMUNICATED?

WHAT IS INTERPRETED?

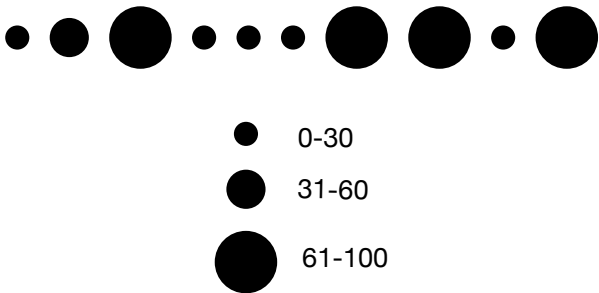
RAW DATA

12 54 82 29 25 22 67 61 23 79

NO CLEAR MESSAGE.

UNKNOWN. READER IS ON THEIR OWN.

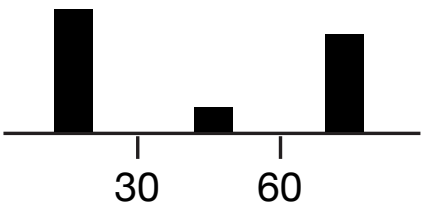
DISCRETIZED



SCALE

THREE RANGES ARE IMPORTANT. INDIVIDUAL VALUES WITHIN A RANGE ARE NOT.

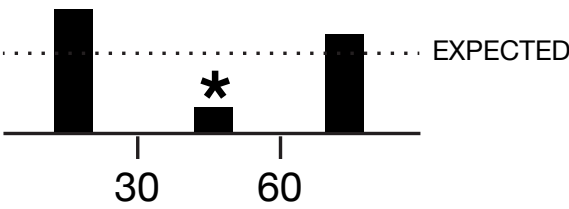
BINNED



DISTRIBUTION

THERE ARE FEWER MEDIUM-SIZED VALUES.

TREND



SIGNIFICANCE

THERE ARE SIGNIFICANTLY FEWER MEDIUM-SIZED VALUES.

# CRAFTING A MESSAGE

WHAT IS SHOWN?

WHAT IS COMMUNICATED?

WHAT IS INTERPRETED?

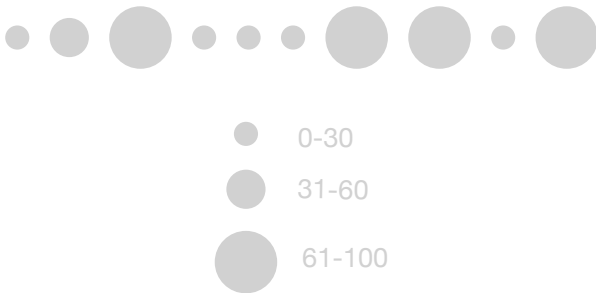
RAW DATA

12 54 82 29 25 22 67 61 23 79

NO CLEAR MESSAGE.

UNKNOWN. READER IS ON THEIR OWN.

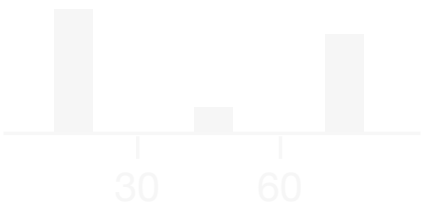
DISCRETIZED



SCALE

THREE RANGES ARE IMPORTANT. INDIVIDUAL VALUES WITHIN A RANGE ARE NOT.

BINNED



DISTRIBUTION

THERE ARE FEWER MEDIUM-SIZED VALUES.

TREND



SIGNIFICANCE

THERE ARE SIGNIFICANTLY FEWER MEDIUM-SIZED VALUES



BC Cancer Agency  
CARE + RESEARCH



CANADA'S MICHAEL SMITH  
GENOME  
SCIENCES  
CENTRE

# CRAFTING A MESSAGE

WHAT IS SHOWN?

WHAT IS COMMUNICATED?

WHAT IS INTERPRETED?

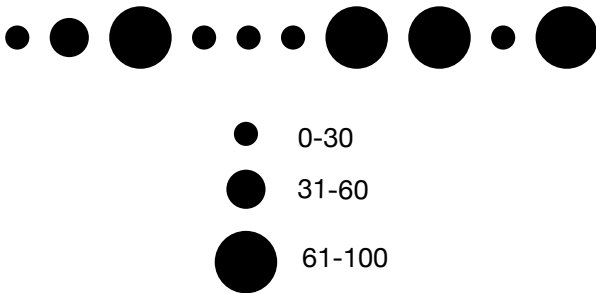
RAW DATA

12 54 82 29 25 22 67 61 23 79

NO CLEAR MESSAGE.

UNKNOWN. READER IS ON THEIR OWN.

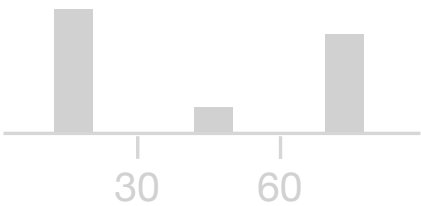
DISCRETIZED



SCALE

THREE RANGES ARE IMPORTANT. INDIVIDUAL VALUES WITHIN A RANGE ARE NOT.

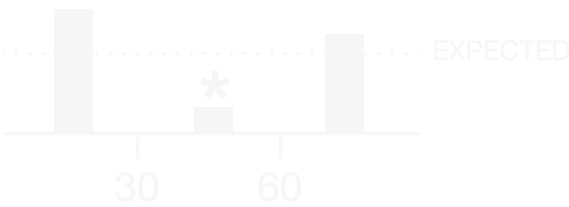
BINNED



DISTRIBUTION

THERE ARE FEWER MEDIUM-SIZED VALUES.

TREND



SIGNIFICANCE

THERE ARE SIGNIFICANTLY FEWER MEDIUM-SIZED VALUES



BC Cancer Agency  
CARE + RESEARCH



CANADA'S MICHAEL SMITH  
GENOME  
SCIENCES  
CENTRE

# CRAFTING A MESSAGE

WHAT IS SHOWN?

WHAT IS COMMUNICATED?

WHAT IS INTERPRETED?

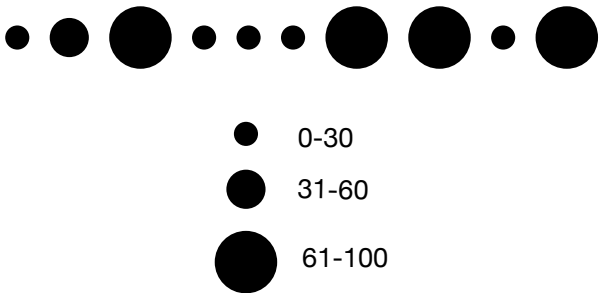
RAW DATA

12 54 82 29 25 22 67 61 23 79

NO CLEAR MESSAGE.

UNKNOWN. READER IS ON THEIR OWN.

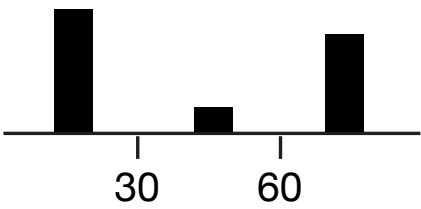
DISCRETIZED



SCALE

THREE RANGES ARE IMPORTANT. INDIVIDUAL VALUES WITHIN A RANGE ARE NOT.

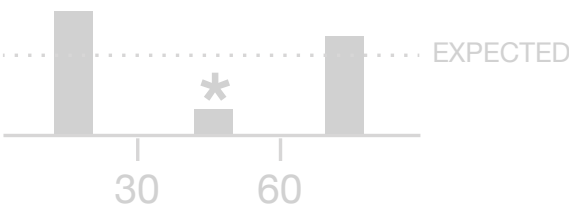
BINNED



DISTRIBUTION

THERE ARE FEWER MEDIUM-SIZED VALUES.

TREND



SIGNIFICANCE

THERE ARE SIGNIFICANTLY FEWER MEDIUM-SIZED VALUES.



# CRAFTING A MESSAGE

WHAT IS SHOWN?

WHAT IS COMMUNICATED?

WHAT IS INTERPRETED?

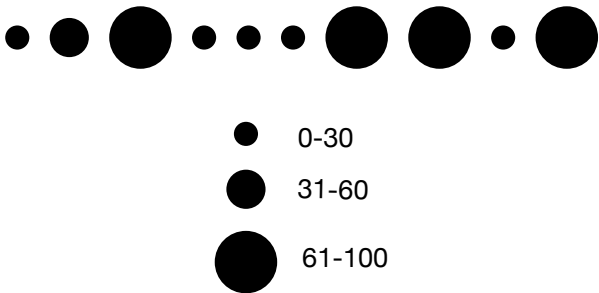
RAW DATA

12 54 82 29 25 22 67 61 23 79

NO CLEAR MESSAGE.

UNKNOWN. READER IS ON THEIR OWN.

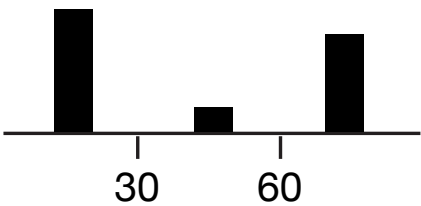
DISCRETIZED



SCALE

THREE RANGES ARE IMPORTANT. INDIVIDUAL VALUES WITHIN A RANGE ARE NOT.

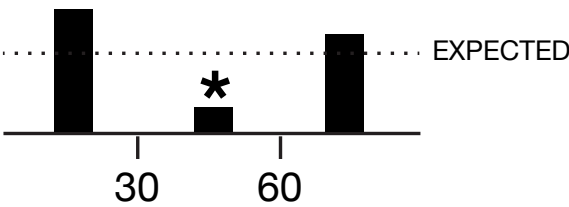
BINNED



DISTRIBUTION

THERE ARE FEWER MEDIUM-SIZED VALUES.

TREND

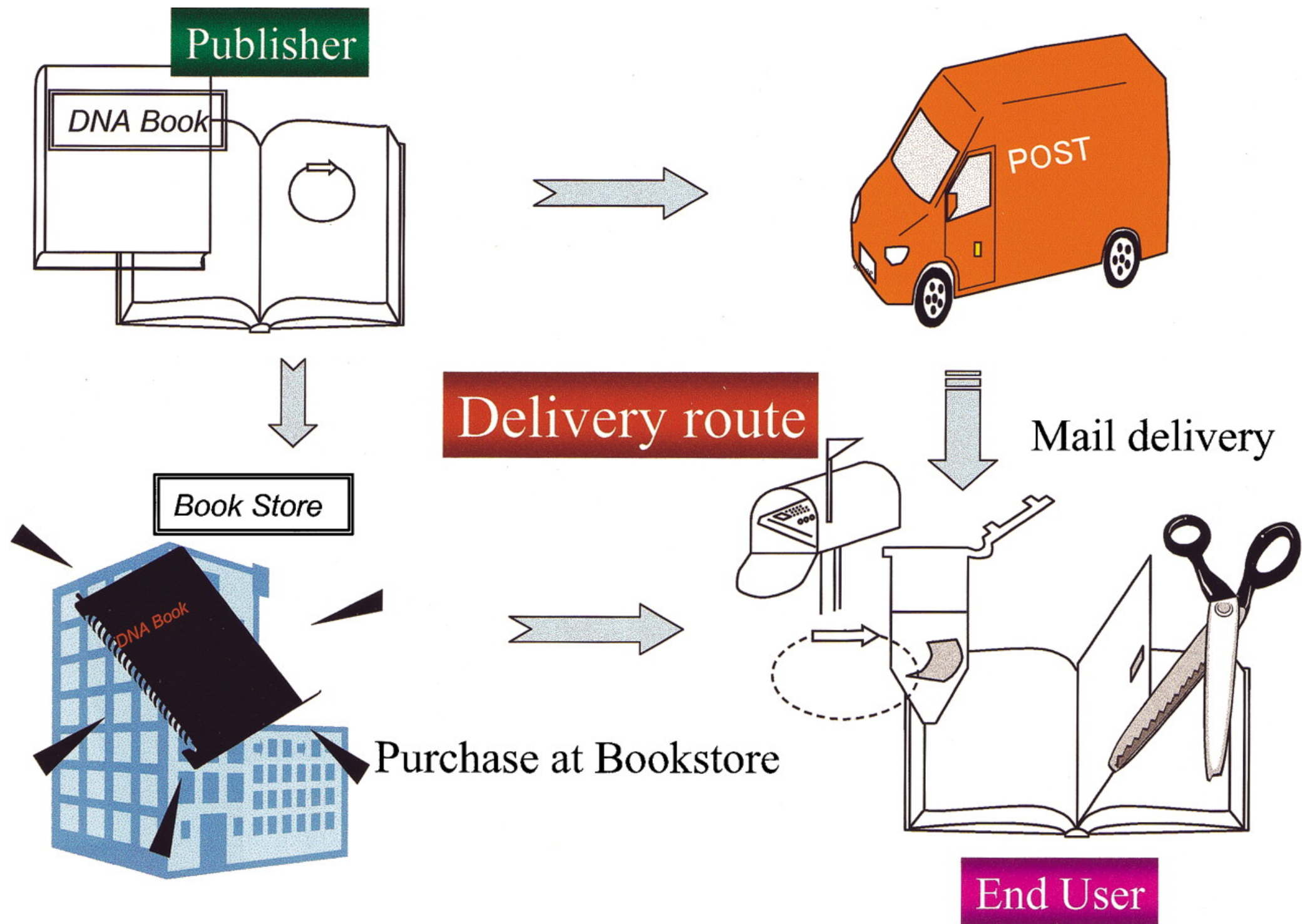


SIGNIFICANCE

THERE ARE SIGNIFICANTLY FEWER MEDIUM-SIZED VALUES.

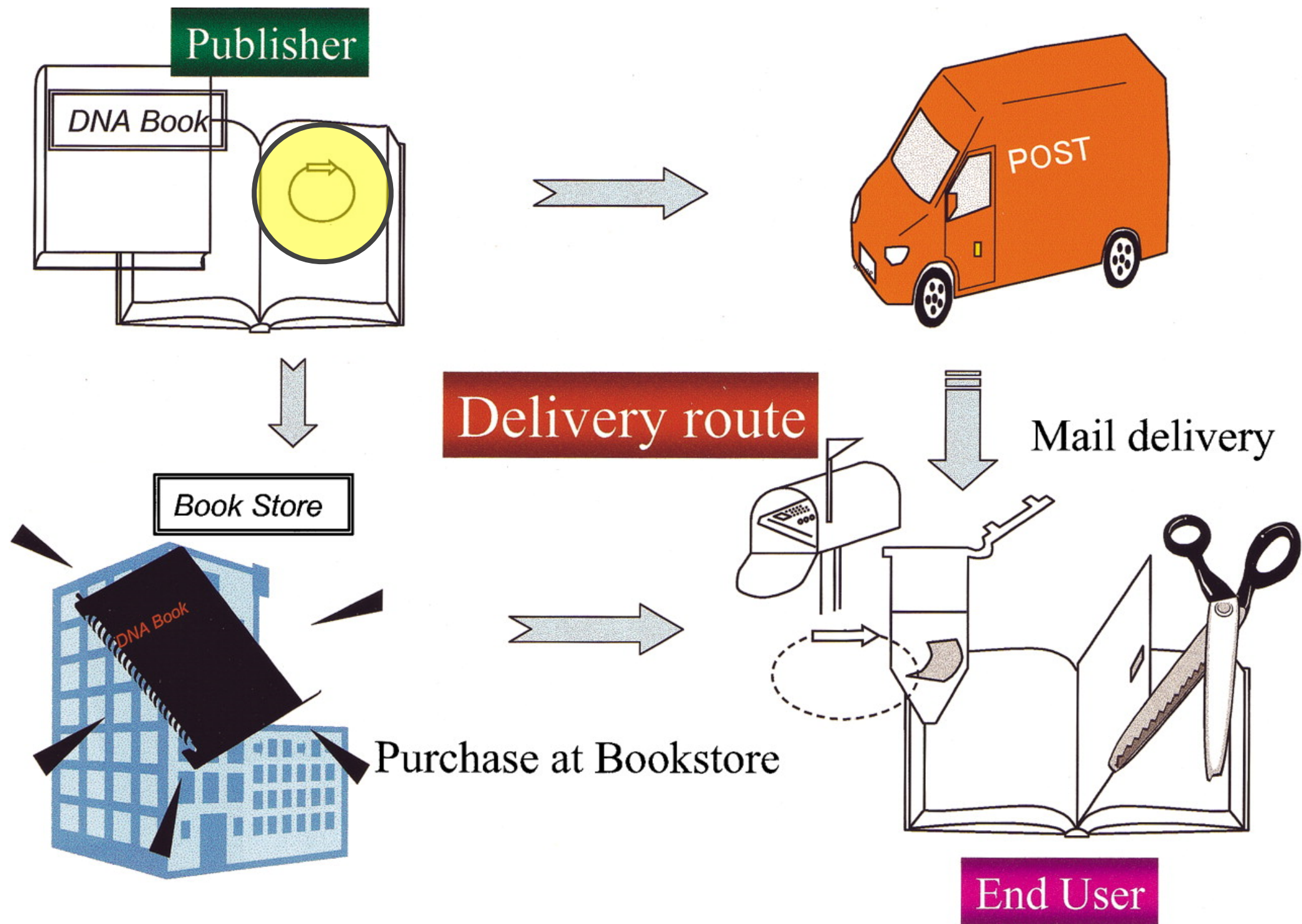


# WHAT IS THE CORE MESSAGE?



Concept of the "DNA Book." Kawai, J. and Y. Hayashizaki, DNA book. Genome Res, 2003. 13(6B): p. 1488-95.

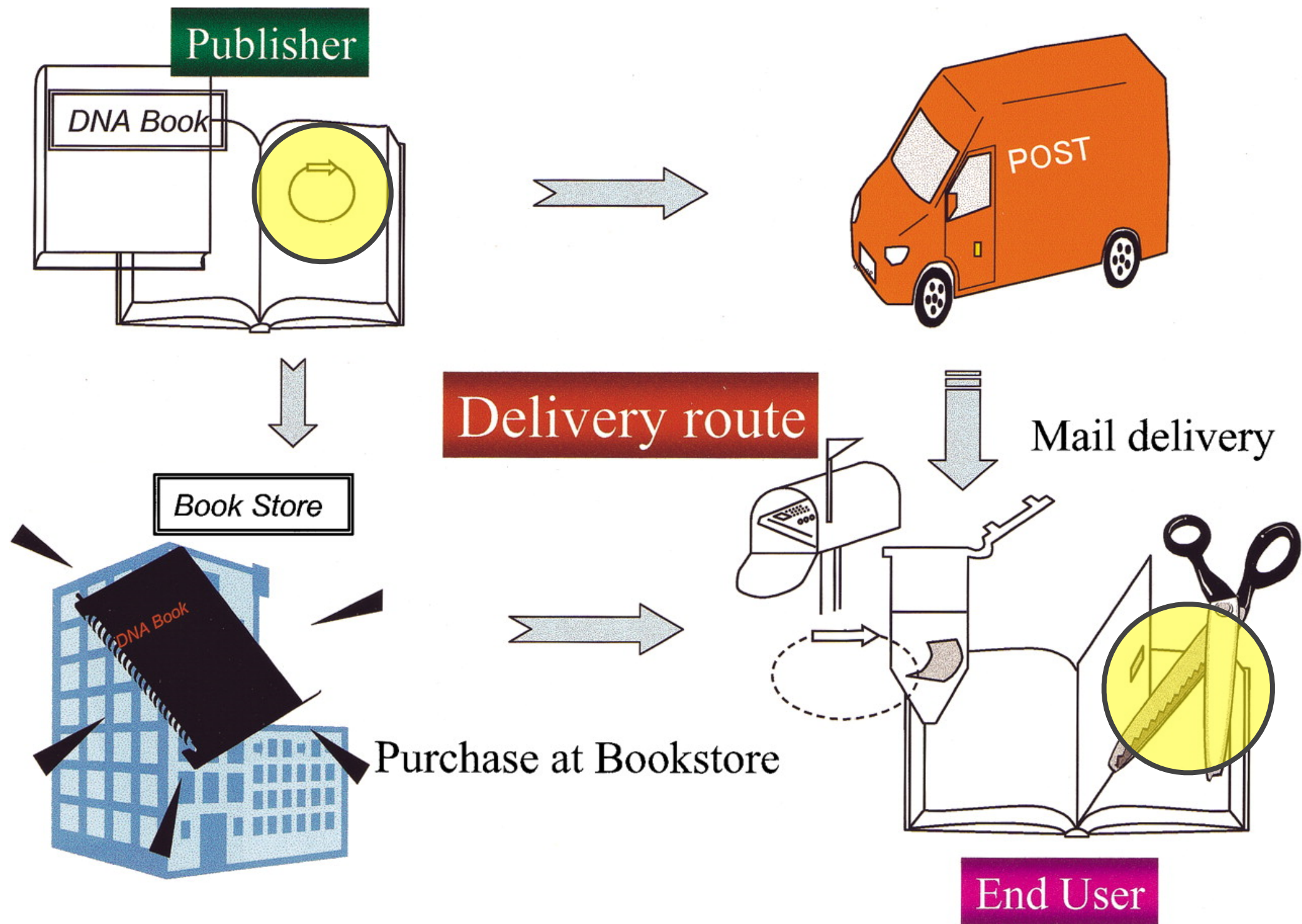
# WHAT IS THE CORE MESSAGE?



Concept of the "DNA Book." Kawai, J. and Y. Hayashizaki, DNA book. Genome Res, 2003. 13(6B): p. 1488-95.

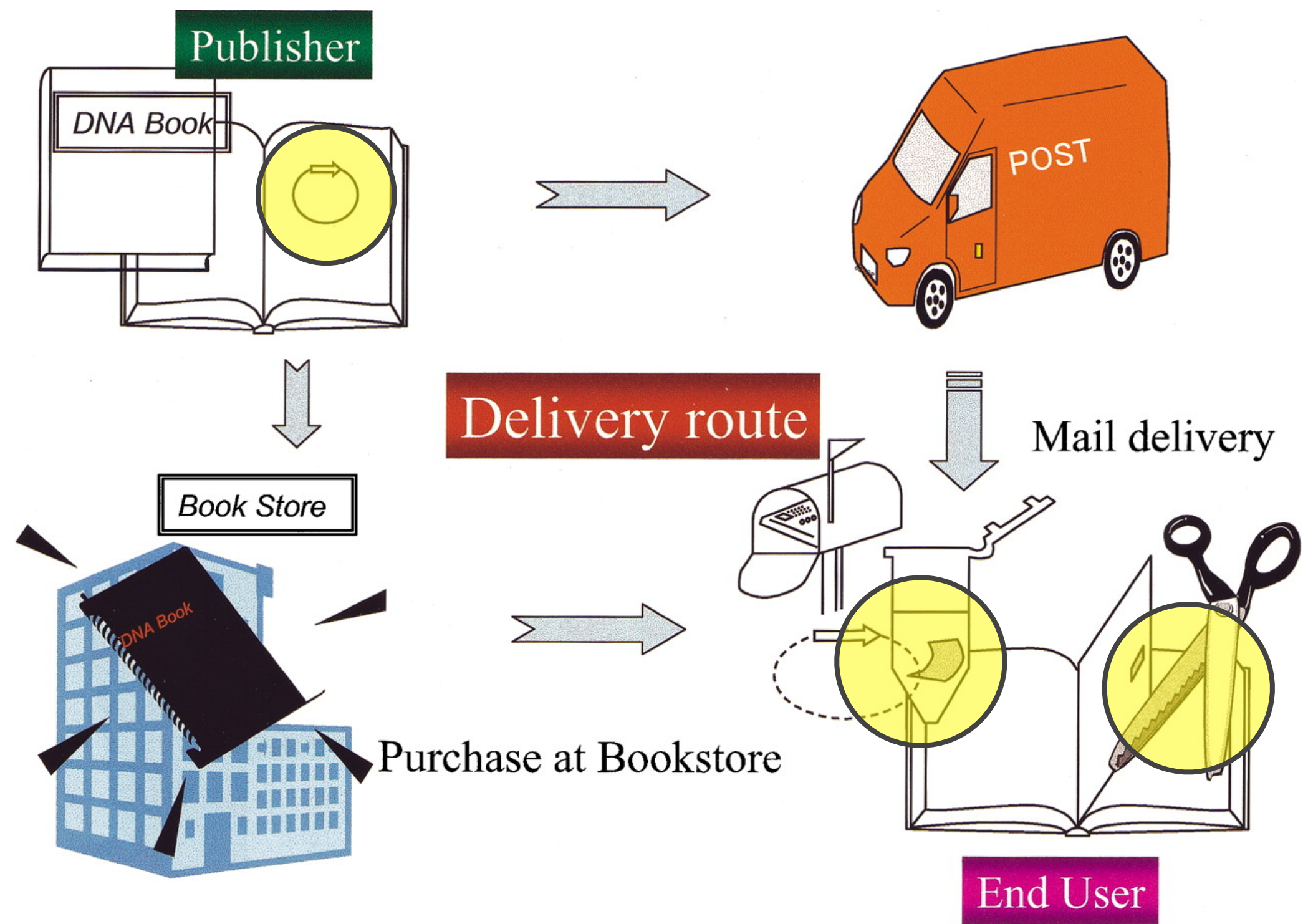


# WHAT IS THE CORE MESSAGE?



Concept of the “DNA Book.” Kawai, J. and Y. Hayashizaki, DNA book. *Genome Res*, 2003. 13(6B): p. 1488-95.

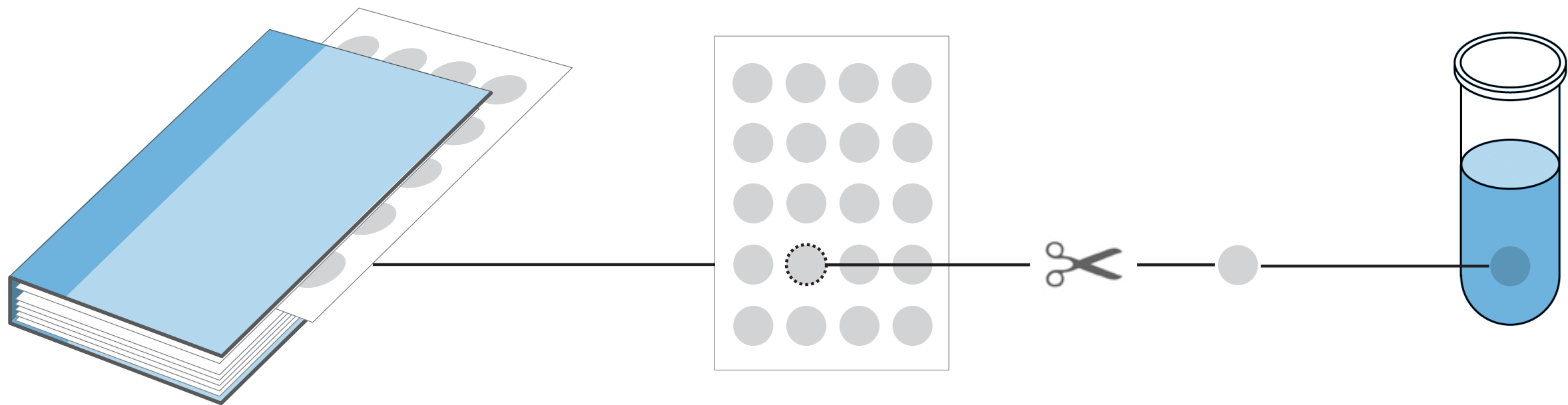
# WHAT IS THE CORE MESSAGE?



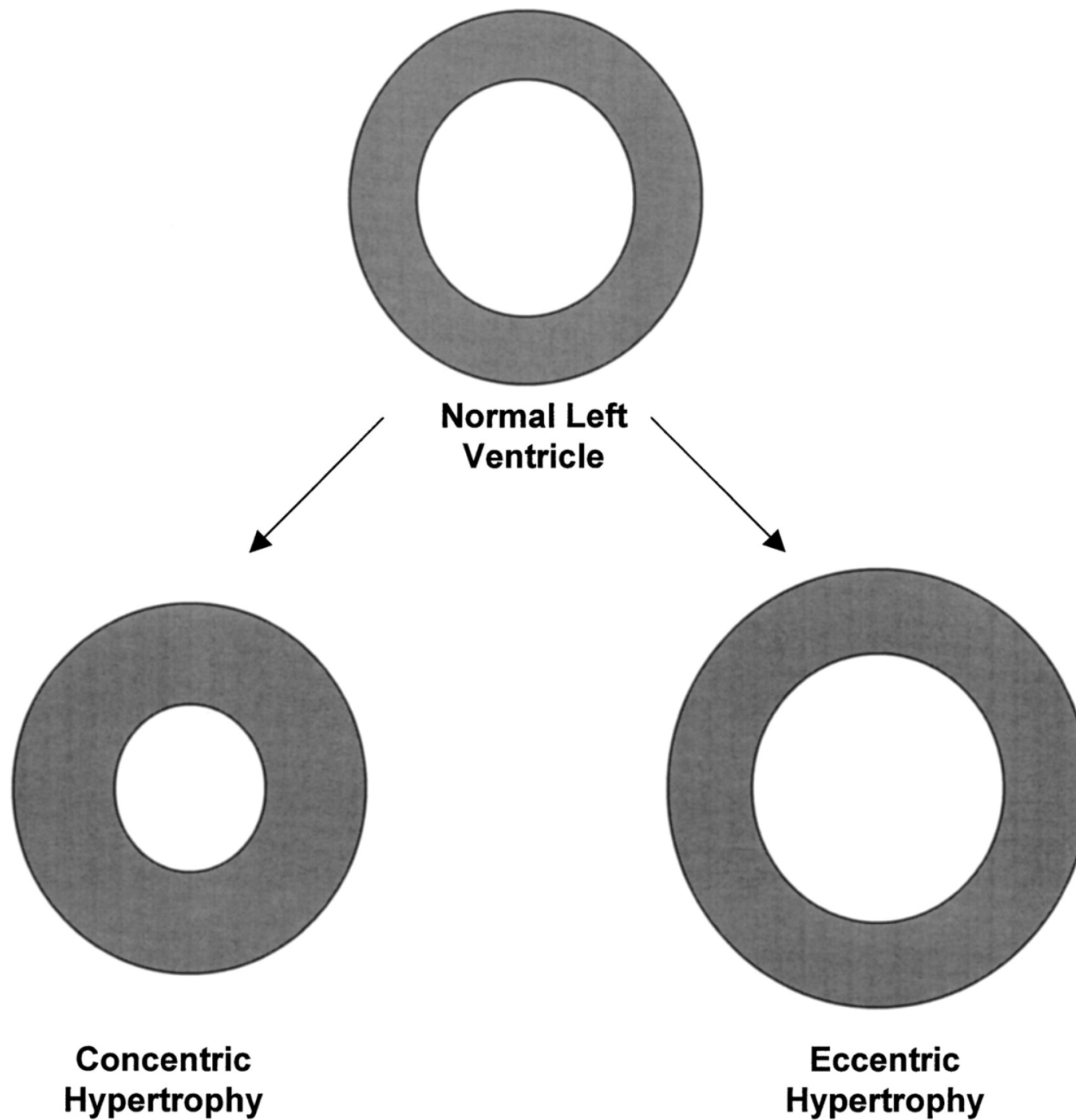
Concept of the "DNA Book." Kawai, J. and Y. Hayashizaki, DNA book. Genome Res, 2003. 13(6B): p. 1488-95.



# EXTRACTING THE CORE MESSAGE

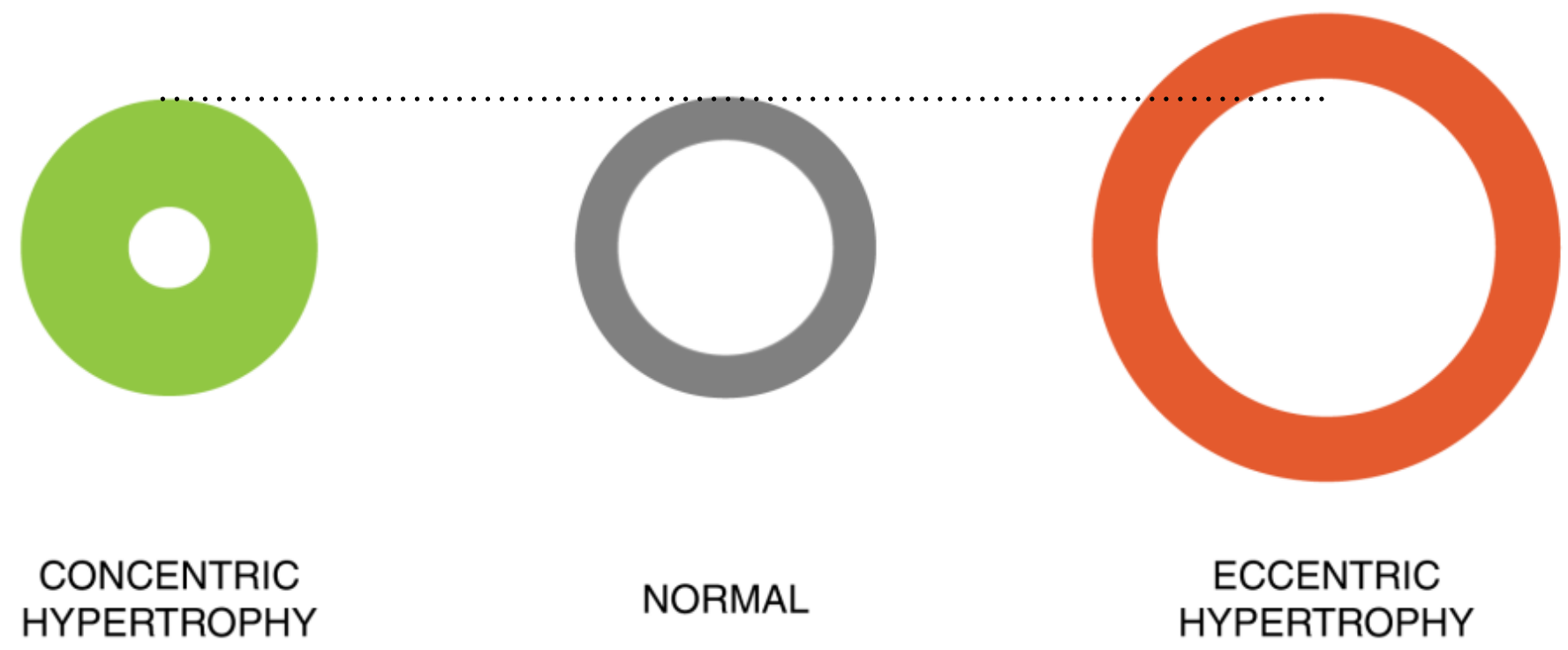
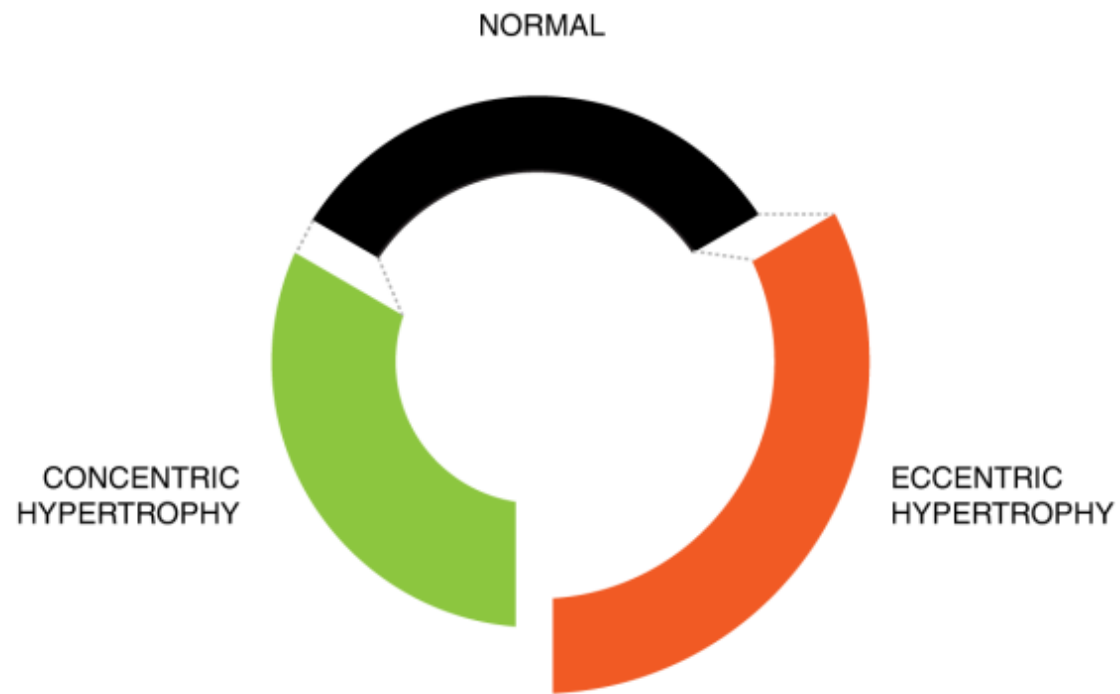
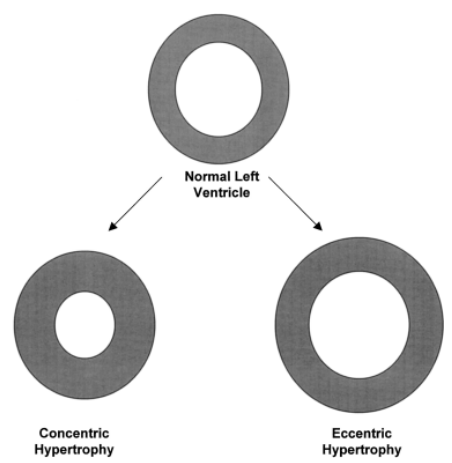


# AMBIGUOUS PROPORTIONS

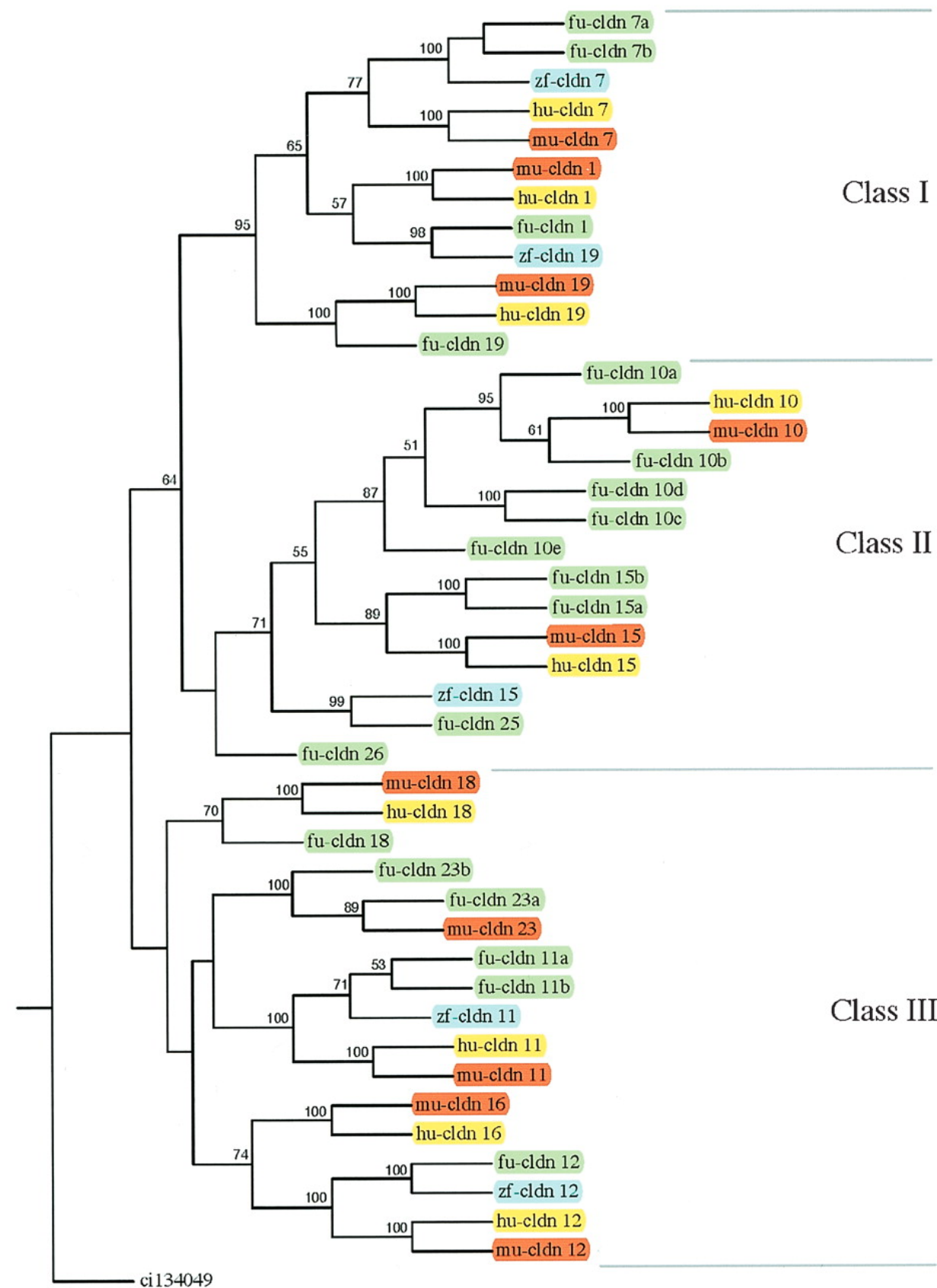


Nadeau, J.H., et al., Pleiotropy, homeostasis, and functional networks based on assays of cardiovascular traits in genetically randomized populations. *Genome Res*, 2003. 13(9): p. 2082-91.

# CLEAR PROPORTIONS



# REDUNDANCY



All labels are of the form

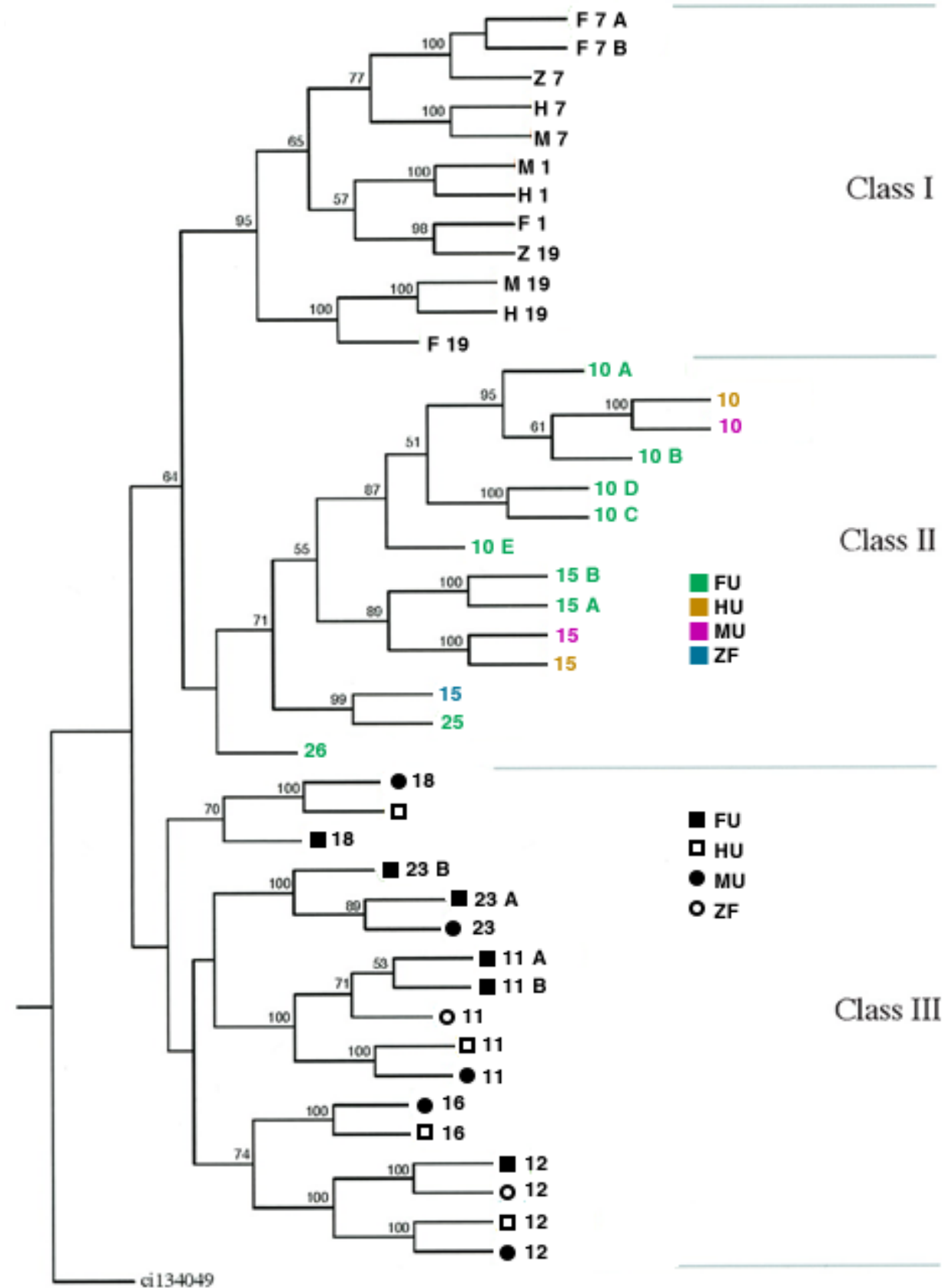
fu-cldn 7a  
fu-cldn 7b  
zf-cldn 7  
mu-cldn 19  
hu-cldn 15  
...

Instead, consider

F7a  
F7b  
Z7  
M19  
H15  
...

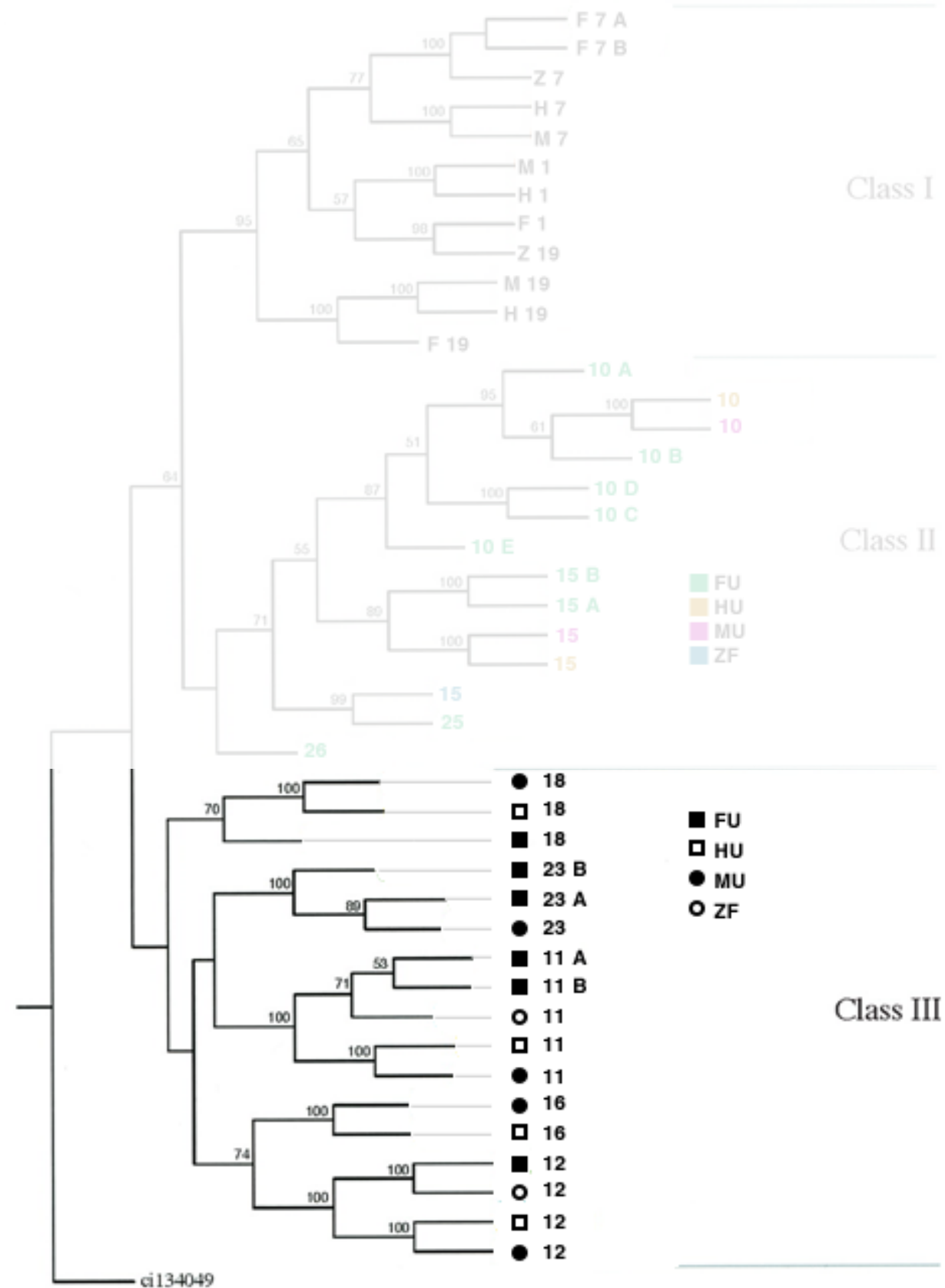
Y. H. Loh, A. Christoffels, S. Brenner, W. Hunziker, B. Venkatesh, Genome Res 14, 1248 (Jul, 2004).

# REFACTORED REDUNDANCY



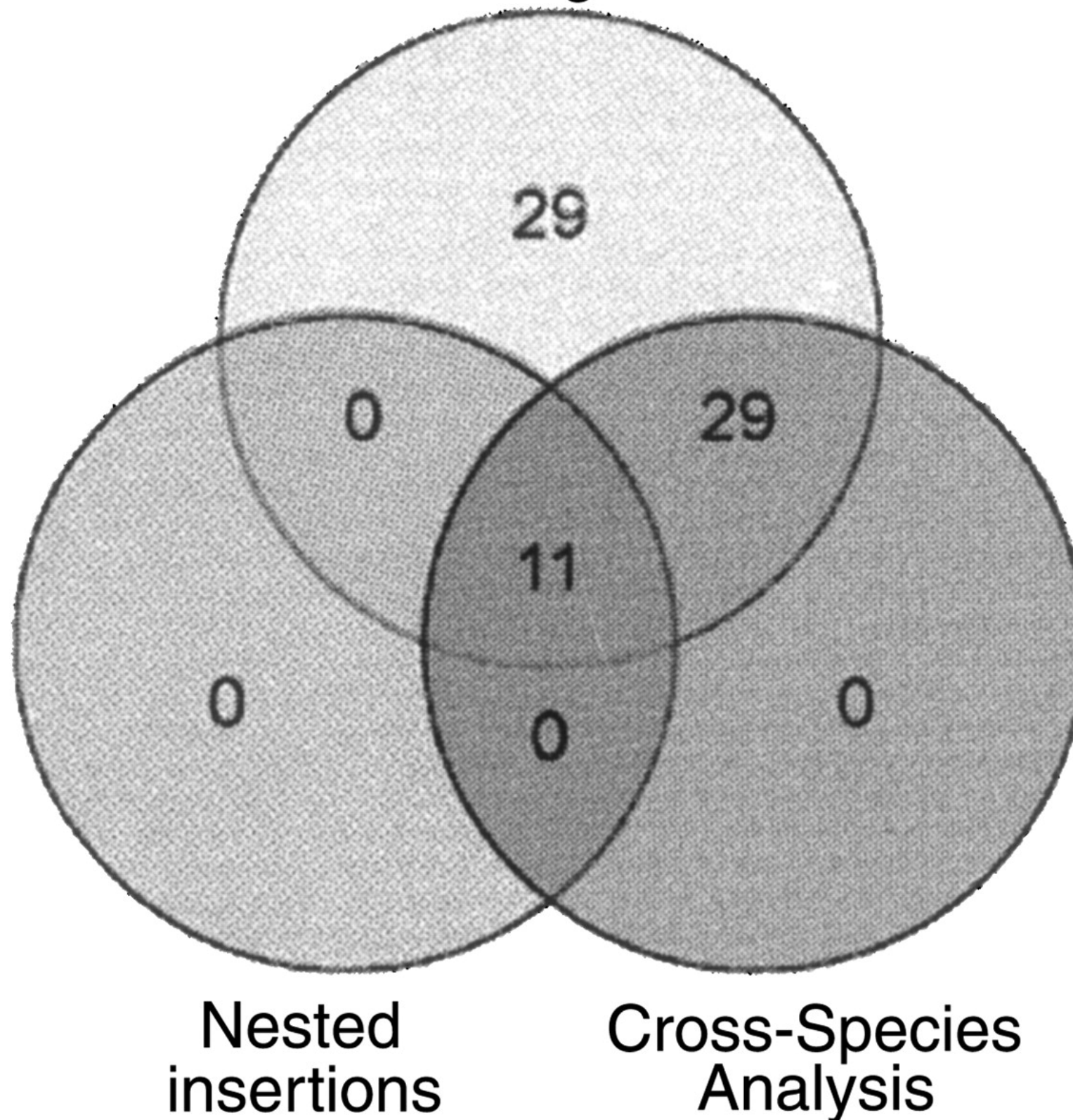


# REFACTORED REDUNDANCY



# INAPPROPRIATE REPRESENTATION

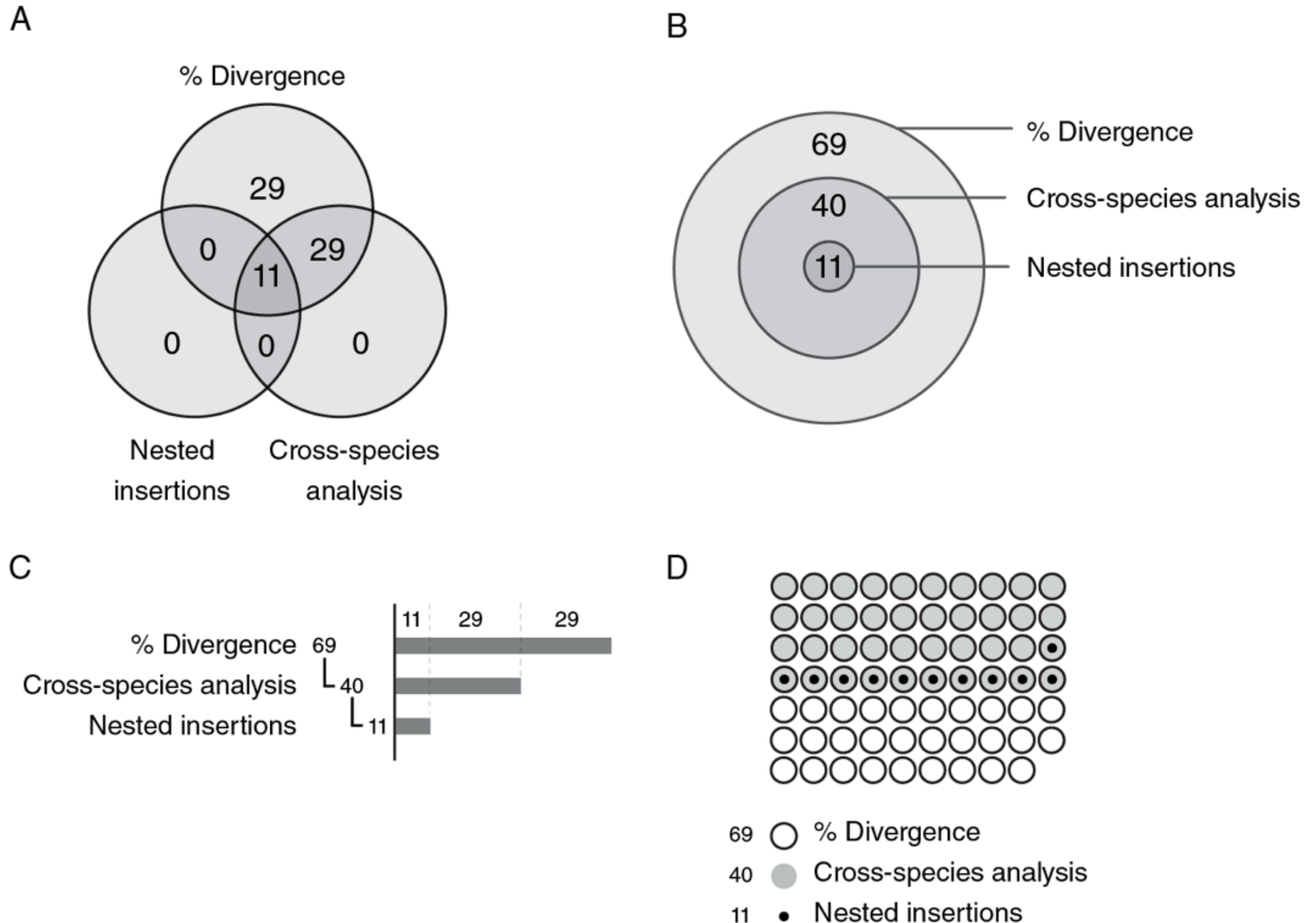
% Divergence



The traditional Venn diagram representation is too complex for this data set. It is not obvious that the data are nested.

Pace, J.K., 2nd and C. Feschotte, The evolutionary history of human DNA transposons: evidence for intense activity in the primate lineage. *Genome Res*, 2007. 17(4): p. 422-32.

# CHOOSE A REPRESENTATION FOR THE TASK

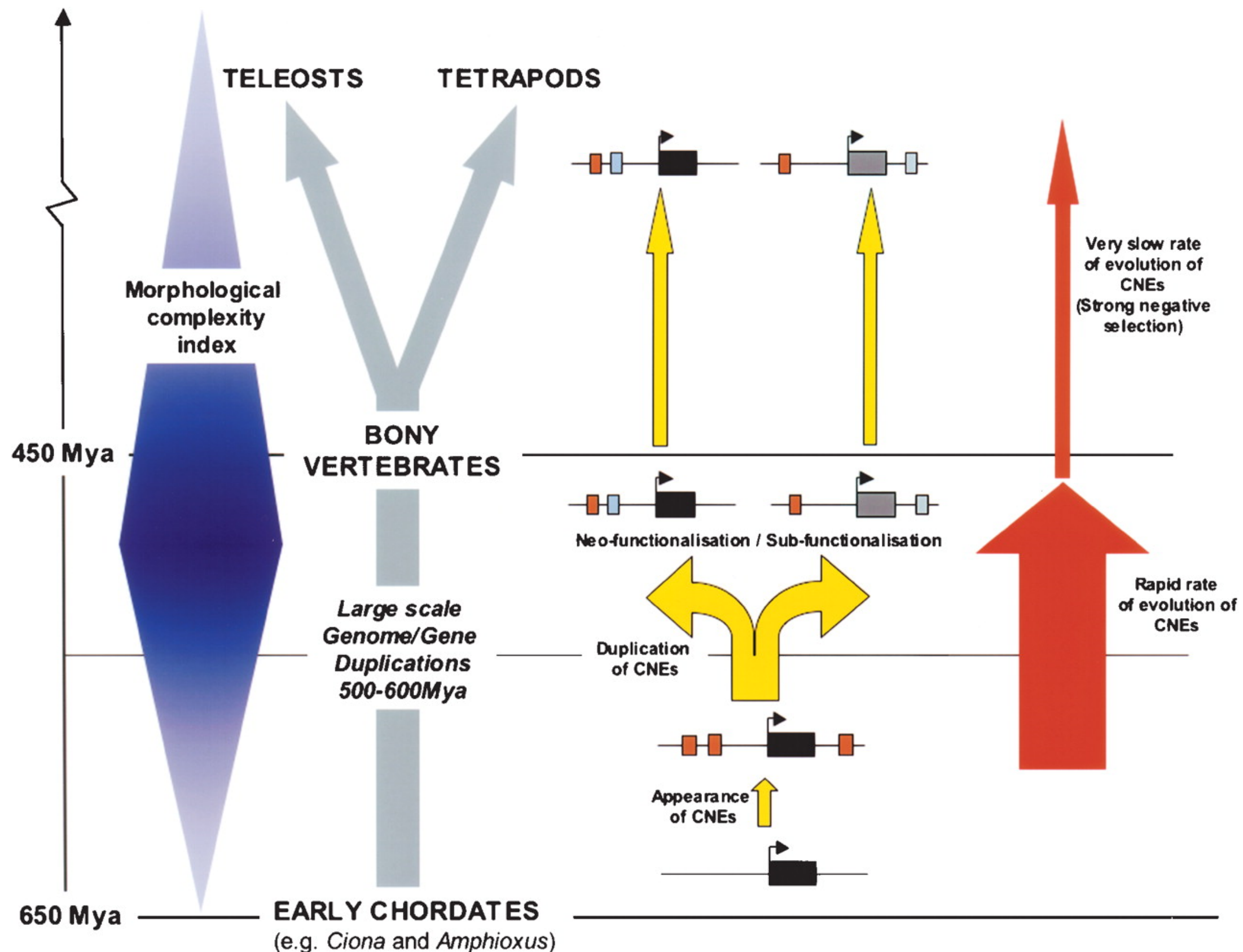


The original figure (A) can be represented to (B) emphasize its nested property, (C) the quantitative difference between values and (D) the qualitative difference between values.

Panel (A) from Pace, J.K., 2nd and C. Feschotte, The evolutionary history of human DNA transposons: evidence for intense activity in the primate lineage. *Genome Res*, 2007. 17(4): p. 422-32.



# CONSISTENCY BURDENS THE BRAIN

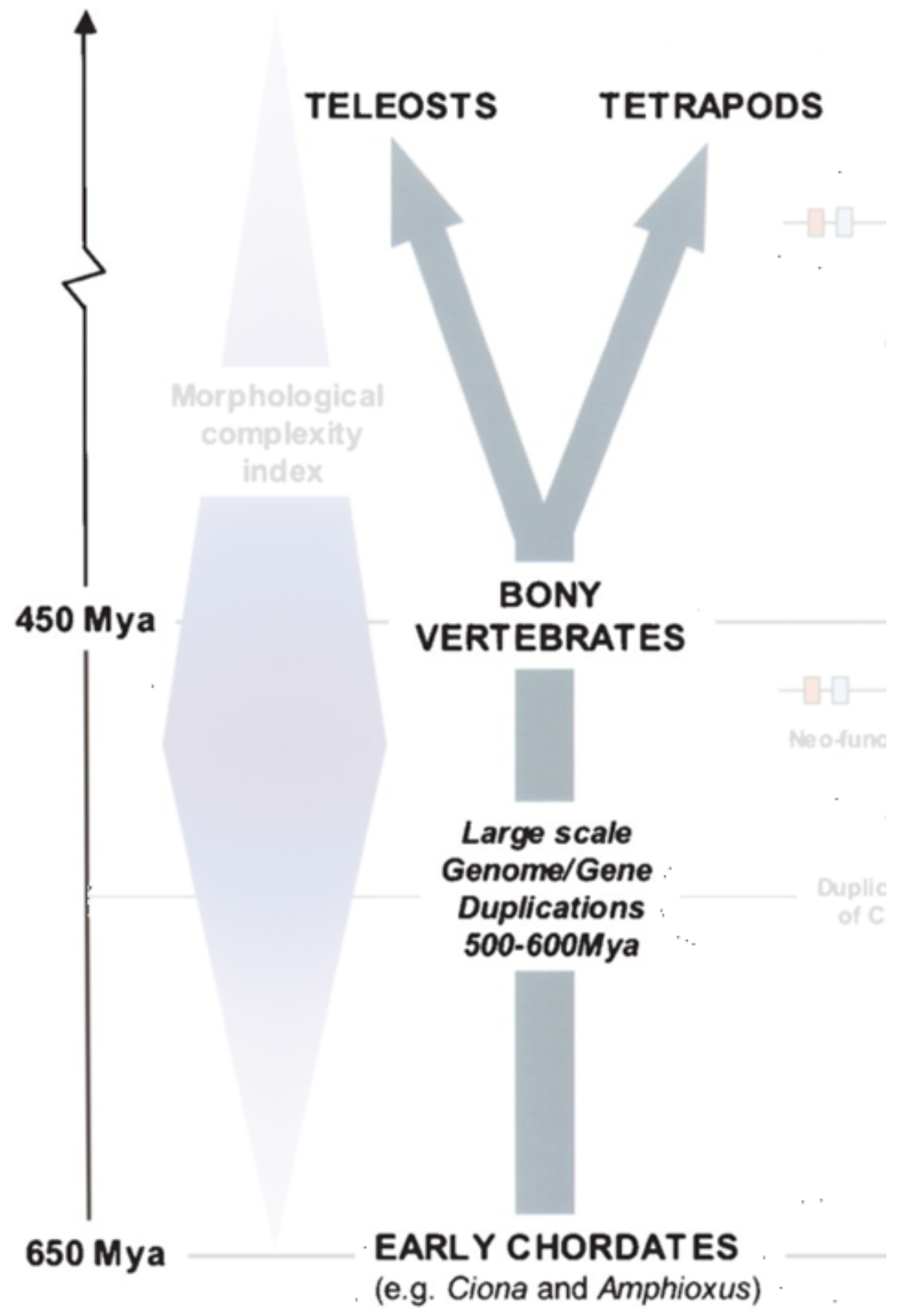
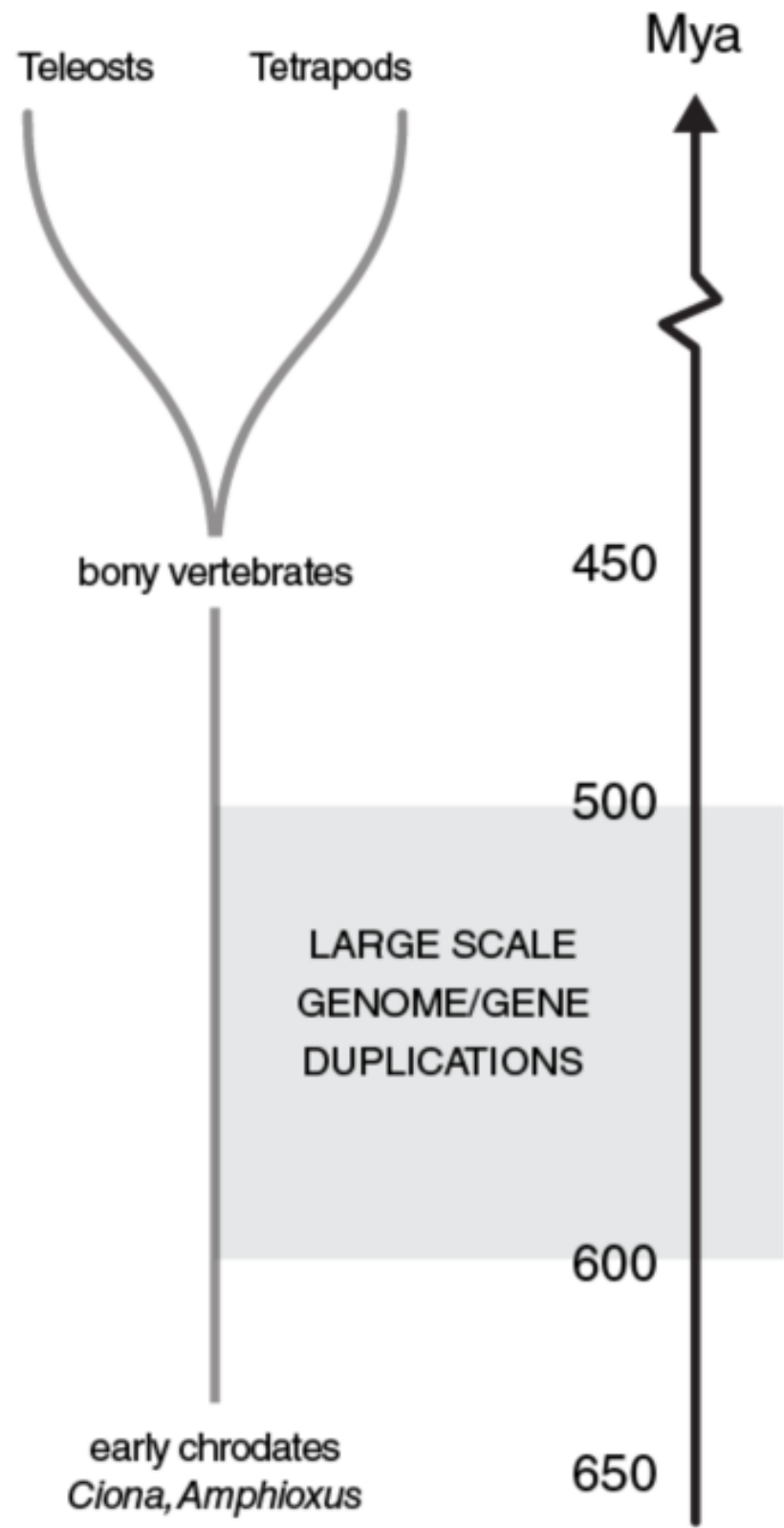


When a figure uses different visual elements to indicate the same type of relationship, process or callout, the reader is left to reduce the complexity while parsing the figure. This imposes a large cognitive burden and interferes with assimilating the information.

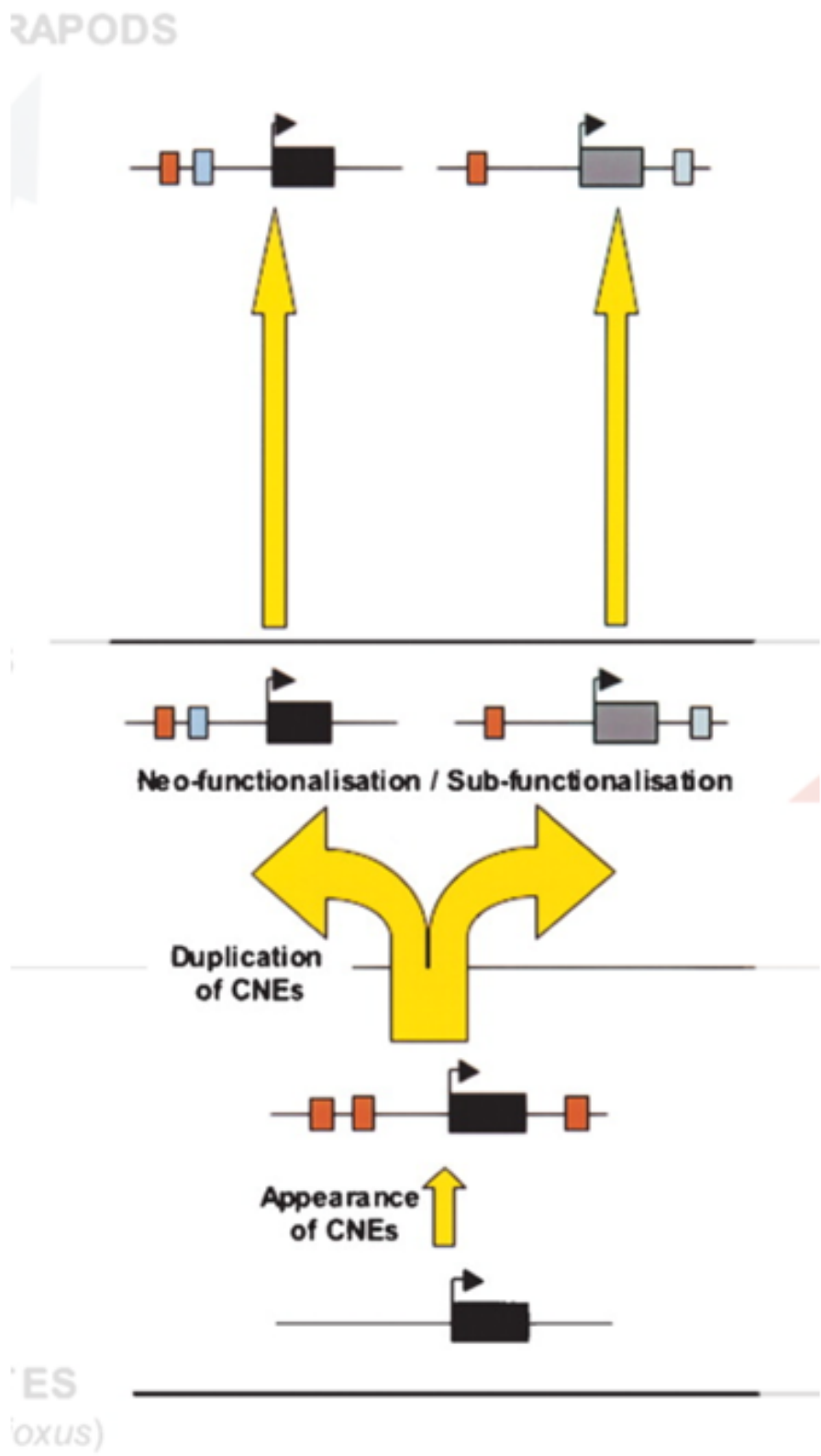
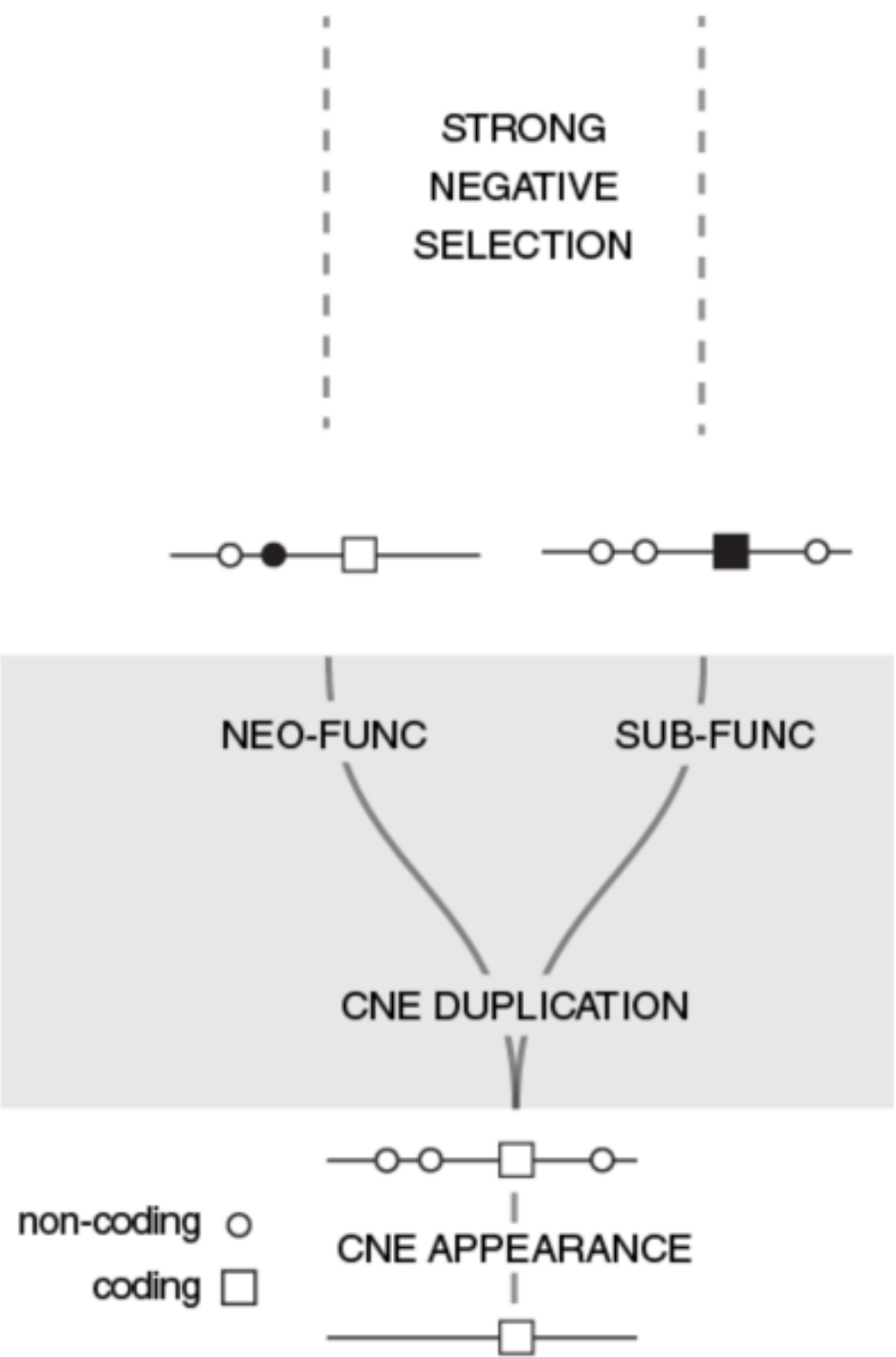
McEwen, G.K., et al., Ancient duplicated conserved noncoding elements in vertebrates: a genomic and functional analysis. *Genome Res*, 2006. 16(4): p. 451-65.



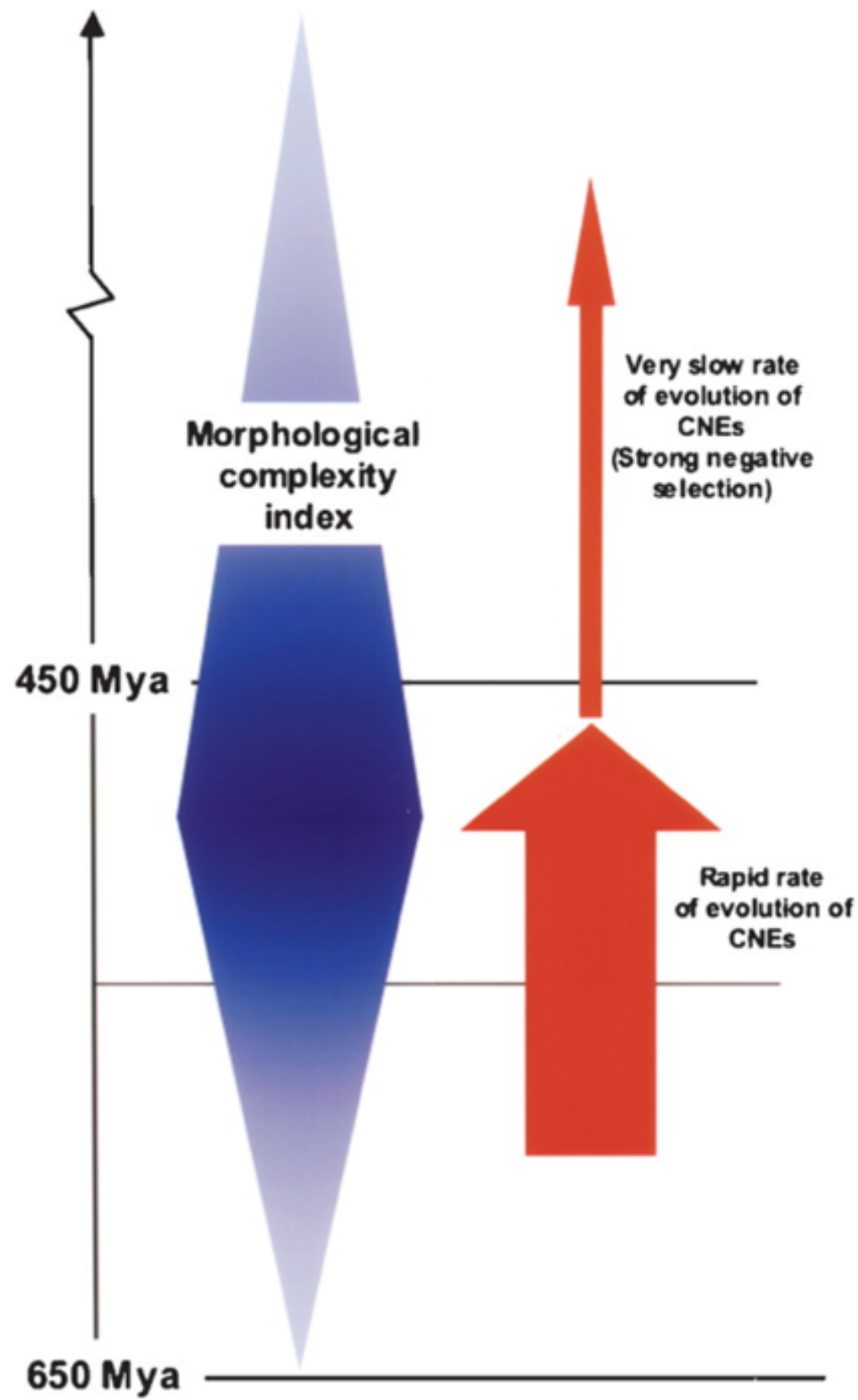
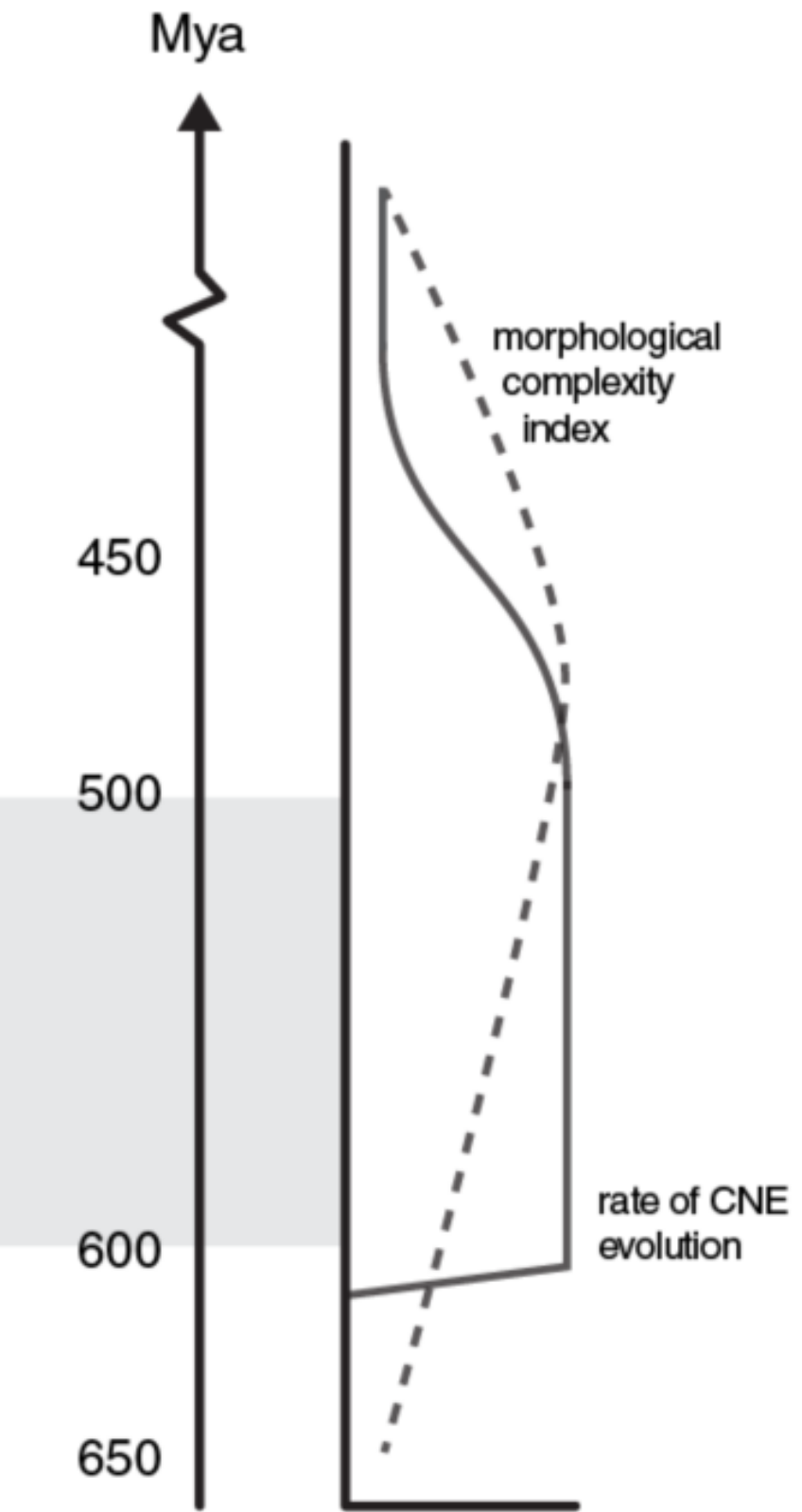
# SEPARATE CONTEXT FROM DATA



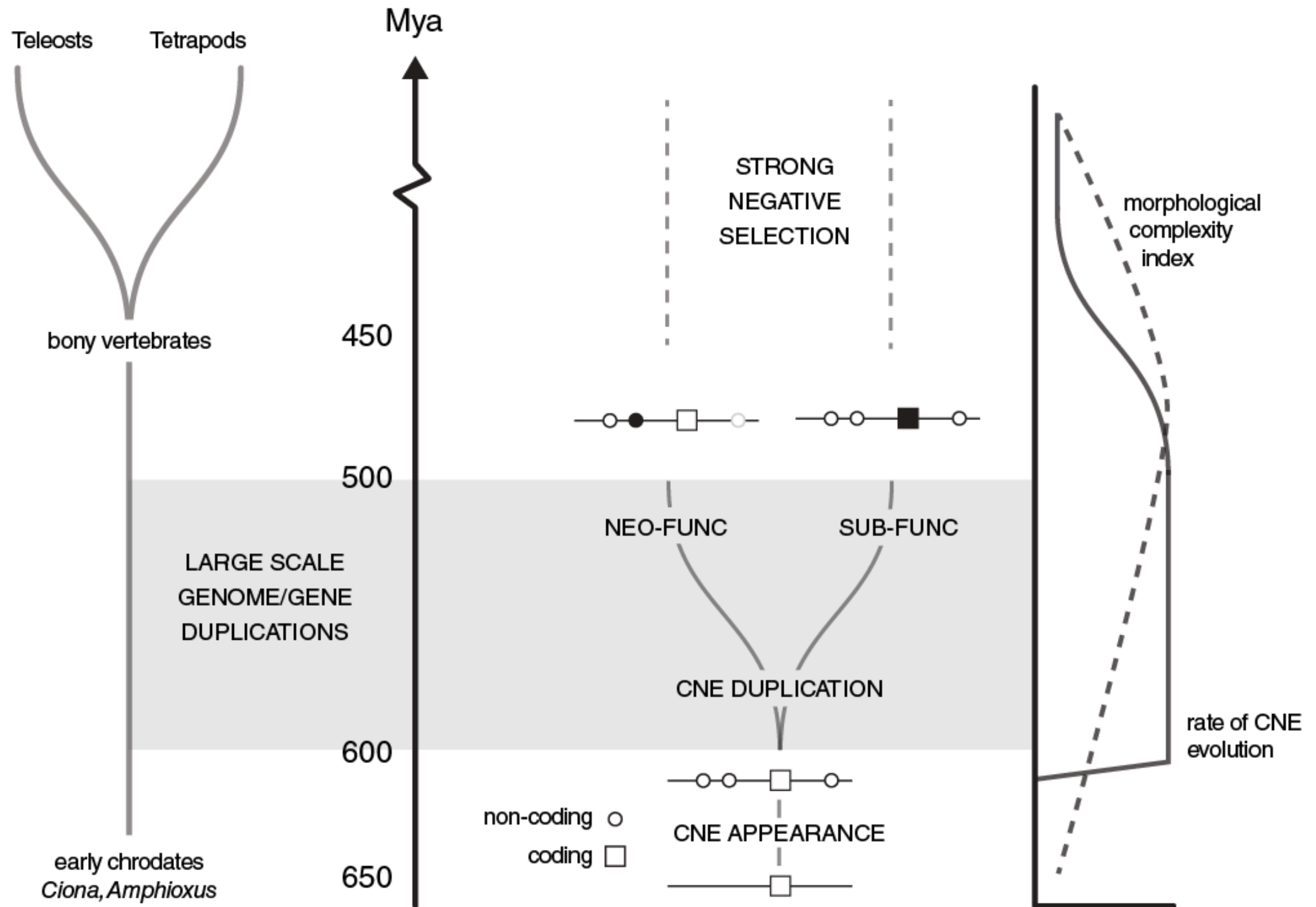
# SIMPLIFY VOCABULARY - NO SYNONYMS!



# SAME DATA TYPE - SAME REPRESENTATION



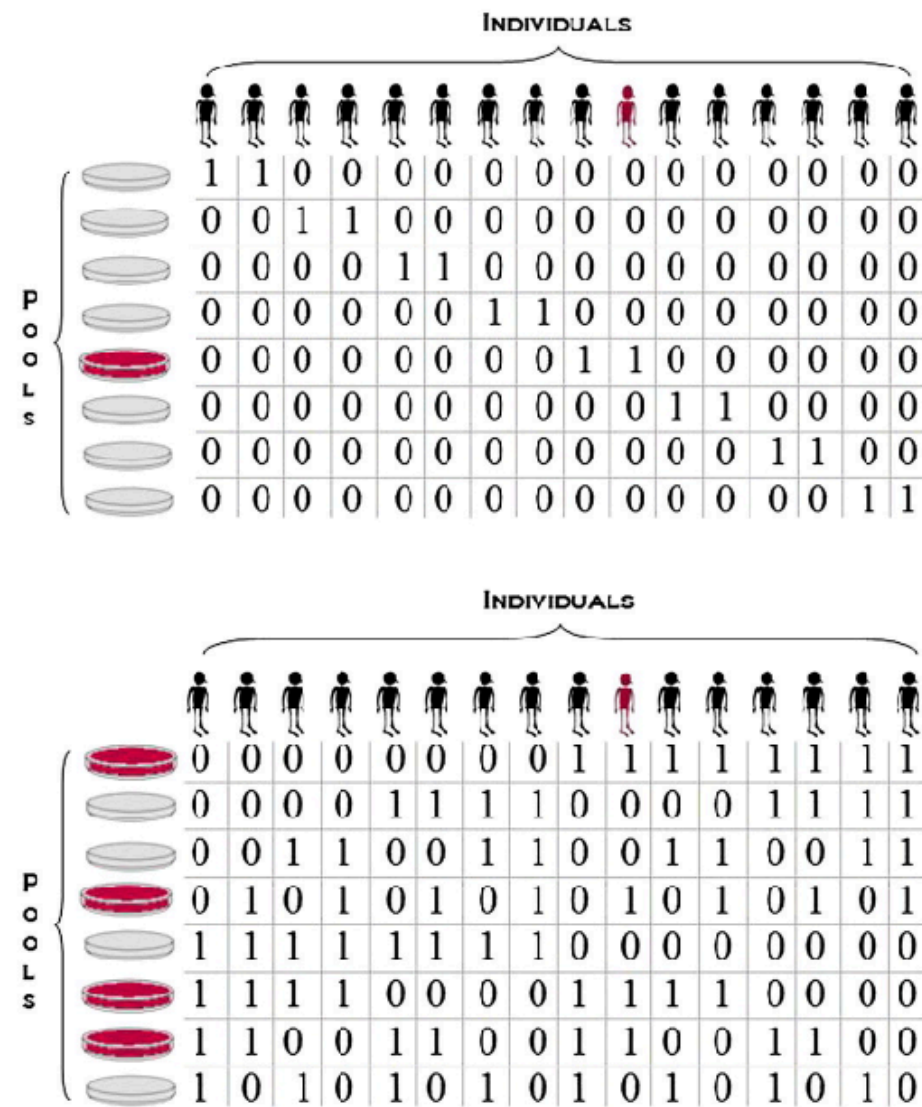
# REFACTORED INCONSISTENCY



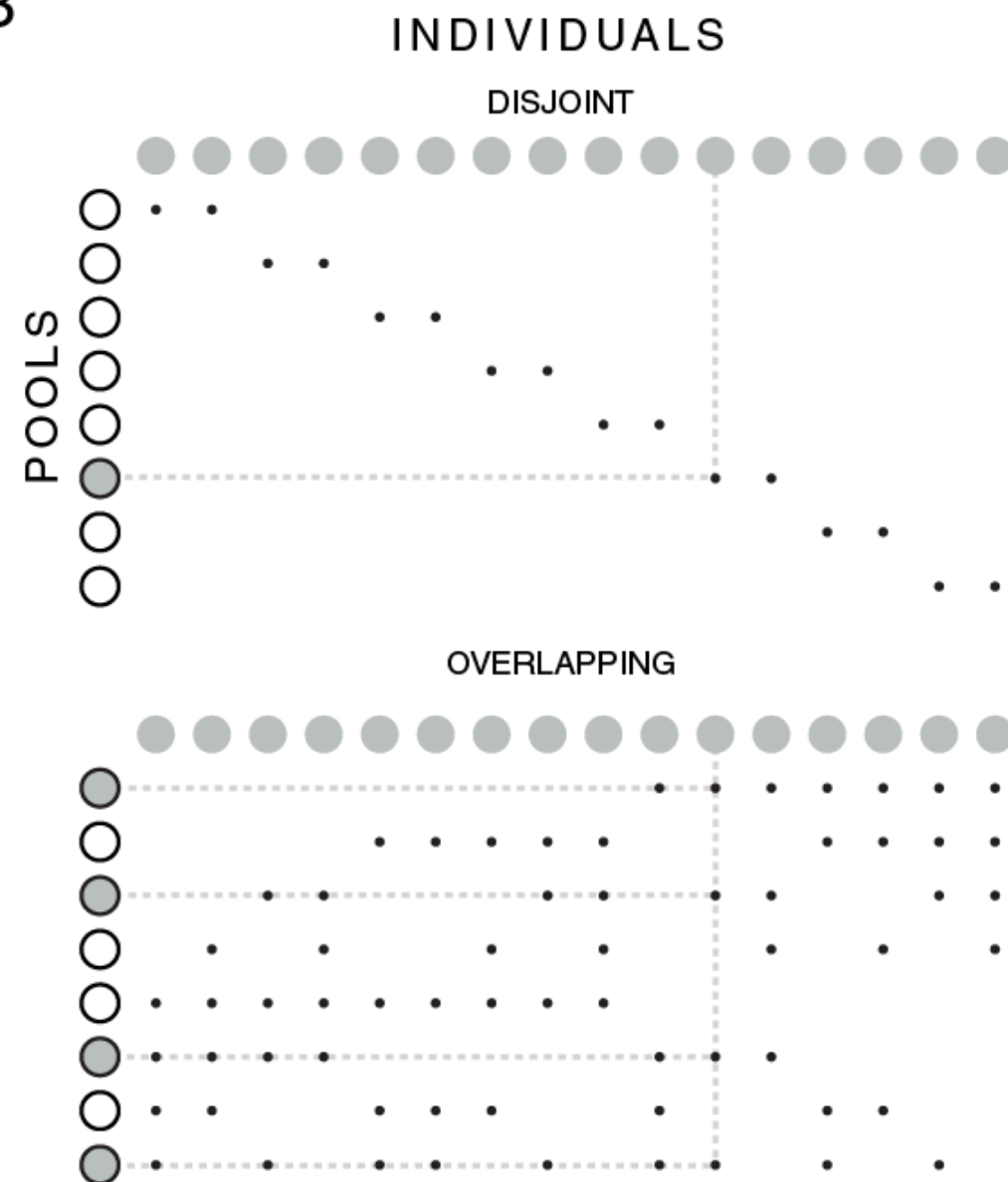


# EXCESS INK

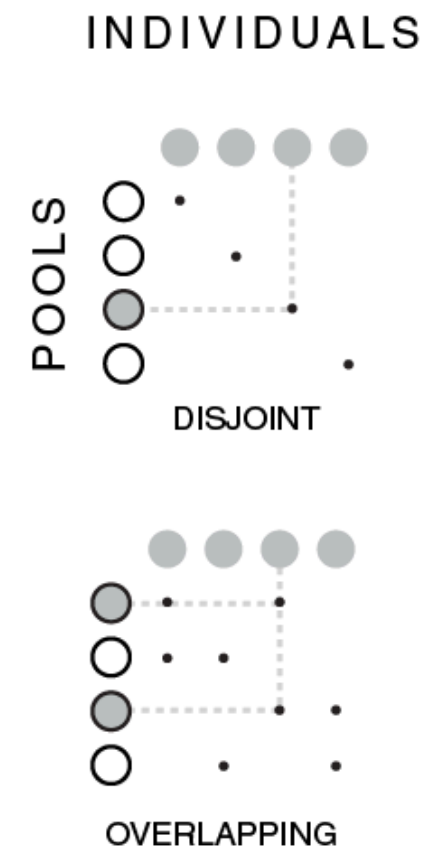
A



B



C

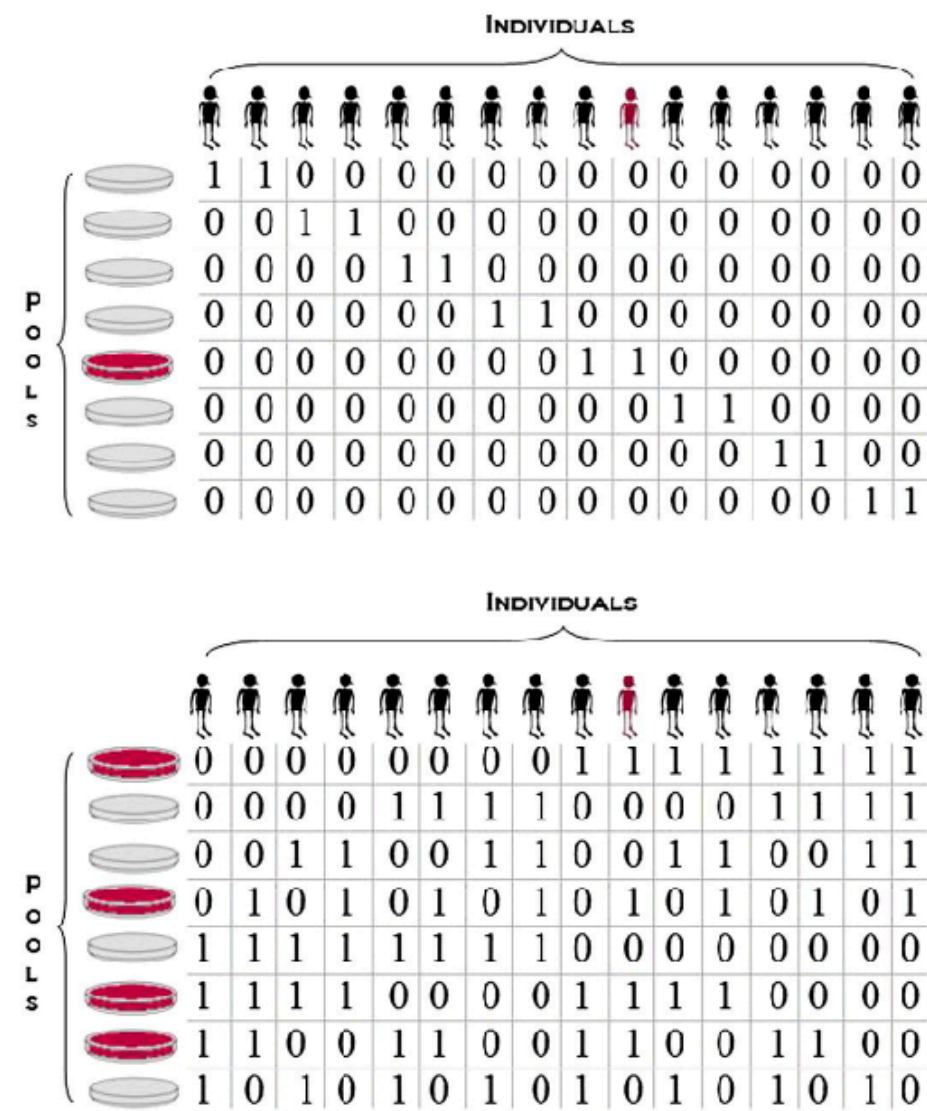


Limit the amount of ink to show data. Use the bare minimum to keep the message clear and figure uncluttered. Make use of space to indicate missing (or in this case zero) data.

Panel (A) from Prabhu, S. and I. Pe'er, Overlapping pools for high-throughput targeted resequencing. Genome Res, 2009. 19(7): p. 1254-61.

# EXCESS INK

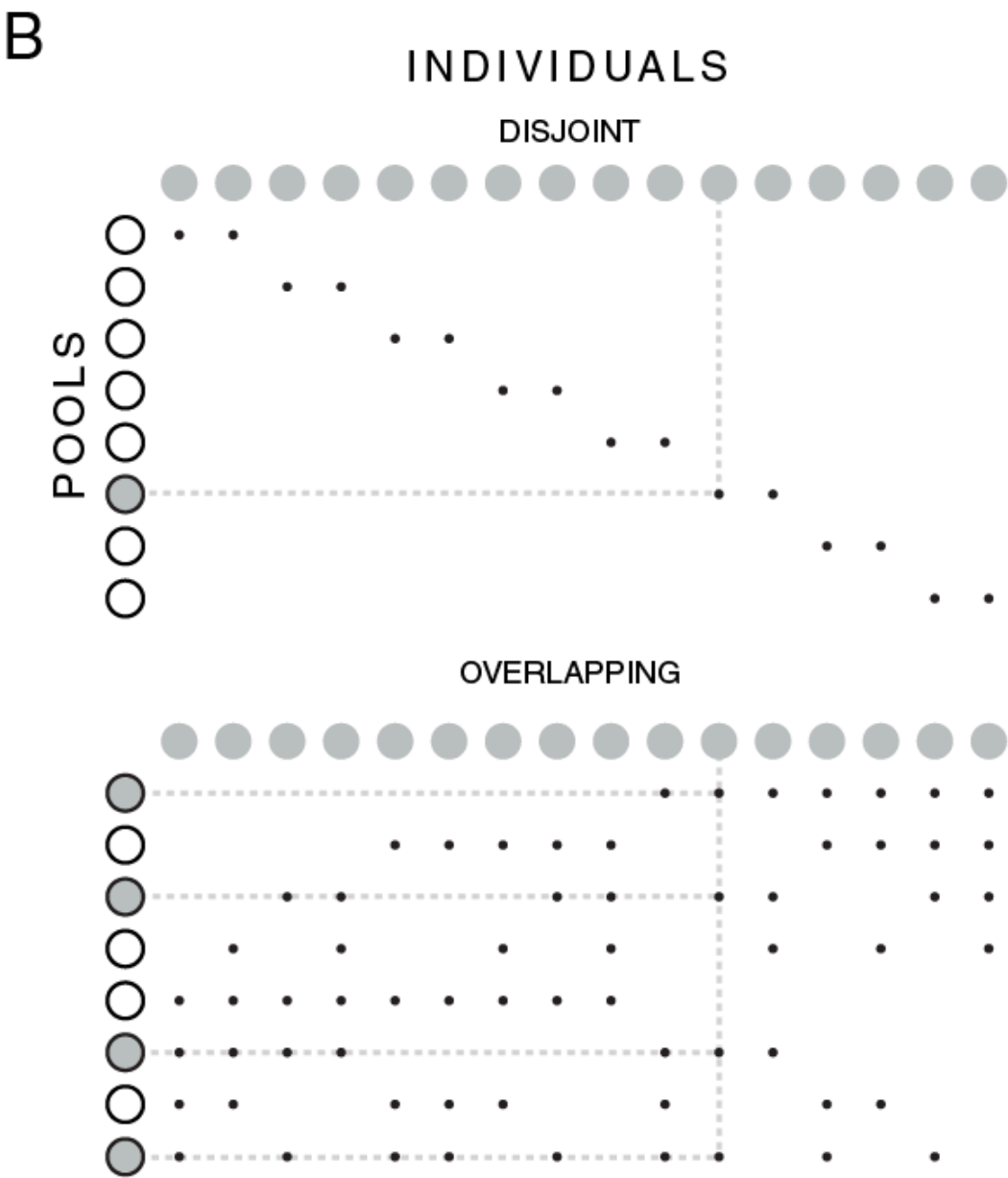
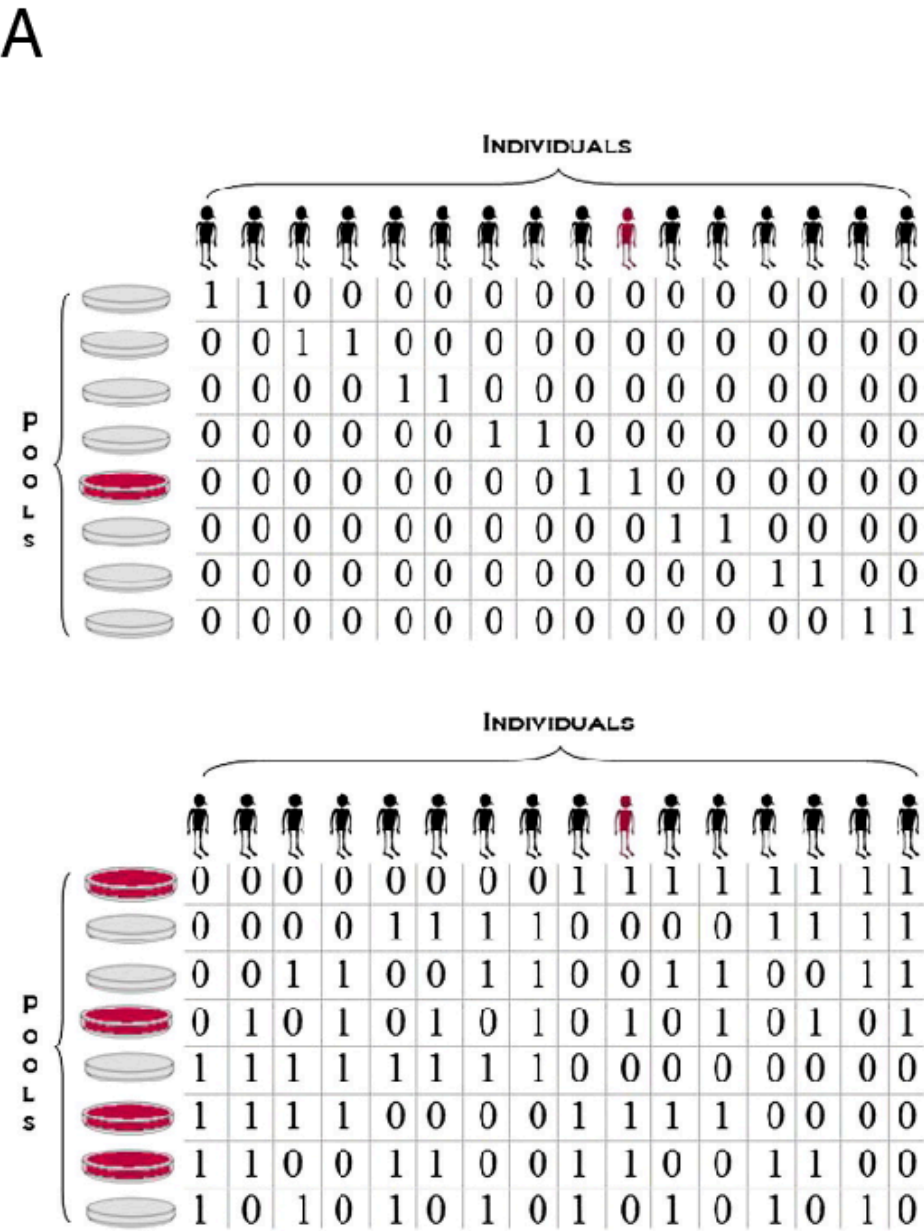
A



Limit the amount of ink to show data. Use the bare minimum to keep the message clear and figure uncluttered. Make use of space to indicate missing (or in this case zero) data.

Panel (A) from Prabhu, S. and I. Pe'er, Overlapping pools for high-throughput targeted resequencing. Genome Res, 2009. 19(7): p. 1254-61.

# EXCESS INK

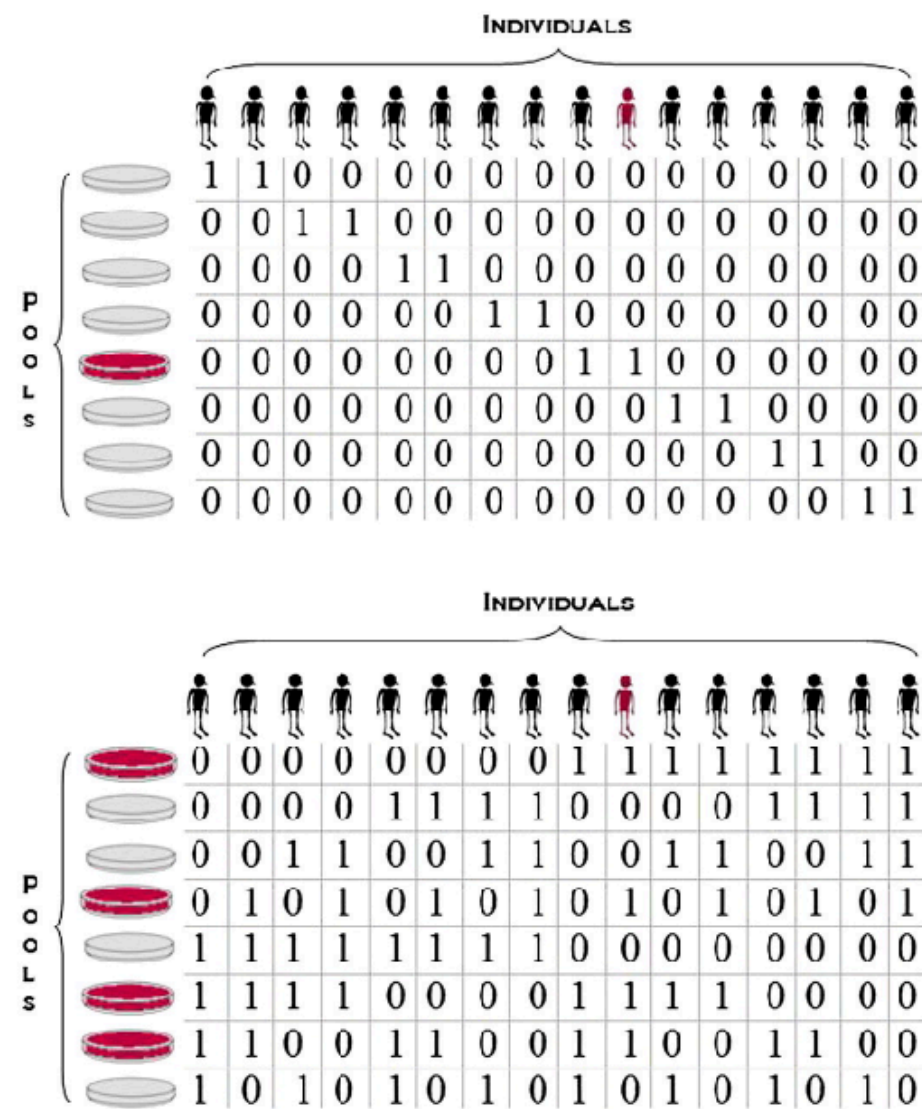


Limit the amount of ink to show data. Use the bare minimum to keep the message clear and figure uncluttered. Make use of space to indicate missing (or in this case zero) data.

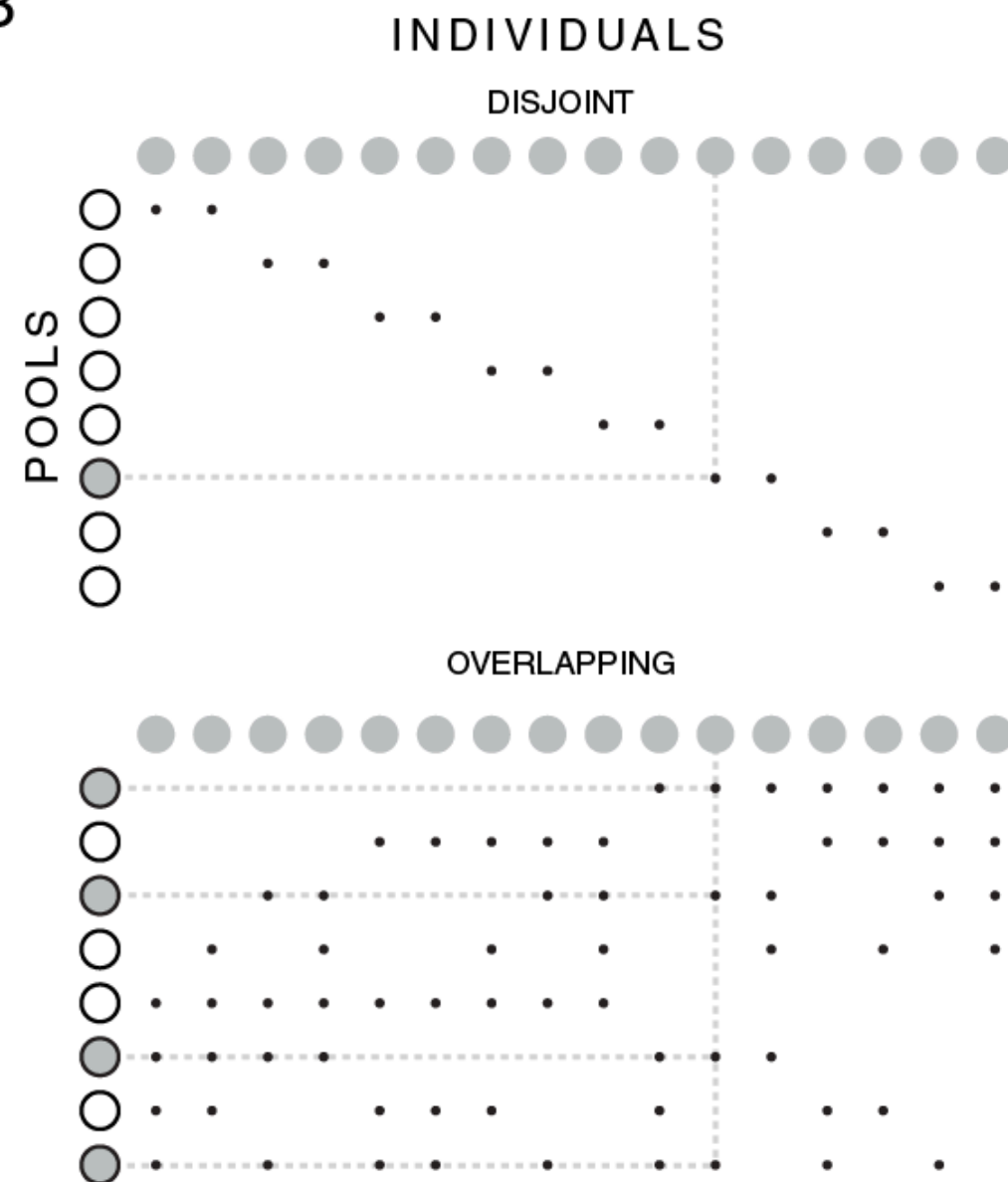
Panel (A) from Prabhu, S. and I. Pe'er, Overlapping pools for high-throughput targeted resequencing. Genome Res, 2009. 19(7): p. 1254-61.

# EXCESS INK

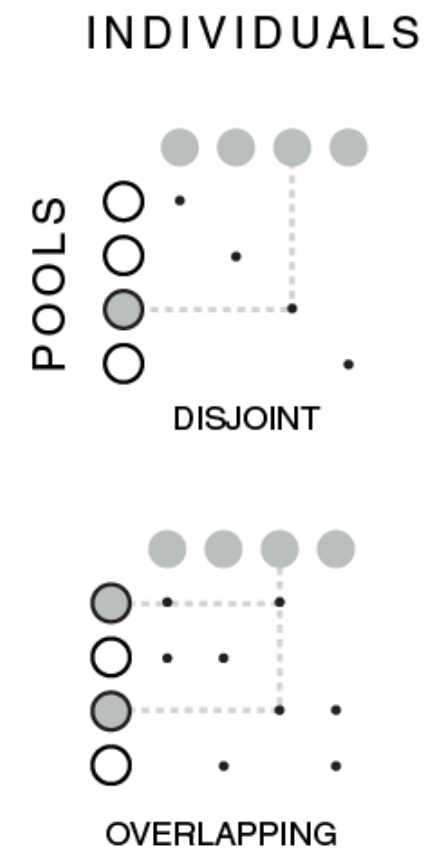
A



B



C

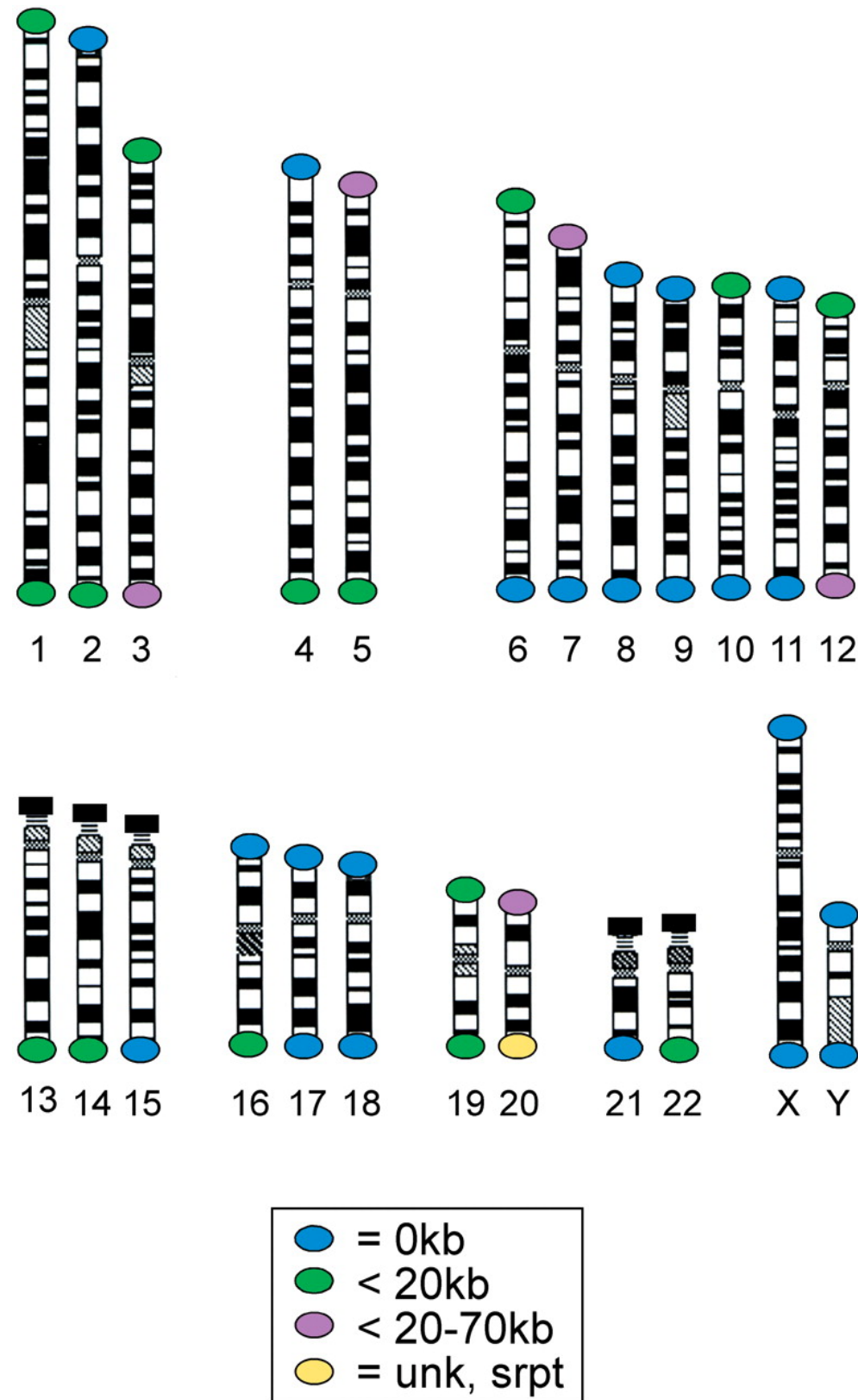


Limit the amount of ink to show data. Use the bare minimum to keep the message clear and figure uncluttered. Make use of space to indicate missing (or in this case zero) data.

Panel (A) from Prabhu, S. and I. Pe'er, Overlapping pools for high-throughput targeted resequencing. Genome Res, 2009. 19(7): p. 1254-61.



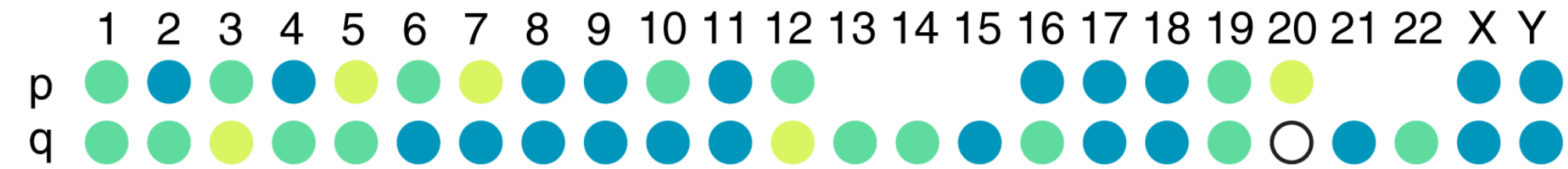
# EXCESS INK



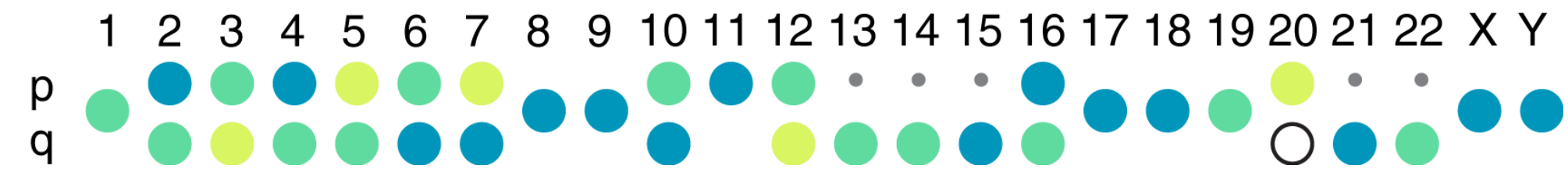
Telomere sequence gaps. Riethman, H., et al., Mapping and initial analysis of human subtelomeric sequence assemblies. Genome Res, 2004. 14(1): p. 18-28.

# REMOVING EXCESS INK

## OPTION 1

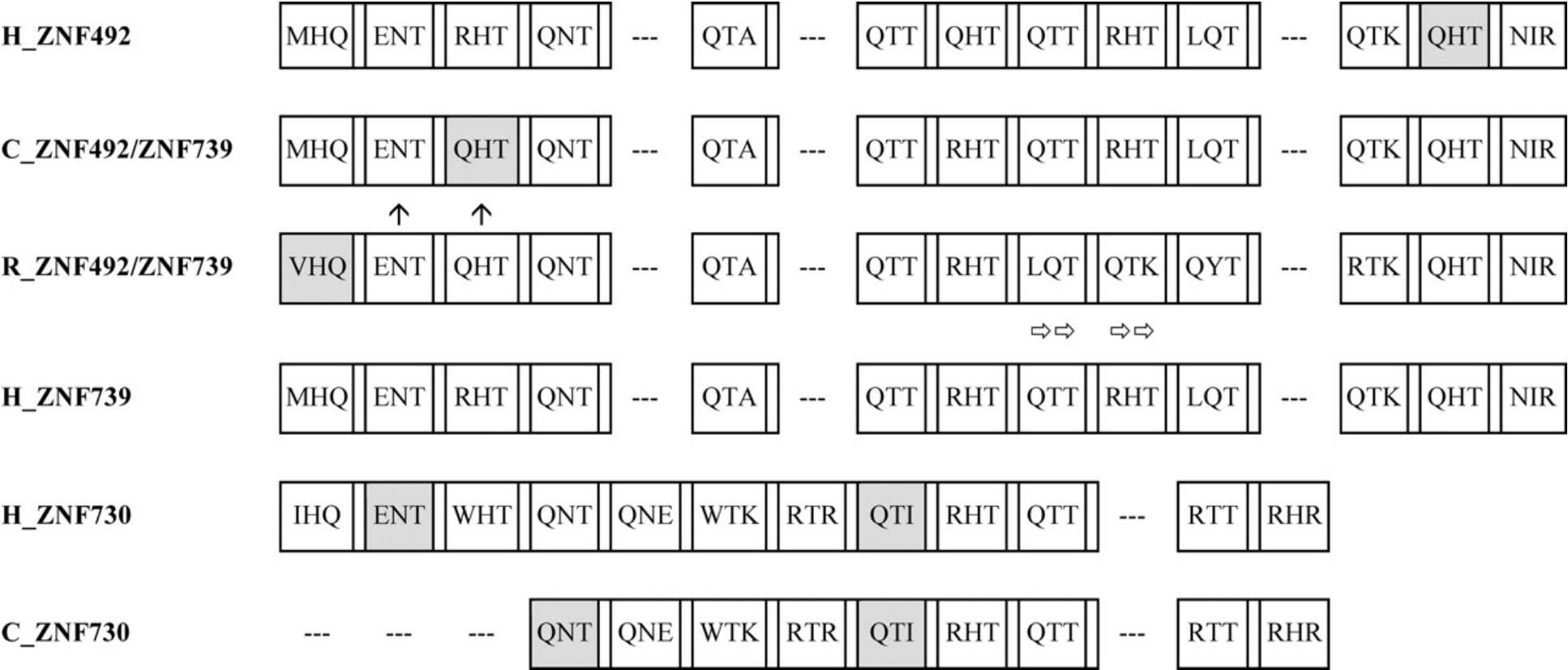


## OPTION 2



- no gap ●
- <20kb ●
- <20-70kb ●
- unk/srp ○
- acrocentric ●

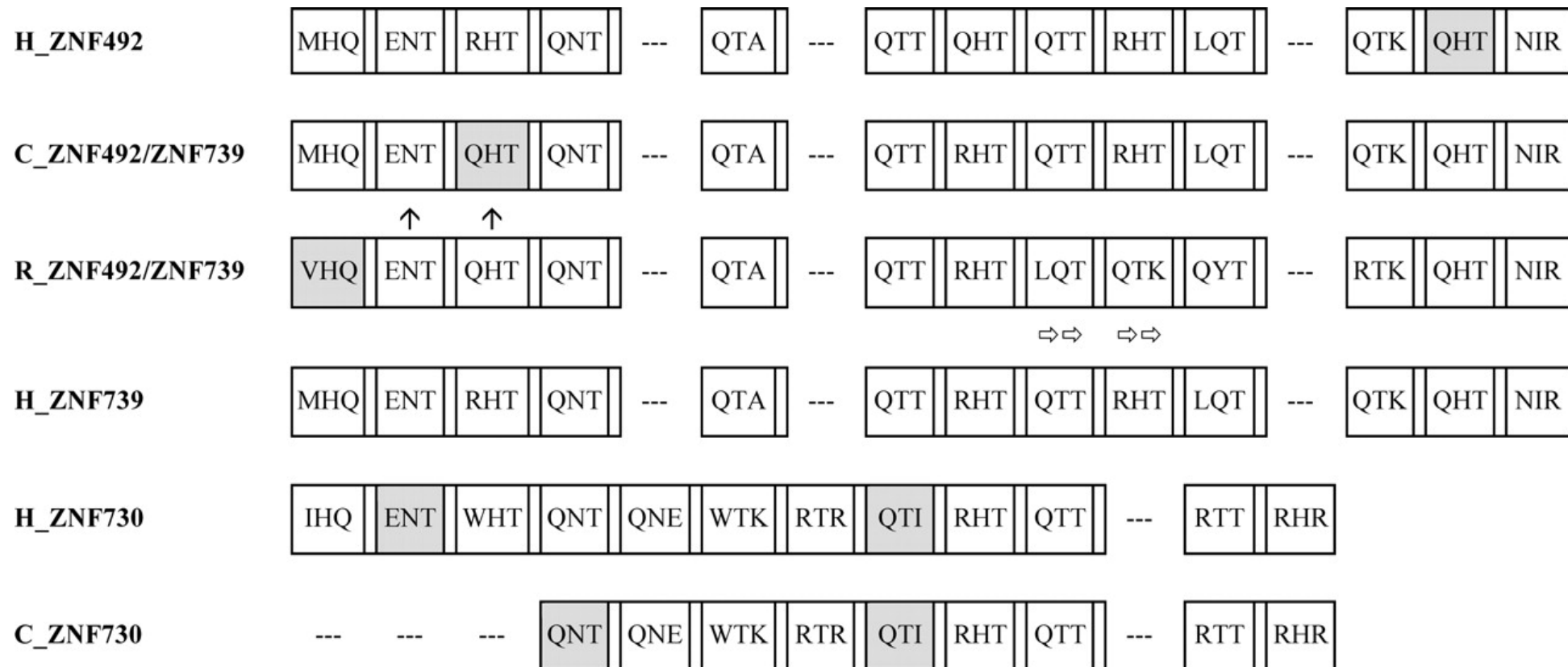
# EXCESS INK



	MHQ	ENT	RHT	QNT	QNE	QTA	RTR	QTT	QHT	QTT	RHT	LQT	RHR	QTK	QHT	NIR
H492	...	...	...	...		...		...	...	...	...	...		...	...	...
C492/739	...	...	Q..	...		...		...	R..	...	...	...		...	...	...
R492/739	V..	...	Q..	...		...		...	R..	LQ.	QTK	QY.		R..	...	...
H739	M..	...	...	...		...		...	R..	...	QTK	...		...	...	...
H730	I..	...	W..	...	...	W.K	...	..I	R..	...		RT.	...			
C730				...	...	W.K	...	..I	R..	...		RT.	...			

Zinc finger exon analysis for ZNF493 and ZNF738, two divergent genes from the ZNF431 clade. Hamilton, A.T., et al., Evolutionary expansion and divergence in the ZNF91 subfamily of primate-specific zinc finger genes. Genome Res, 2006. 16(5): p. 584-94.

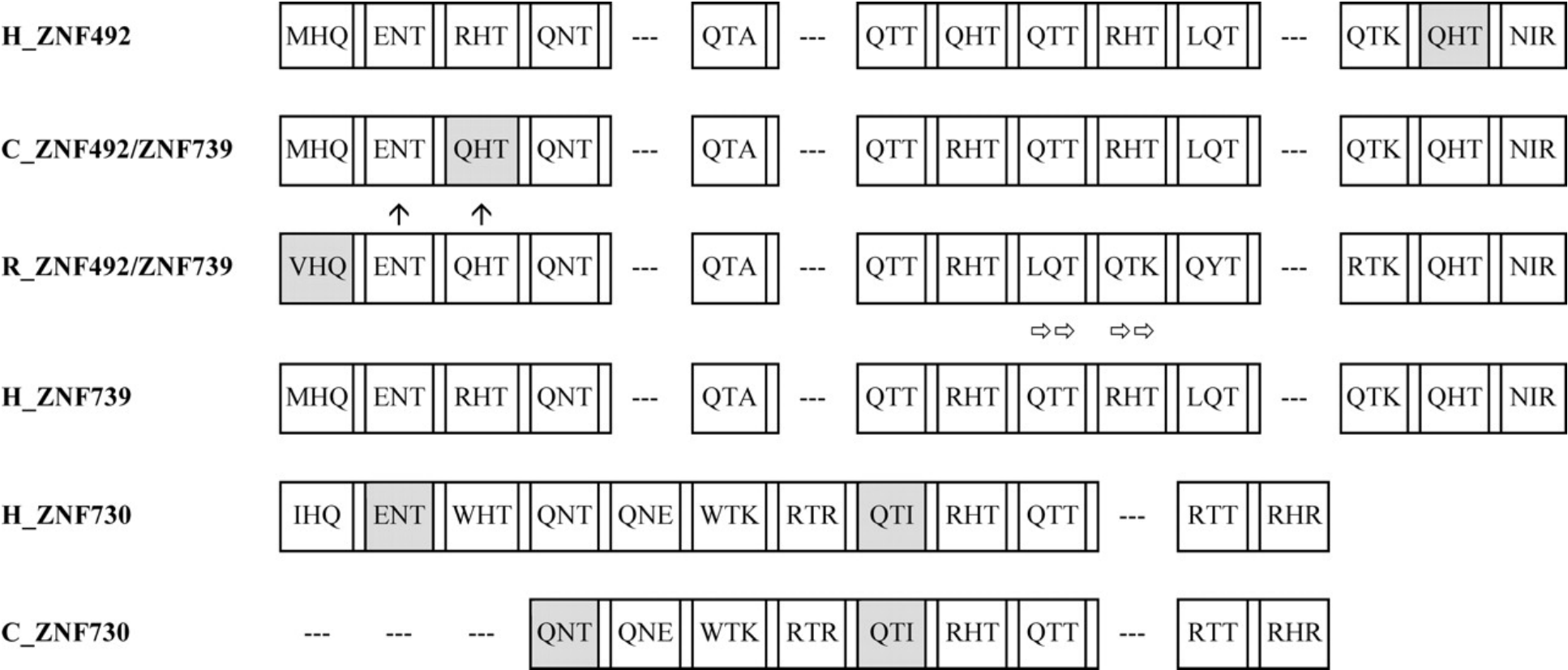
# EXCESS INK



Zinc finger exon analysis for ZNF493 and ZNF738, two divergent genes from the ZNF431 clade. Hamilton, A.T., et al., Evolutionary expansion and divergence in the ZNF91 subfamily of primate-specific zinc finger genes. *Genome Res*, 2006. 16(5): p. 584-94.



# EXCESS INK



	MHQ	ENT	RHT	QNT	QNE	QTA	RTR	QTT	QHT	QTT	RHT	LQT	RHR	QTK	QHT	NIR
H492	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
C492/739	...	...	Q..	...	...	...	...	...	R..	...	...	...	...	...	...	...
R492/739	V..	...	Q..	...	...	...	...	...	R..	LQ.	QTK	QY.	...	R..	...	...
H739	M..	...	...	...	...	...	...	...	R..	...	QTK	...	...	...	...	...
H730	I..	...	W..	...	...	W.K	...	..I	R..	...	...	RT.	...	...	...	...
C730	...	...	...	...	W.K	...	..I	R..	...	...	RT.	...	...	...	...	...

Zinc finger exon analysis for ZNF493 and ZNF738, two divergent genes from the ZNF431 clade. Hamilton, A.T., et al., Evolutionary expansion and divergence in the ZNF91 subfamily of primate-specific zinc finger genes. Genome Res, 2006. 16(5): p. 584-94.

# THE LEGEND IS THE MAP

- all genes
- ▲ specific in situ result
- △ selective in situ result
- ▽ nonselective in situ result
- ▲ specific in literature
- × nonspecific in literature

ISH results in relation to profile similarity to SM-MHC (Y-axis) and to expression level (X-axis). Nelander, S., P. Mostad, and P. Lindahl, Prediction of cell type-specific gene modules: identification and initial characterization of a core set of smooth muscle-specific genes. *Genome Res*, 2003. 13(8): p. 1838-54.

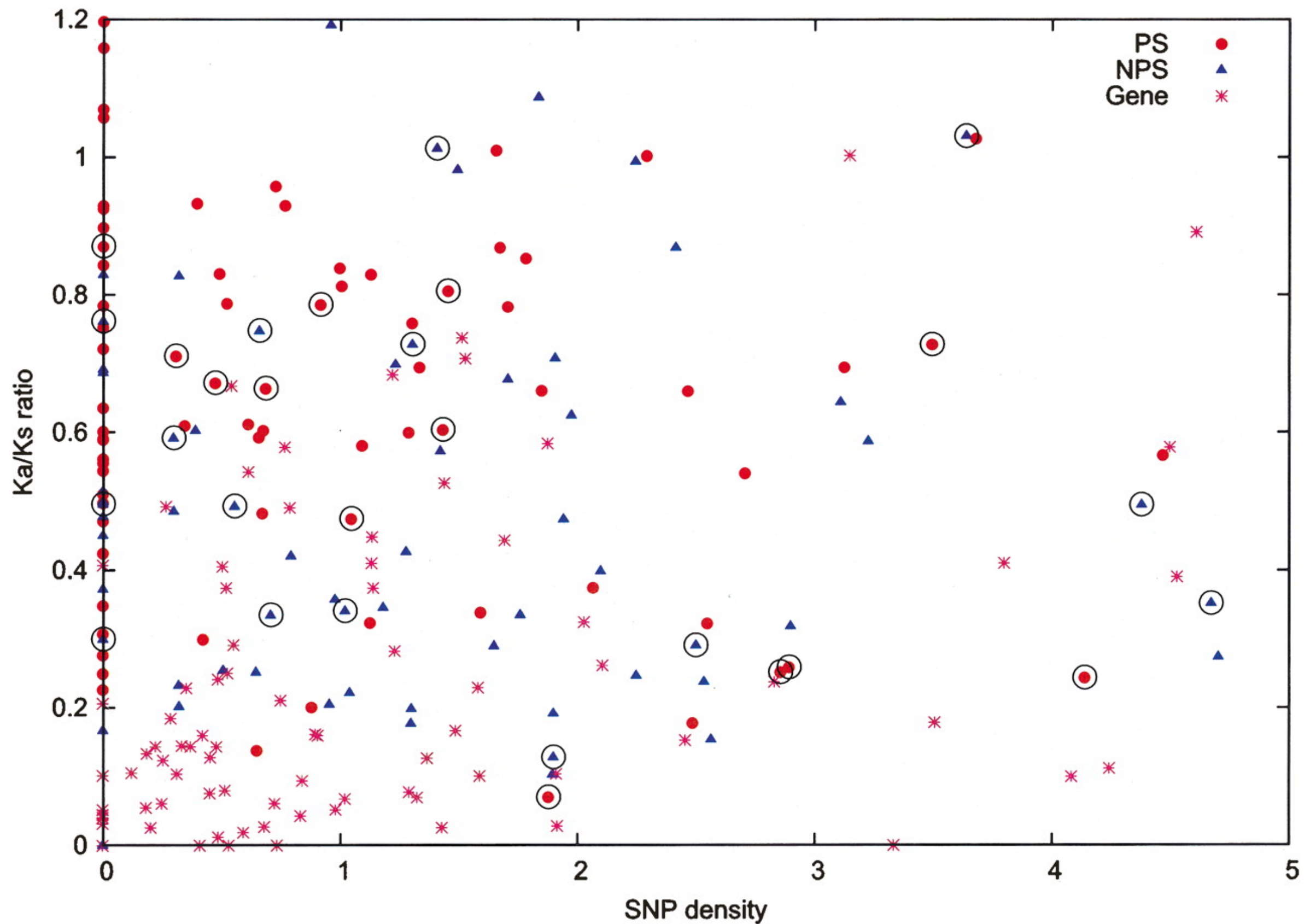
# THE LEGEND IS THE MAP

- all genes
- ▲ specific in situ result
- △ selective in situ result
- ▽ nonselective in situ result
- ▲ specific in literature
- × nonspecific in literature

GENE		
	IN SITU	IN LIT
SPECIFIC	●	▲
SELECTIVE	●	
NON-SELECTIVE	○	
NON-SPECIFIC		△

ISH results in relation to profile similarity to SM-MHC (Y-axis) and to expression level (X-axis). Nelander, S., P. Mostad, and P. Lindahl, Prediction of cell type-specific gene modules: identification and initial characterization of a core set of smooth muscle-specific genes. Genome Res, 2003. 13(8): p. 1838-54.

# THE GLYPHS ARE CUES



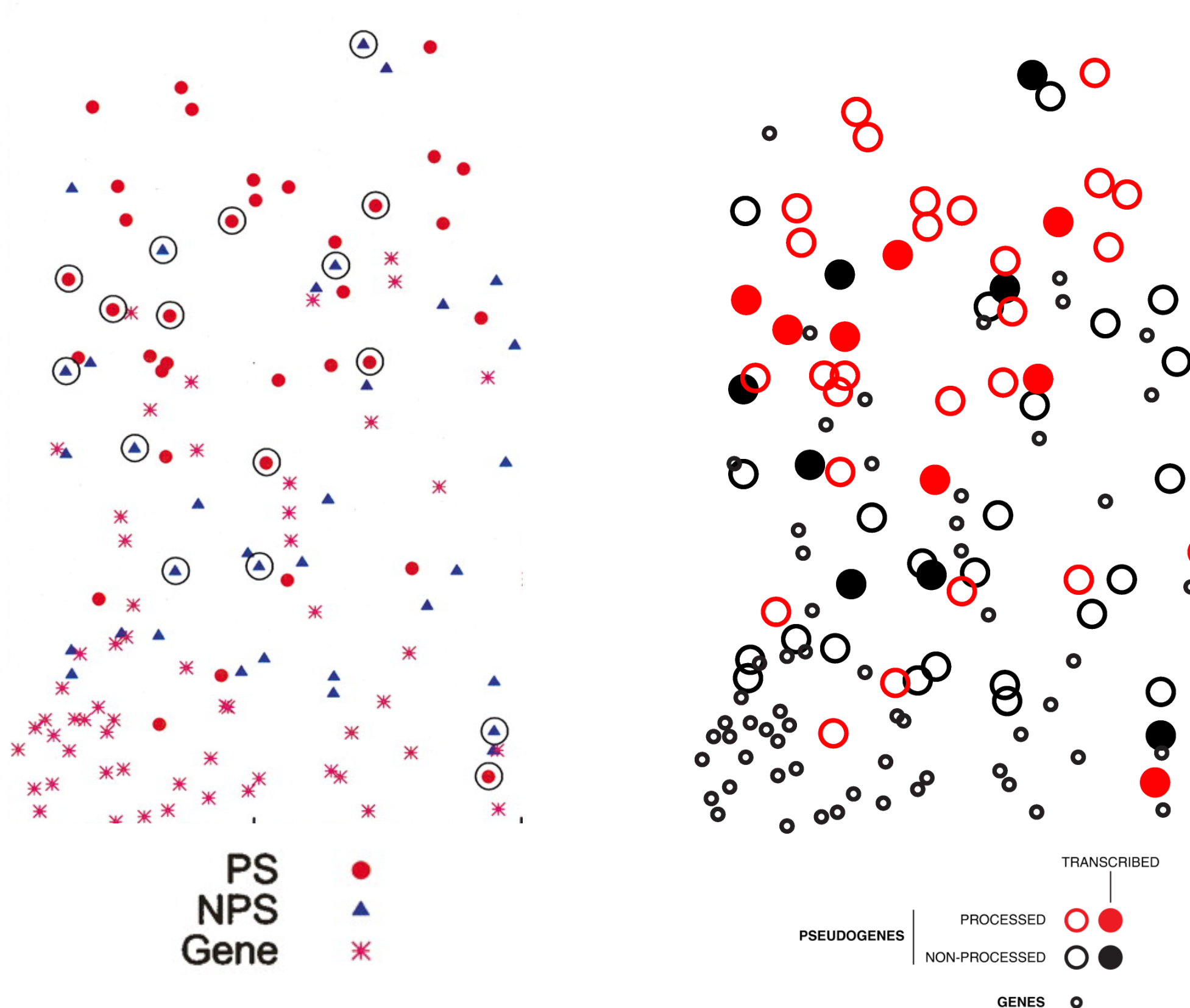
Comparison of Ka/Ks ratio and SNP density for genes and pseudogenes. Zheng, D., et al., Pseudogenes in the ENCODE regions: consensus annotation, analysis of transcription, and evolution. Genome Res, 2007. 17(6): p. 839-51.



# HIERARCHICAL GLYPHS

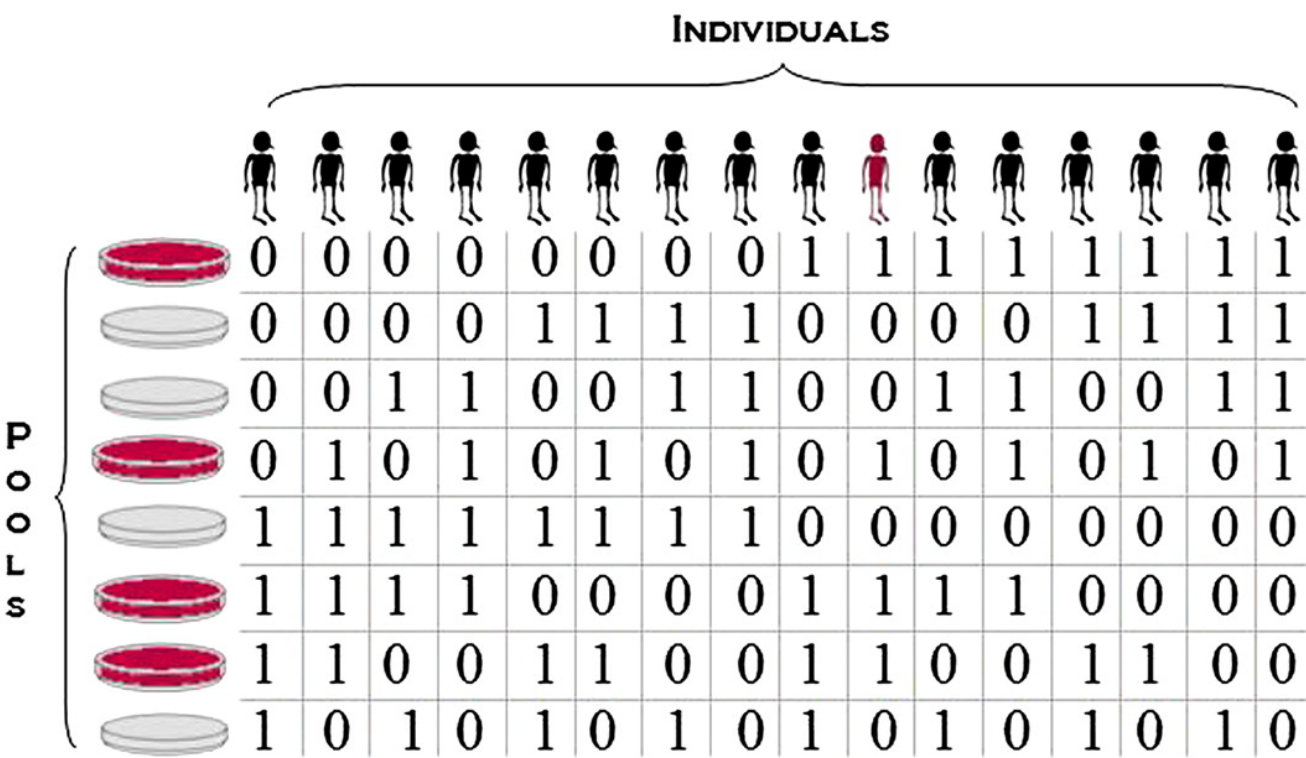
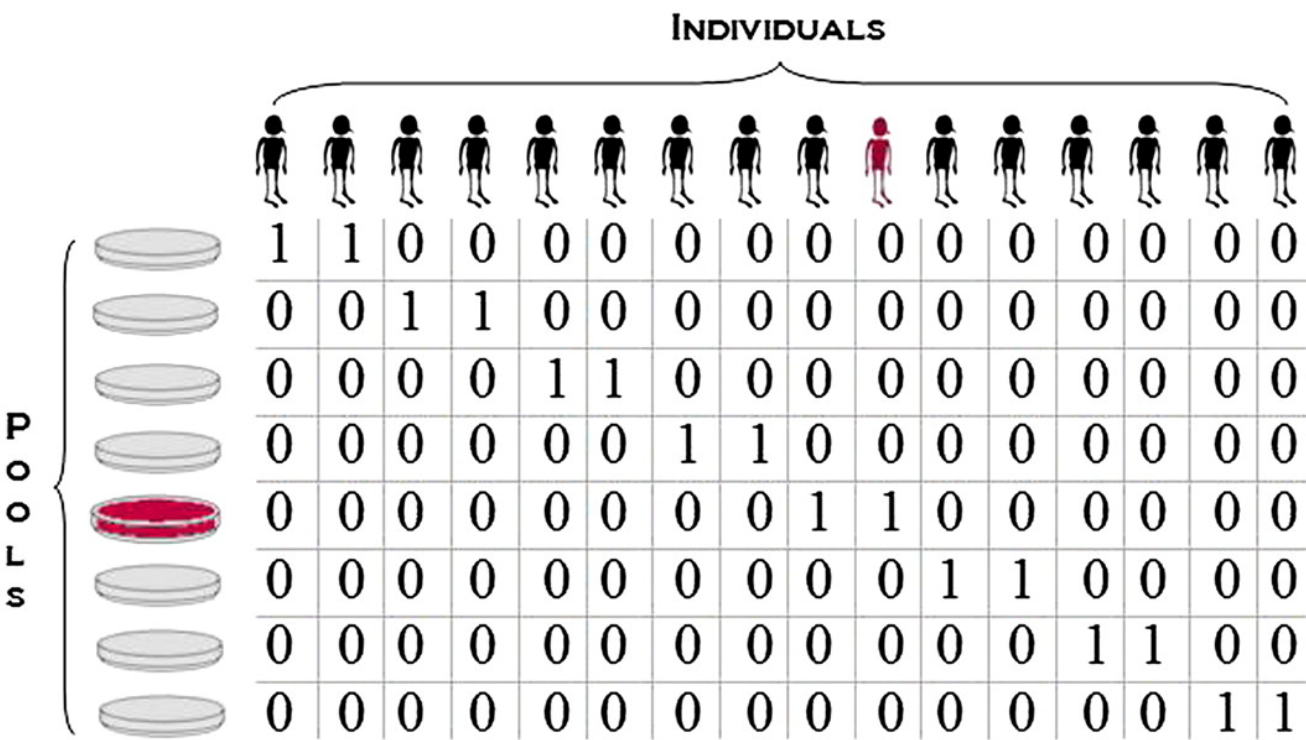


# HIERARCHICAL GLYPHS



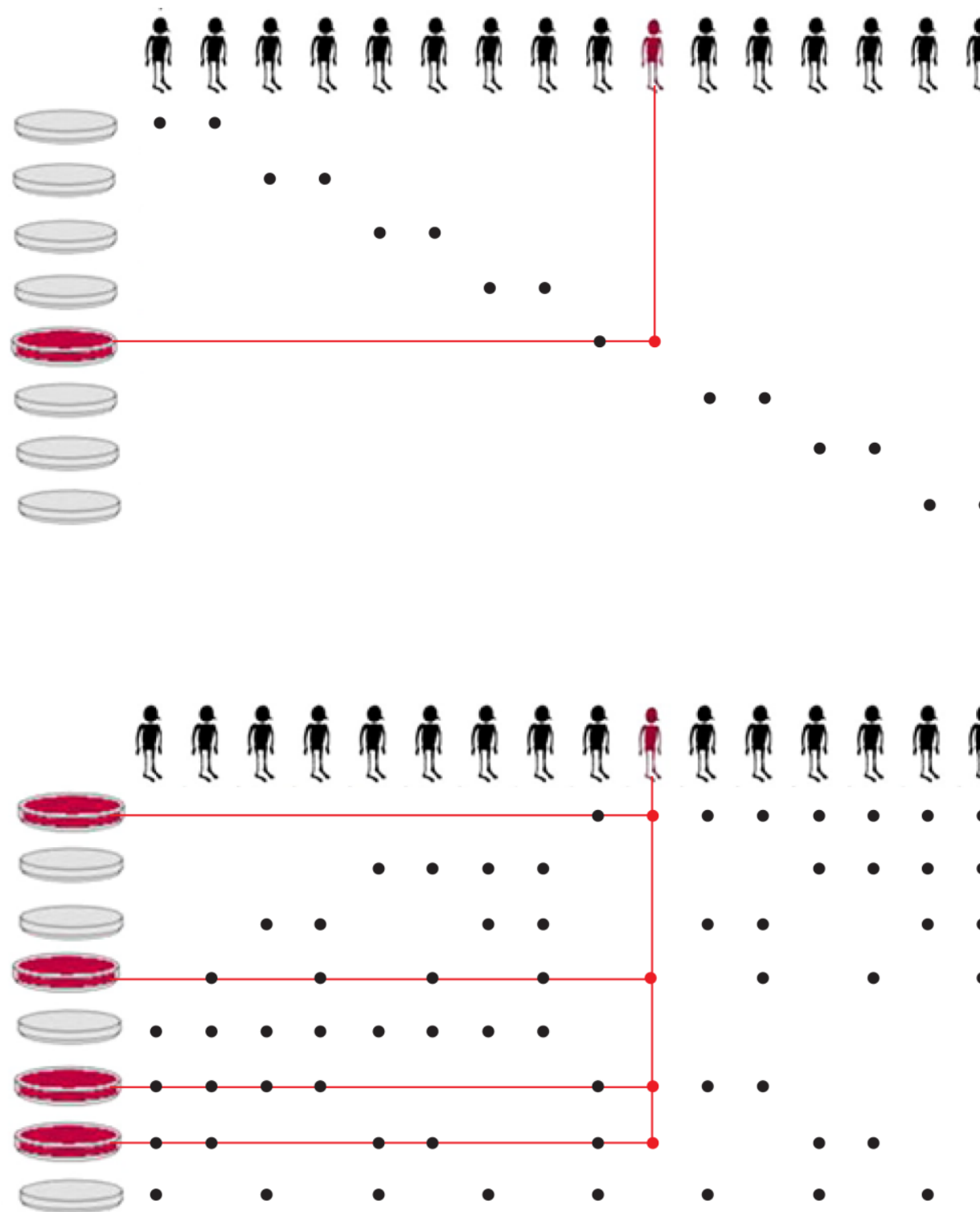
Panel (A) from Comparison of Ka/Ks ratio and SNP density for genes and pseudogenes. Zheng, D., et al., Pseudogenes in the ENCODE regions: consensus annotation, analysis of transcription, and evolution. *Genome Res*, 2007. 17(6): p. 839-51.

# EXCESS INK COMPETES WITH MESSAGE



Resequencing with naïve and log pool designs. Prabhu, S. and I. Pe'er, Overlapping pools for high-throughput targeted resequencing. Genome Res, 2009. 19 (7): p. 1254-61.

# REMOVING EXCESS INK

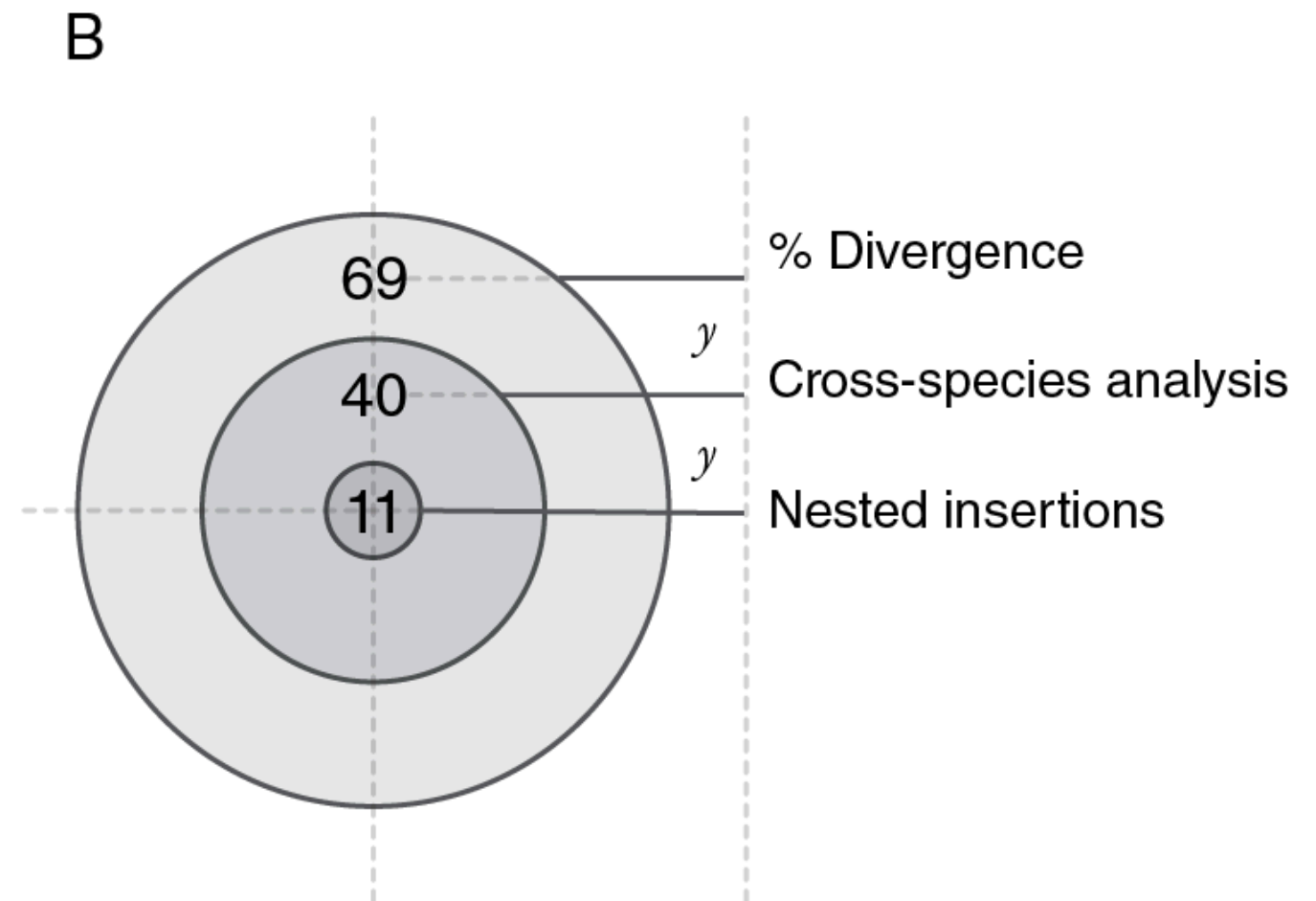
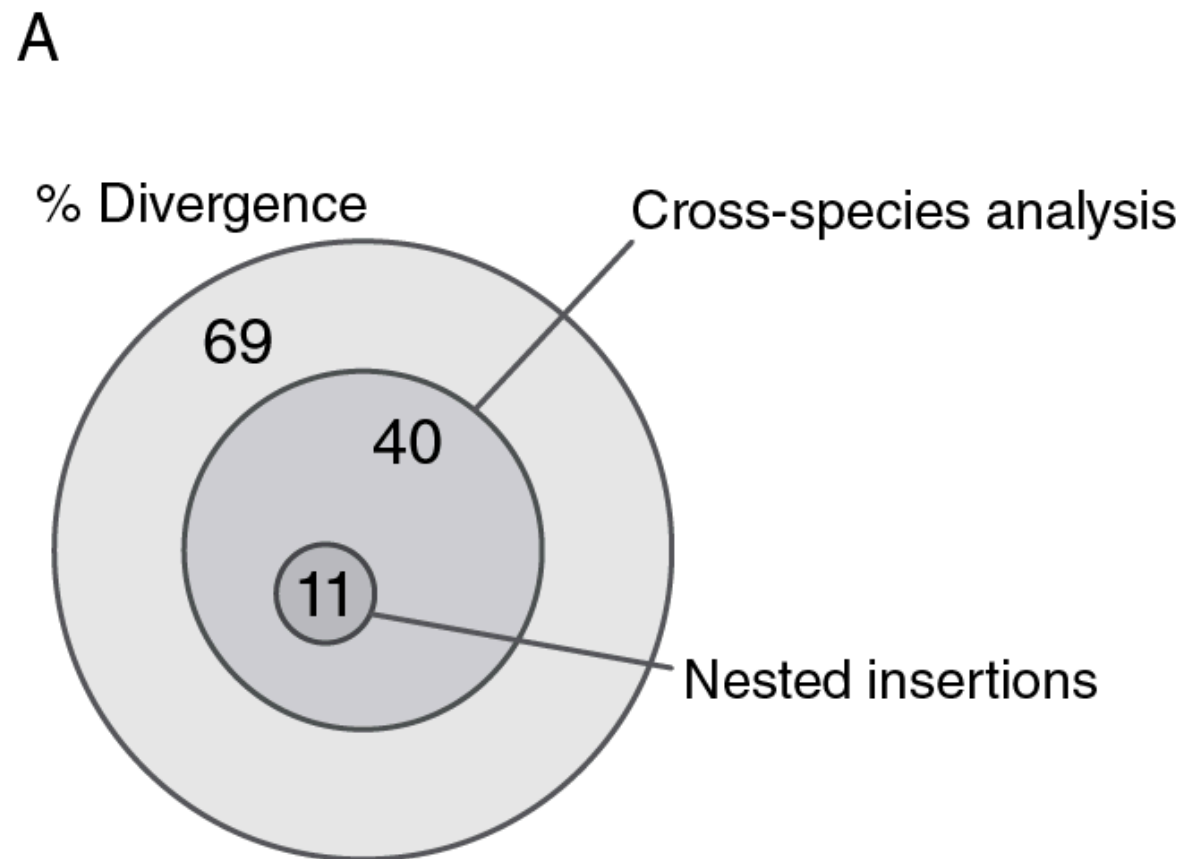




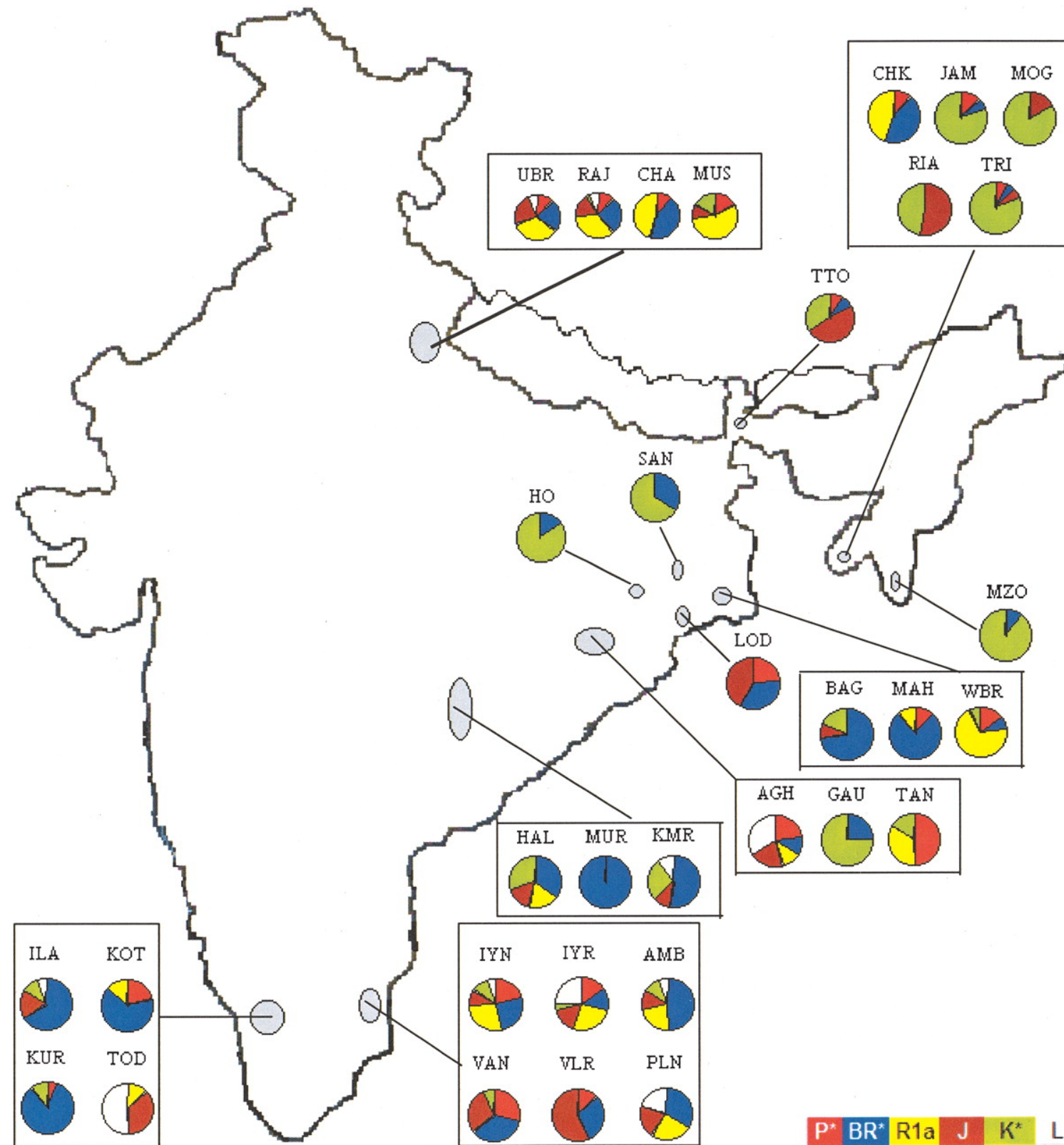
# legibilityclarity**attractiveness**

## GRID LAYOUT

# TITLE



# GRID LAYOUT



Frequencies (%) of Y-chromosomal haplogroups among ethnic populations. Basu, A., et al., Ethnic India: a genomic view, with special reference to peopling and structure. *Genome Res*, 2003. 13(10): p. 2277-90.

# GRID LAYOUT



R = G = B = 210 215 220 225 230 235 240 245 250

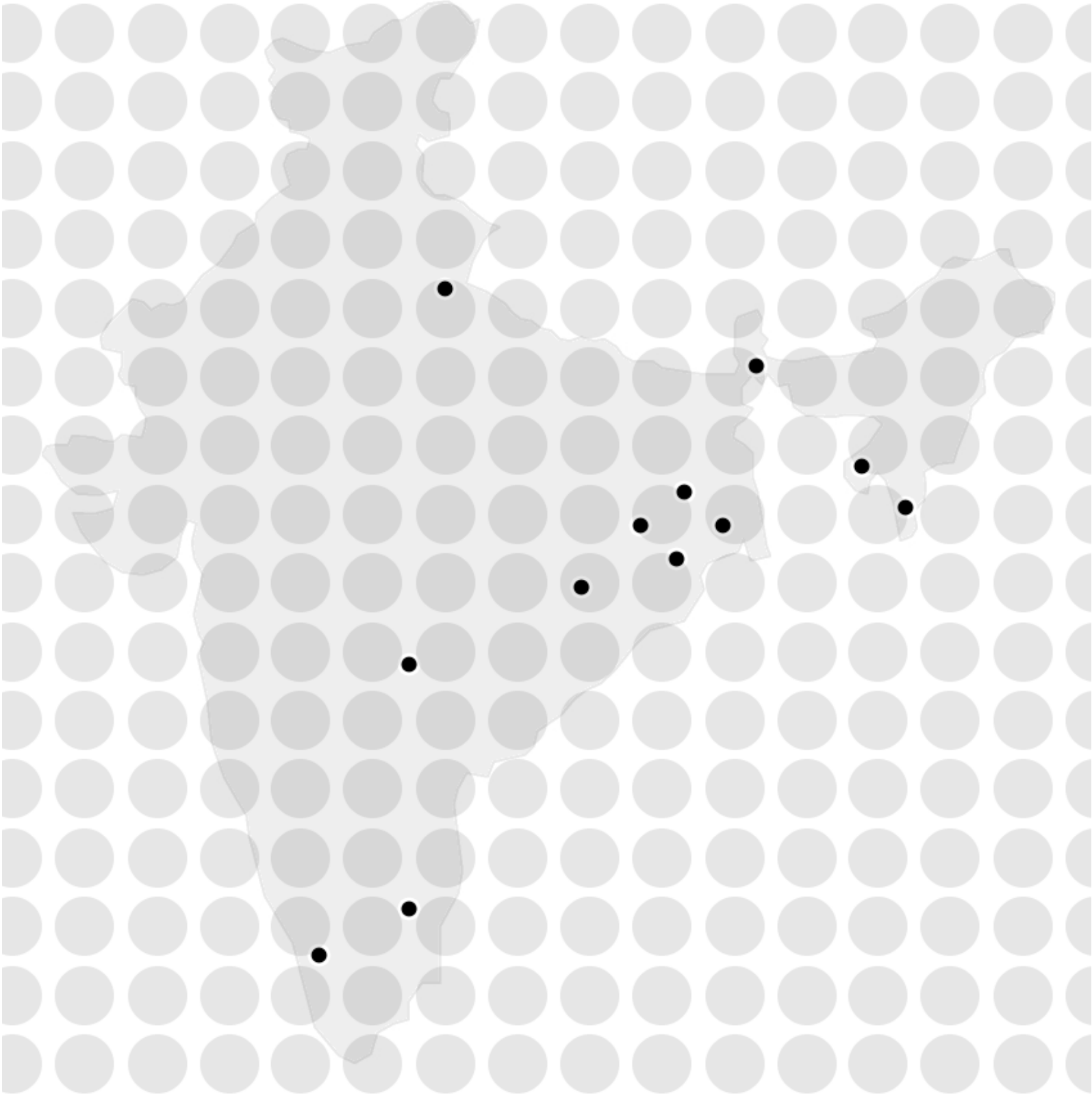
Be careful when using very light colors. They frequently do not render well on LCD projectors and appear washed out.



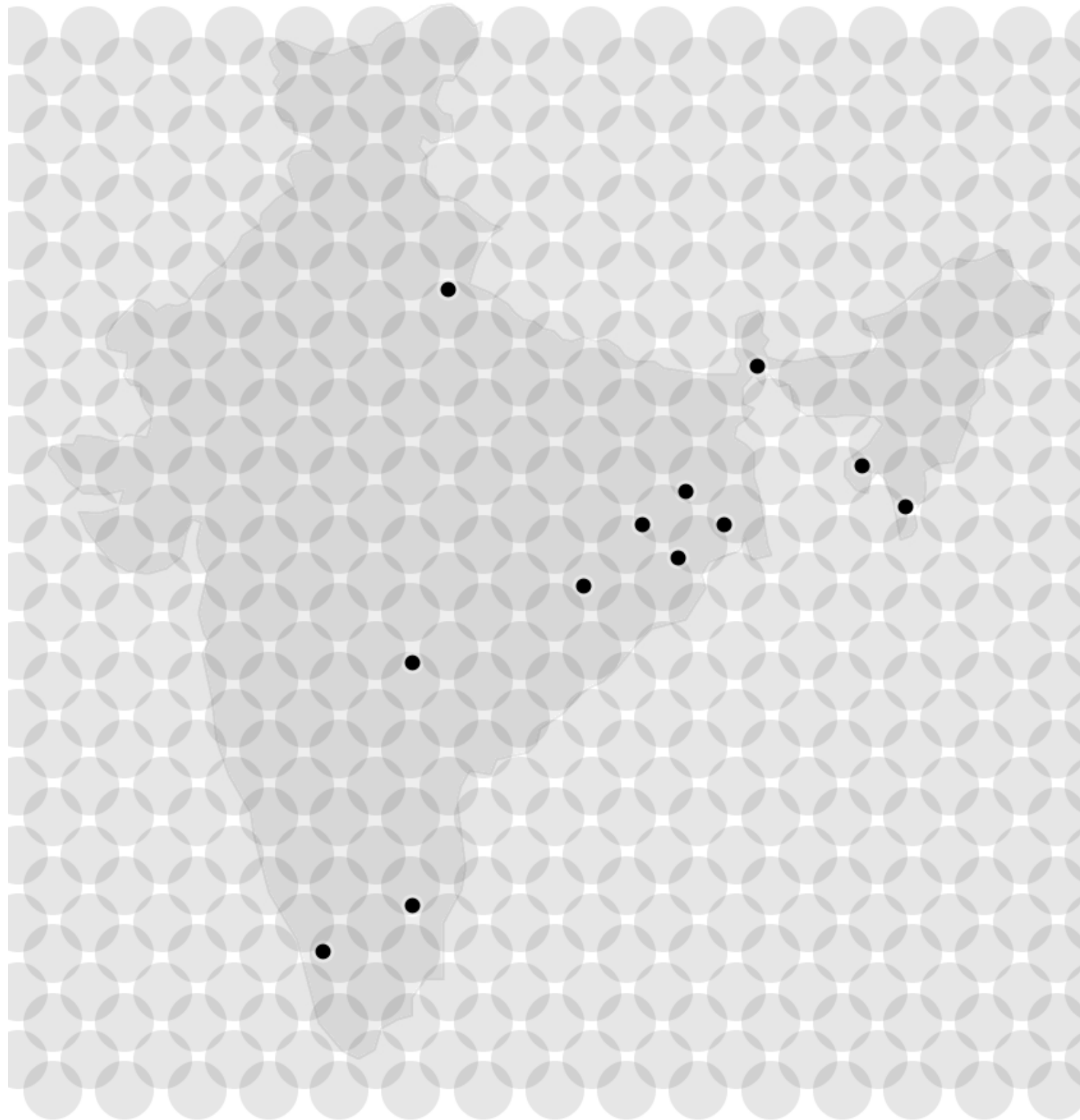
# GRID LAYOUT



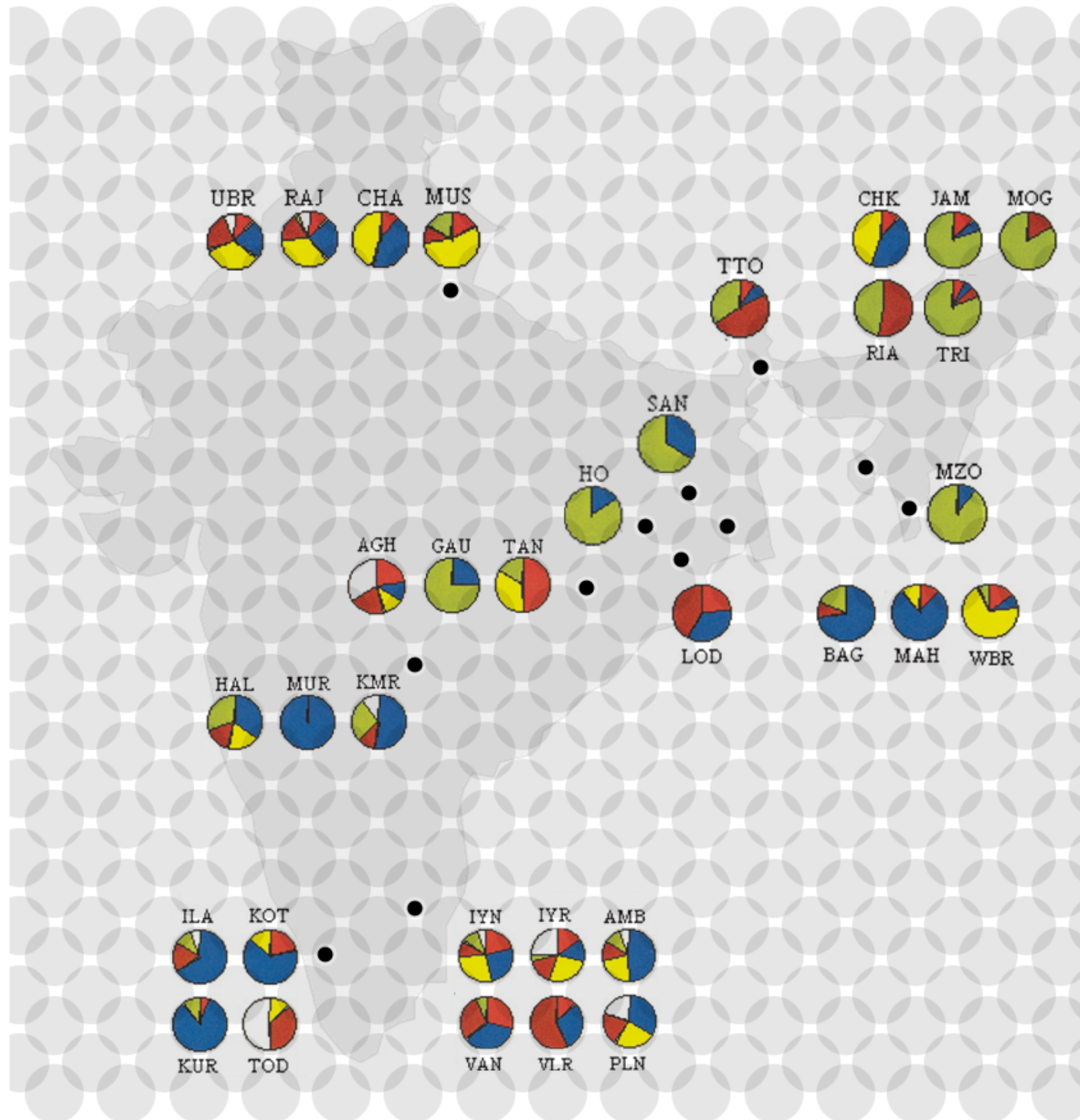
# GRID LAYOUT



# GRID LAYOUT

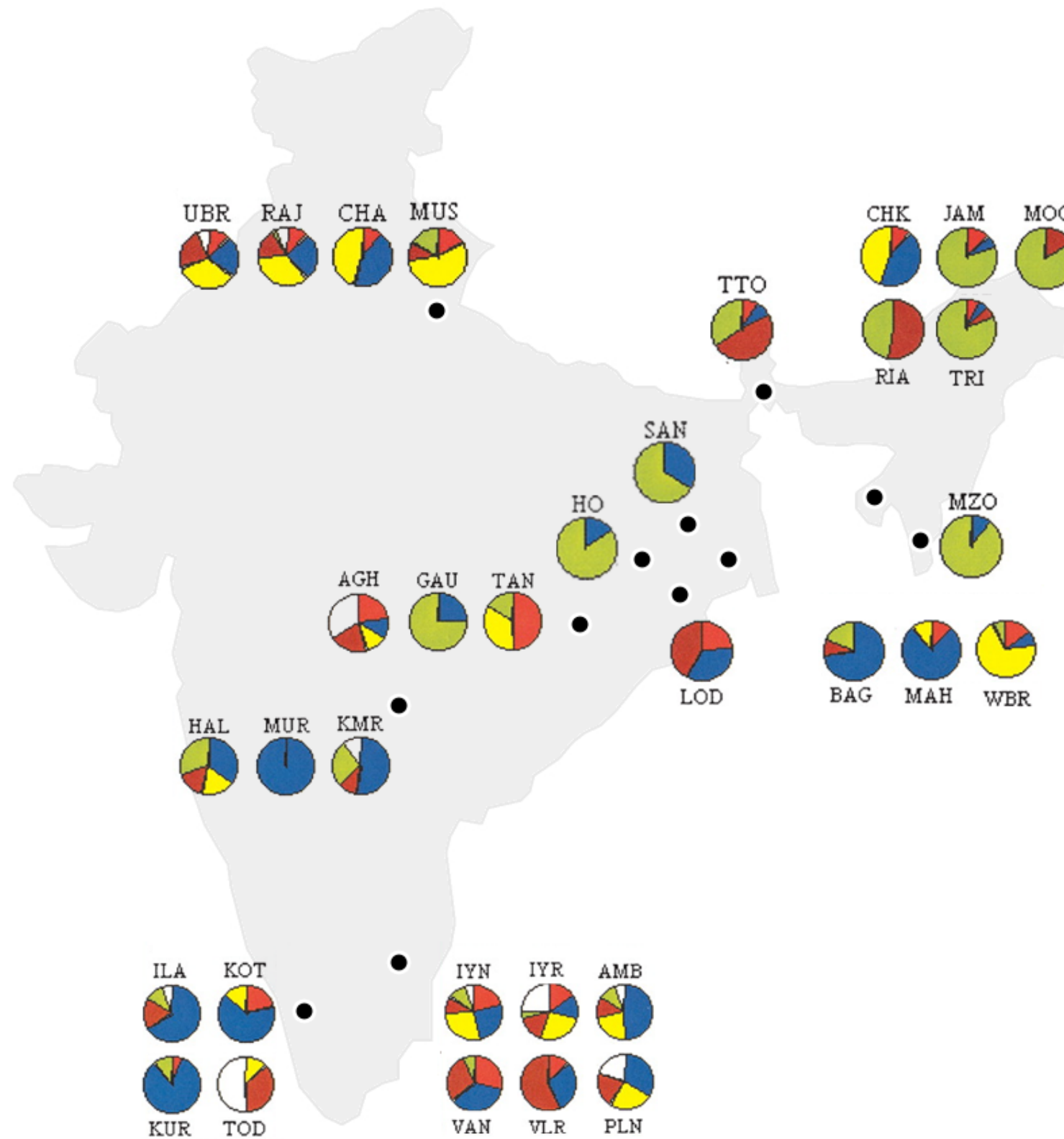


# GRID LAYOUT

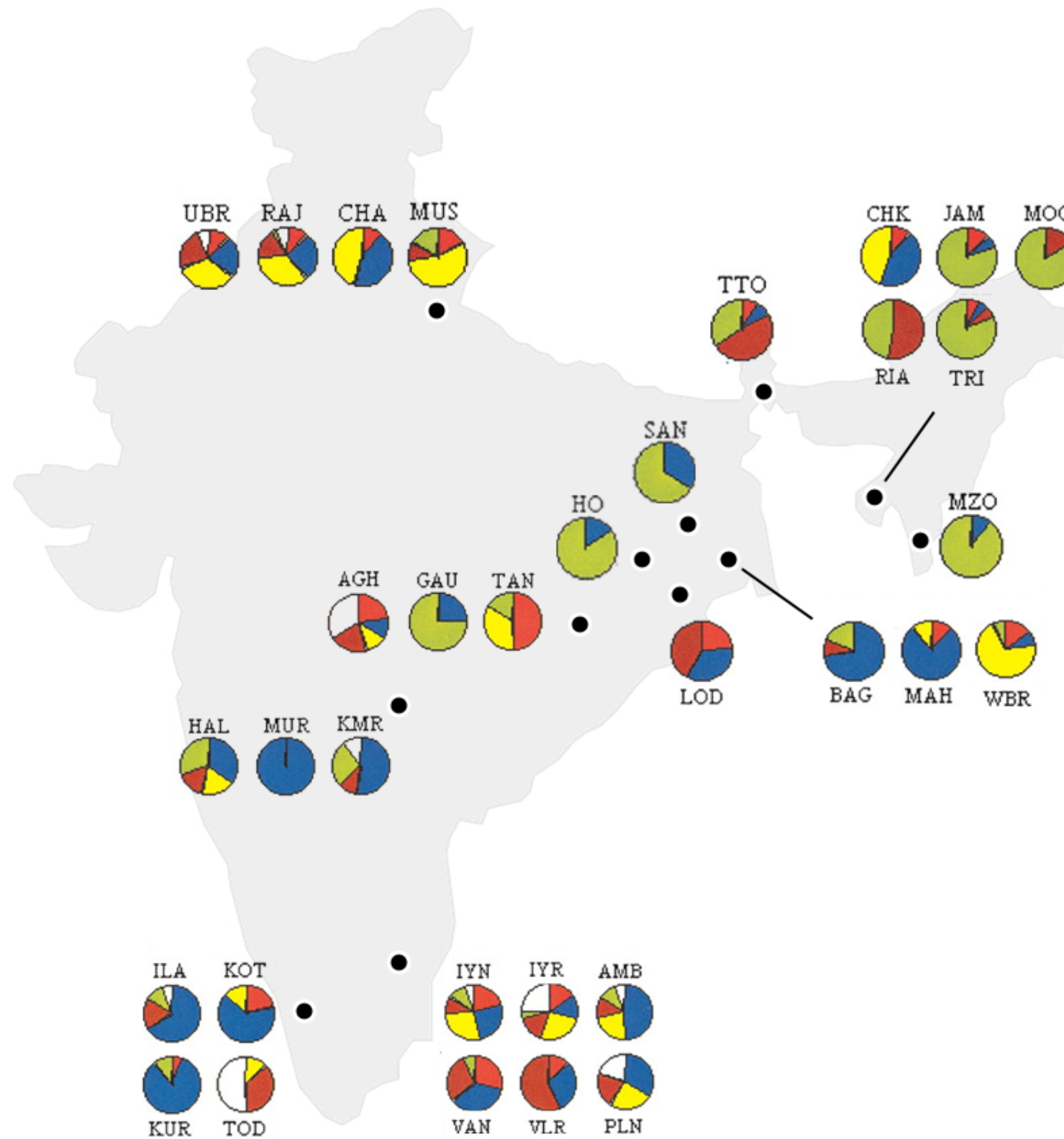




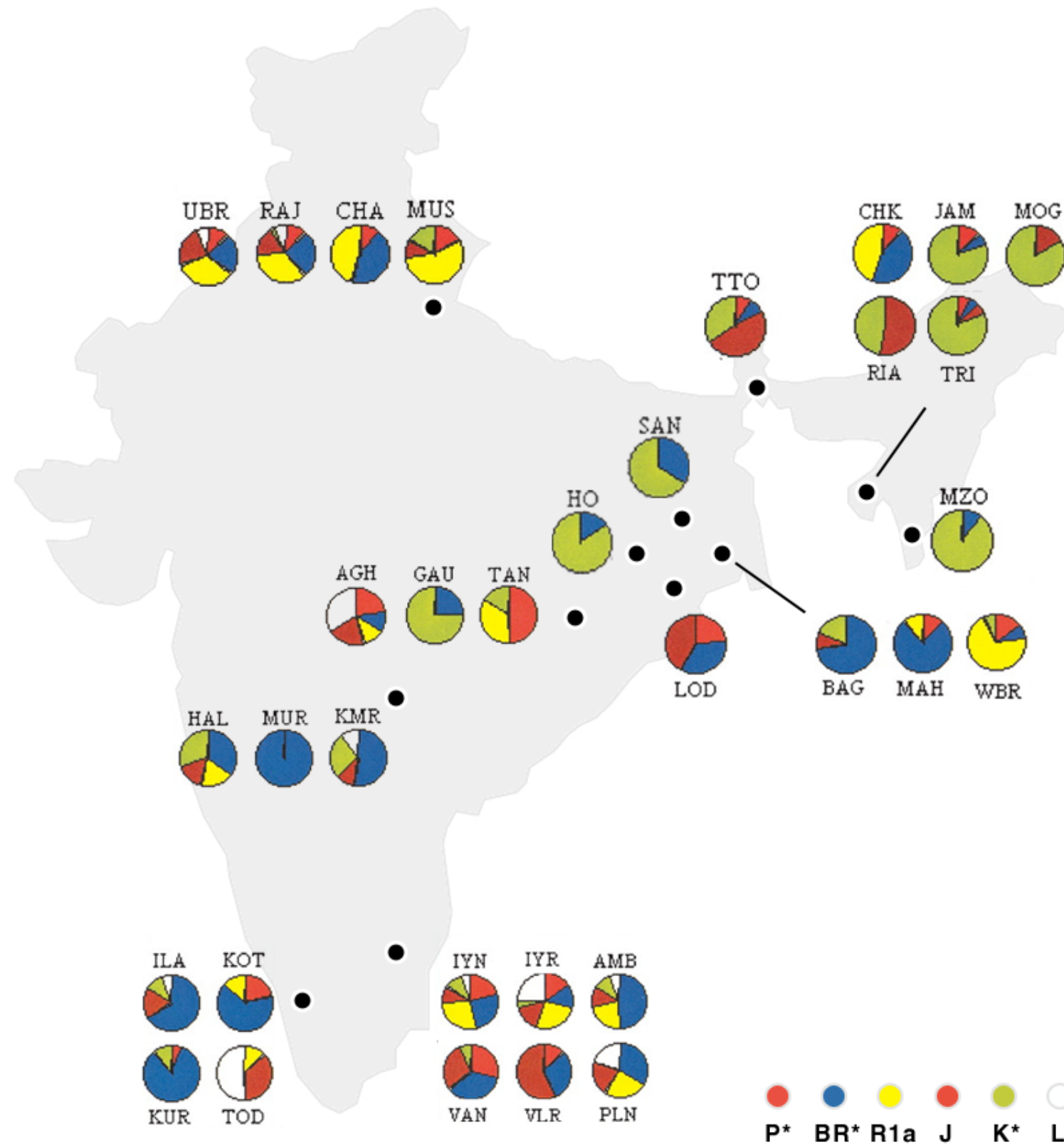
# GRID LAYOUT



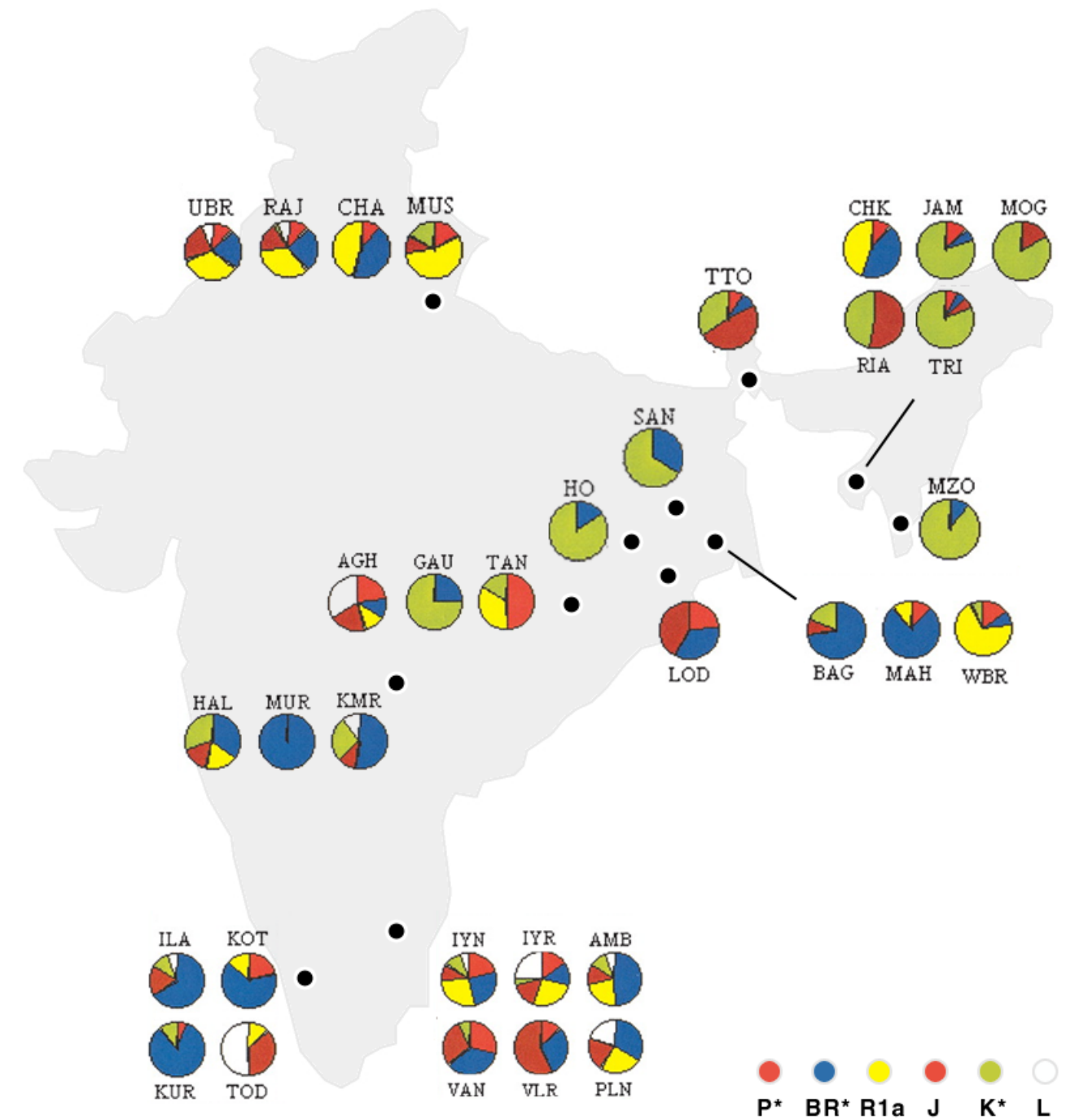
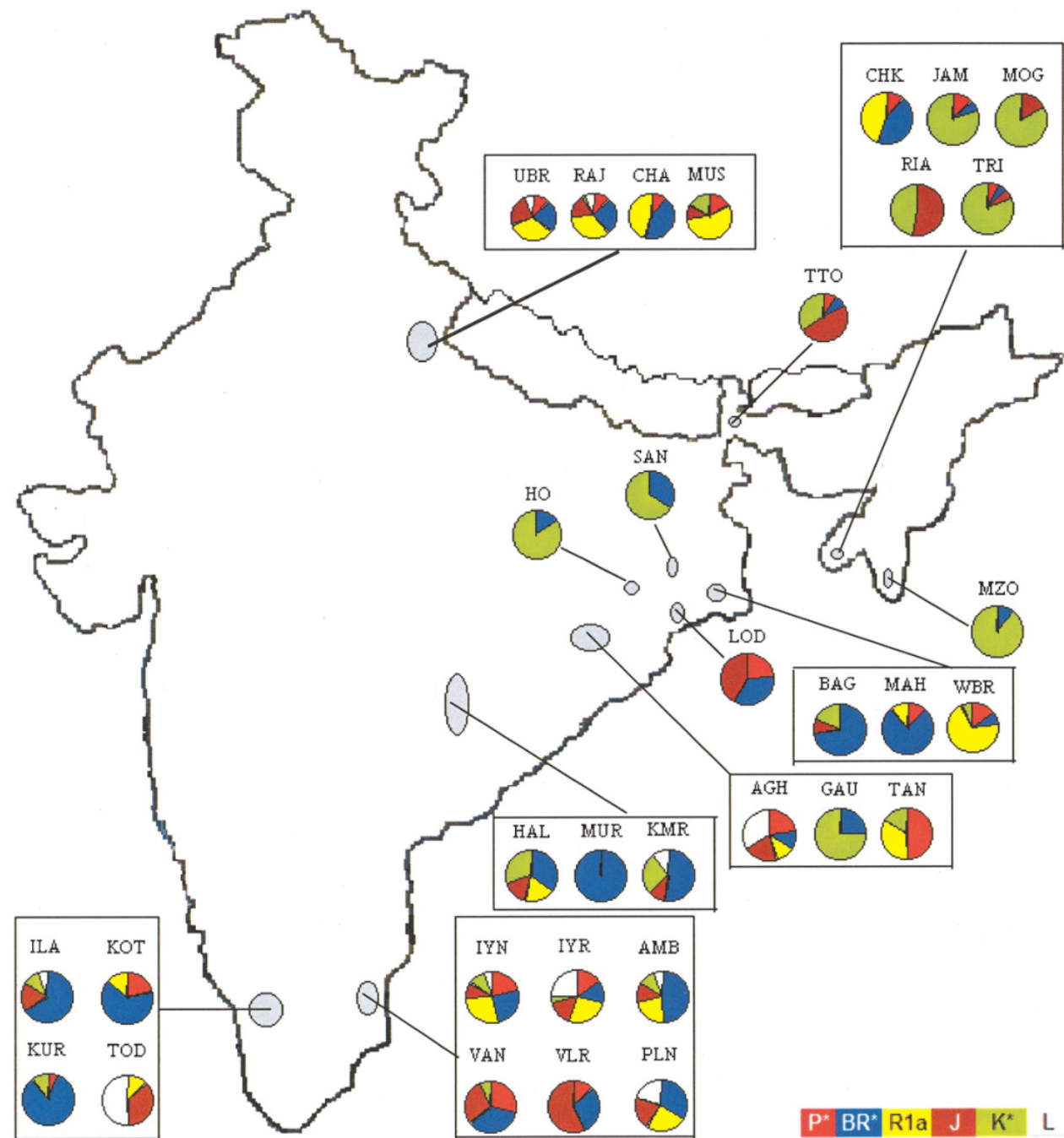
# GRID LAYOUT



# GRID LAYOUT



# GRID LAYOUT



Left panel from Frequencies (%) of Y-chromosomal haplogroups among ethnic populations. Basu, A., et al., Ethnic India: a genomic view, with special reference to peopling and structure. Genome Res, 2003. 13(10): p. 2277-90.



# REQUIRED READING

