

Equating OTUs with Species Diversity

Micah Dunthorn



Natural History
Museum



OTUS

OTUs $\stackrel{?}{=}$ species

clustering as a step

clustering

= grouping of similar sequences

clustering

sampling,
storage



nucleotide
extraction



amplification,
sequencing

clustering

Opinion

Replicate or lie

James I. Prosser*

*Institute of Biological and Environmental Sciences,
University of Aberdeen, Cruickshank Building, St.
Machar Drive, Aberdeen, AB24 3UU, UK.*

2017 8:1188



ARTICLE

program: LULU

DOI: [10.1038/s41467-017-01312-x](https://doi.org/10.1038/s41467-017-01312-x)

OPEN

Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates

Tobias Guldberg Frøslev ^{1,2}, Rasmus Kjøller¹, Hans Henrik Bruun ¹, Rasmus Ejrnæs³, Ane Kirstine Brunbjerg³, Carlotta Pietroni² & Anders Johannes Hansen²

sampling,
storage



nucleotide
extraction



amplification,
sequencing

cleaning,
dereplication



clustering



more
cleaning

sampling,
storage



nucleotide
extraction



amplification,
sequencing

cleaning,
dereplication



clustering



more
cleaning

taxonomic assignments,
phylogenetic placements



α, β diversities,
co-occurrences

what is a species

1997 in "Species: The units of biodiversity"

A hierarchy of species concepts: the denouement in the saga of the species problem

R. L. Mayden

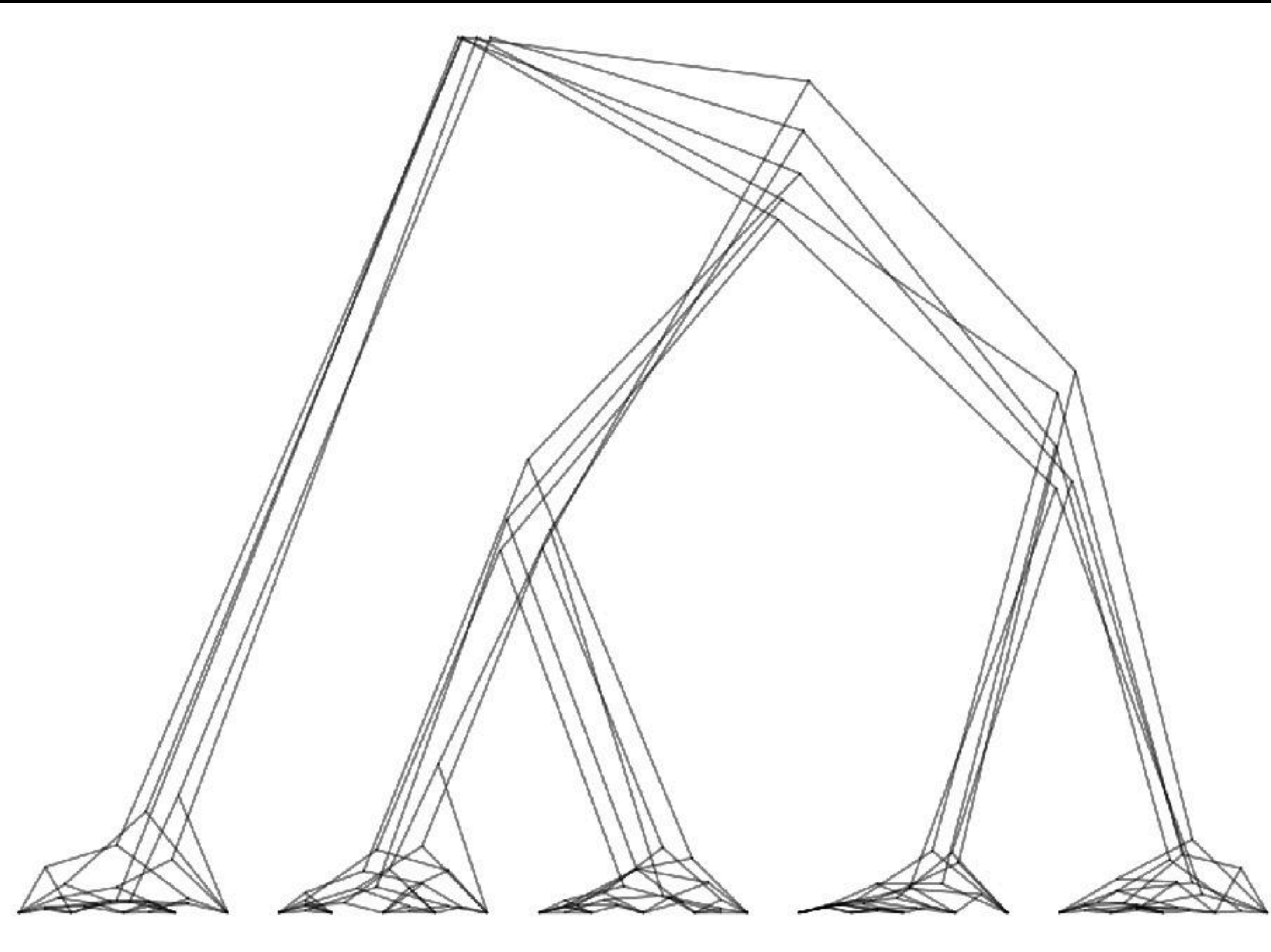
Contacting address: Department of Biological Sciences, P.O. Box 0344, University of Alabama, Tuscaloosa, AL 35487, USA

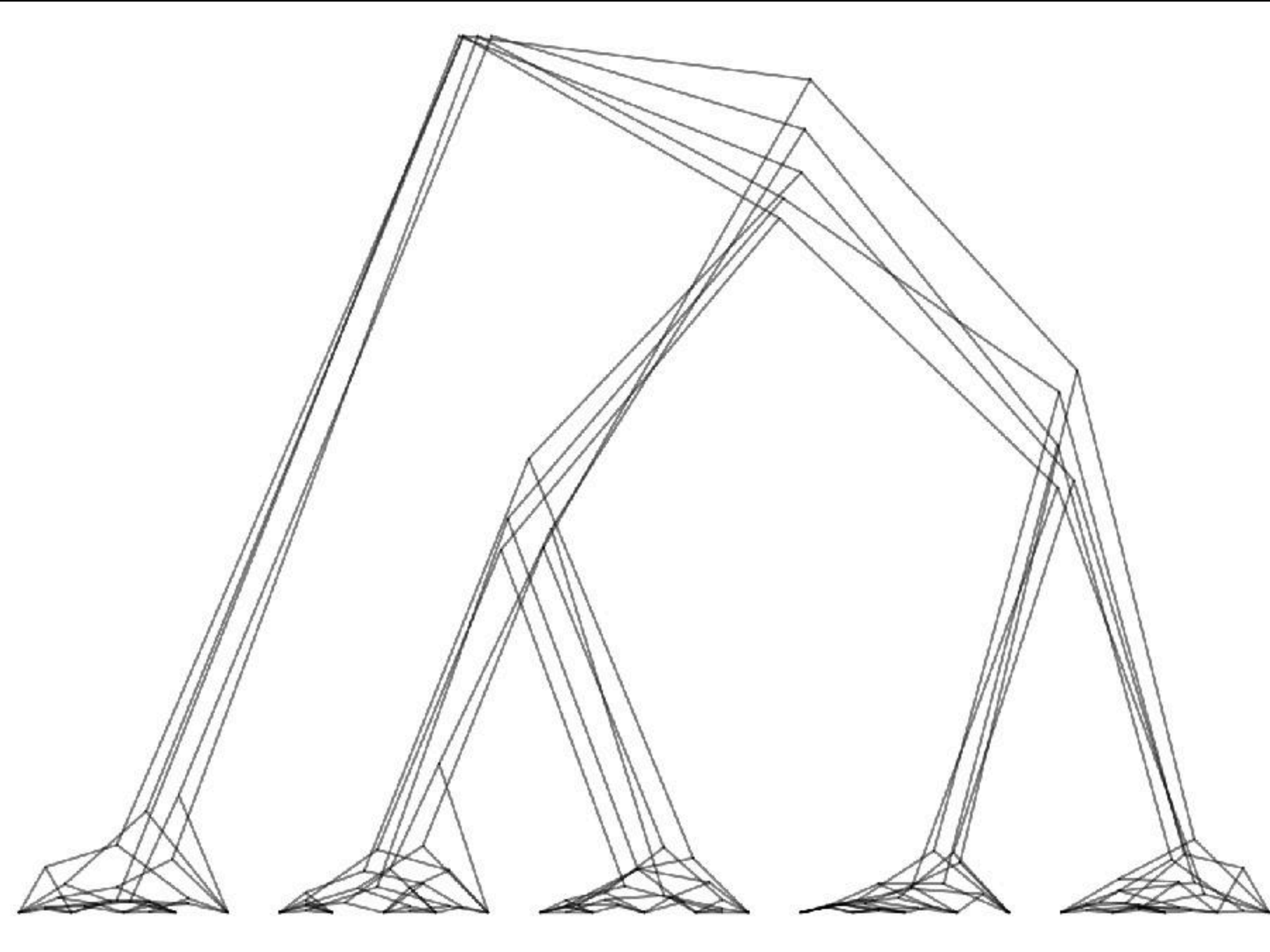
eukaryote

vs.

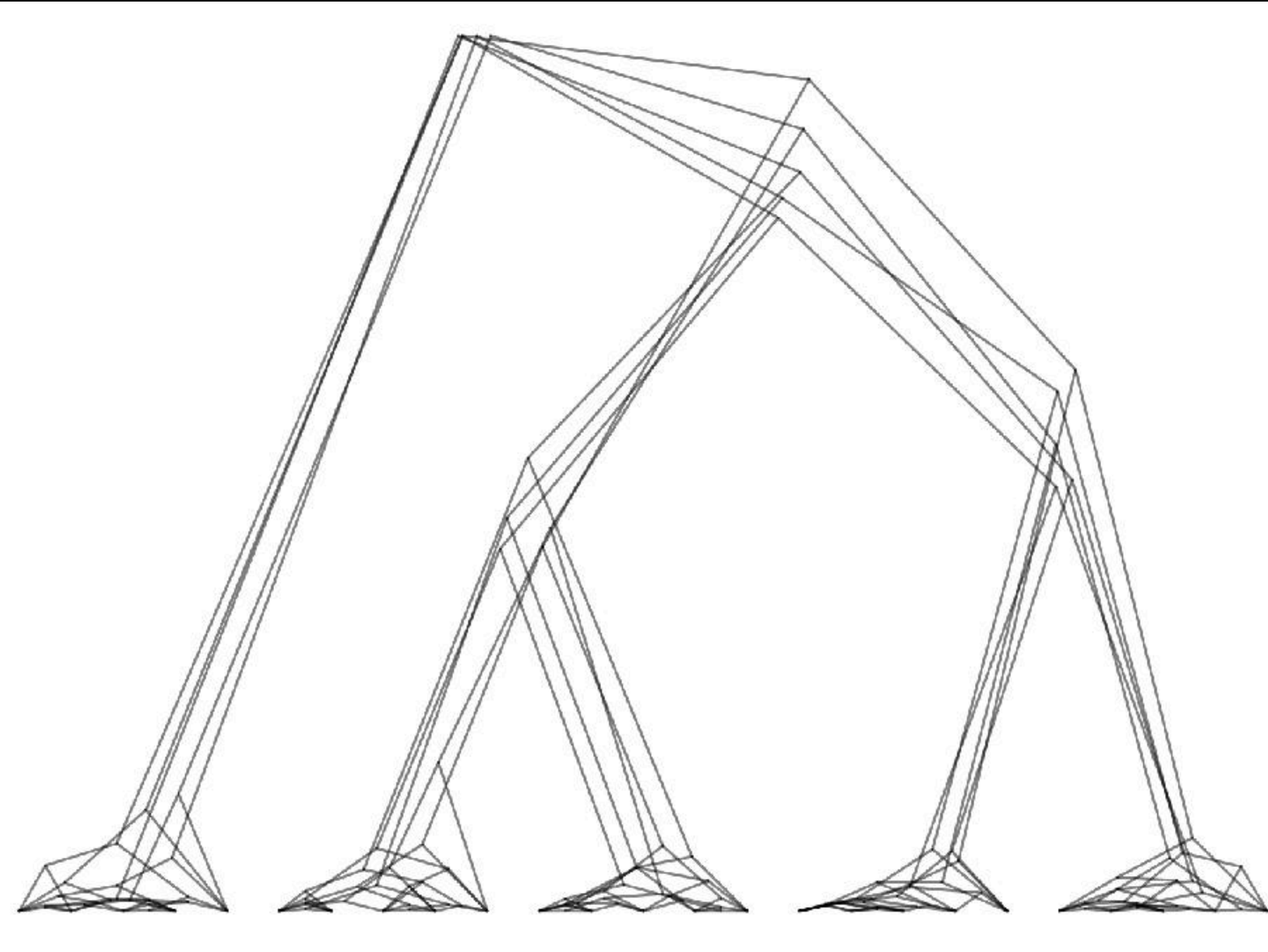
bacteria and archaea

sexual
vs.
asexual



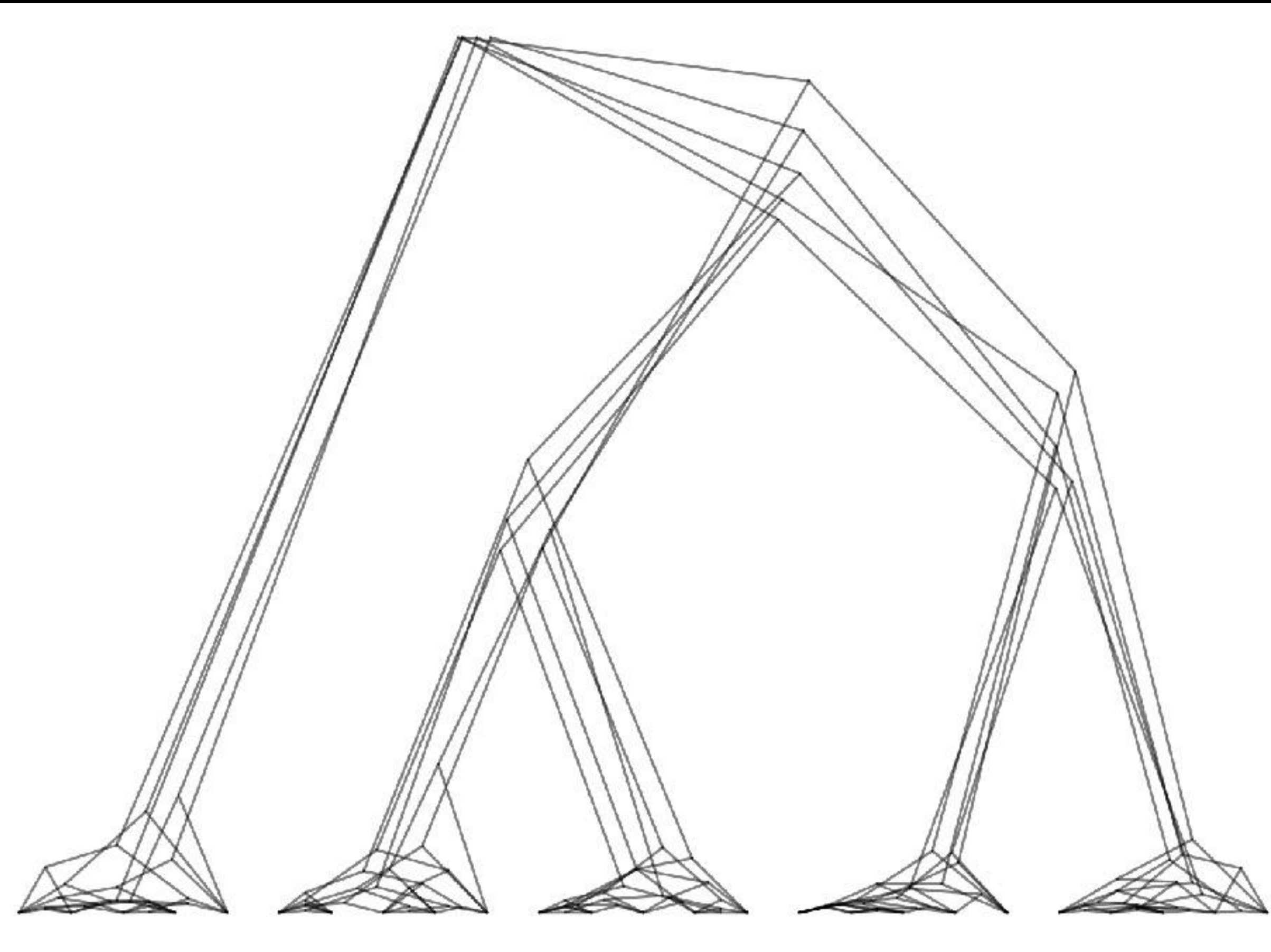


**macro-evolutionary
processes of phylogeny**

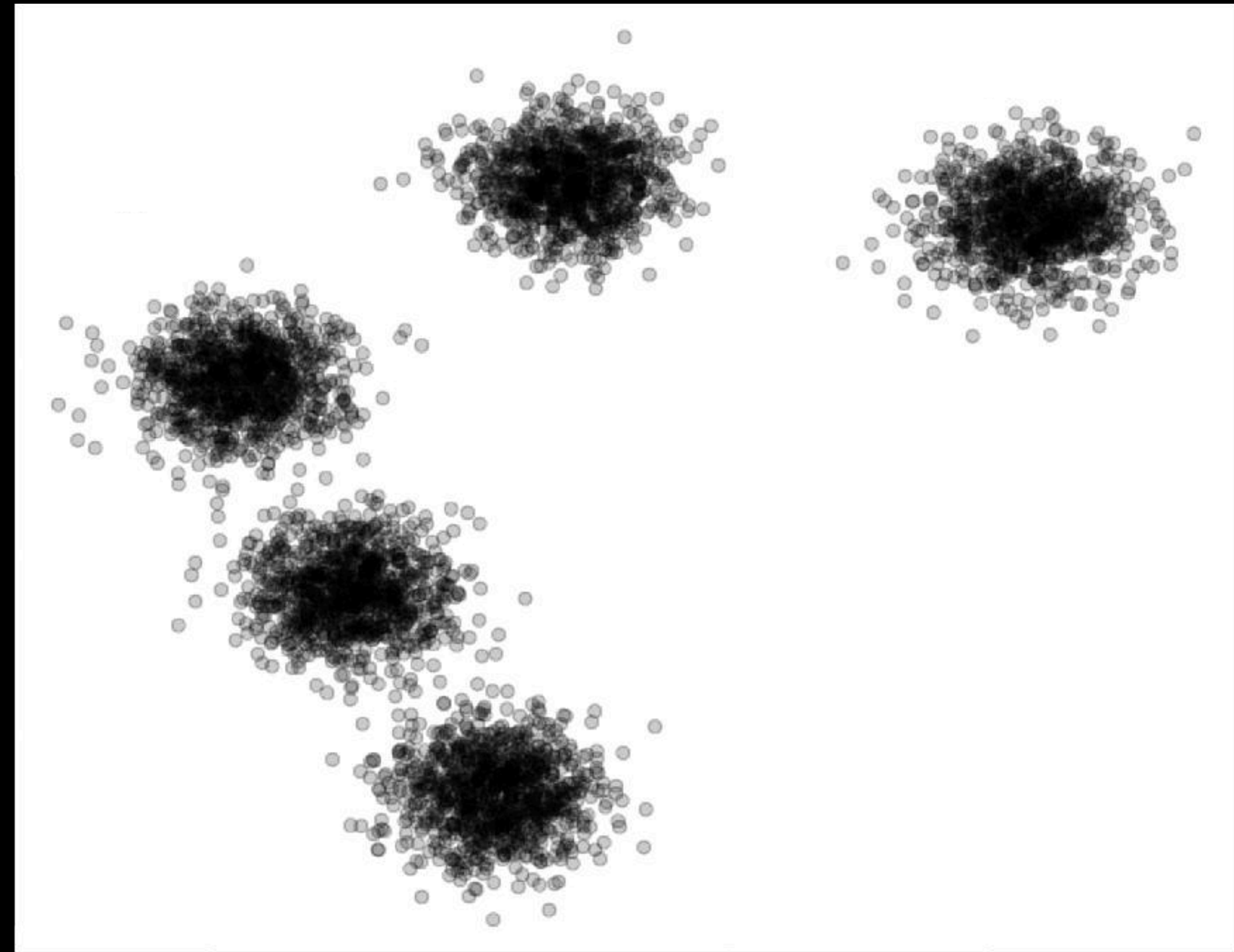


**macro-evolutionary
processes of phylogeny**

**micro-evolutionary
processes of population
genetics**



trait 2



trait 1

biological species

= actually or potentially inter-breeding
natural populations, which are
reproductively isolated

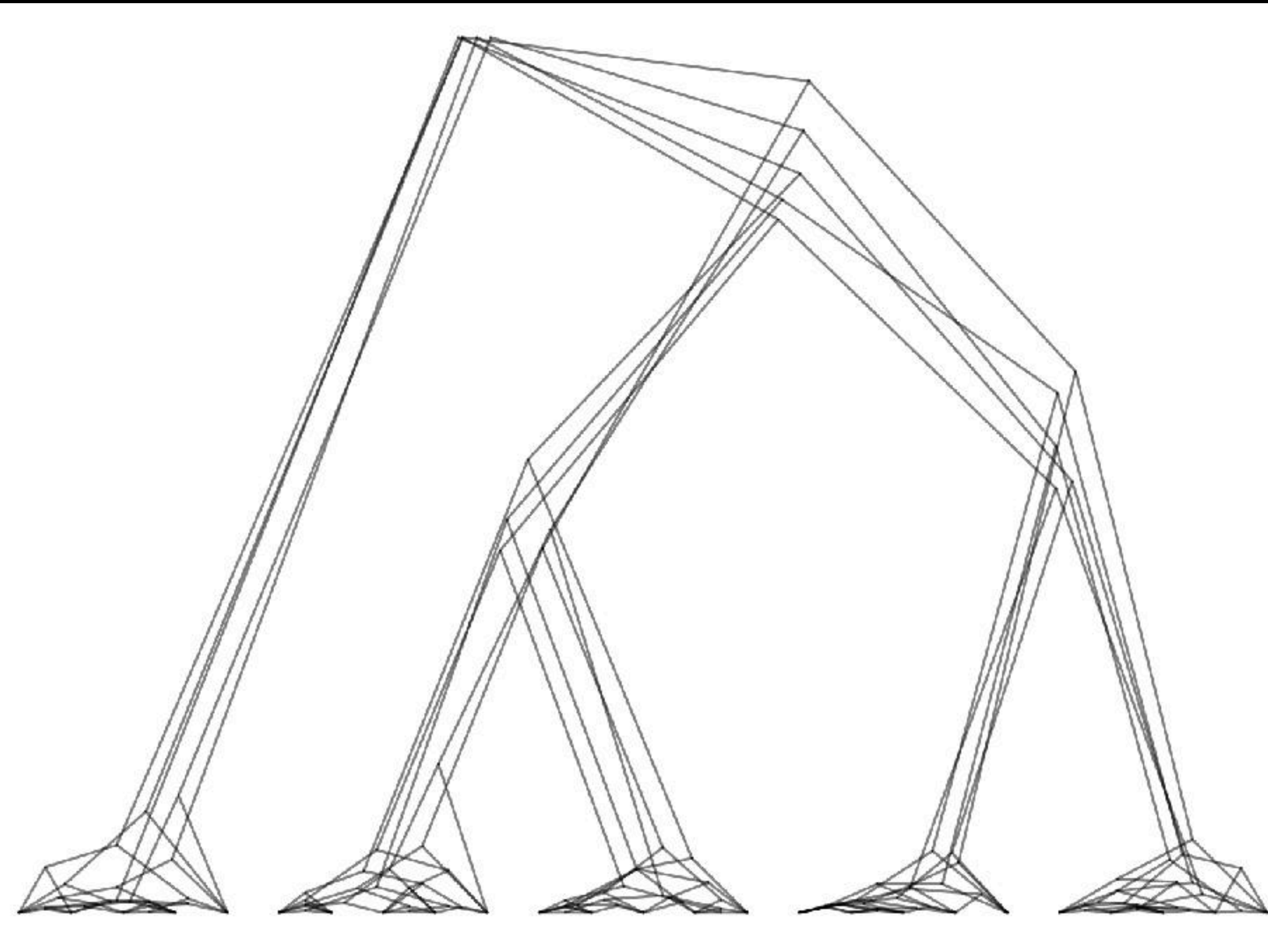
- Mayr 1942



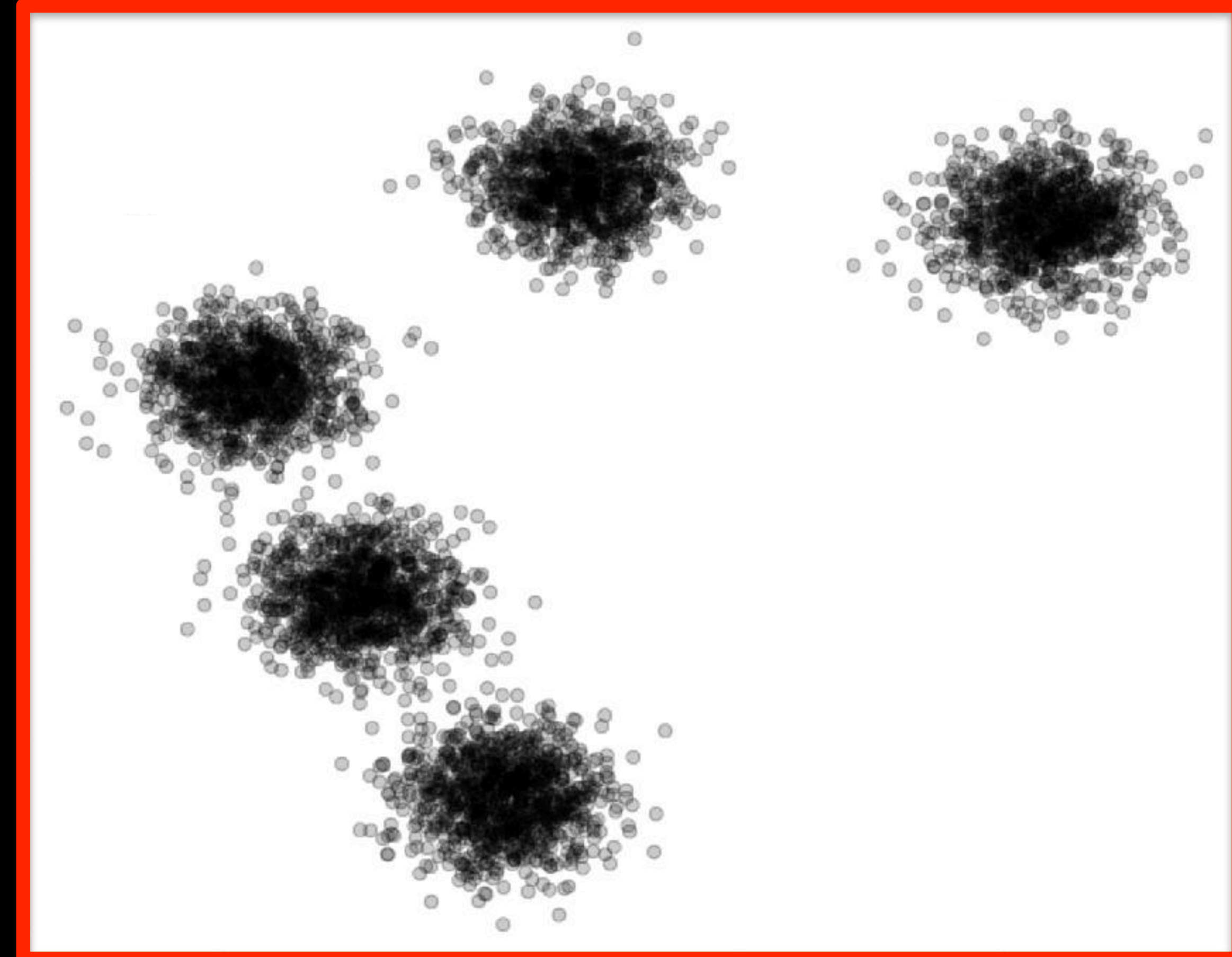
ciliate sex



Symbiodinium



trait 2



trait 1

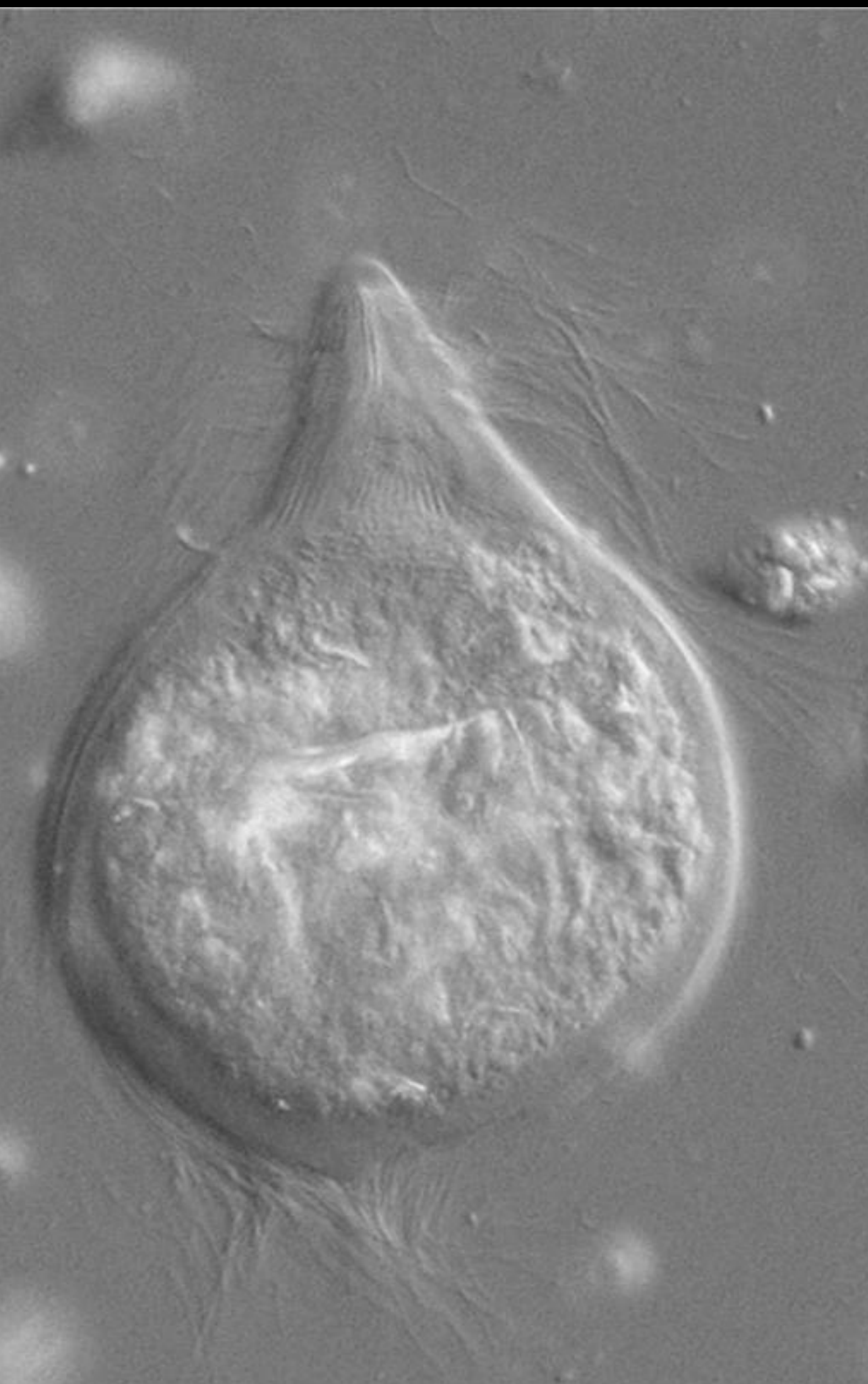
morphological species

= the smallest groups that are
consistently and persistently distinct

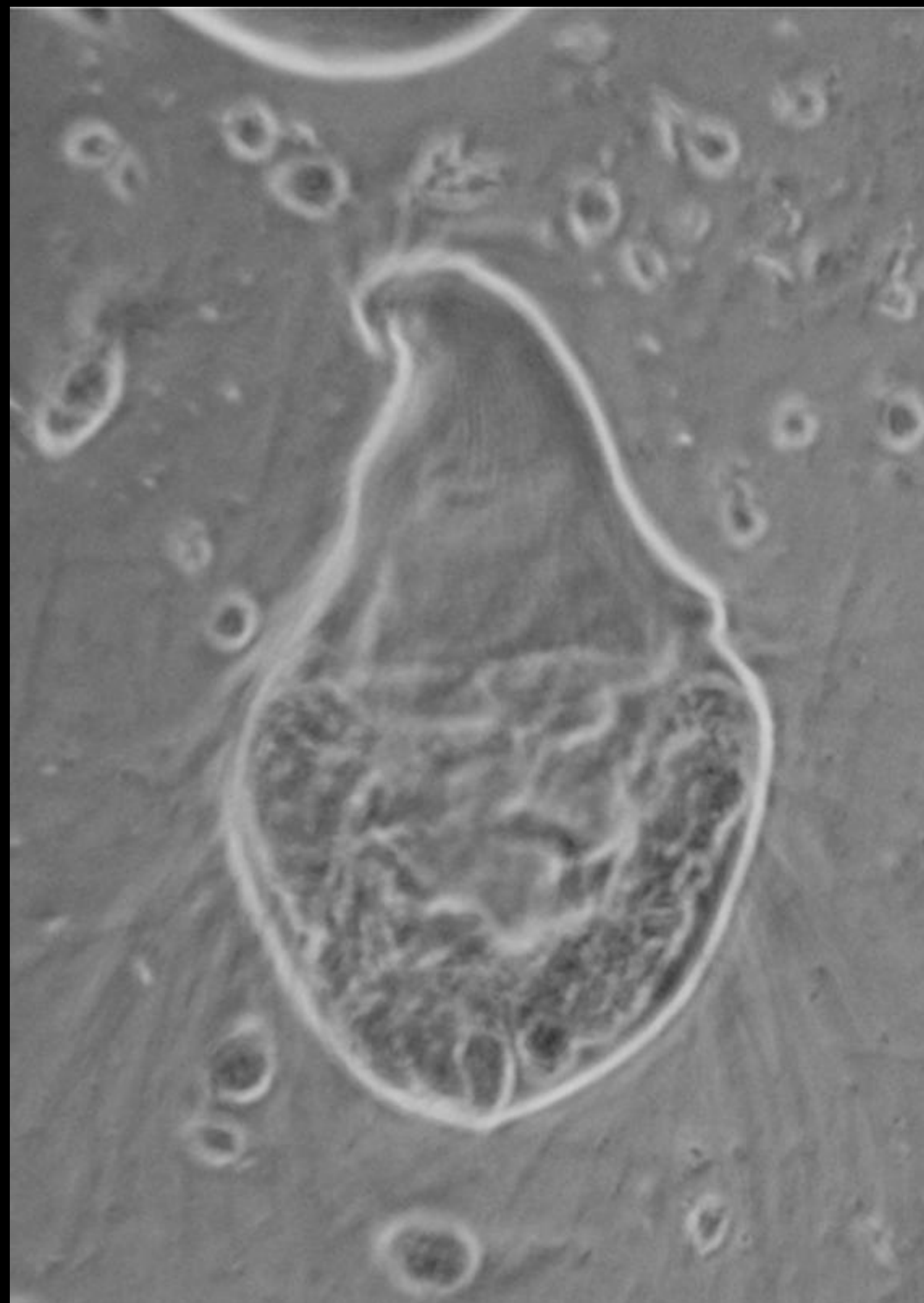
- Cronquist 1978



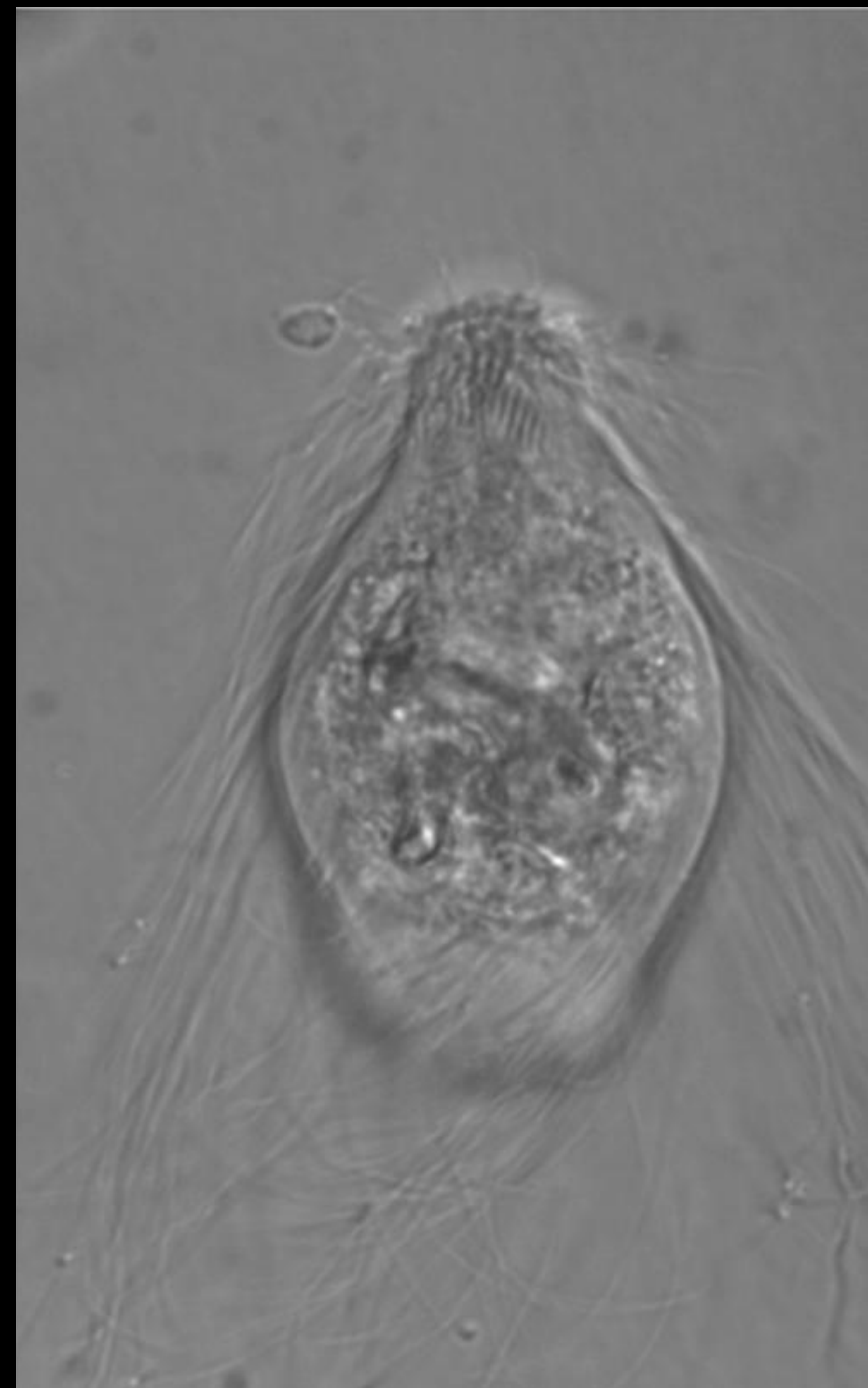
ciliates



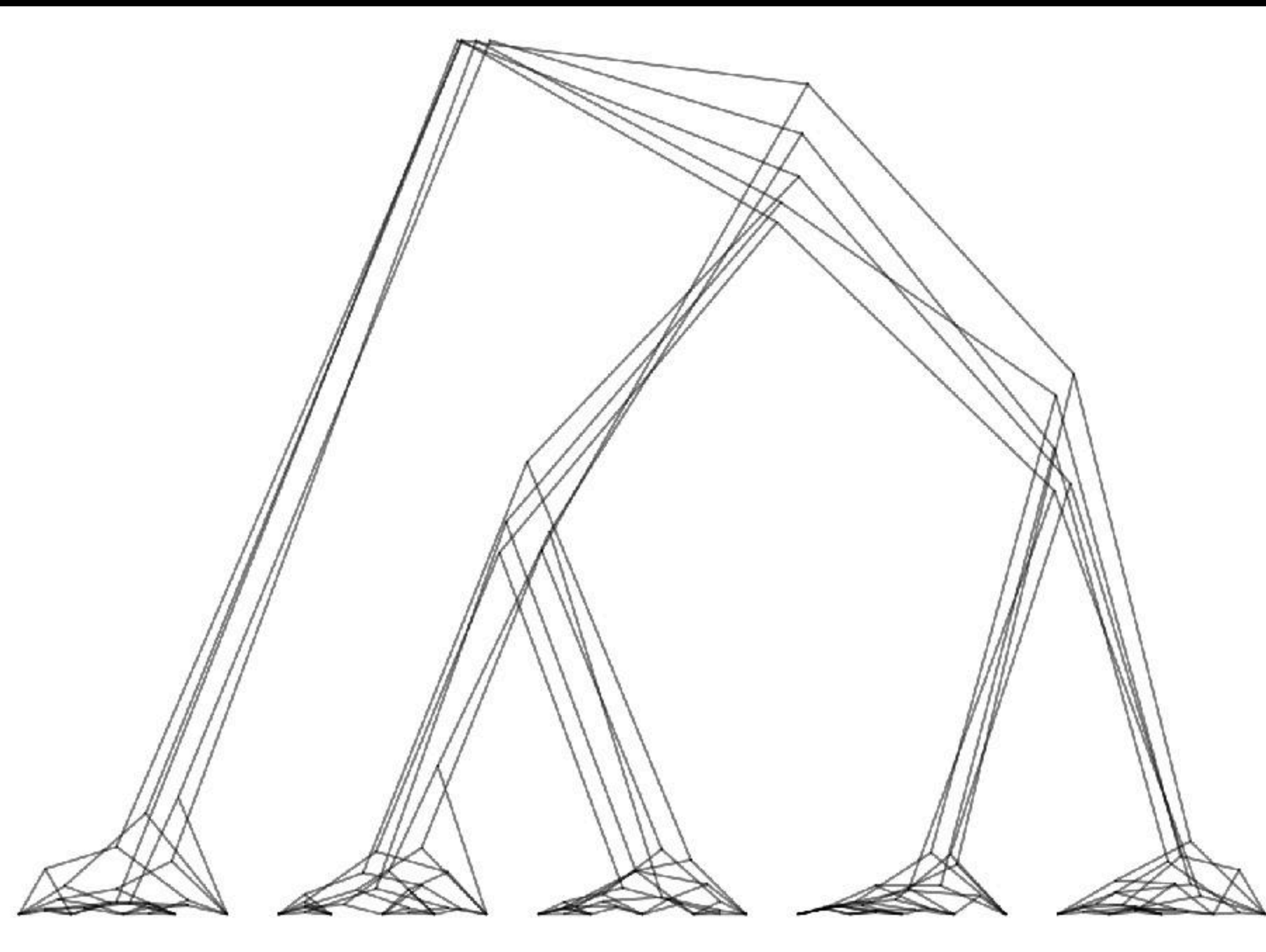
Trichonympha hueyi



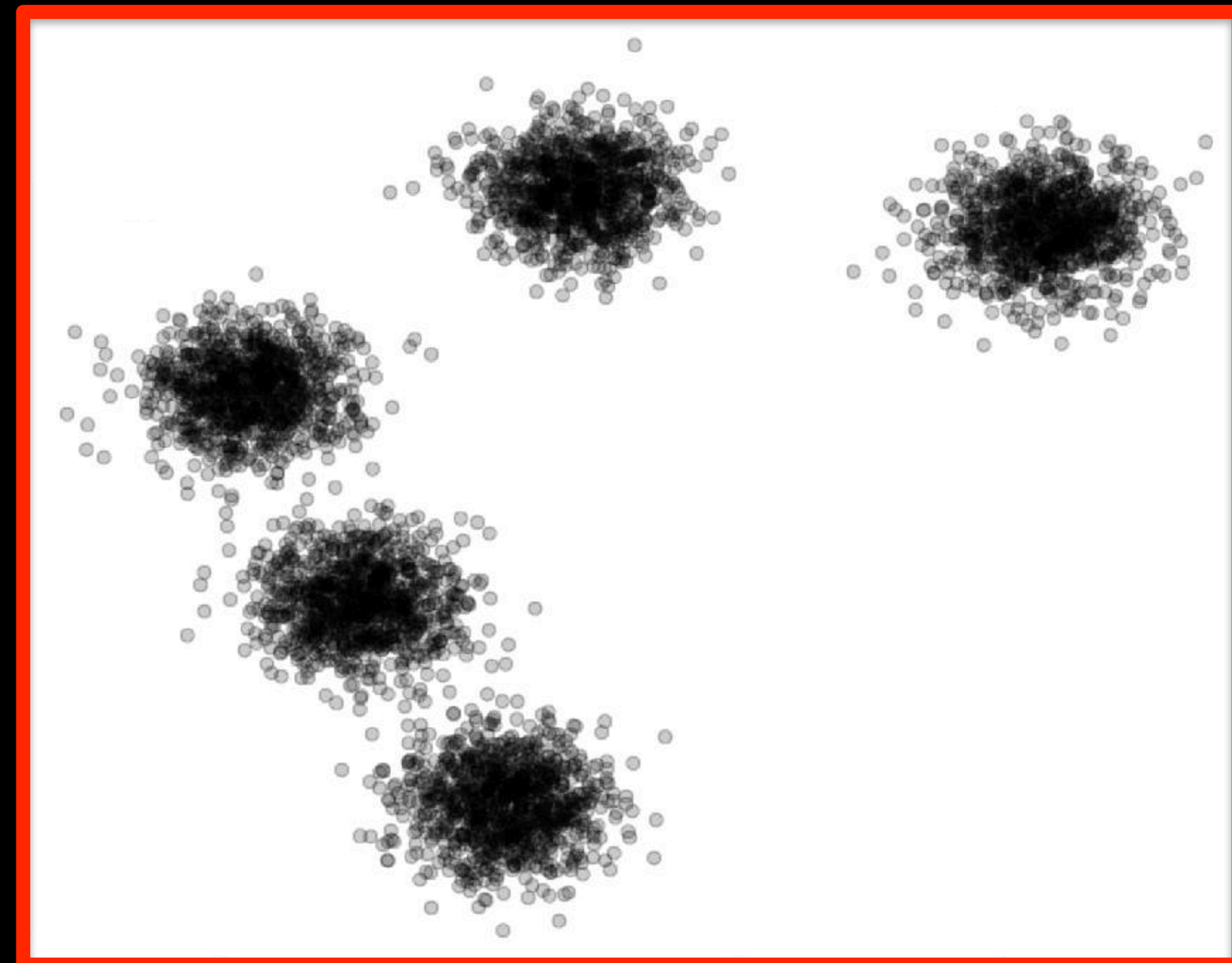
Trichonympha deweyi



Trichonympha louiei



trait 2



trait 1

phylogenetic species

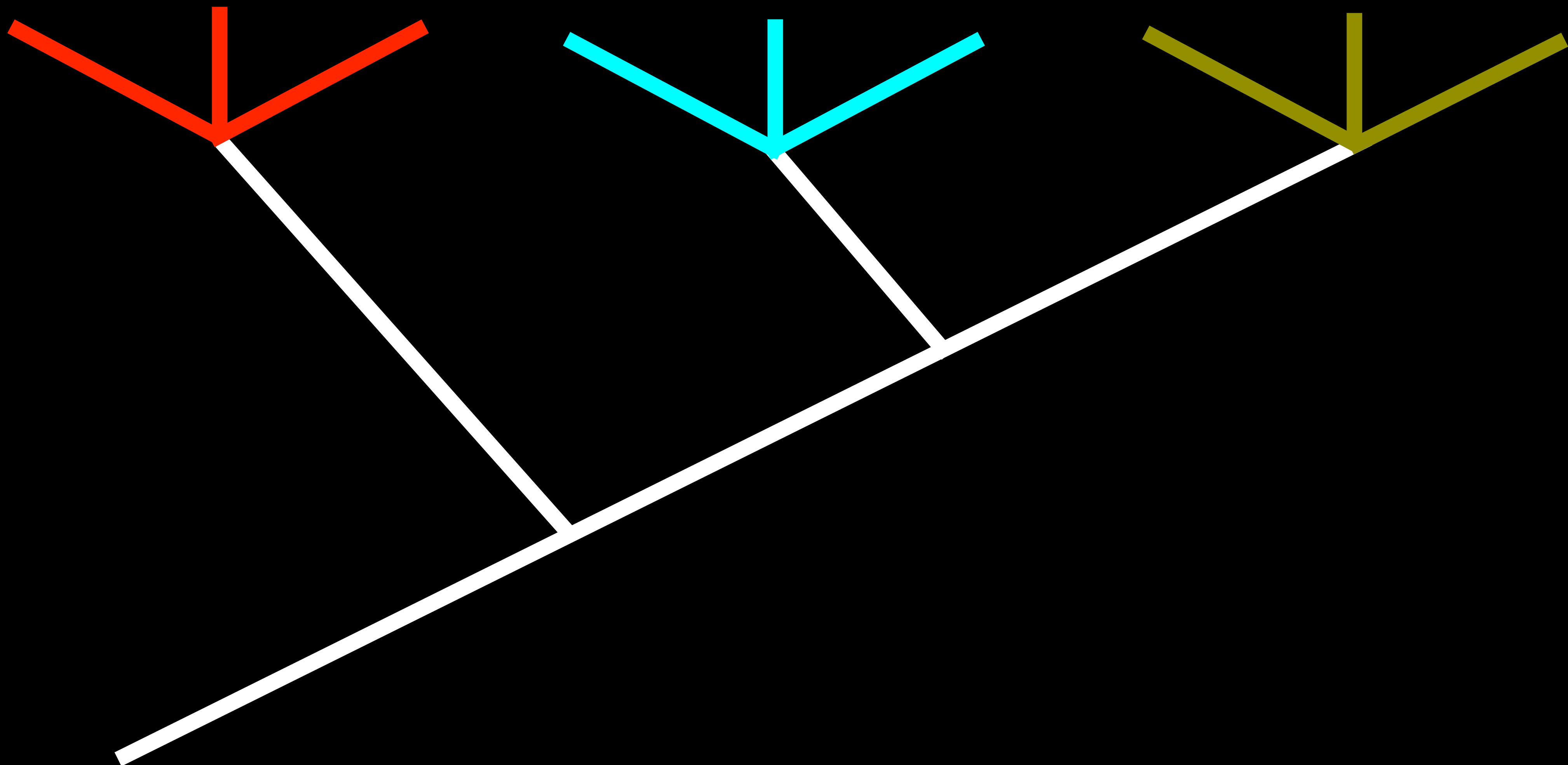
**= the smallest biological entities that
are diagnosable and/or monophyletic**

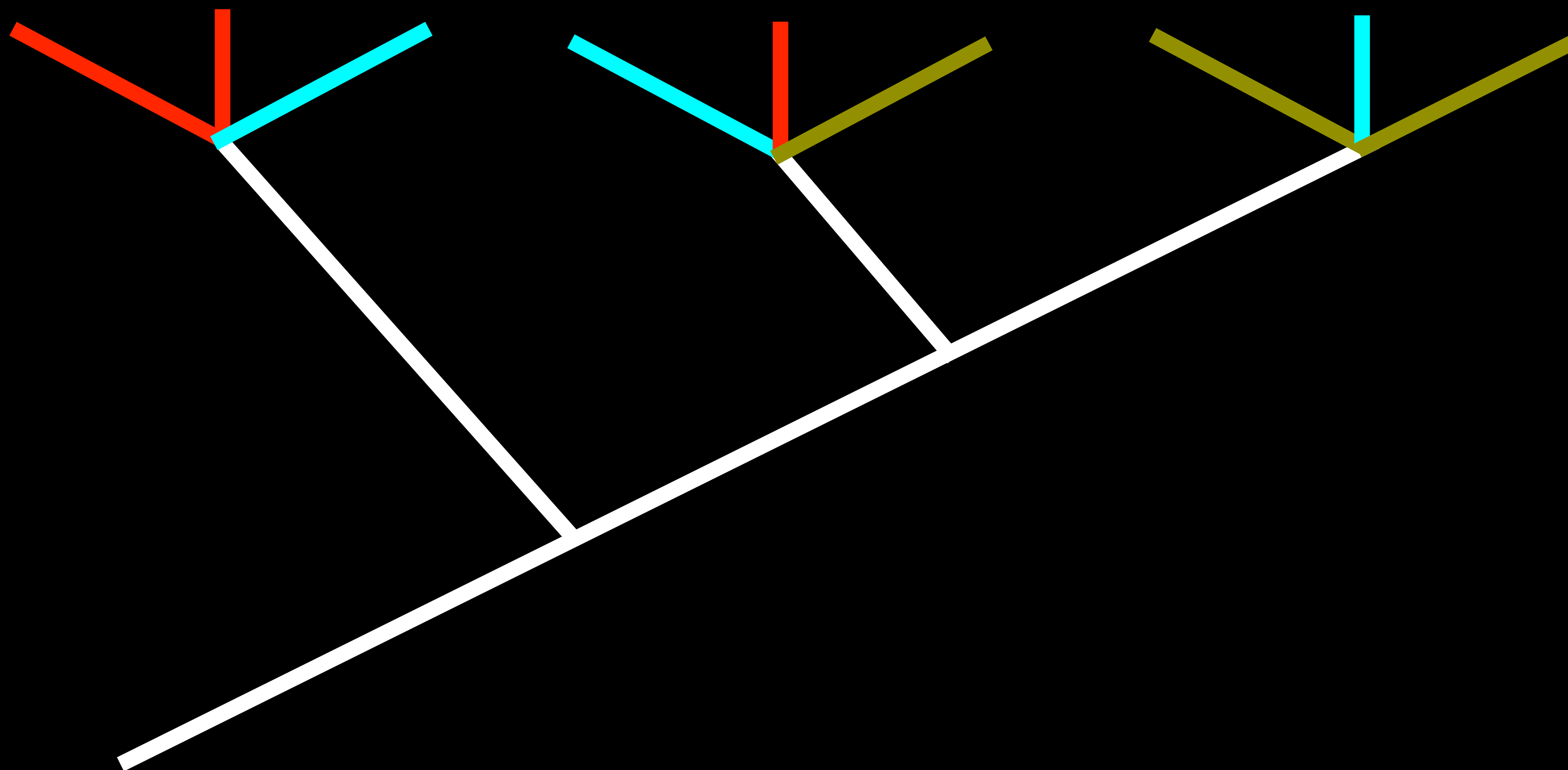
- Mayden 1997

species 1

species 2

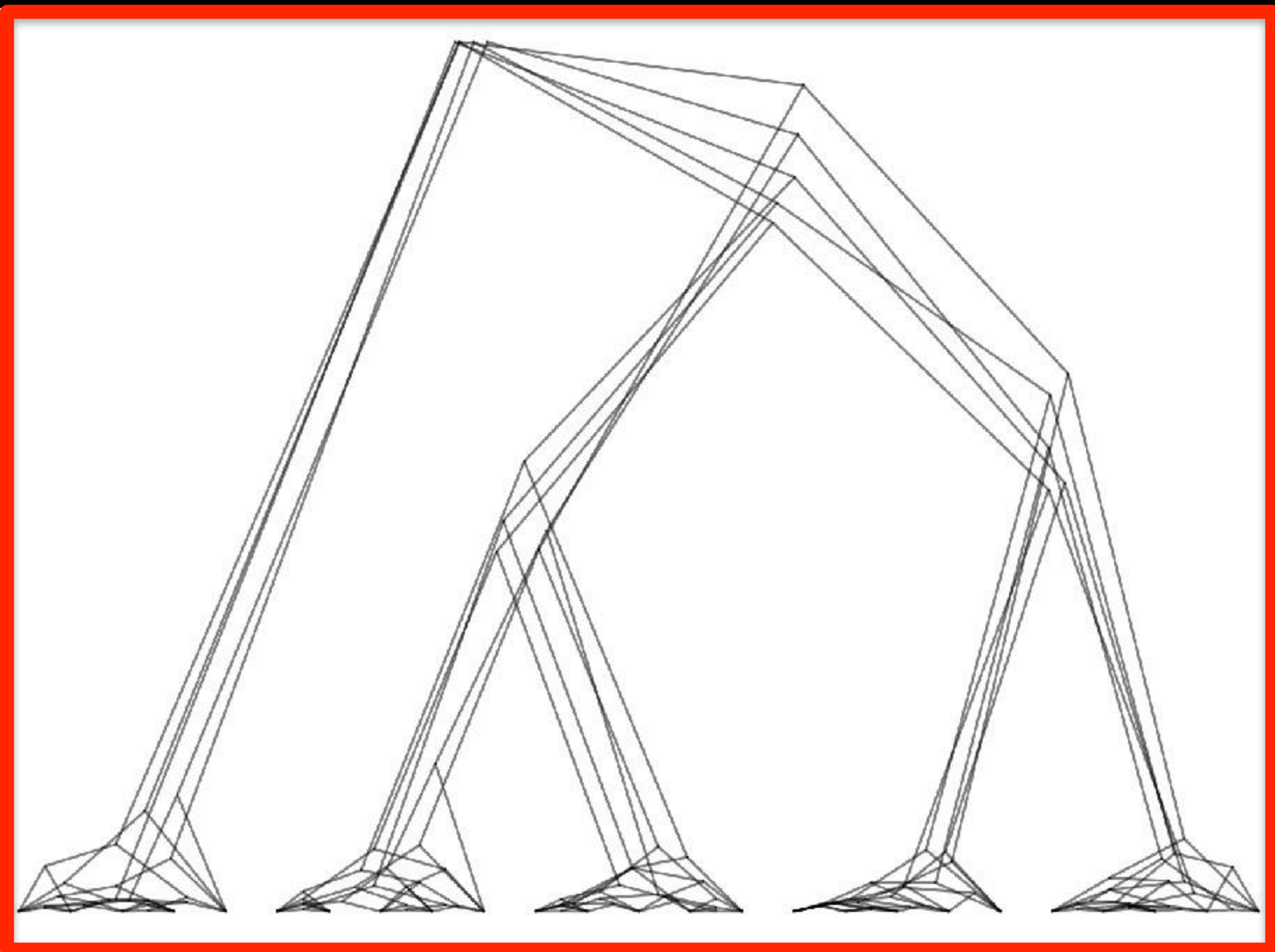
species 3



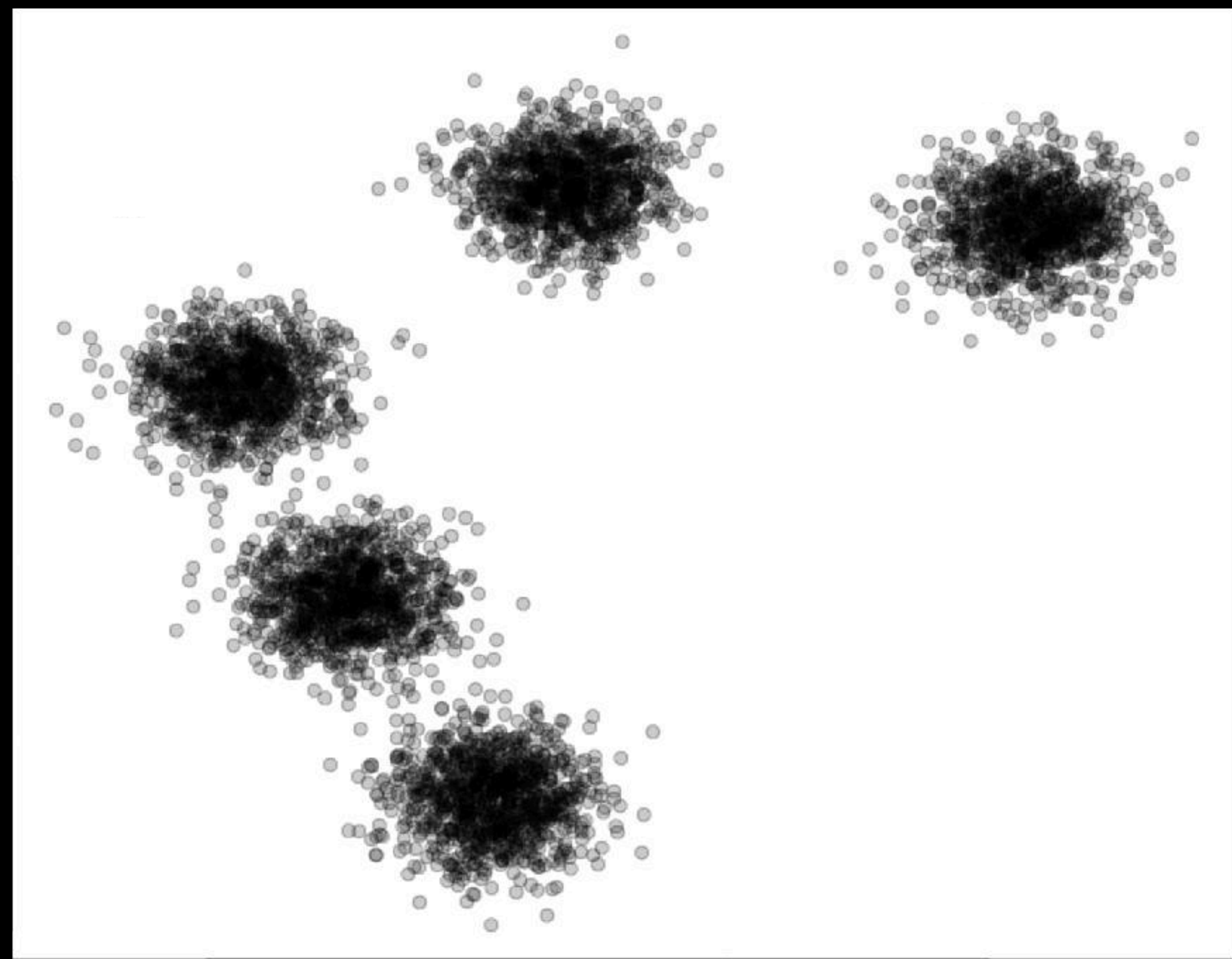




Paramecium



trait 2



trait 1

Syst. Biol. 56(6):879–886, 2007

Copyright © Society of Systematic Biologists

ISSN: 1063-5157 print / 1076-836X online

DOI: 10.1080/10635150701701083

Species Concepts and Species Delimitation

KEVIN DE QUEIROZ

*Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington,
DC 20560-0162, USA; E-mail: dequeirozk@si.edu*

concept of a species

= separately evolving metapopulation
lineages

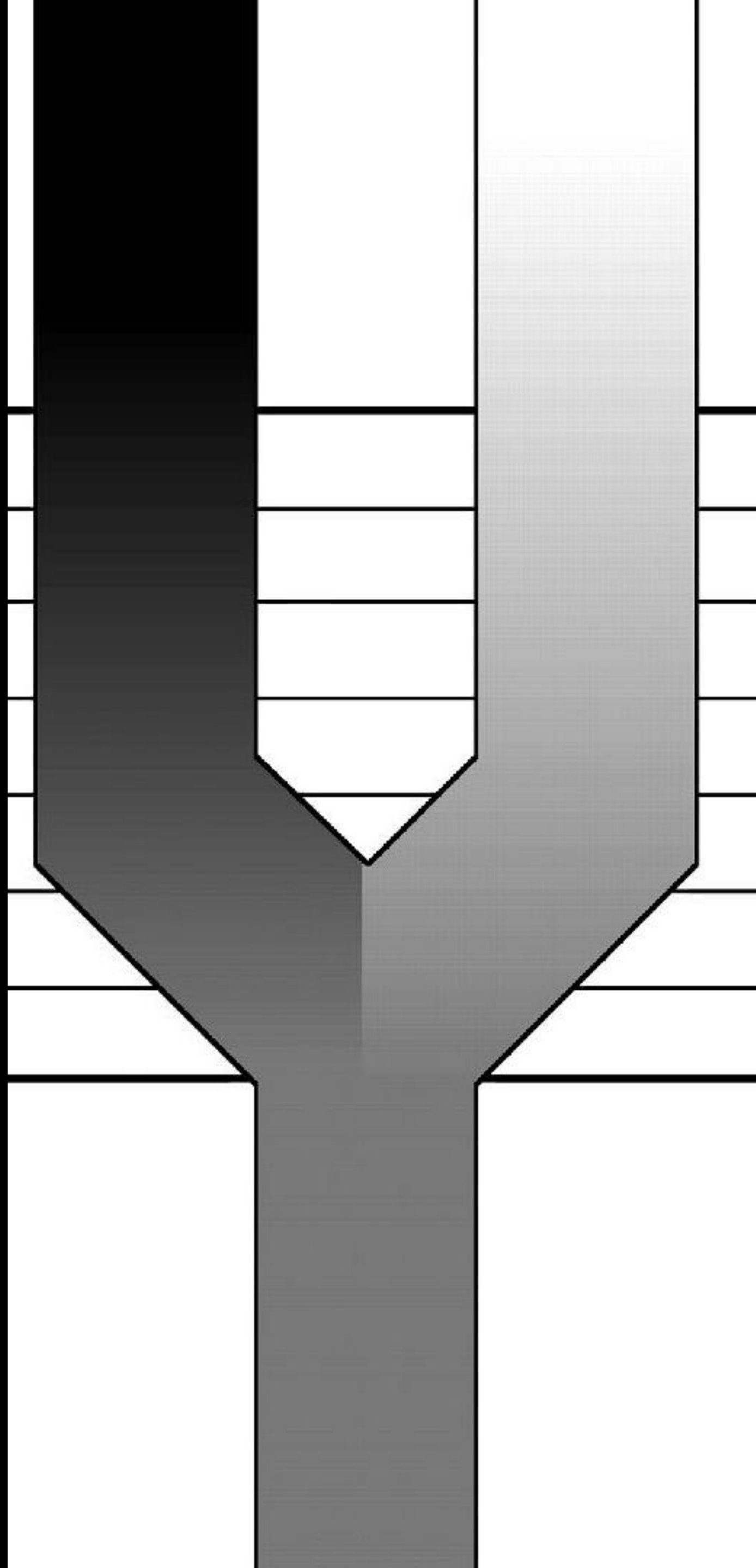
concept of a species

= separately evolving metapopulation
lineages

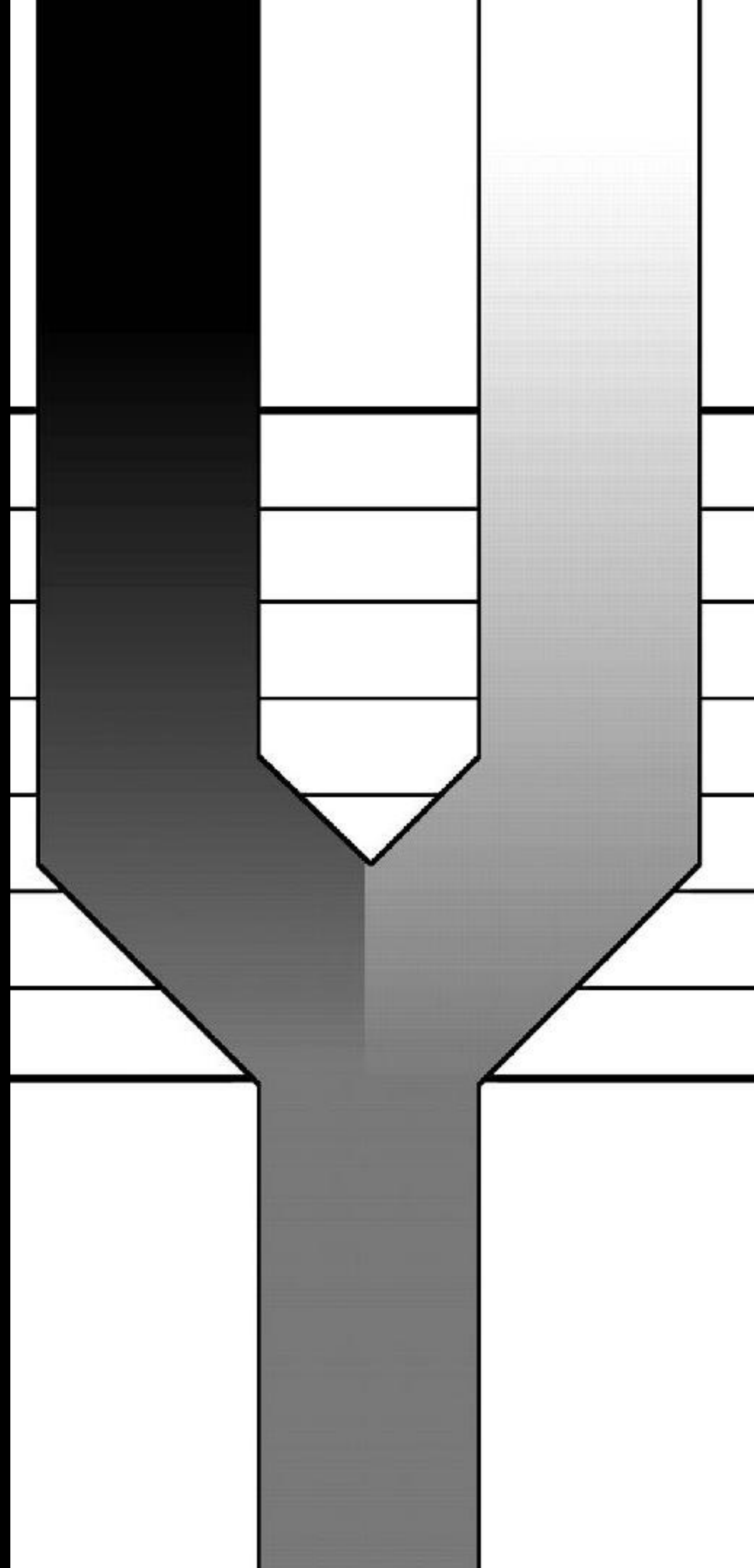
operational criteria for delimitation

= Biological Species, Morphological
Species, Phylogenetic Species, etc.

time \uparrow



time ↑



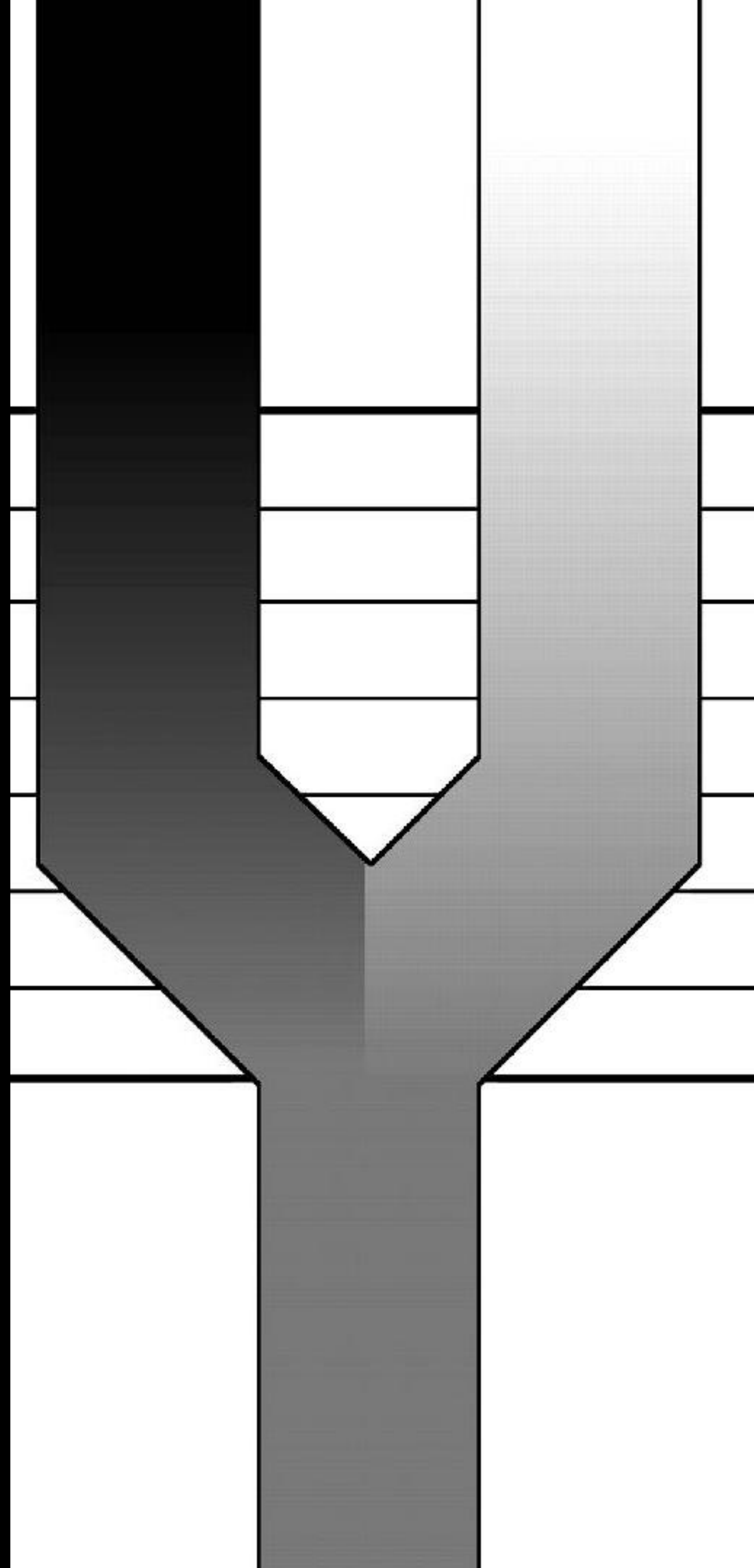
operational criteria

Phylogenetic

Morphological

Biological

time ↑



operational criteria

Morphological

Biological

Phylogenetic

clustering methods

just clustering

vs.

clustering and cleaning

de novo clustering

vs.

closed-reference clustering

metabarcoding data

vs.

metagenomic/

metatranscriptomic

2002 75:509-516

Biol. J. Linn. Soc.

On the use of genetic divergence for identifying species

J. WILLEM H. FERGUSON*

Department Zoology and Entomology, University of Pretoria, 0002 Pretoria, South Africa

**global clustering
thresholds**

**local clustering
thresholds**

pairwise
comparisons

phylogenetic
comparisons

global clustering
thresholds

local clustering
thresholds

**pairwise
comparisons**

**phylogenetic
comparisons**

**global clustering
thresholds**

**VSEARCH/
Mothur**

DADA2

**local clustering
thresholds**

Swarm

**multi-rate
PTP**

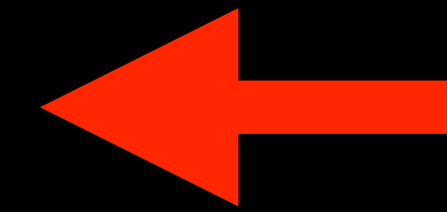
**pairwise
comparisons**

**phylogenetic
comparisons**

**global clustering
thresholds**

**VSEARCH/
Mothur**

DADA2



**local clustering
thresholds**

Swarm

**multi-rate
PTP**



VSEARCH: a versatile open source tool for metagenomics

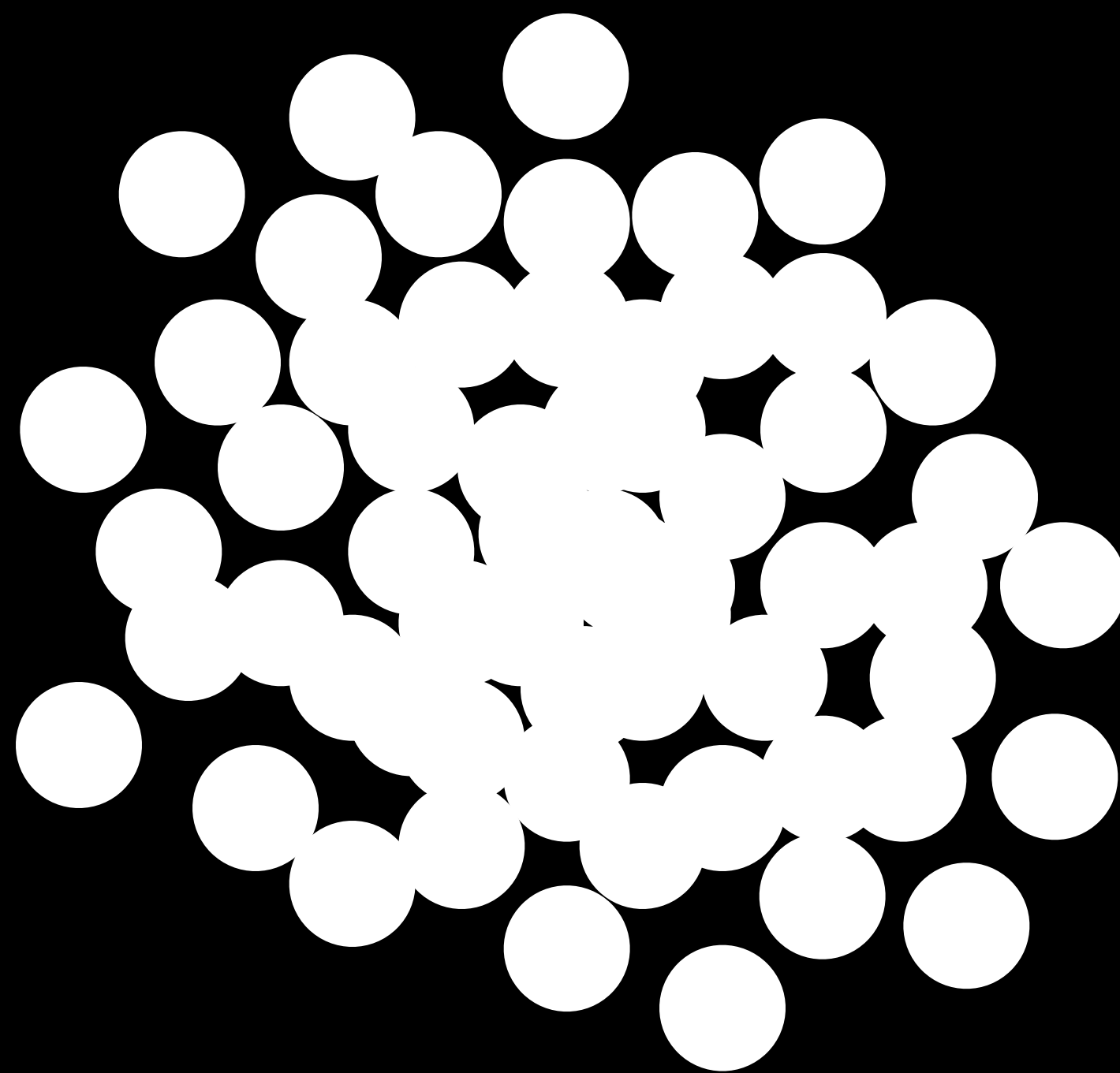
Torbjørn Rognes^{1,2}, Tomáš Flouri^{3,4}, Ben Nichols⁵, Christopher Quince^{5,6} and Frédéric Mahé^{7,8}

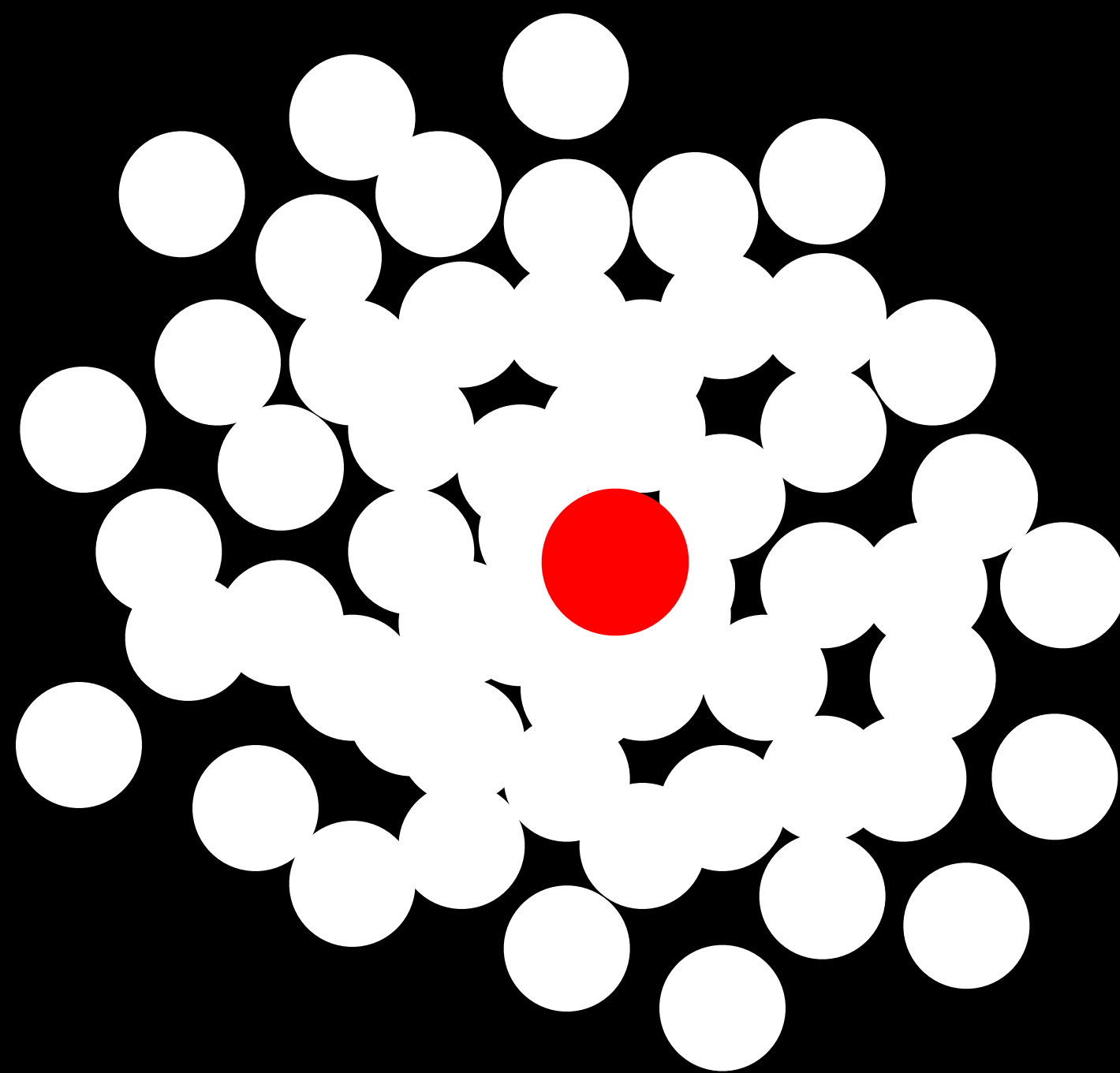
APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Dec. 2009, p. 7537–7541
0099-2240/09/\$12.00 doi:10.1128/AEM.01541-09
Copyright © 2009, American Society for Microbiology. All Rights Reserved.

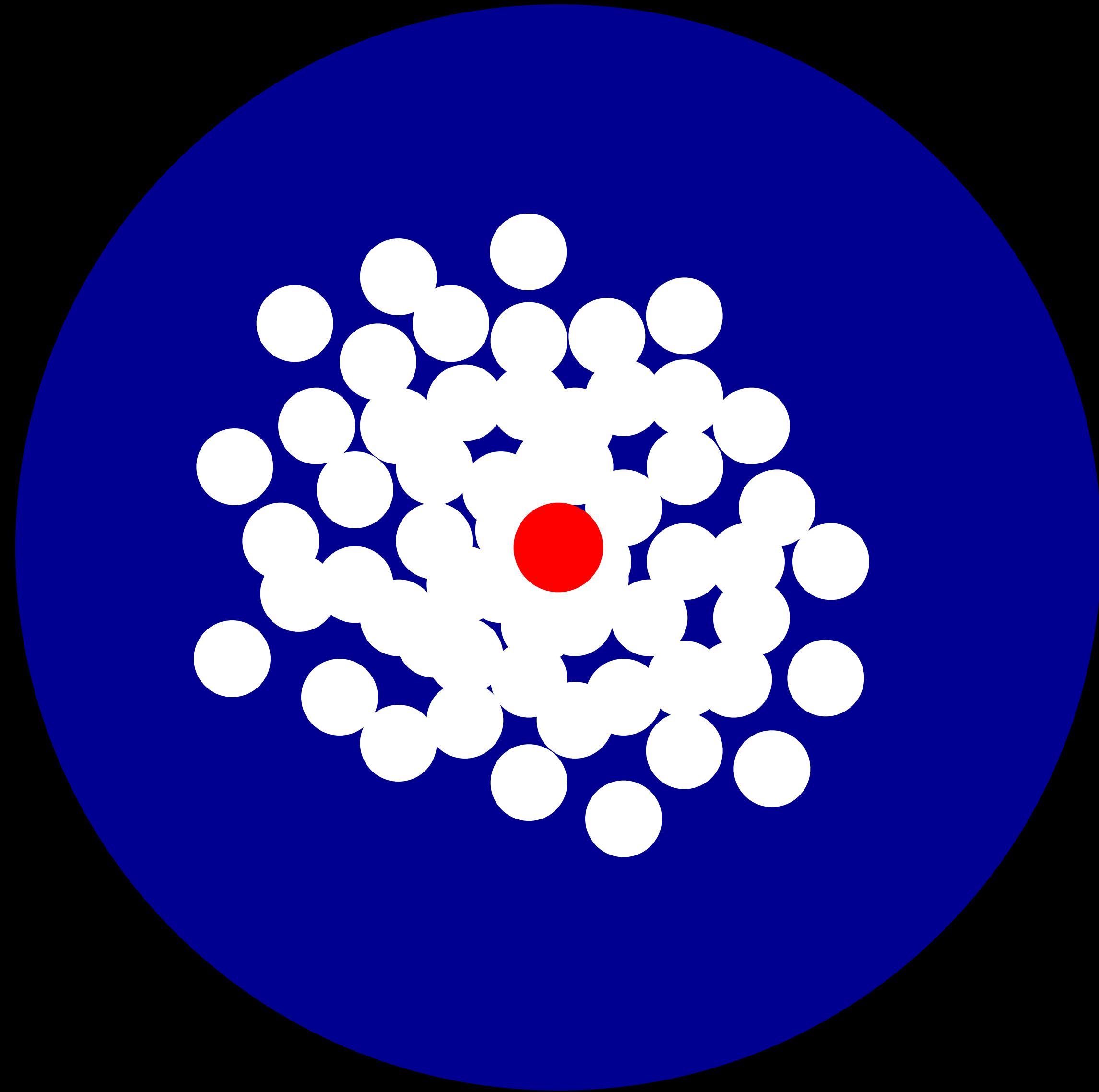
Vol. 75, No. 23

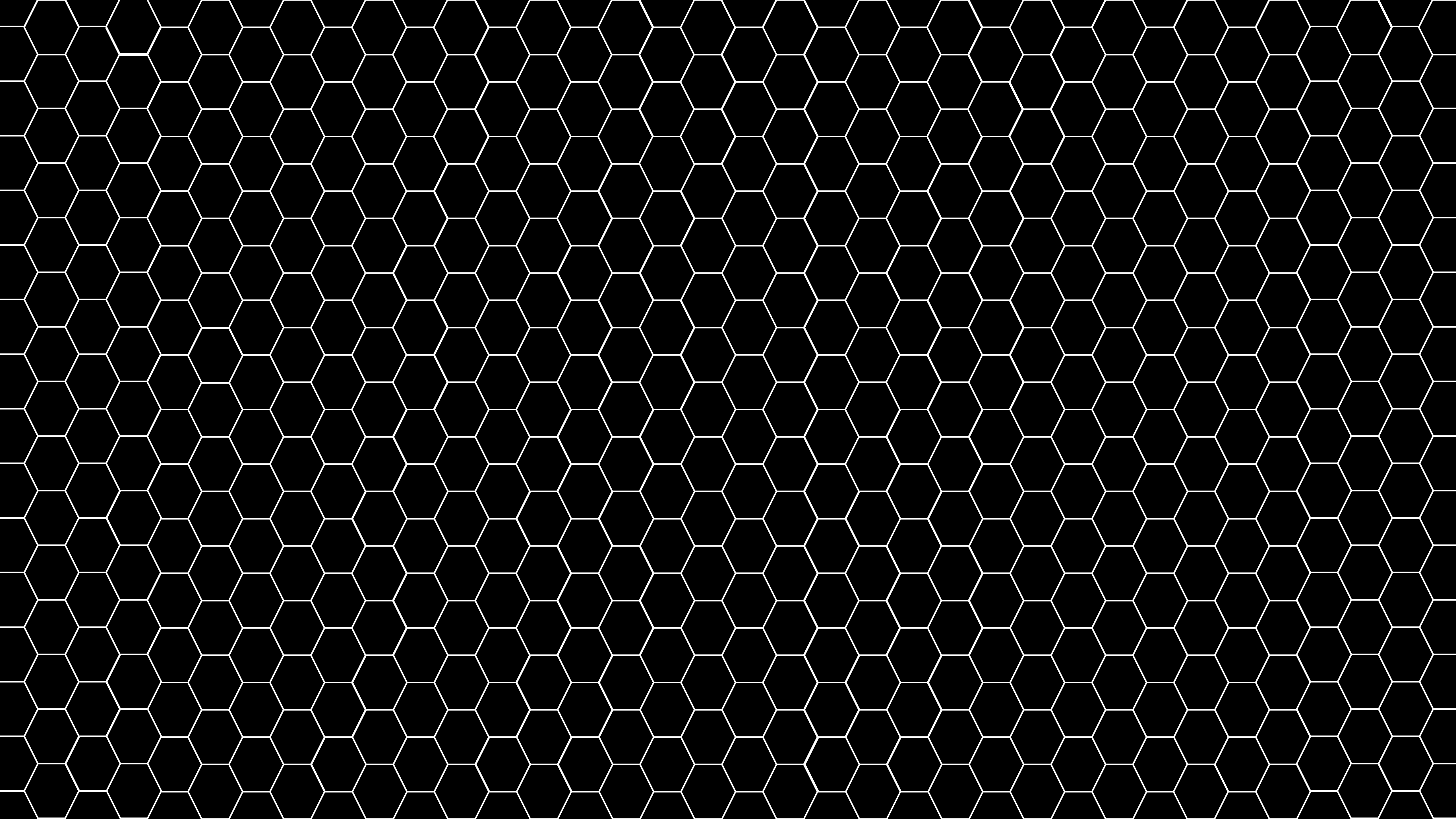
Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities[▽]

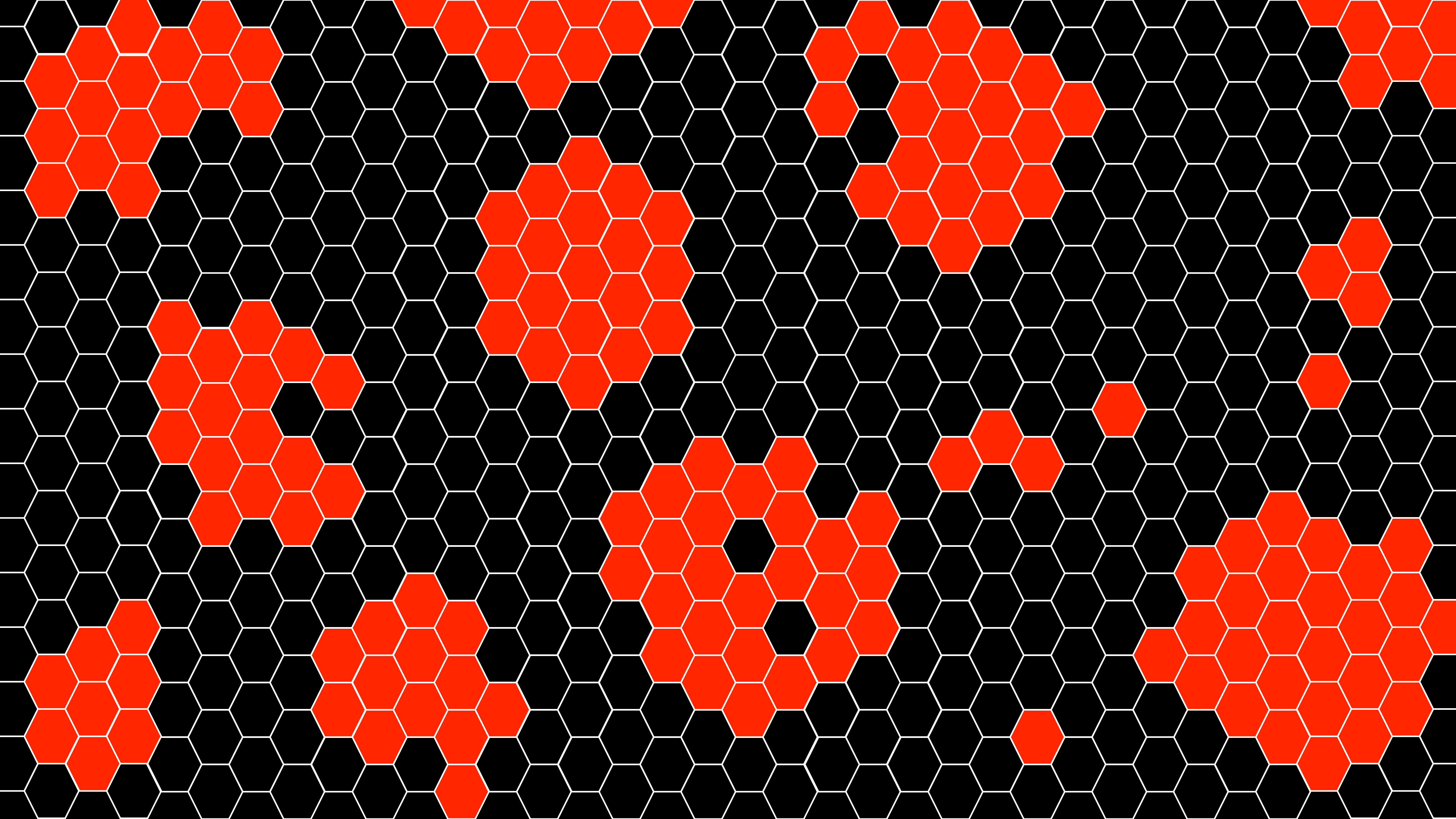
Patrick D. Schloss,^{1,2*} Sarah L. Westcott,^{1,2} Thomas Ryabin,¹ Justine R. Hall,³ Martin Hartmann,⁴ Emily B. Hollister,⁵ Ryan A. Lesniewski,⁶ Brian B. Oakley,⁷ Donovan H. Parks,⁸ Courtney J. Robinson,² Jason W. Sahl,⁹ Blaz Stres,¹⁰ Gerhard G. Thallinger,¹¹ David J. Van Horn,² and Carolyn F. Weber¹²

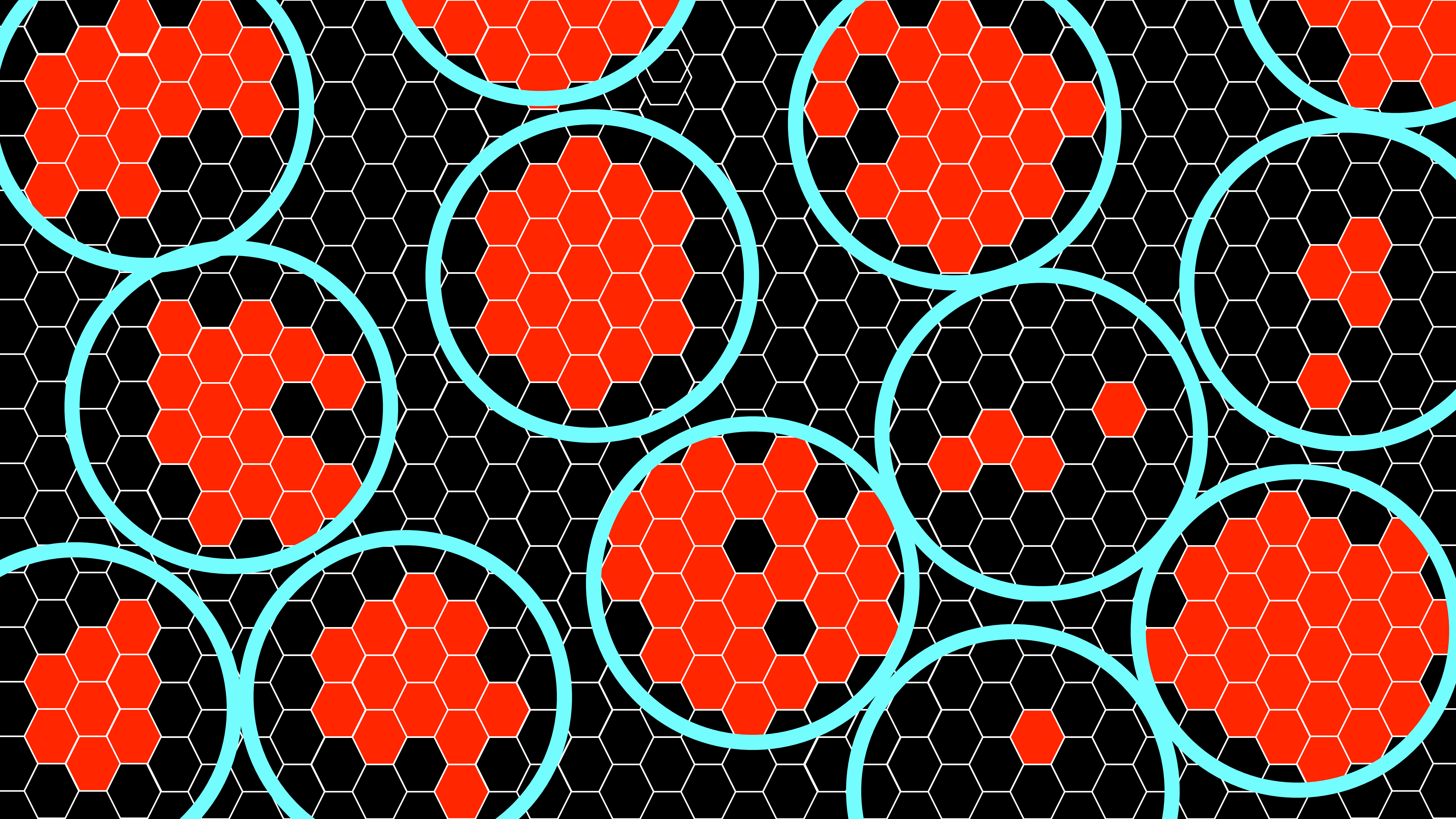


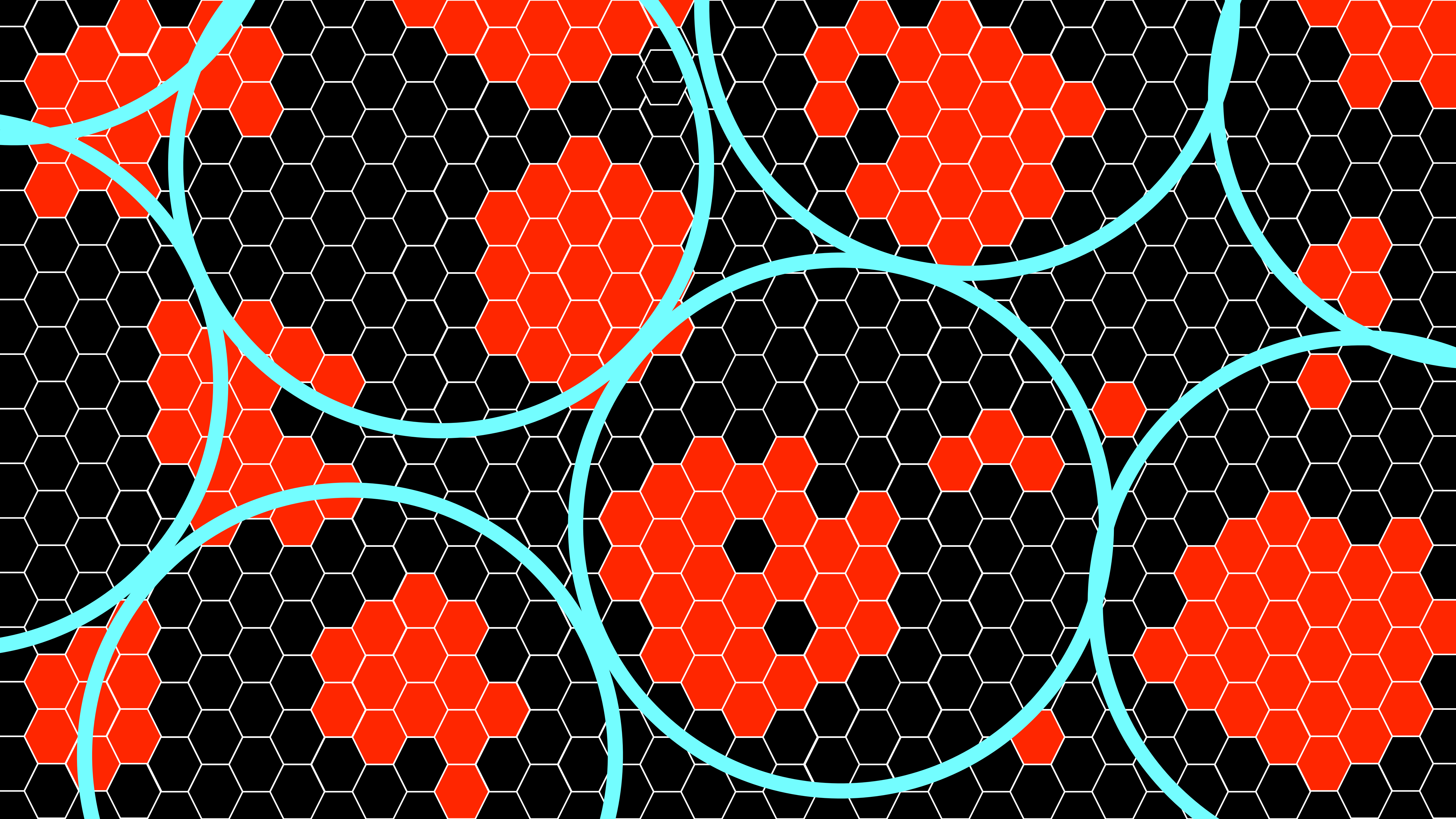


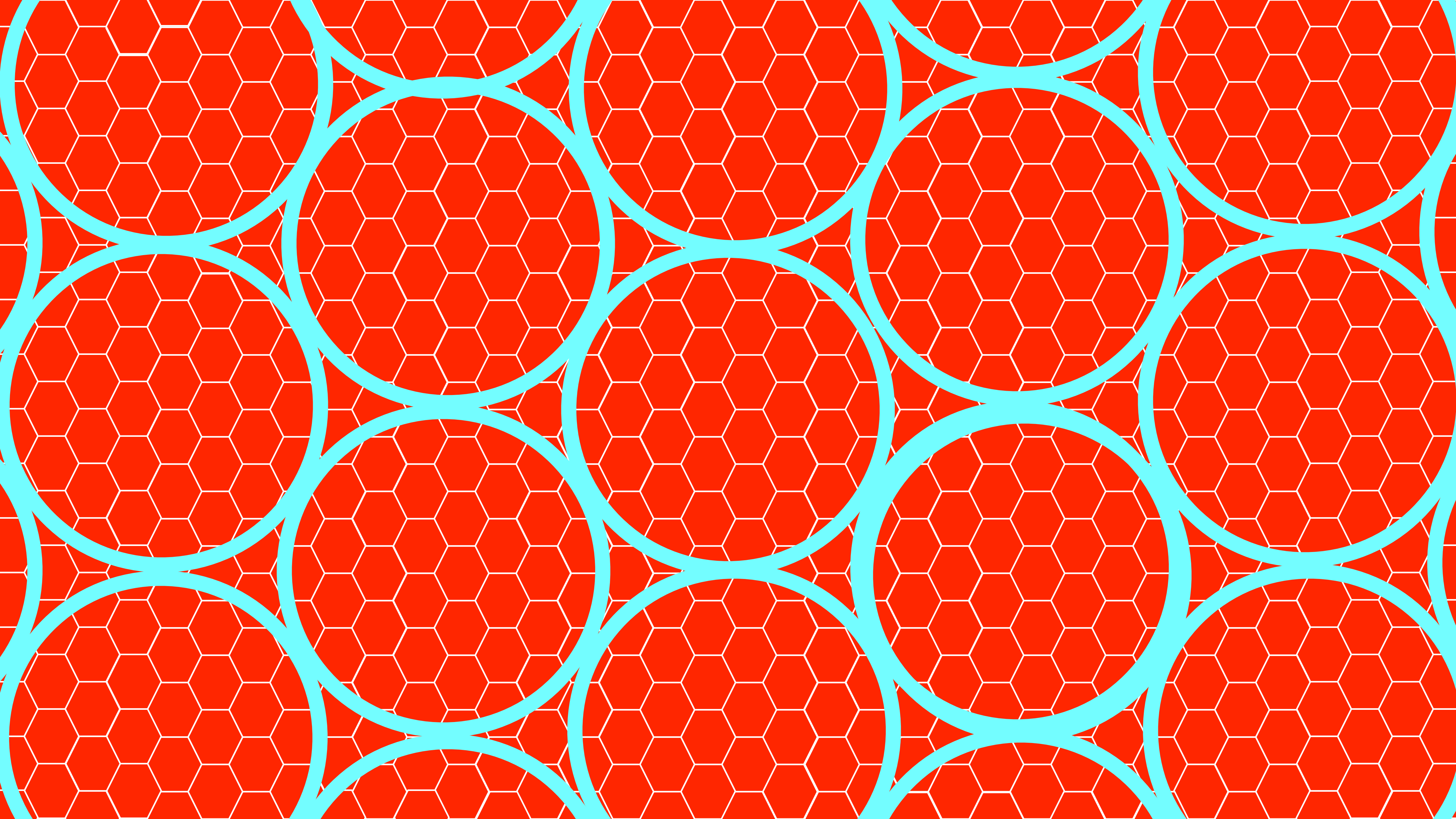


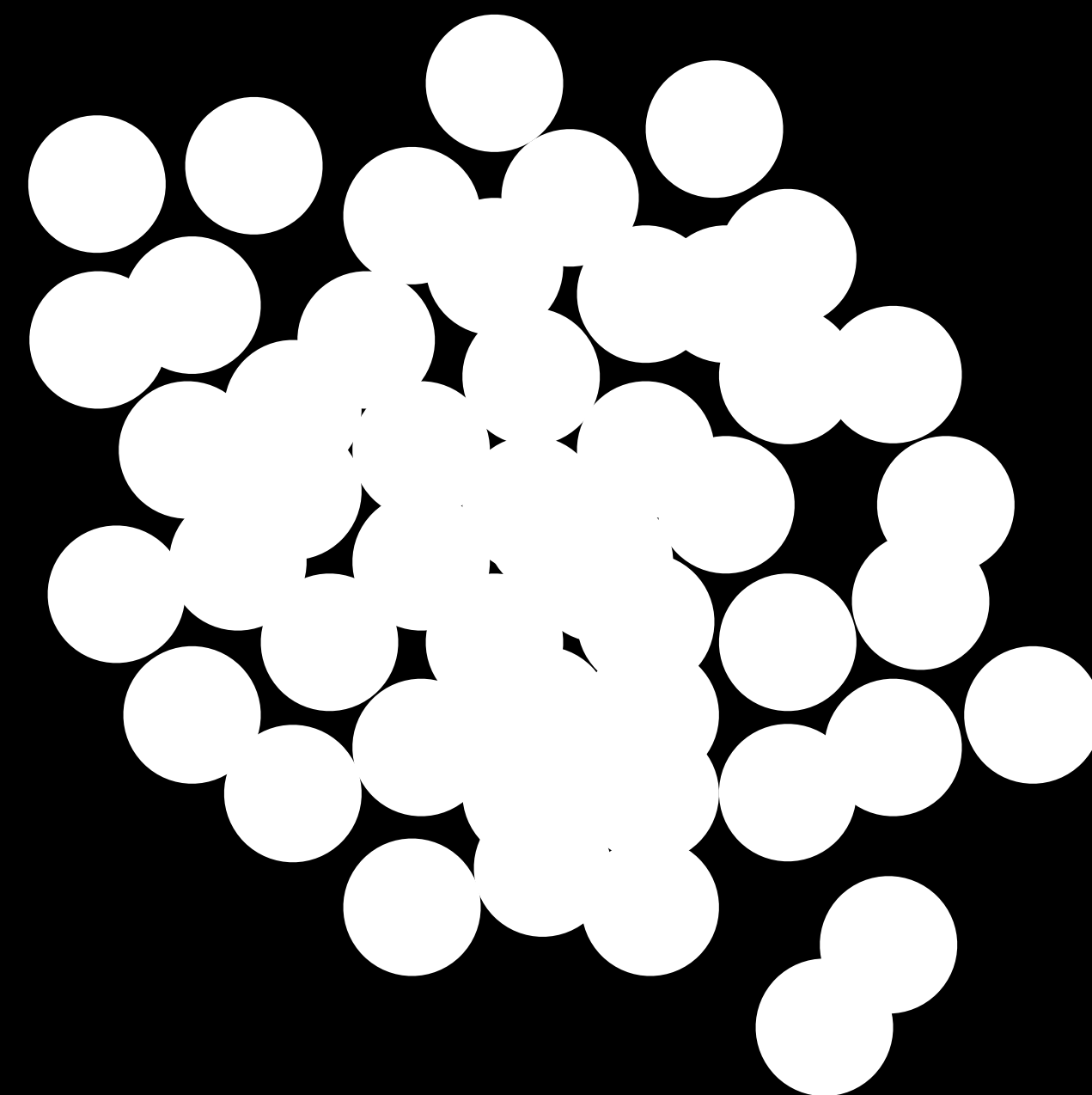
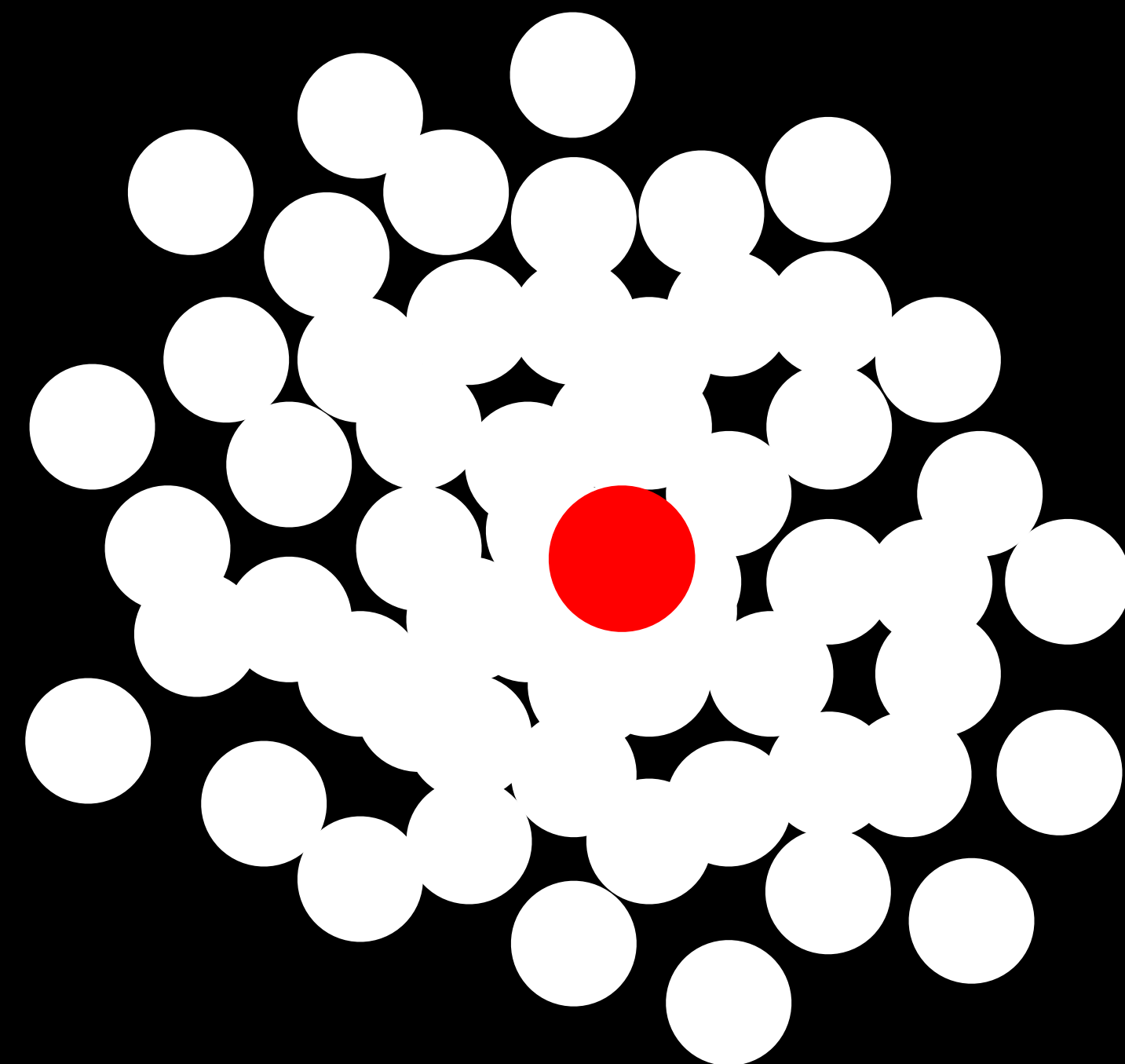


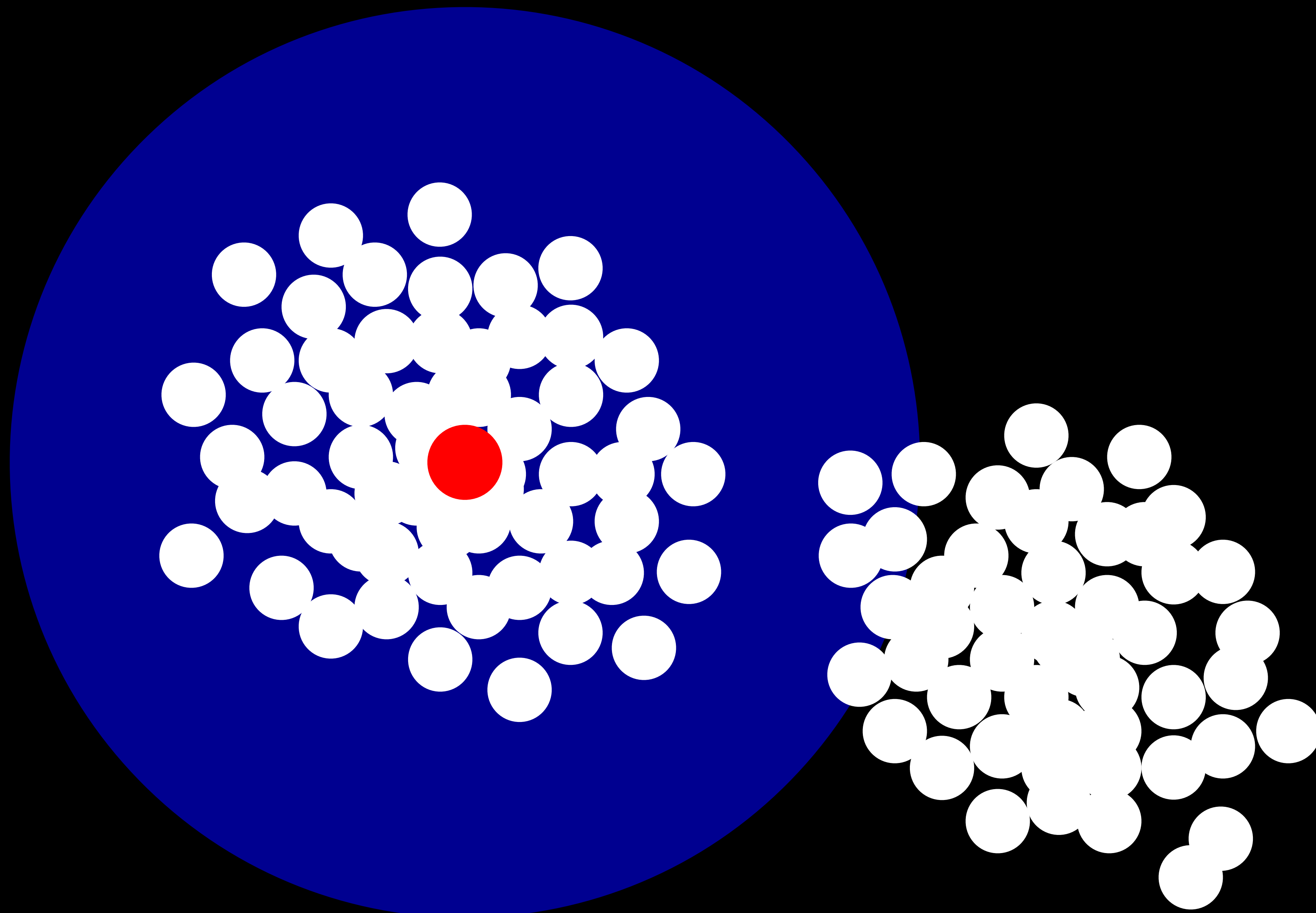


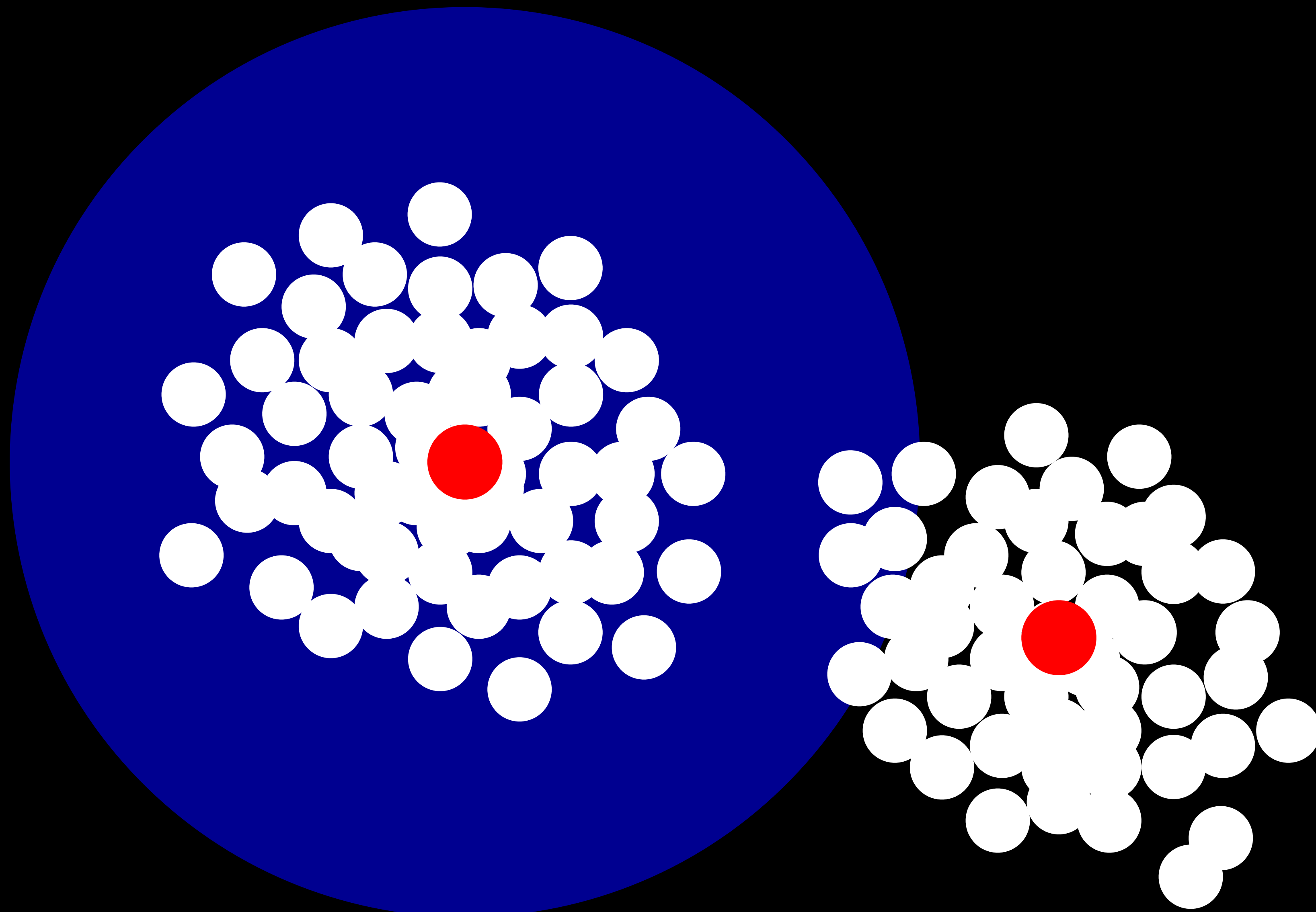


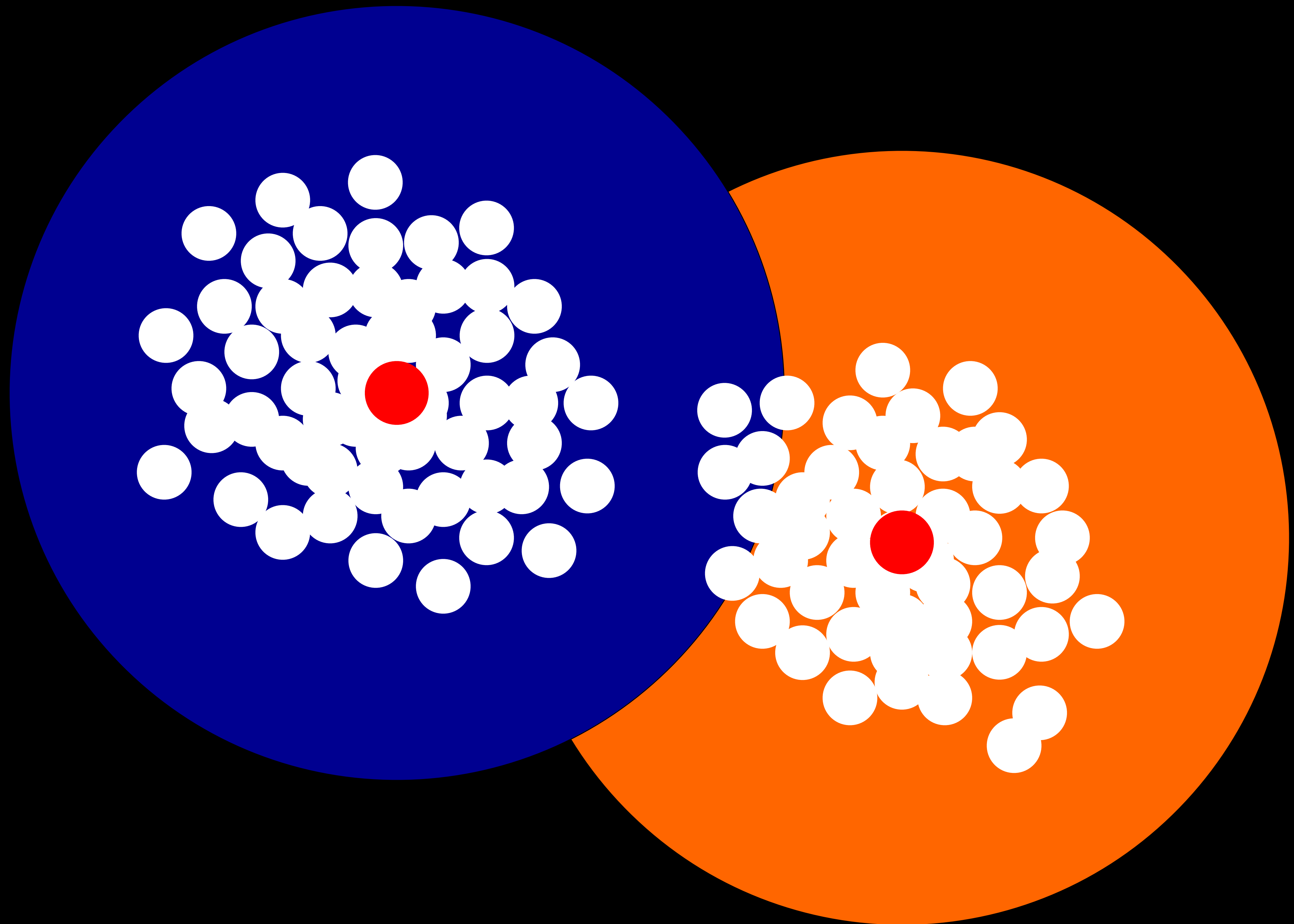


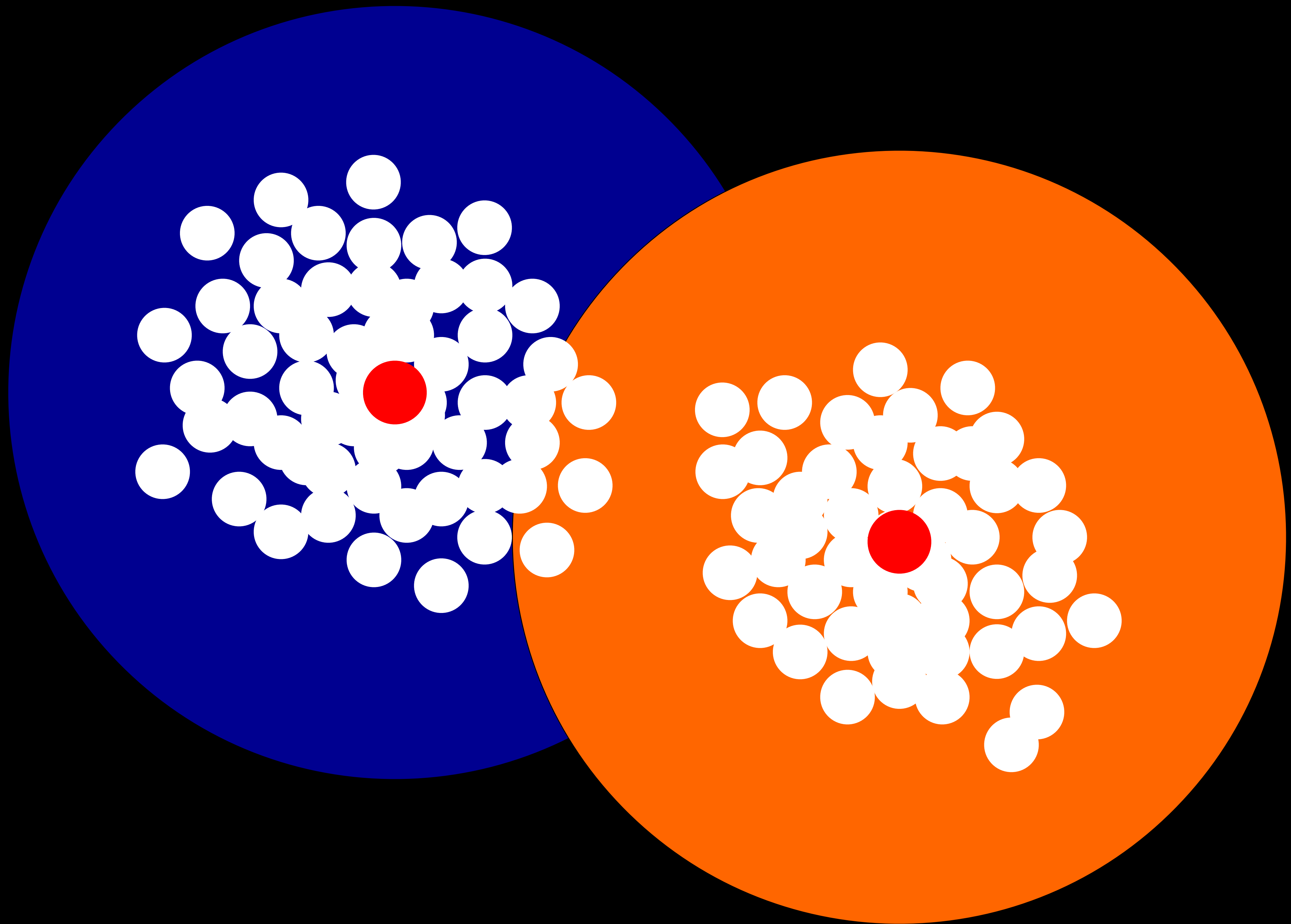




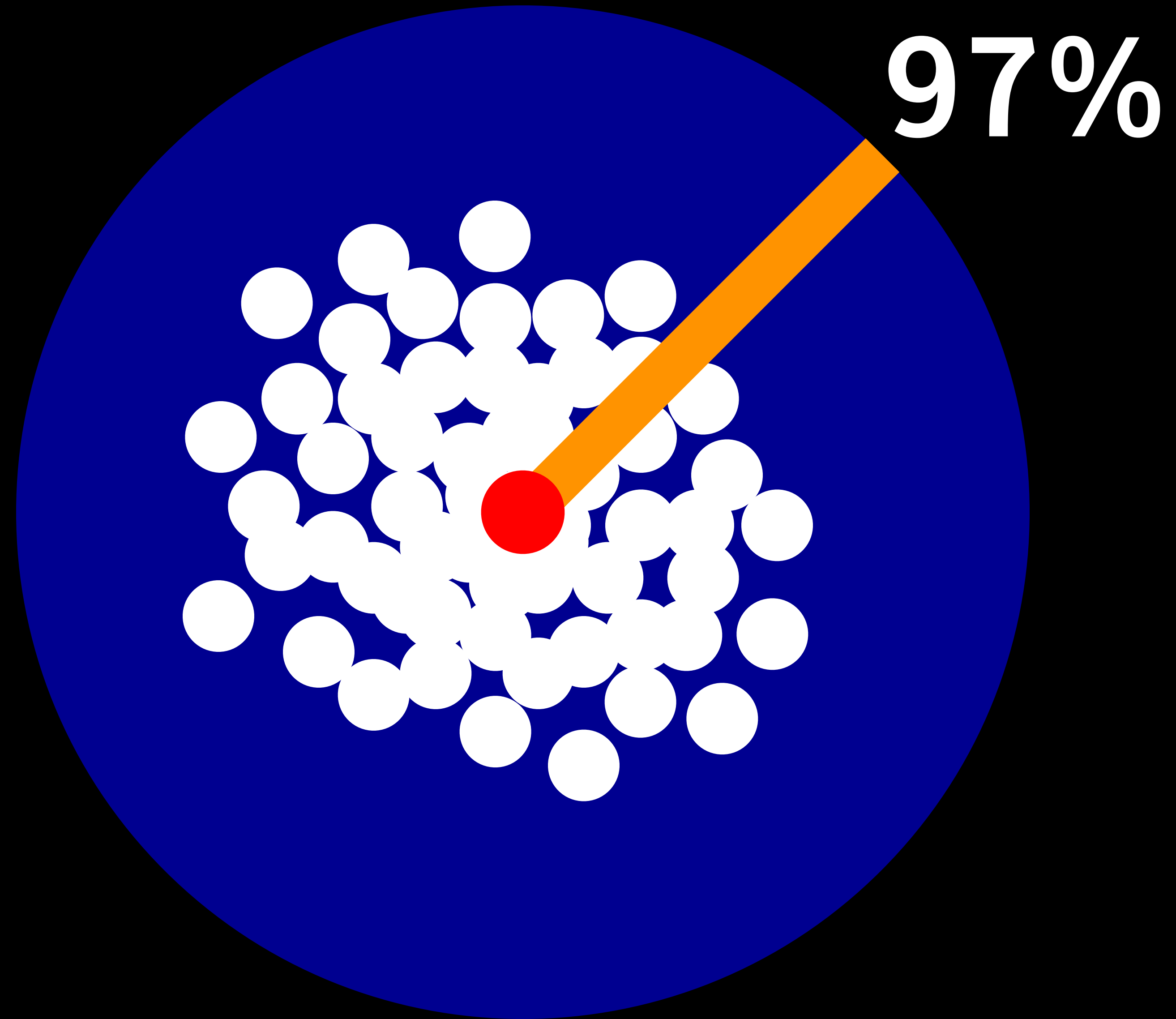


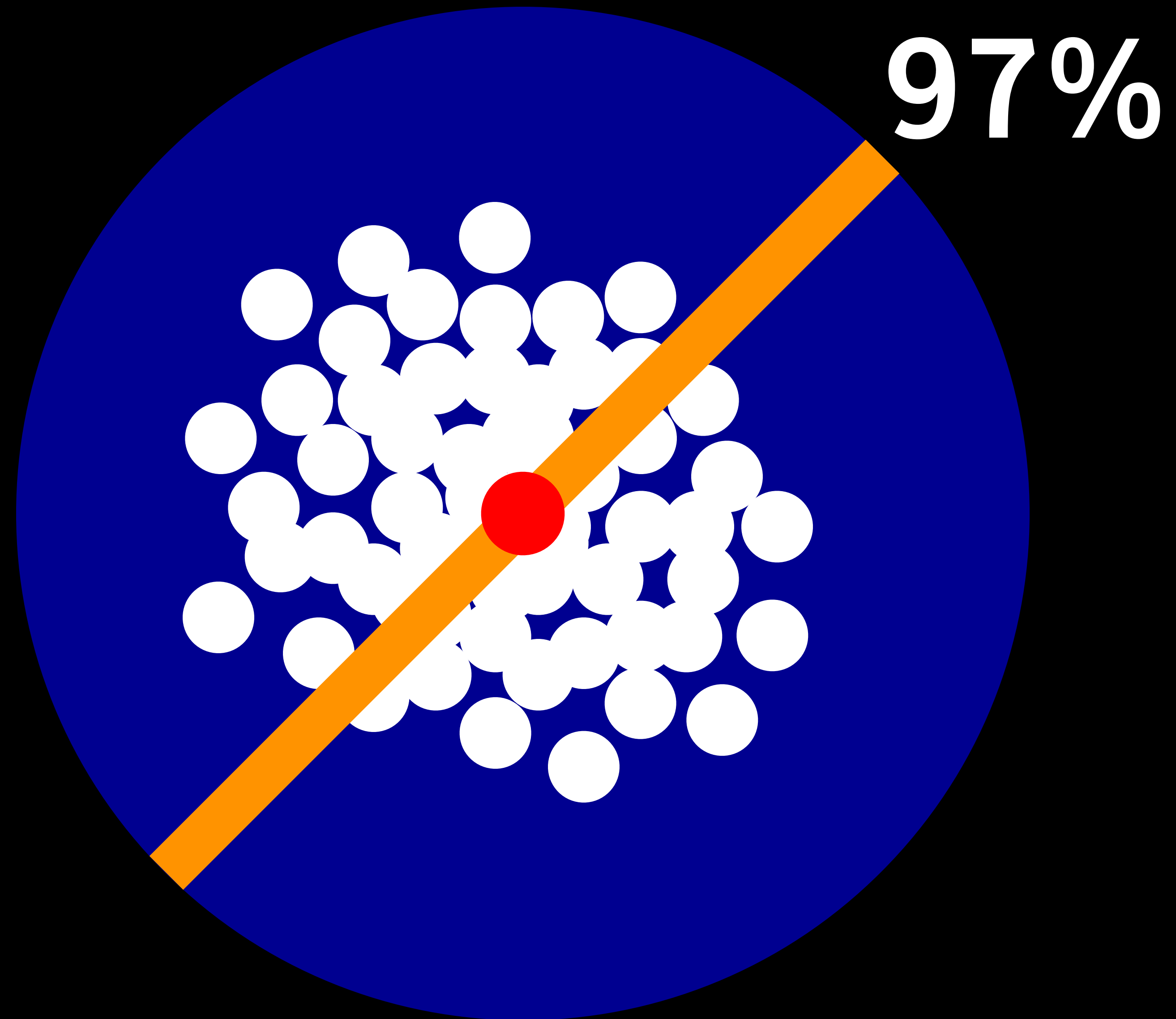


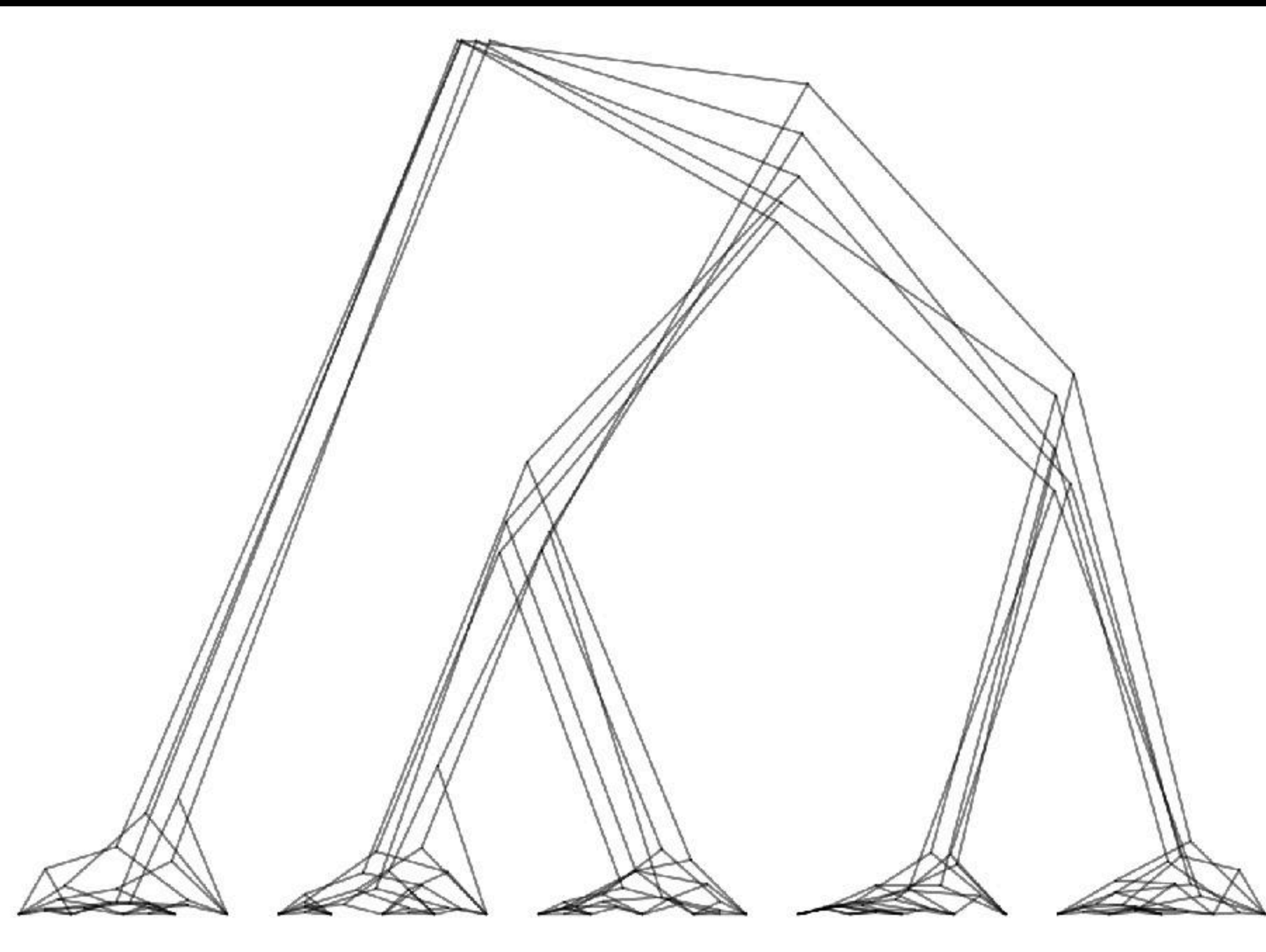




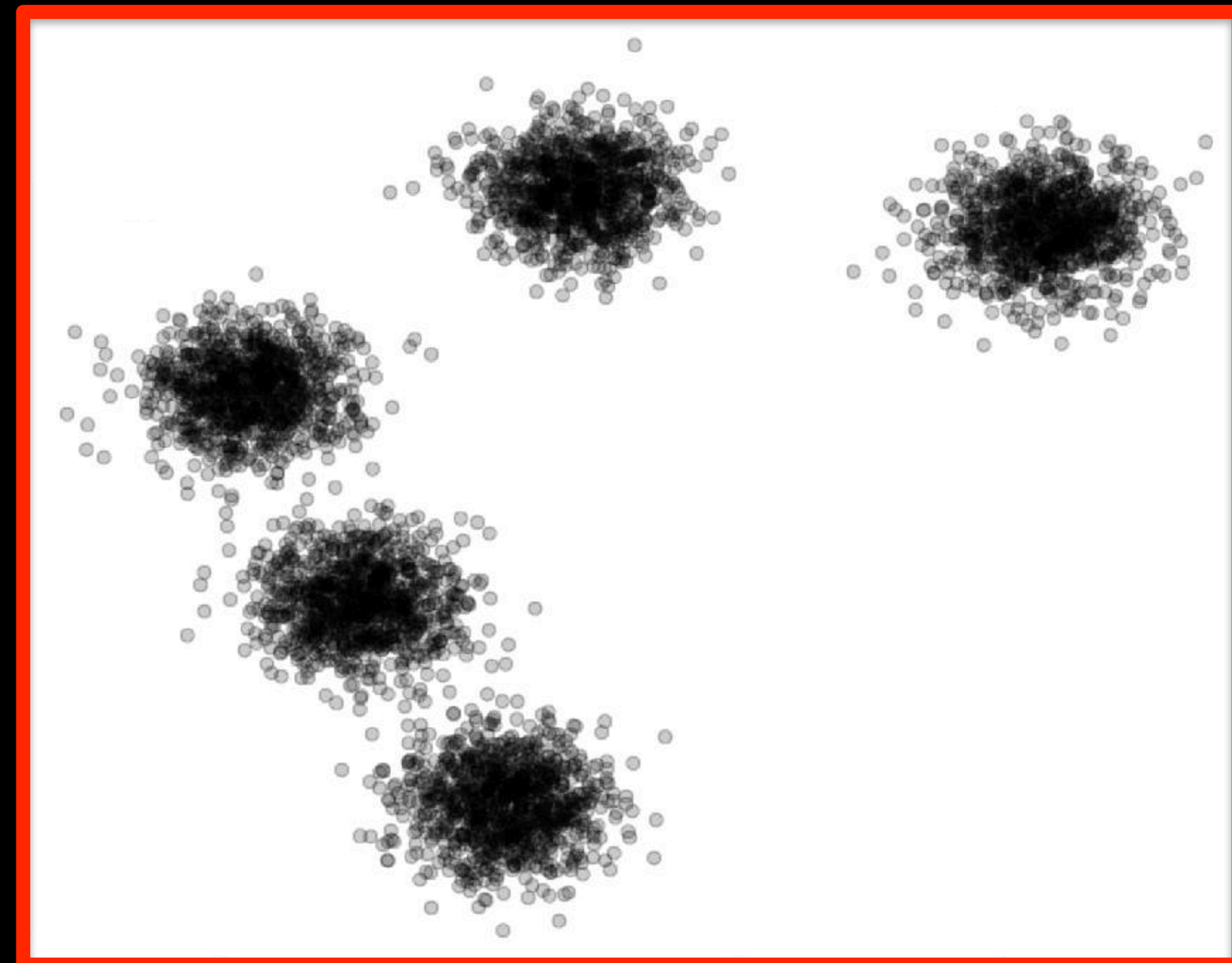
97%







trait 2



trait 1

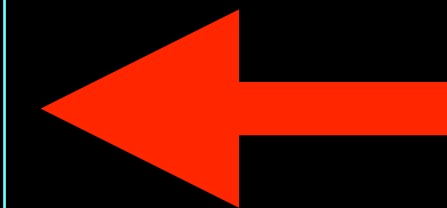
**pairwise
comparisons**

**phylogenetic
comparisons**

**global clustering
thresholds**

**VSEARCH/
Mothur**

DADA2



**local clustering
thresholds**

Swarm

**multi-rate
PTP**

2016 13:581-583

DADA2: High-resolution sample inference from Illumina amplicon data

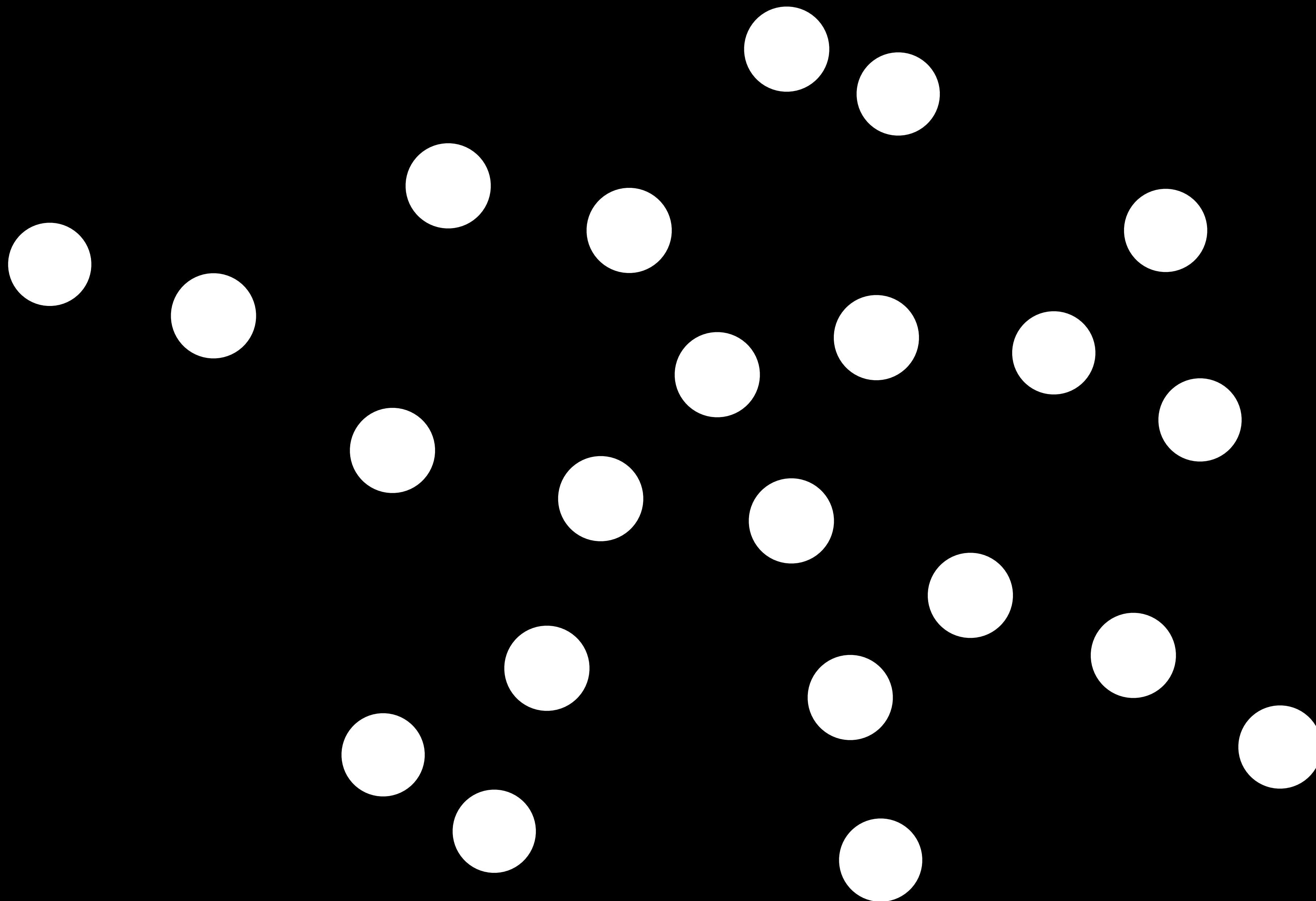
Benjamin J Callahan¹, Paul J McMurdie²,
Michael J Rosen³, Andrew W Han², Amy Jo A Johnson² &
Susan P Holmes¹

Nature Methods

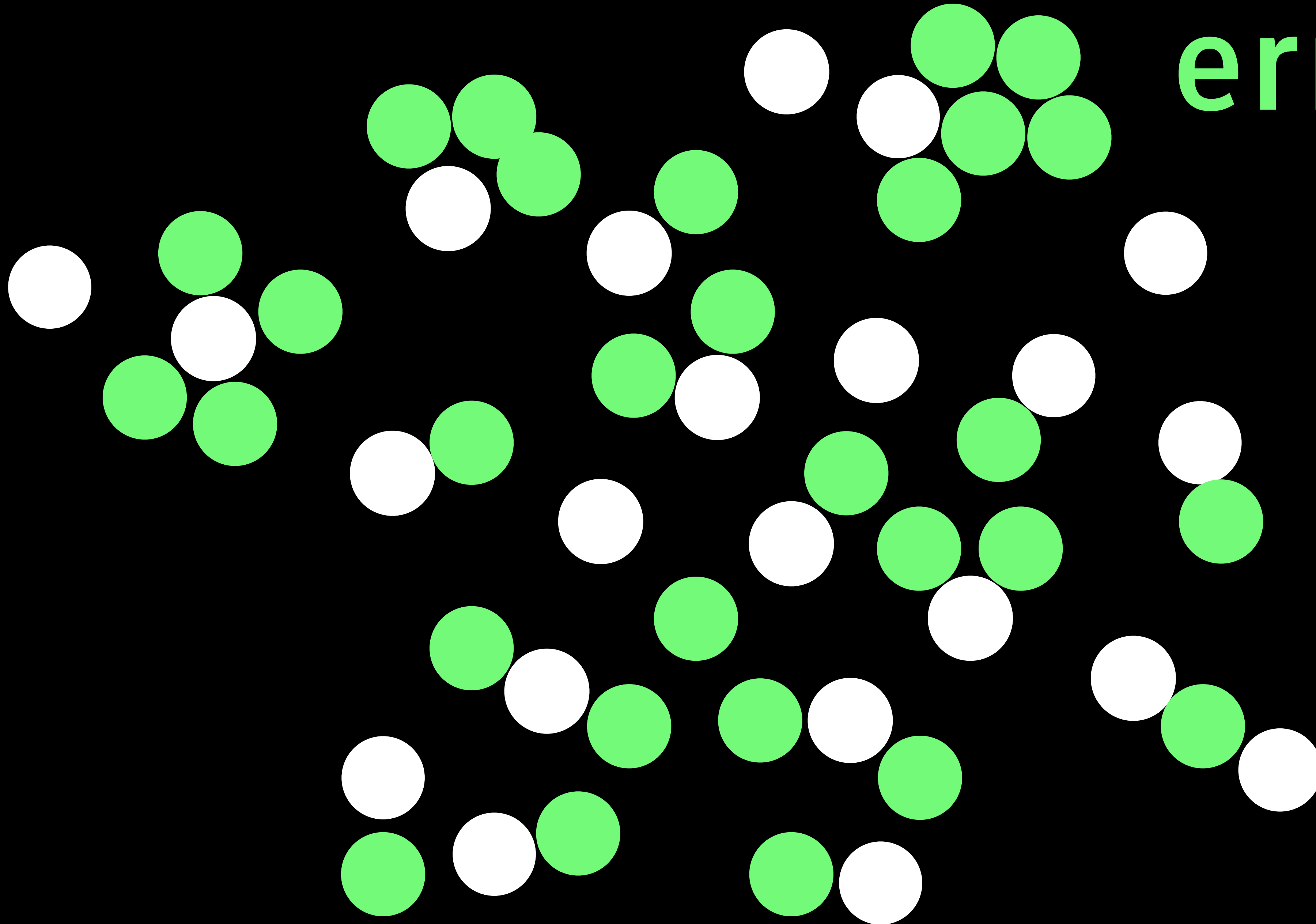
amplicon sequence variants
(ASVs)

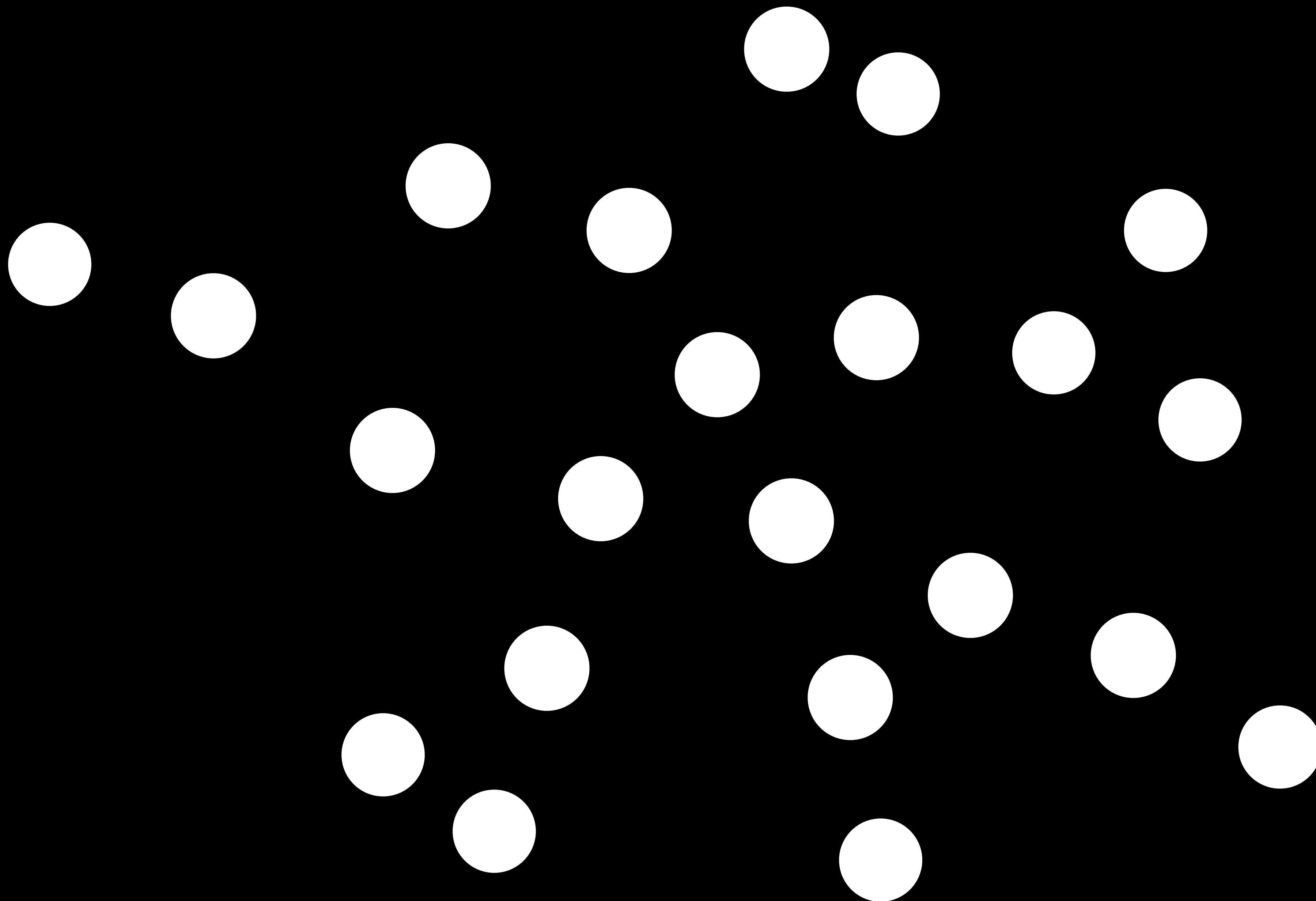
**amplicon sequence variants
(ASVs)**

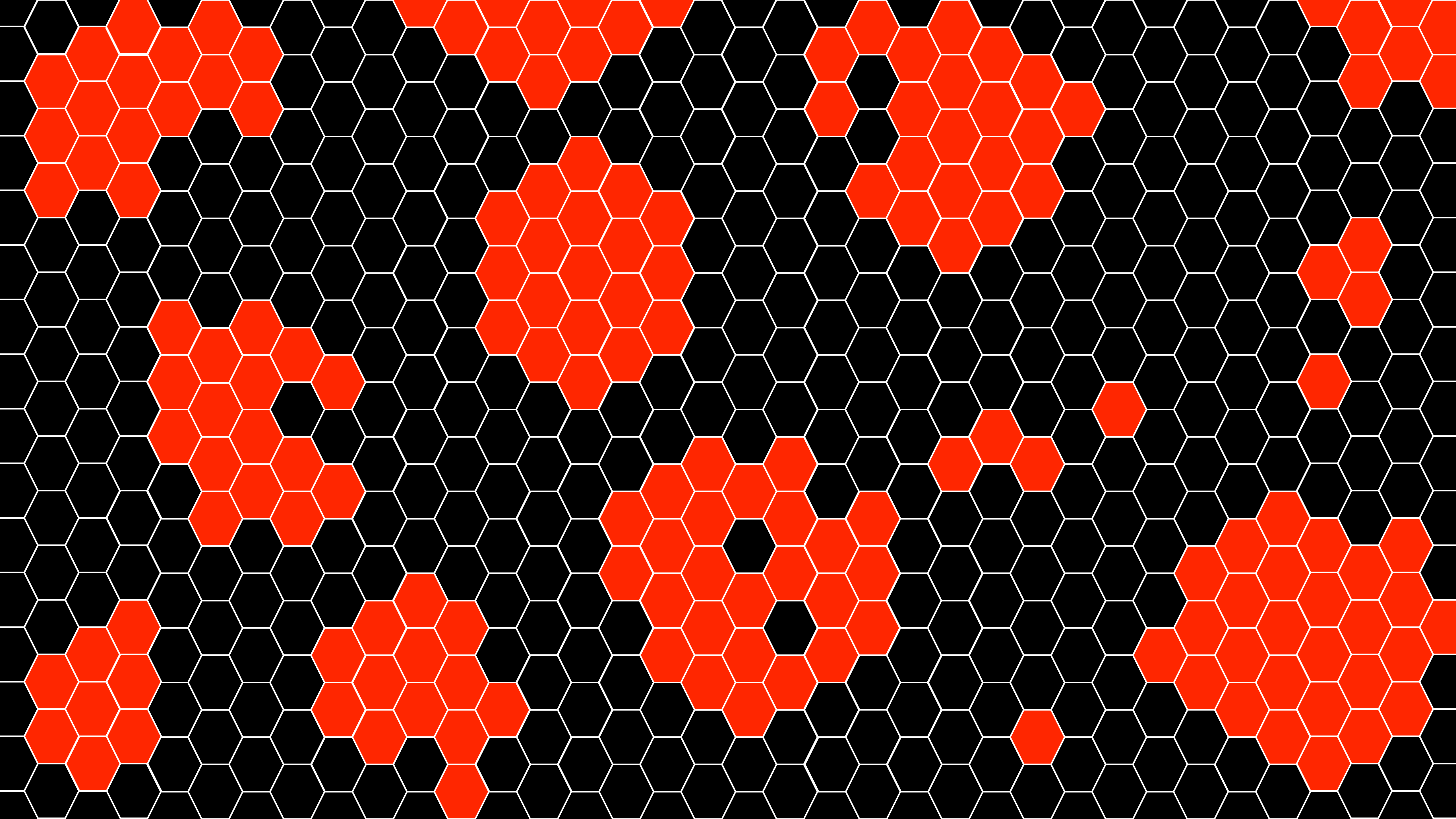
clustering by another name

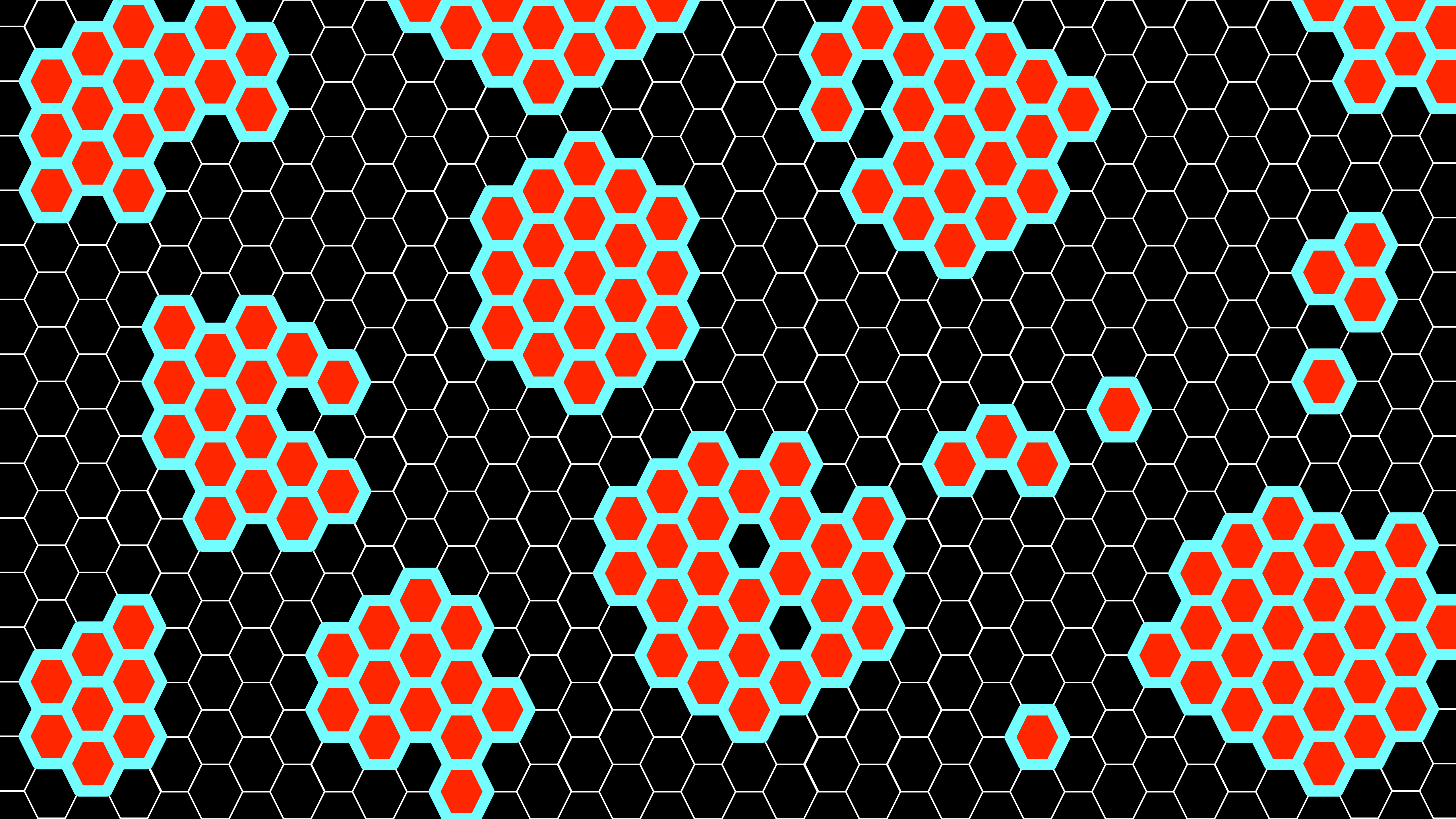


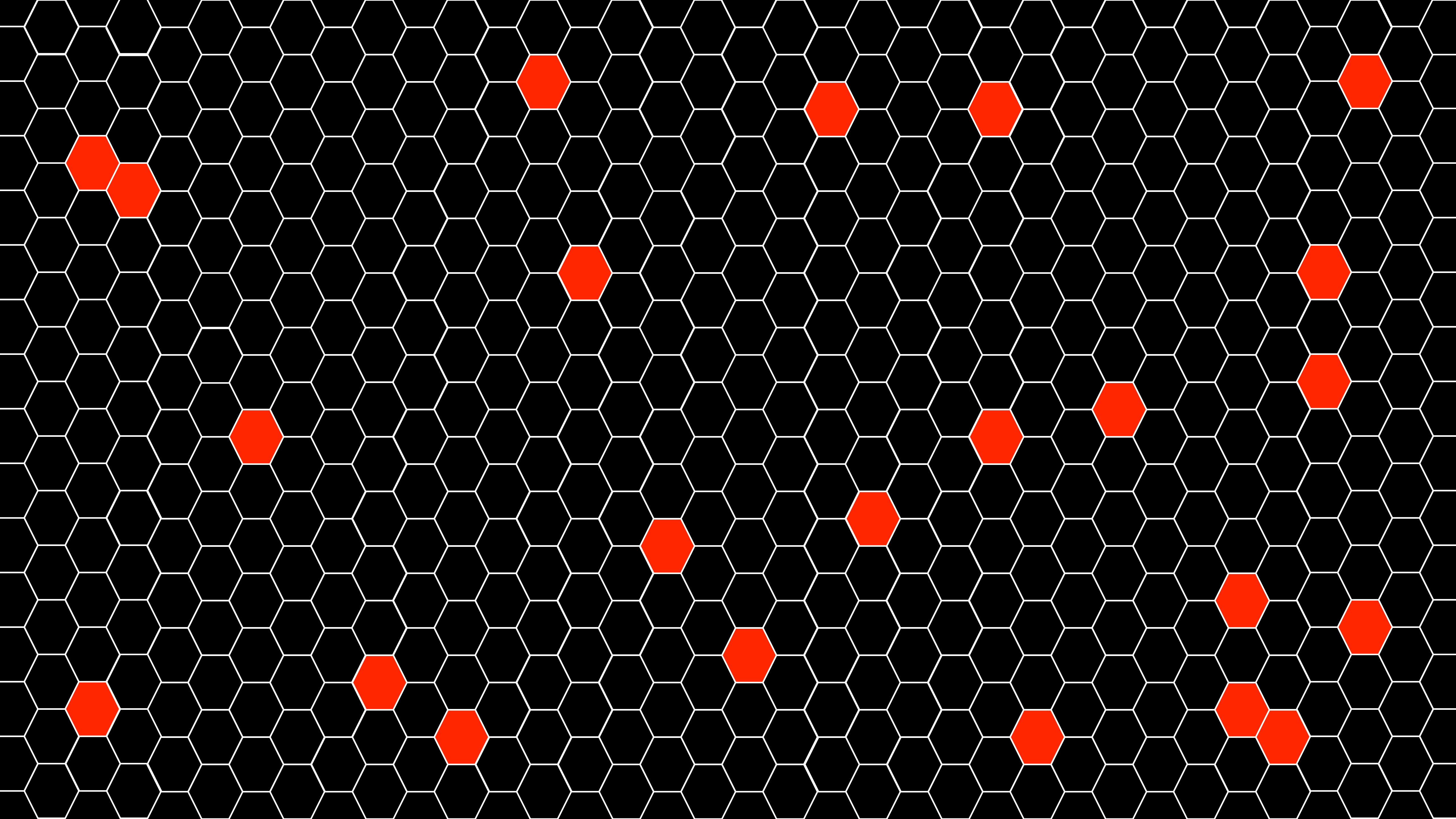
errors

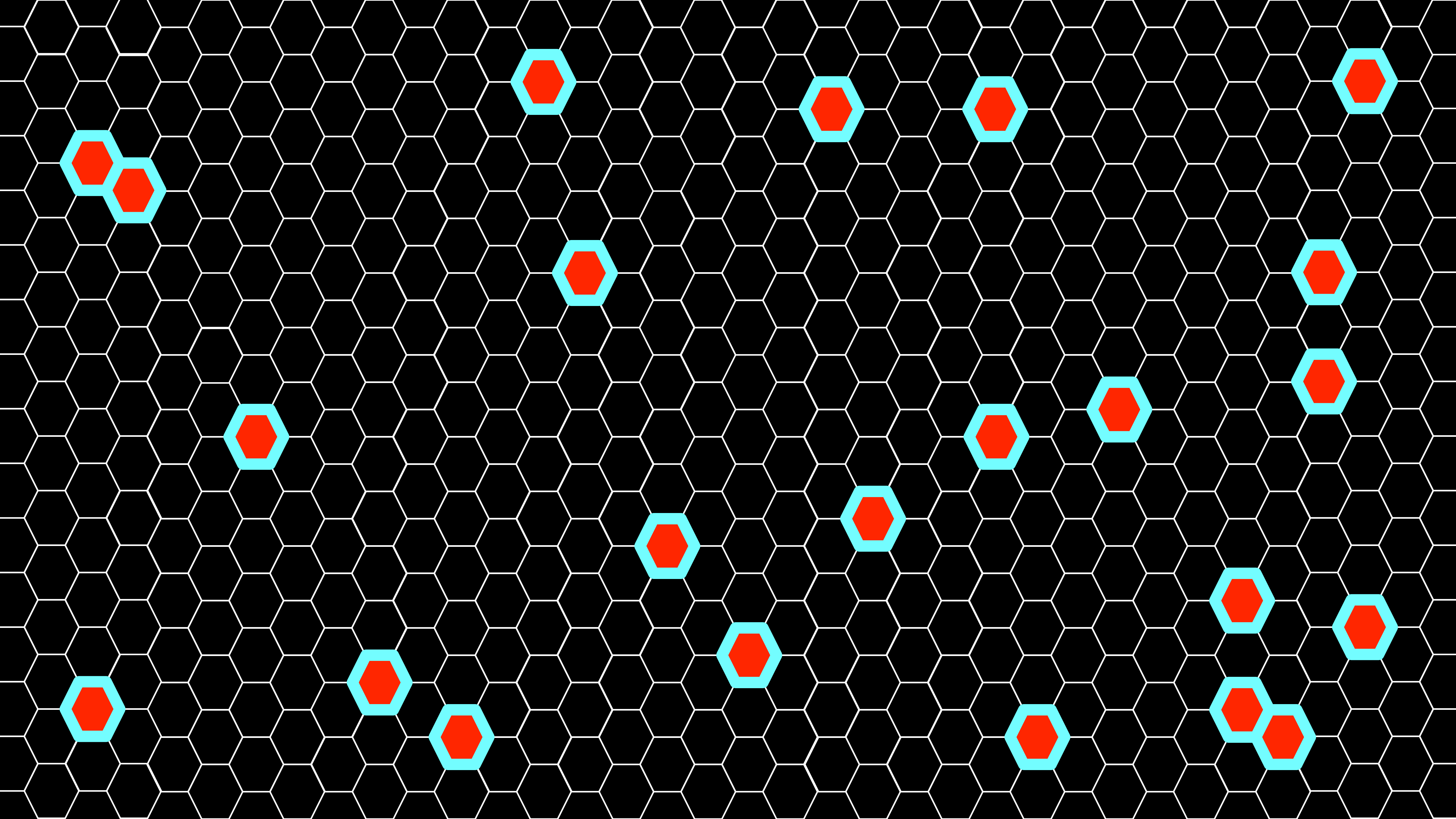












100%

population genetics doesn't matter

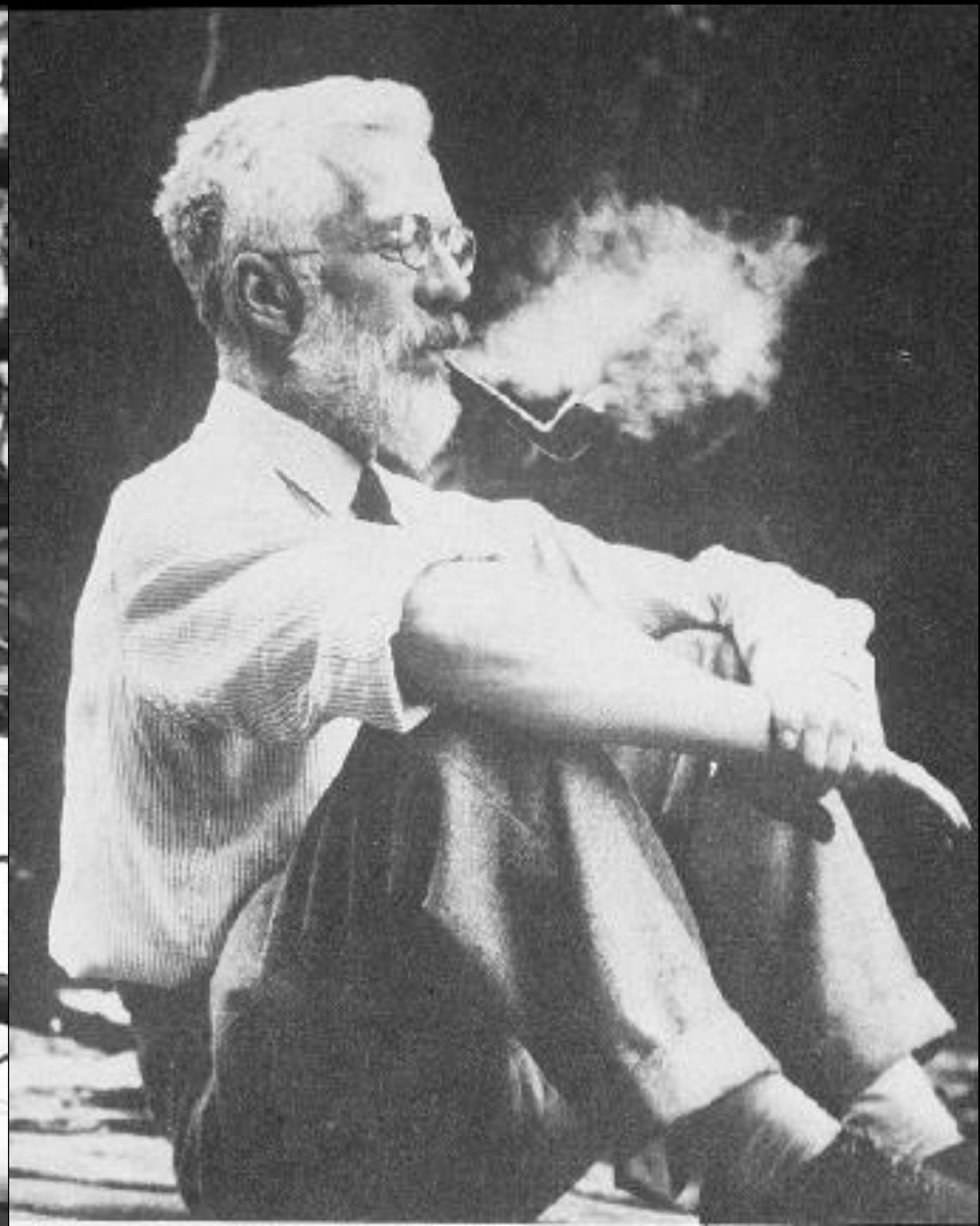
population genetics doesn't matter



Sewall Wright



J.B.S. Haldane



Ronald Fisher

sampling,
storage



nucleotide
extraction



amplification,
sequencing

cleaning,
dereplication



clustering



more
cleaning

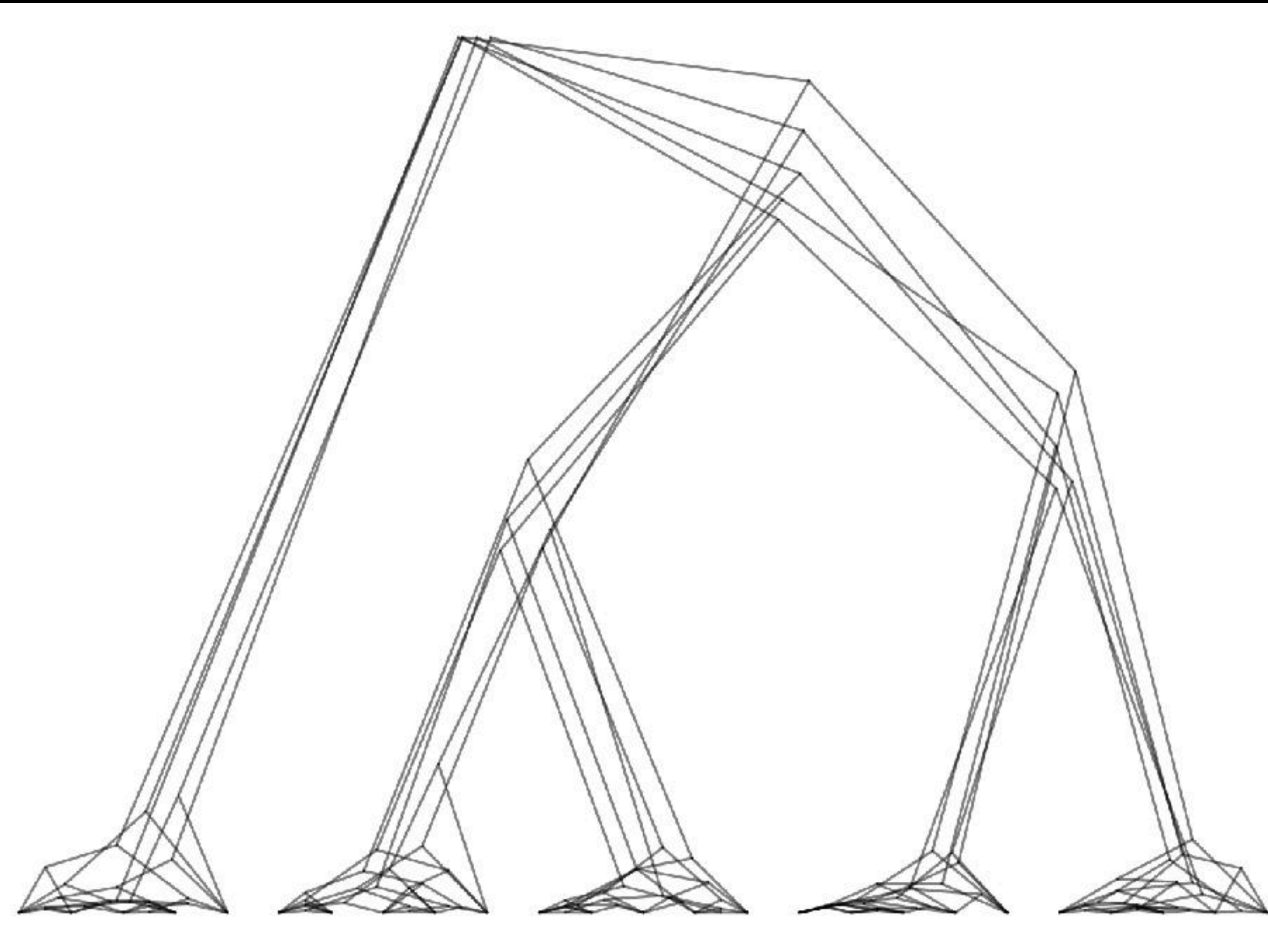
taxonomic
assignments



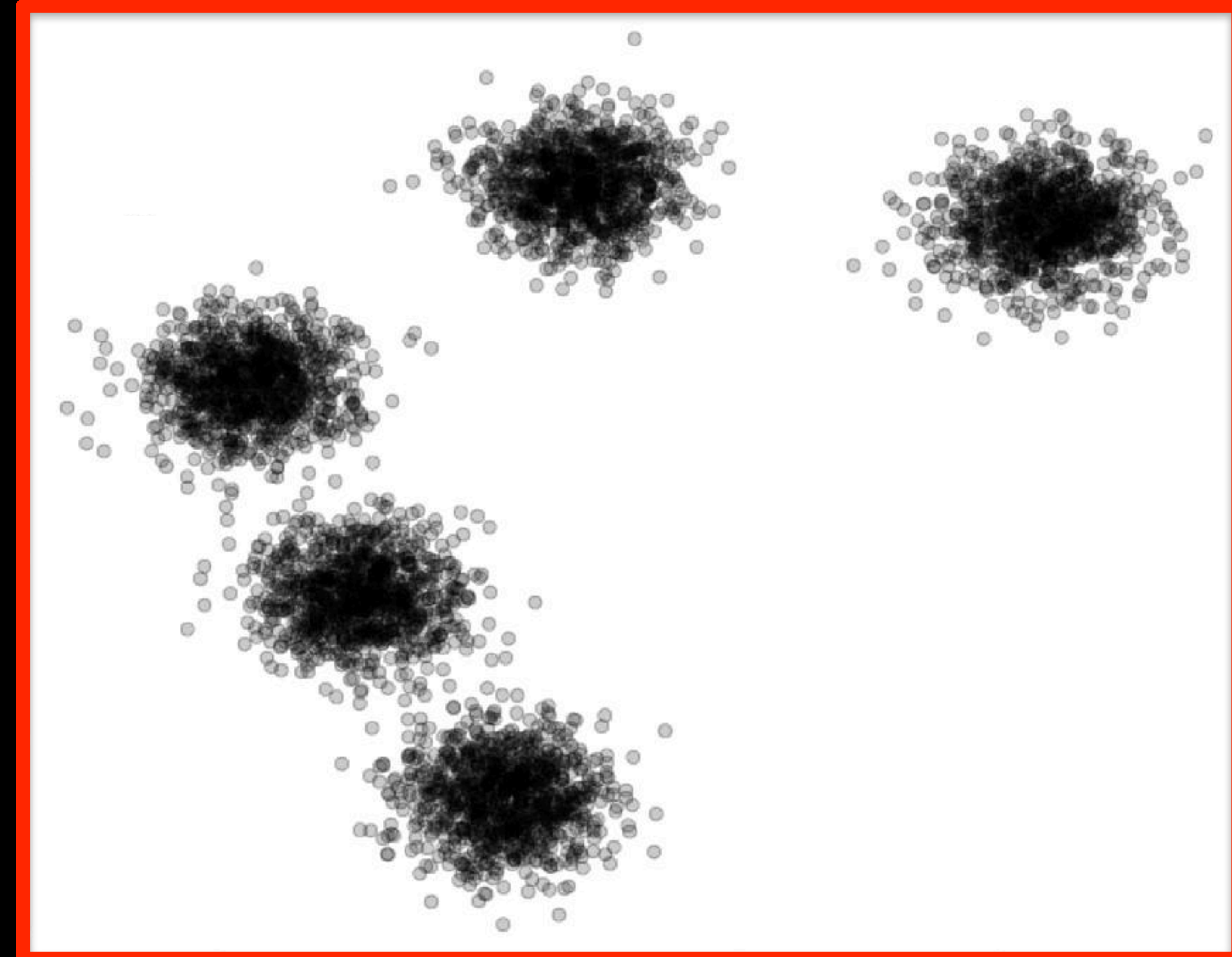
phylogenetic
placements



α, β diversities,
co-occurrences



trait 2



trait 1

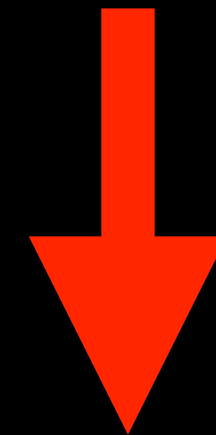
**pairwise
comparisons**

**phylogenetic
comparisons**

**global clustering
thresholds**

**VSEARCH/
Mothur**

DADA2



**local clustering
thresholds**

Swarm

**multi-rate
PTP**

Bioinformatics, 38(1), 2022, 267–269

doi: 10.1093/bioinformatics/btab493

Advance Access Publication Date: 9 July 2021

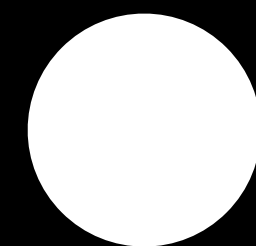
Applications Note

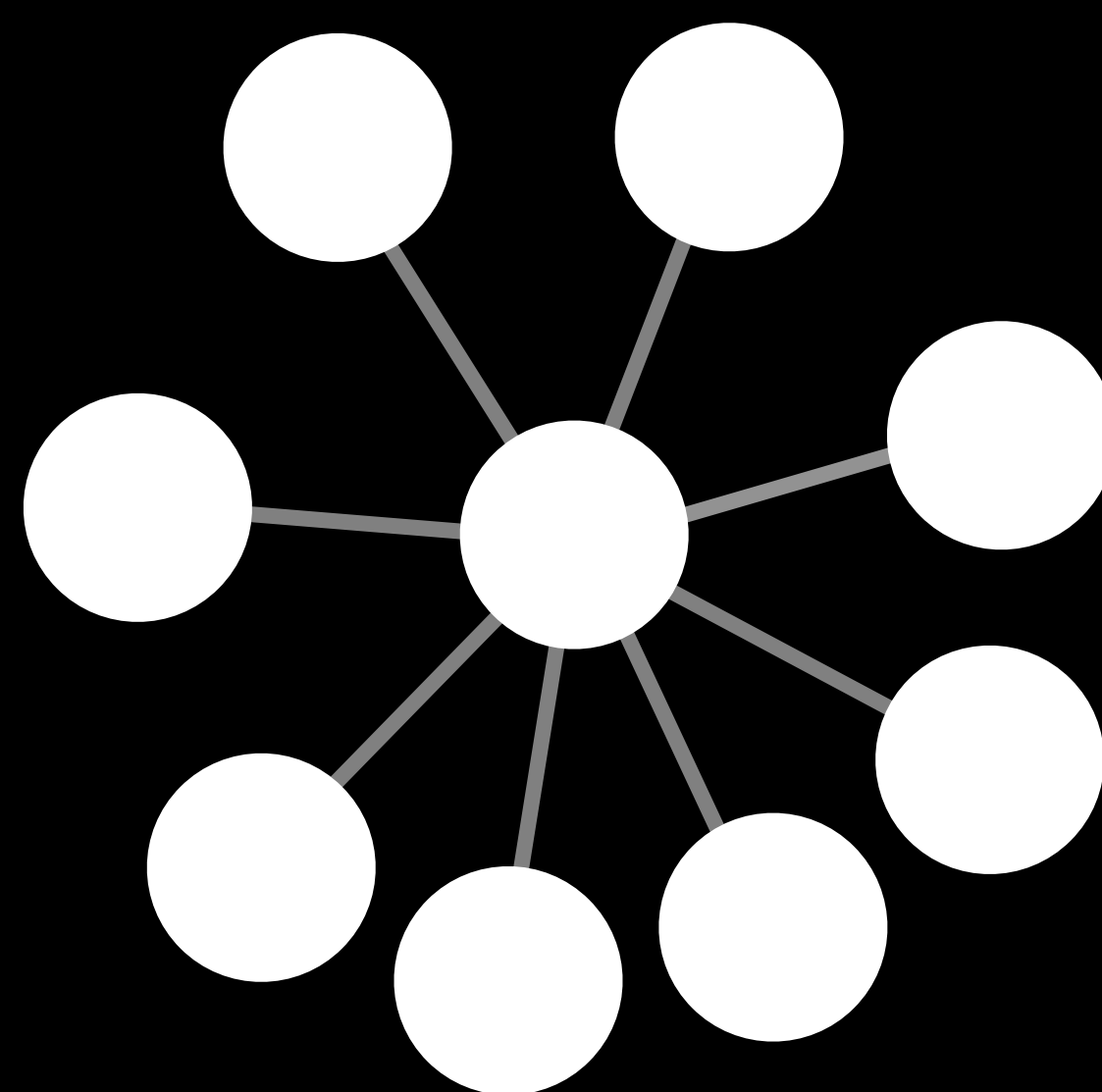
OXFORD

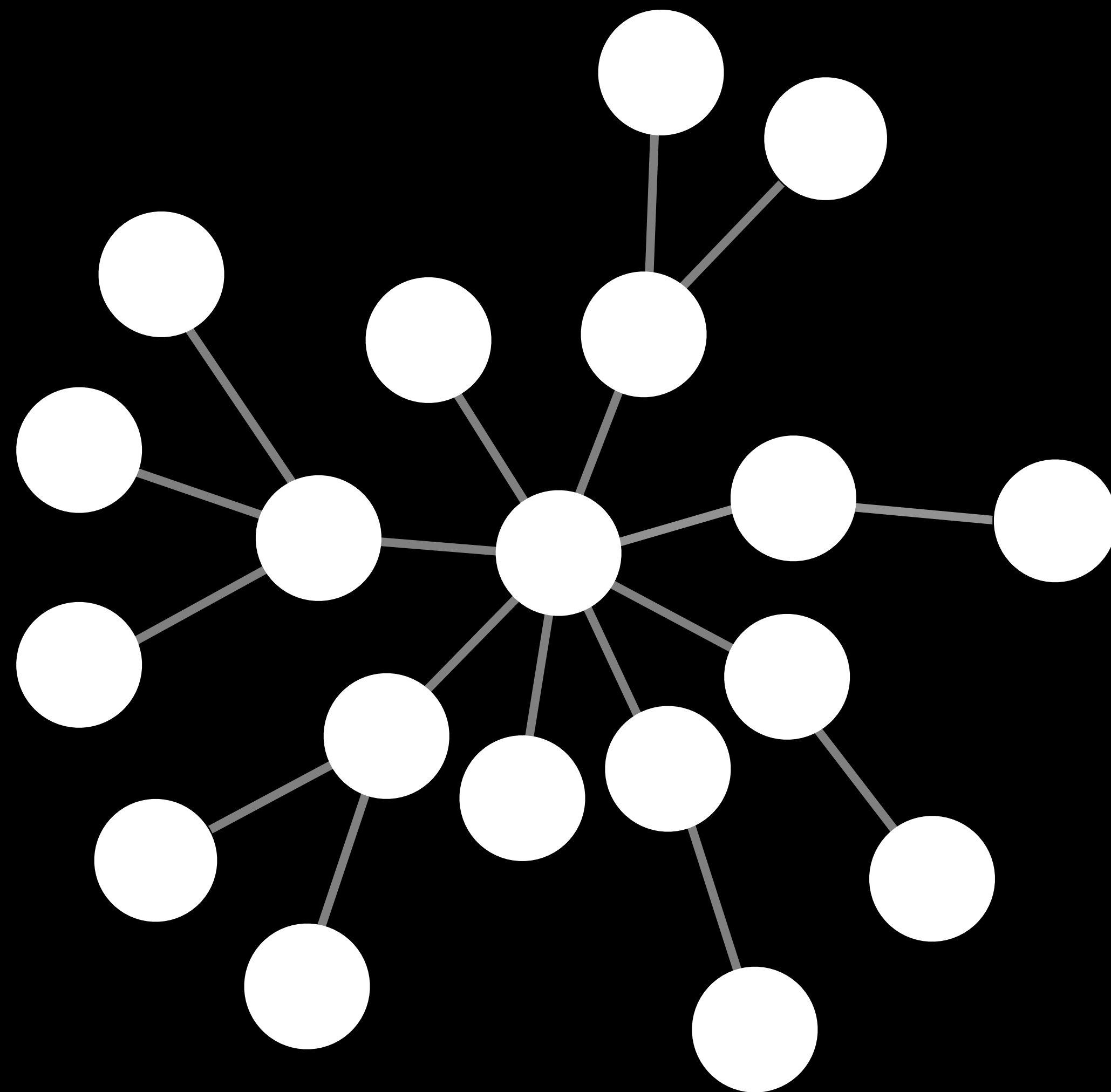
Sequence analysis

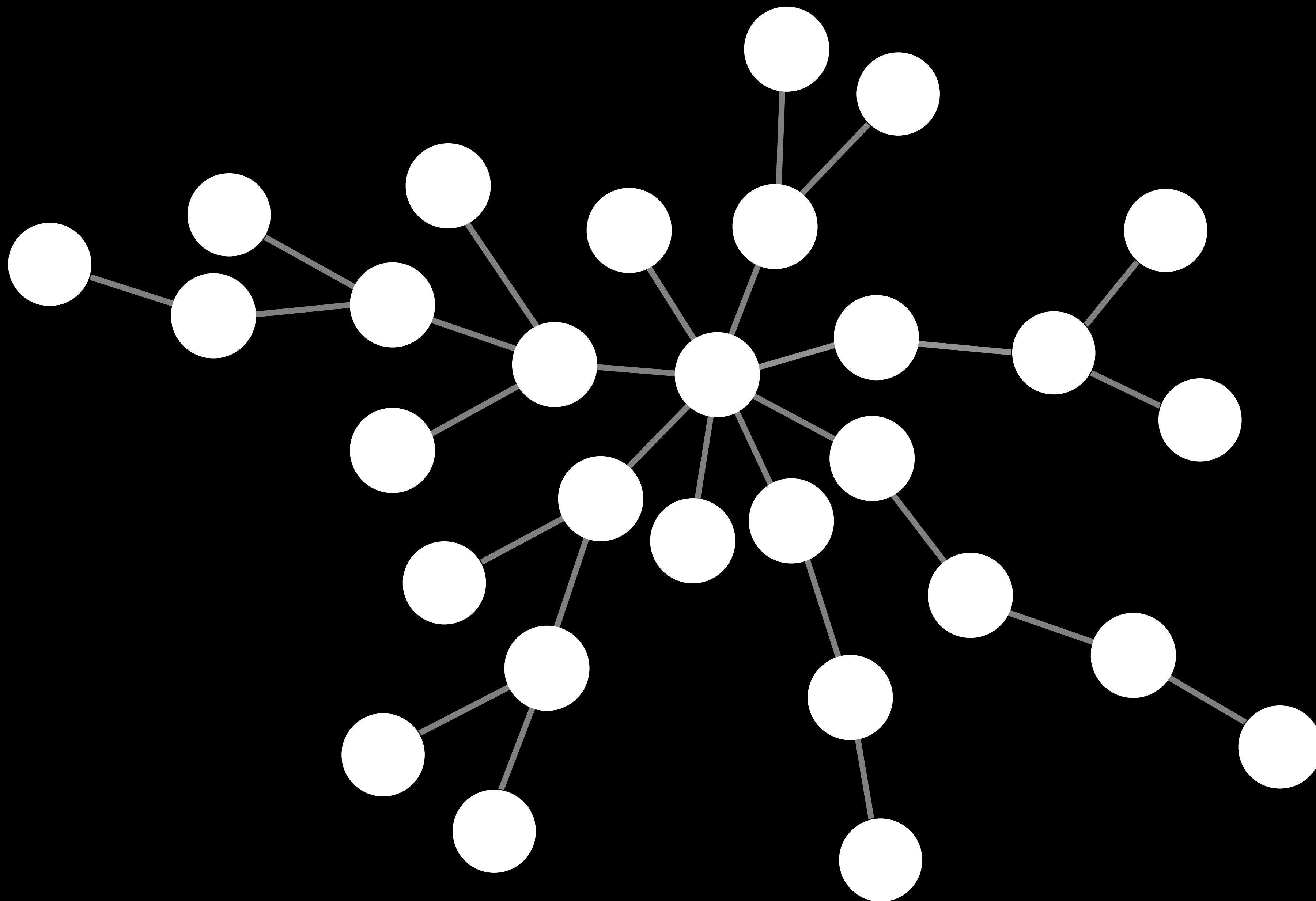
Swarm v3: towards tera-scale amplicon clustering

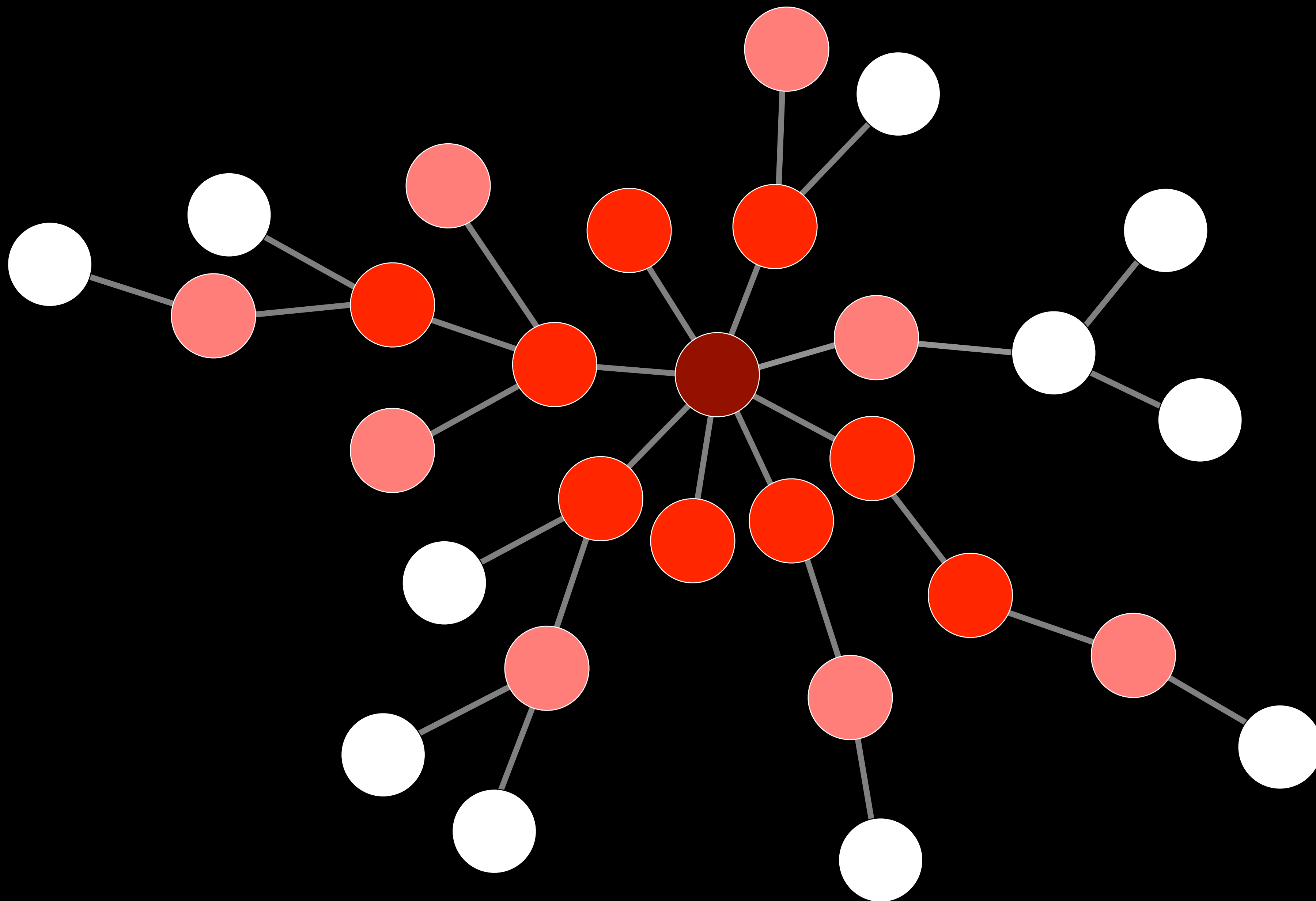
Frédéric Mahé^{1,2,*}, Lucas Czech ^{3,4}, Alexandros Stamatakis^{3,5},
Christopher Quince^{6,7,8}, Colomban de Vargas^{9,10}, Micah Dunthorn ^{11,12} and
Torbjørn Rognes^{13,14}

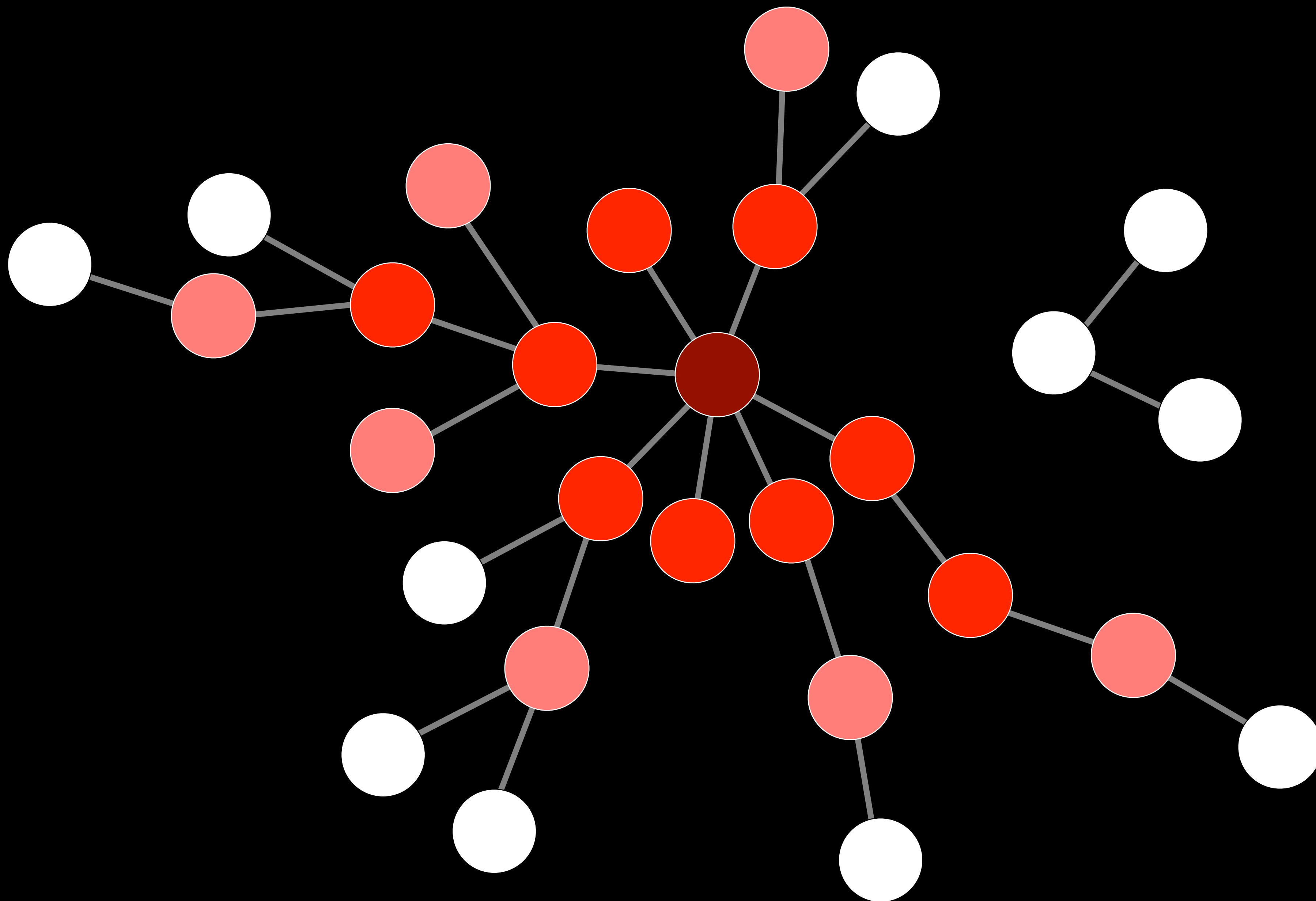


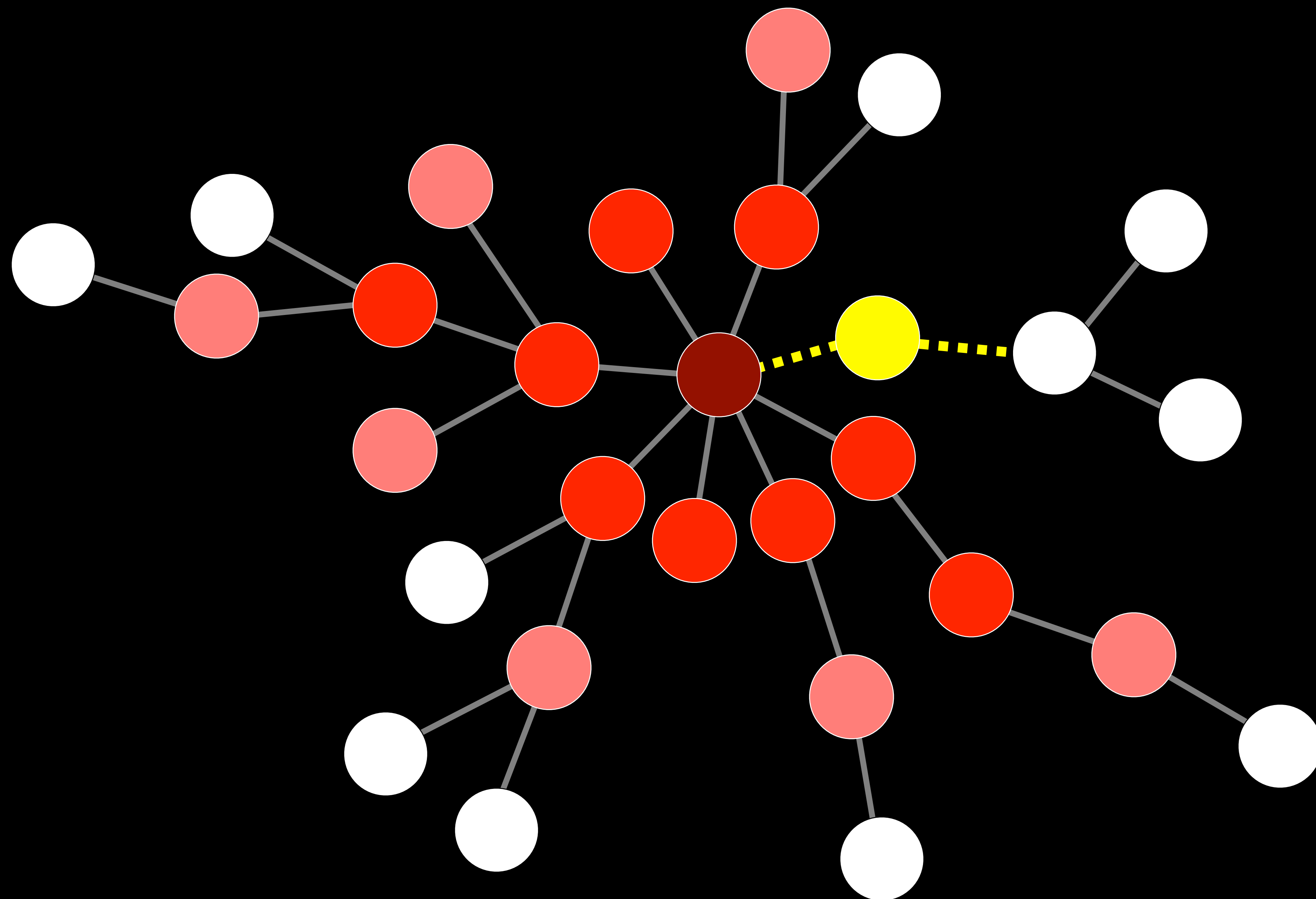


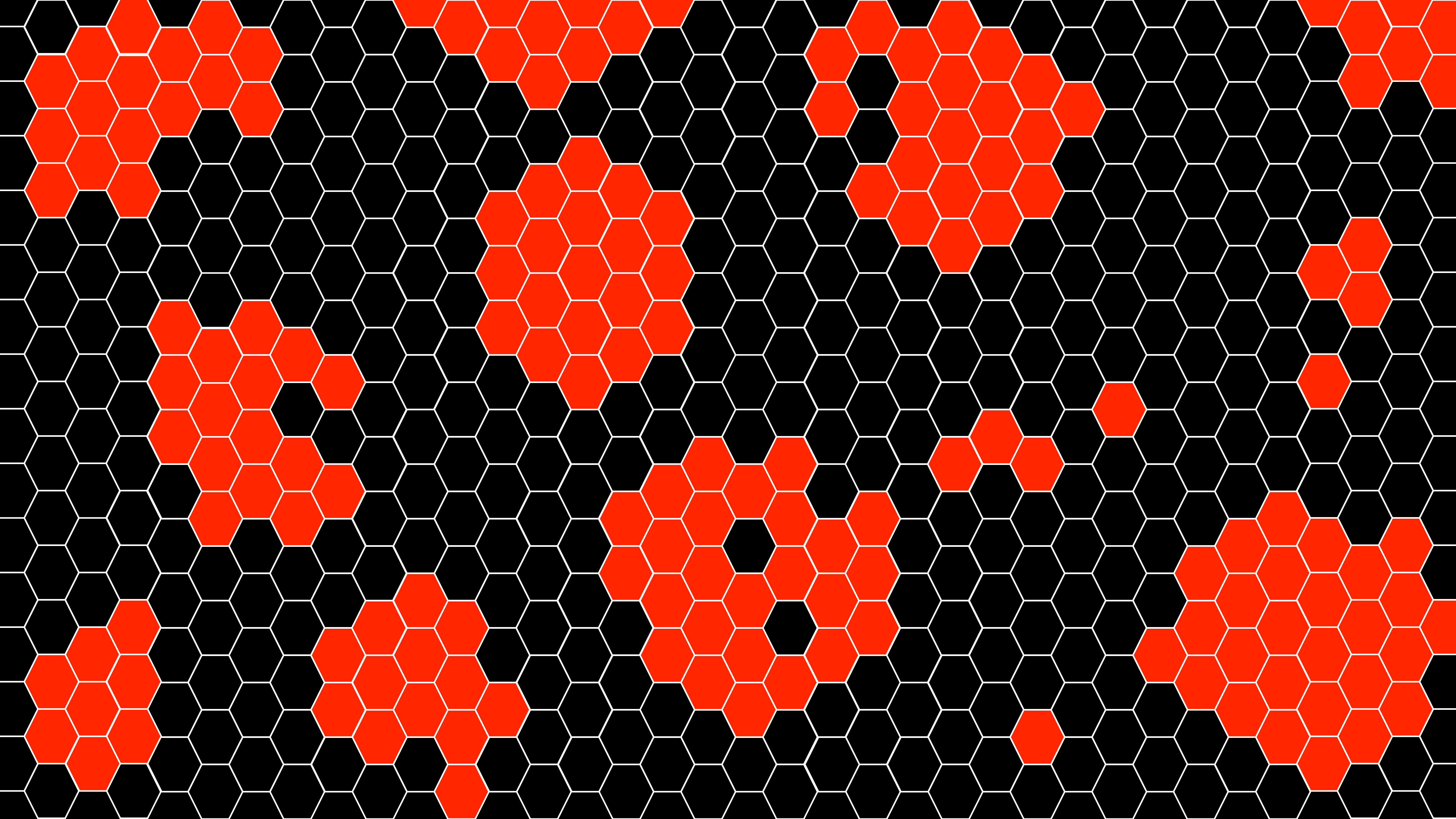


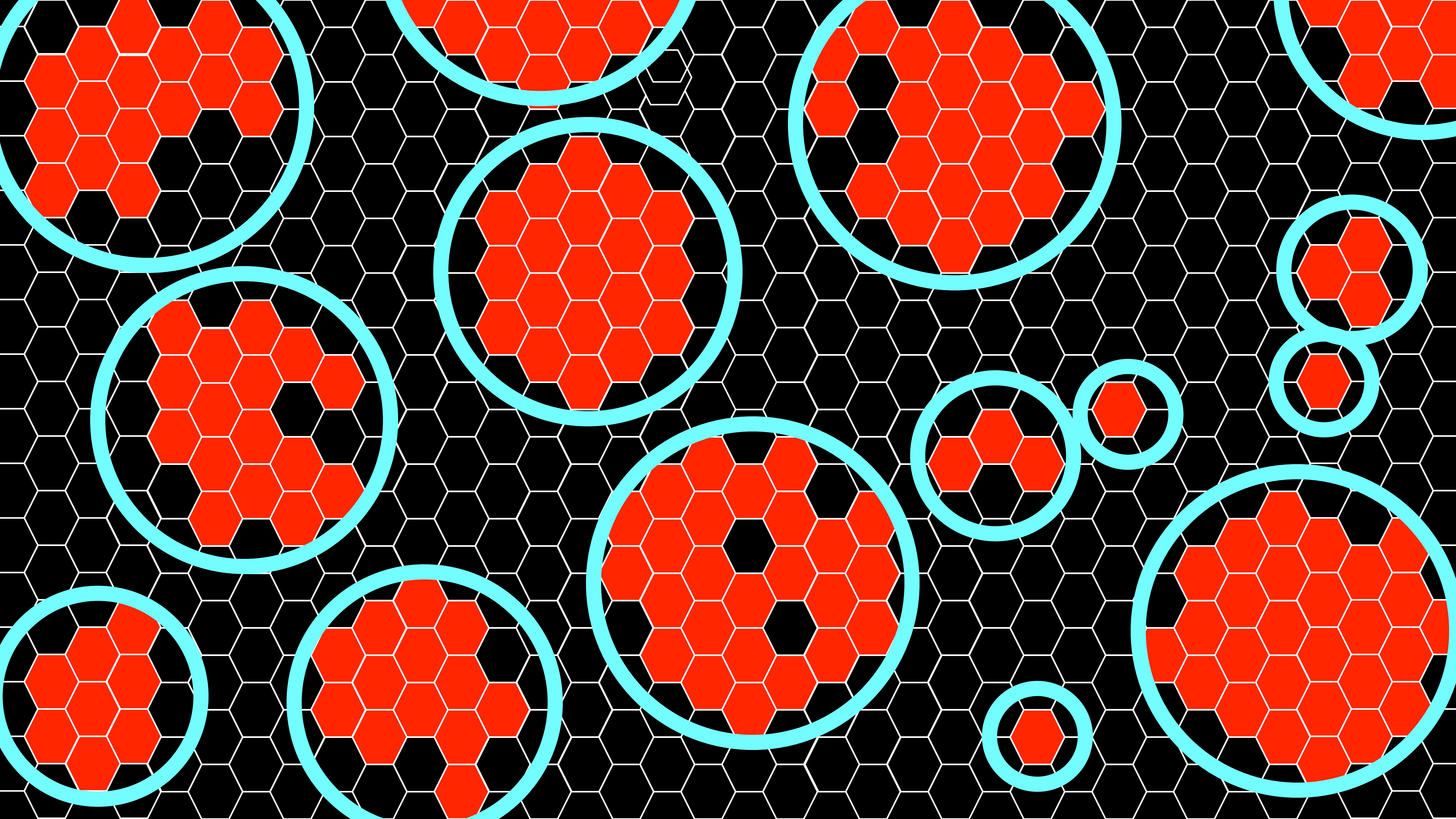


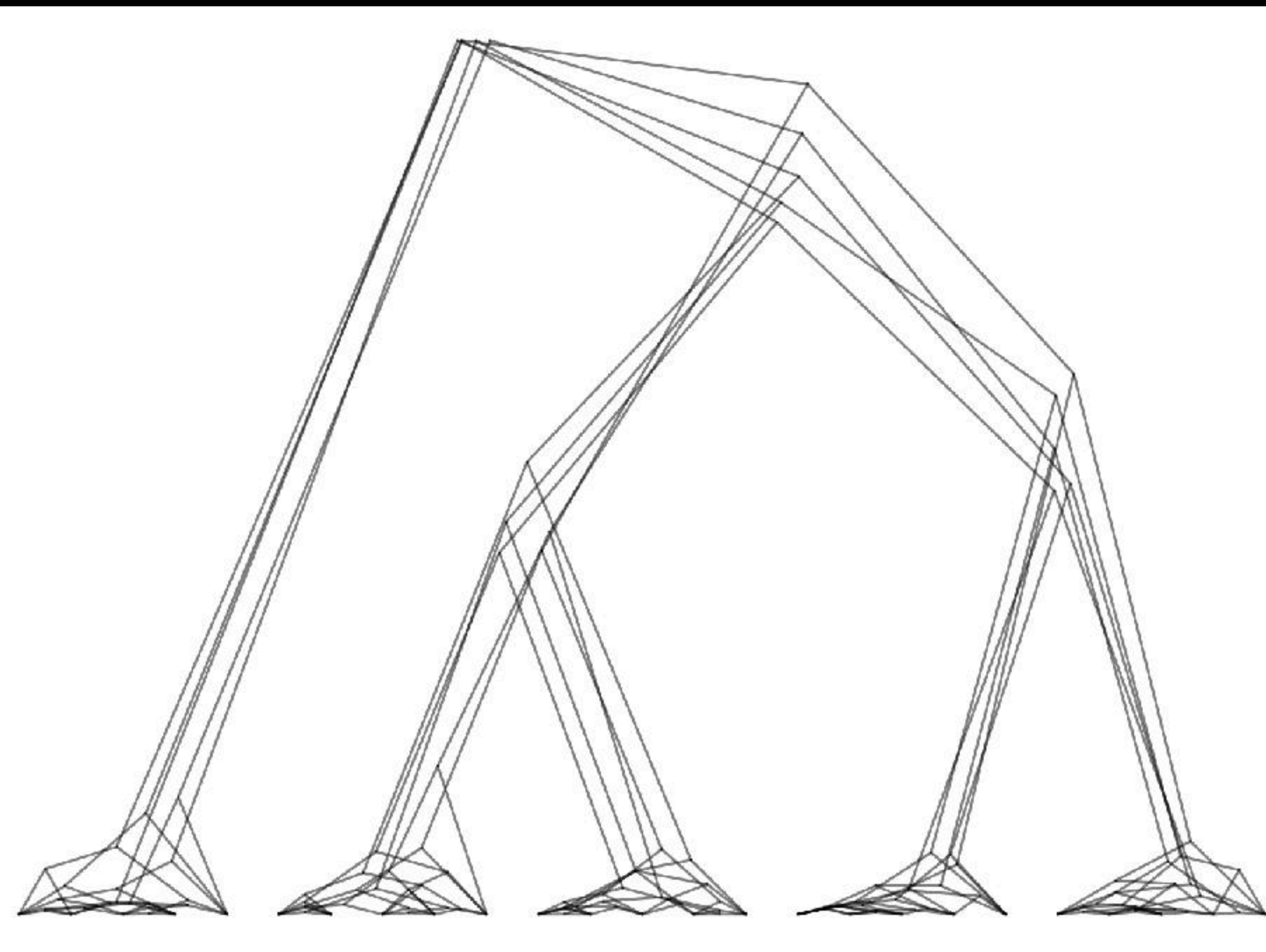




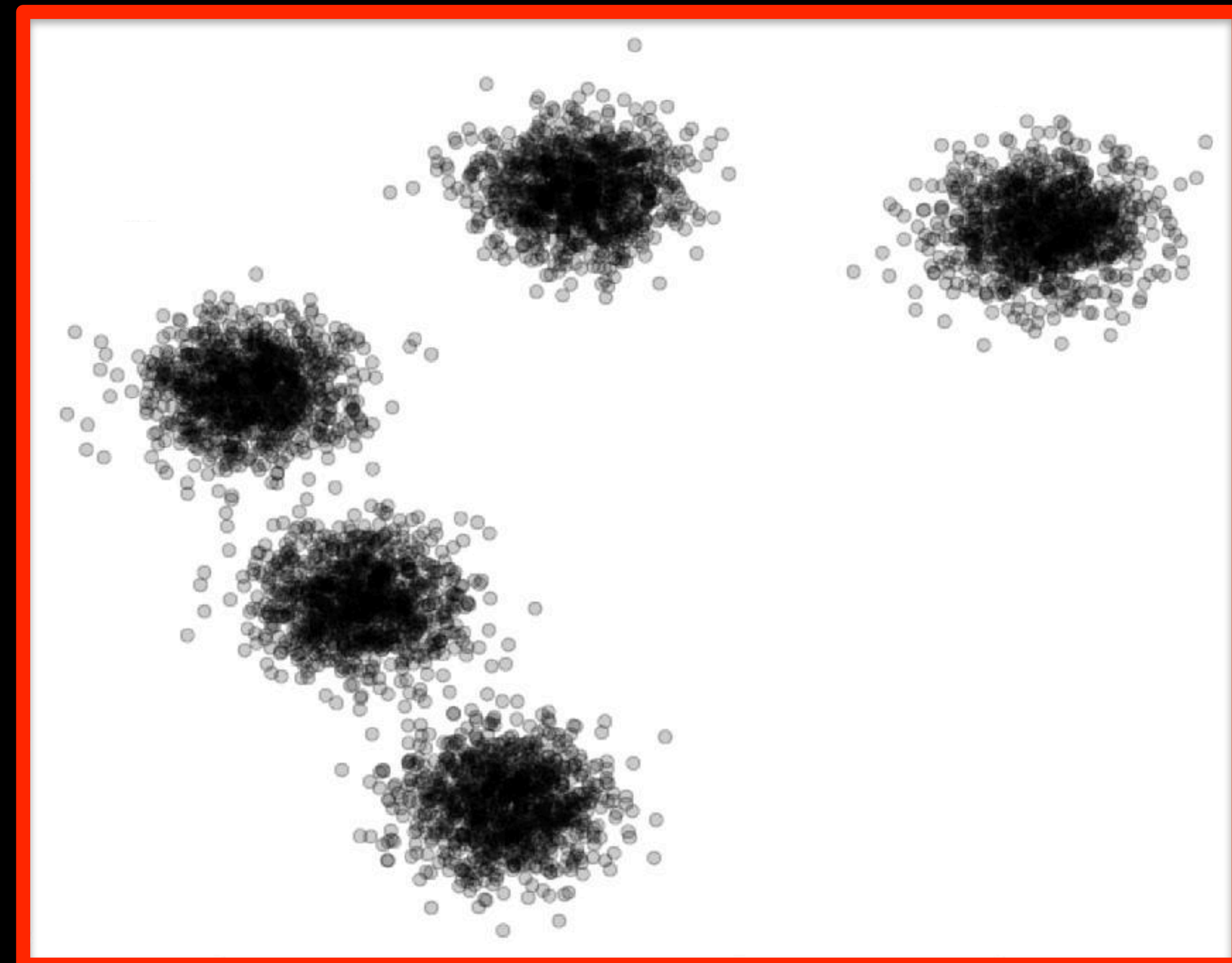








trait 2



trait 1

**pairwise
comparisons**

**phylogenetic
comparisons**

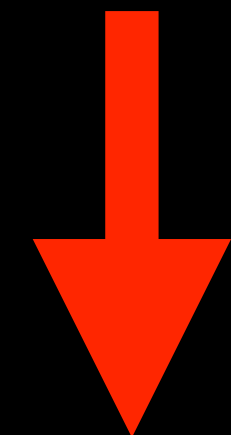
**global clustering
thresholds**

**VSEARCH/
Mothur**

DADA2

**local clustering
thresholds**

Swarm



**multi-rate
PTP**

Bioinformatics, 33(11), 2017, 1630–1638

doi: 10.1093/bioinformatics/btx025

Advance Access Publication Date: 20 January 2017

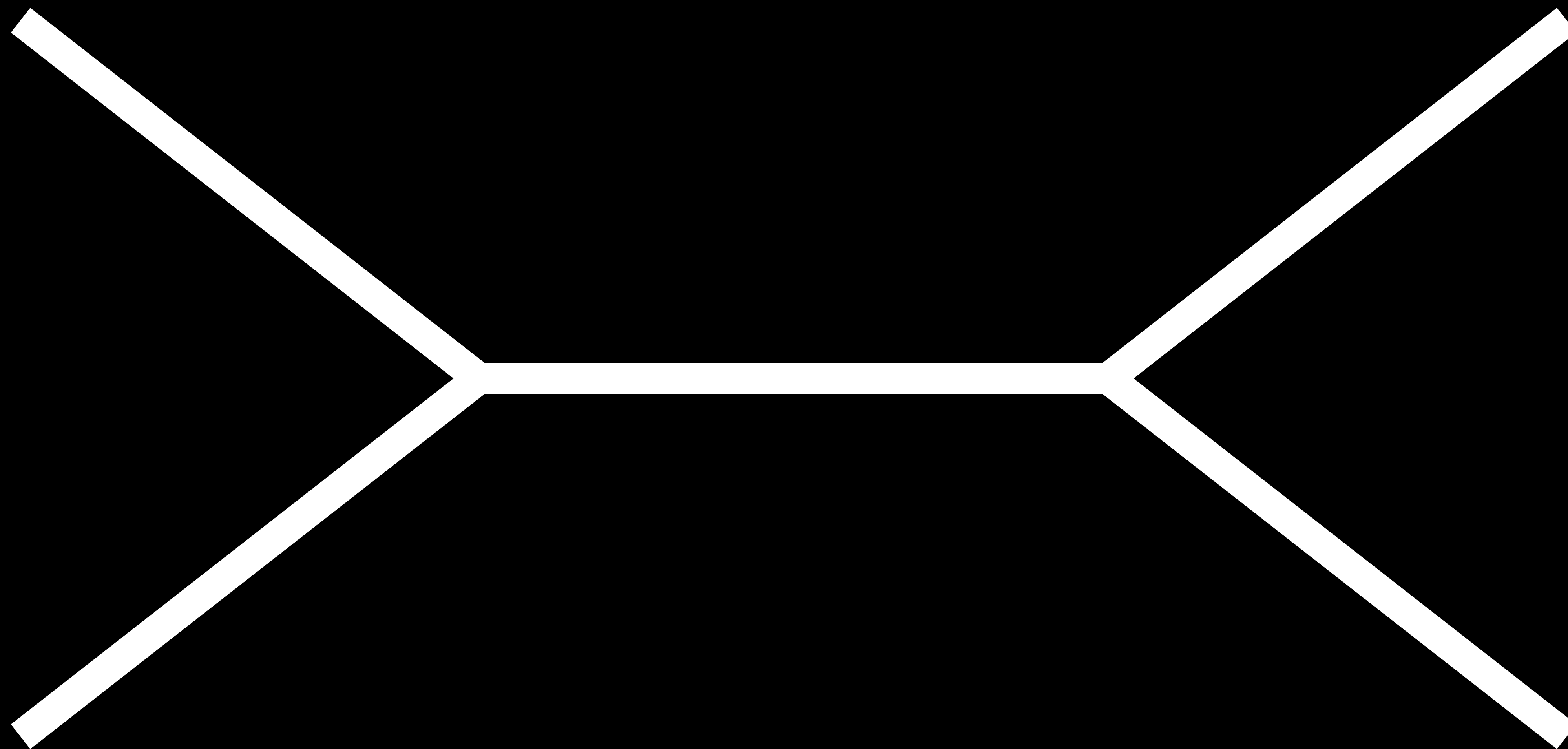
Original Paper

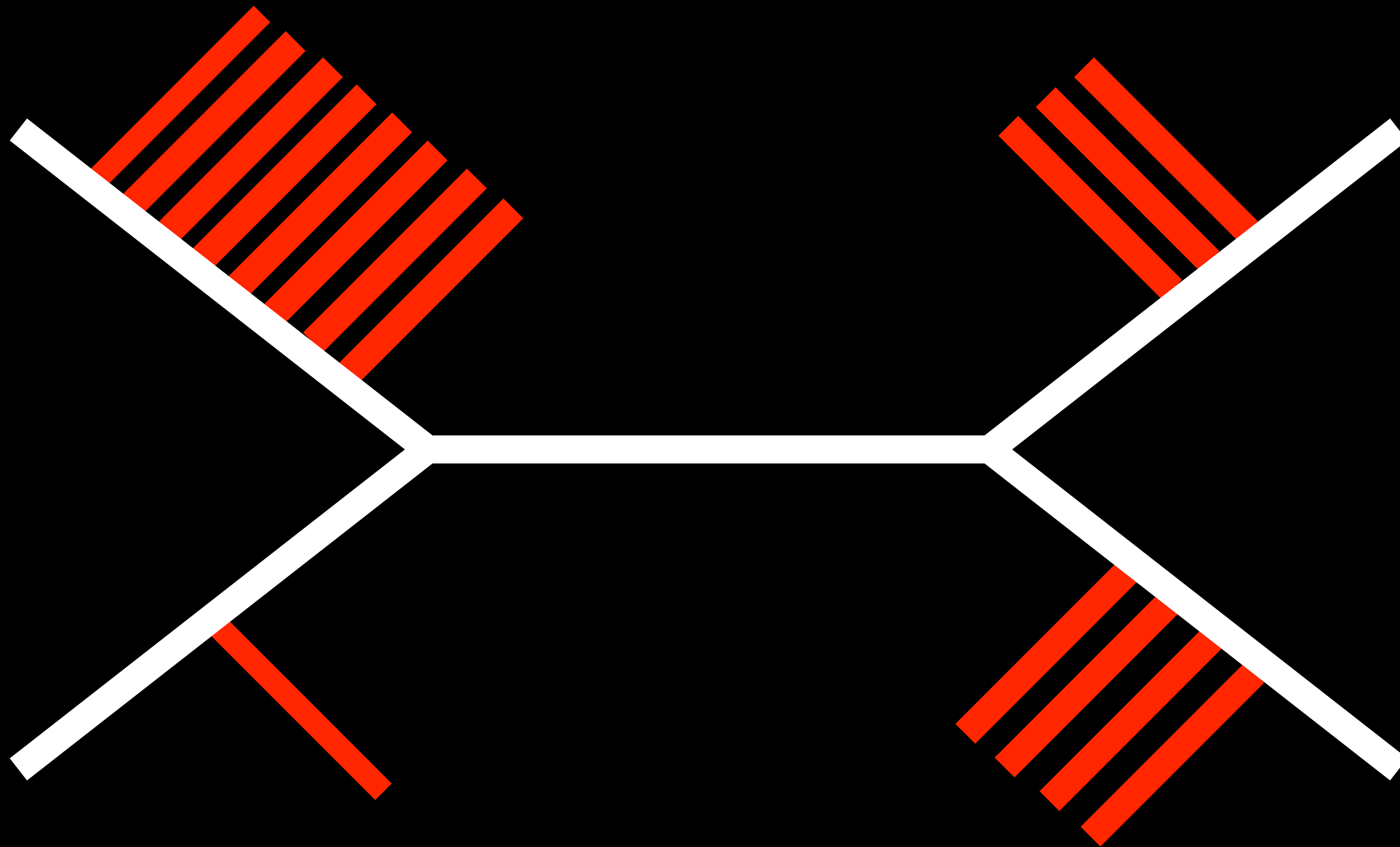
OXFORD

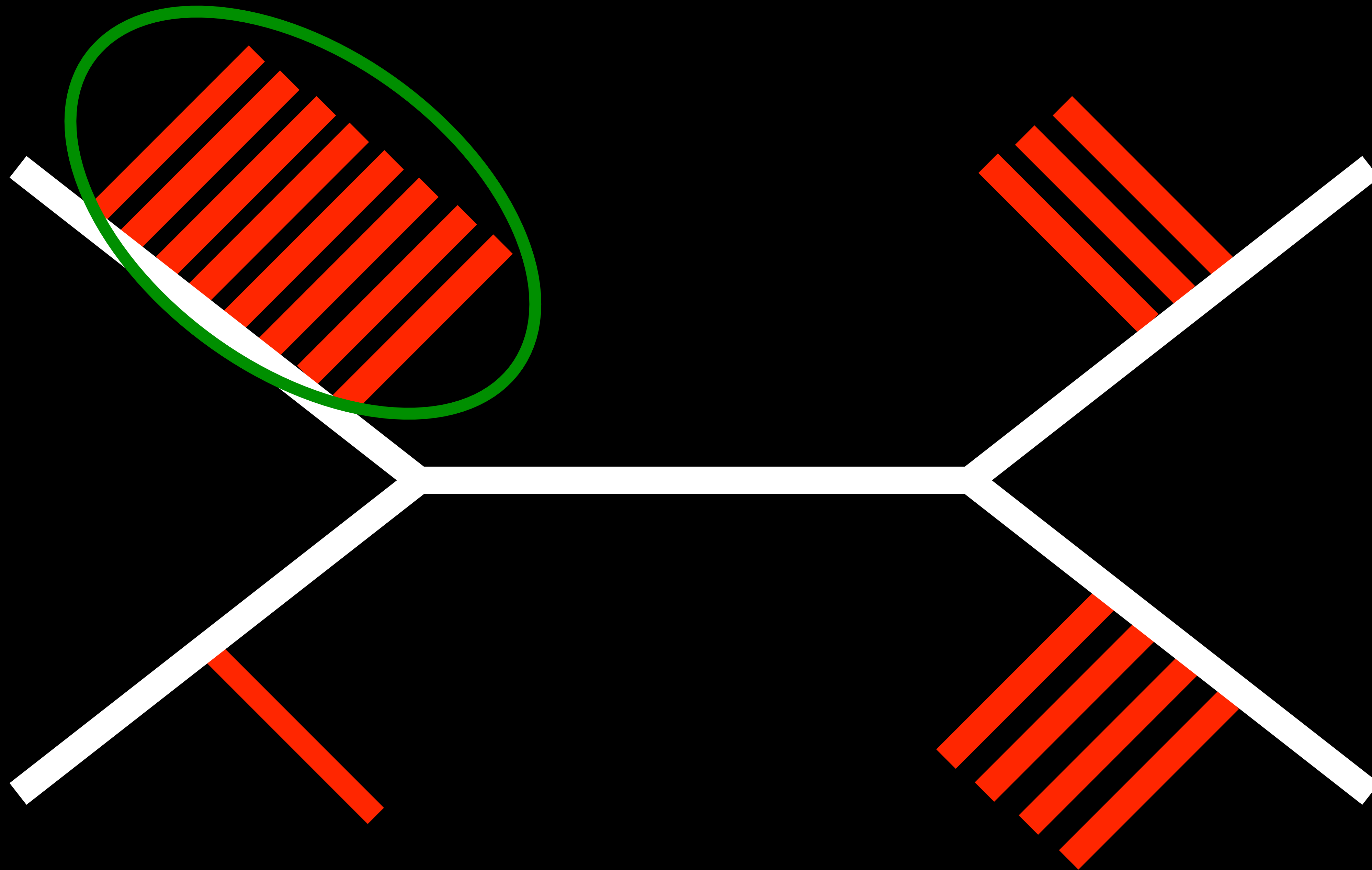
Phylogenetics

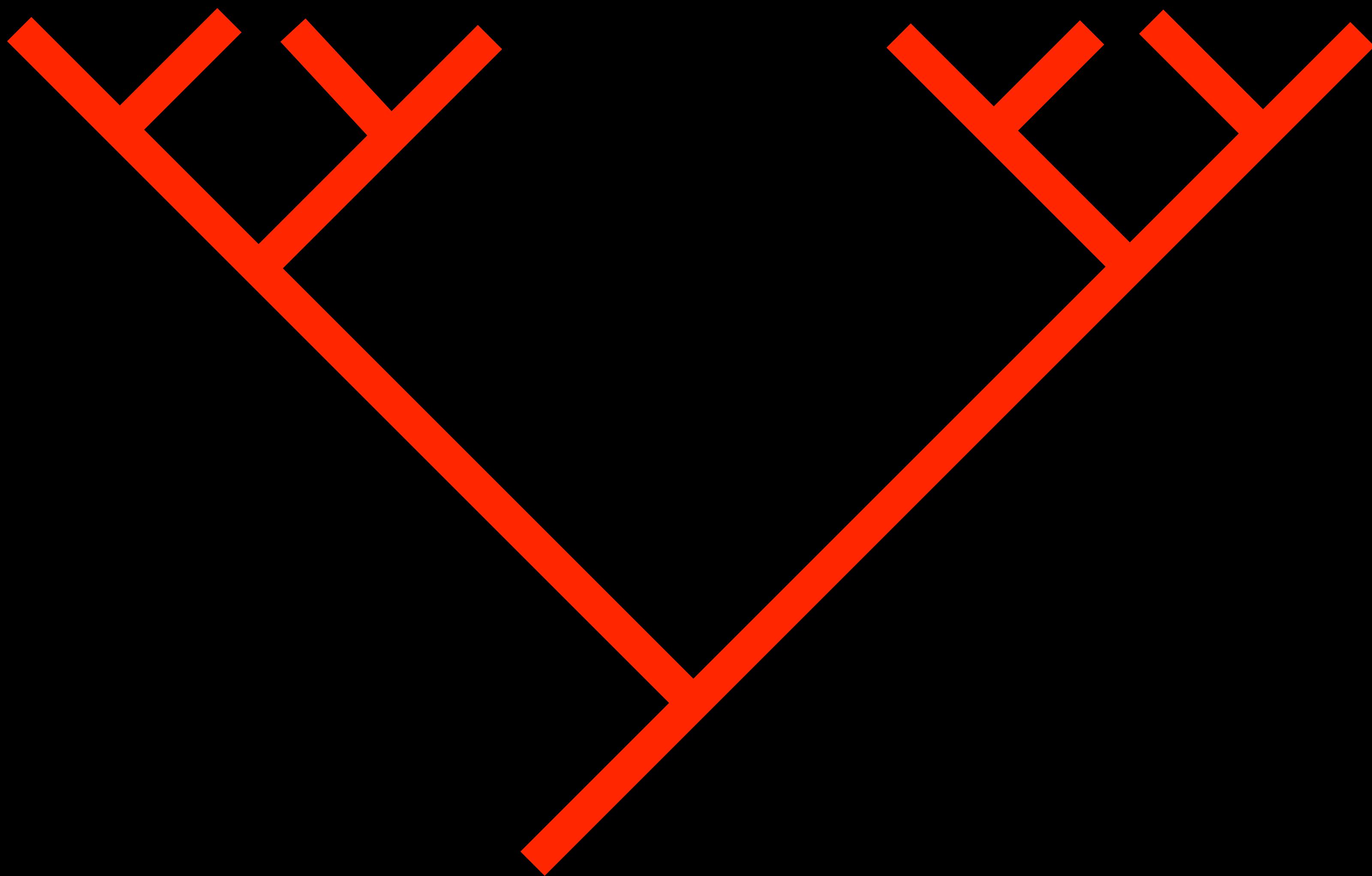
Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo

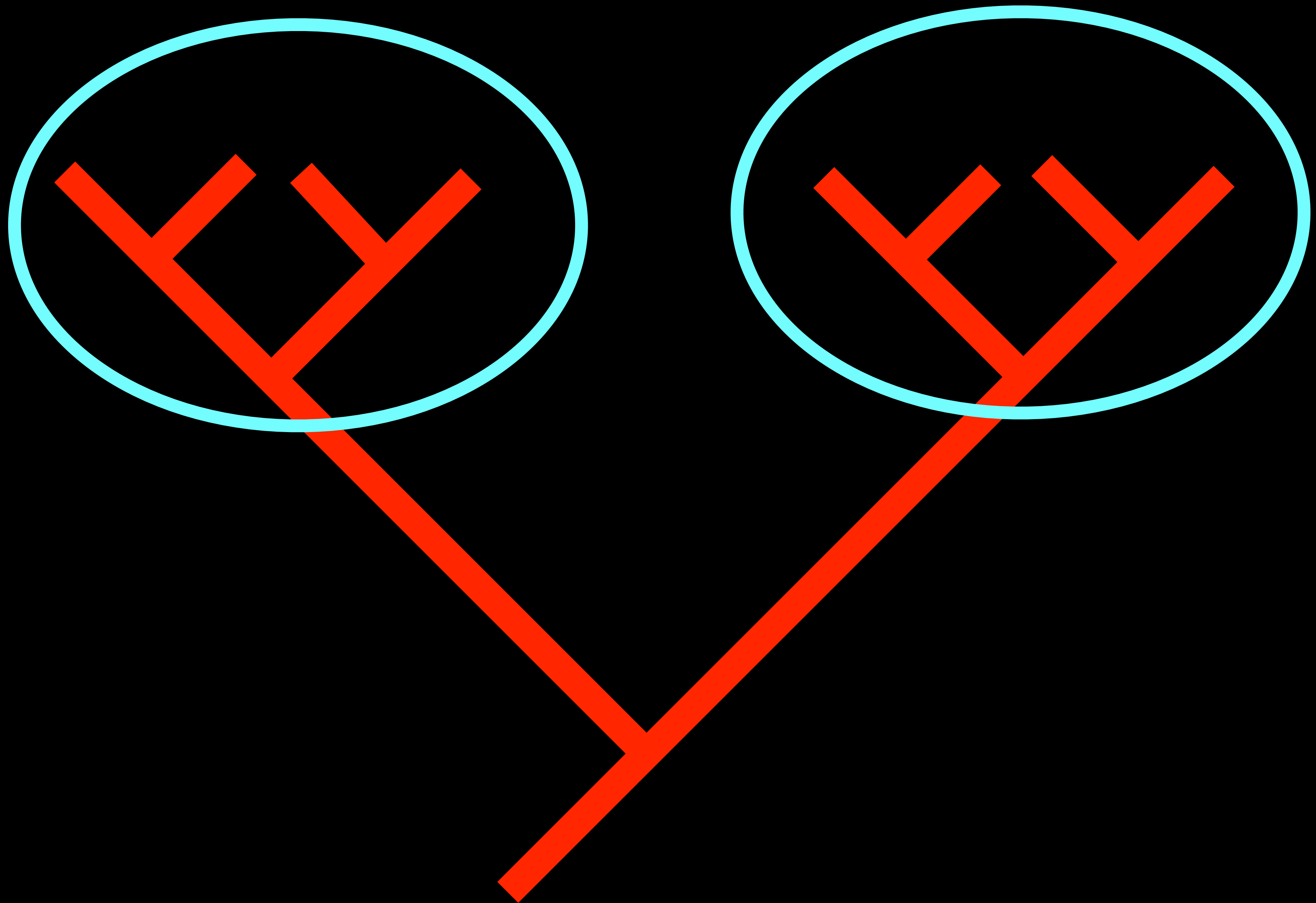
**P. Kapli^{1,*}, S. Lutteropp^{1,2}, J. Zhang¹, K. Kobert¹, P. Pavlidis³,
A. Stamatakis^{1,2,*} and T. Flouri^{1,2,*}**

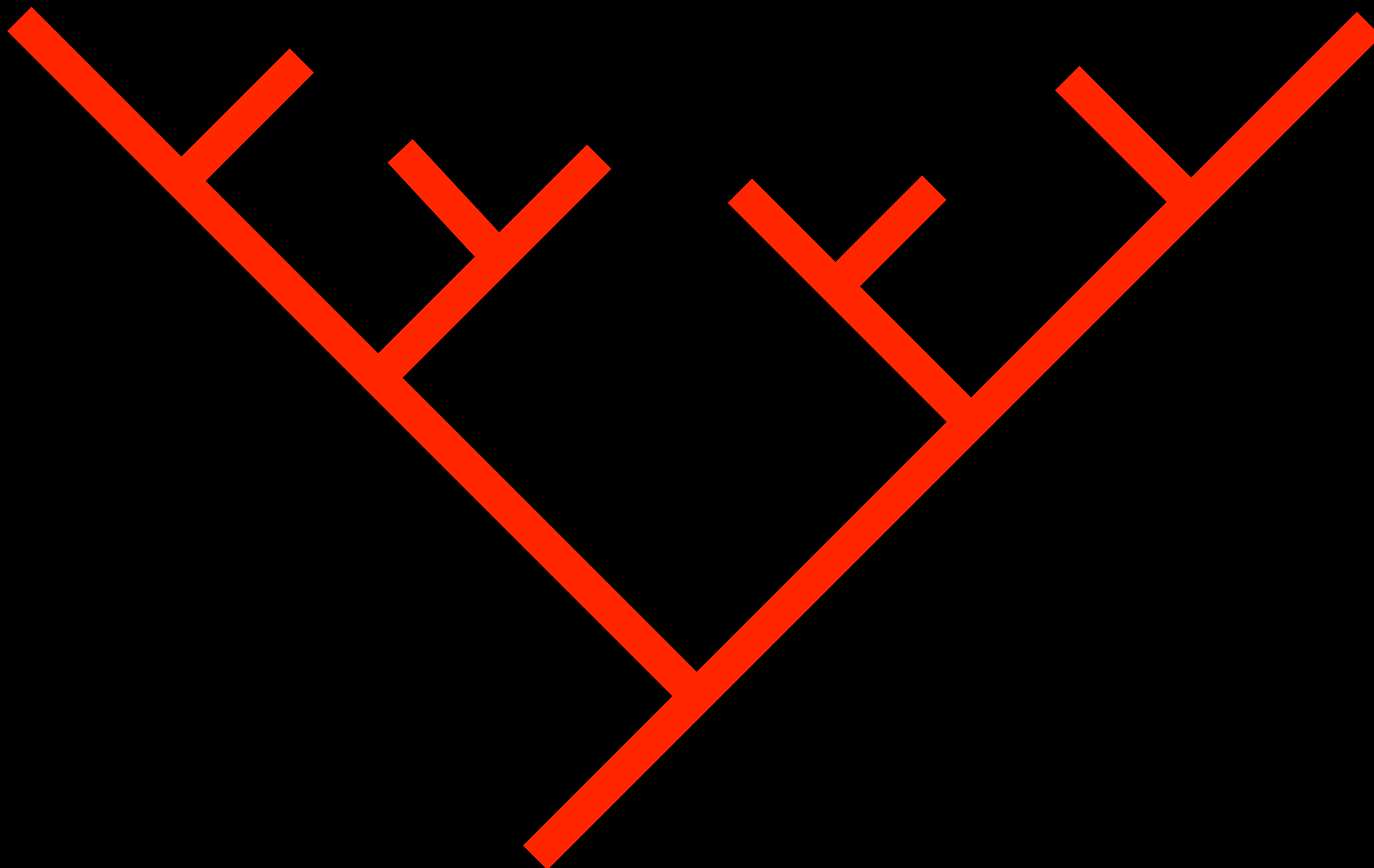


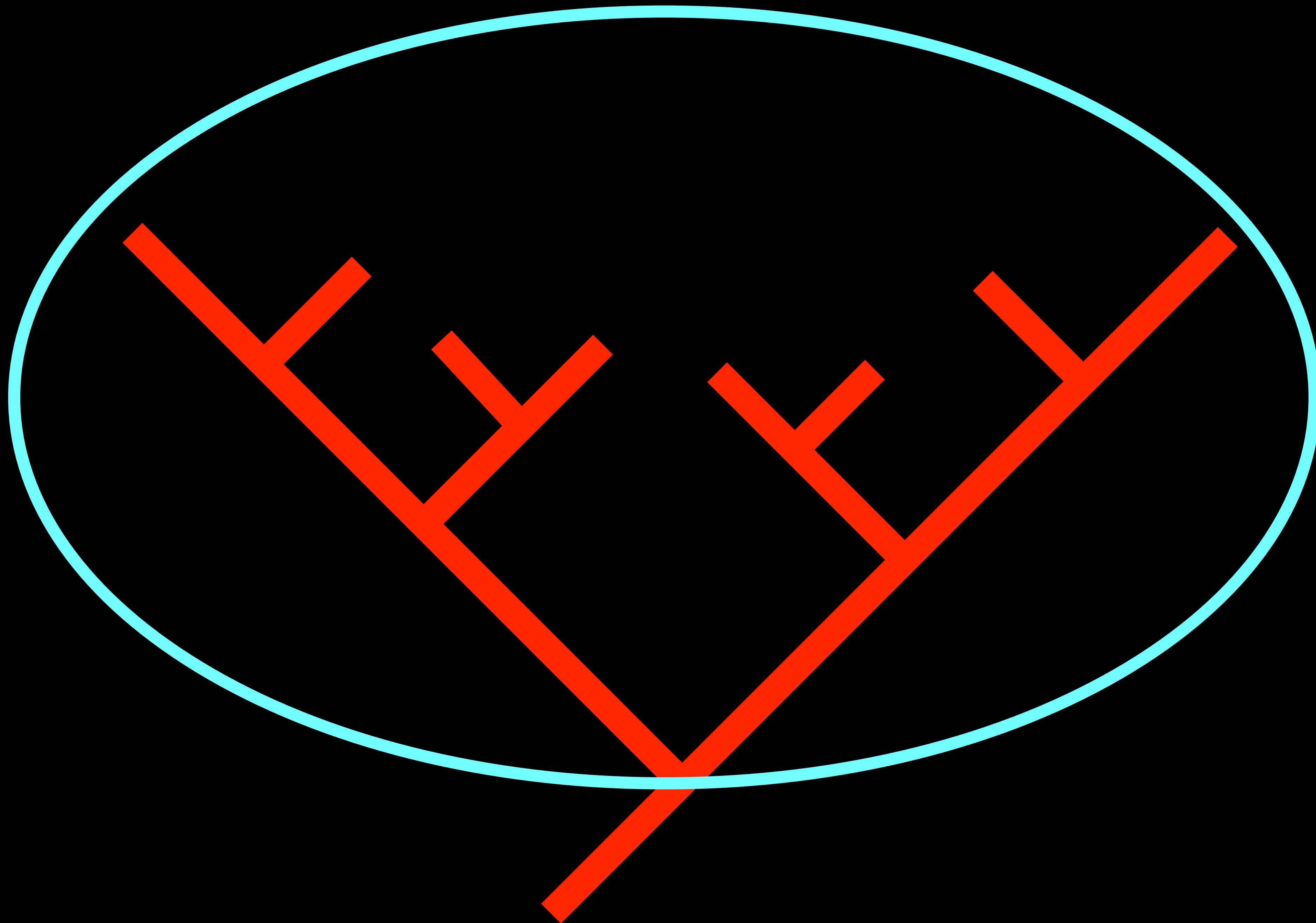


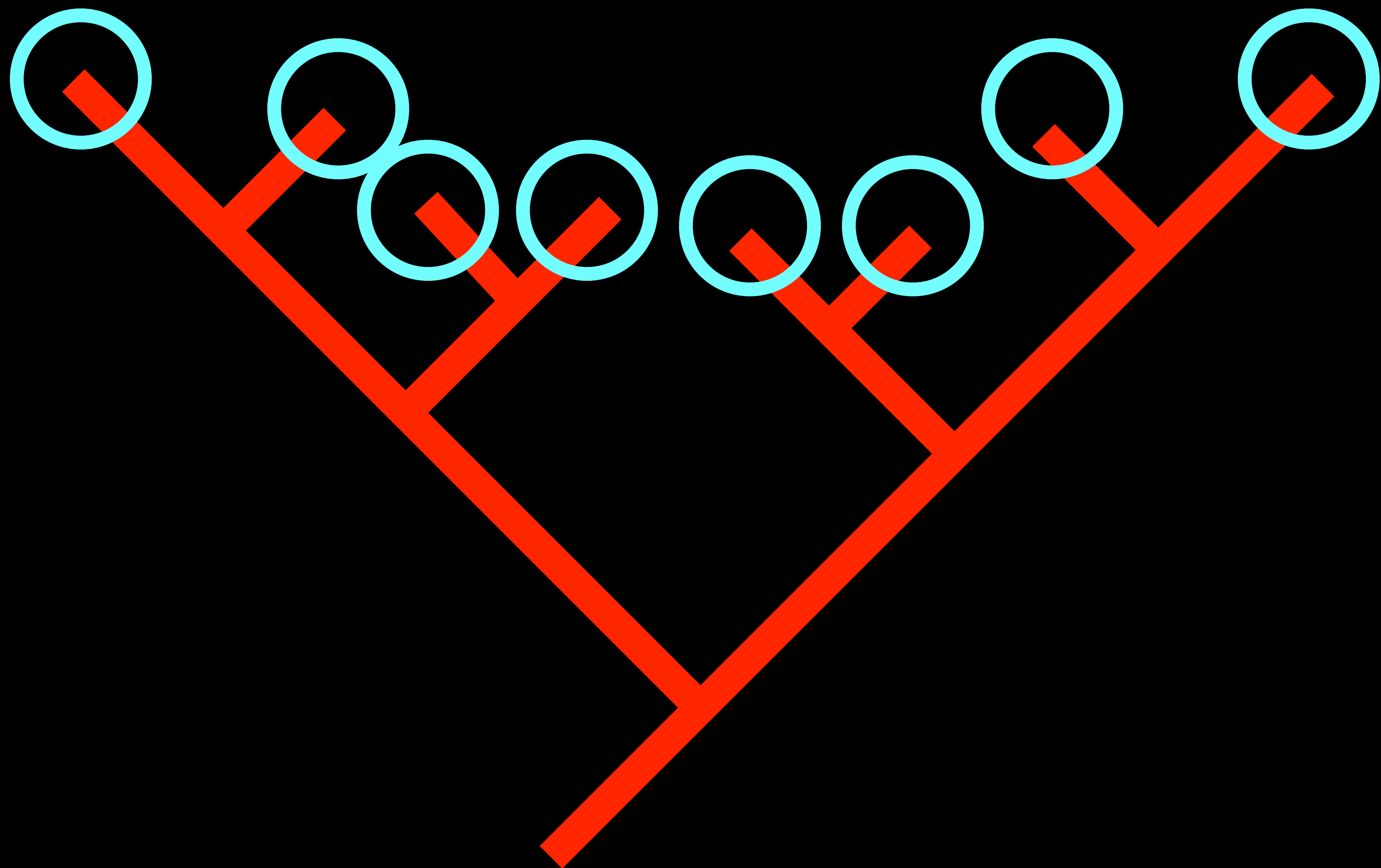


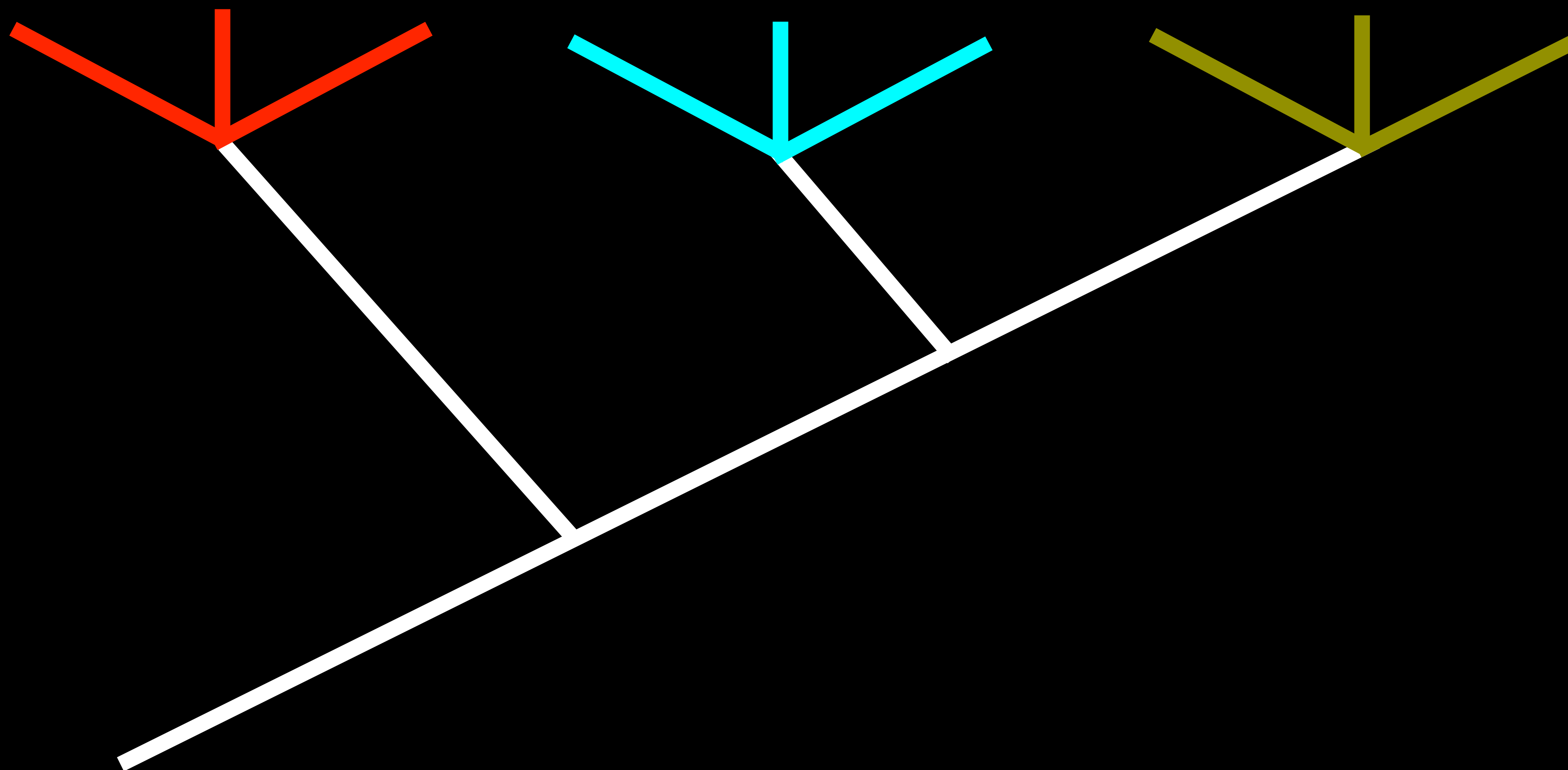


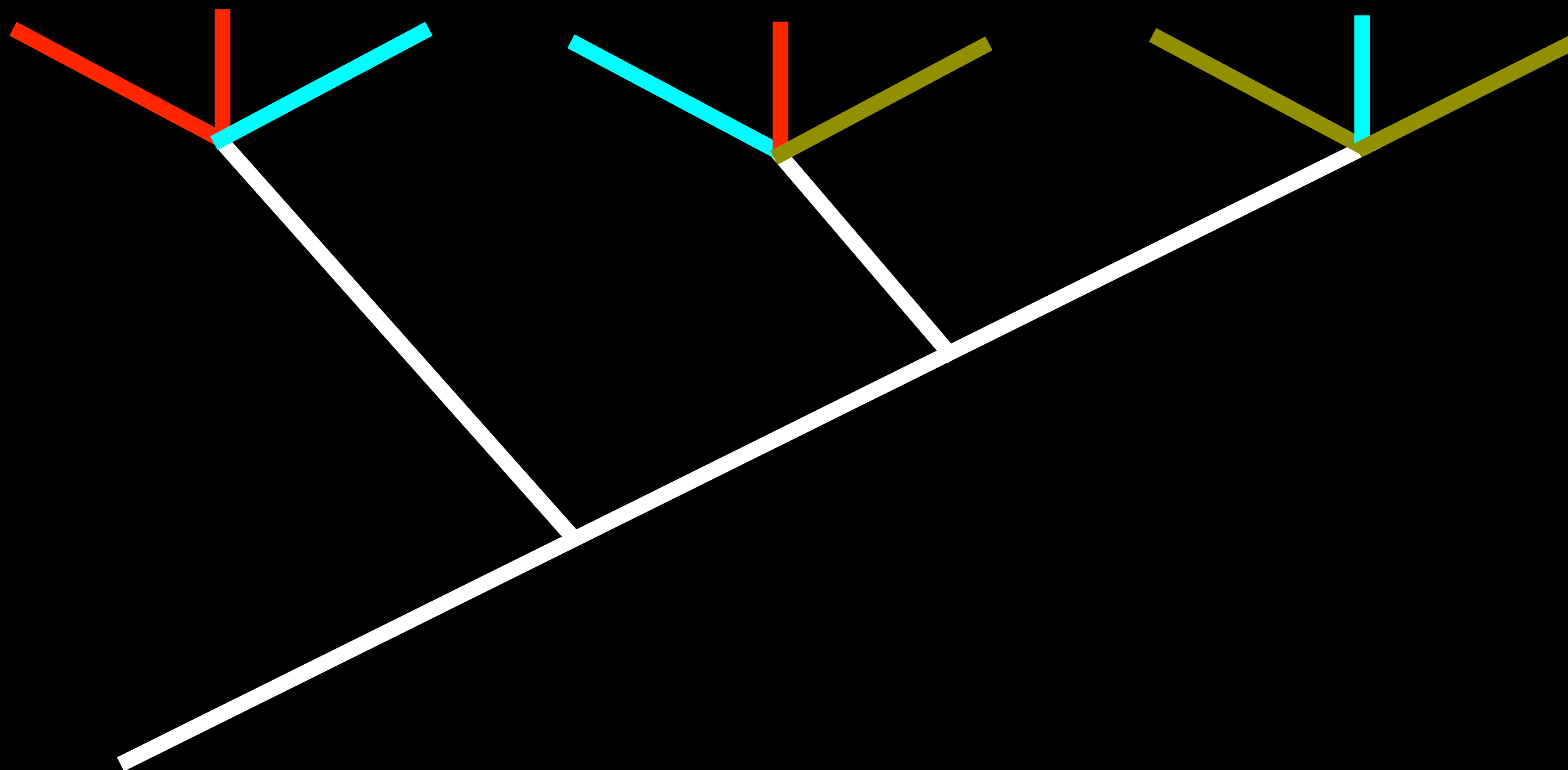


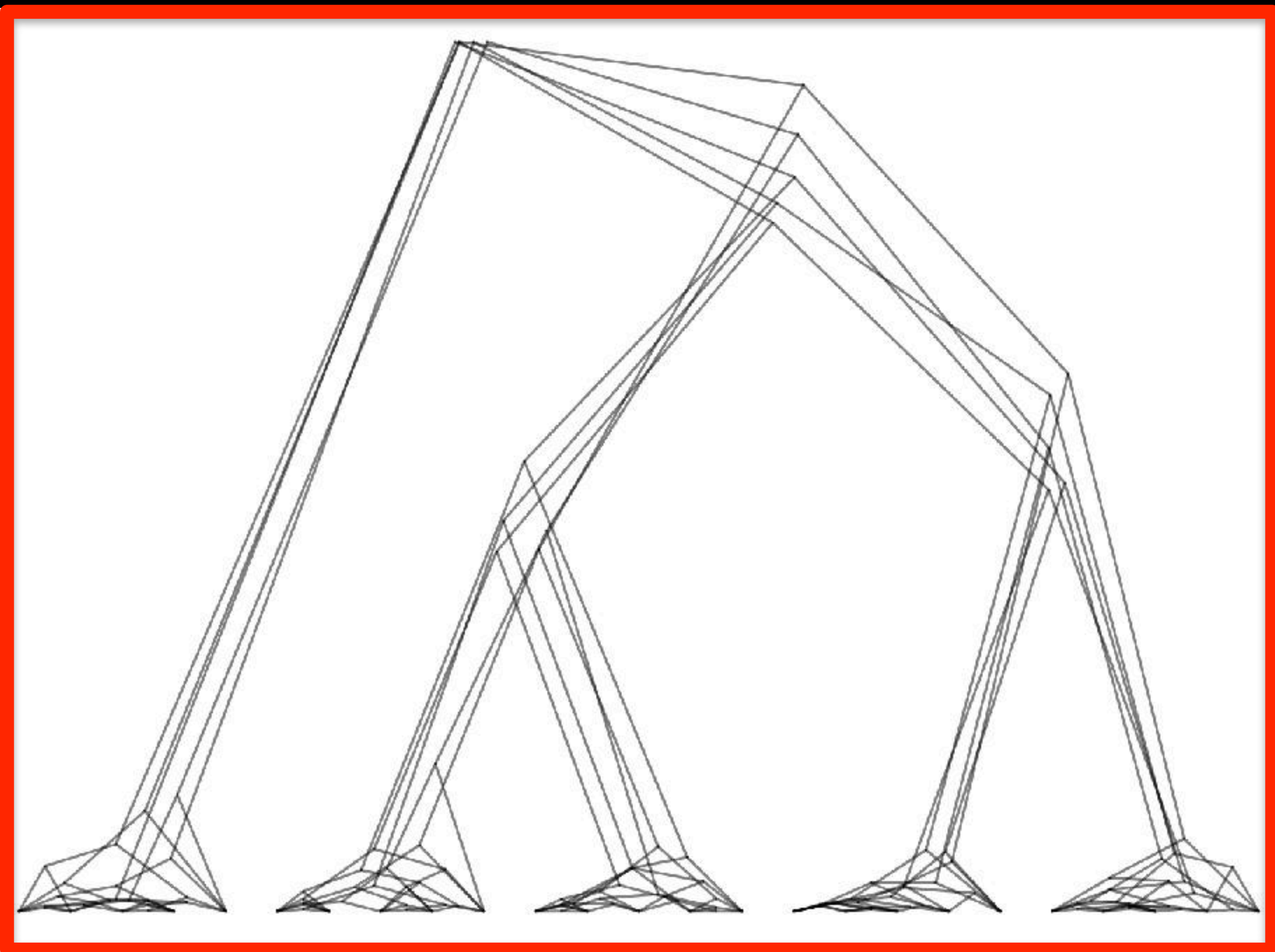




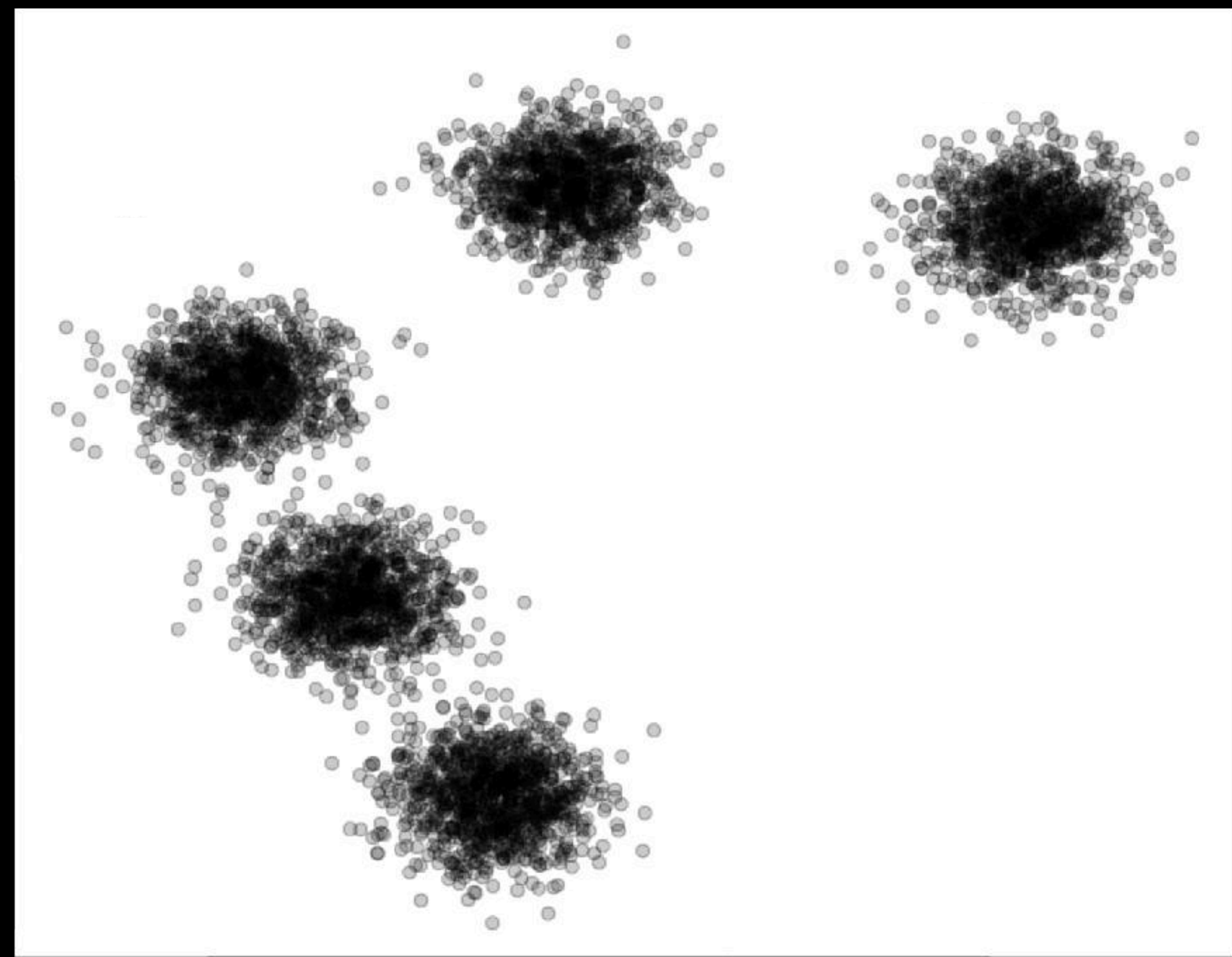








trait 2



trait 1

where to now with OTUs

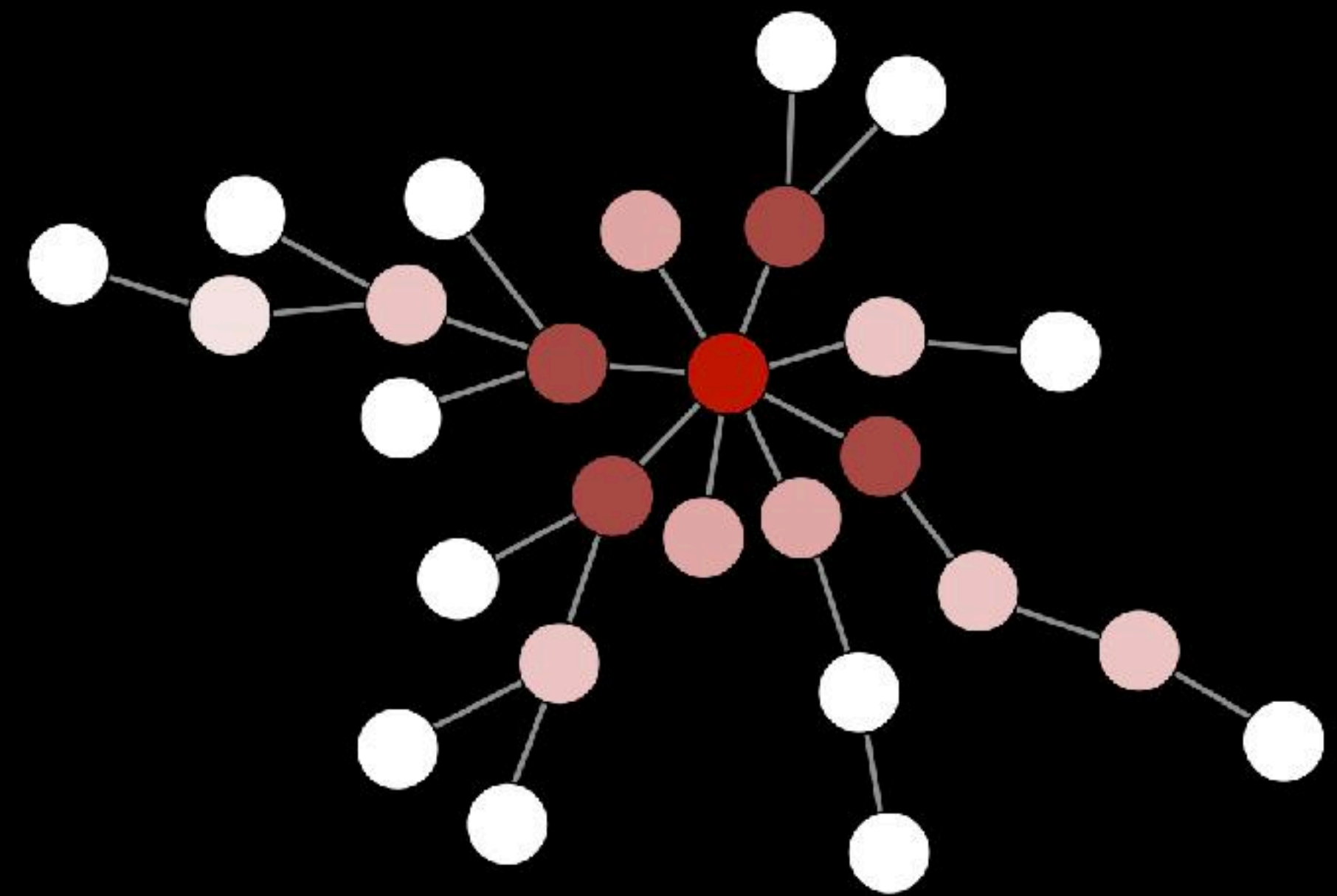
operational taxonomic units

operational taxonomic units

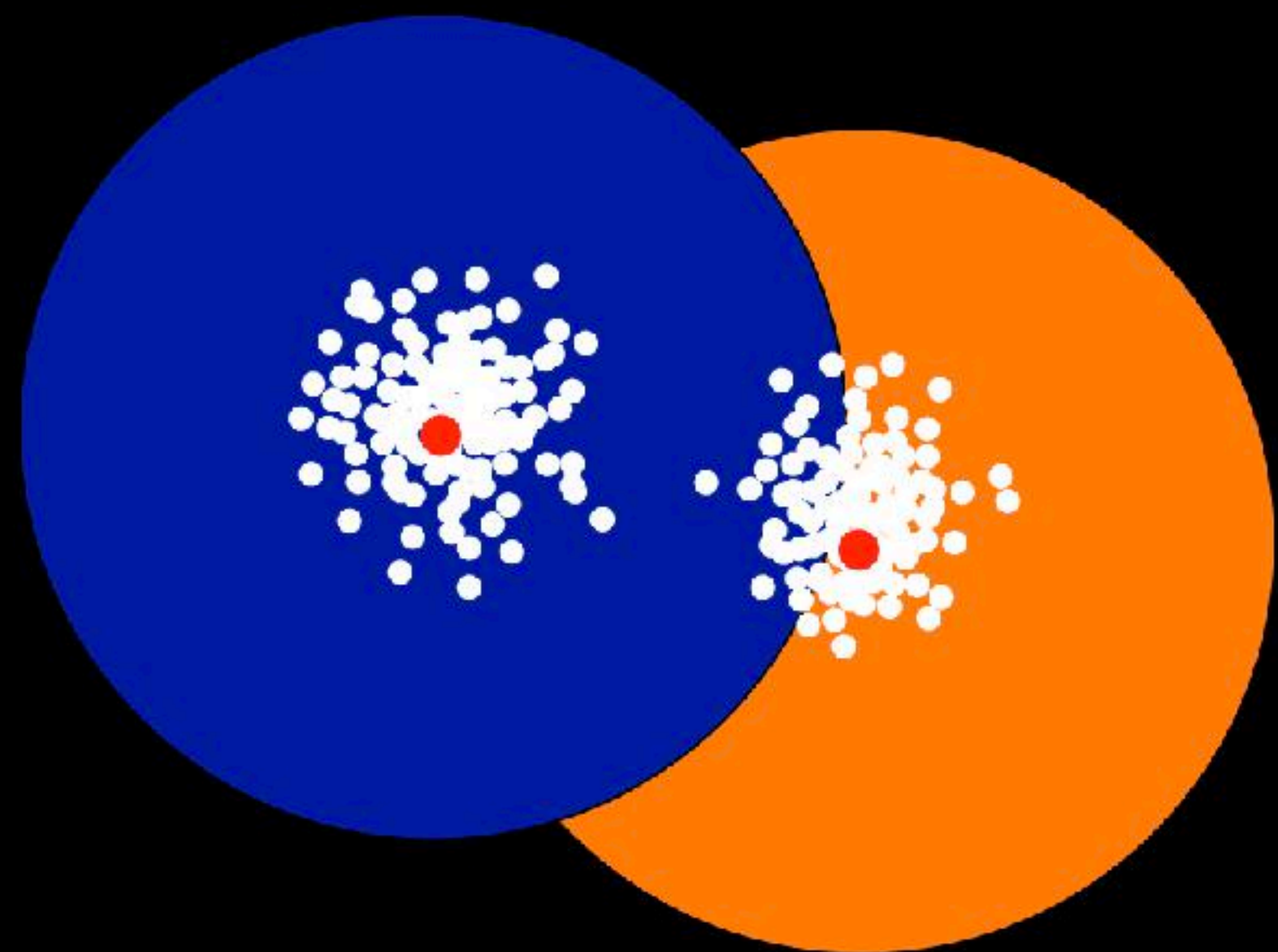
are

operational criteria for

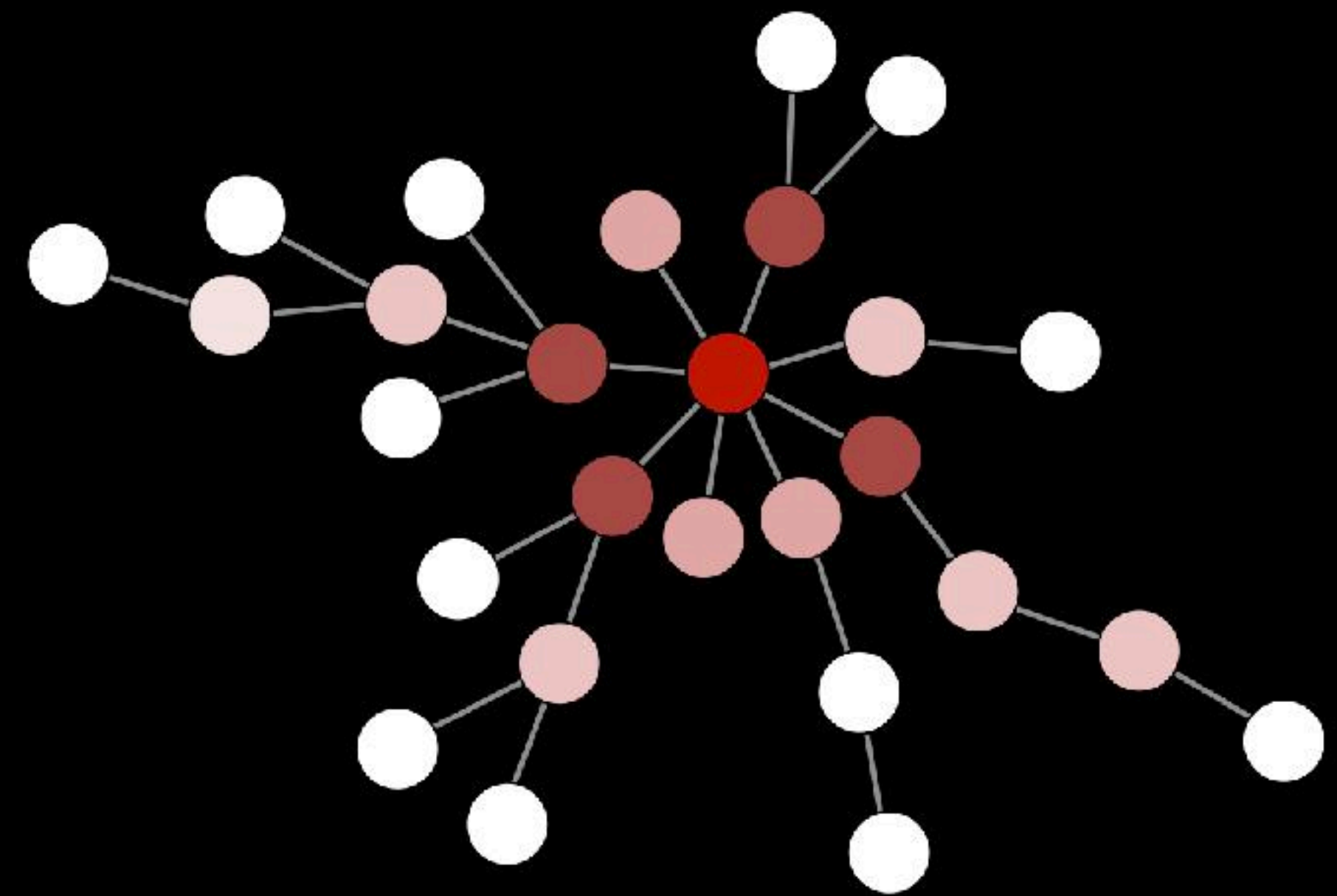
species delimitation



≠

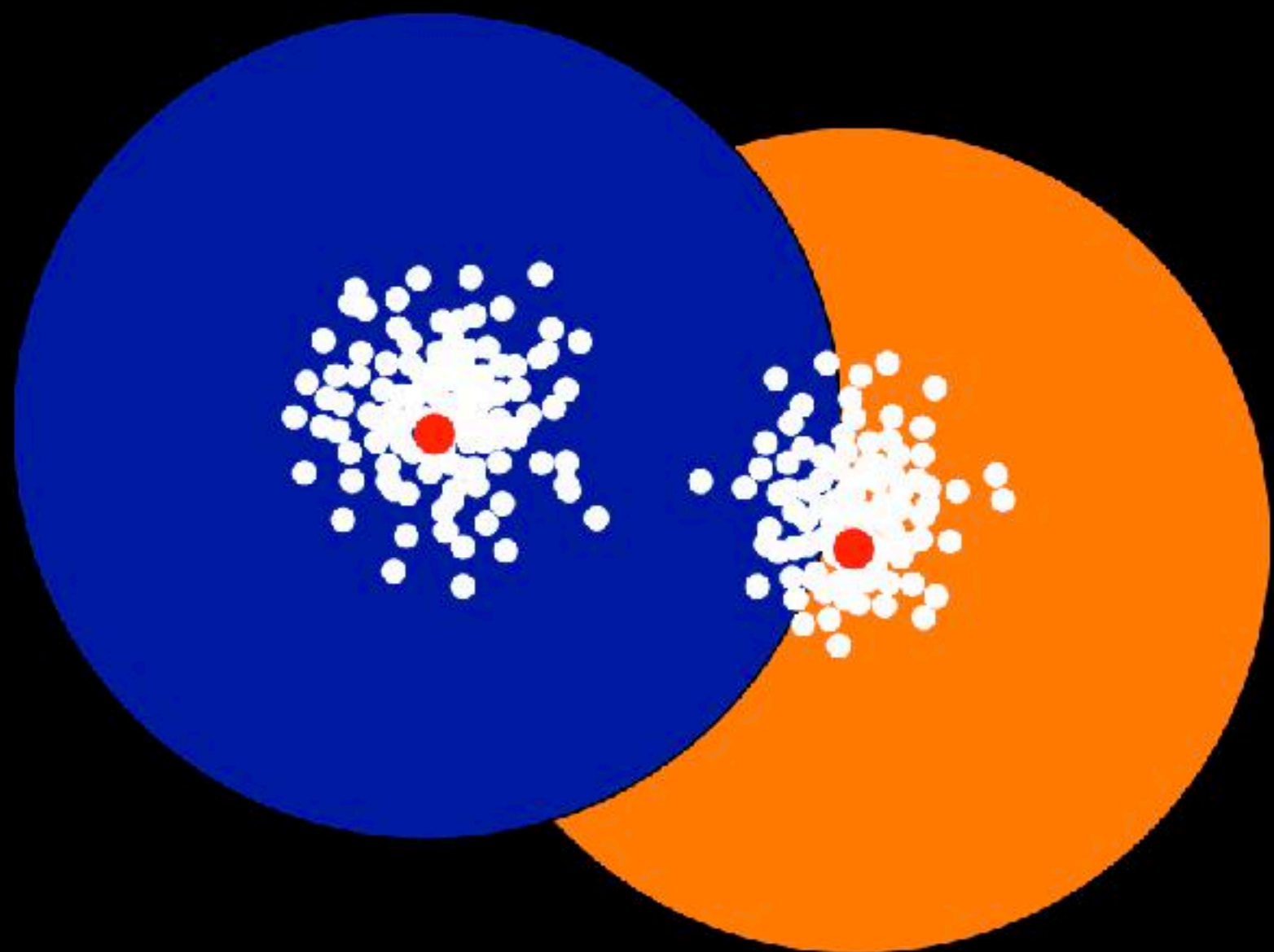


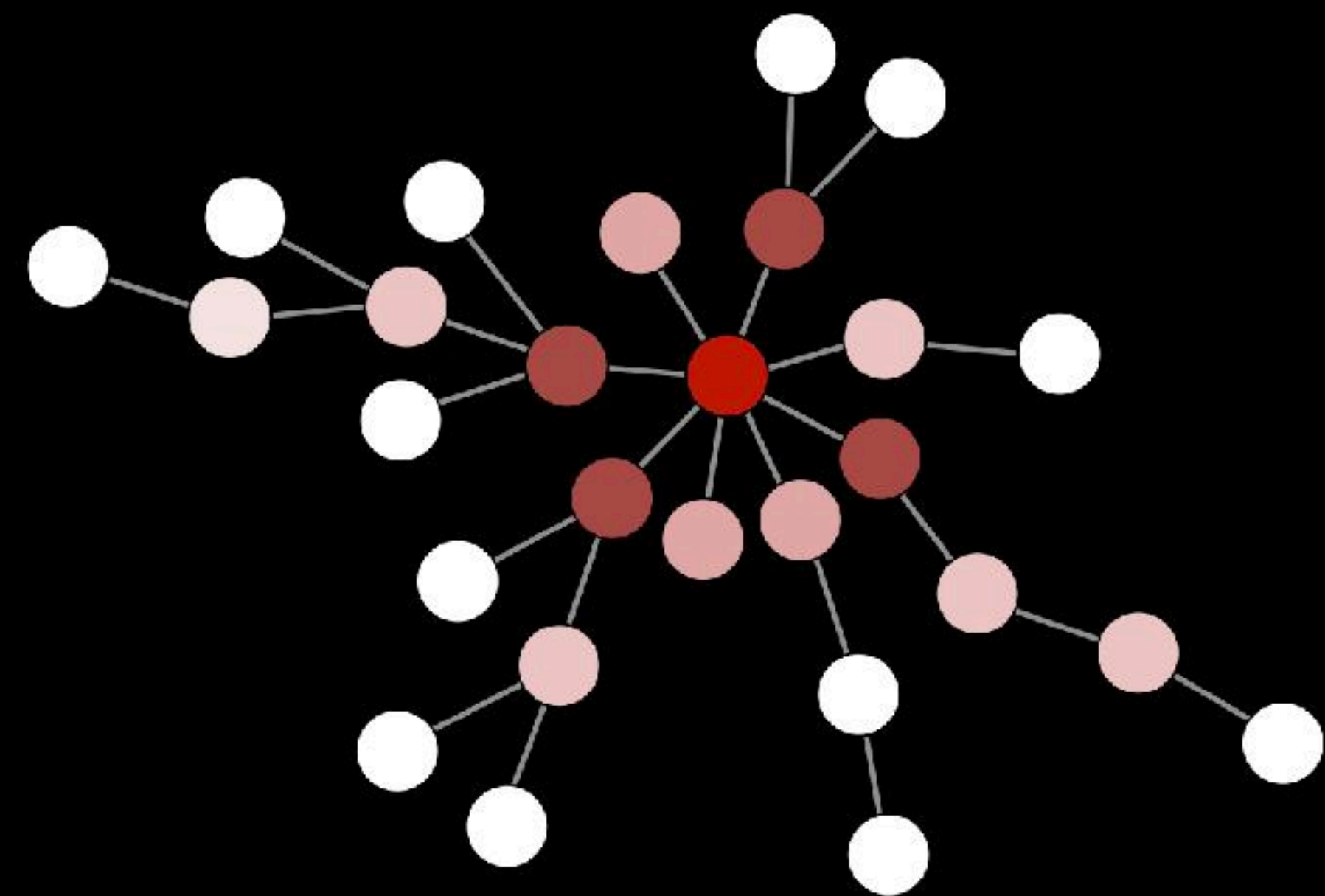
morpho
species



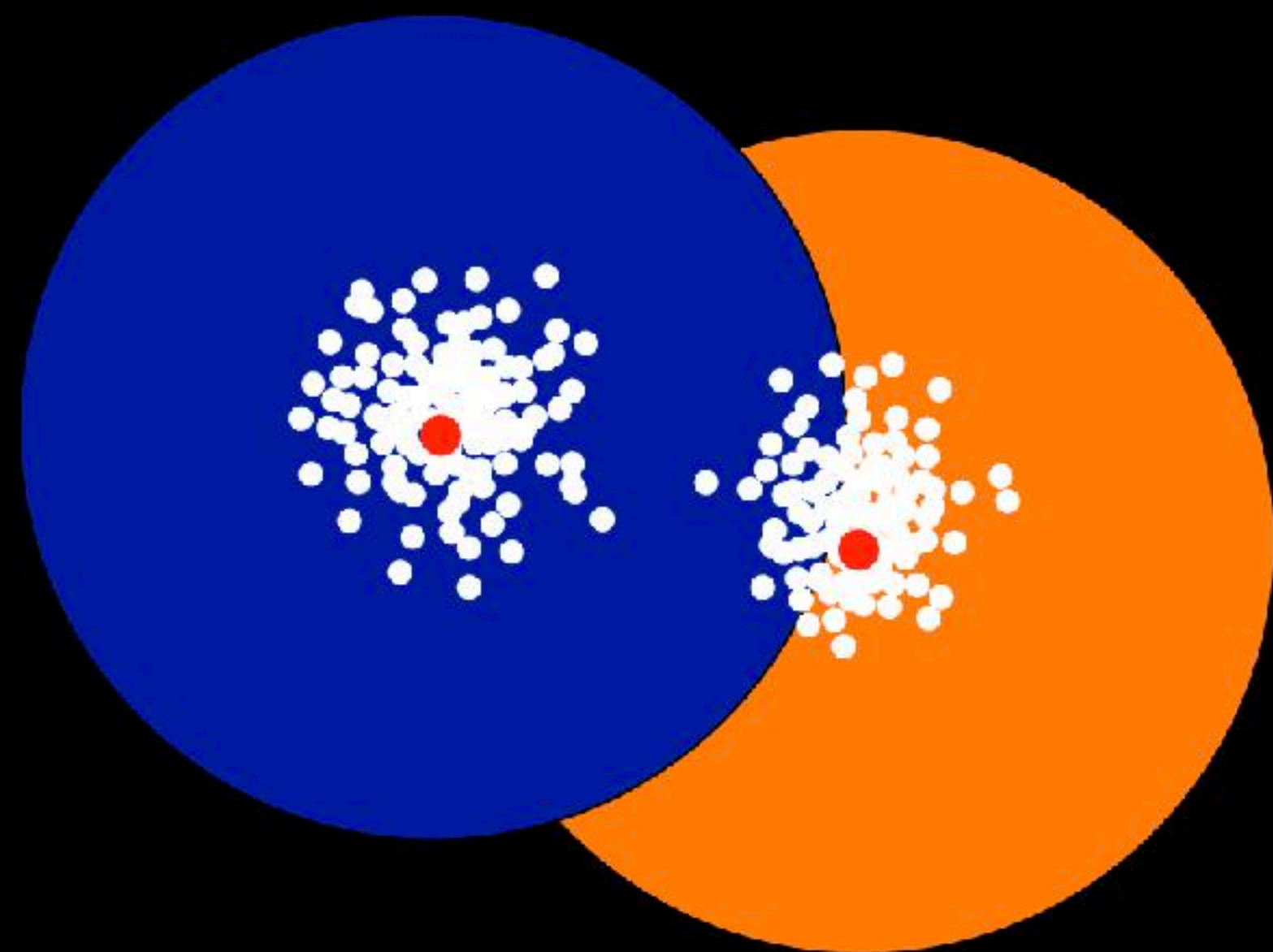
≠

biological
species

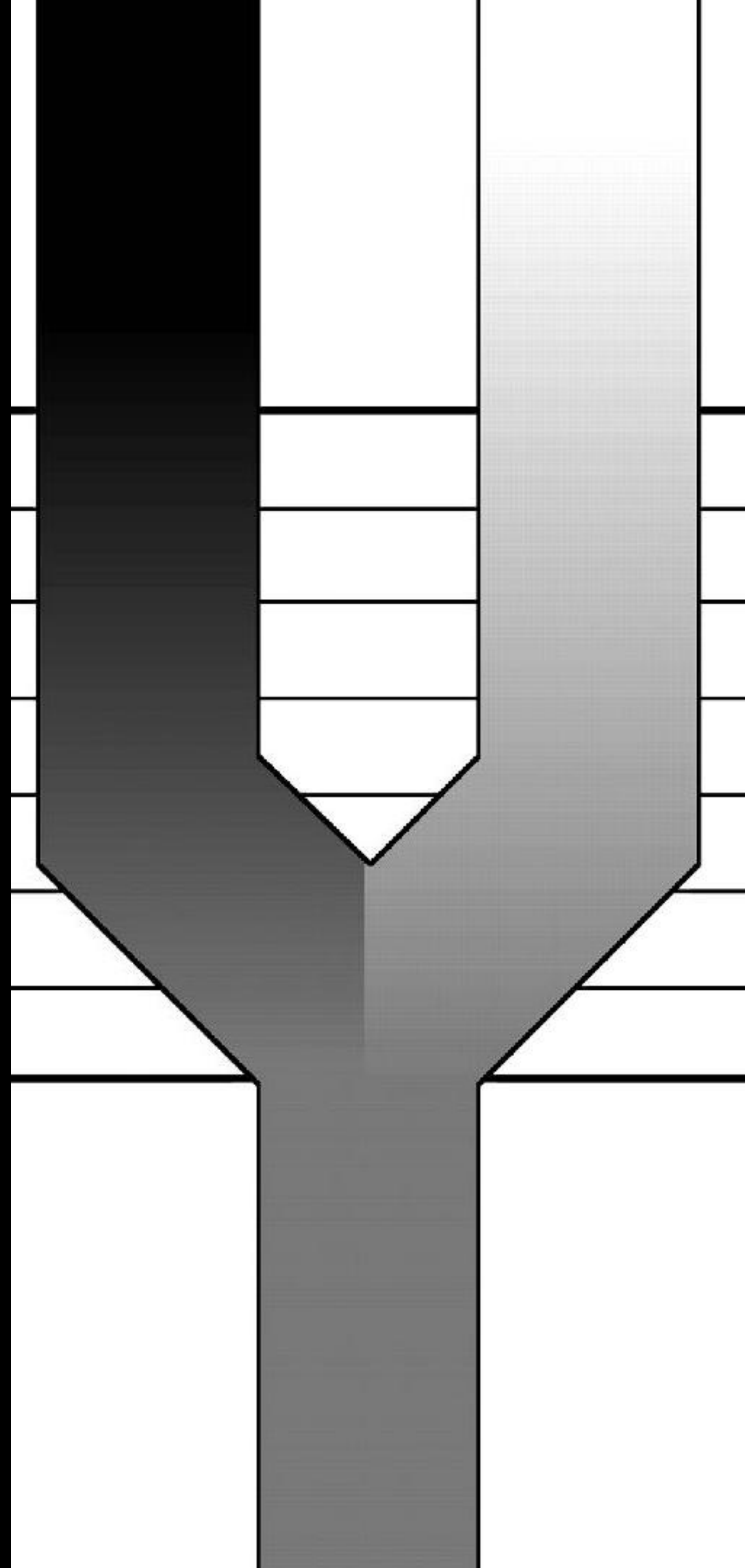




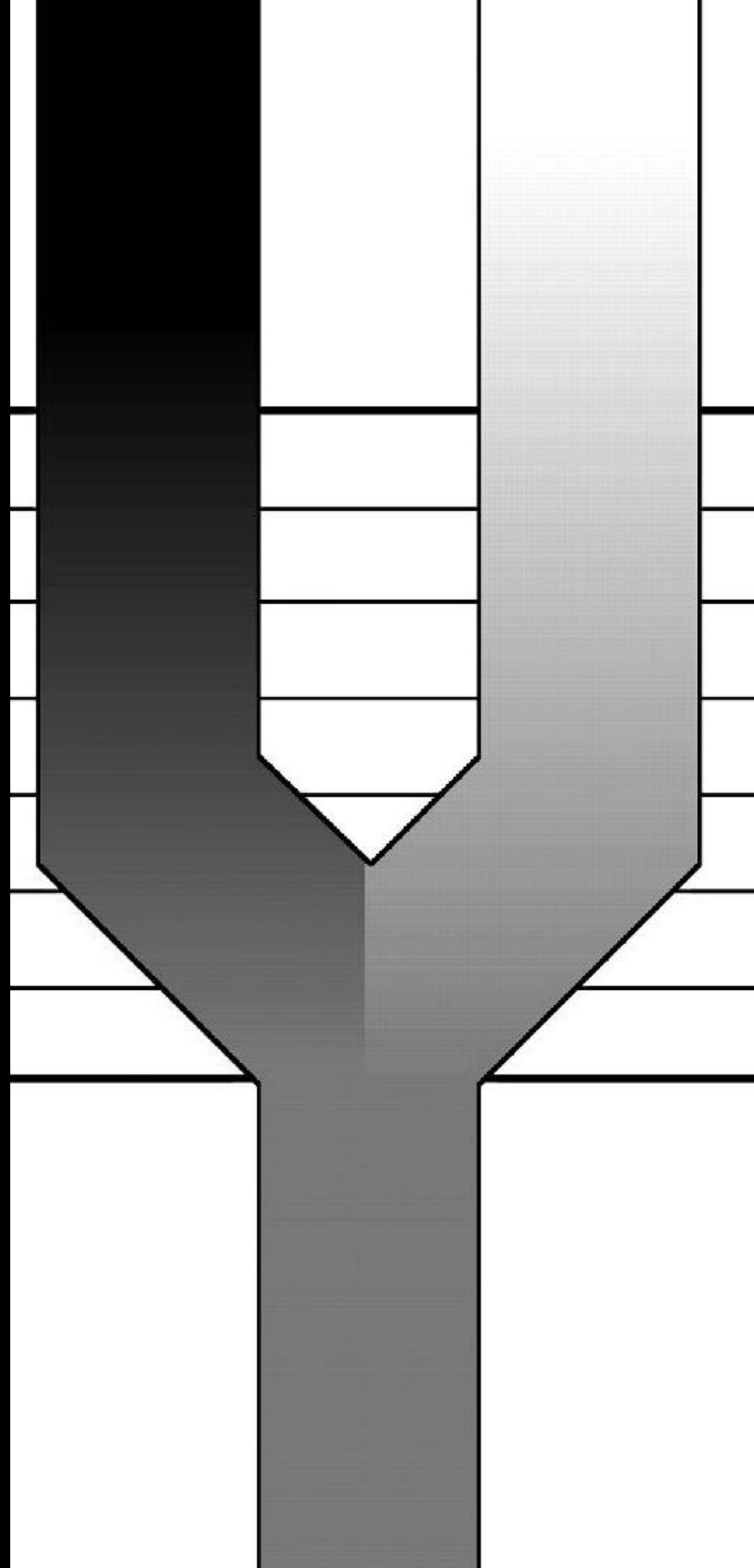
=



operational
criteria



time ↑



operational criteria

Phylogenetic

Morphological

Biological

time ↑

operational criteria

Phylogenetic
Multi-rate PTP

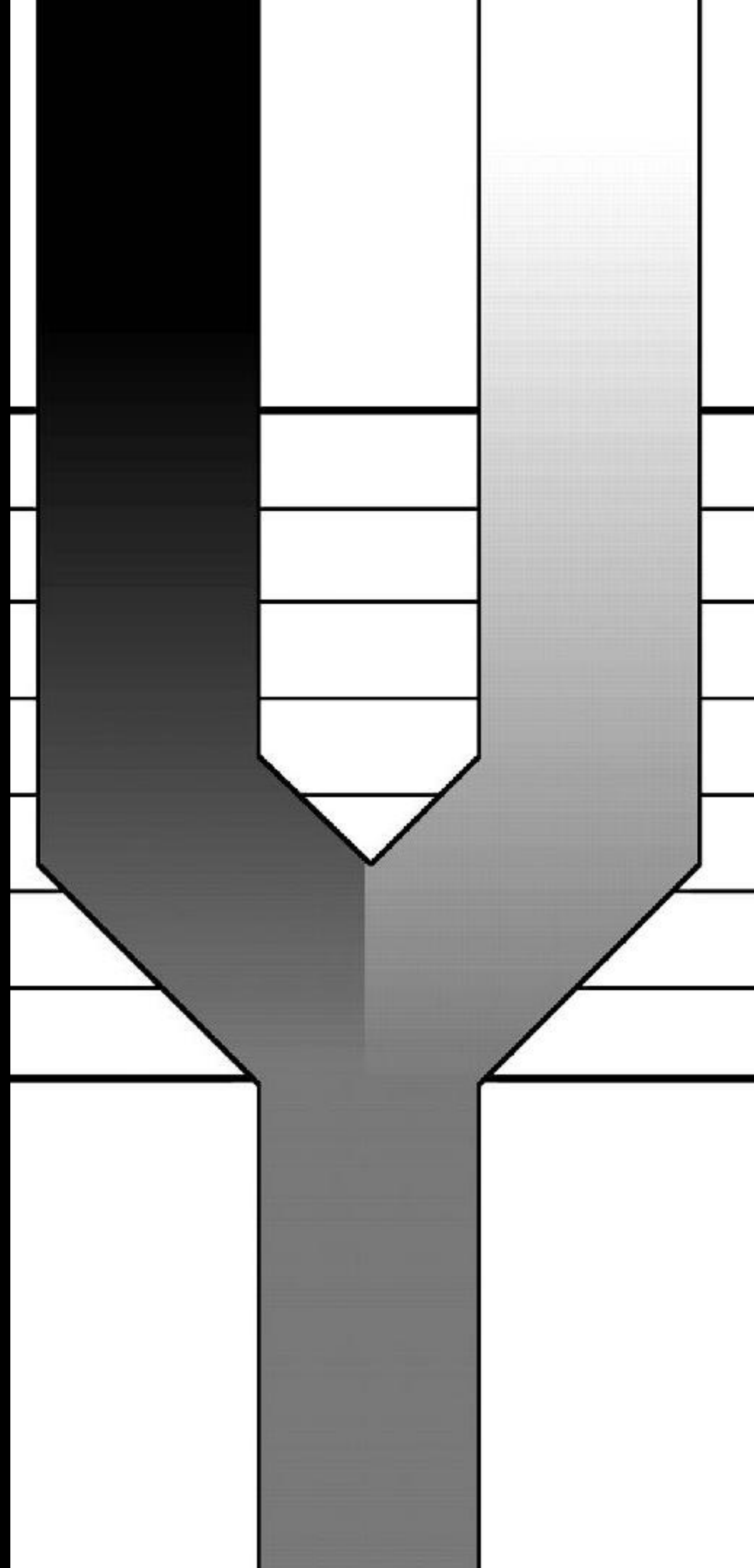
Morphological

Swarm

Biological

DADA

time ↑



operational criteria

Phylogenetic
Multi-rate PTP

Morphological

Swarm

Biological

DADA

in conclusion

OTUs $\stackrel{?}{=}$ species

OTUS ?
= separately
evolving
metapopulation
lineages

thanks to:

CIRAD

Frédéric Mahé

Heidelberg Institute for Theoretical Studies

Alexandros Stamatakis

Pierre Barbera

Institut de Ciències del Mar

Ramiro Logares



**Heisenberg-
Programm**

Deutsche
Forschungsgemeinschaft

