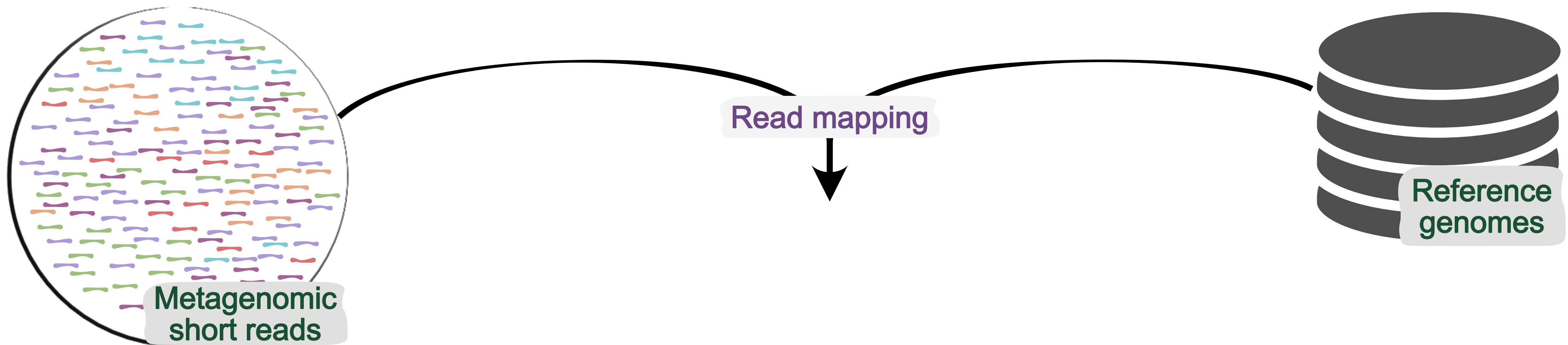


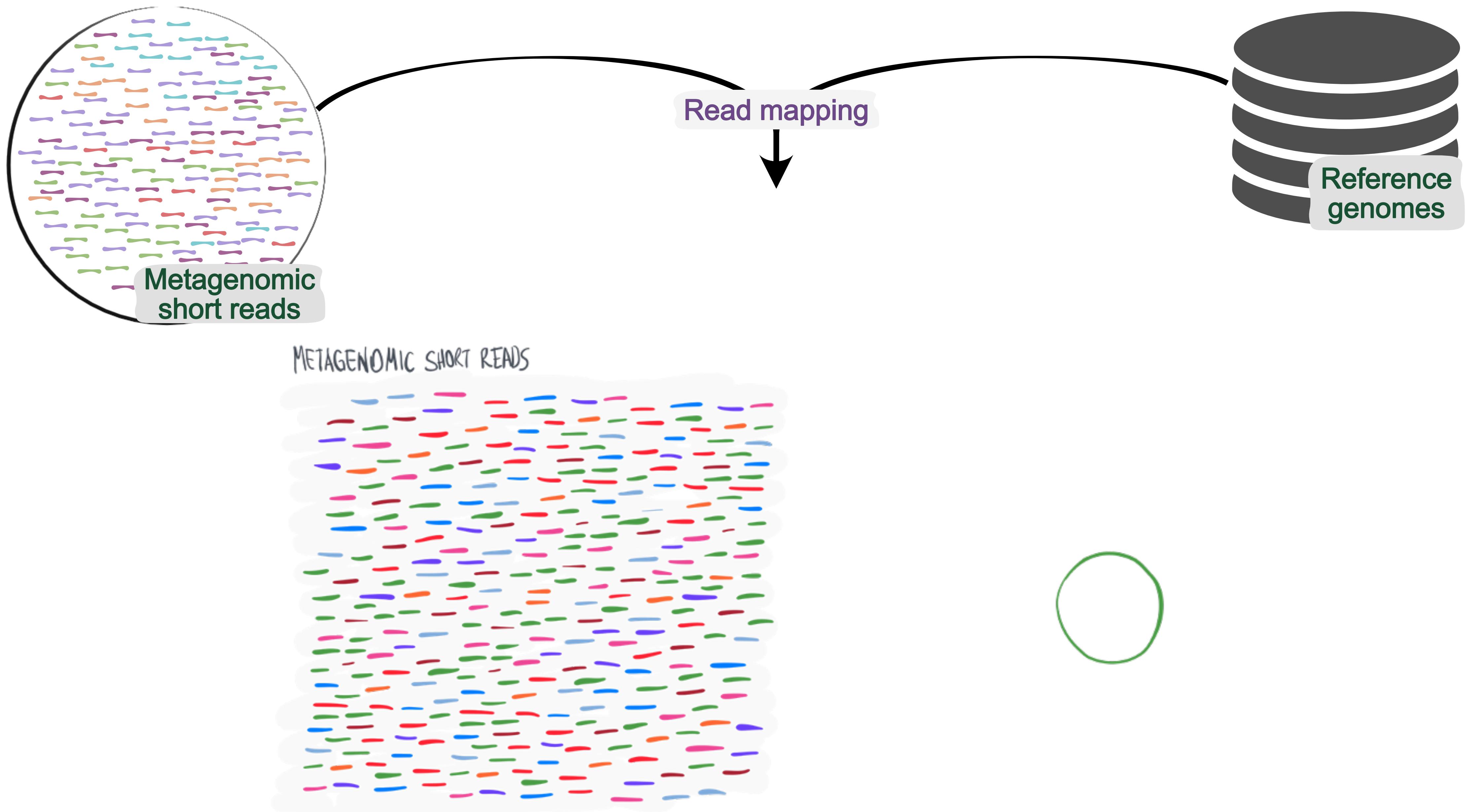
# Methodological details

# Metagenomic read mapping

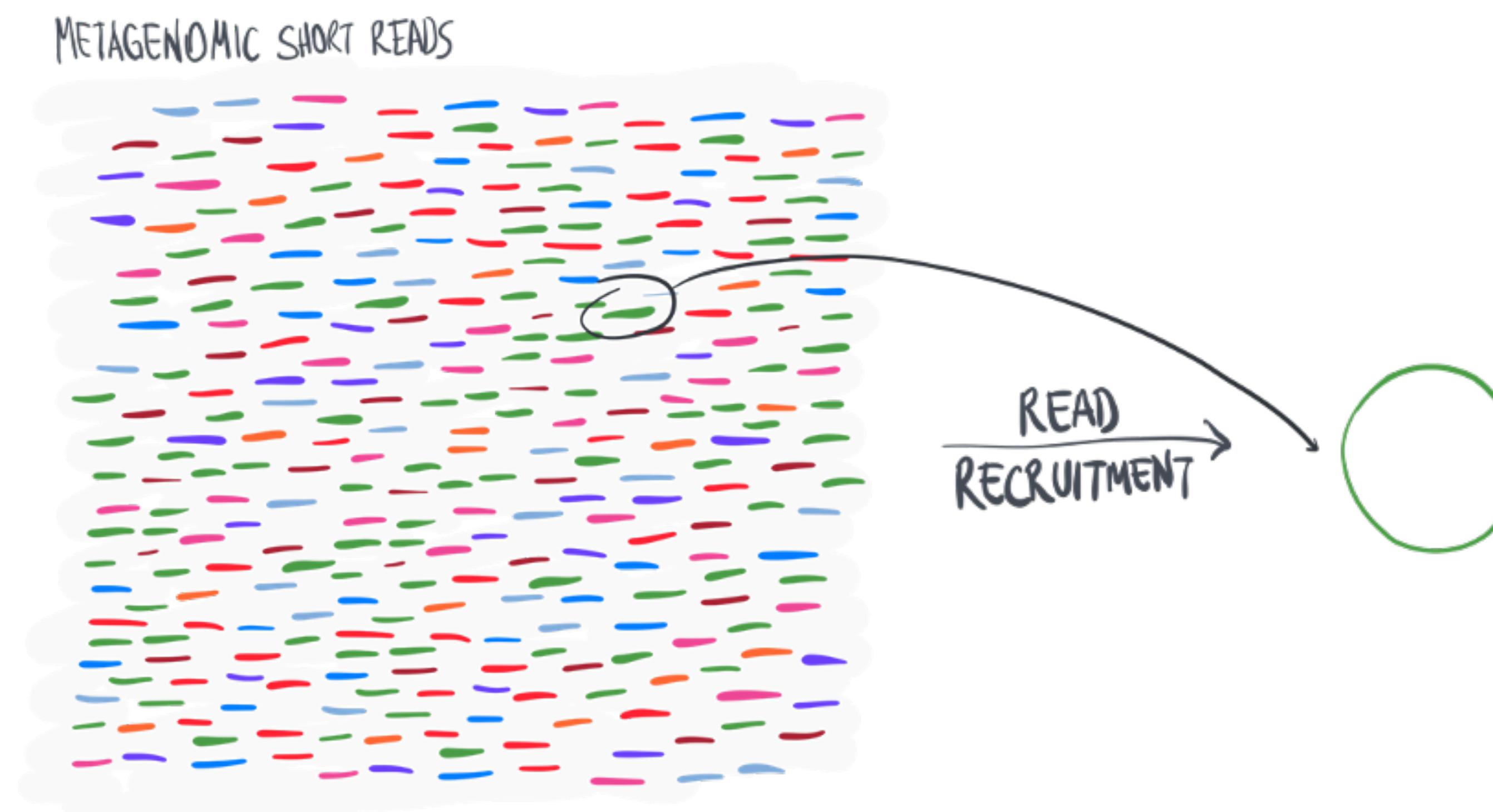
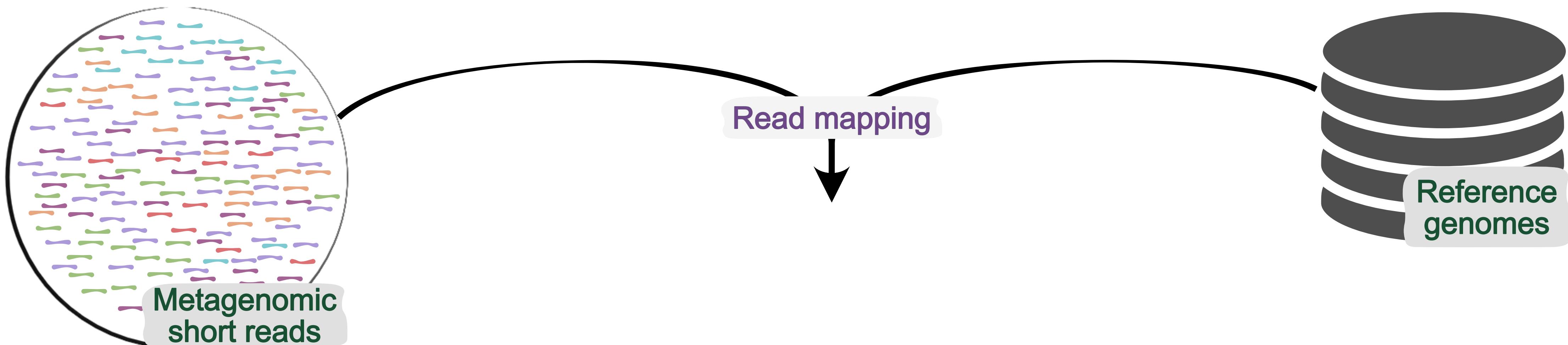
# Metagenomics in a reference context



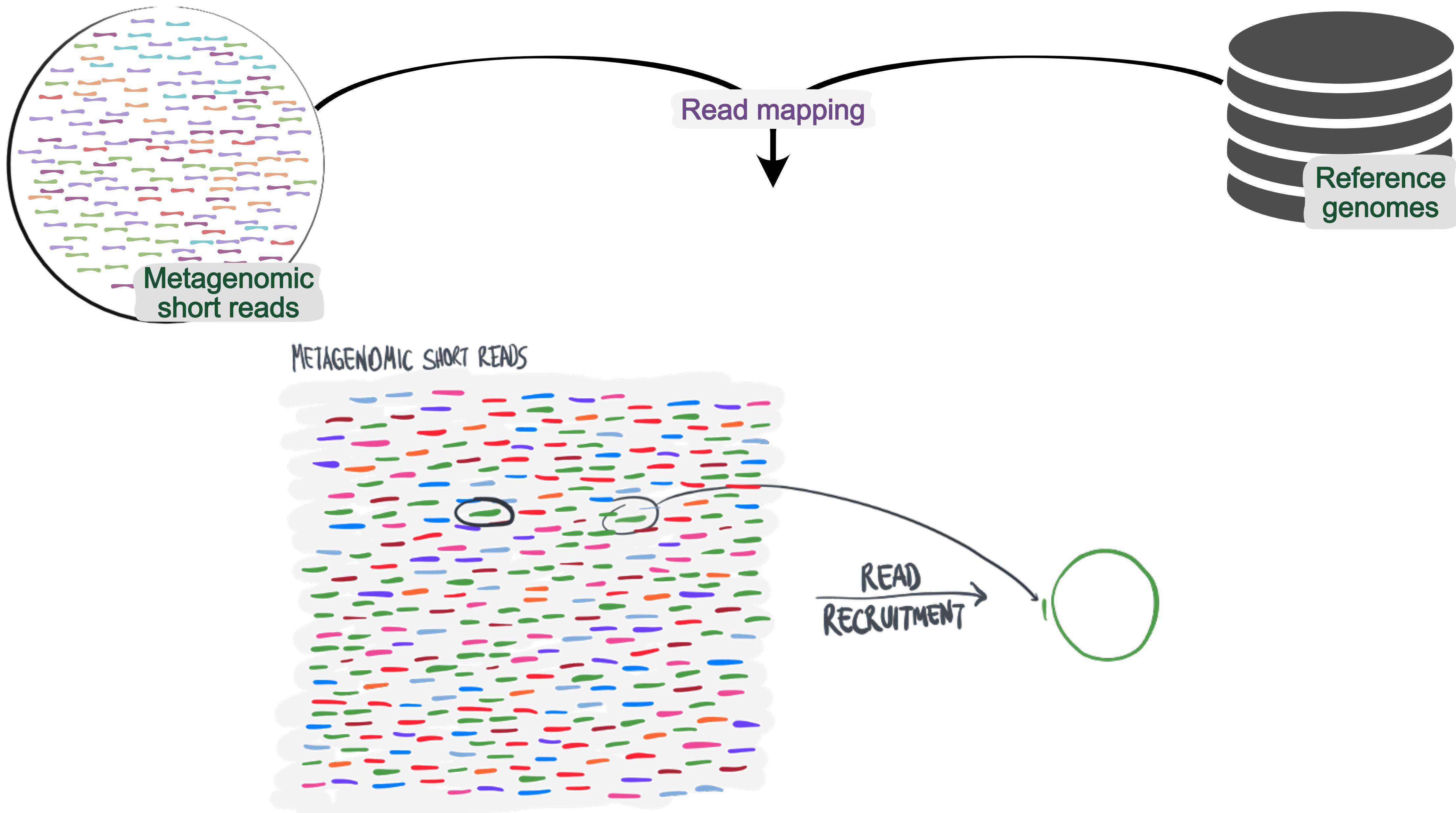
# Metagenomics in a reference context



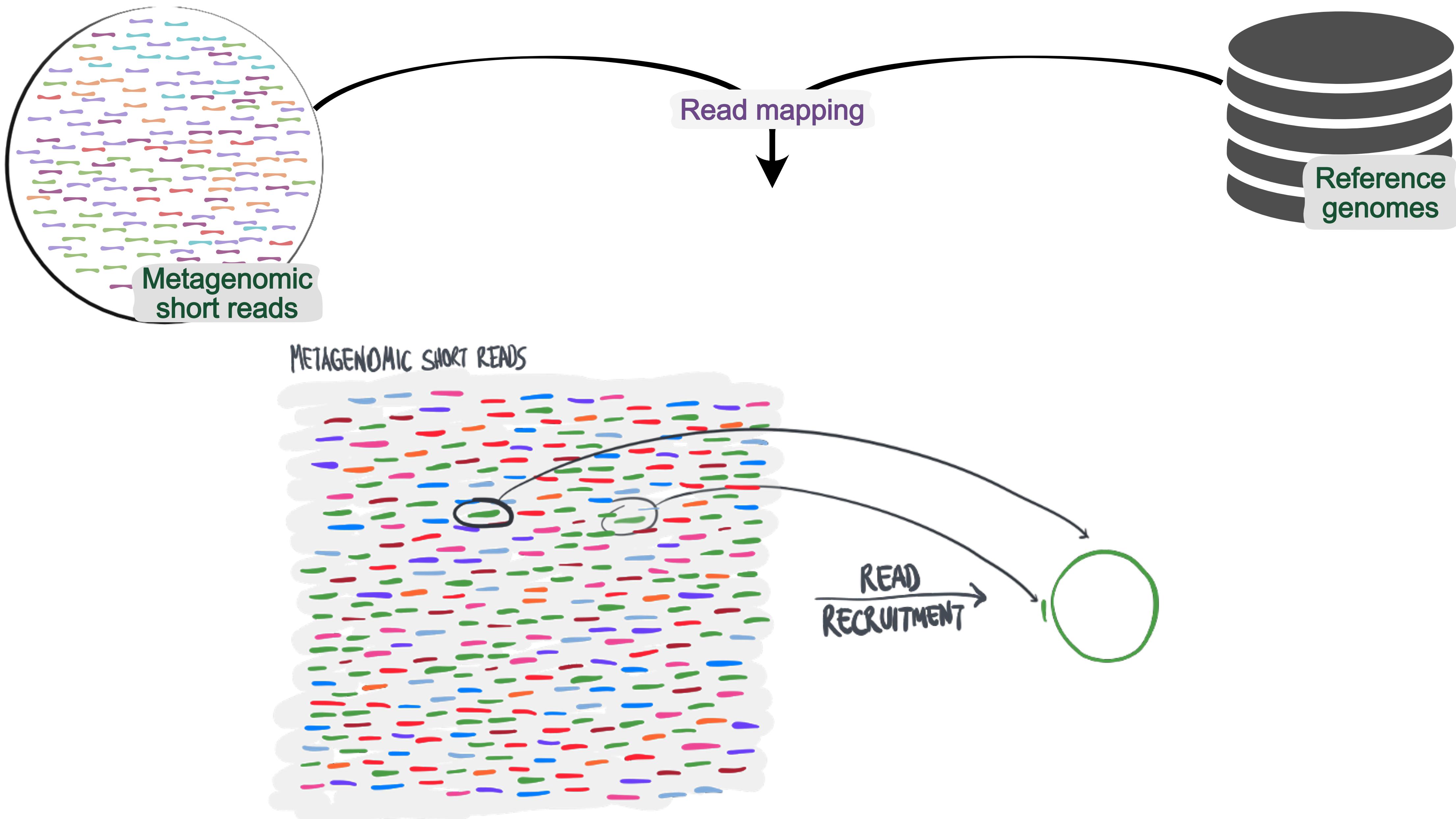
# Metagenomics in a reference context



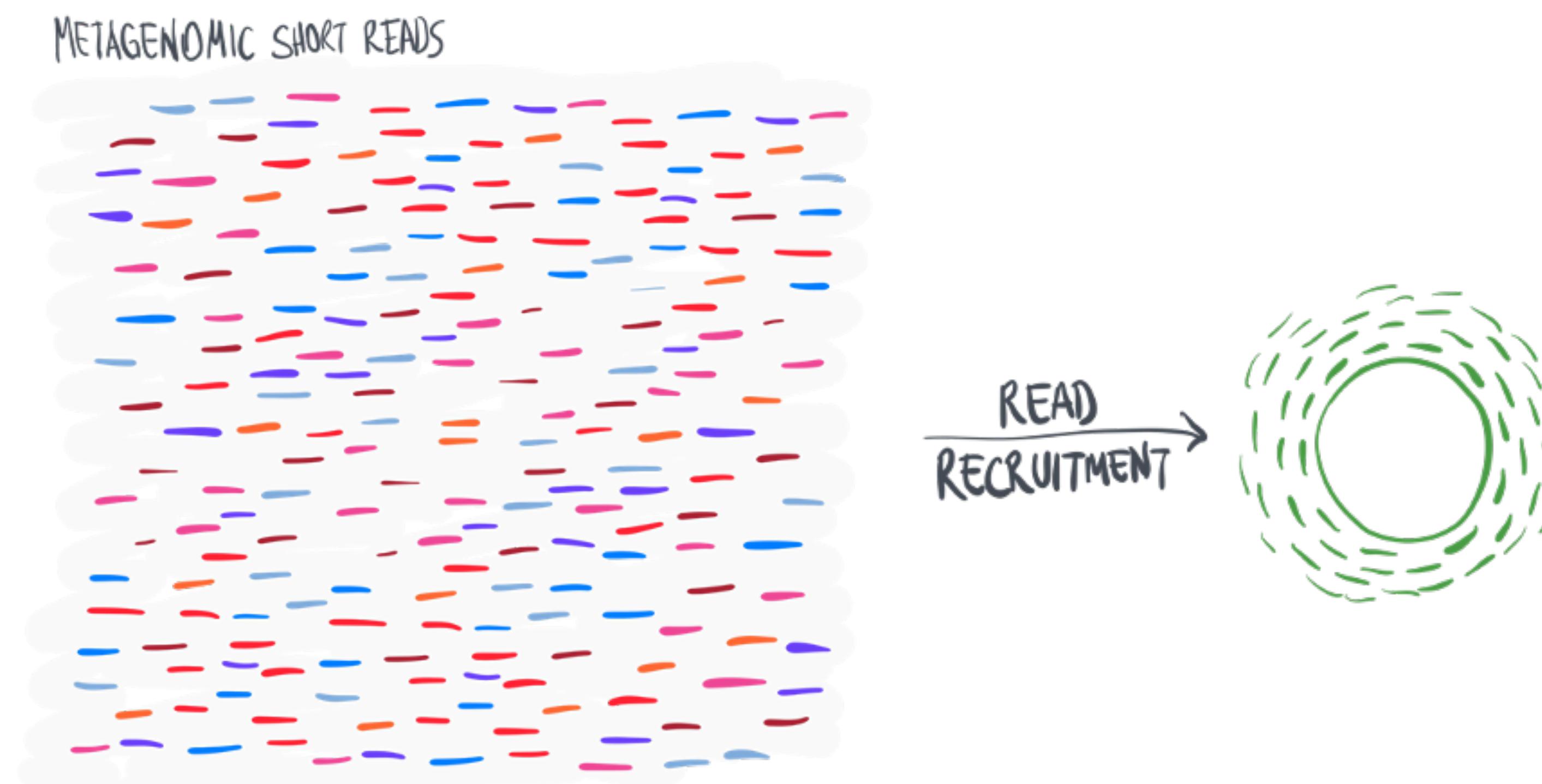
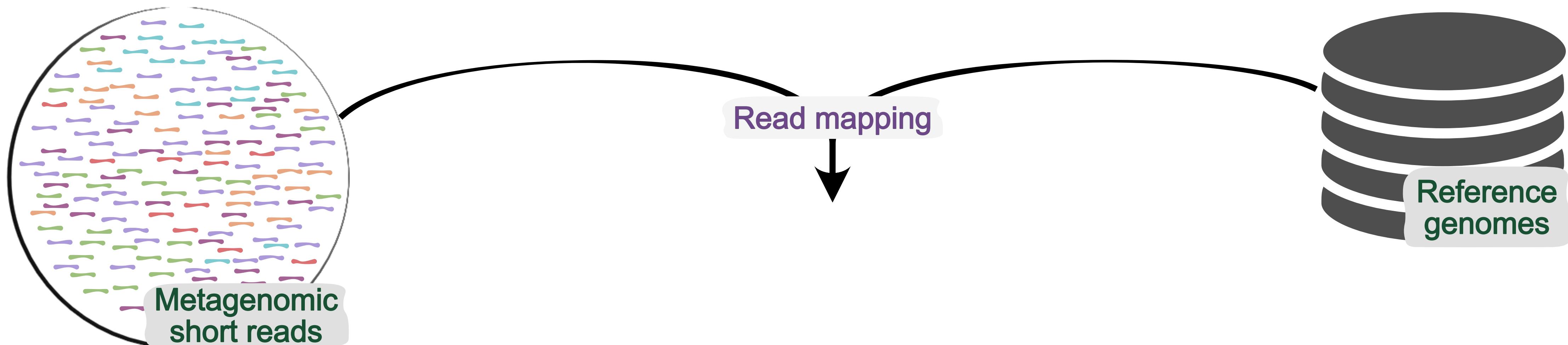
# Metagenomics in a reference context



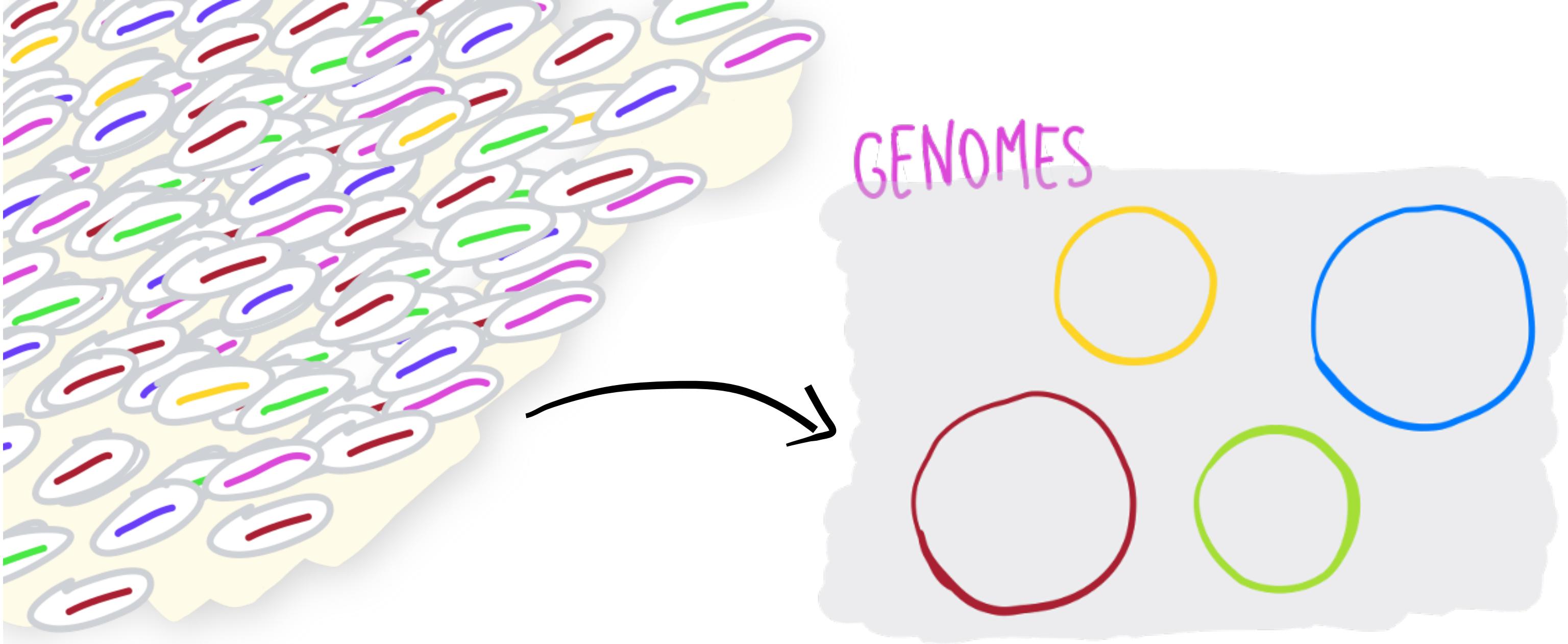
# Metagenomics in a reference context

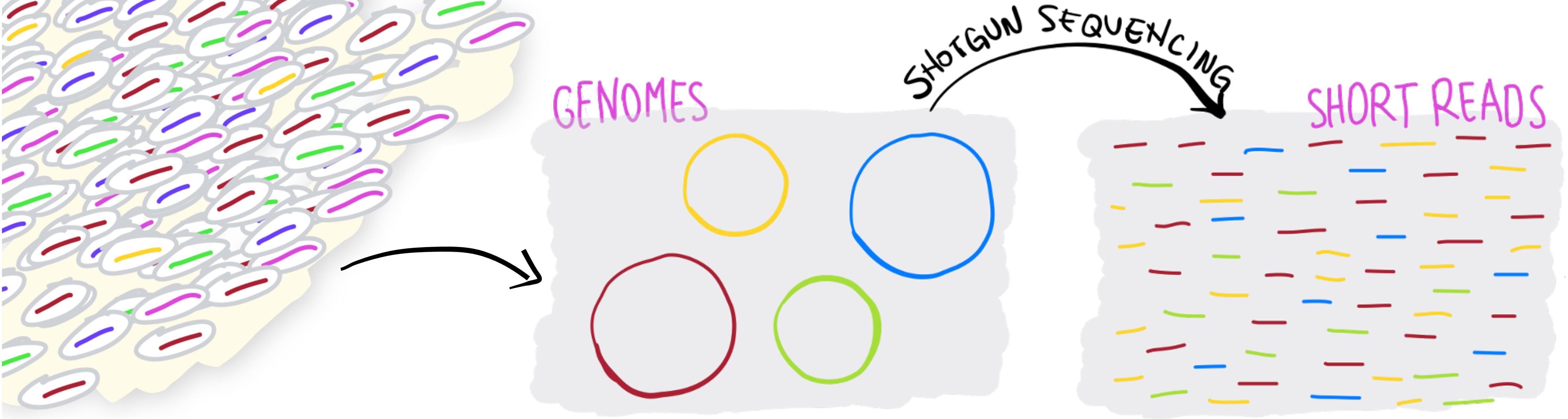


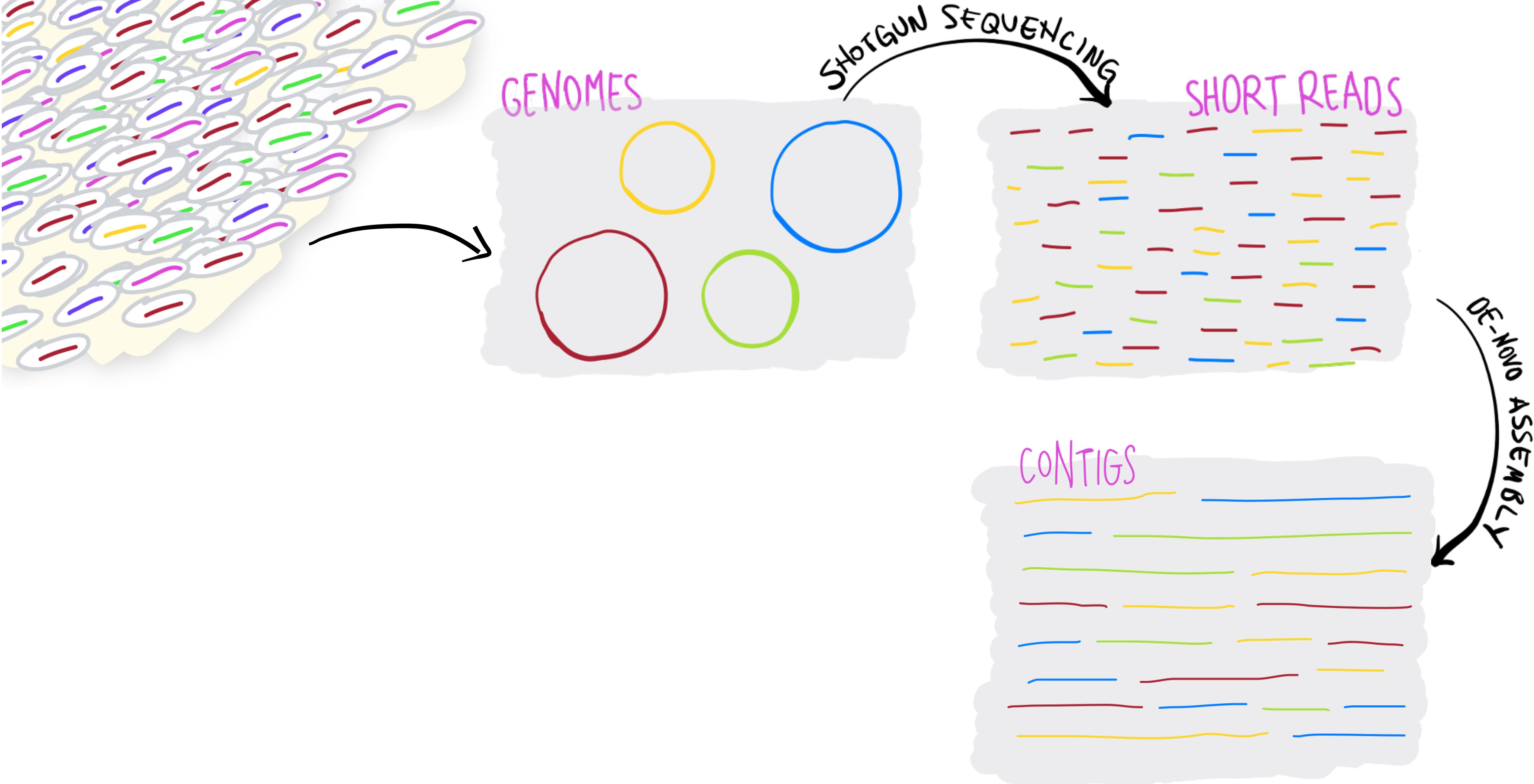
# Metagenomics in a reference context

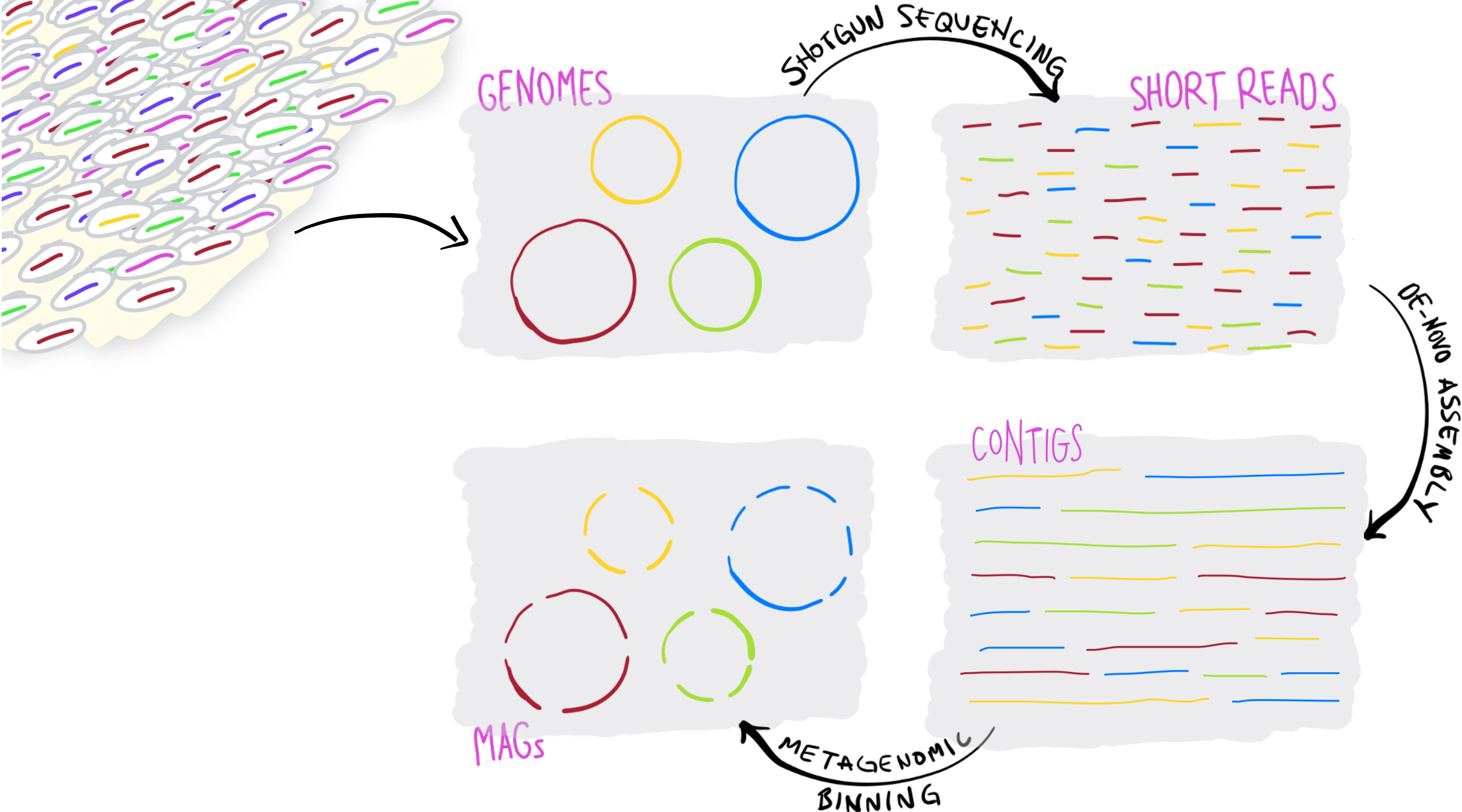


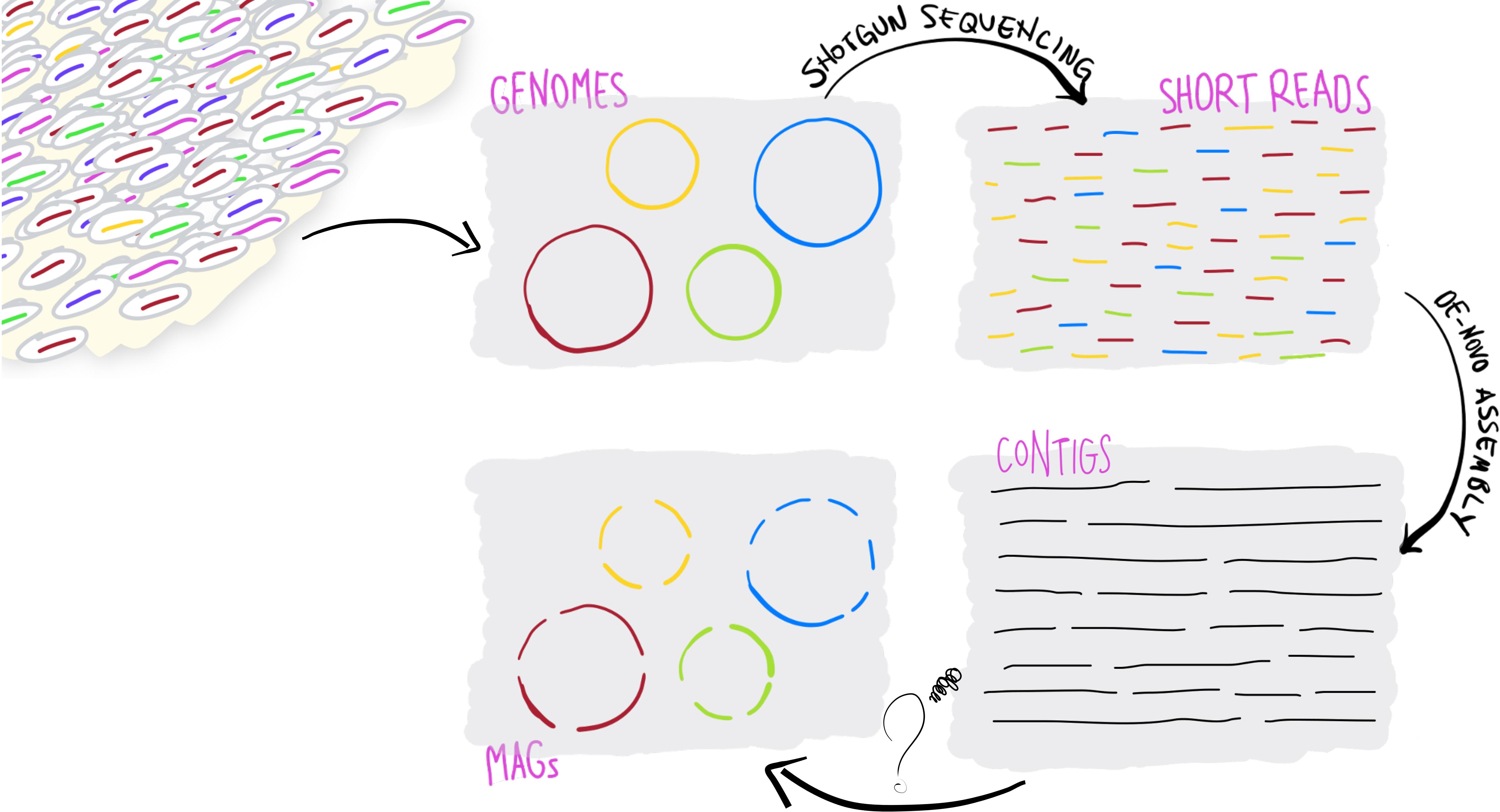
# Metagenomic binning











# Sequence composition

## Computing k-mer frequencies

GTTTGGCATGATTAAGGAGTTCTTTGTGCTTC

k=2

GTTTGGCATGATTAAGGAGTTCTTTGTGCTTC

AA	AC	AG	AT	CA	CC	CG	CT	GA	GC	GG	GT	TA	TC	TG	TT
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

k=2

GT	TTGGCATGATTAAAGGAGTTCTTTGTGCTTC														
AA	AC	AG	AT	CA	CC	CG	CT	GA	GC	GG	GT	TA	TC	TG	TT
0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0

k=2

 TT TGGCATGATTAAGGGAGTTCTTTGTGCTTC

AA	AC	AG	AT	CA	CC	CG	CT	GA	GC	GG	GT	TA	TC	TG	TT	
0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1

k=2

GTTTGGCATGATTAAGGAGTTCTTTGTGCTTC

AA	AC	AG	AT	CA	CC	CG	CT	GA	GC	GG	GT	TA	TC	TG	TT
1	0	2	2	1	0	0	2	2	2	2	3	1	2	4	10

k=2

GTTTGGCATGATTAAGGAGTTCTTTGTGCTTC

AA	AC	AG	AT	CA	CC	CG	CT	GA	GC	GG	GT	TA	TC	TG	TT
1	0	2	2	1	0	0	2	2	2	2	3	1	2	4	10

9

GAAGCACAAAAGAAA**ACTCCTTAATCATGCCAAAAAC**

AA	AC	AG	AT	CA	CC	CG	CT	GA	GC	GG	GT	TA	TC	TG	TT
10	3	2	2	4	2	0	2	2	2	6	0	1	2	1	1

GTTTGGCATGATTAAGGAGTTCTTTGTGCTTC  
GAAGCACAAAAGAAA**ACTCCTTAATCATGCCAAAAAC**

AA	AC	AG	GA	CA	CC	CG	GC	AT	TA	→ PALINDROMES :)				
11	3	4	4	5	2	0	2	2	1					

k=2

GTTTGGCATGATTAAGGAGTTCTTTGTGCTTC

	AA	AC	AG	GA	CA	CC	CG	GC	AT	TA
X	11	3	4	4	5	2	0	2	2	1
Y										
Z										
L										
K										
M										

k=2

ACTTCCGCAGTCGGGCATTACGCGTTGTGGAATGA

	AA	AC	AG	GA	CA	CC	CG	GC	AT	TA
X	11	3	4	4	5	2	0	2	2	1
Y	4	5	2	4	5	4	4	3	2	1
Z										
L										
K										
M										

k=2

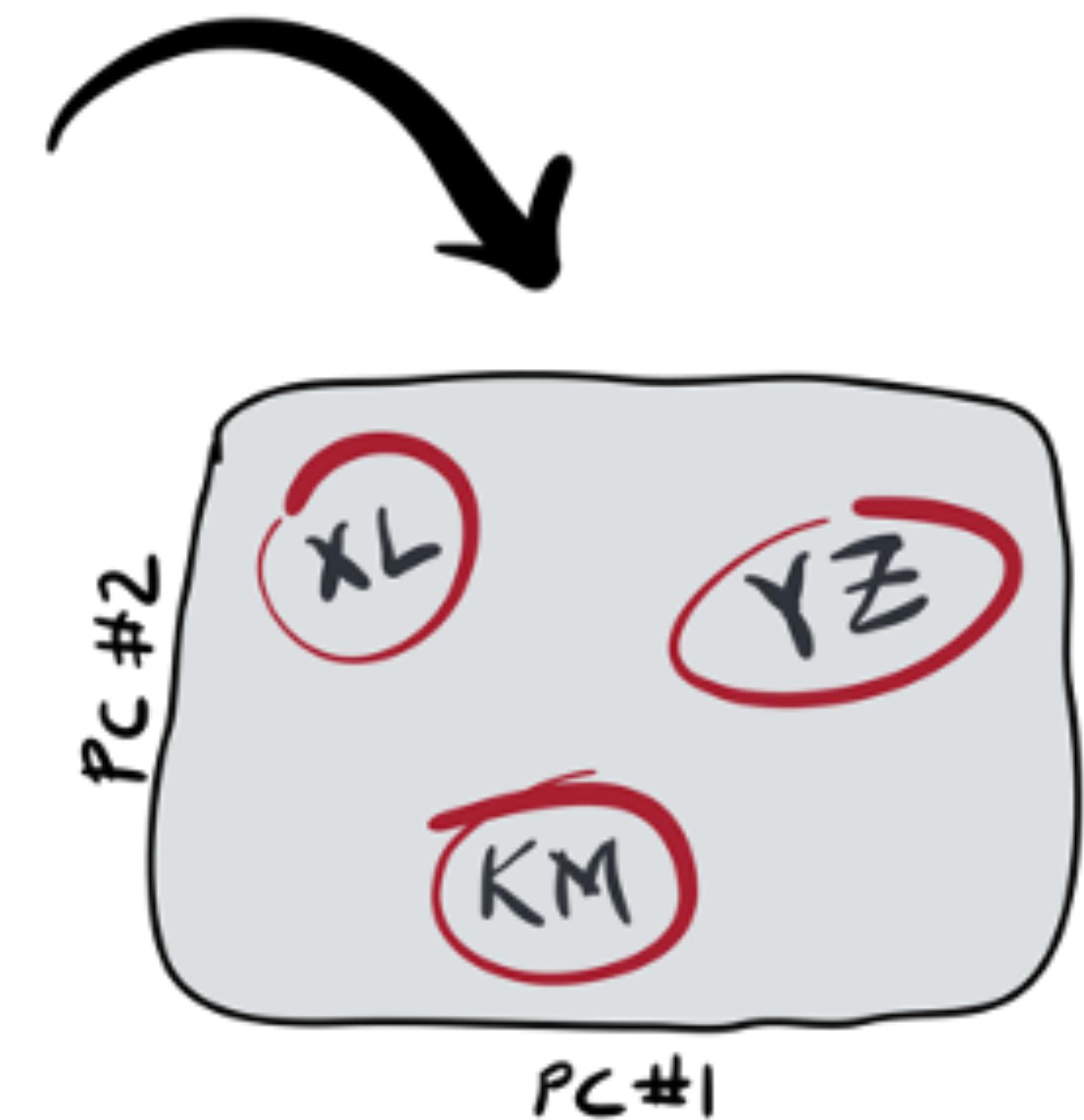
GGGCCCTGCGCCGGTCCAGTCACCCGGCTGCGACCT

	AA	AC	AG	GA	CA	CC	CG	GC	AT	TA
X	11	3	4	4	5	2	0	2	2	1
Y	4	5	2	4	5	4	4	3	2	1
Z	4	5	3	2	4	1	5	5	2	3
L	11	6	3	2	2	3	2	1	1	4
K	1	1	2	2	1	8	9	10	0	0
M	0	4	4	3	4	10	4	5	0	0

k=2

	AA	AC	AG	GA	CA	CC	CG	GC	AT	TA
X	11	3	4	4	5	2	0	2	2	1
Y	4	5	2	4	5	4	4	3	2	1
Z	4	5	3	2	4	1	5	5	2	3
L	11	6	3	2	2	3	2	1	1	4
K	1	1	2	2	1	8	9	10	0	0
M	0	4	4	3	4	10	4	5	0	0

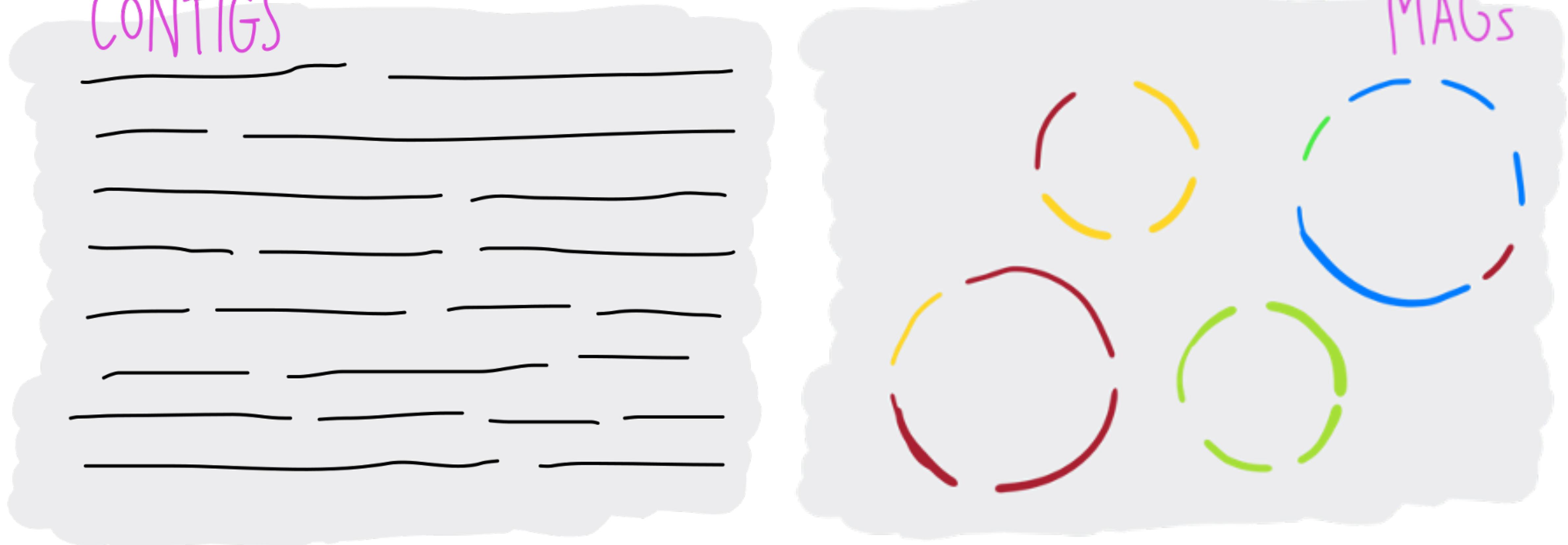
k=2



SEQUENCE COMPOSITION

CONTIGS

MAGs



# Abundance correlation

Counting stuff across samples

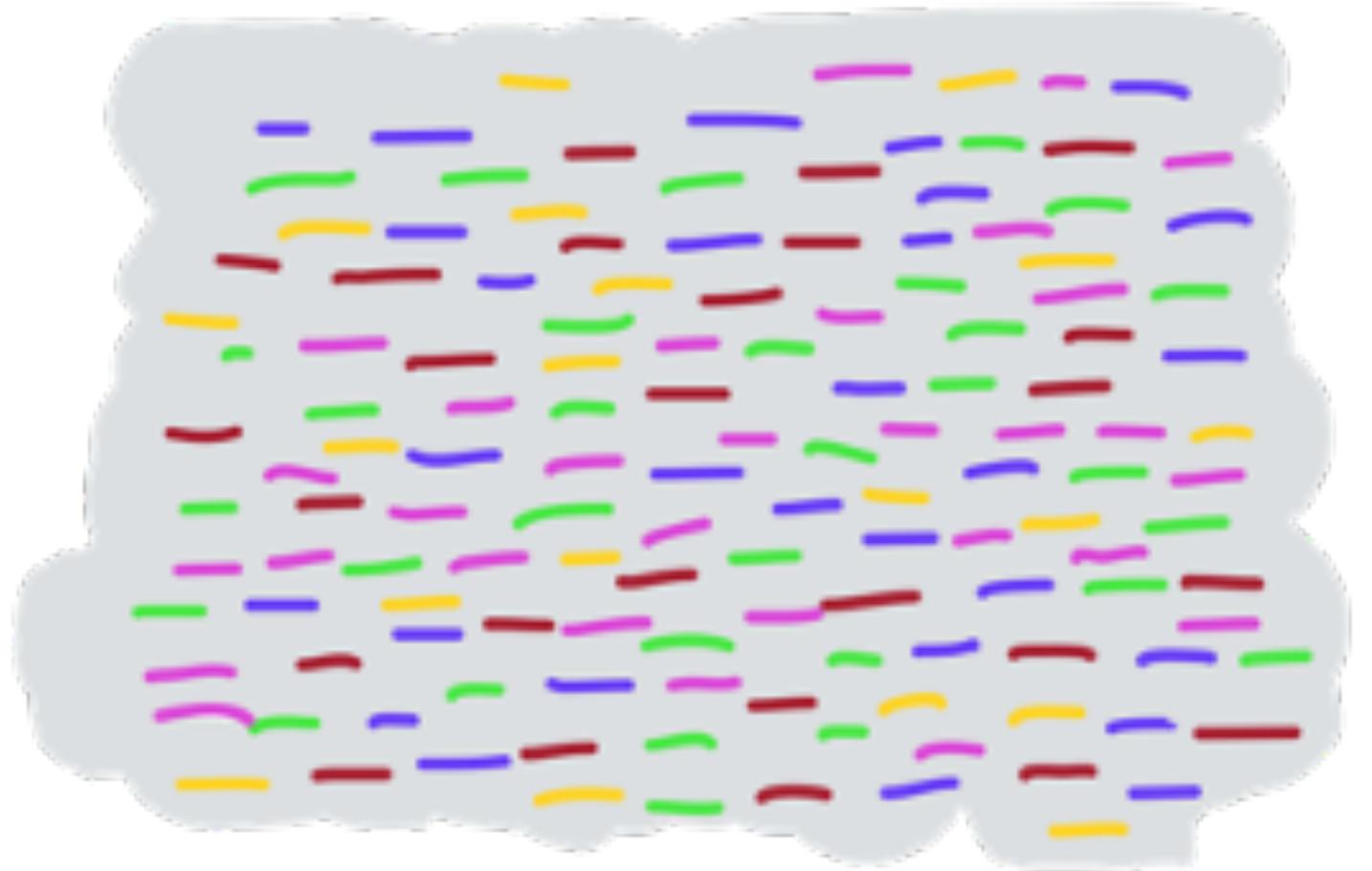
CONTIG #1

CONTIG #2

CONTIG #1



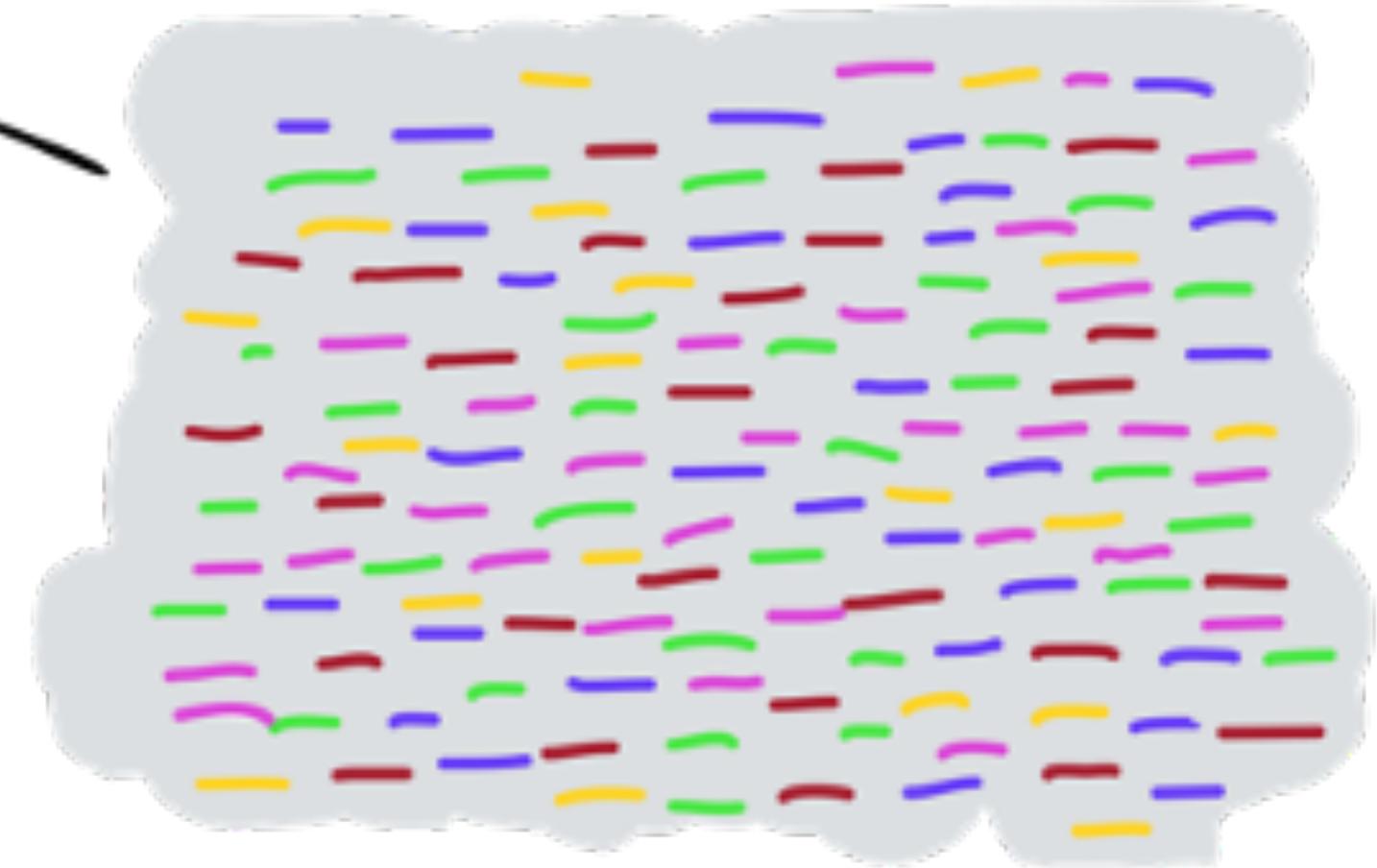
CONTIG #2



METAGENOMIC READS

CONTIG #1

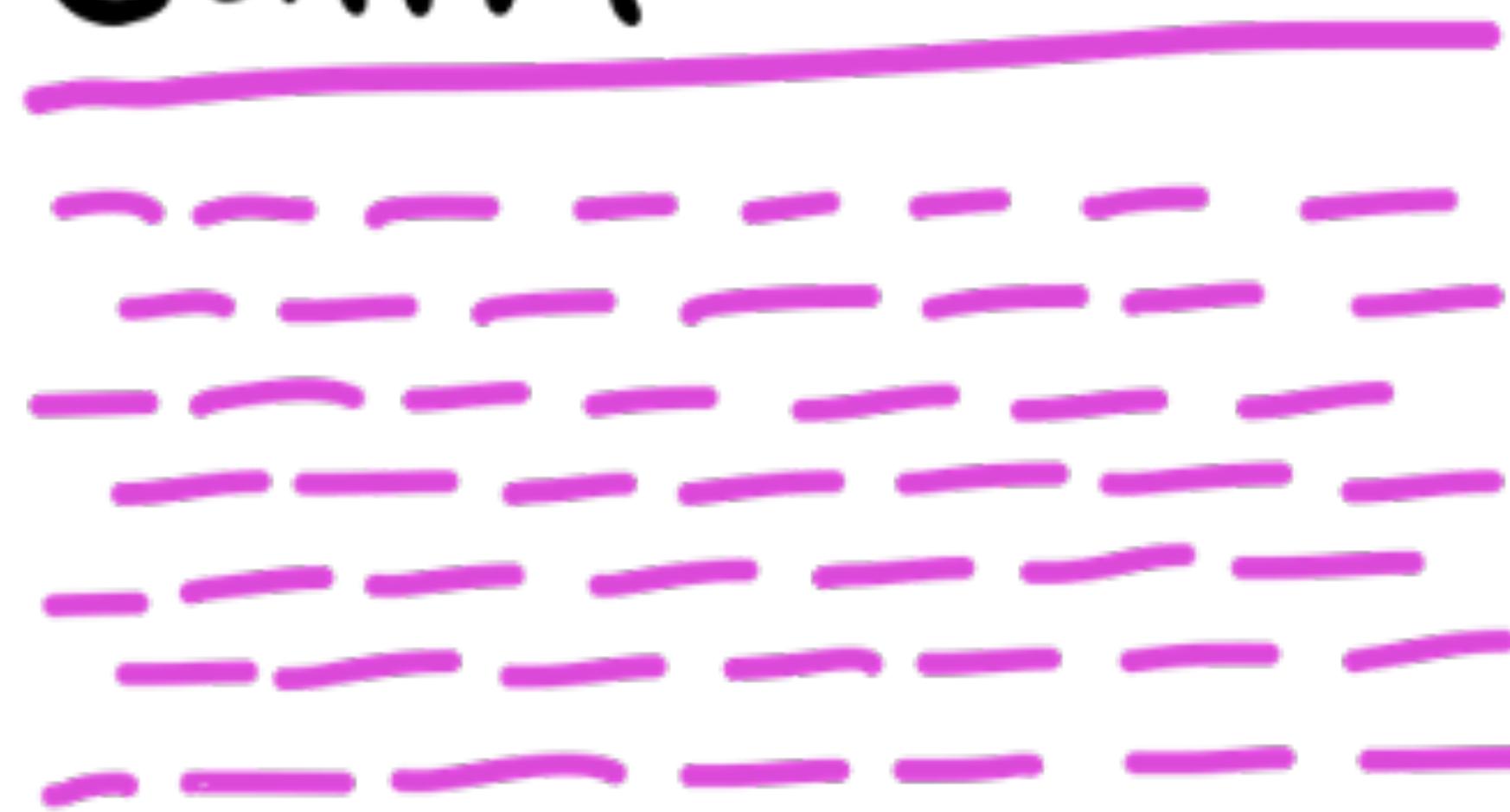
MAPPING



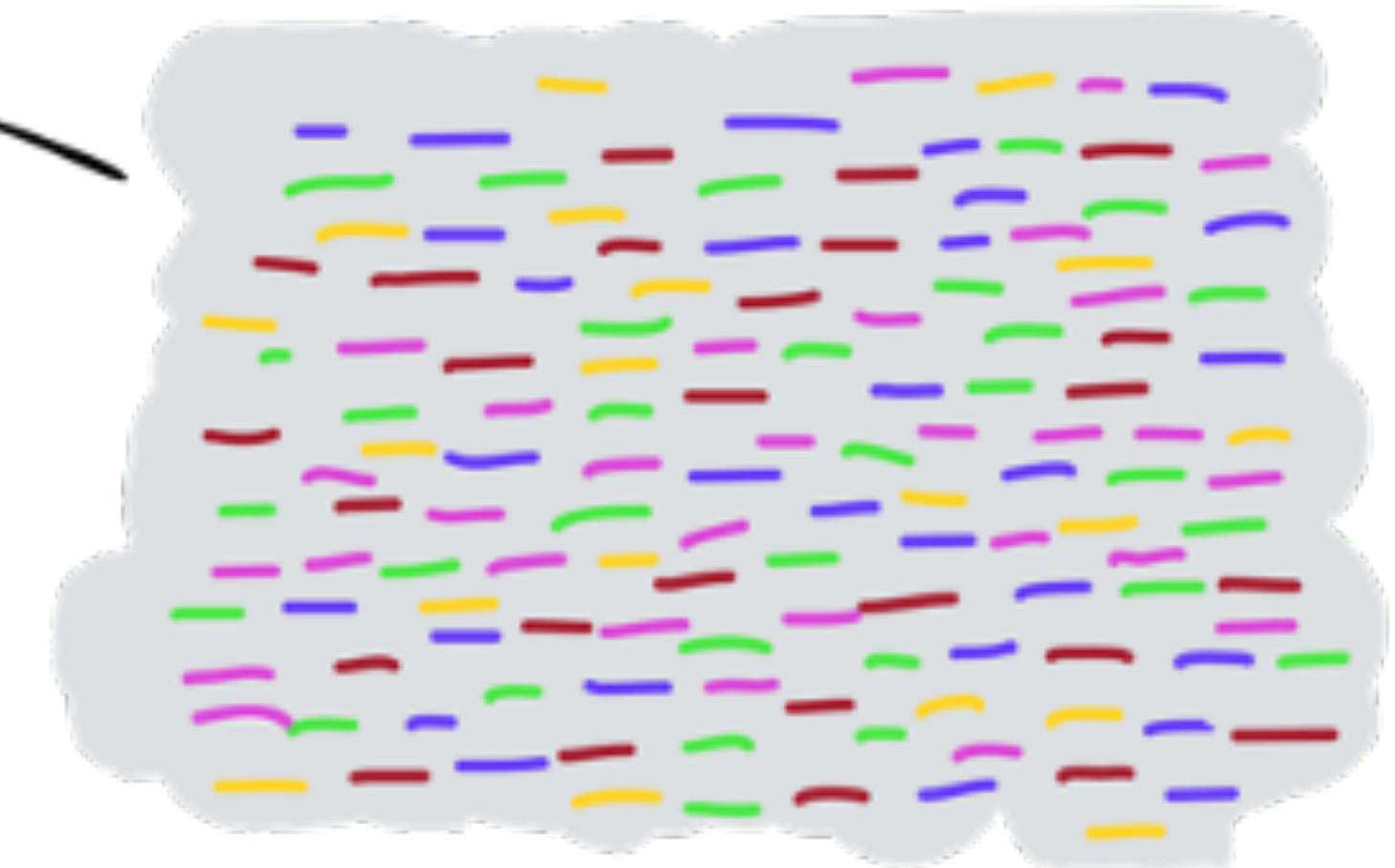
METAGENOMIC READS

CONTIG #2

CONTIG #1



MAPPING

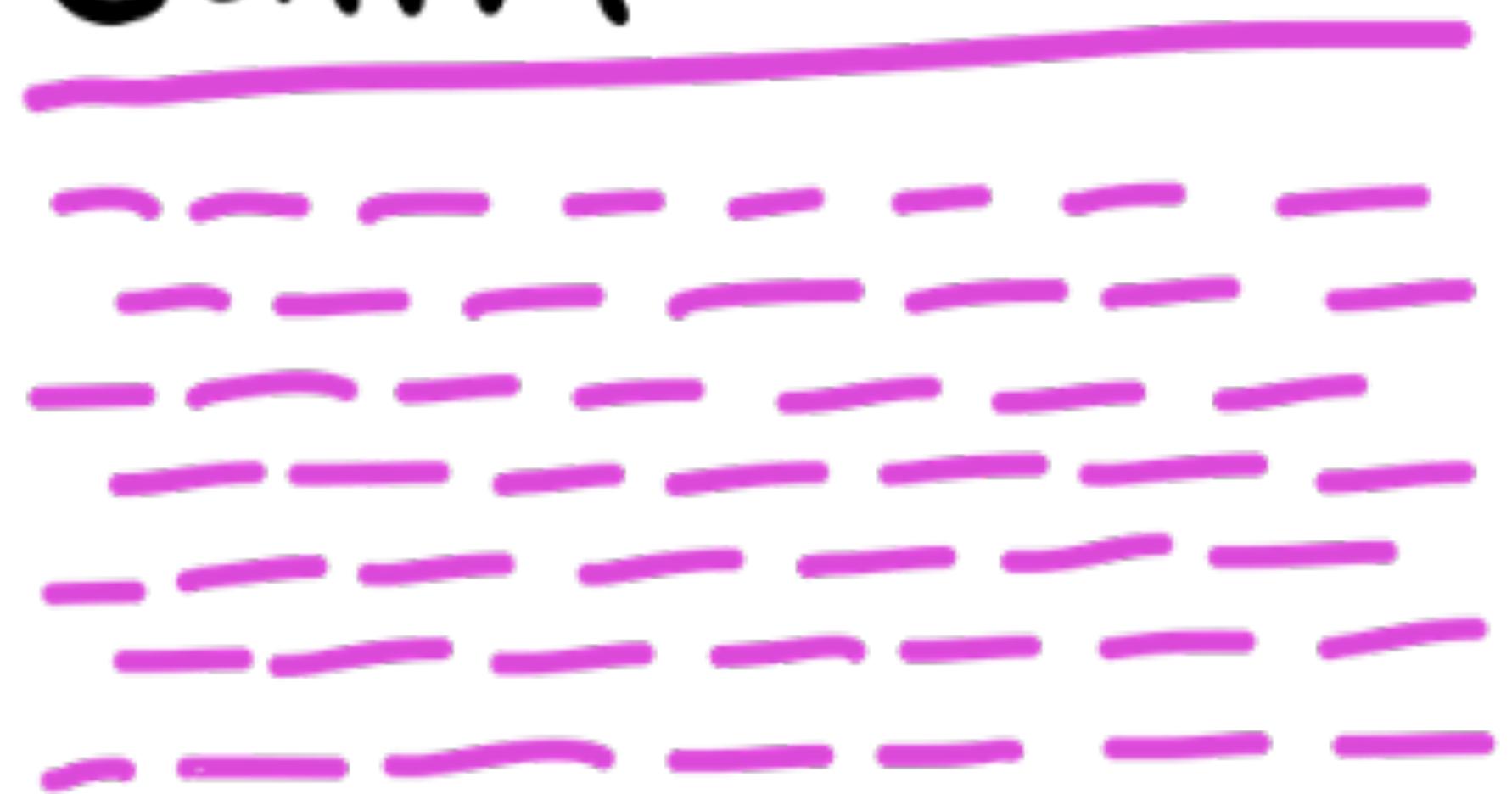


METAGENOMIC READS

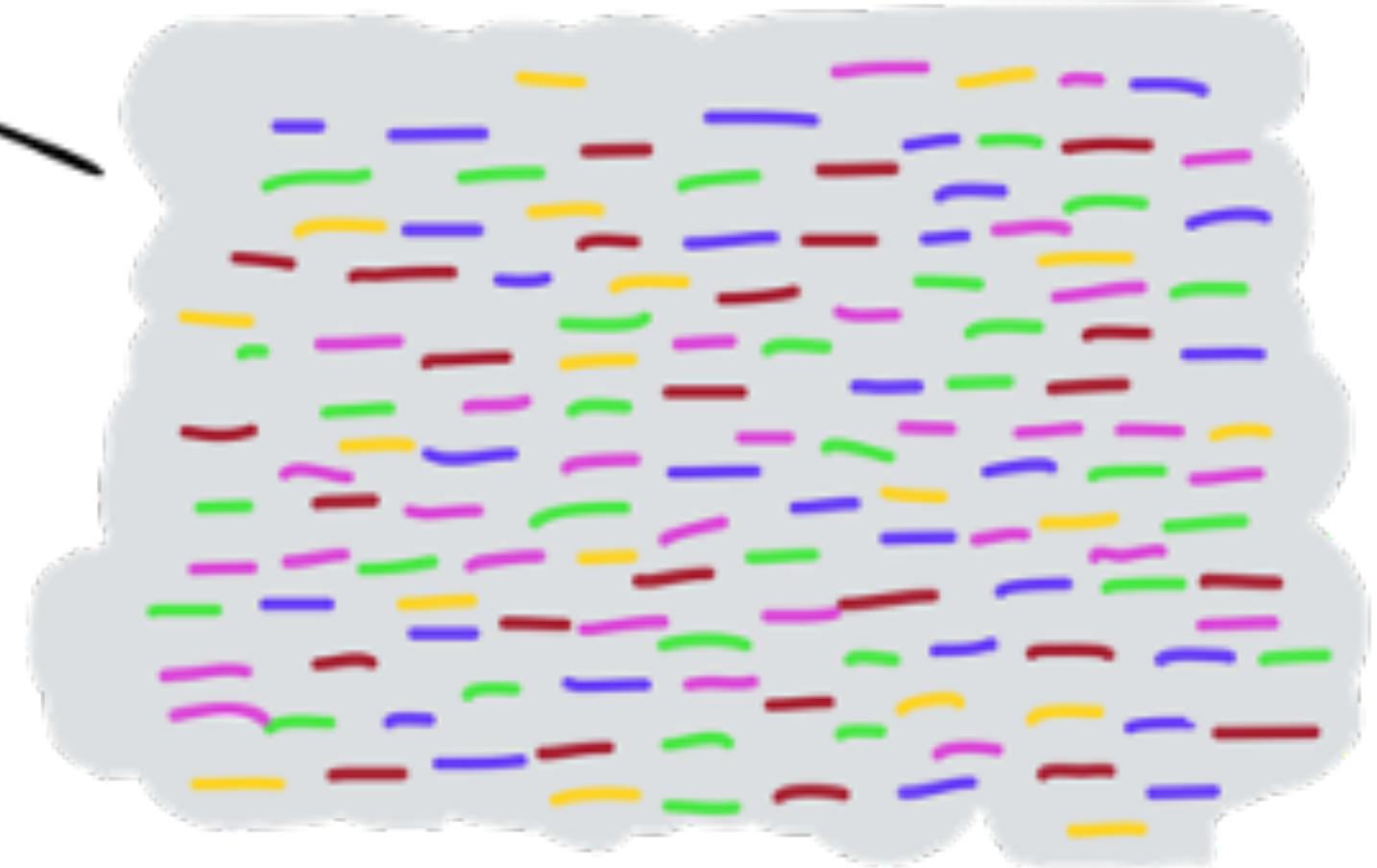
CONTIG #2



**CONTIG #1**



MAPPING

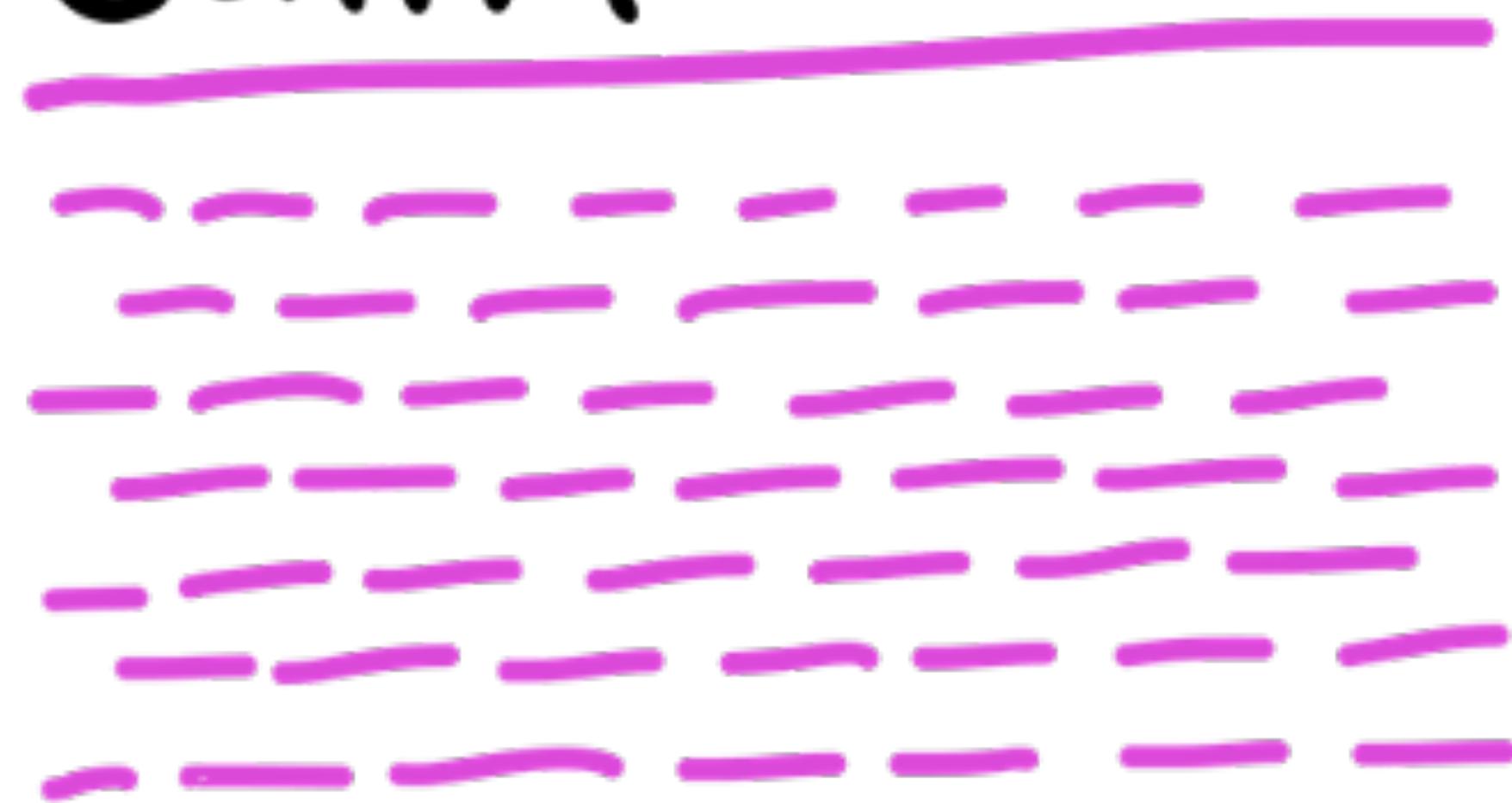


**METAGENOMIC READS**

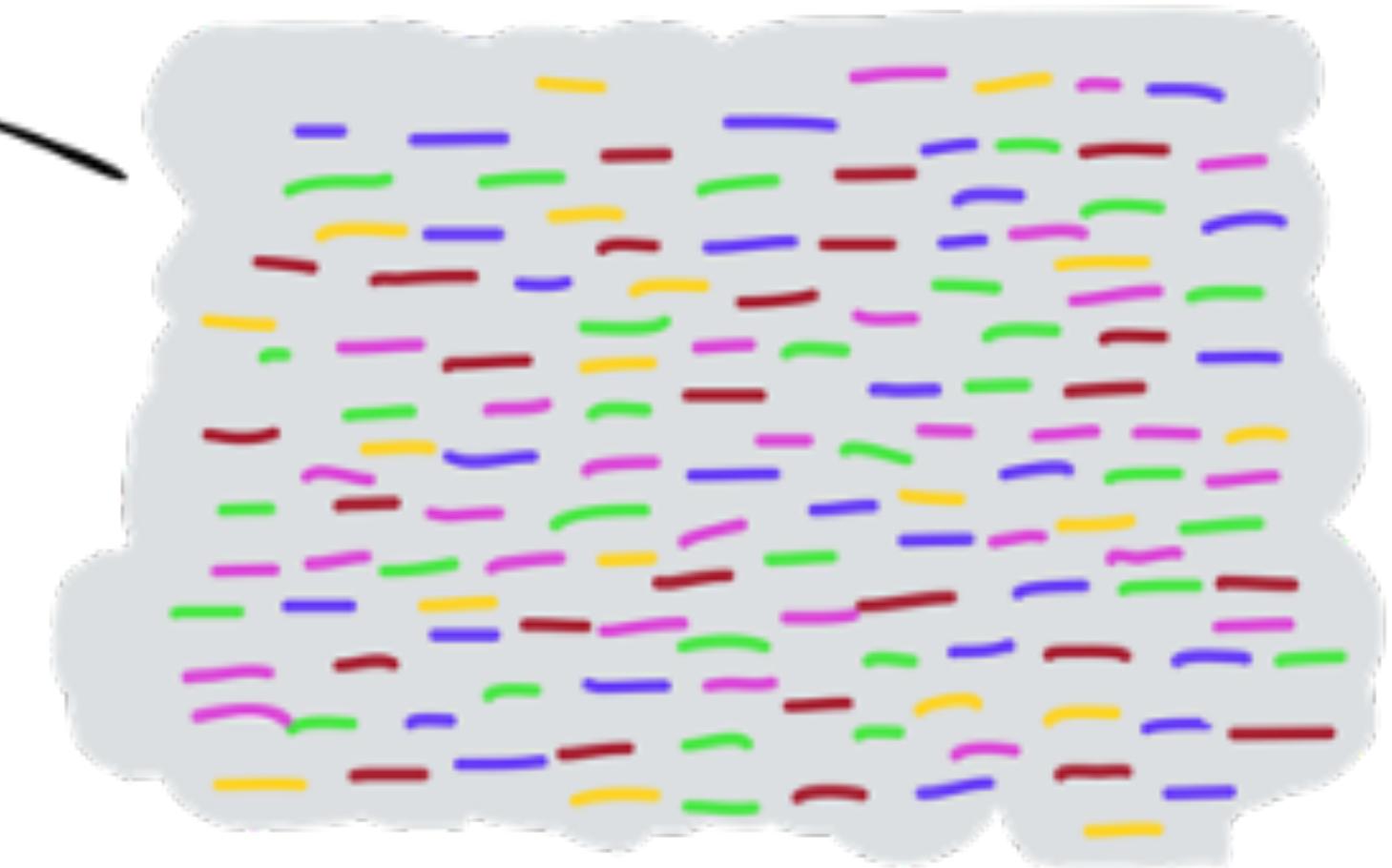
**CONTIG #2**



CONTIG #1



MAPPING



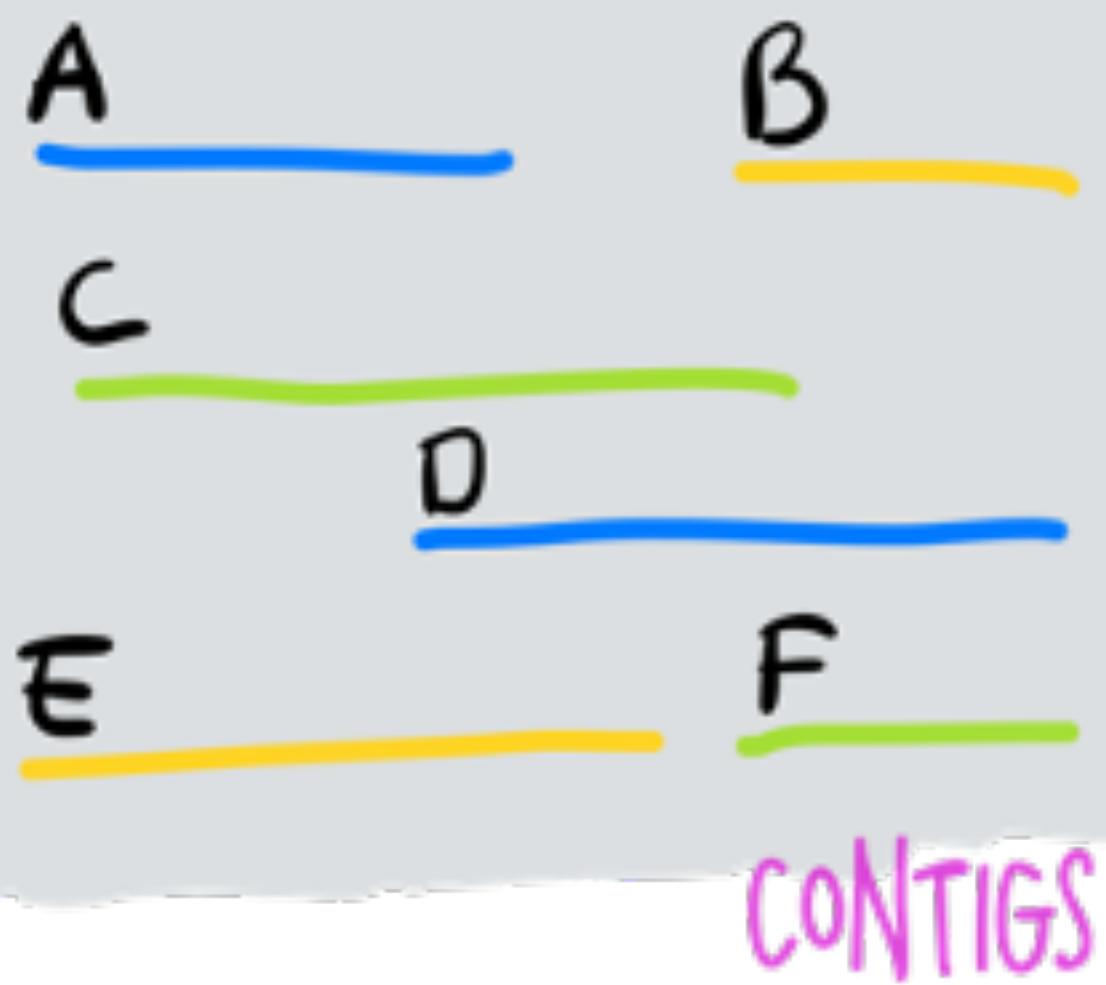
METAGENOMIC READS

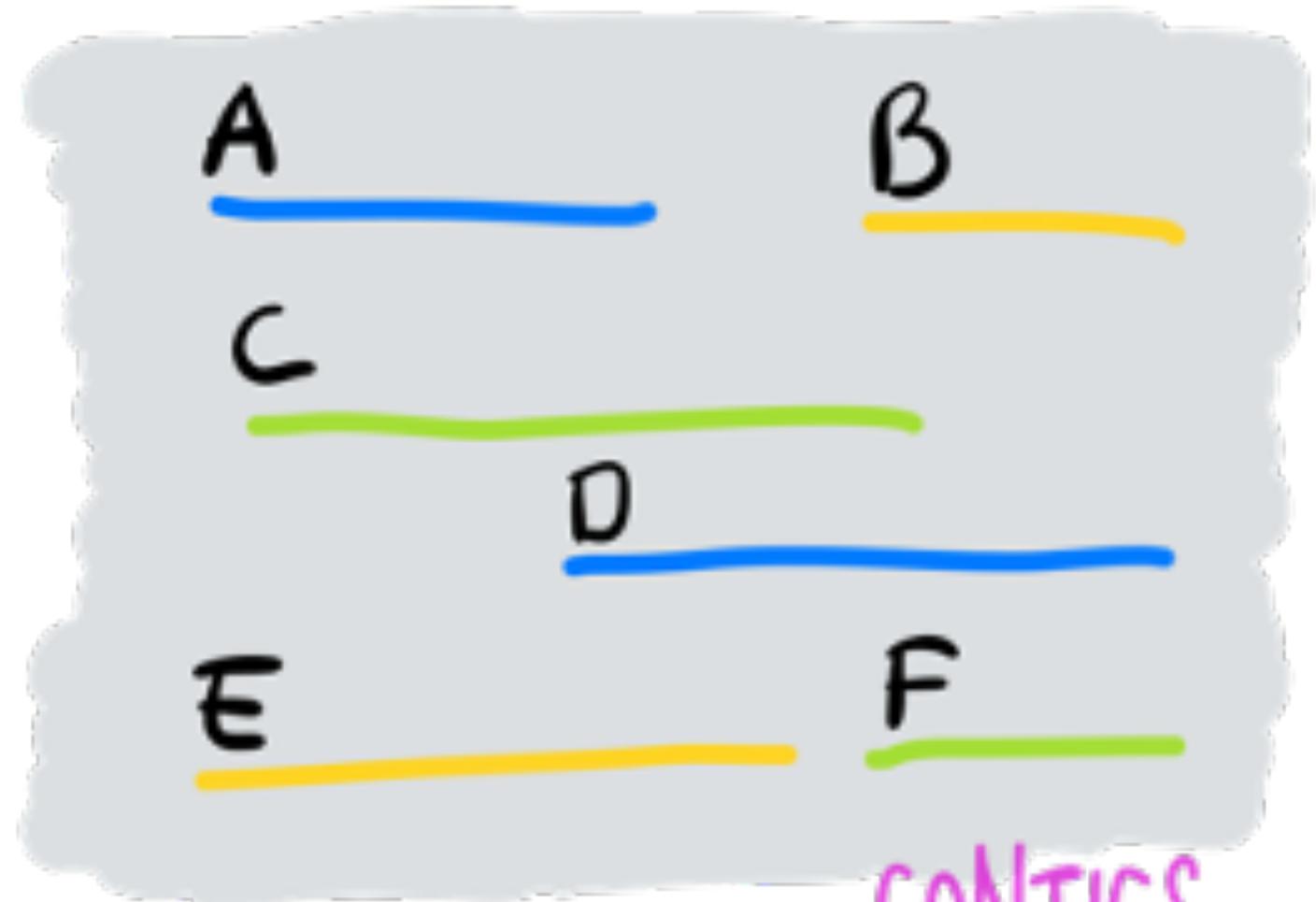
COVERAGE : ~7X

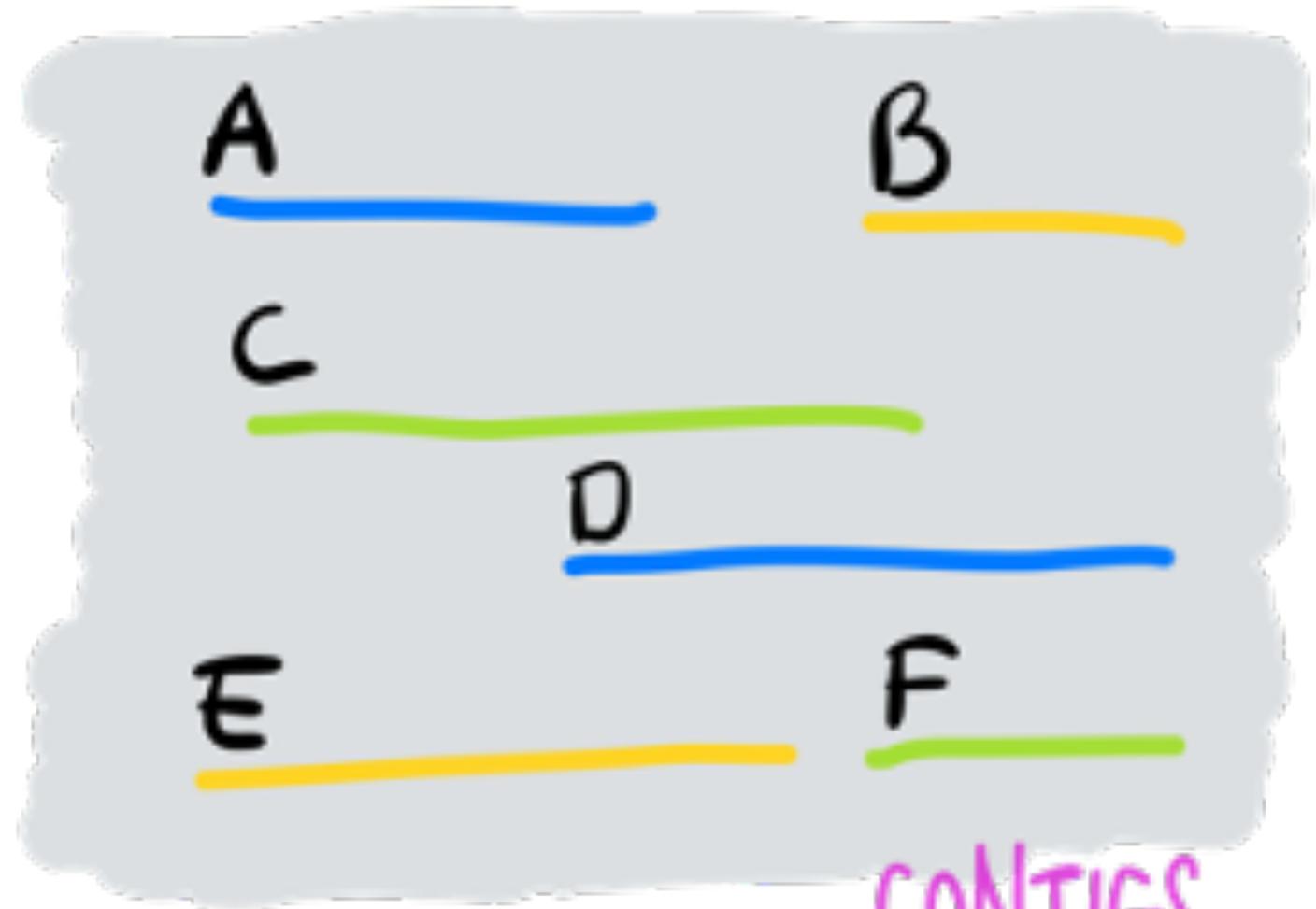
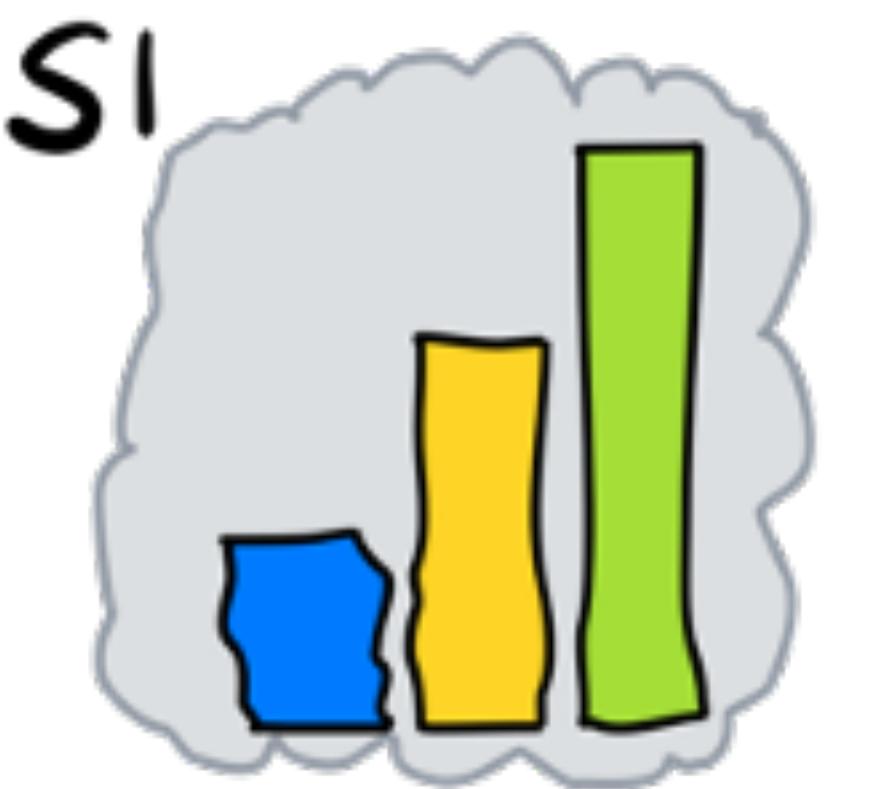
CONTIG #2

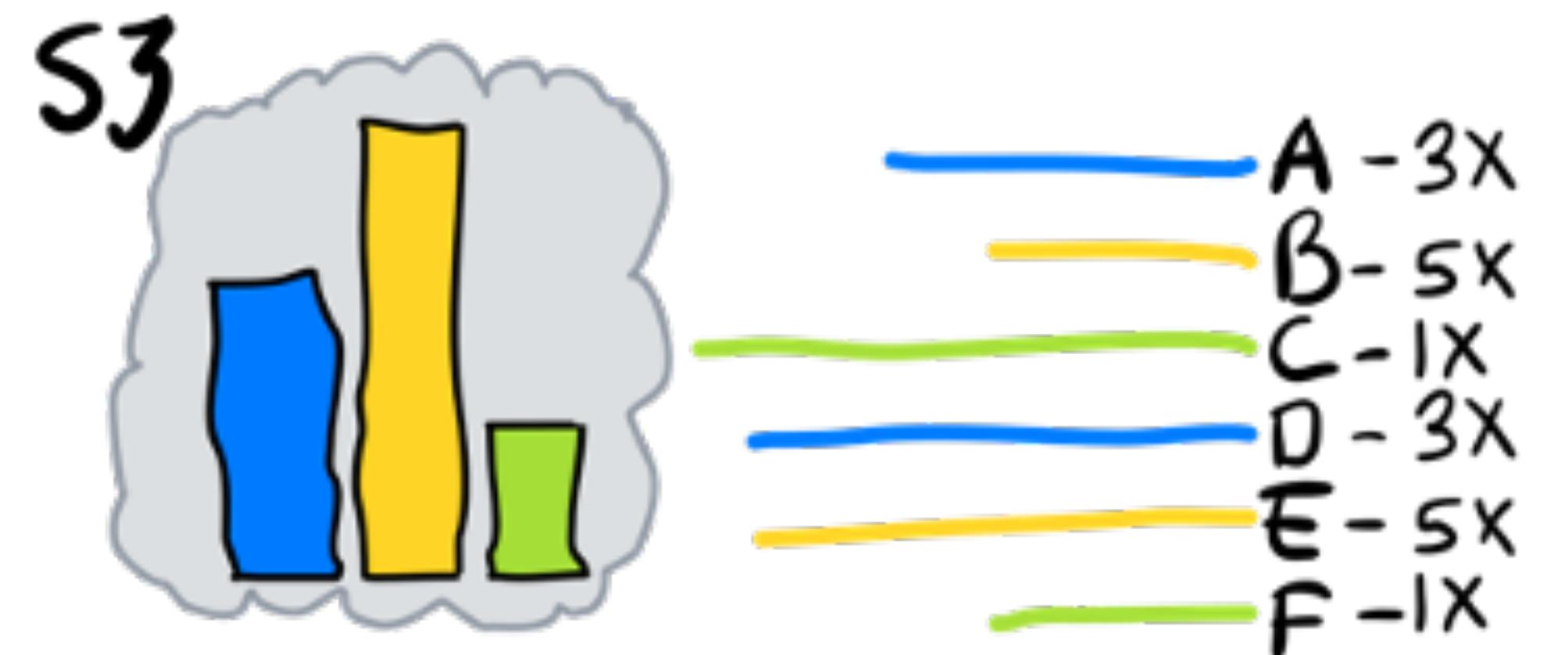
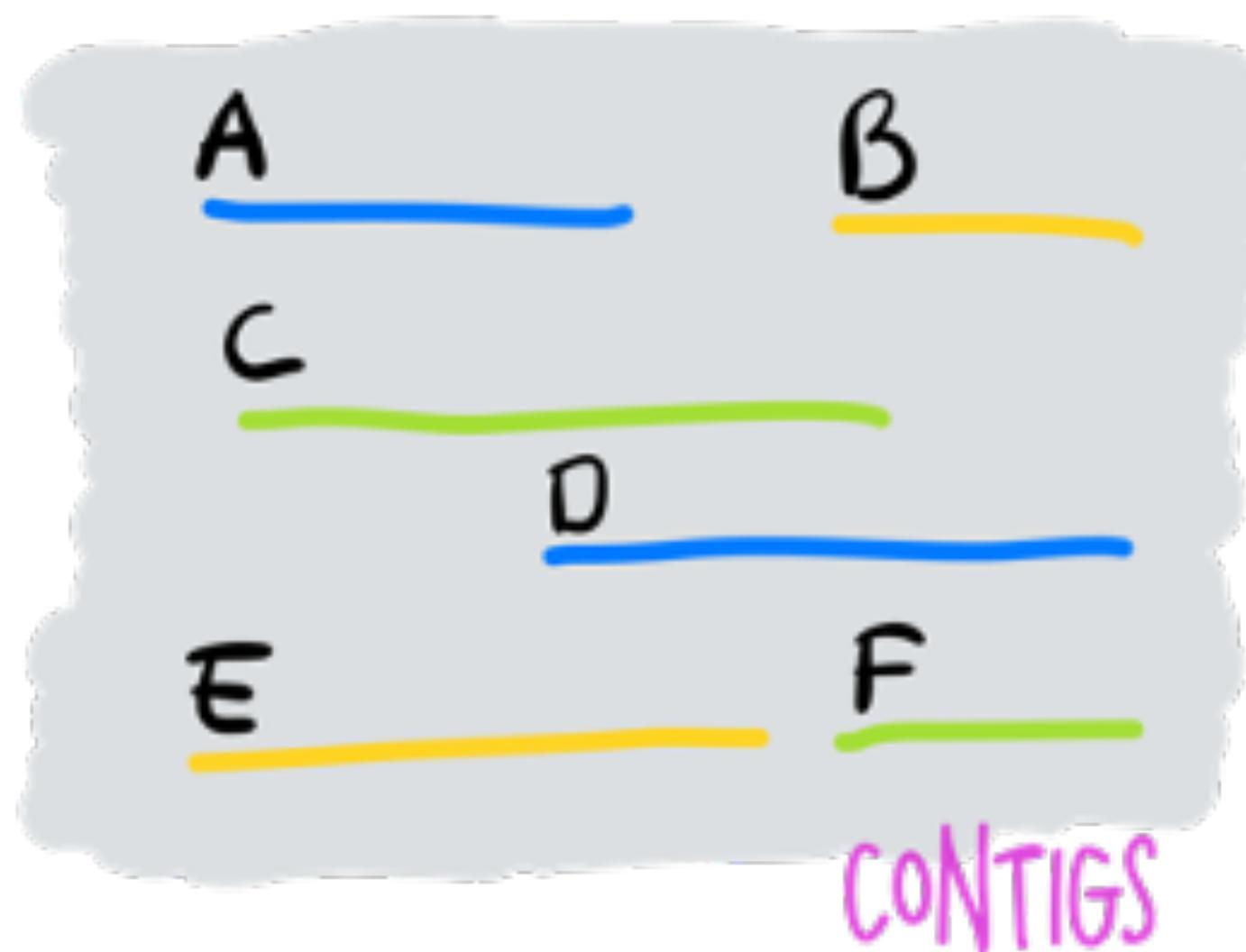
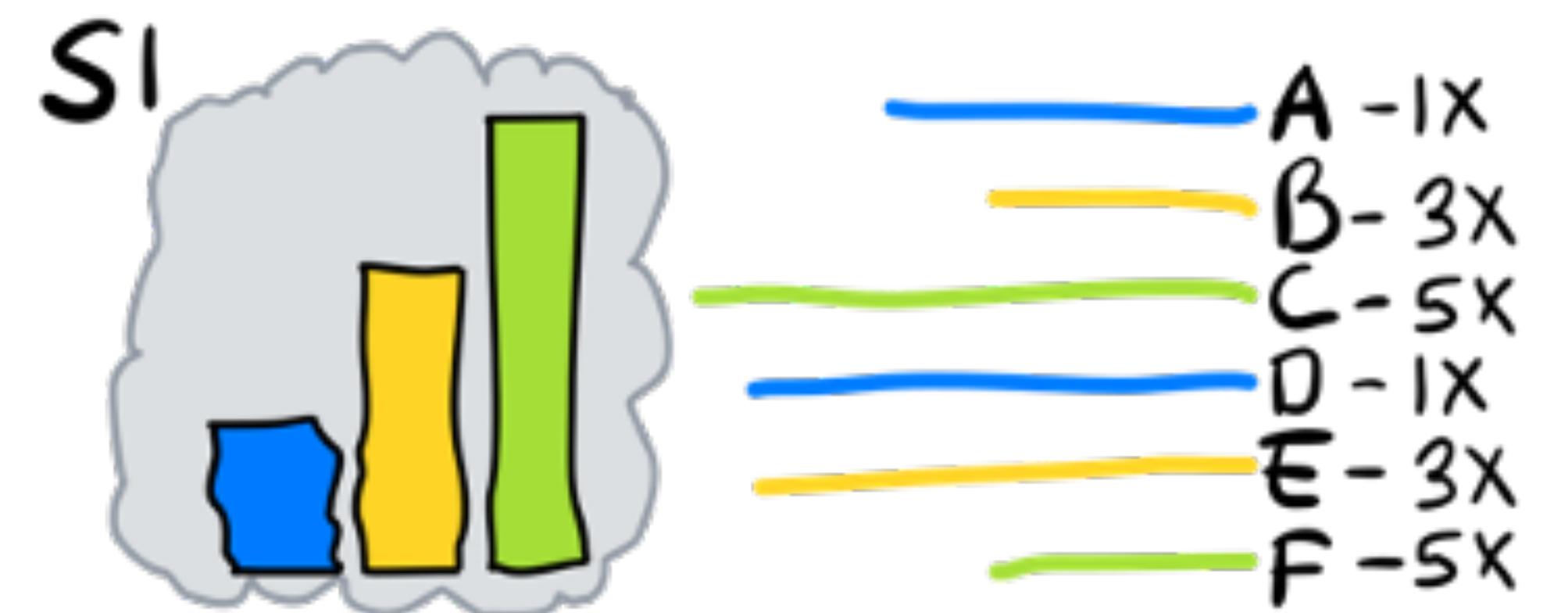


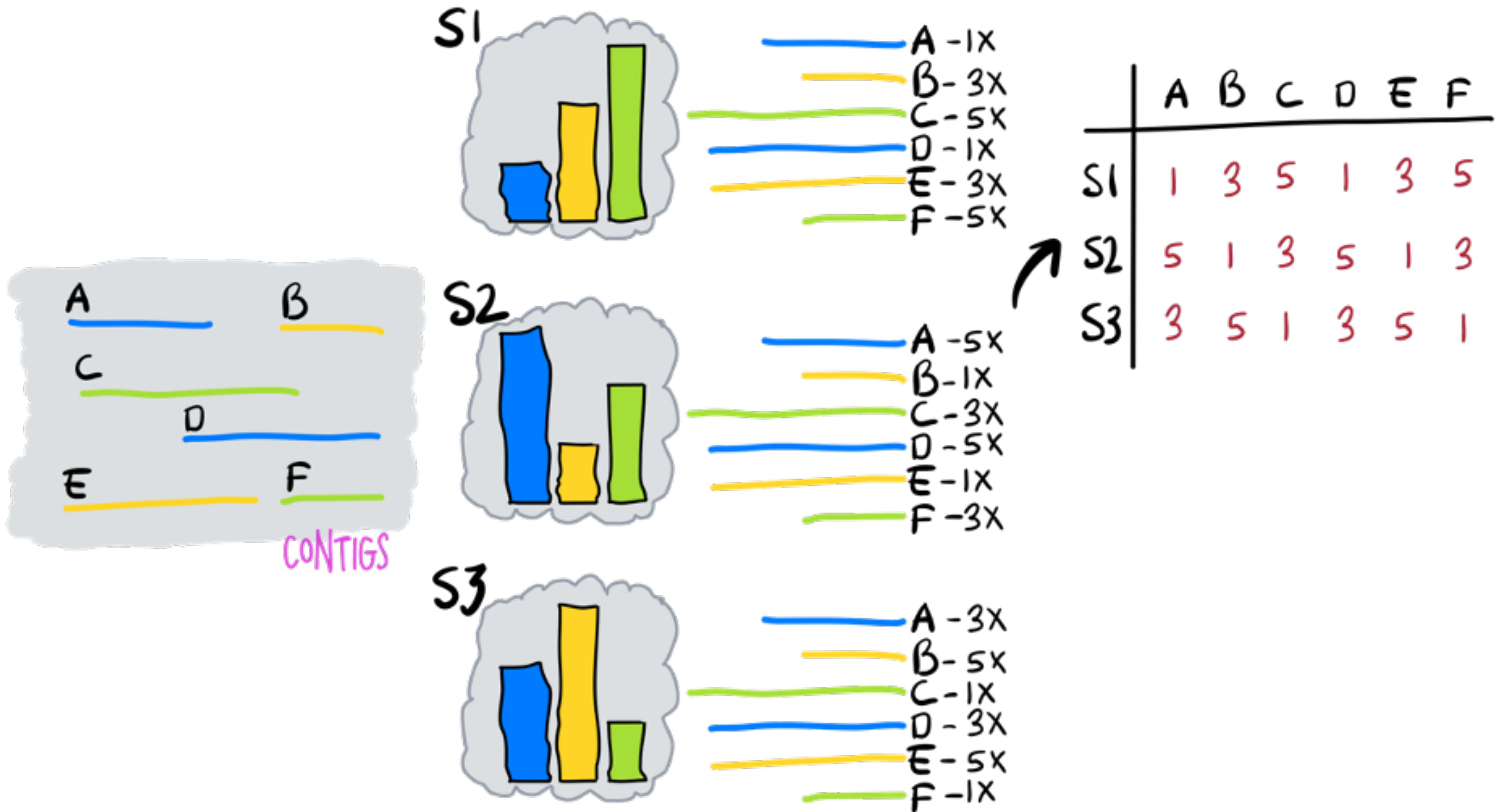
COVERAGE : ~4X

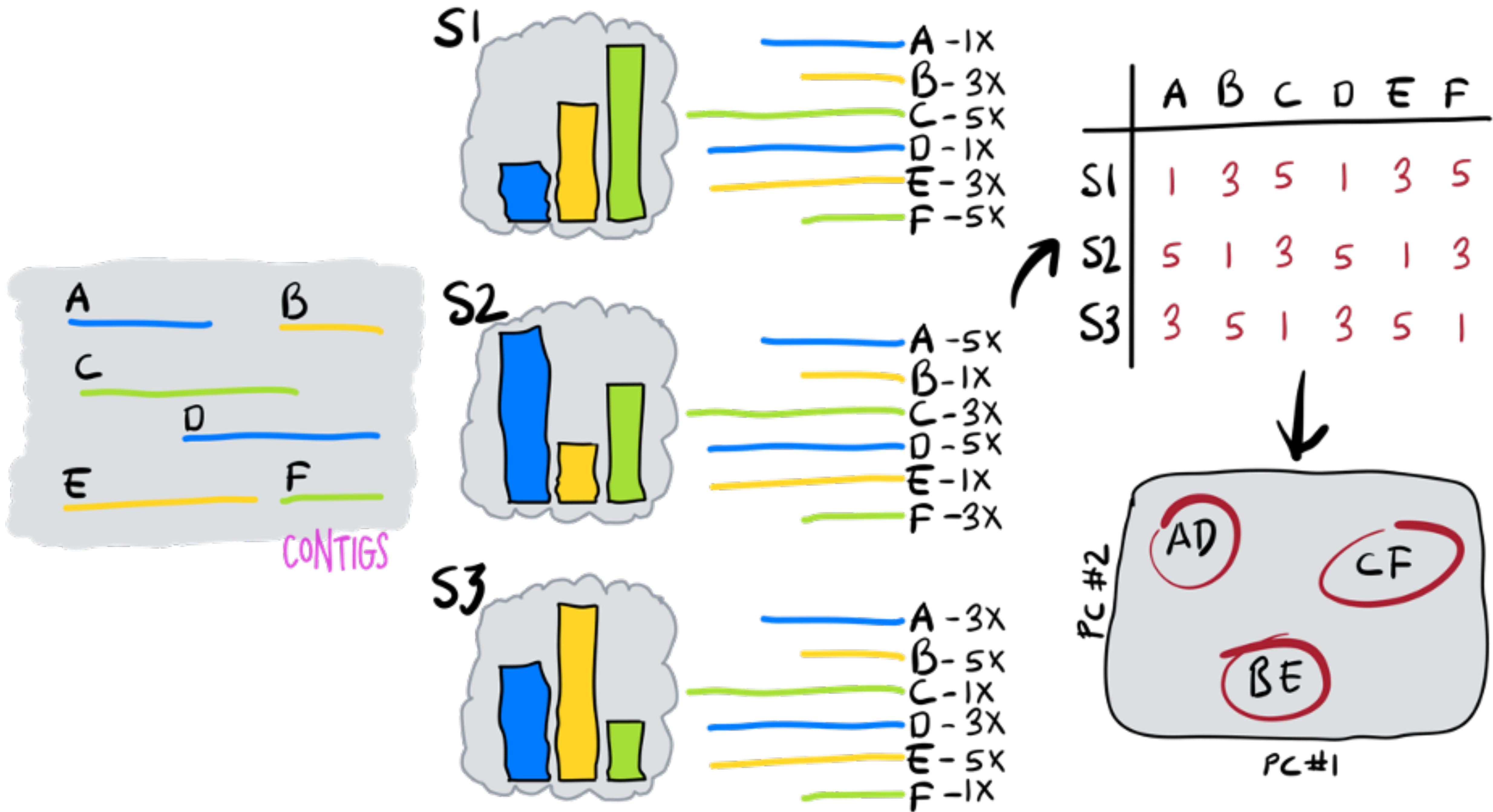






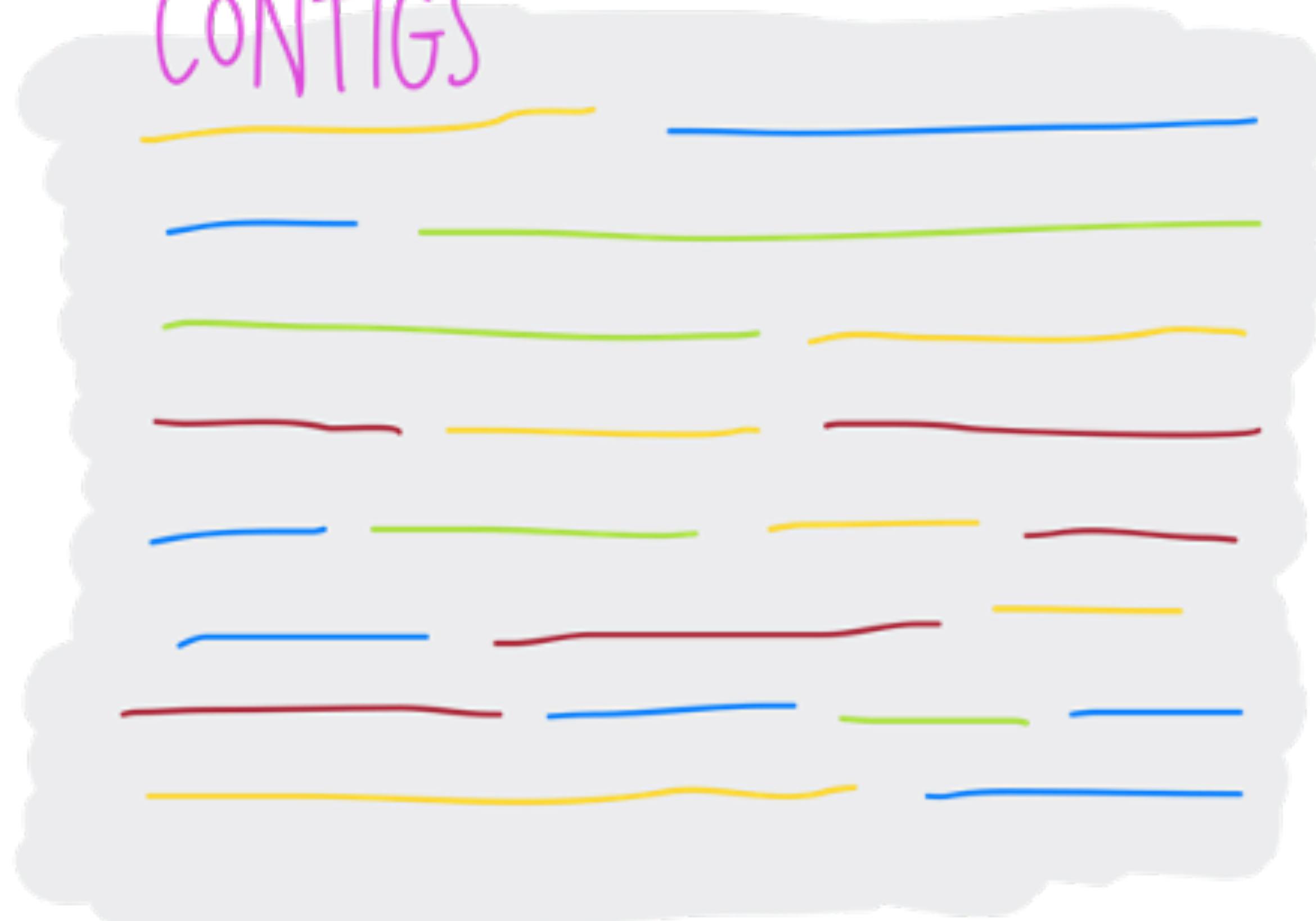




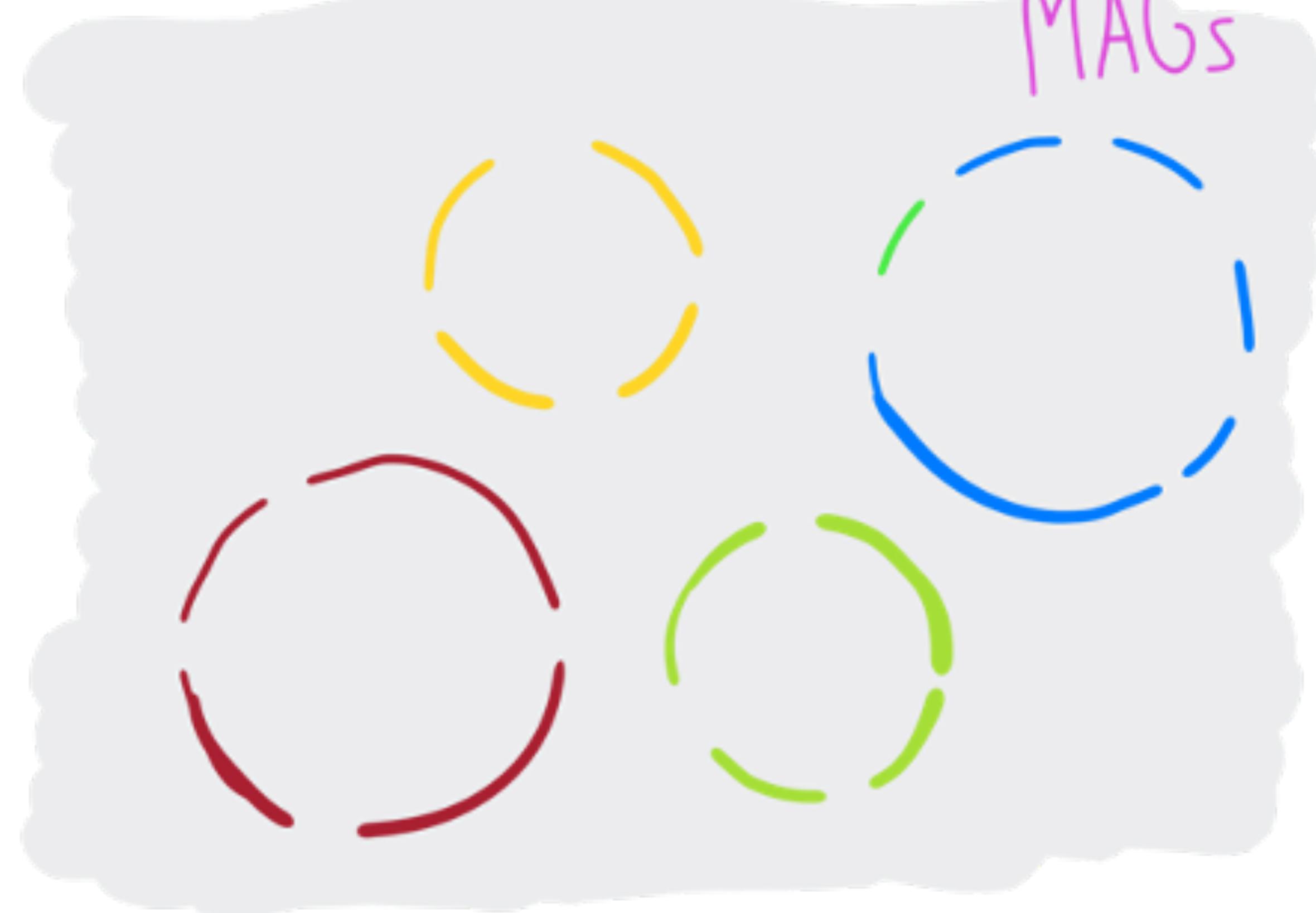


SEQUENCE COMPOSITION

CONTIGS



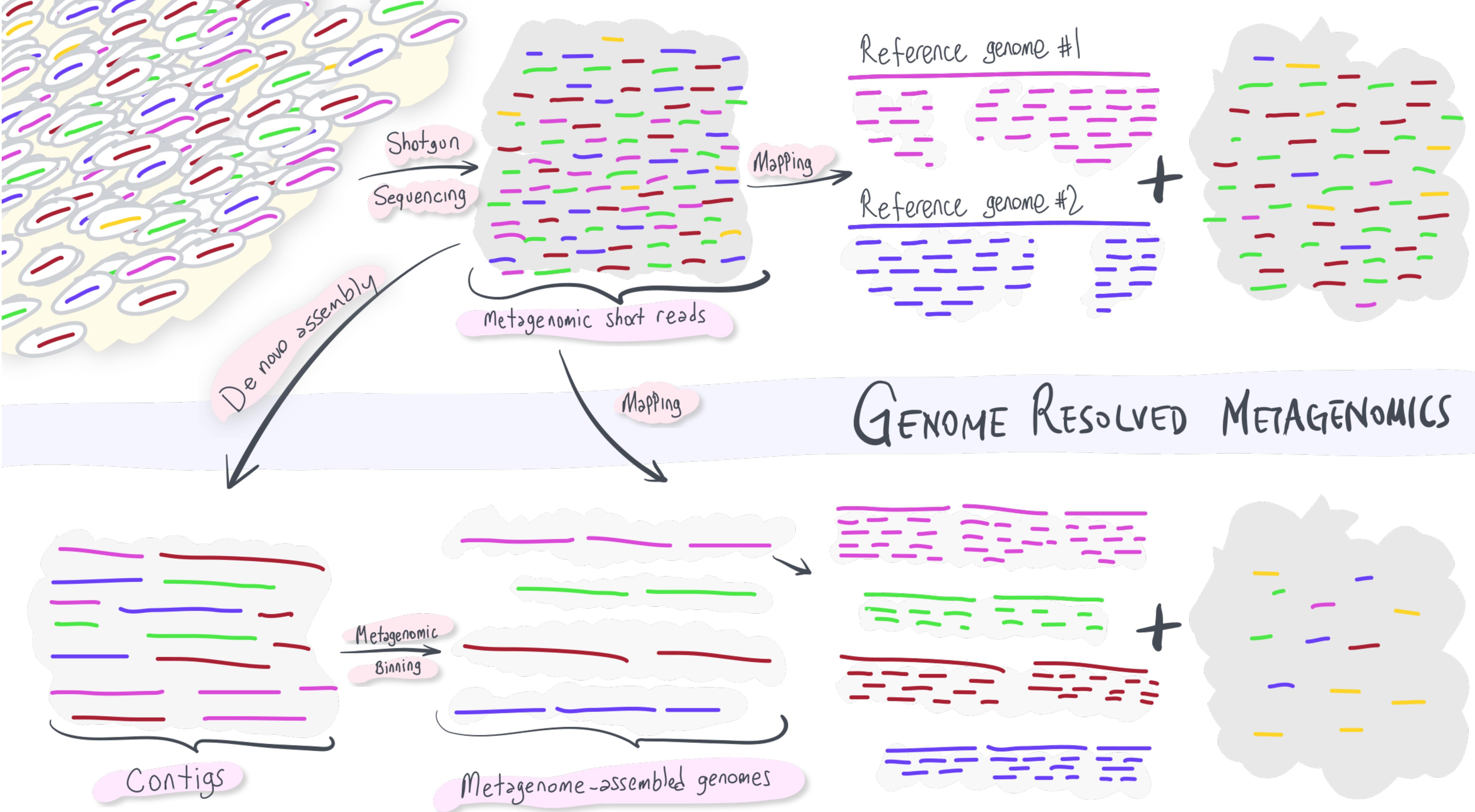
MAGs



DIFFERENTIAL COVERAGE

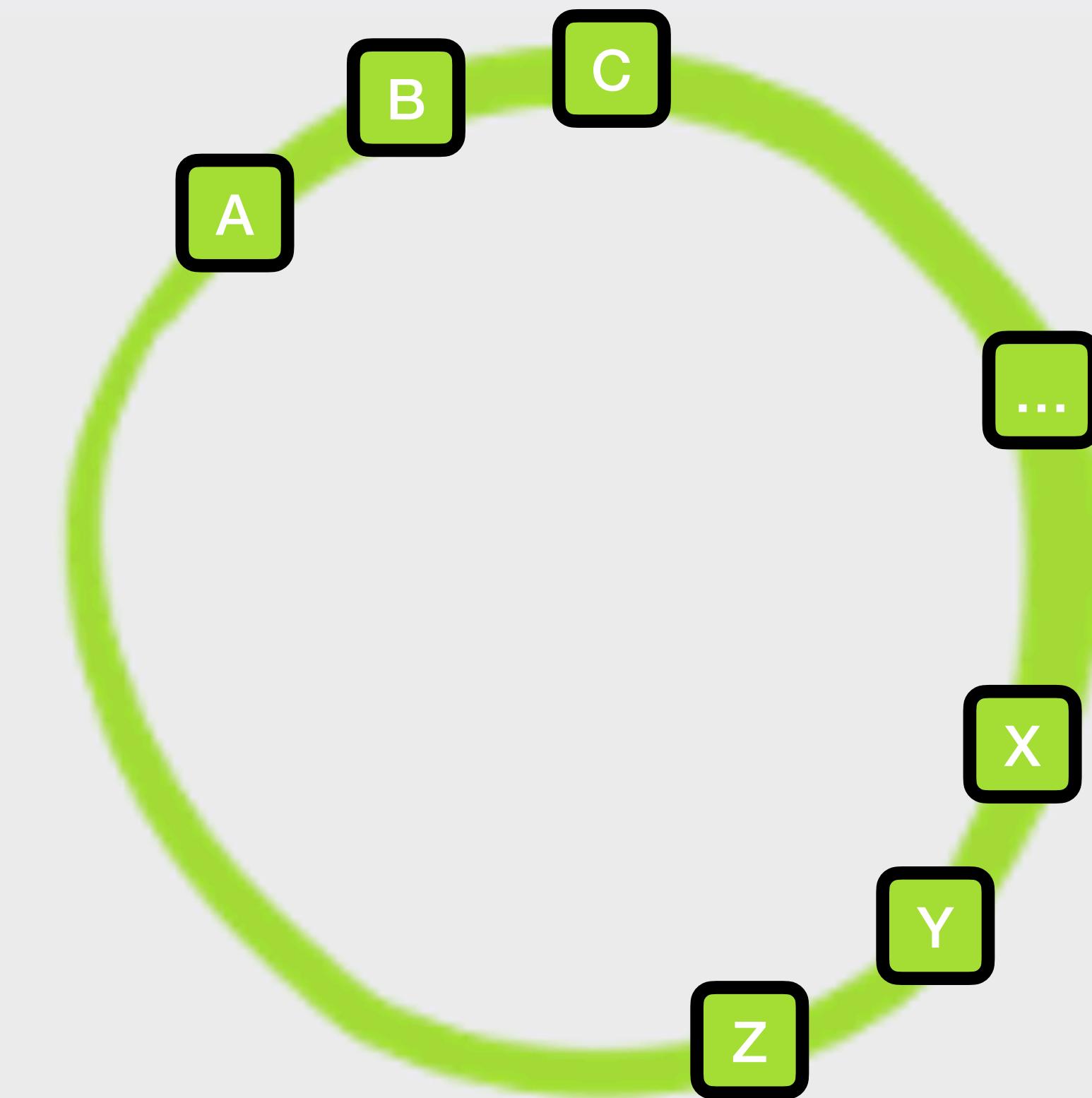
# Cultivation-independent genome-resolved metagenomics

## A summary

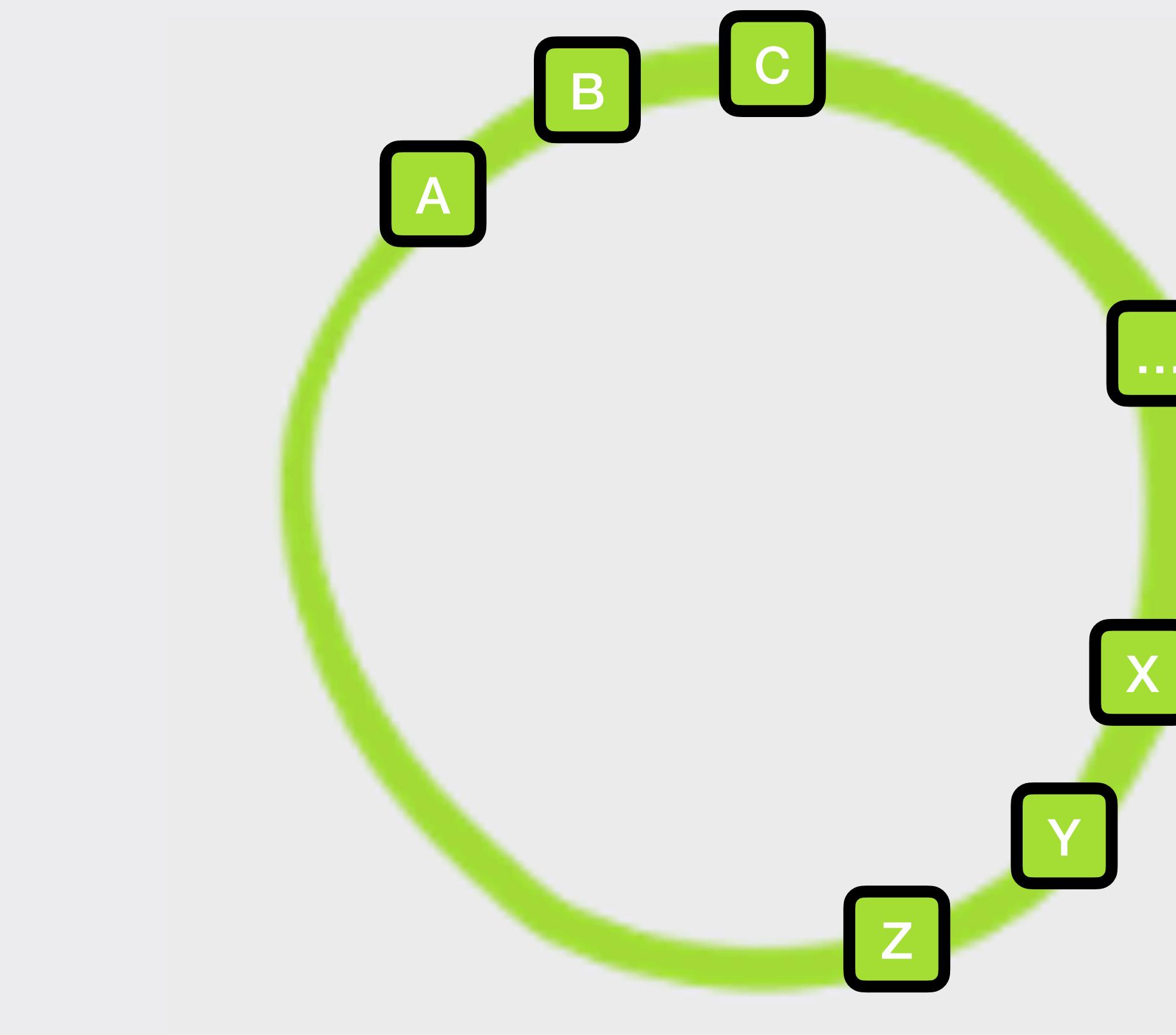


# Evaluation and limits of MAGs

# Universal single-copy marker genes

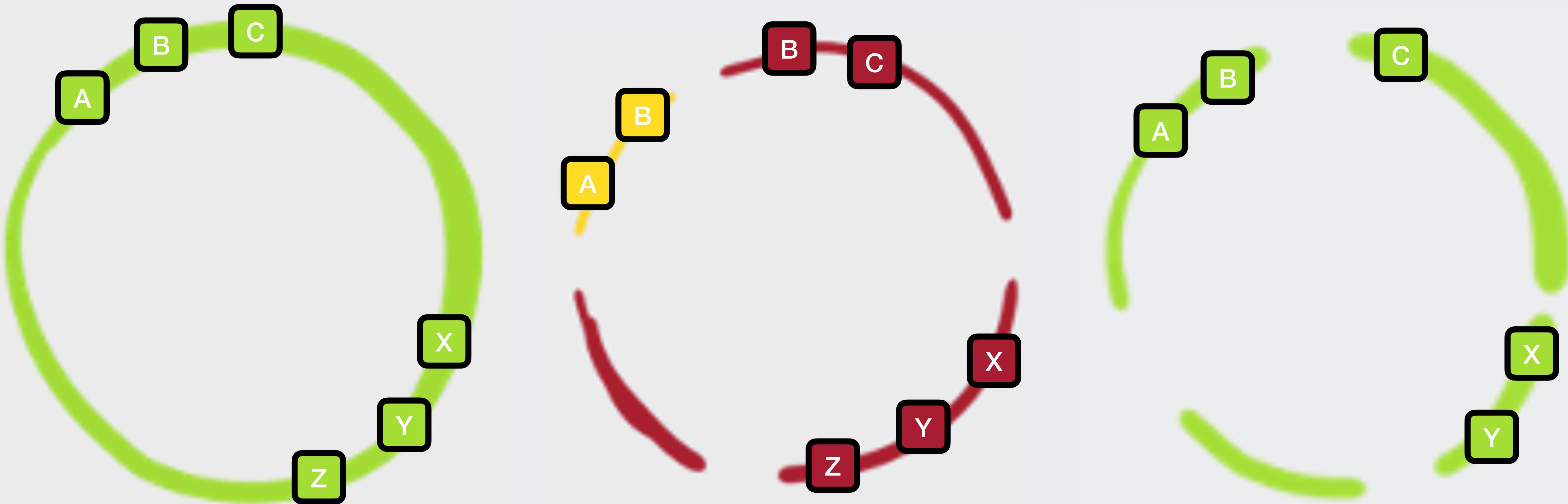


# Universal single-copy marker genes

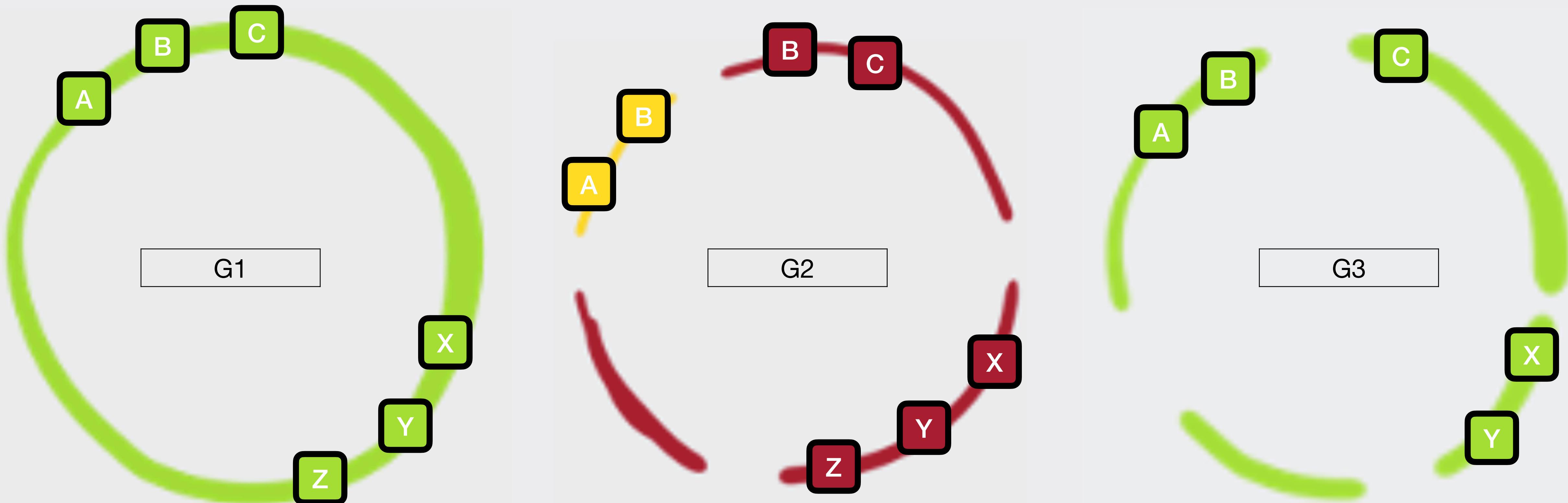


Between 40 and 120 for Bacteria/Archaea depending on cutoffs

# Universal single-copy marker genes

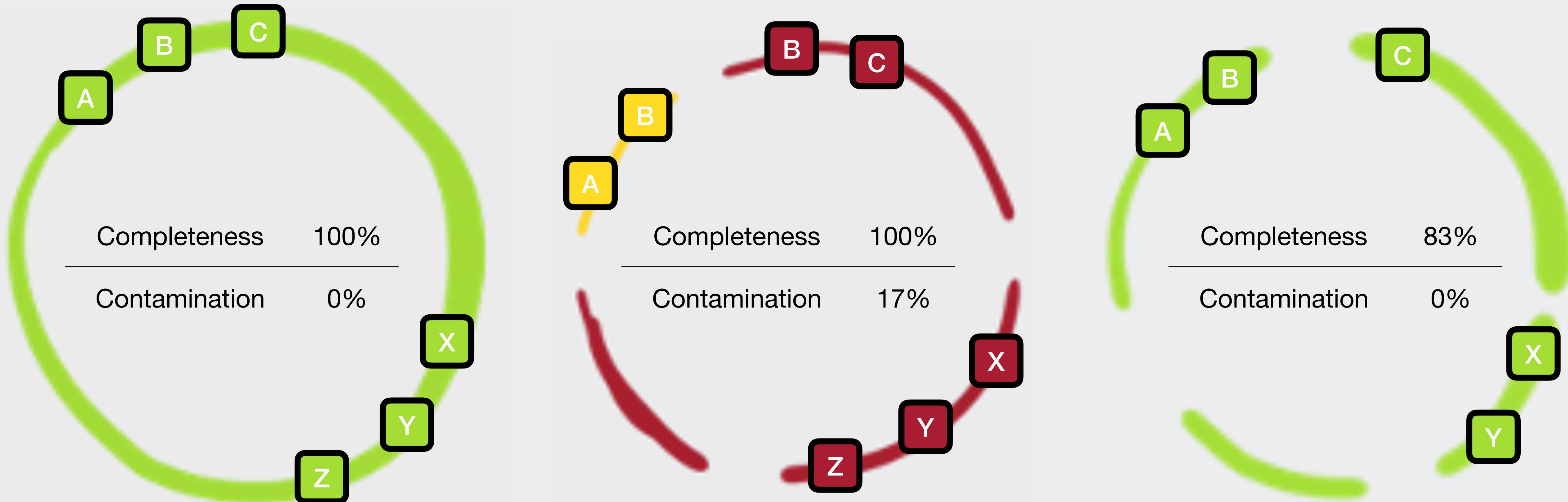


# Universal single-copy marker genes



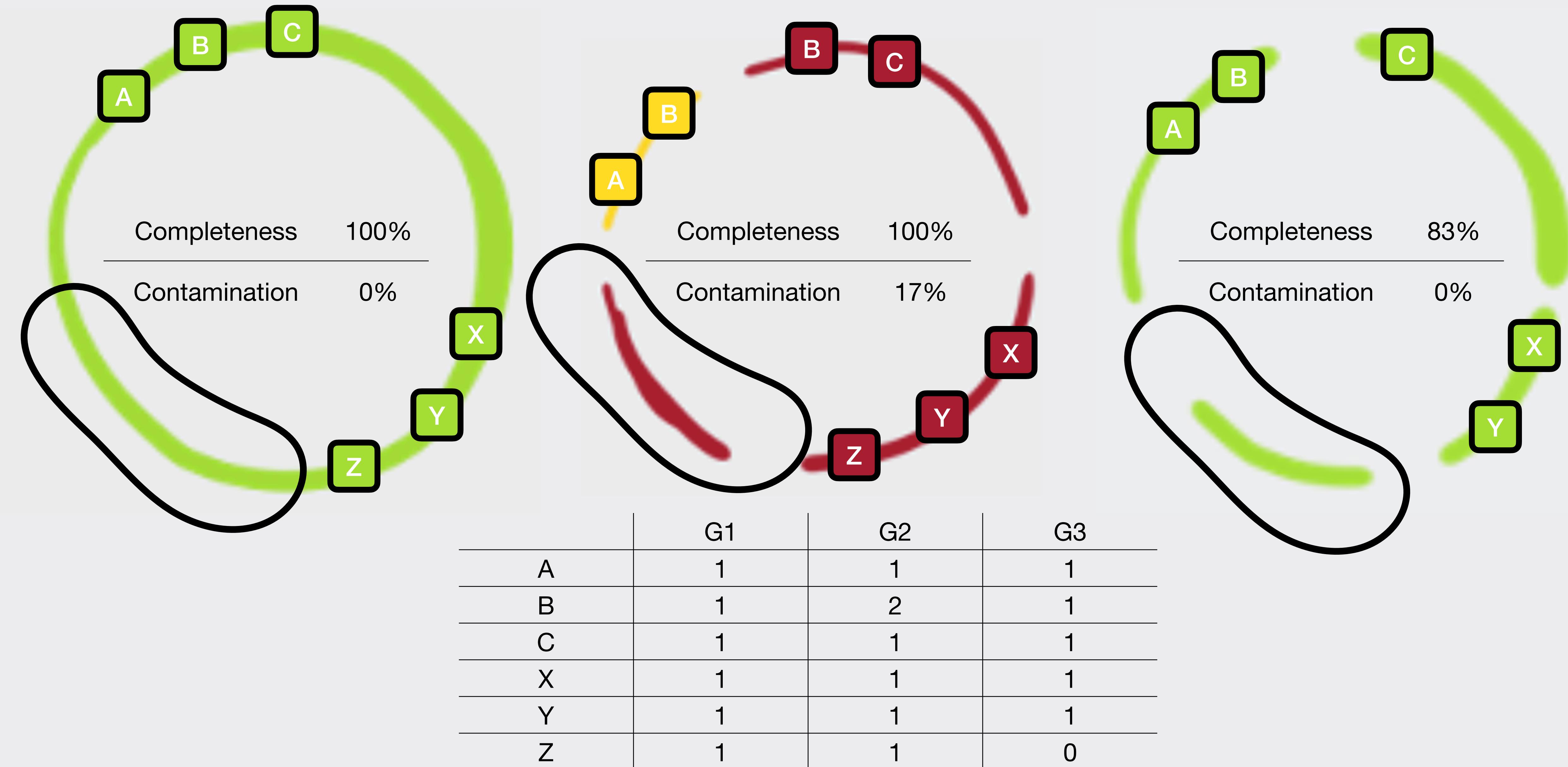
	G1	G2	G3
A	1	1	1
B	1	2	1
C	1	1	1
X	1	1	1
Y	1	1	1
Z	1	1	0

# Universal single-copy marker genes



	G1	G2	G3
A	1	1	1
B	1	2	1
C	1	1	1
X	1	1	1
Y	1	1	1
Z	1	1	0

# Universal single-copy marker genes



# Material

This course uses a lot of material from <https://merenlab.org/momics/>, I invite you to have a look.

If you want details on the bioinformatics behind you can start by having a look here:  
[https://astrobiomike.github.io/genomics/metagen\\_anvio](https://astrobiomike.github.io/genomics/metagen_anvio)



**A. Murat Eren (Meren) (PI)**

- Web  Email  Twitter  LinkedIn  Github  ORCiD
- Address: Knapp Center for Biomedical Discovery, 900 E. 57th St., MB 9, RM 9118, Chicago, IL 60637 USA
- Phone: +1-773-702-5935  Fax: +1-773-702-2281

*I am a computer scientist with a deep appreciation for the complexity of life. I design algorithms and experiments to better understand microbes and their ecology. [photos: 1, 2, 3].*

» MBL Fellow, [Marine Biological Laboratory](#).

» Assistant Professor, [The Department of Medicine at the University of Chicago](#).

» Committee on Microbiology, [The Biomedical Sciences Cluster at the University of Chicago](#).



**Mike Lee**

- Web  Email  Twitter  LinkedIn  Github

- » NASA Space Biology Fellow, [NASA Ames Research Center](#).
- » JCVI Research Fellow, [J. Craig Venter Institute](#).

- 👉 Combining reference genome annotations with your own in pangenomes (Sat, Dec 01, 2018)
- 👉 Anvi'o 'views' demystified (Mon, May 08, 2017)
- 👉 Making anvi'o use your own HMM collection (Sat, May 21, 2016)