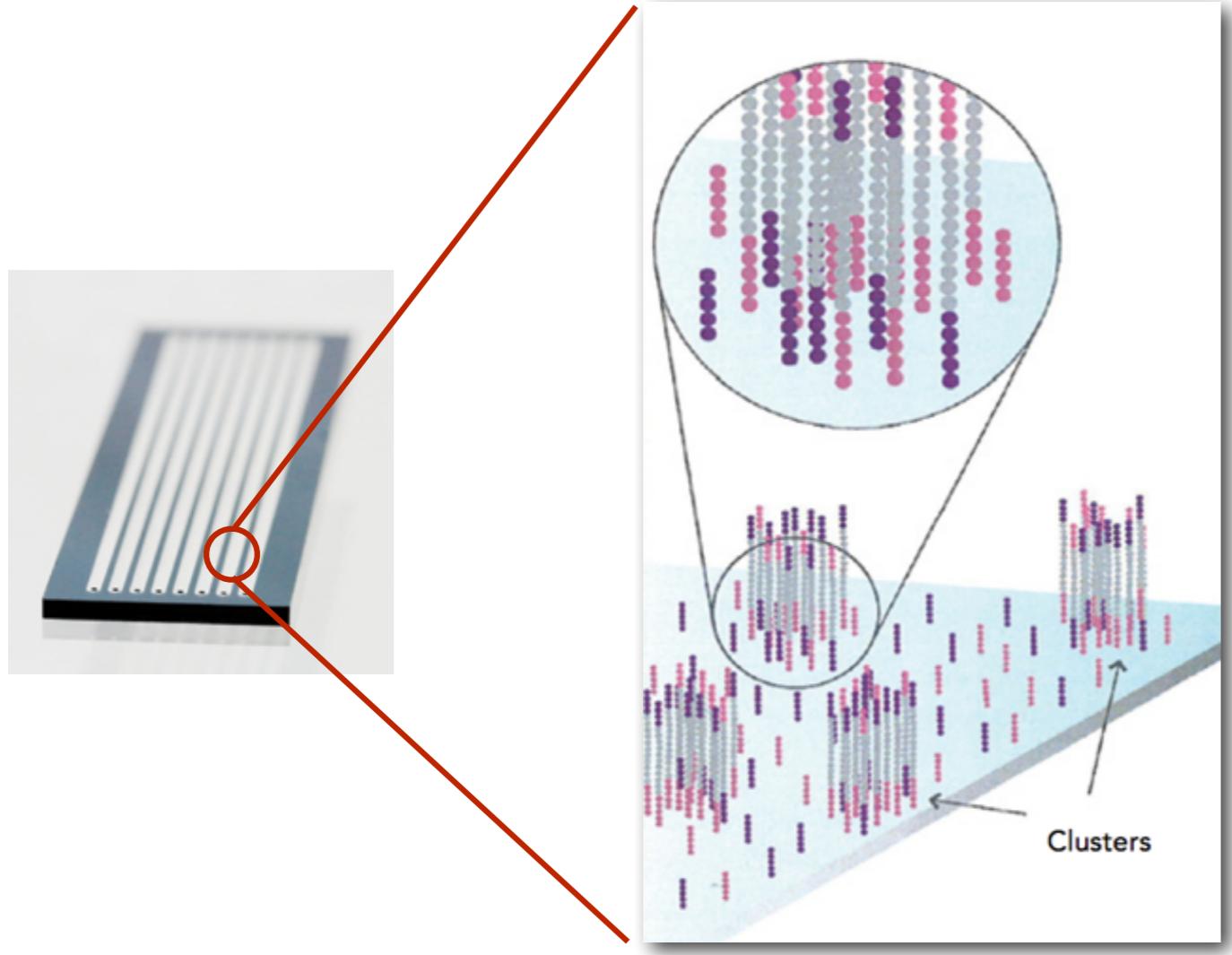


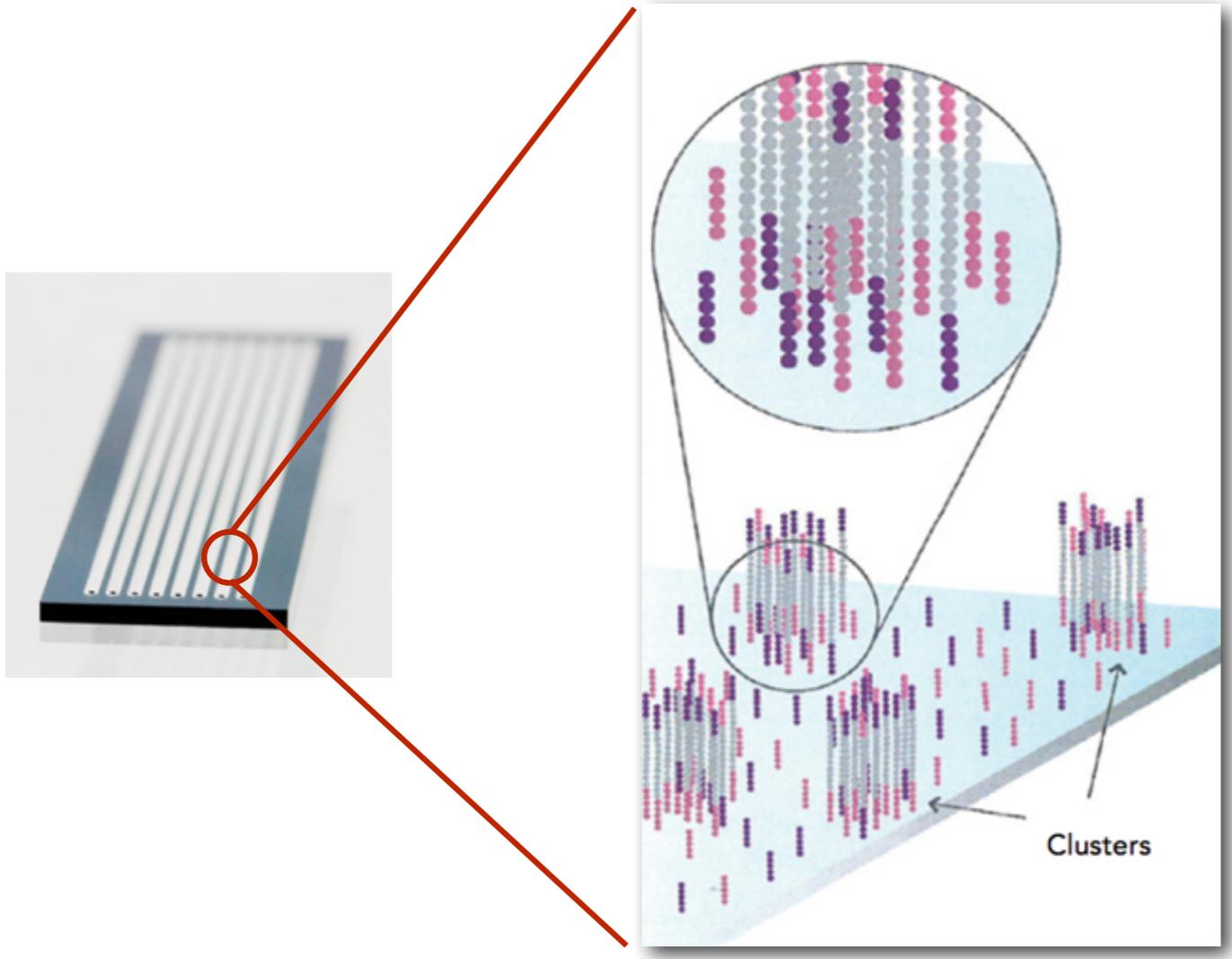
RADseq Variations

Original RAD
double digest RAD (ddRAD)
Genotyping by Sequencing (GBS)
ezRAD
2bRAD

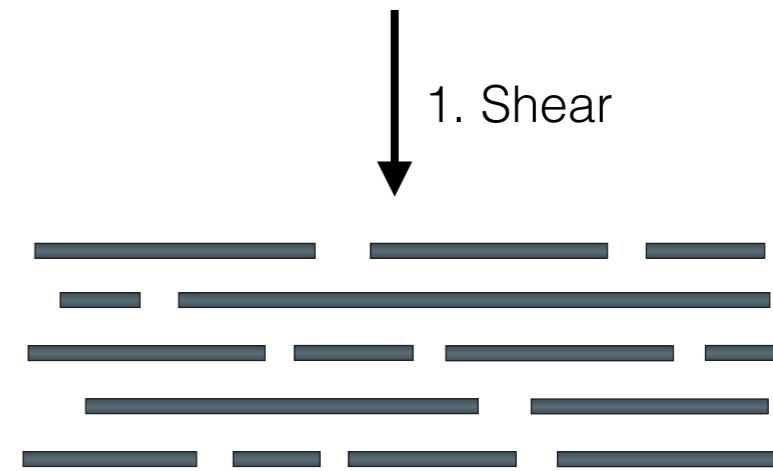
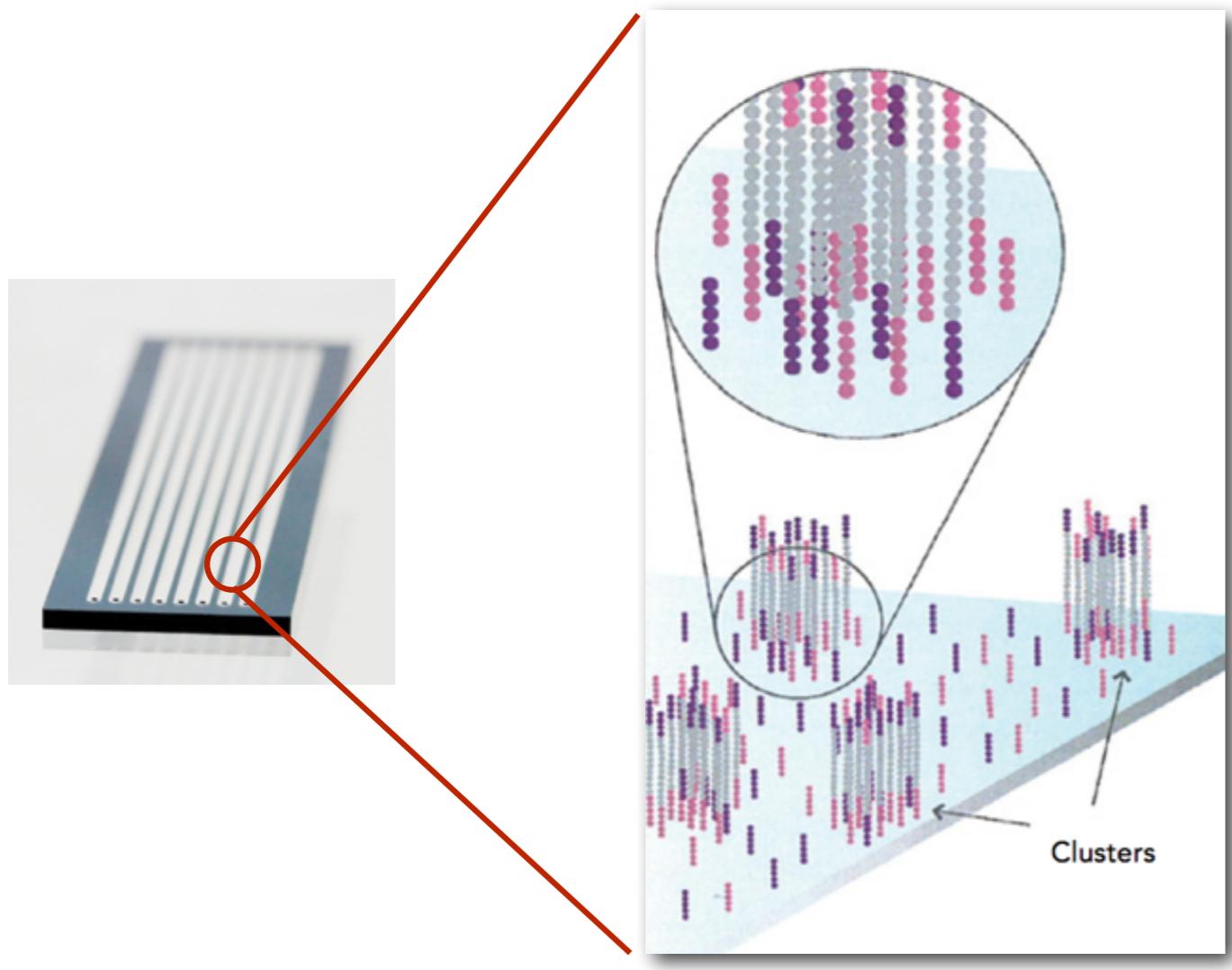
Issues

Allele dropout
PCR duplicates
Paired end contigs
Read depth fragment length bias

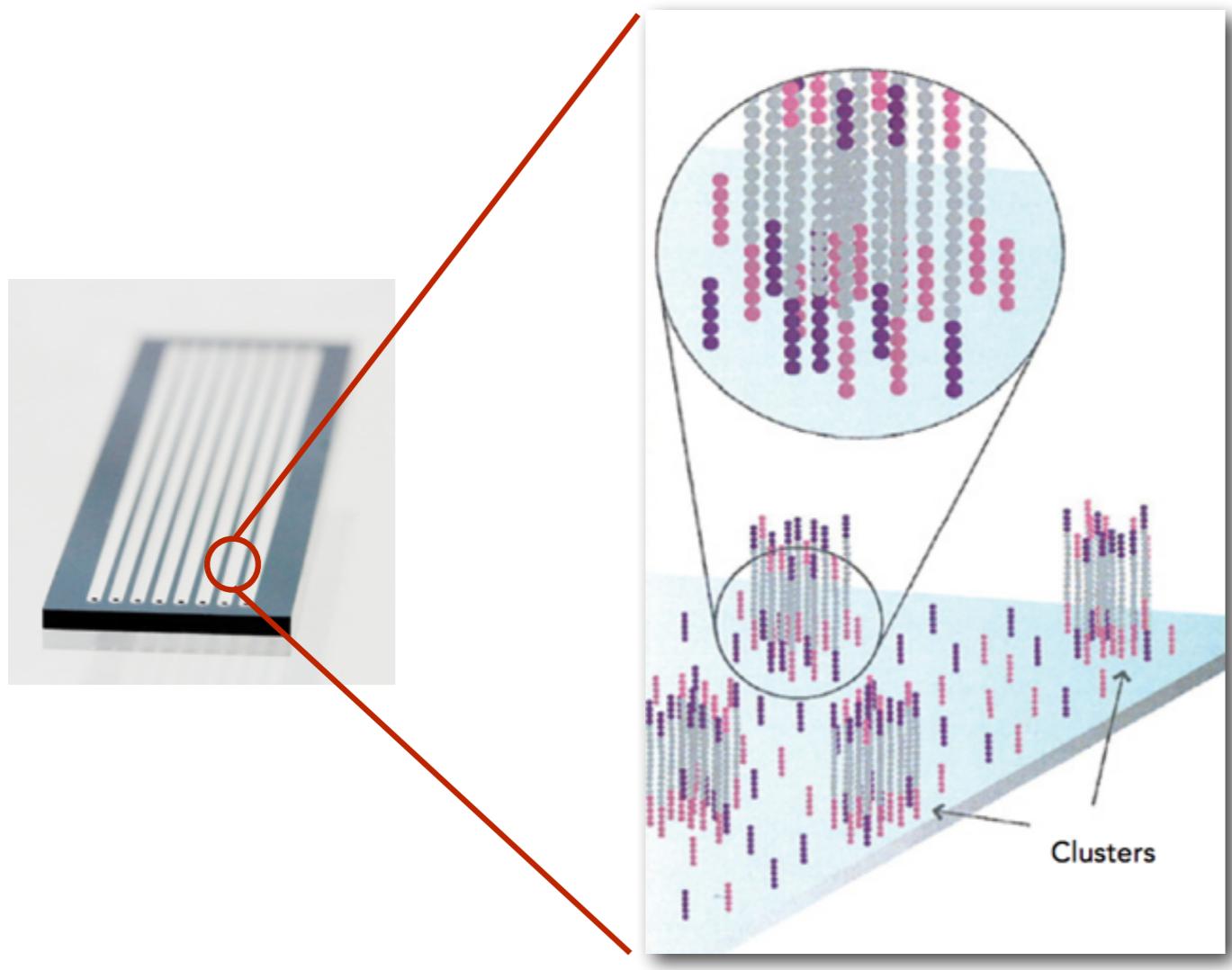




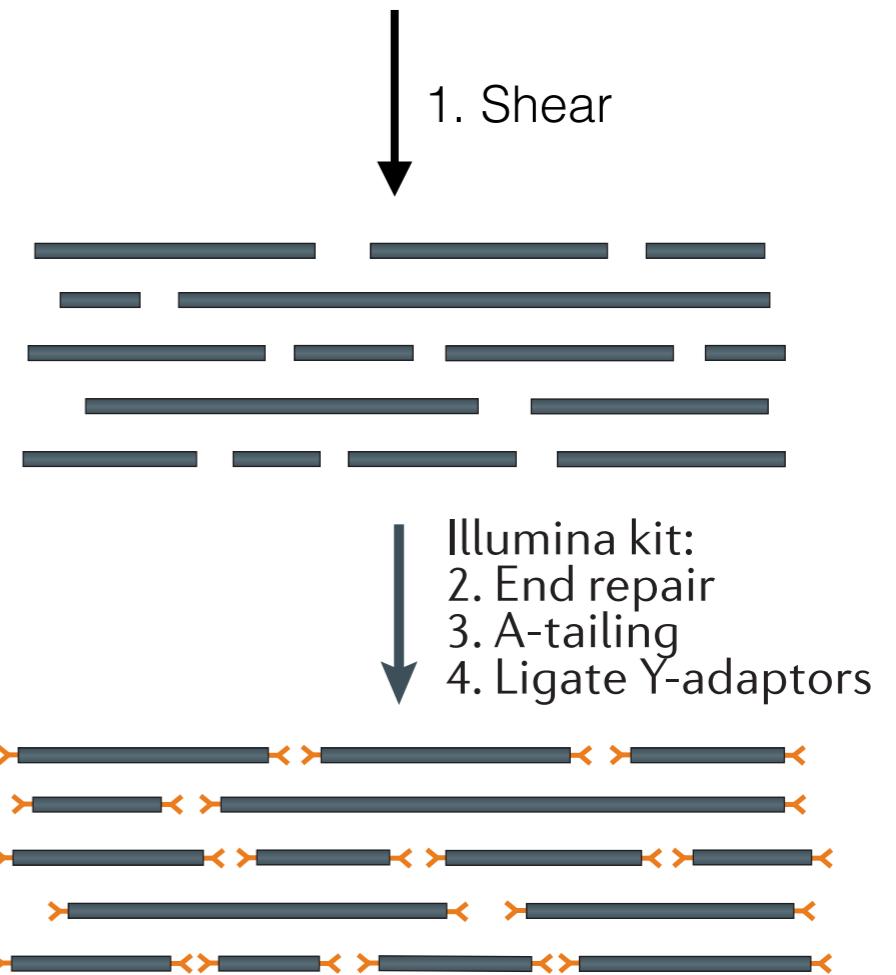
1. Binding to flow cell requires adapters
2. Clustering requires fragments <1kbp

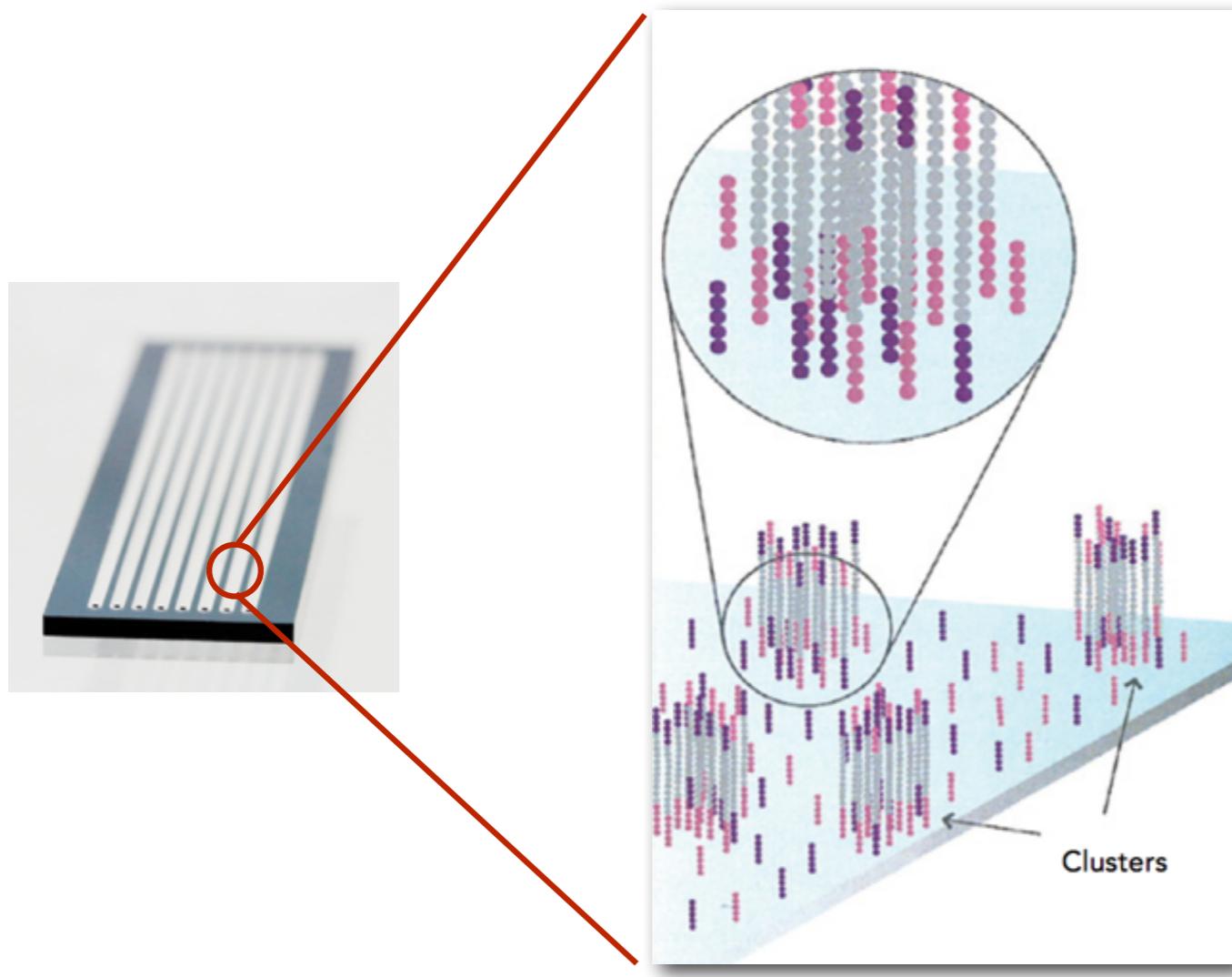


1. Binding to flow cell requires adapters
2. Clustering requires fragments <1kbp

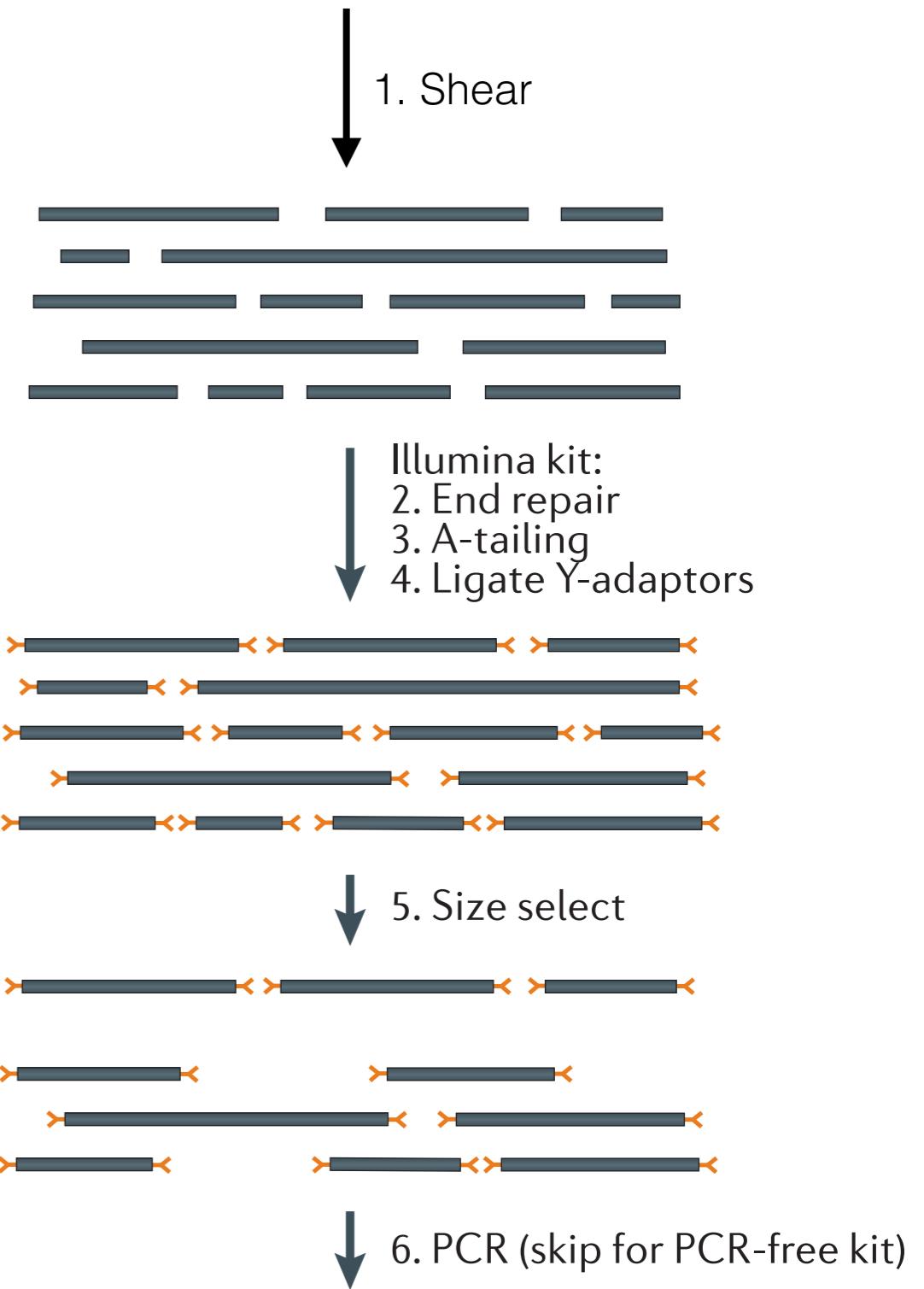


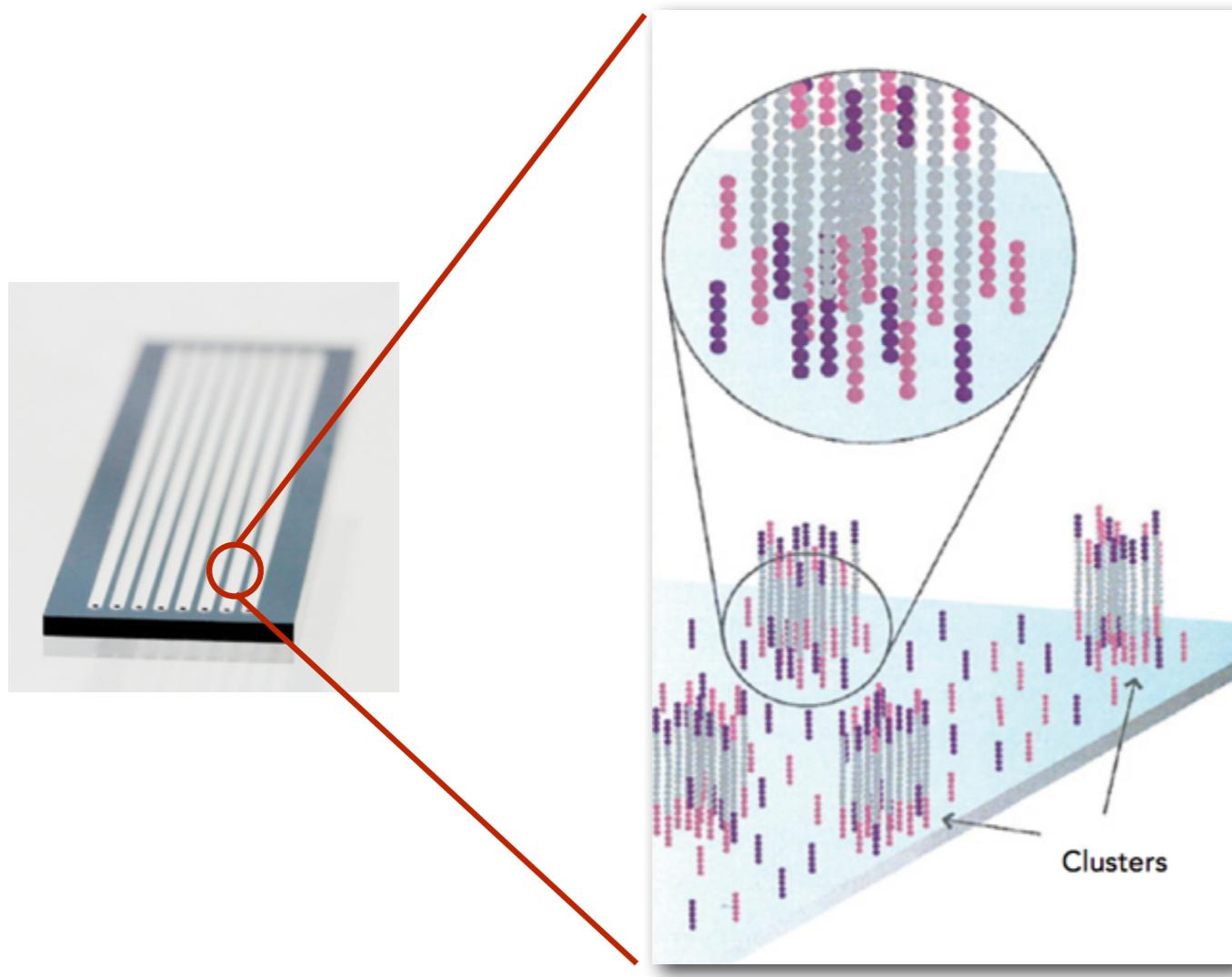
1. Binding to flow cell requires adapters
2. Clustering requires fragments <1kbp



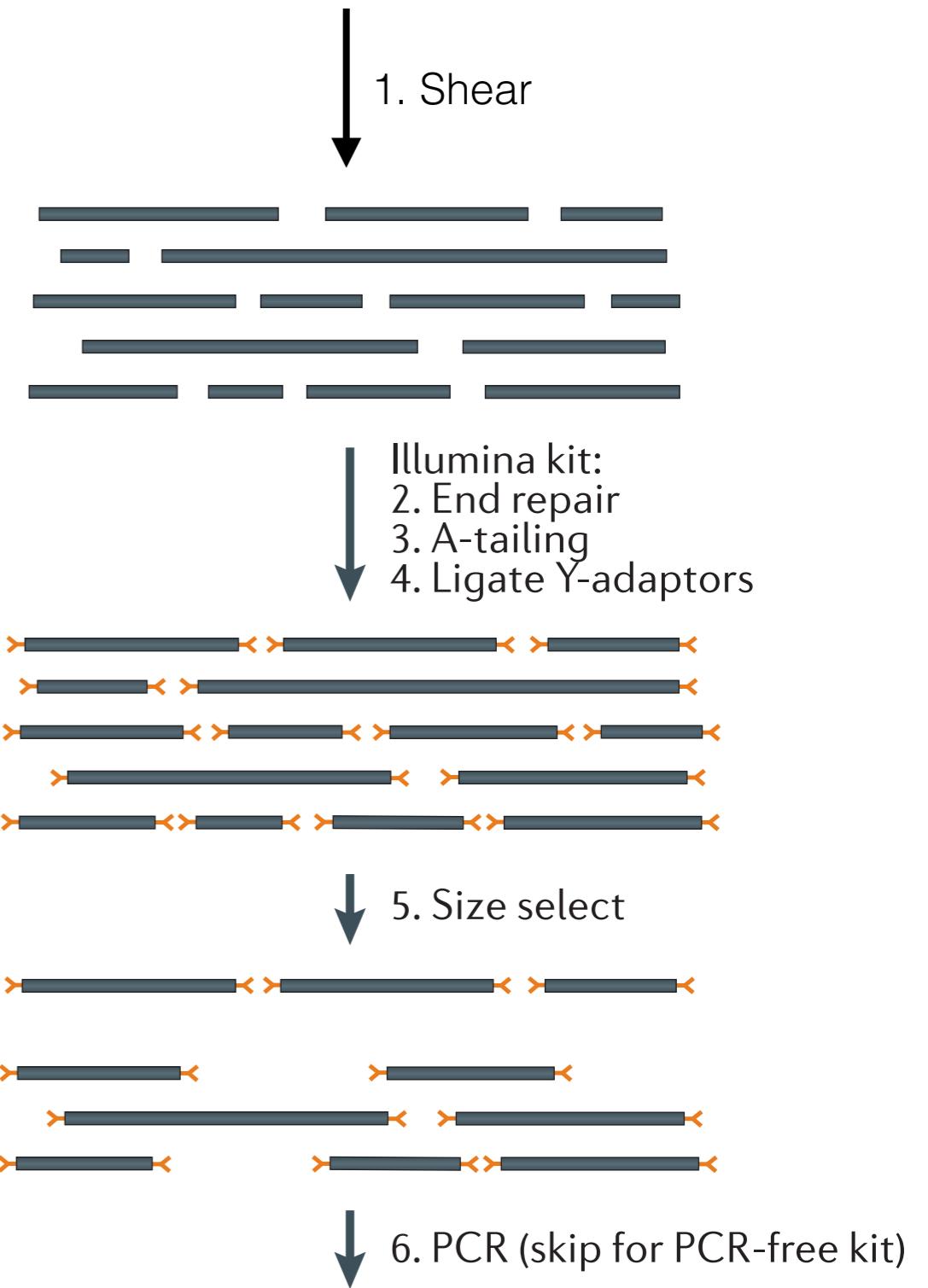


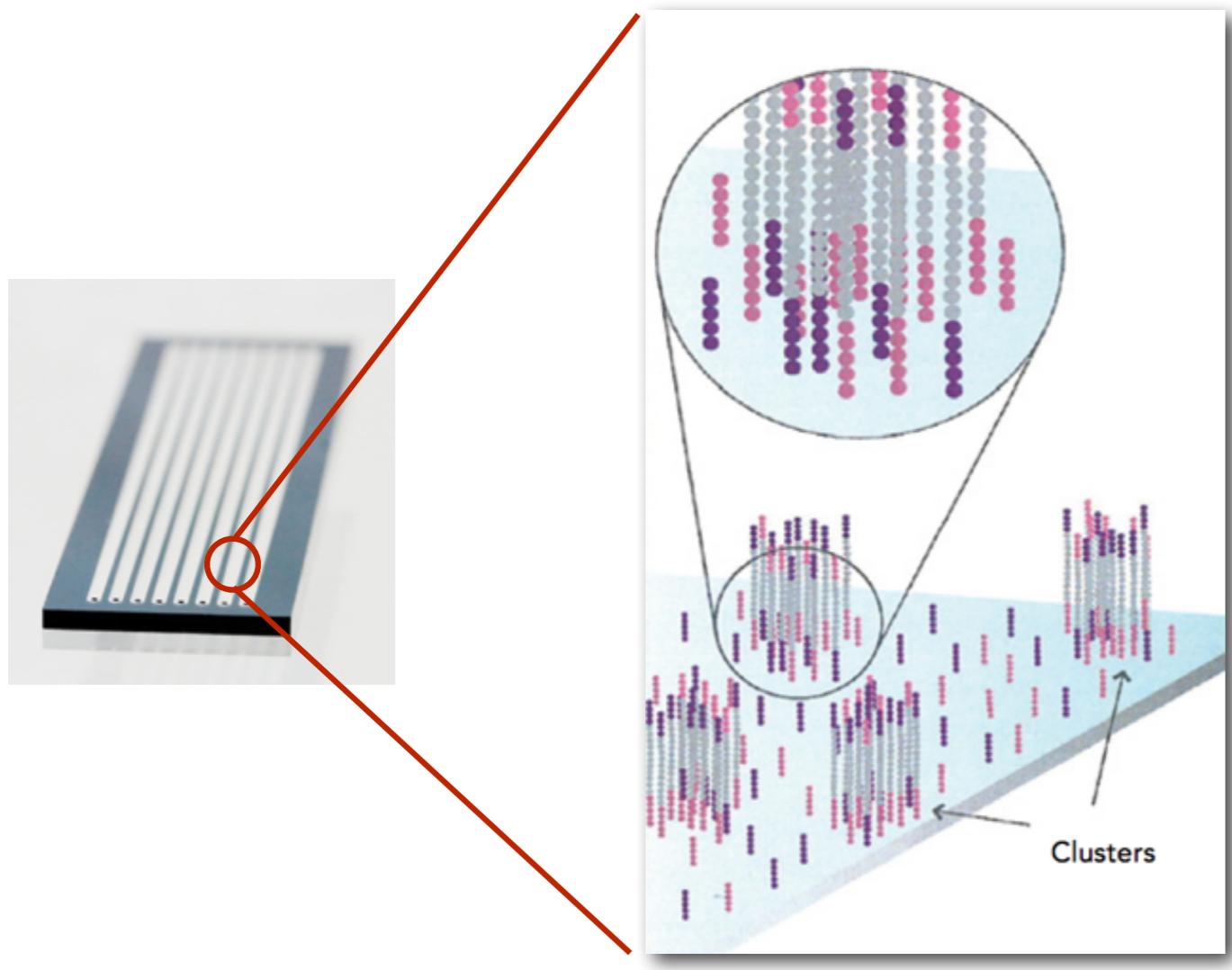
1. Binding to flow cell requires adapters
2. Clustering requires fragments <1kbp



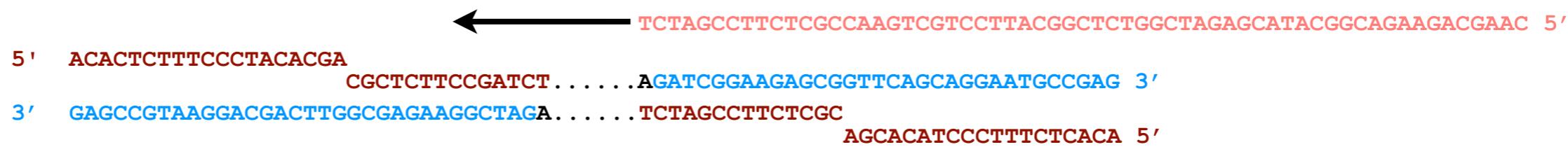
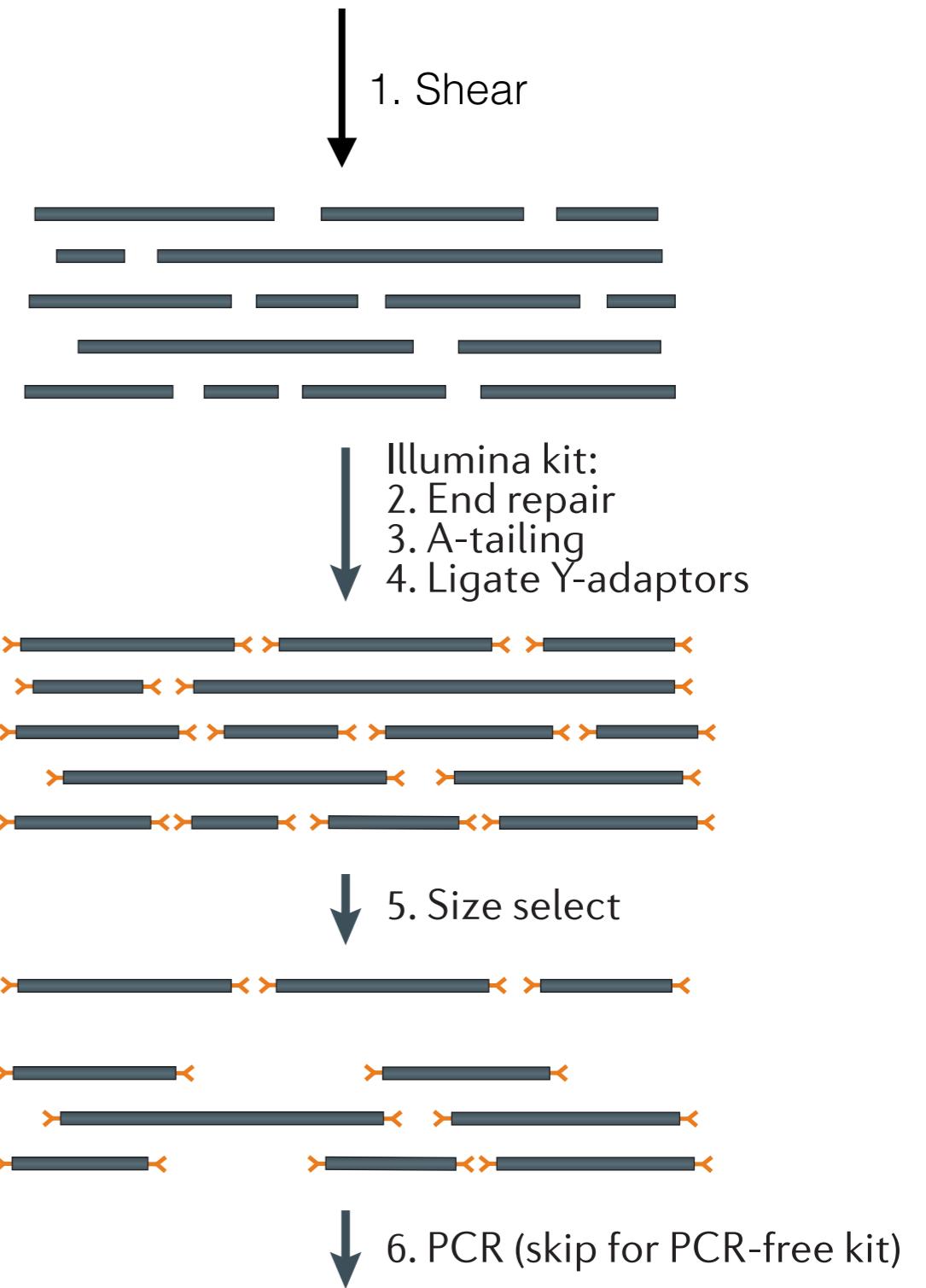


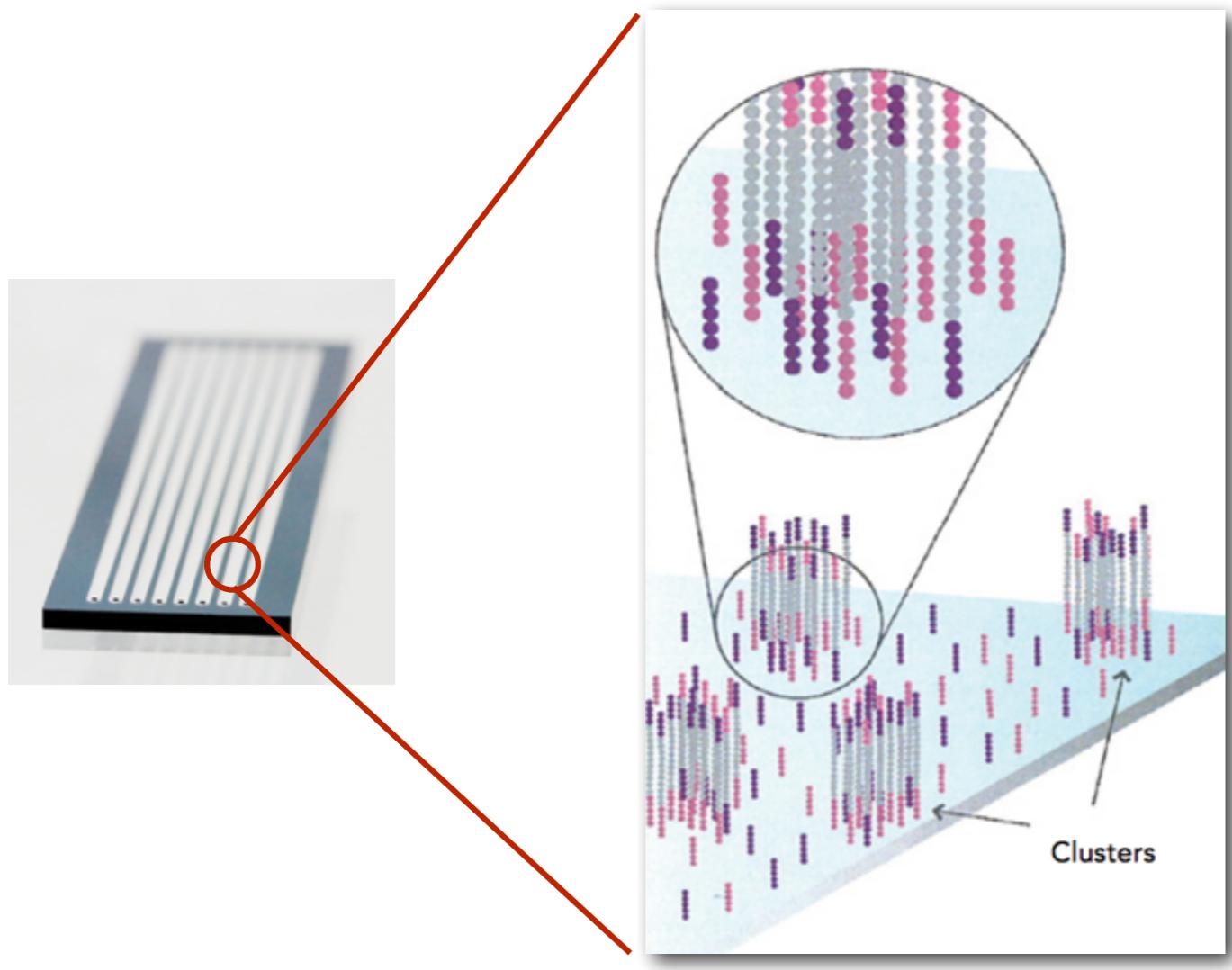
1. Binding to flow cell requires adapters
2. Clustering requires fragments <1kbp



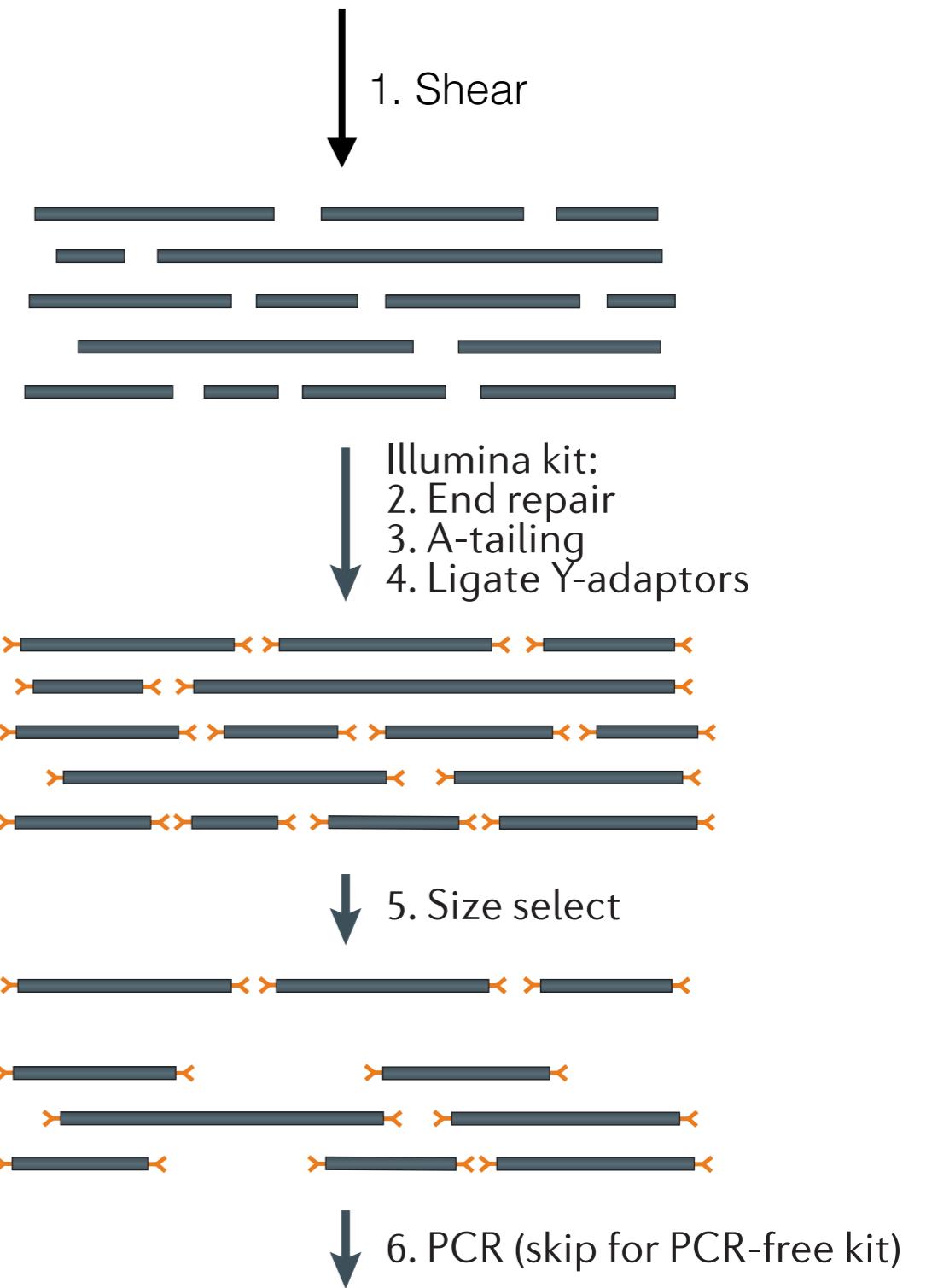


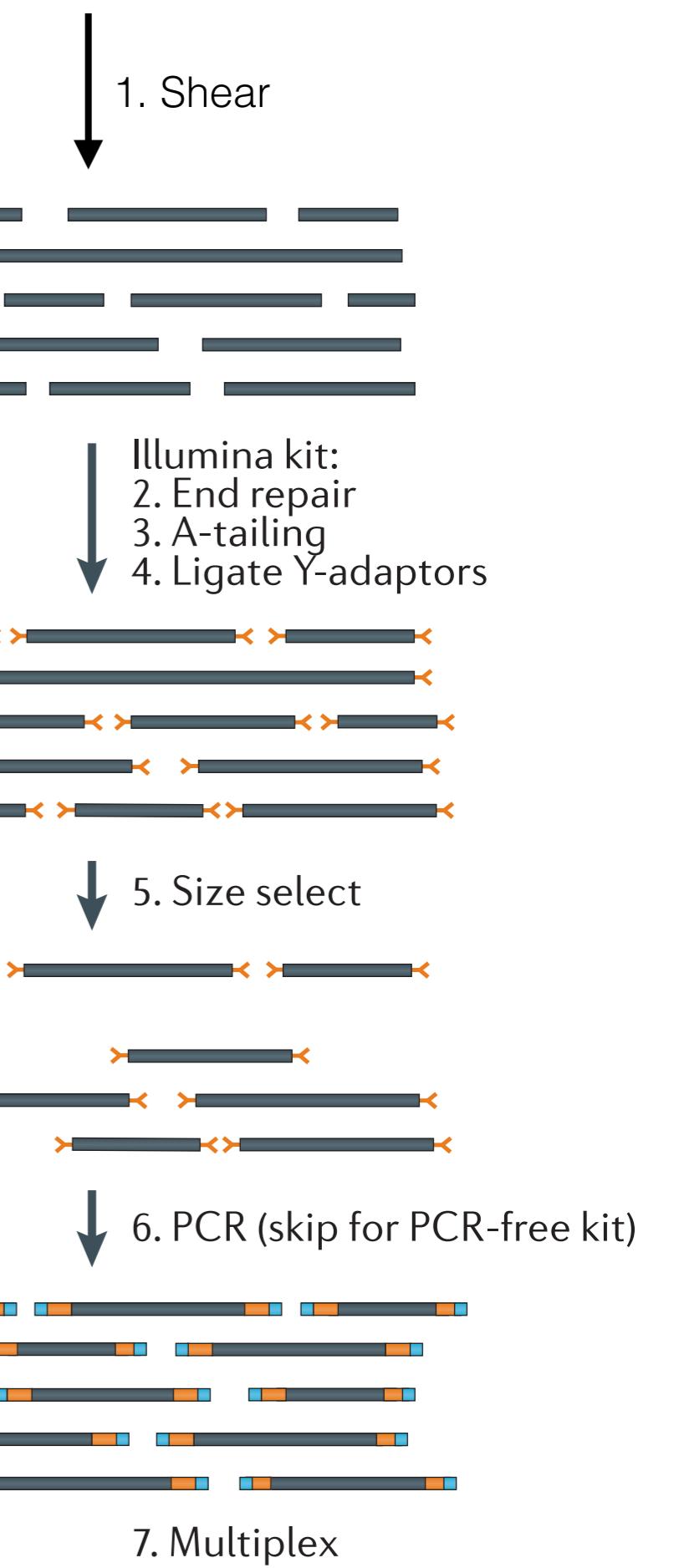
1. Binding to flow cell requires adapters
2. Clustering requires fragments <1kbp

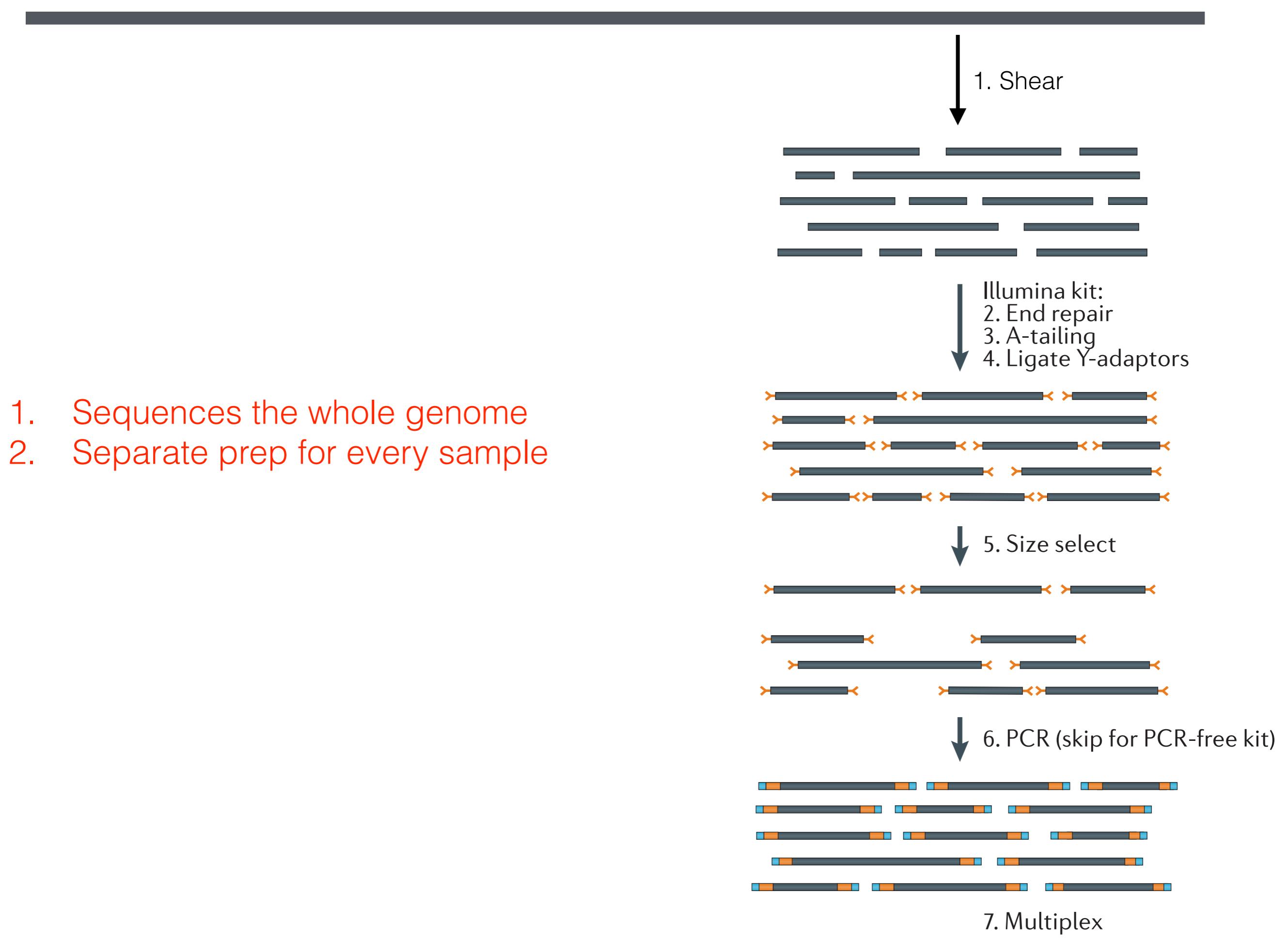




1. Binding to flow cell requires adapters
2. Clustering requires fragments <1kbp



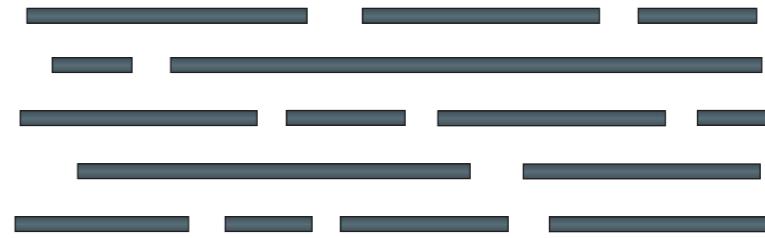




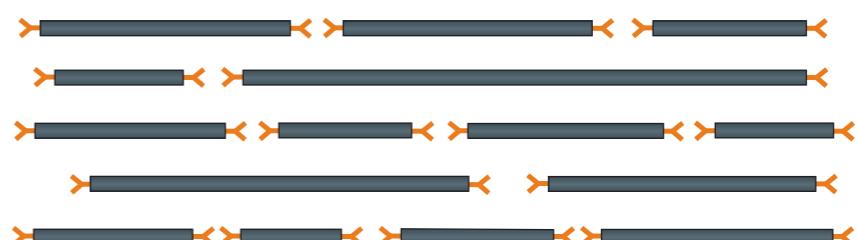
ezRAD

1. Sequences the whole genome
2. Separate prep for every sample

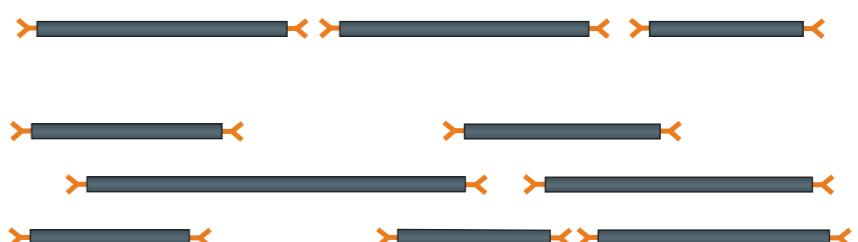
↓ 1. Digest (one or more enzymes)



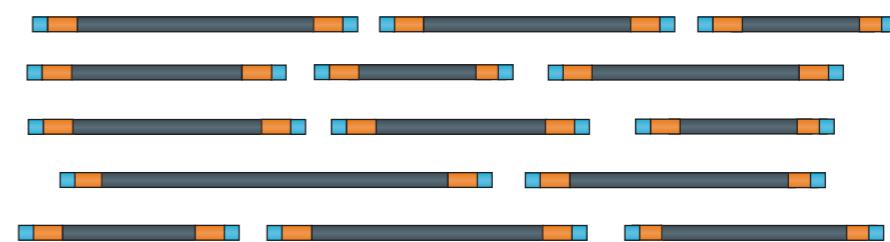
↓ Illumina kit:
2. End repair
3. A-tailing
4. Ligate Y-adaptors



↓ 5. Size select



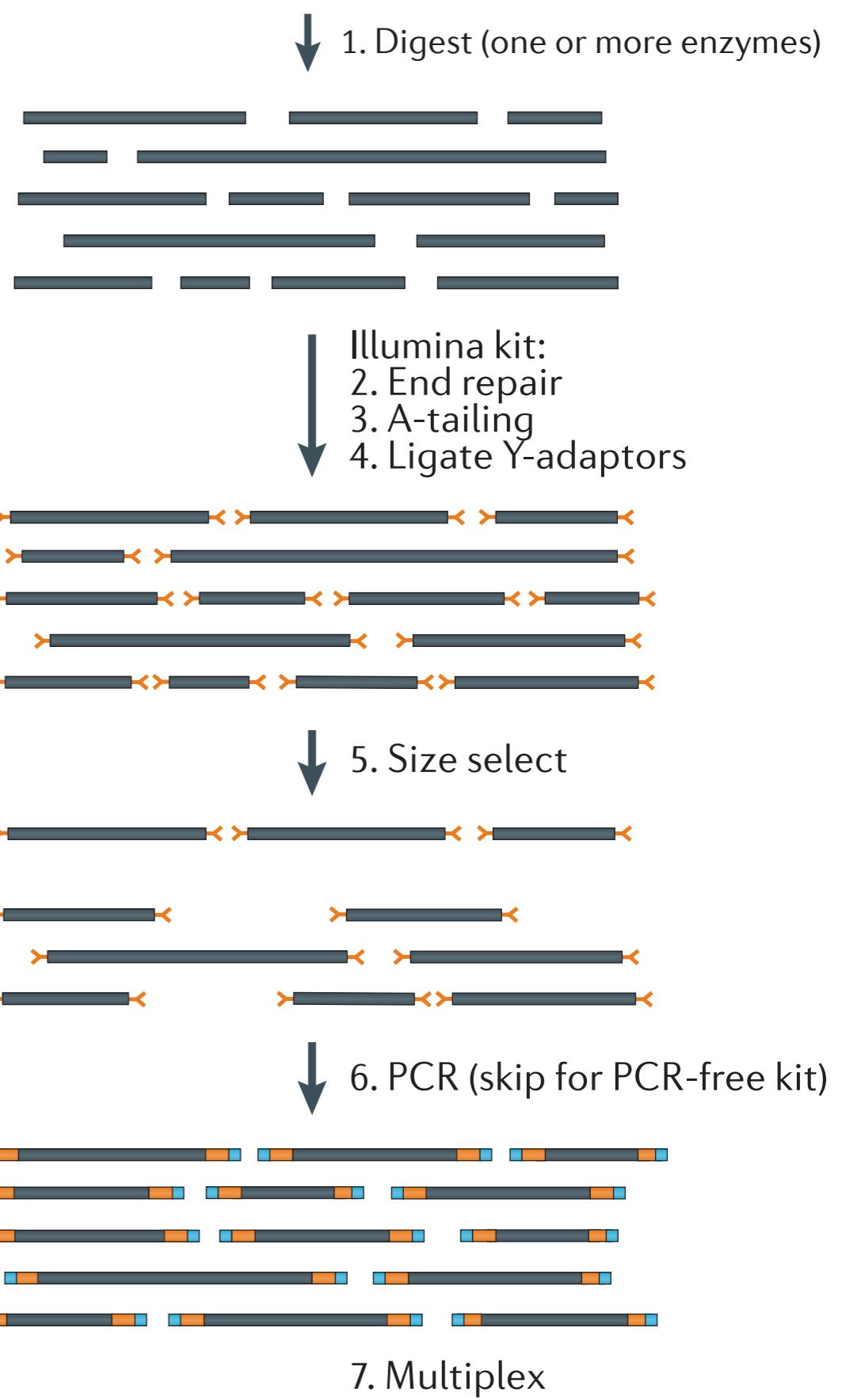
↓ 6. PCR (skip for PCR-free kit)

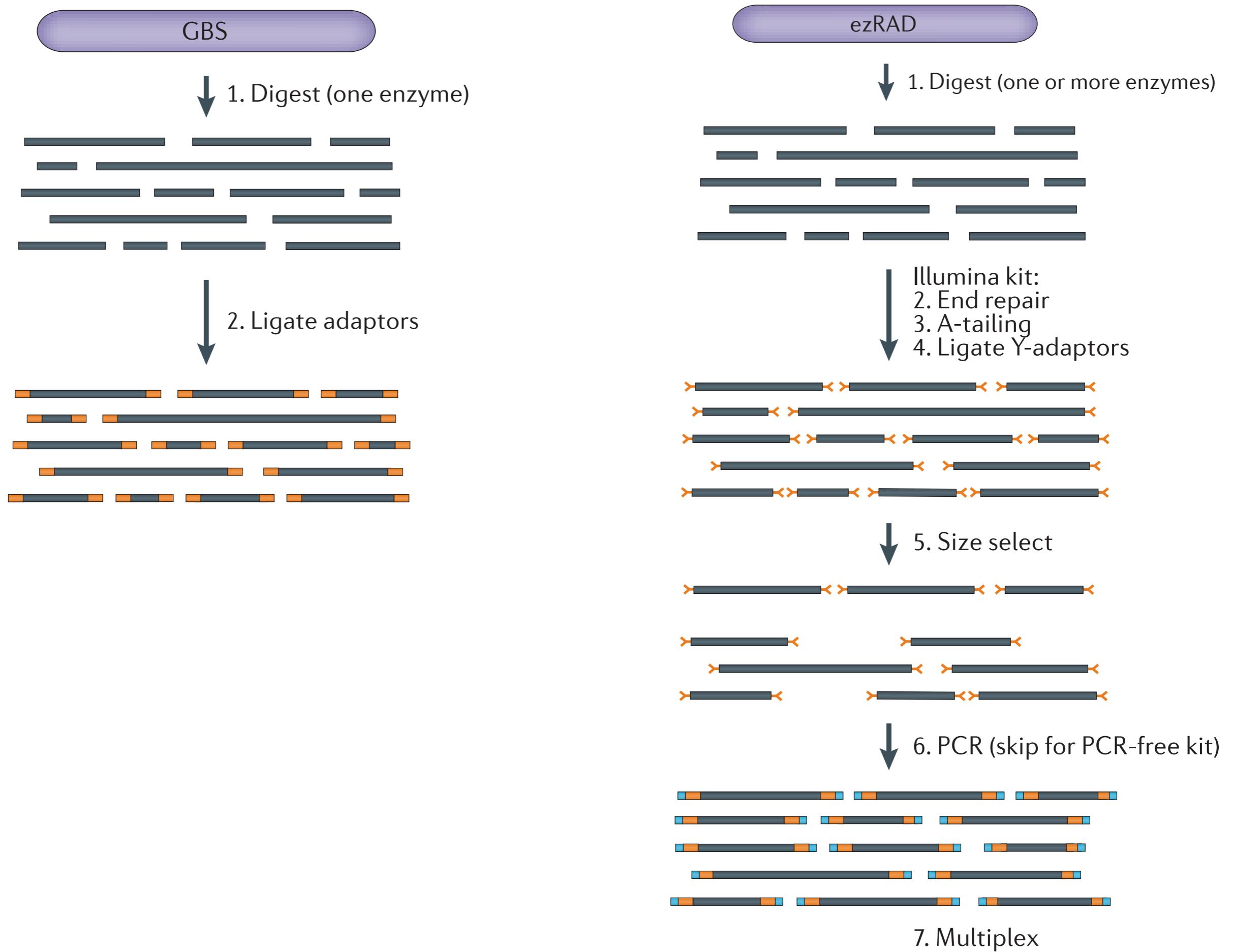


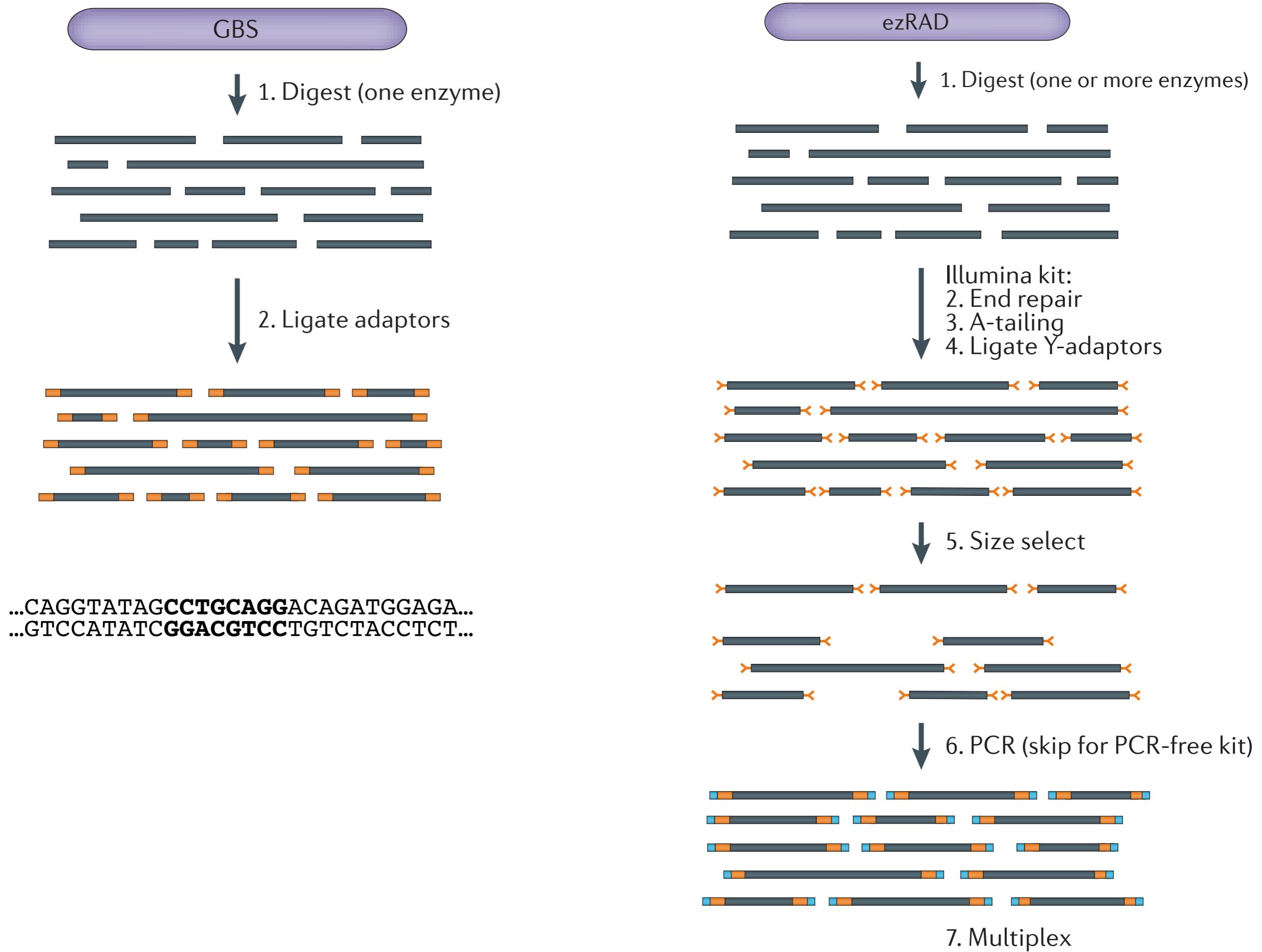
7. Multiplex

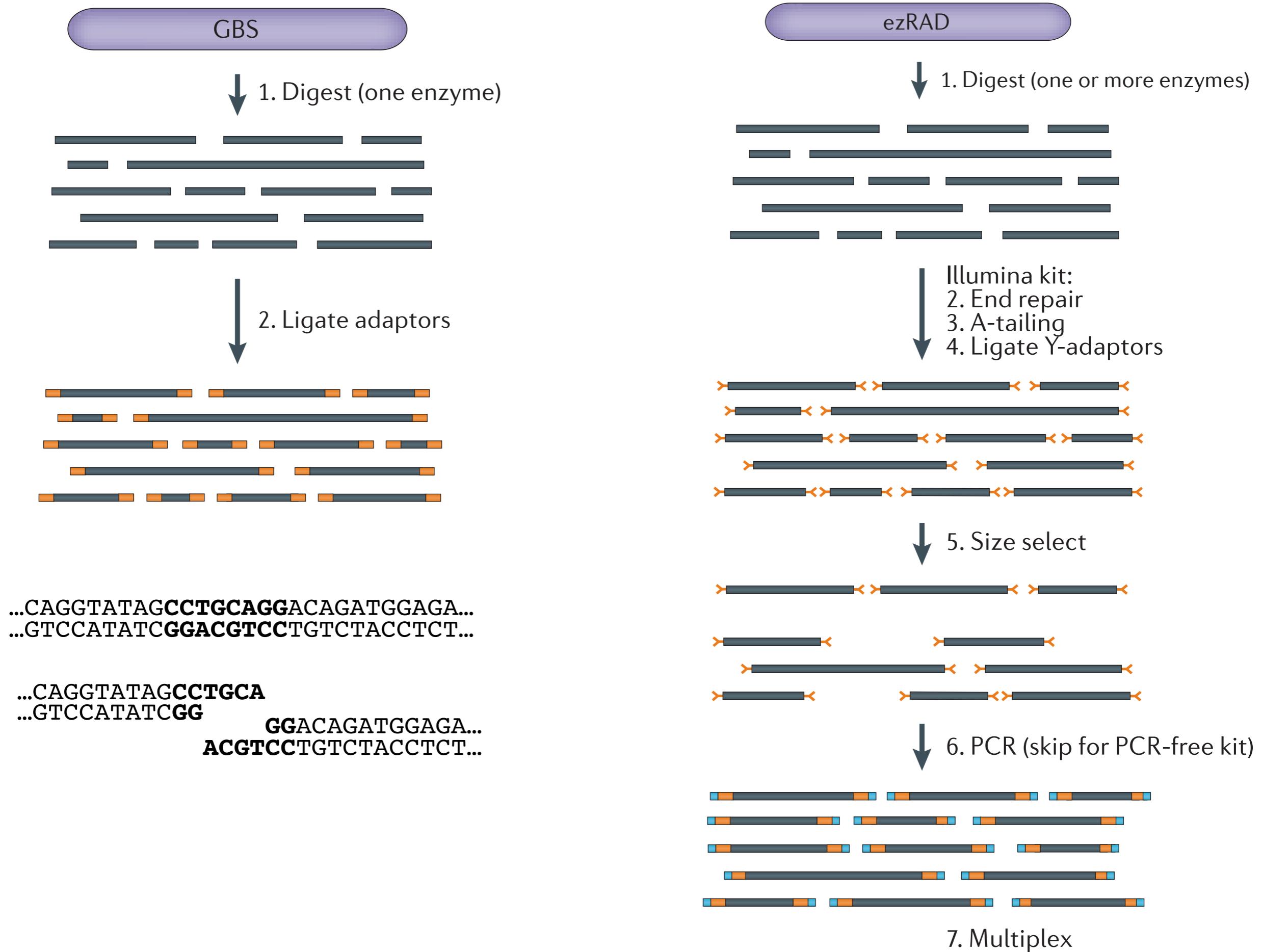
ezRAD

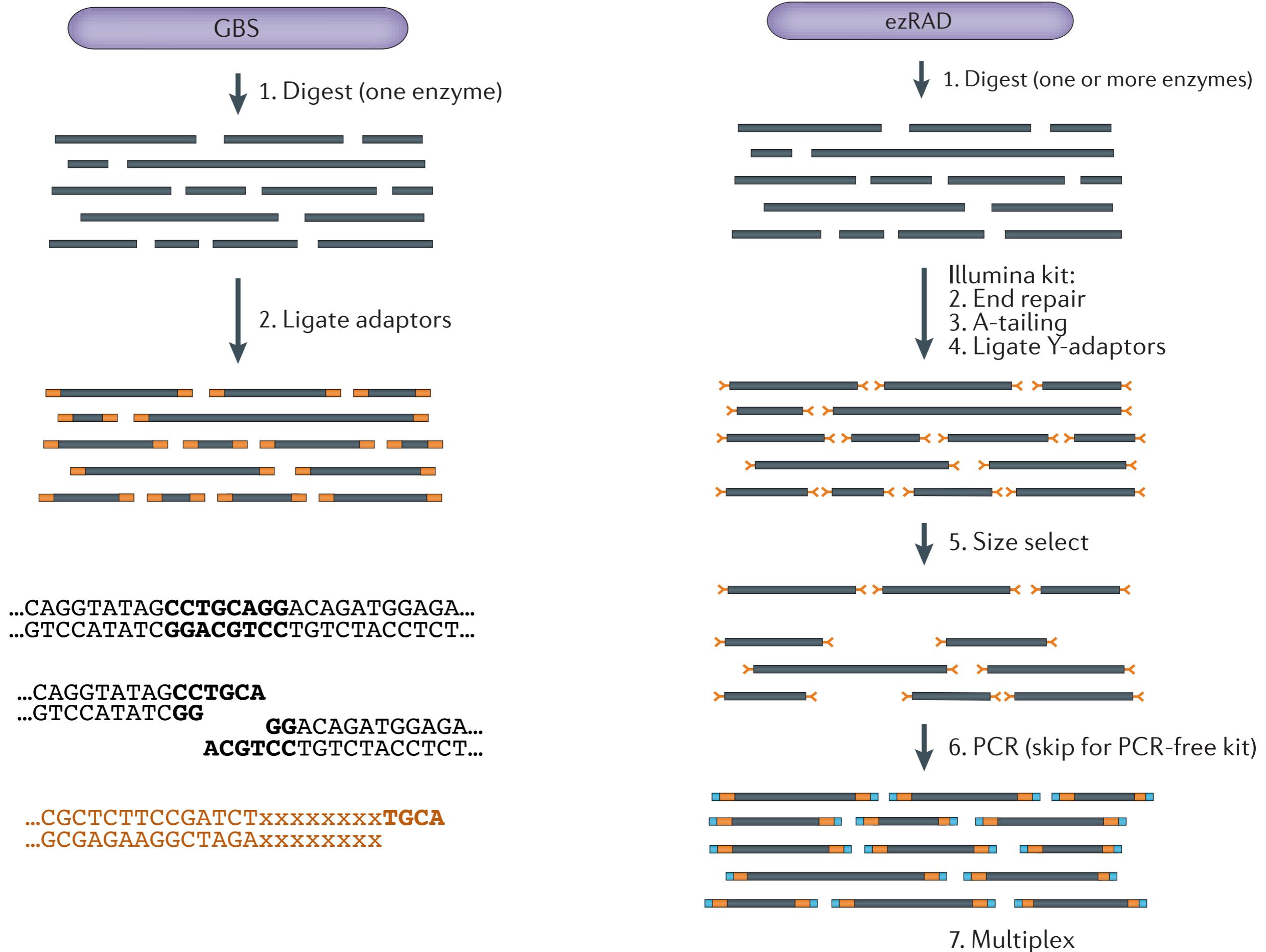
1. Sequences a subset of the genome
2. Separate prep for every sample
3. Will sequence short genomic fragments
4. Uses standard library prep kit

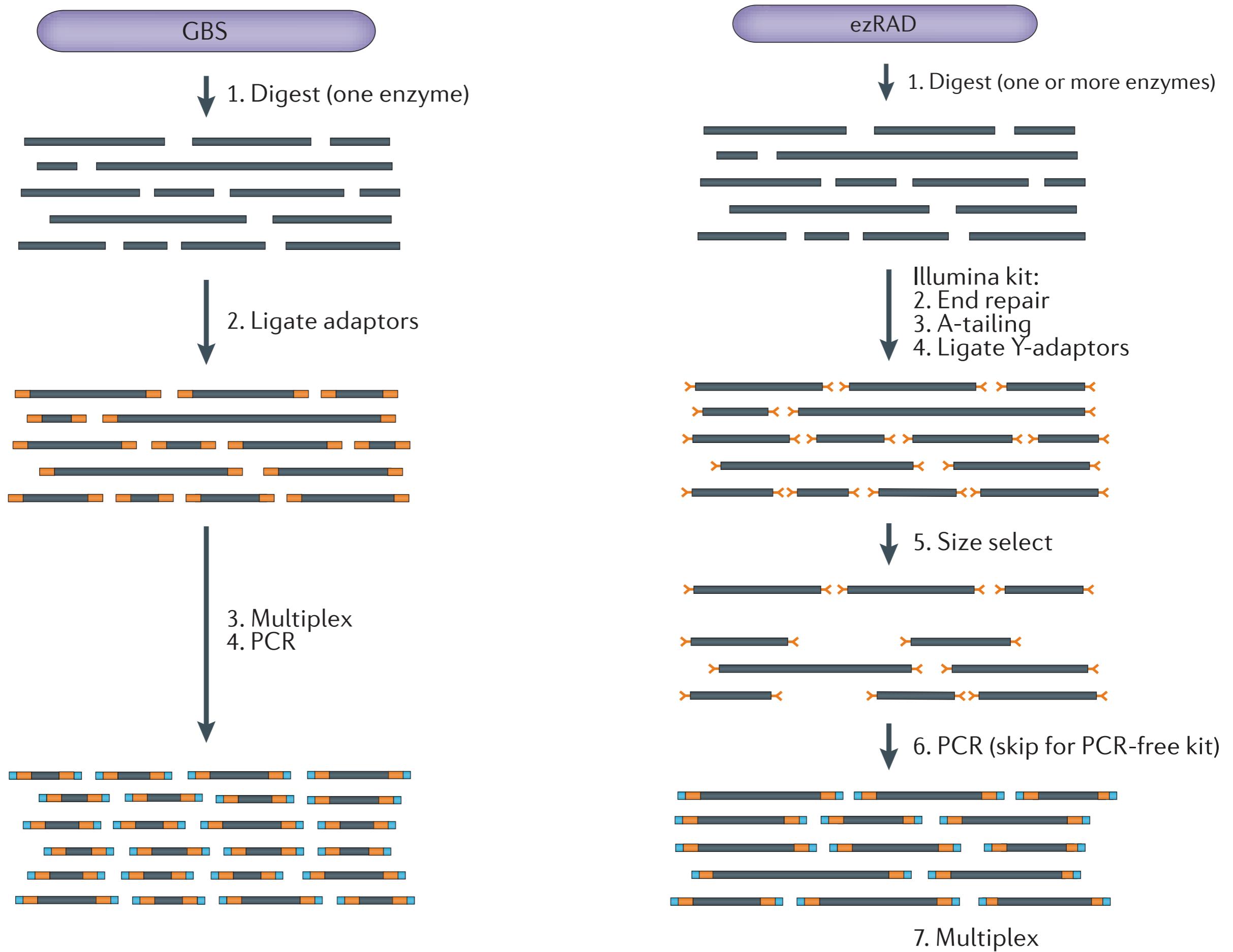


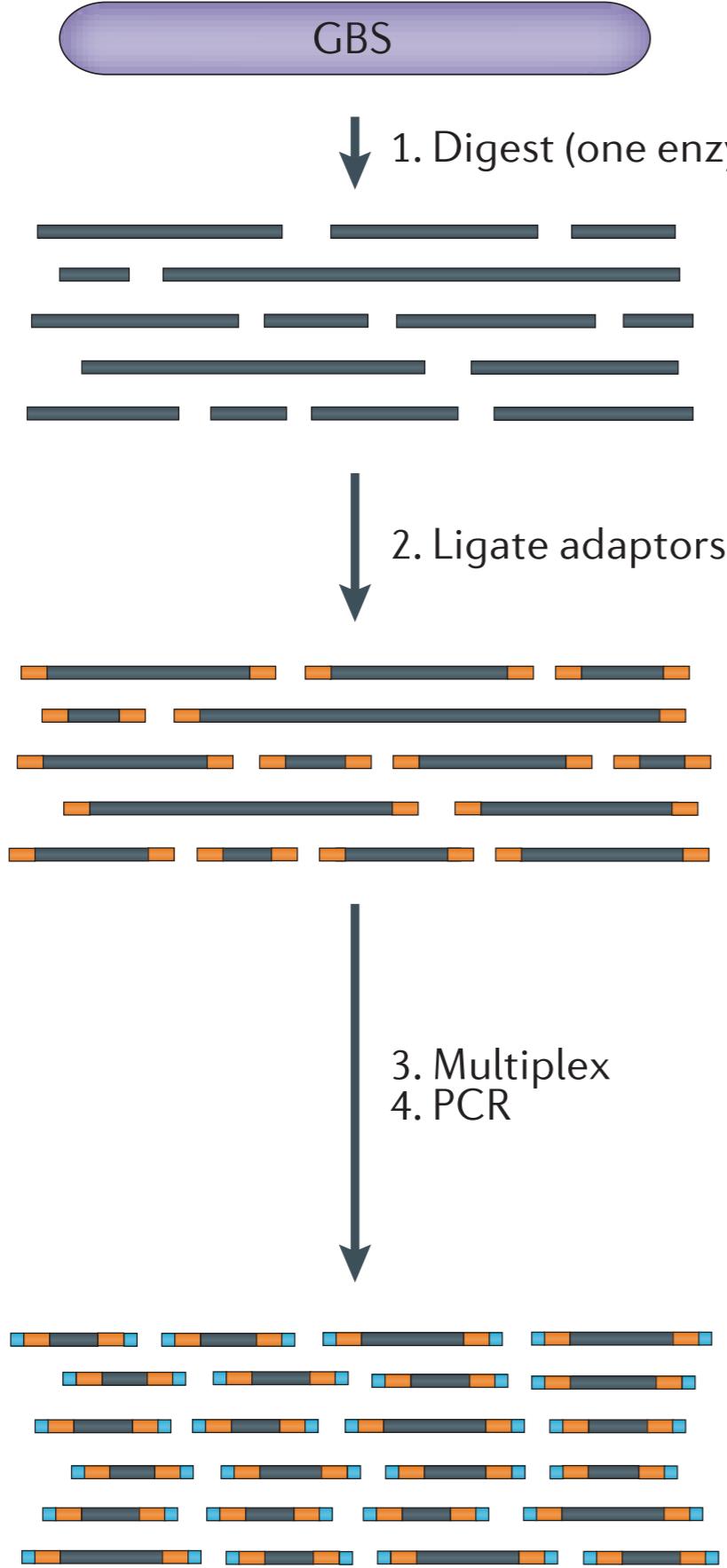




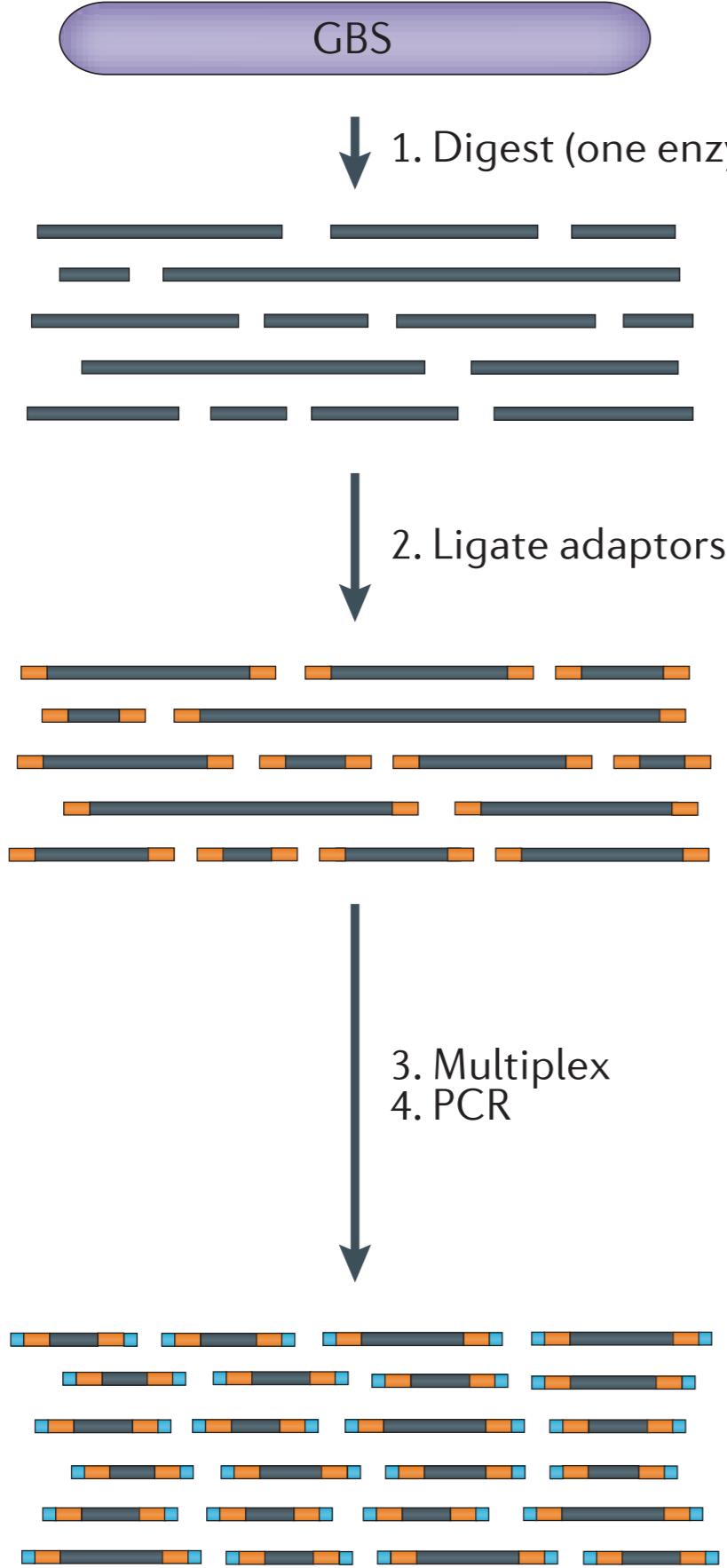




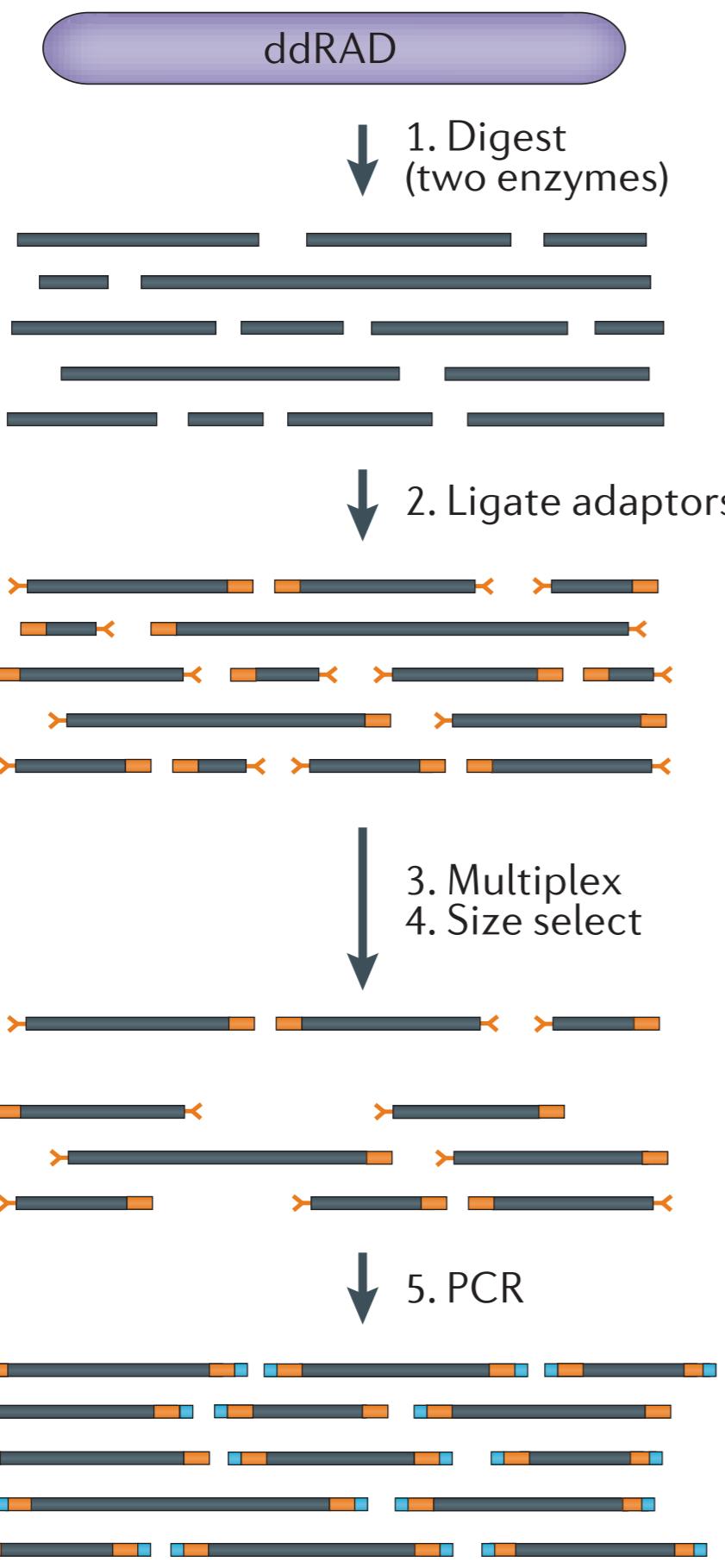


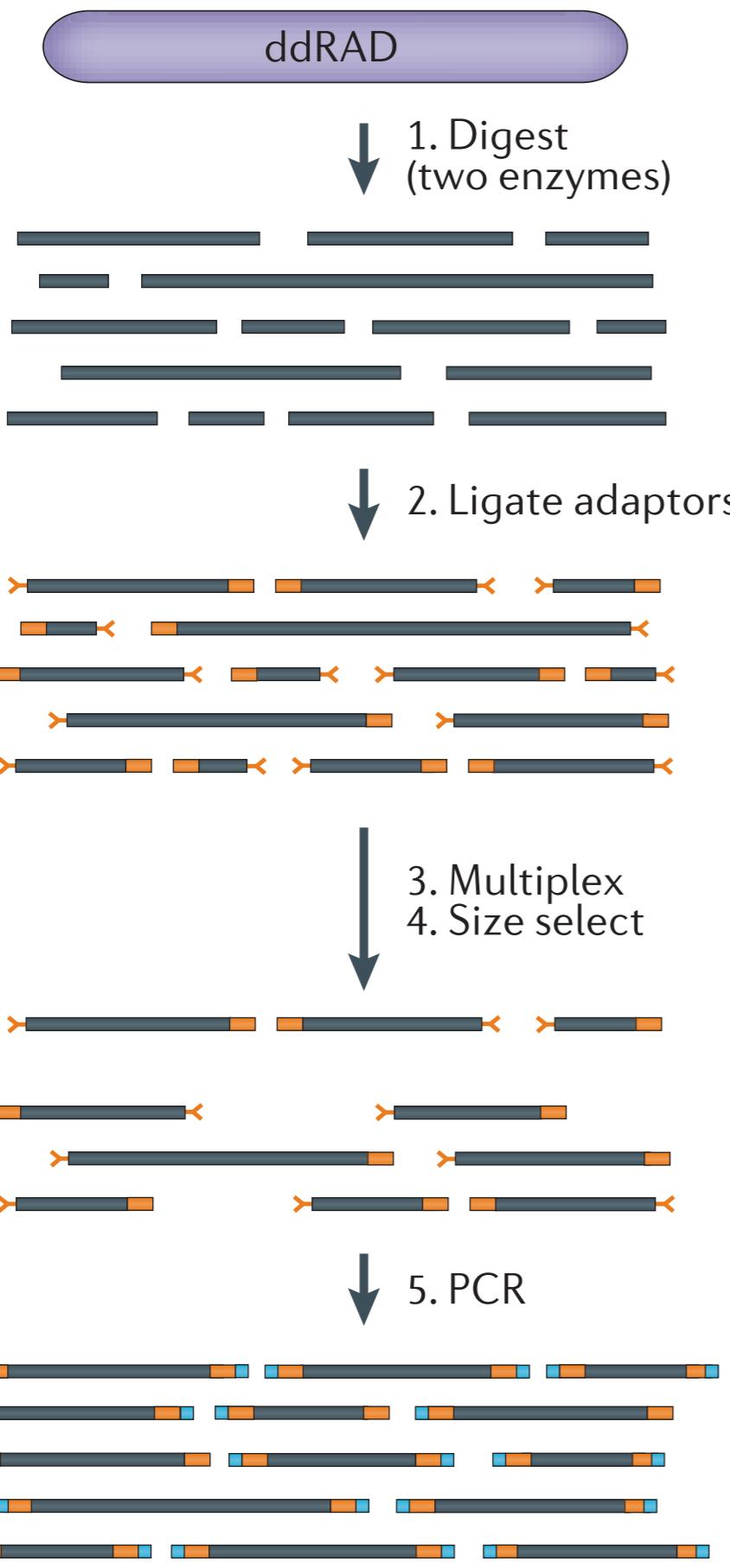


1. Sequences a subset of the genome
2. Separate prep for every sample
3. Will sequence short genomic fragments
4. Uses standard library prep kit



1. Sequences a subset of the genome
2. Multiplex adapters before PCR
3. Difficult enzyme/genome size tradeoff
4. Long fragments on flowcell
5. Requires adapter set

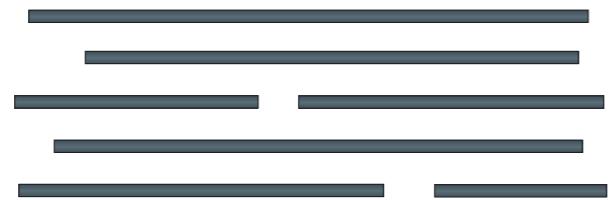




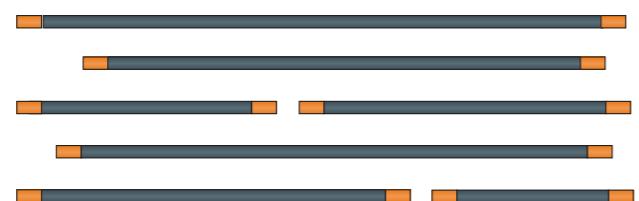
1. Sequences a subset of the genome
2. Multiplex adapters before PCR
3. Free choice of enzymes
4. Requires adapter set(s)
5. More susceptible to allele dropout

Original RAD

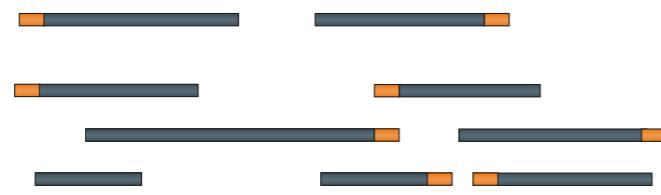
↓ 1. Digest (one enzyme)



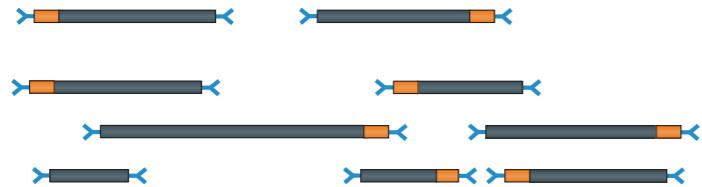
↓ 2. Ligate adaptors



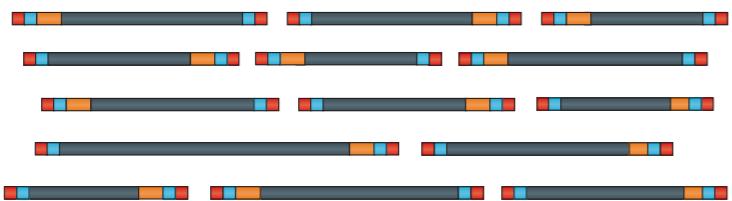
↓ 3. Multiplex
4. Shear
5. Size select

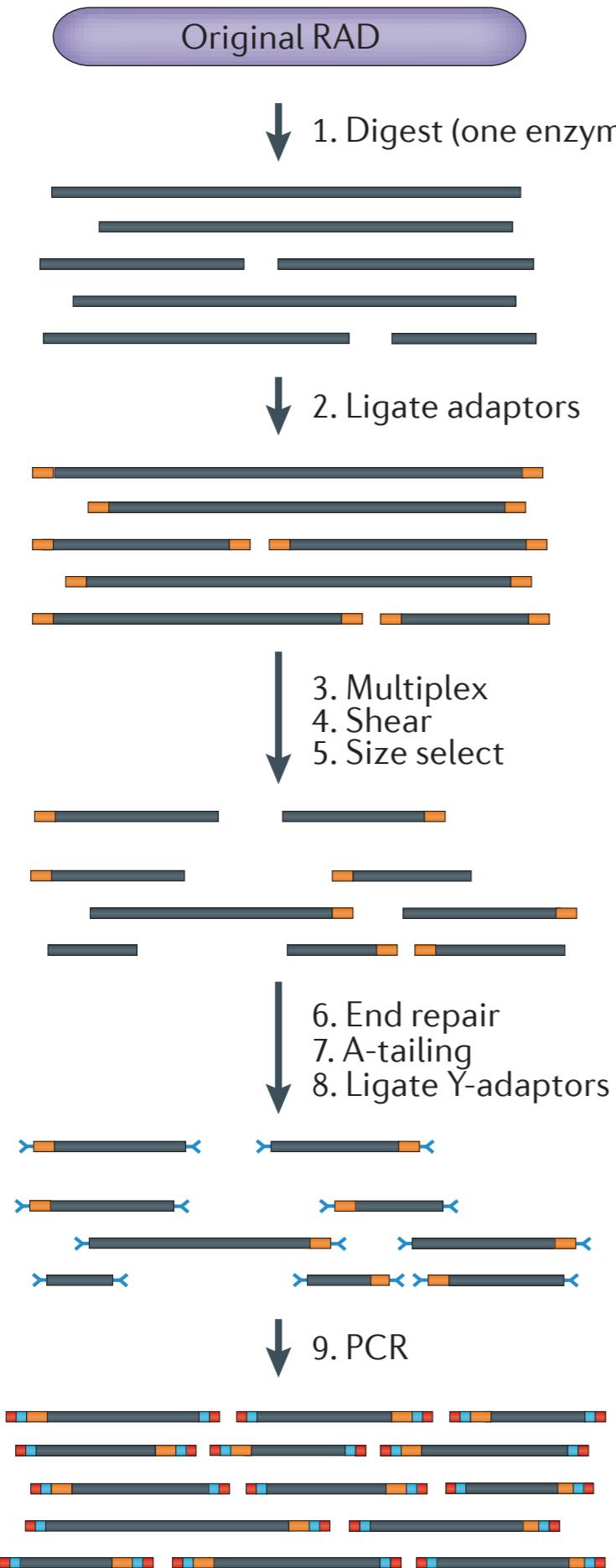


↓ 6. End repair
7. A-tailing
8. Ligate Y-adaptors

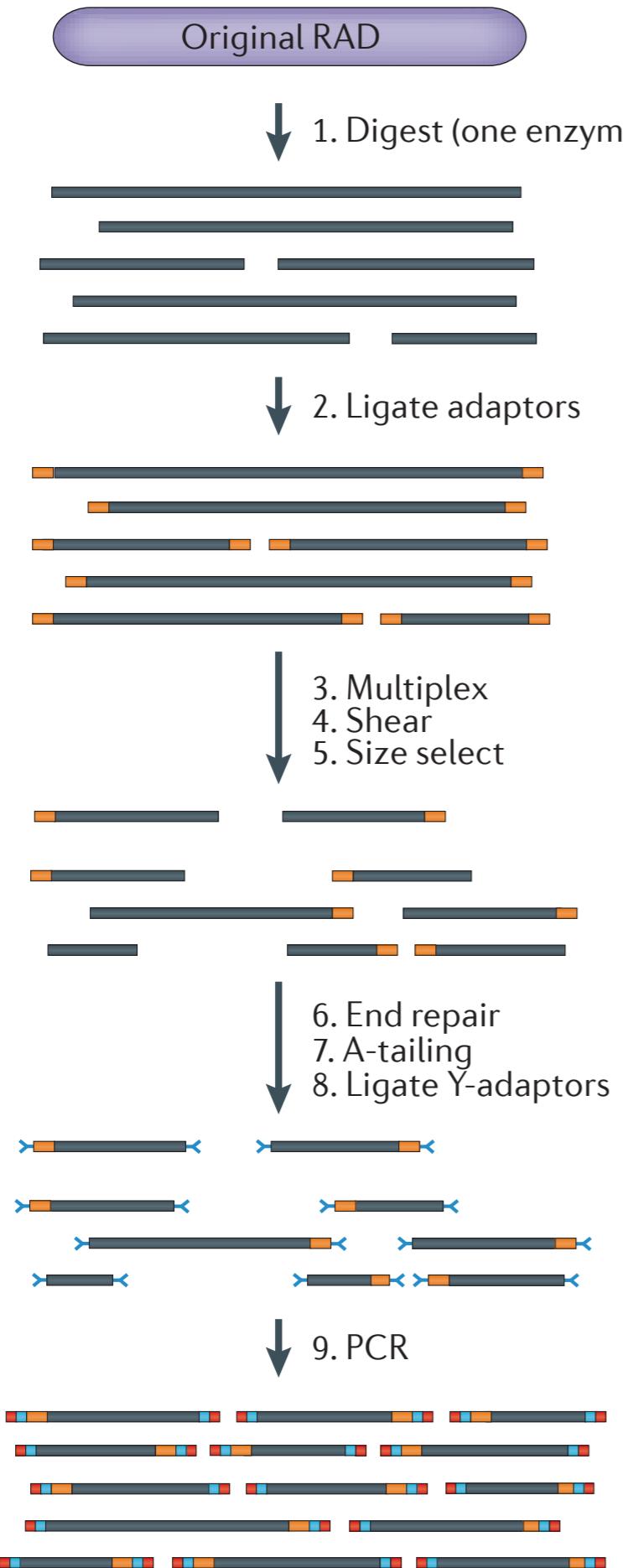


↓ 9. PCR

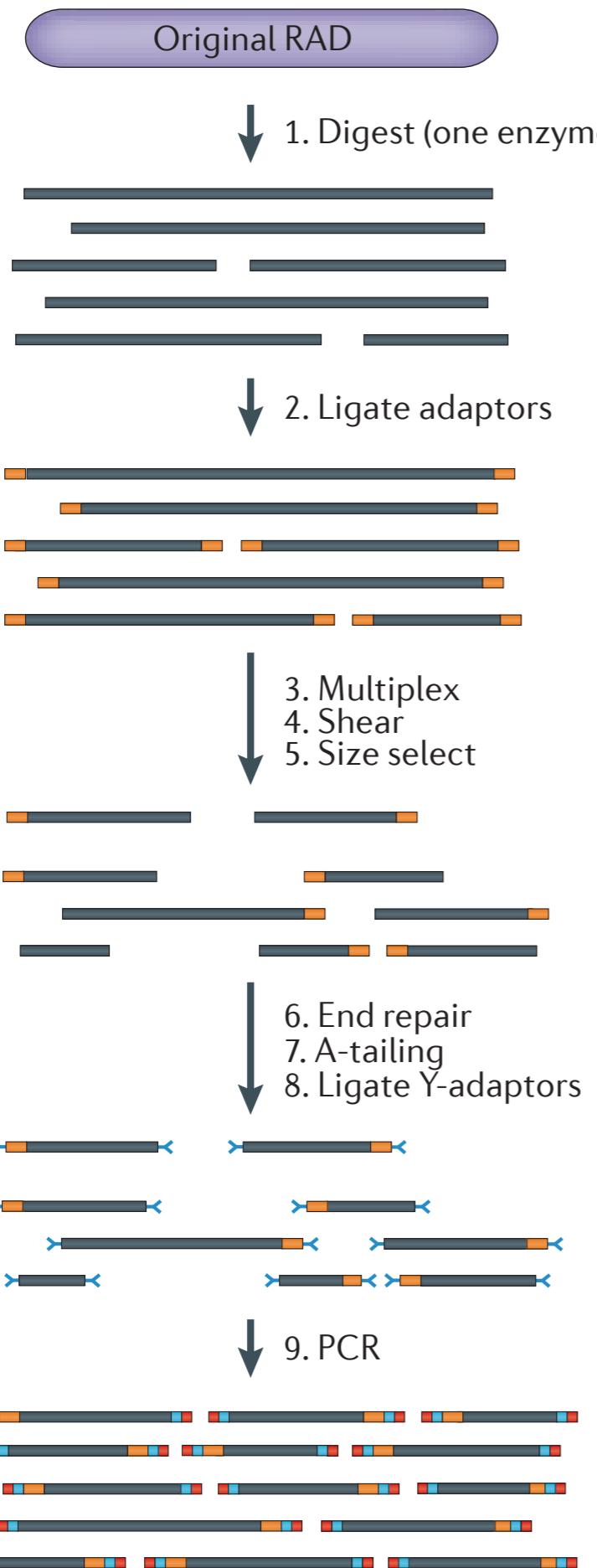




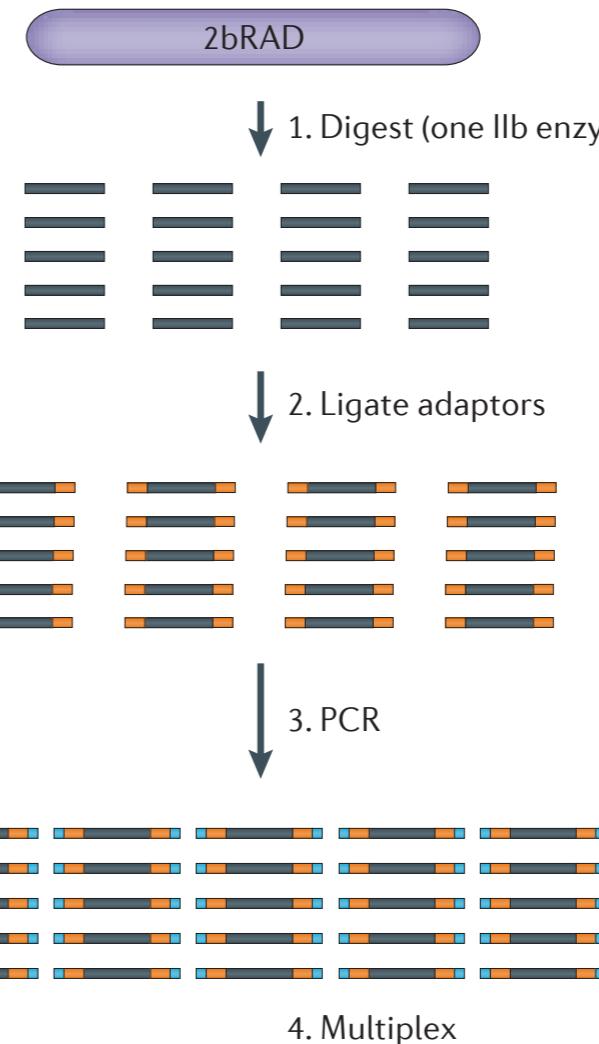
1. Sequences a subset of the genome
2. Multiplex adapters before PCR
3. Free choice of enzymes
4. Requires adapter set(s)
5. More susceptible to allele dropout



1. Sequences a subset of the genome
2. Multiplex adapters before PCR
3. Free choice of enzymes
4. Requires adapter set
5. Less susceptible to allele dropout
6. Long protocol requiring more DNA
7. Can remove PCR duplicates
8. Can produce paired end contigs
9. Read depth bias with frequent cutters



1. Sequences a subset of the genome
2. Multiplex adapters before PCR
3. Free choice of enzymes
4. Requires adapter set
5. Less susceptible to allele dropout
6. Long protocol requiring more DNA
7. Can remove PCR duplicates
8. Can produce paired end contigs
9. Read depth bias with frequent cutters



RADseq Variations

**Original RAD
double digest RAD (ddRAD)**

Genotyping by Sequencing (GBS)

ezRAD

2bRAD

Issues

Allele dropout

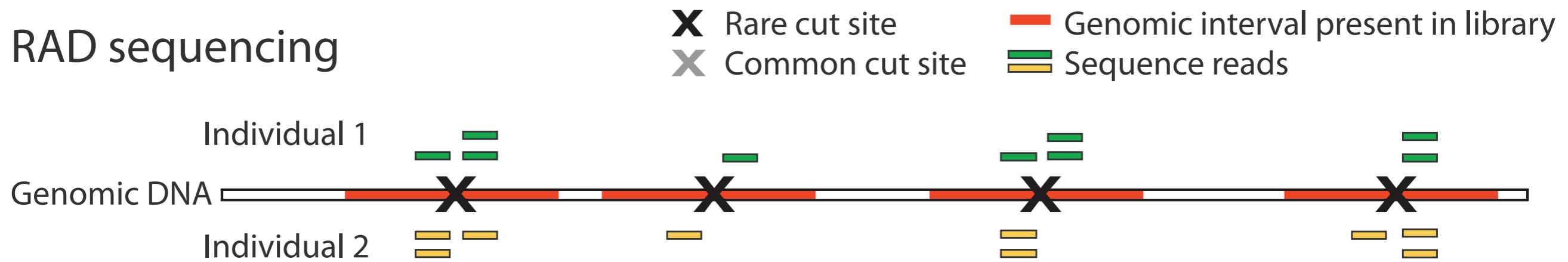
Paired end contigs

PCR duplicates

Read depth fragment length bias

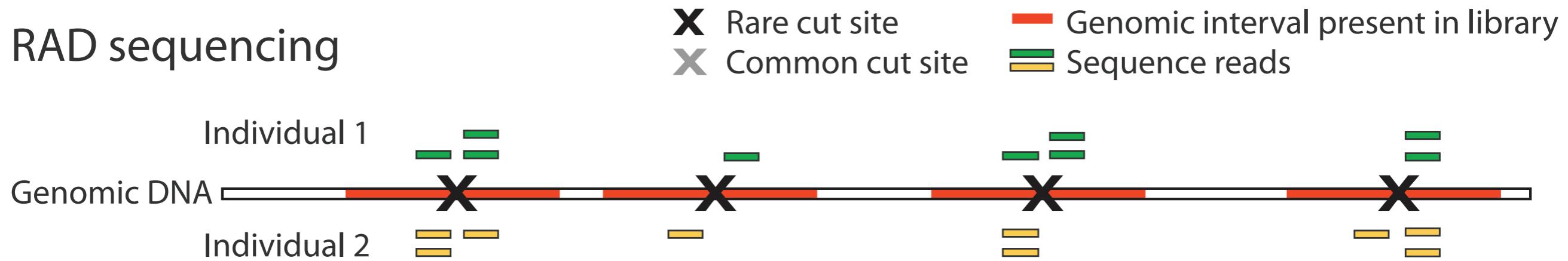
A

RAD sequencing



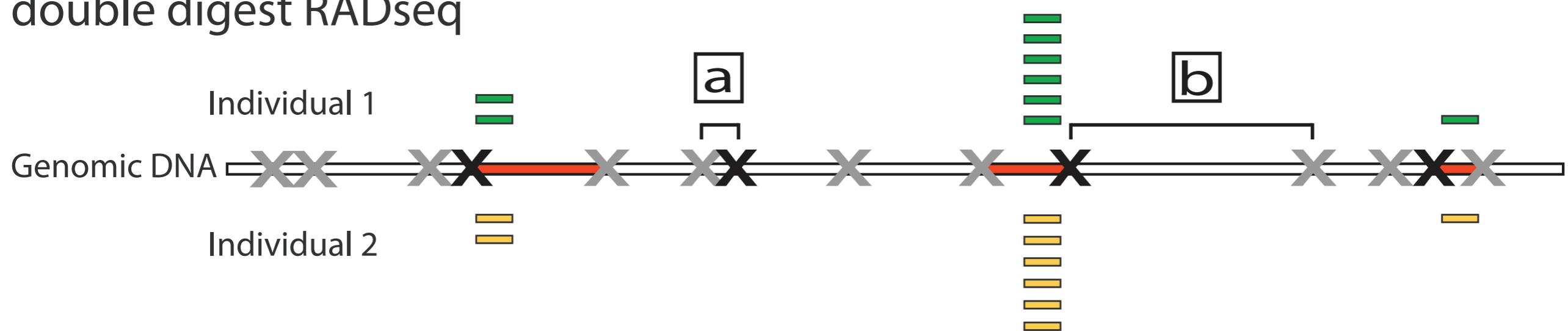
A

RAD sequencing



B

double digest RADseq



...CAGGTATAGCCTGCAGGACAGA**A**TGGAGAGTCCGTCCTGCAGGG**T**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

...CAGGTATAGCCTGCAGGACAG**C**TGGAGAGTCCGTCCTGCAGGG**A**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

...CAGGTATAGCCTGCAGGACAGA**C**AG**A**TGGAGAGTCCGTCCTGCAGGG**T**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

...CAGGTATAGCCTGCAGGACAG**C**TGGAGAGTCCGTCCTGCAGGG**A**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

...CAGGTATAGCCTGCAGGACAGA**C**AG**A**TGGAGAGTCCGTCCTTCAGGG**T**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

...CAGGTATAGCCTGCAGGACAG**C**TGGAGAGTCCGTCCTGCAGGG**A**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

...CAGGTATAGCCTGCAGGACAGA**A**TGGAGAGTCCGTCCTGCAGGG**T**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

...CAGGTATAGCCTGCAGGACAG**C**TGGAGAGTCCGTCCTGCAGGG**A**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

...CAGGTATAGCCTGCAGGACAGA**A**TGGAGAGTCCGTCCTTCAGGG**T**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

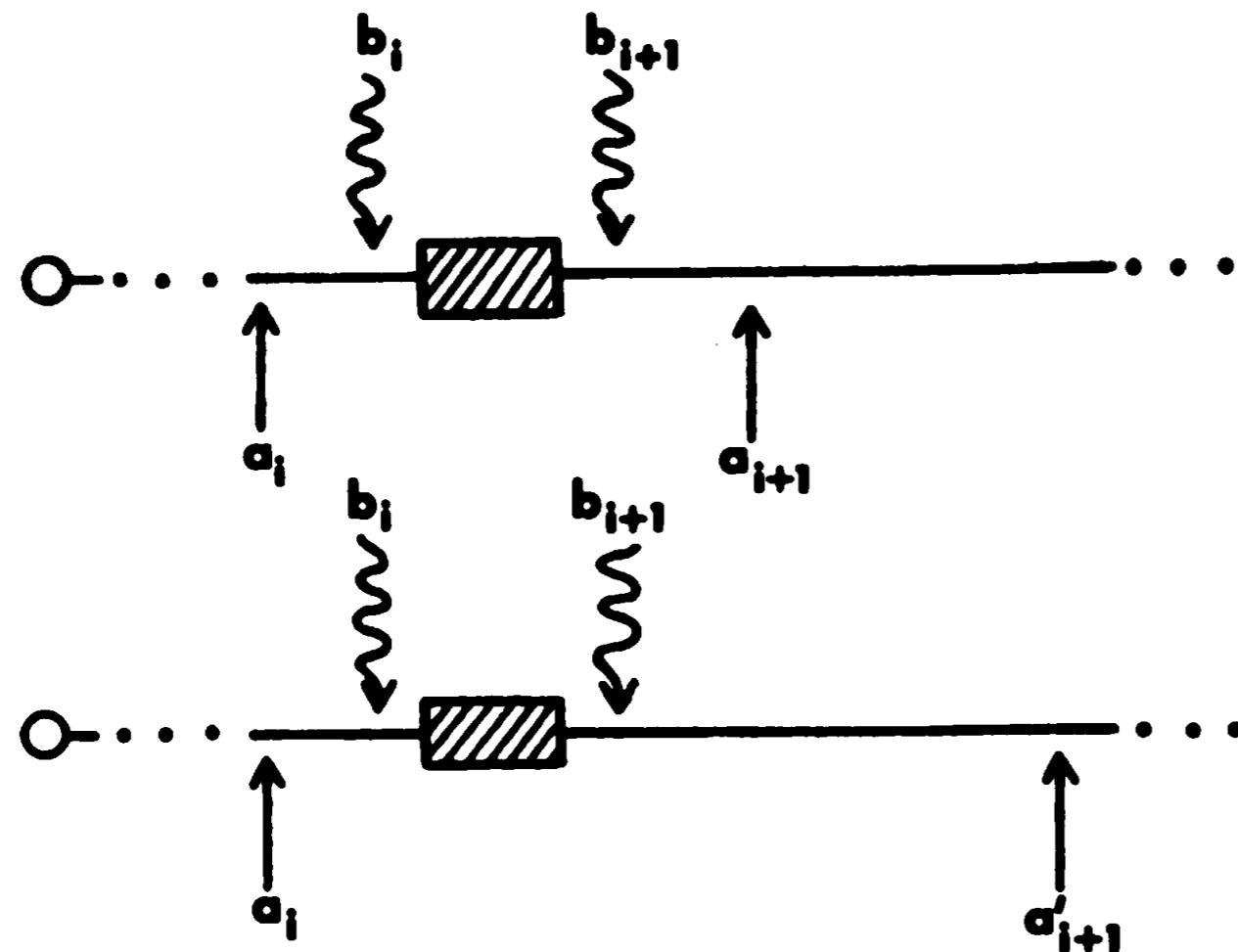
...CAGGTATAGCCTGCAGGACAG**C**TGGAGAGTCCGTCCTGCAGGG**A**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

...CAGGTATAGCCTGCAGGACAGA**A**TGGAGAGTCCGTCCTGCAGGG**T**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

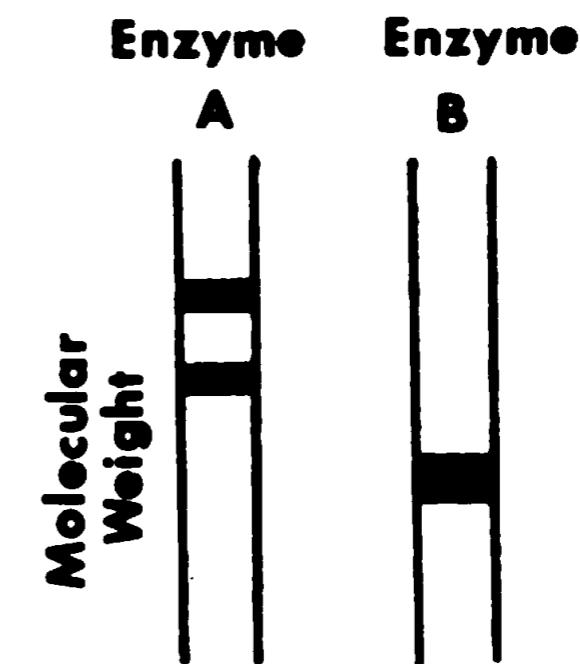
...CAGGTATAGCCTGCAGGACAG**C**TGGAGAGTCCGTCCTGCAGGG**A**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

...CAGGTATAGCCTGCAGGACAGA**B**TGGAGAGTCCGTCCT**T**CAGGGG**T**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

...CAGGTATAGCCTGCAGGACAG**C**TGGAGAGTCCGTCCTGCAGGG**A**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...



a. Chromosomal Arrangement



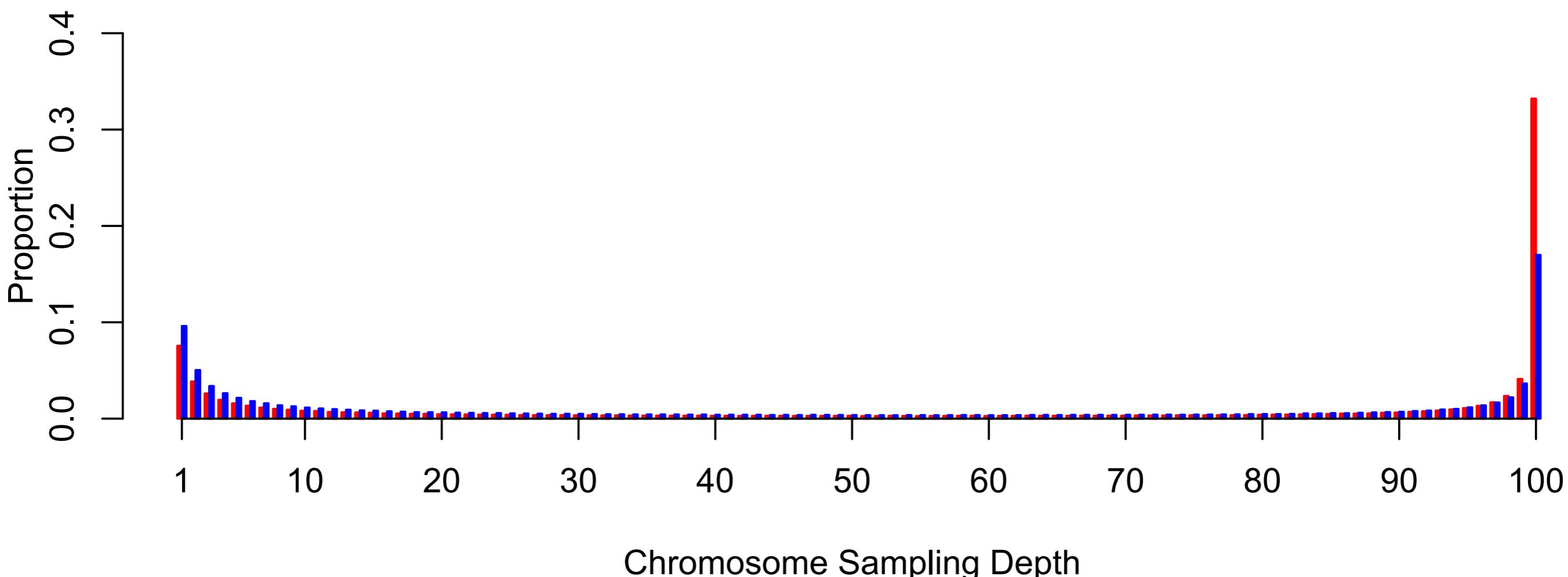
b. Hybridization Pattern

...CAGGTATAGCCTGCAGGACAGACAG**A**TGGAGAGTCCGTCCTGCAGGG**T**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

...CAGGTATAGCCTGCAGGACAG**C**TGGAGAGTCCGTCCTGCAGGG**A**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

...CAGGTATAGCCTGCAGGACAGACAG**A**TGGAGAGTCCGTCCTTCAAGGG**T**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

...CAGGTATAGCCTGCAGGACAG**C**TGGAGAGTCCGTCCTGCAGGG**A**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

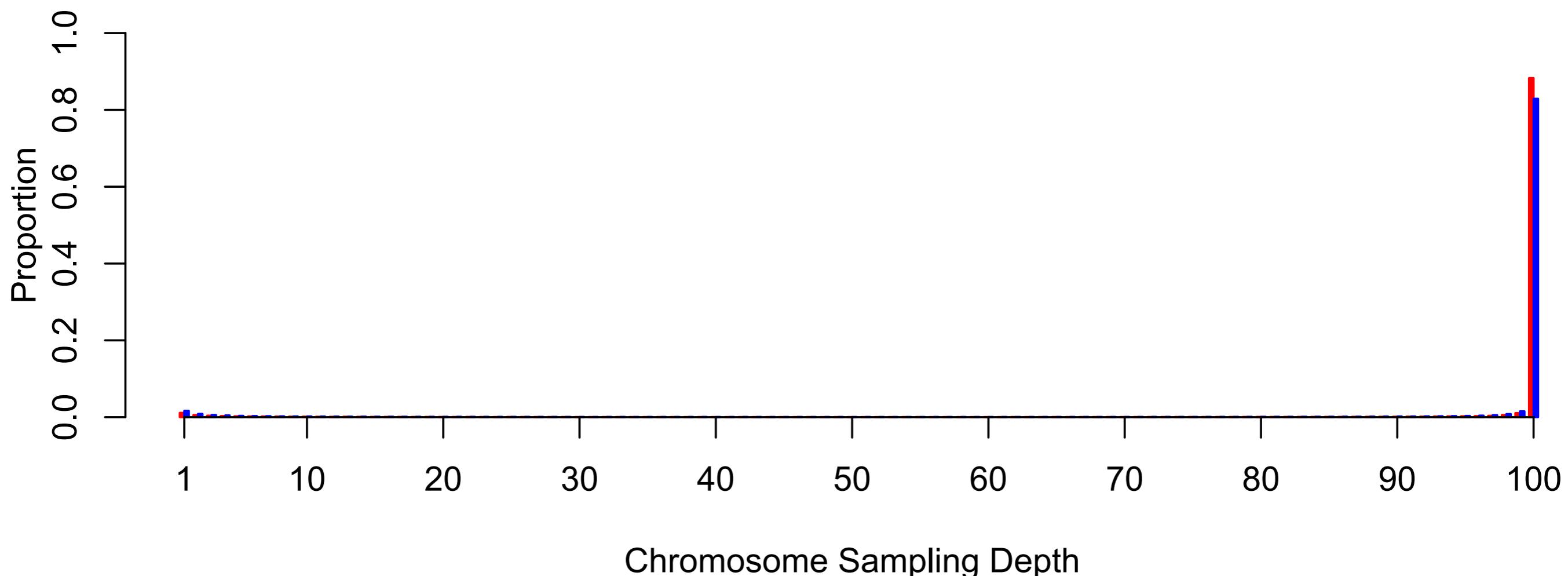


...CAGGTATAGCCTGCAGGACAGA**A**TGGAGAGTCCGTCCTGCAGGG**T**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

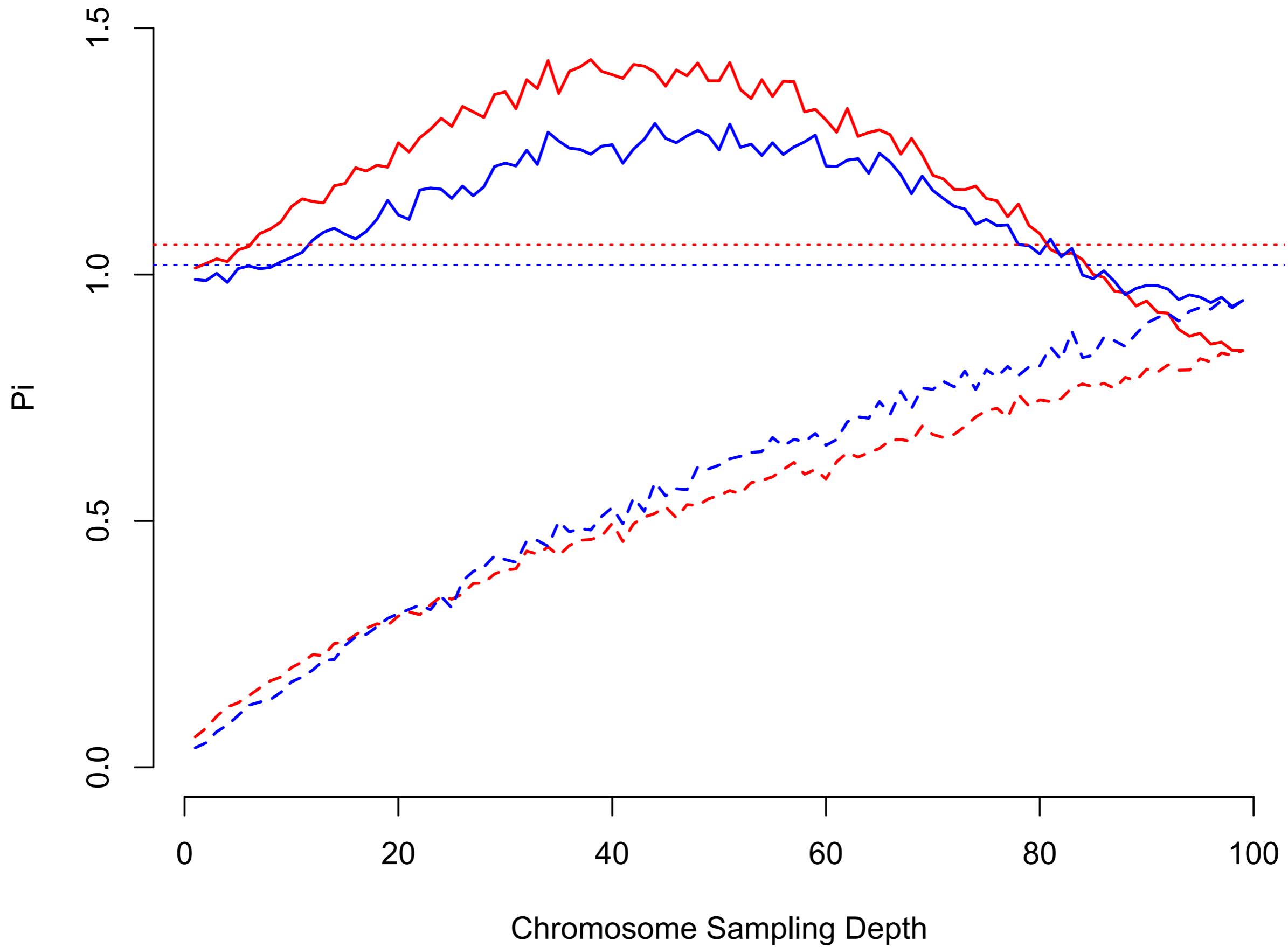
...CAGGTATAGCCTGCAGGACAG**C**TGGAGAGTCCGTCCTGCAGGG**A**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

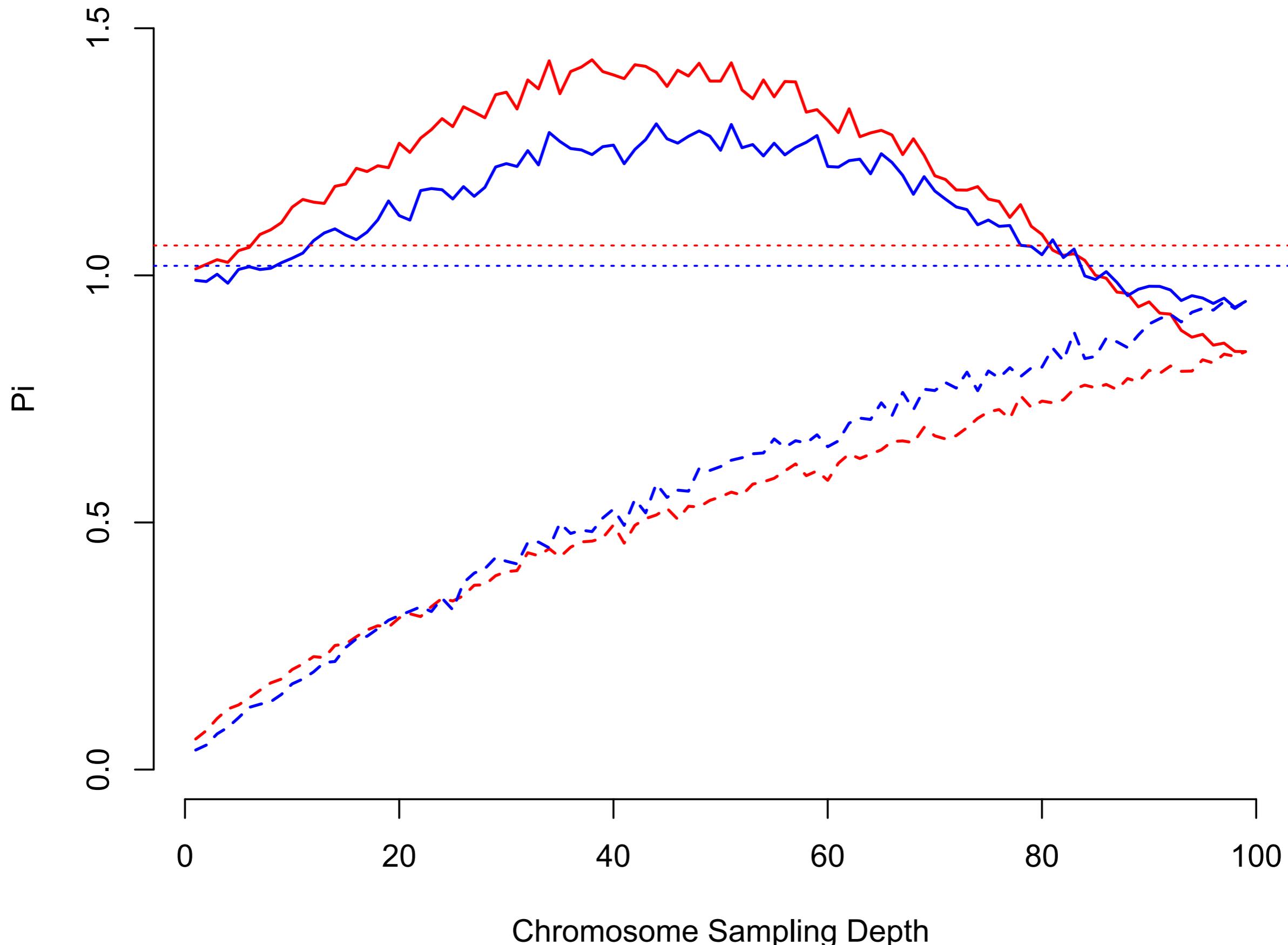
...CAGGTATAGCCTGCAGGACAGA**A**TGGAGAGTCCGTCCT**T**CAGGG**T**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

...CAGGTATAGCCTGCAGGACAG**C**TGGAGAGTCCGTCCTGCAGGG**A**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...



Theta=Rho=0.001/bp





...CAGGTATAGCCTGCAGGACAGA**A**TGGAGAGTCCGTCCTGCAGGG**T**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

...CAGGTATAGCCTGCAGGACAG**C**TGGAGAGTCCGTCCTGCAGGG**A**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

...CAGGTATAGCCTGCAGGACAGA**A**TGGAGAGTCCGTCCTTCAGGG**T**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

...CAGGTATAGCCTGCAGGACAG**C**TGGAGAGTCCGTCCTGCAGGG**A**ACTCTAGACTCCTCAGCCTGCAGGAAATAGACCGTCAGT...

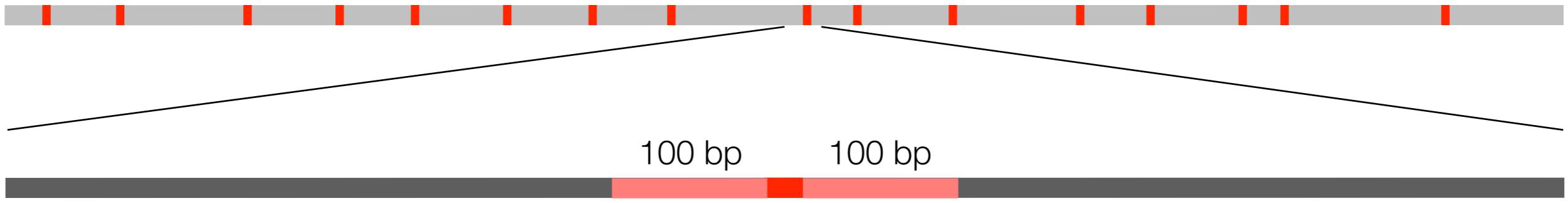
Risks: underestimation of genetic diversity
overestimation of Fst

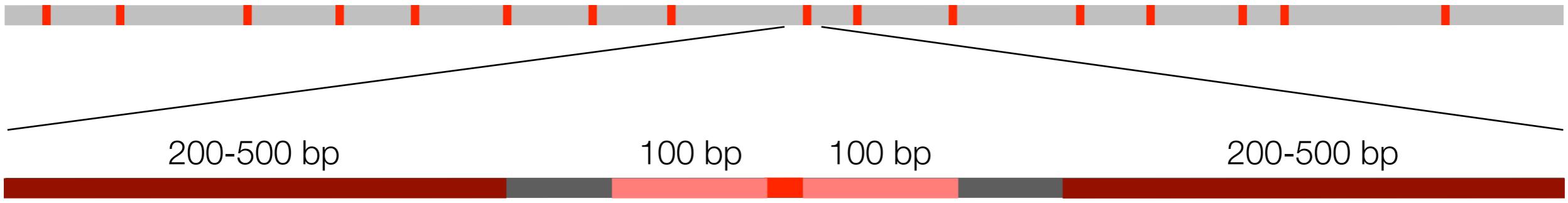
Can attempt to remove these alleles:

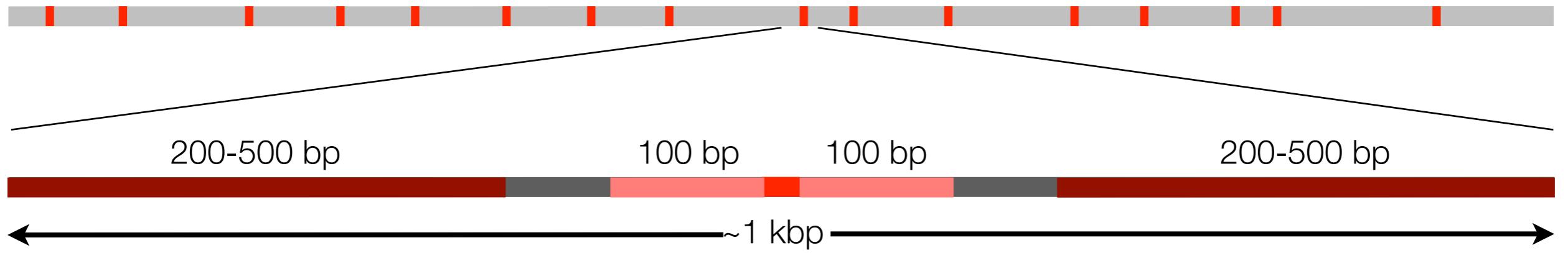
- by coverage
- by presence of any allele at locus across individuals

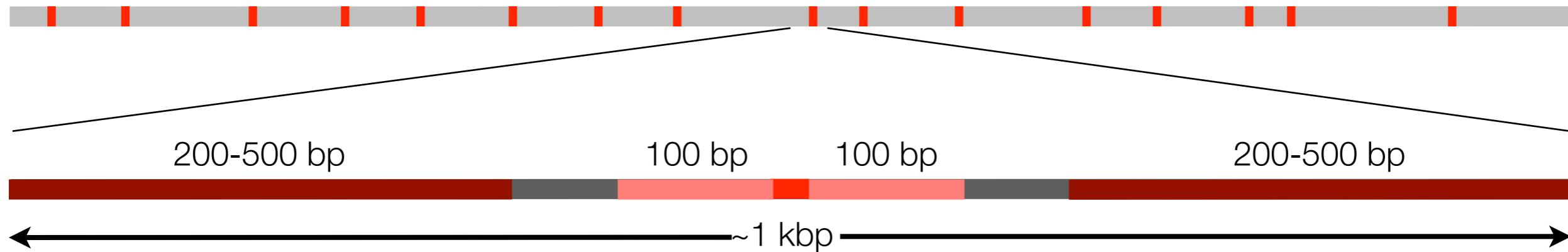


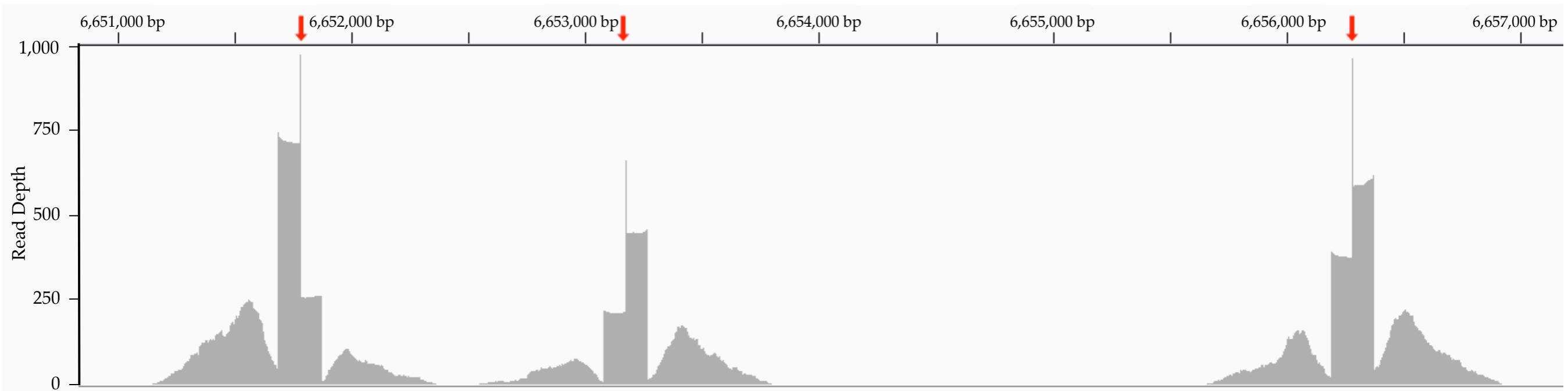
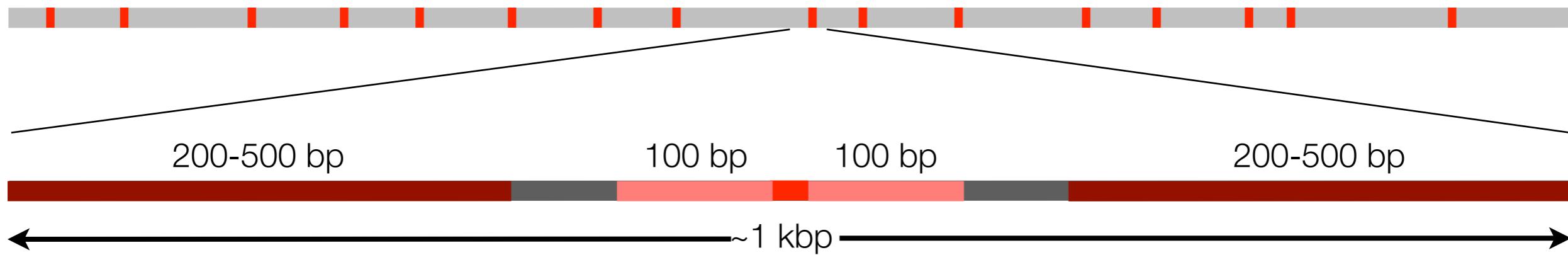


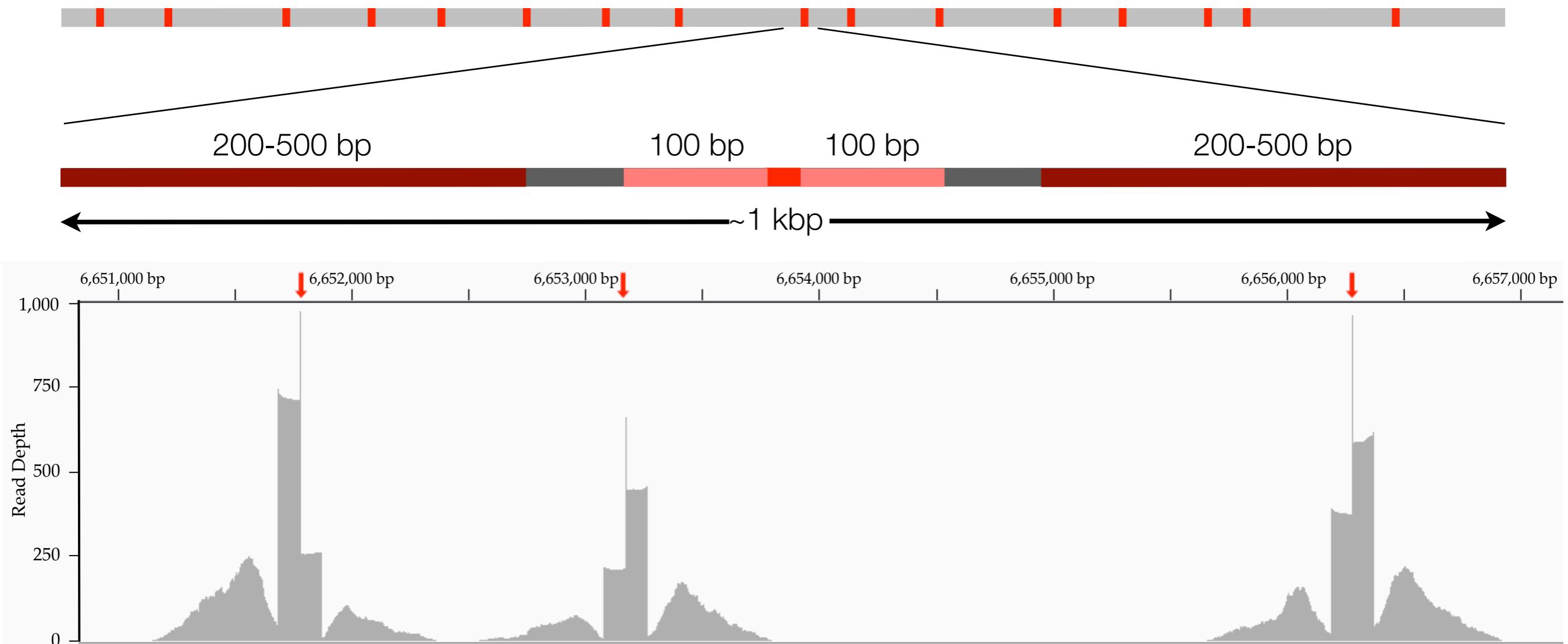










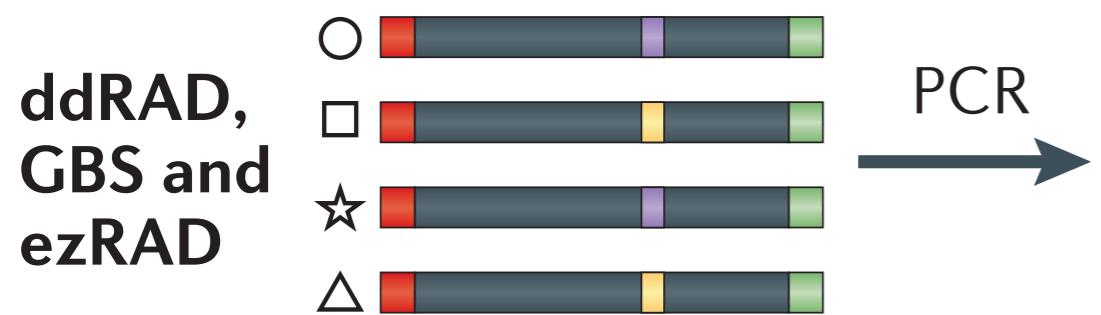


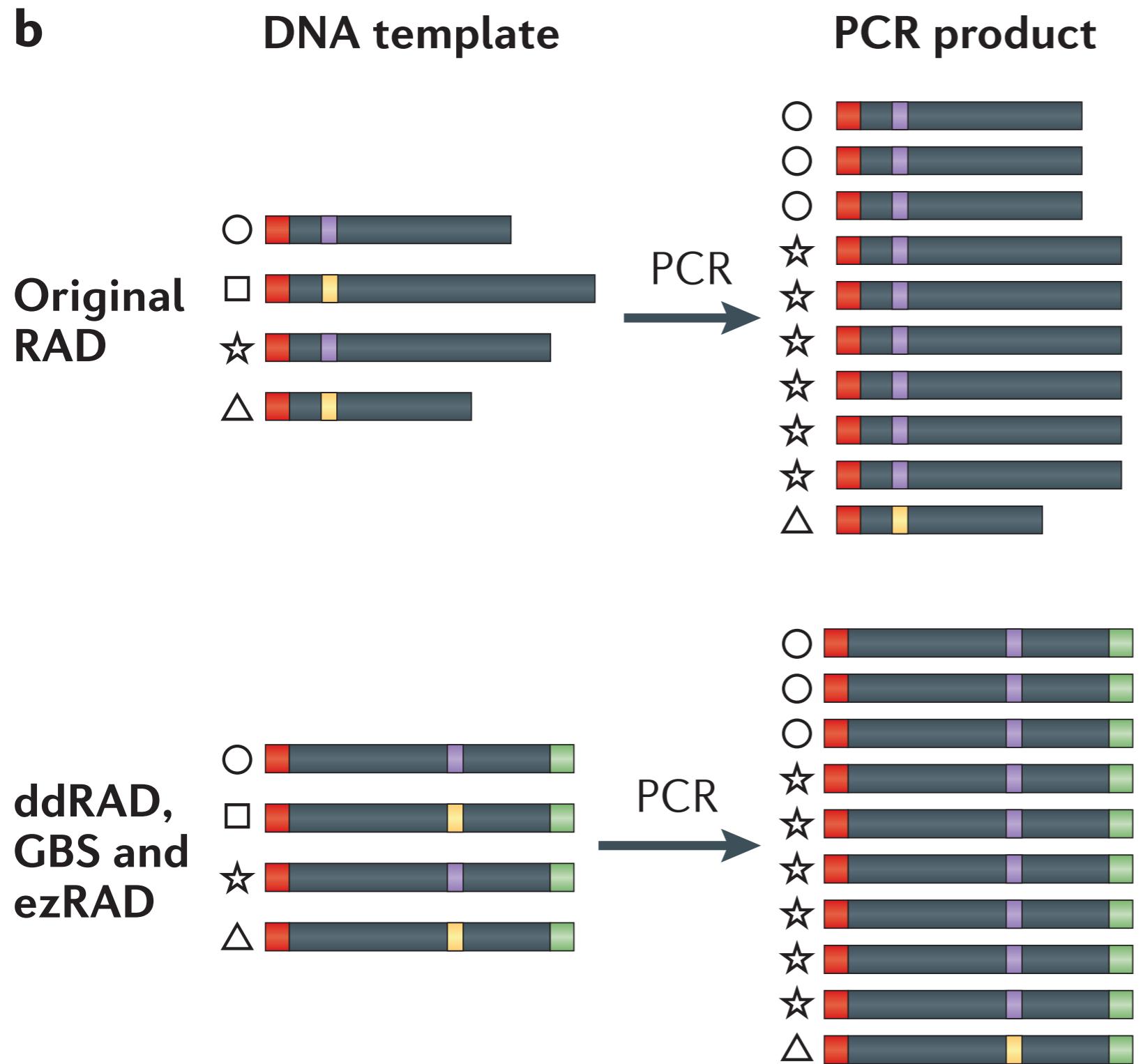
Paired end assemblies can be used for BLASTing:

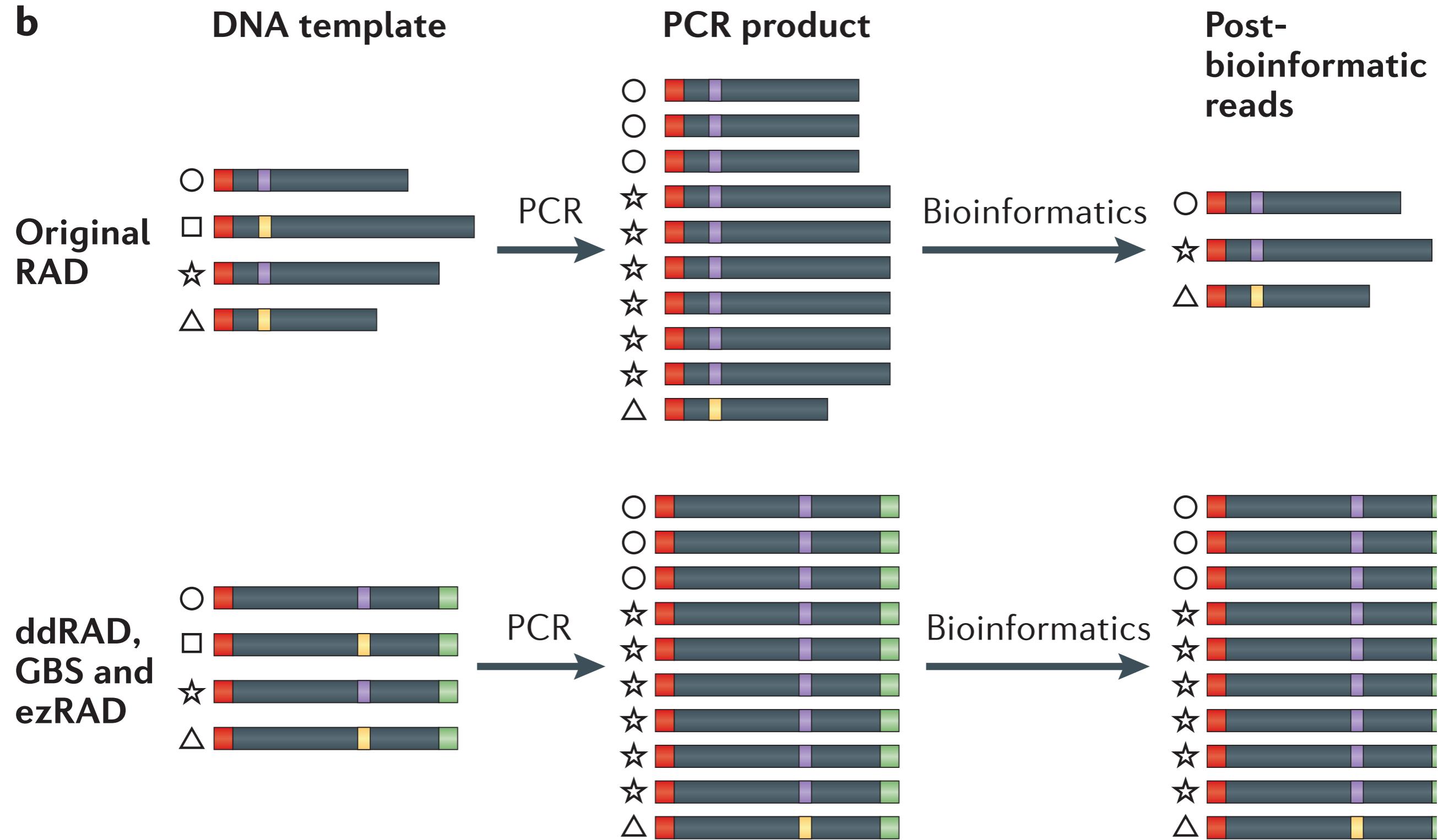
- to a nearby species
- to gene sets
- (and maybe SNP calls, but need very high coverage)

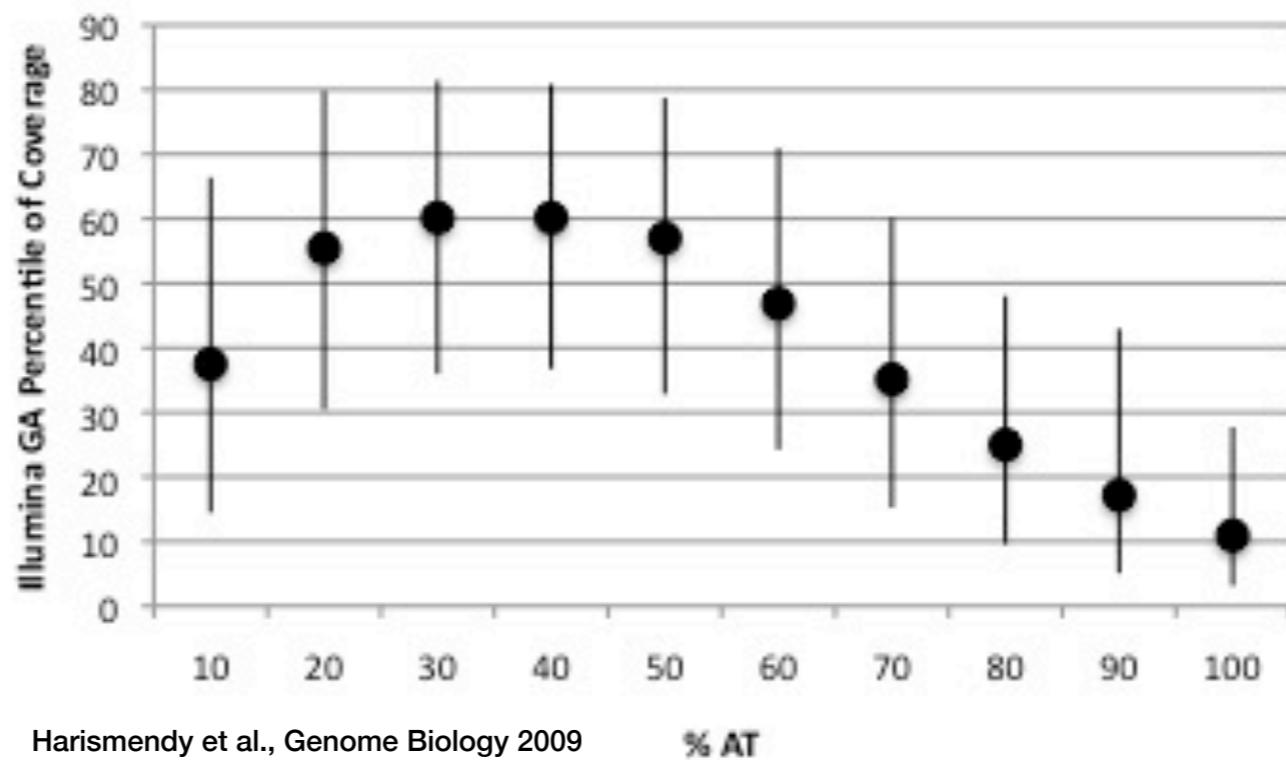
b

DNA template



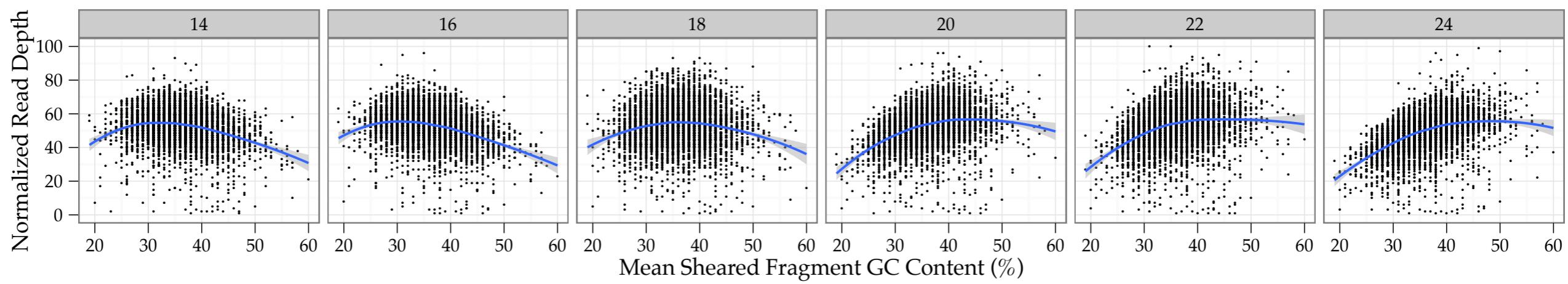
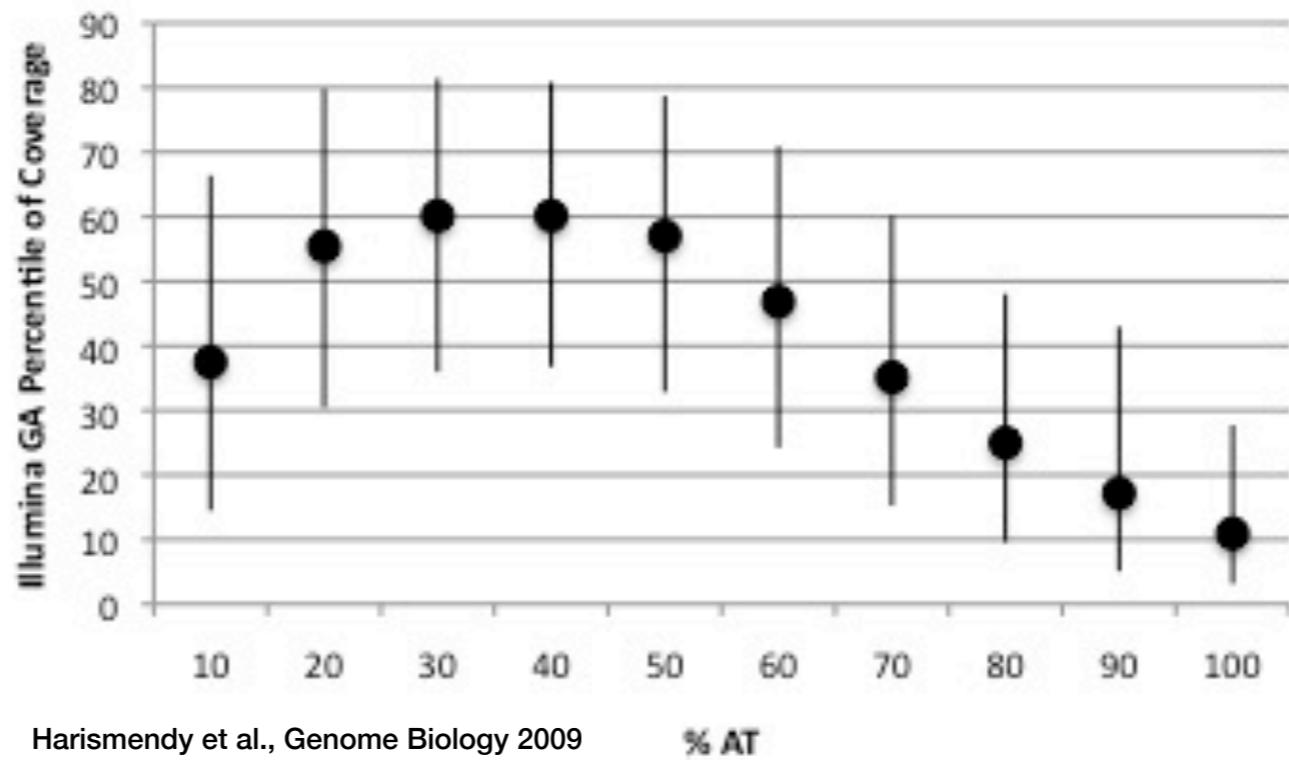
b

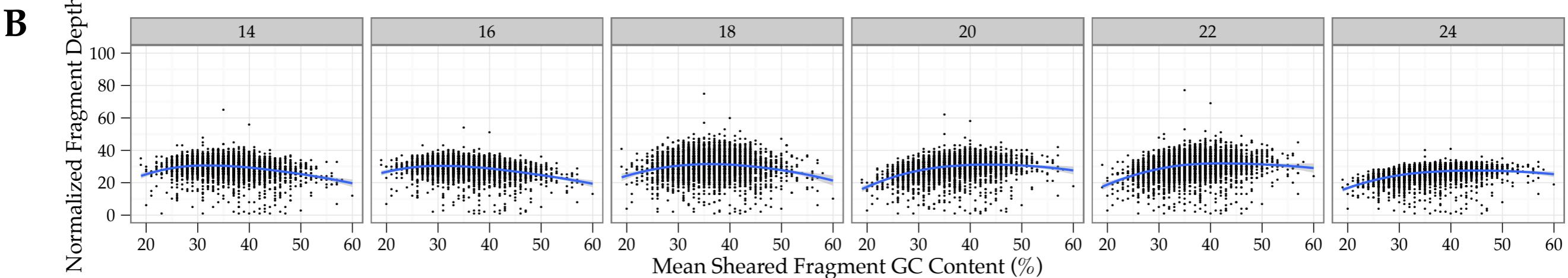
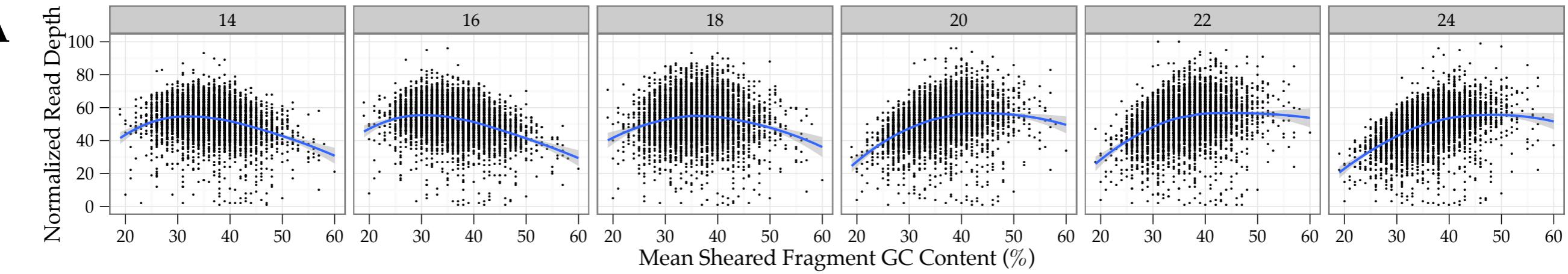
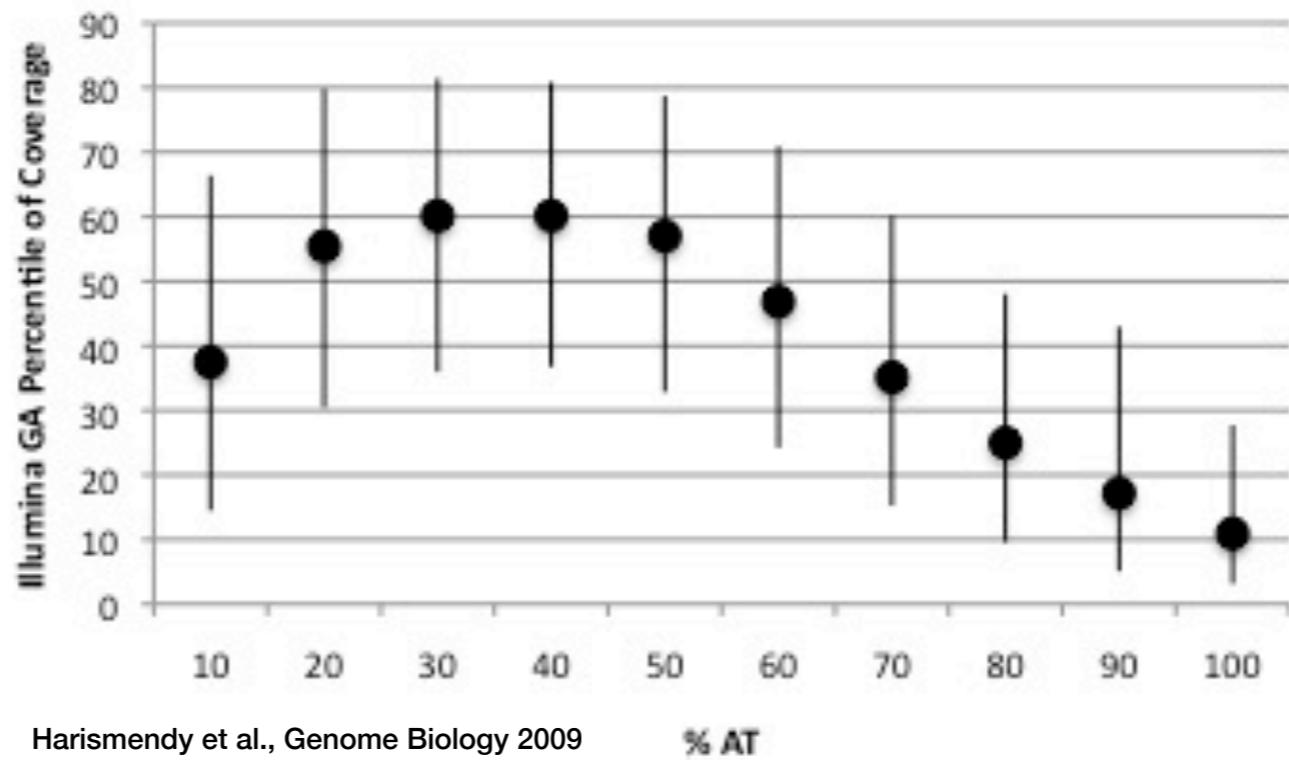
b

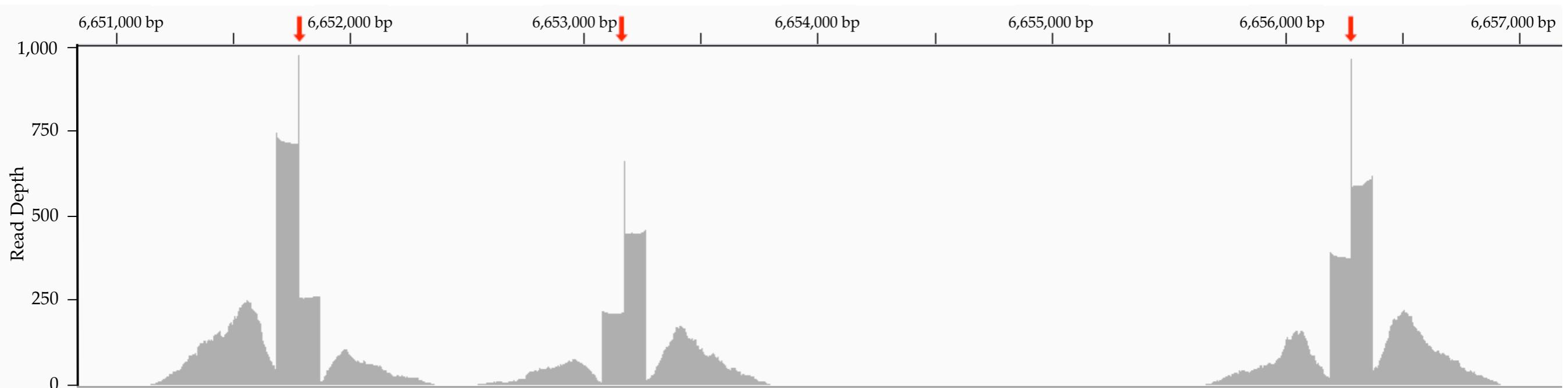


Harismendy et al., Genome Biology 2009

% AT



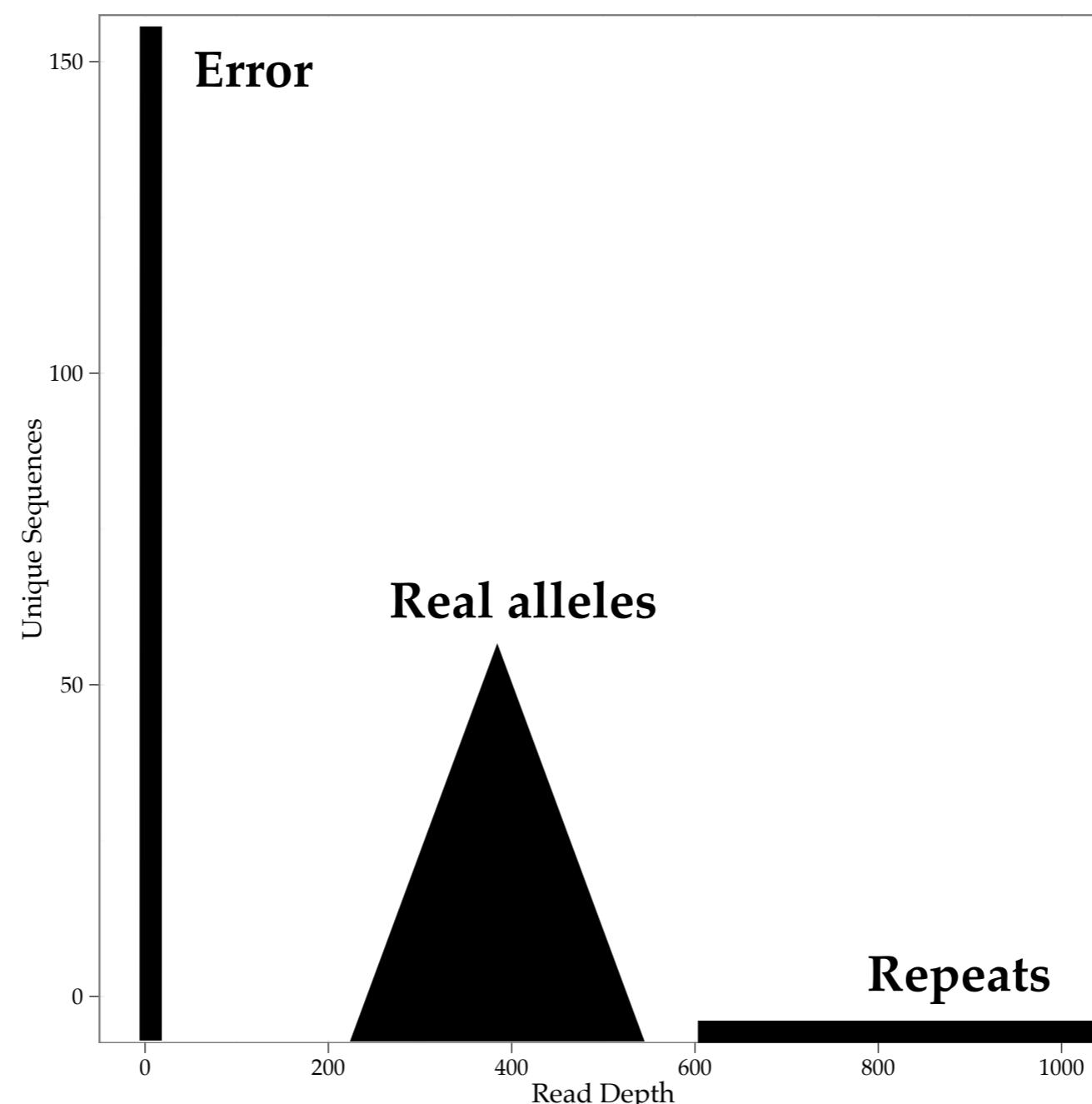
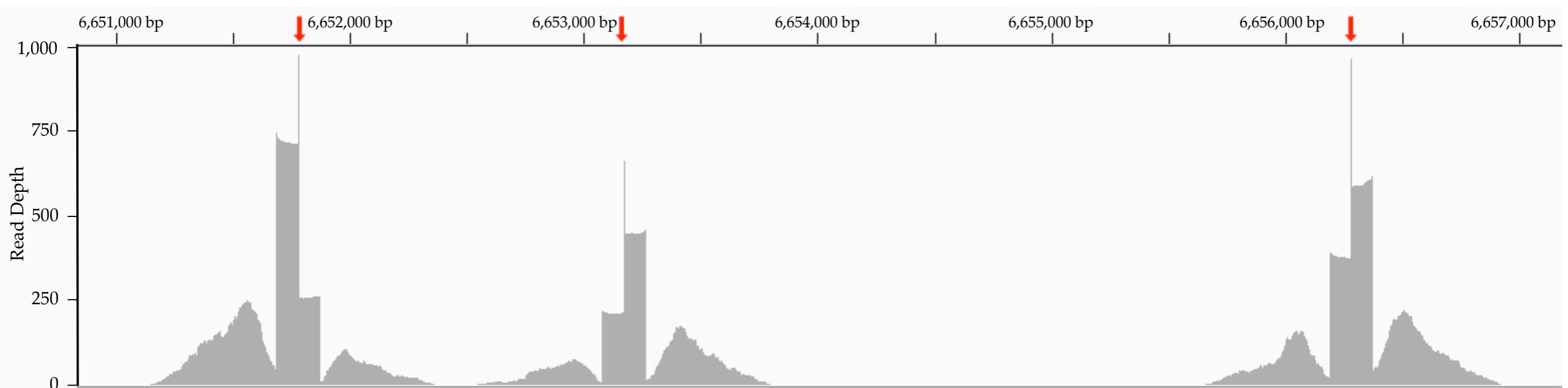


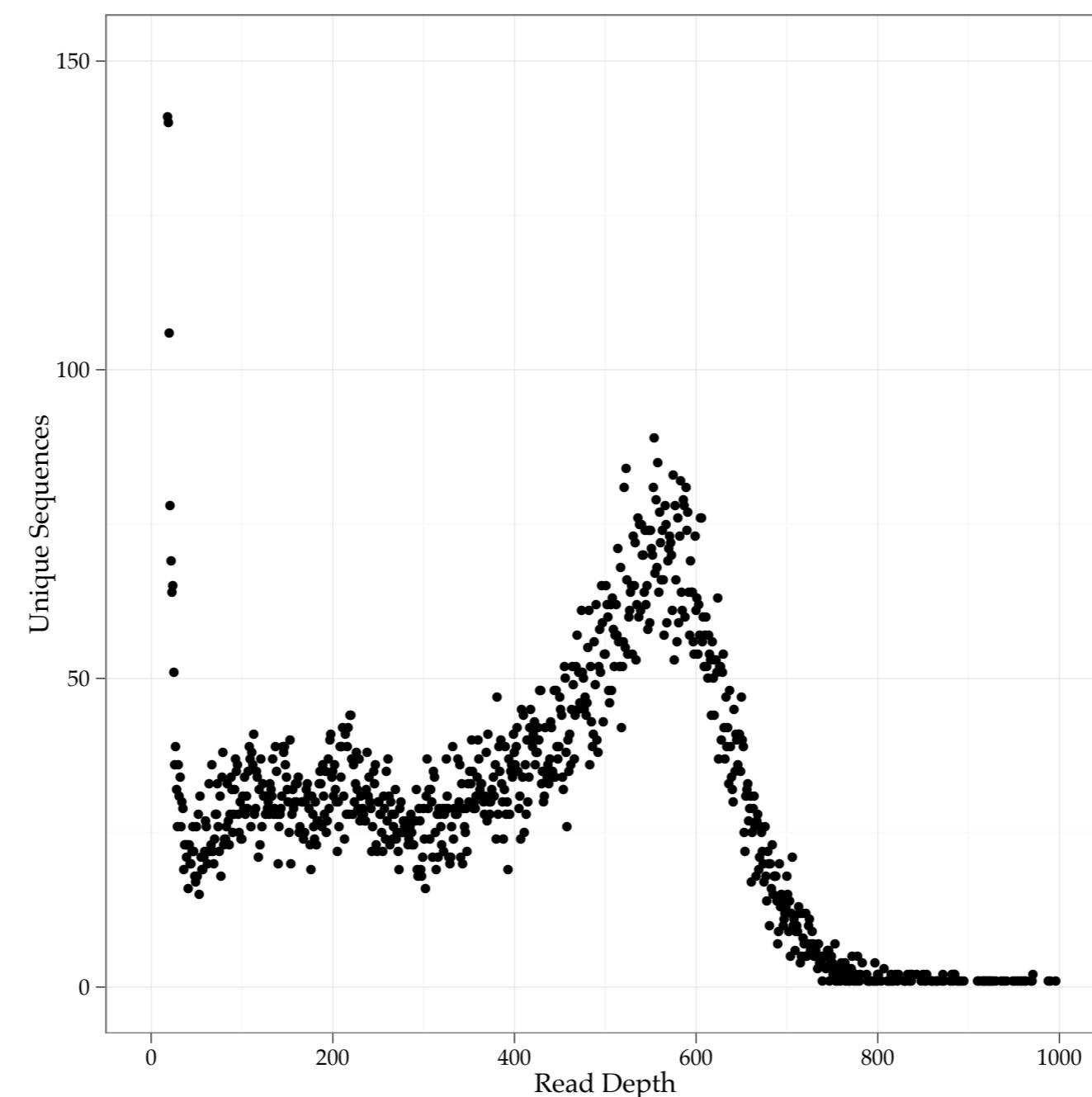
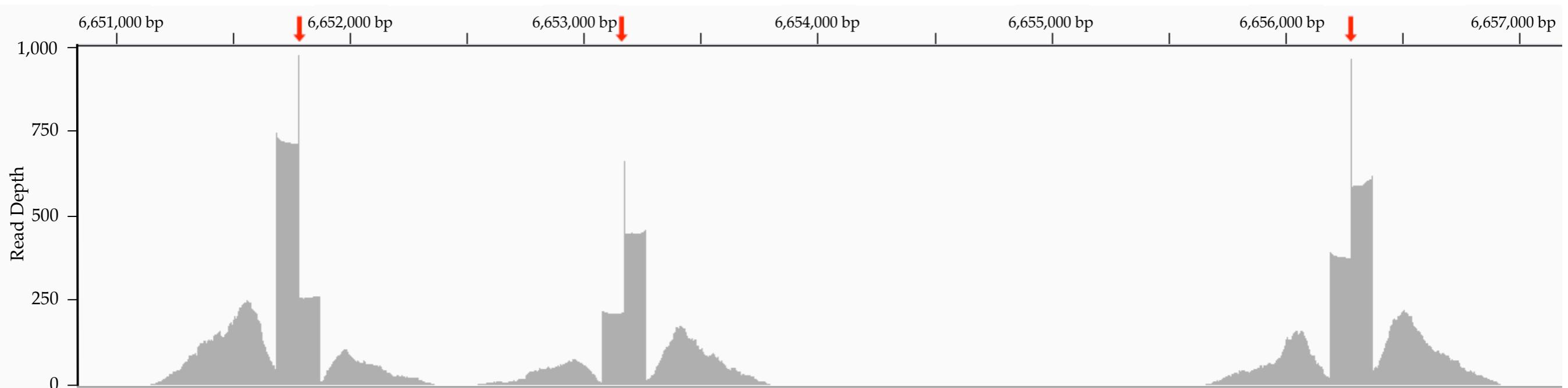


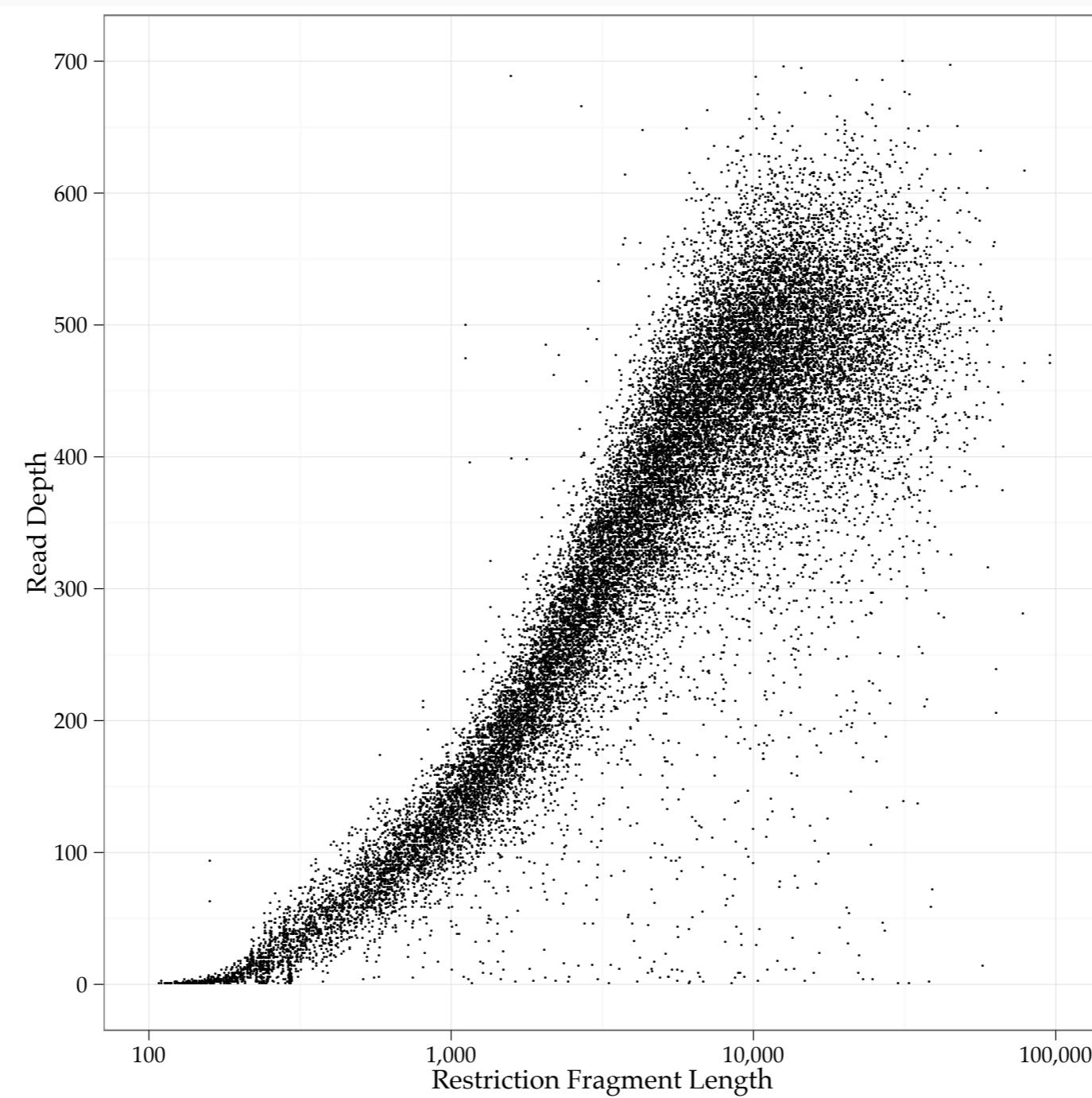
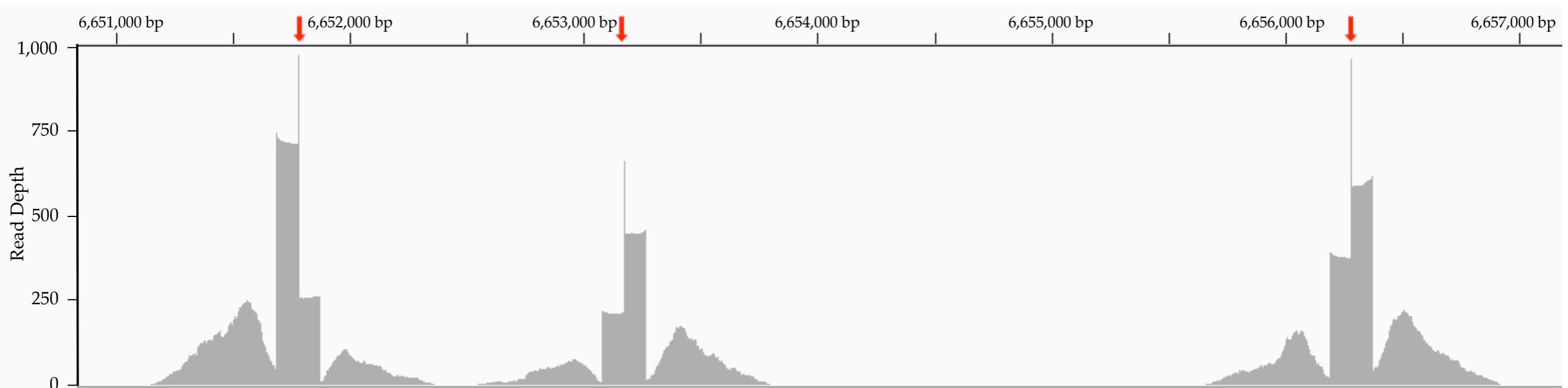
PCR duplicates use up sequencing reads

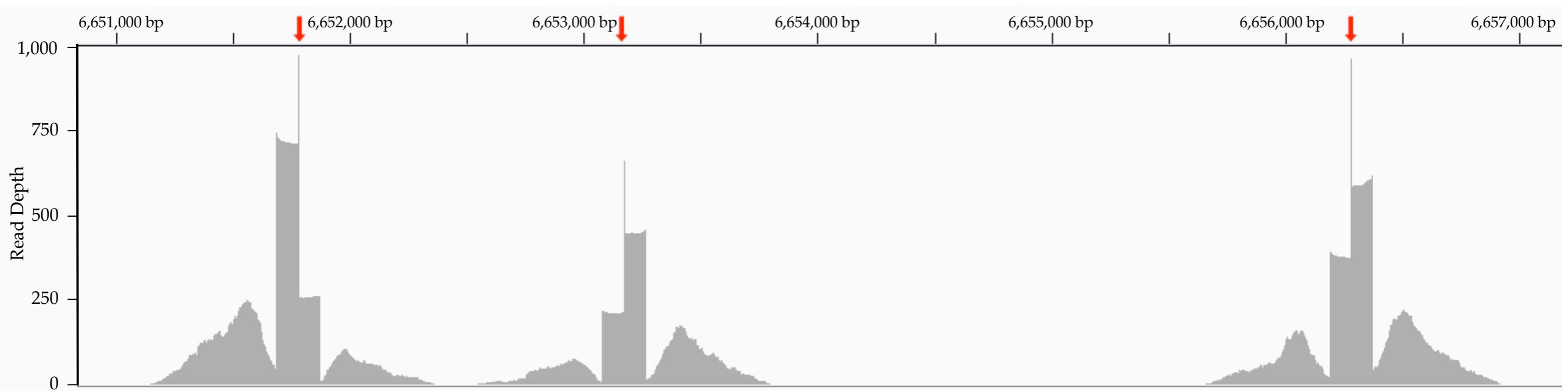
But do they make analysis harder or less correct?

- may make it harder to separate loci by coverage
- may unbalance alleles at a locus
- but 'false SNPs' in one sample are easily ignored







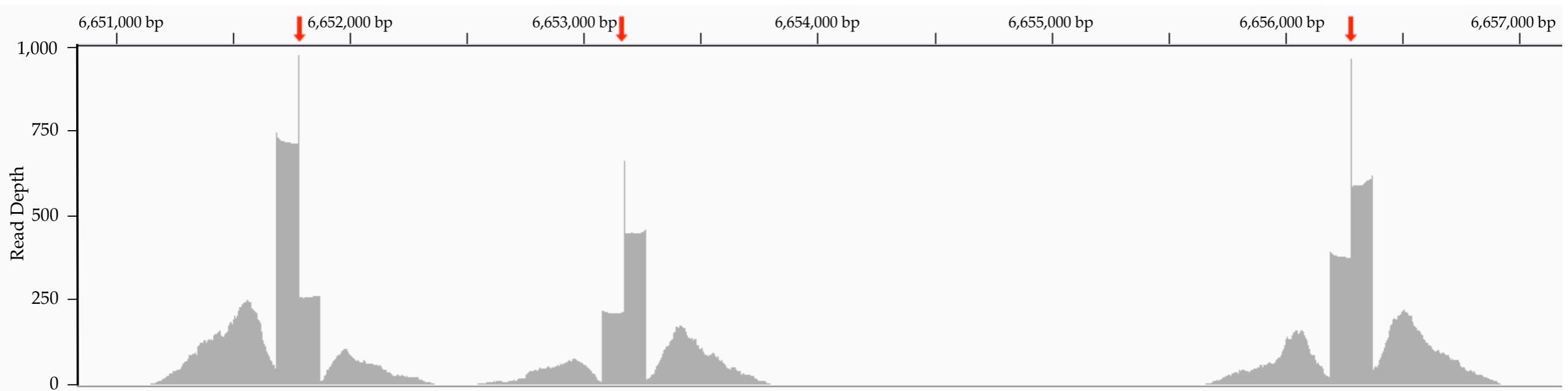


F1 Mother

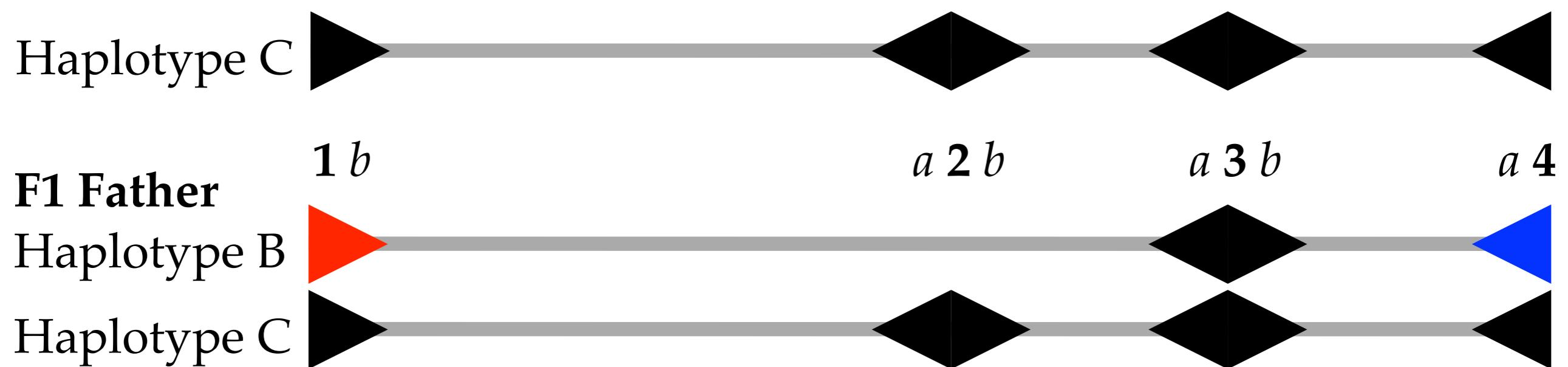


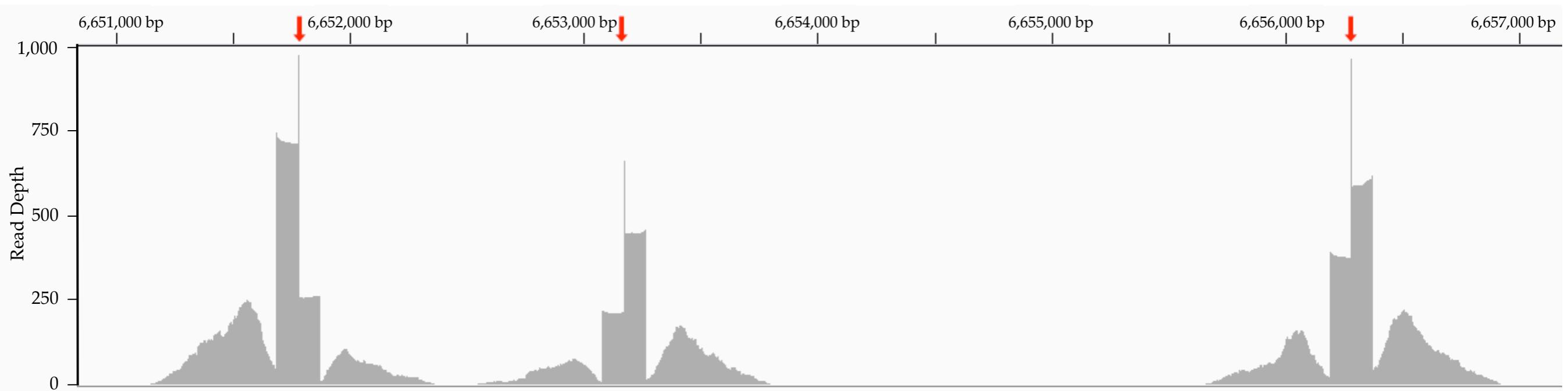
F1 Father





F1 Mother

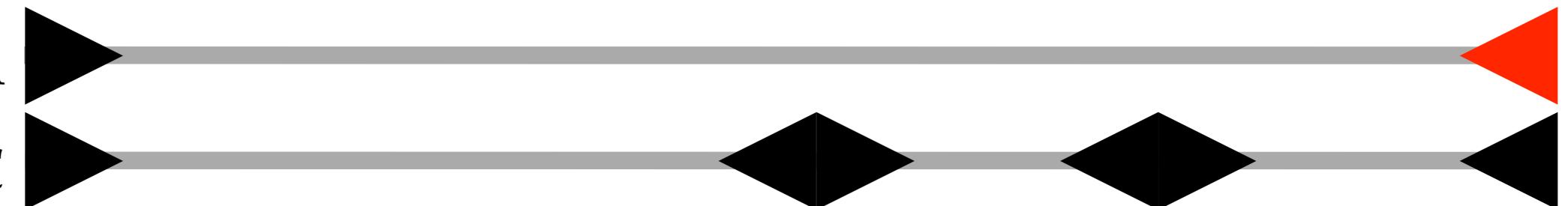




F1 Mother

Haplotype A

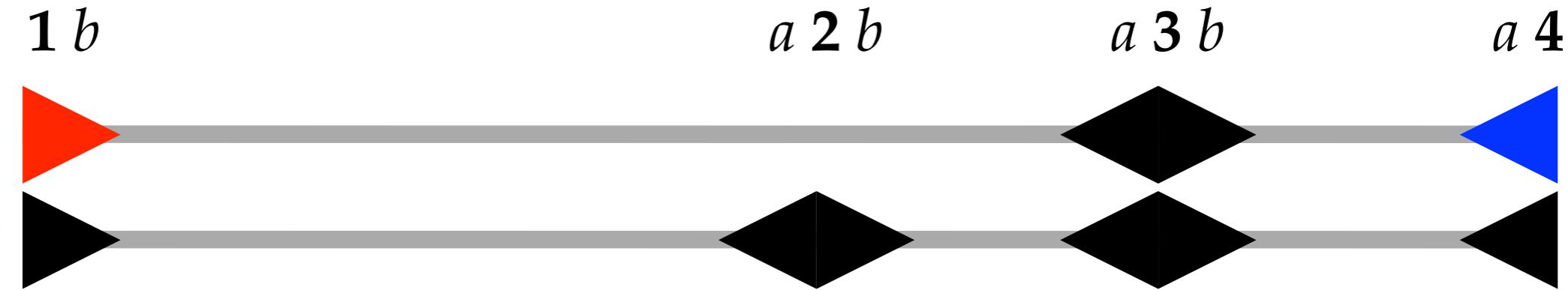
Haplotype C



F1 Father

Haplotype B

Haplotype C



RADseq Variations

Original RAD
double digest RAD (ddRAD)
Genotyping by Sequencing (GBS)
ezRAD
2bRAD

Issues

Allele dropout
PCR duplicates
Paired end contigs
Read depth fragment length bias