20.109
Laboratory Fundamentals in Biological Engineering

Module 1

Nucleic Acid Engineering

Lecture 7

Office hours: by appt.

Fundamental themes in BE

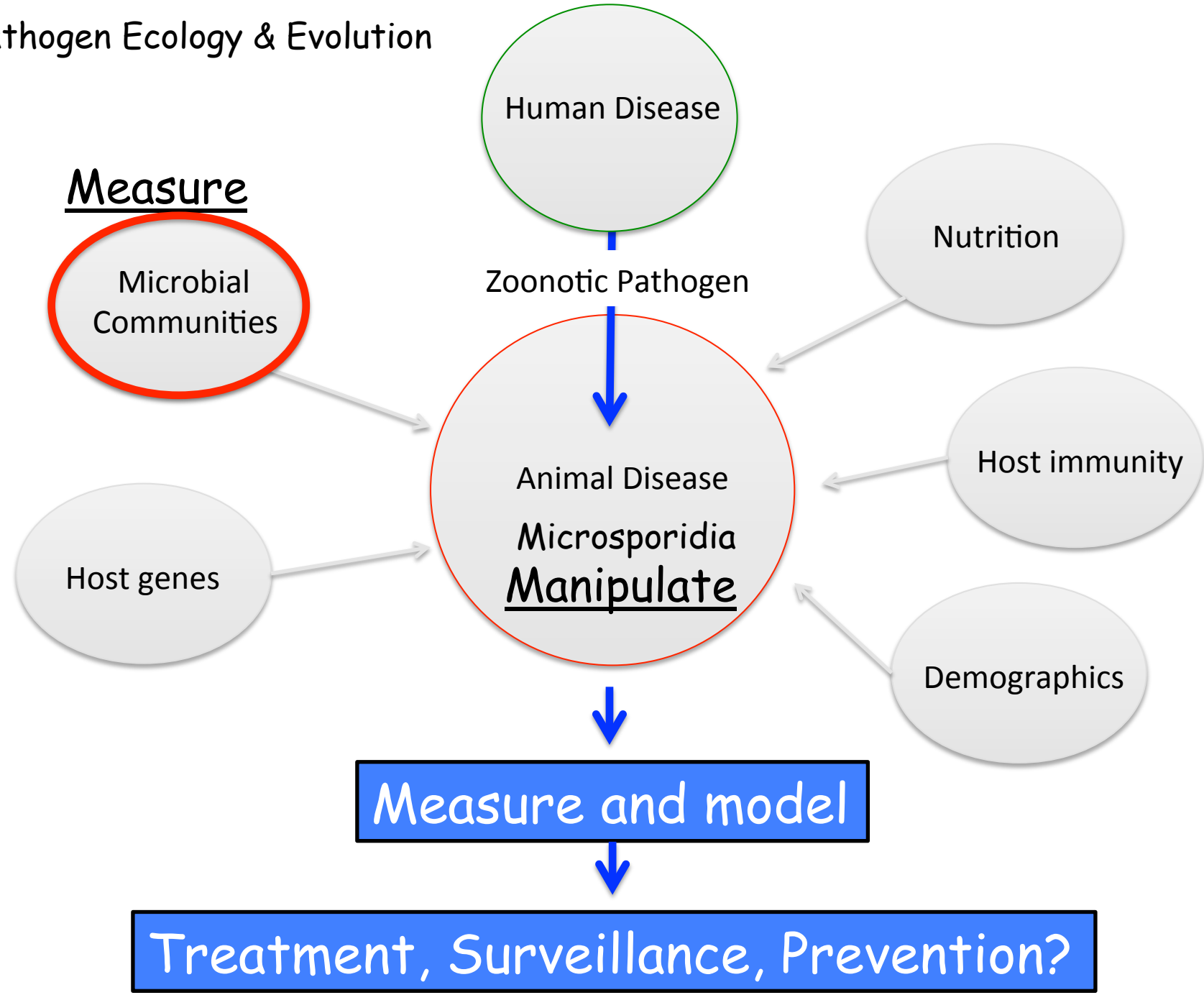
Manipulate → Make

Measure → Model

Module 1 – DNA Objectives

1. Introduce to critical methodology (DNA extraction, PCR, cloning, sequencing, analysis)
2. Do in the context of the major themes

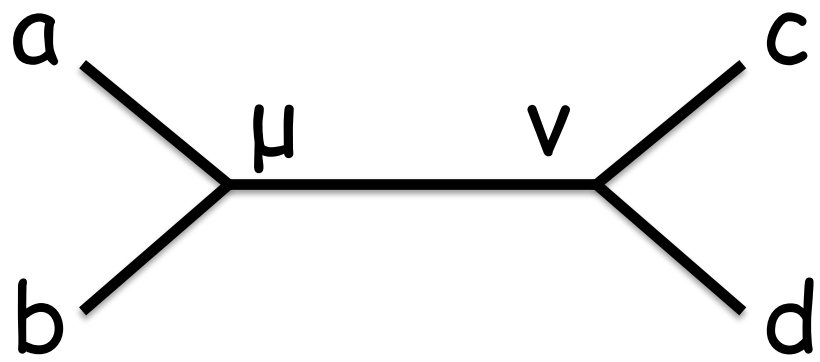
Pathogen Ecology & Evolution

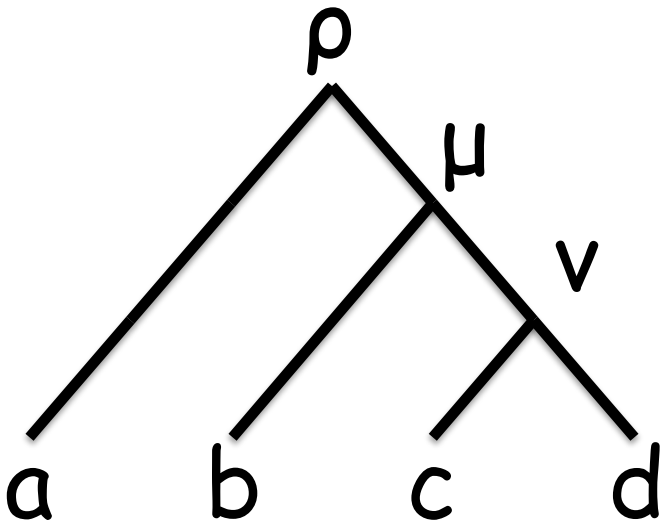


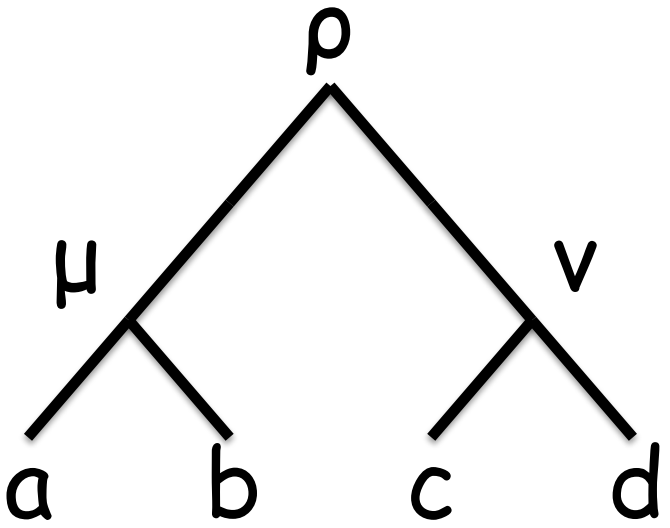
Phylogenetic reconstruction

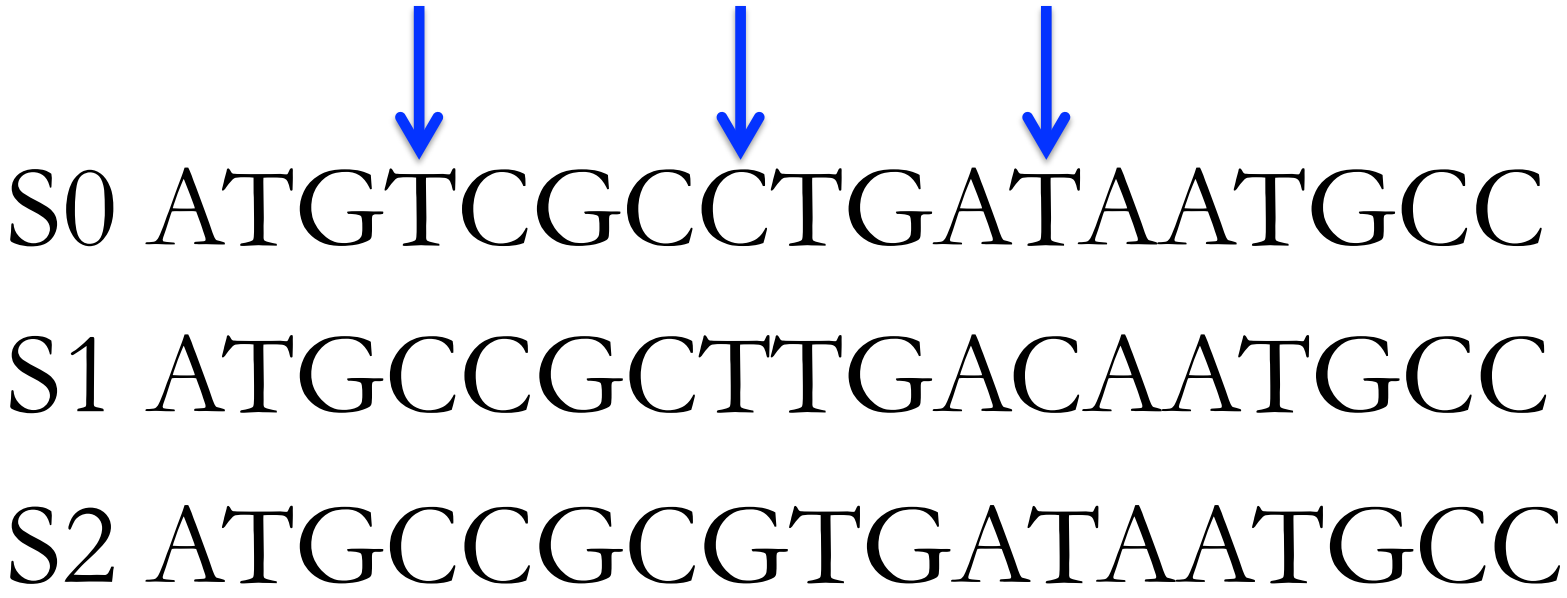
Produce a phylogenetic tree –

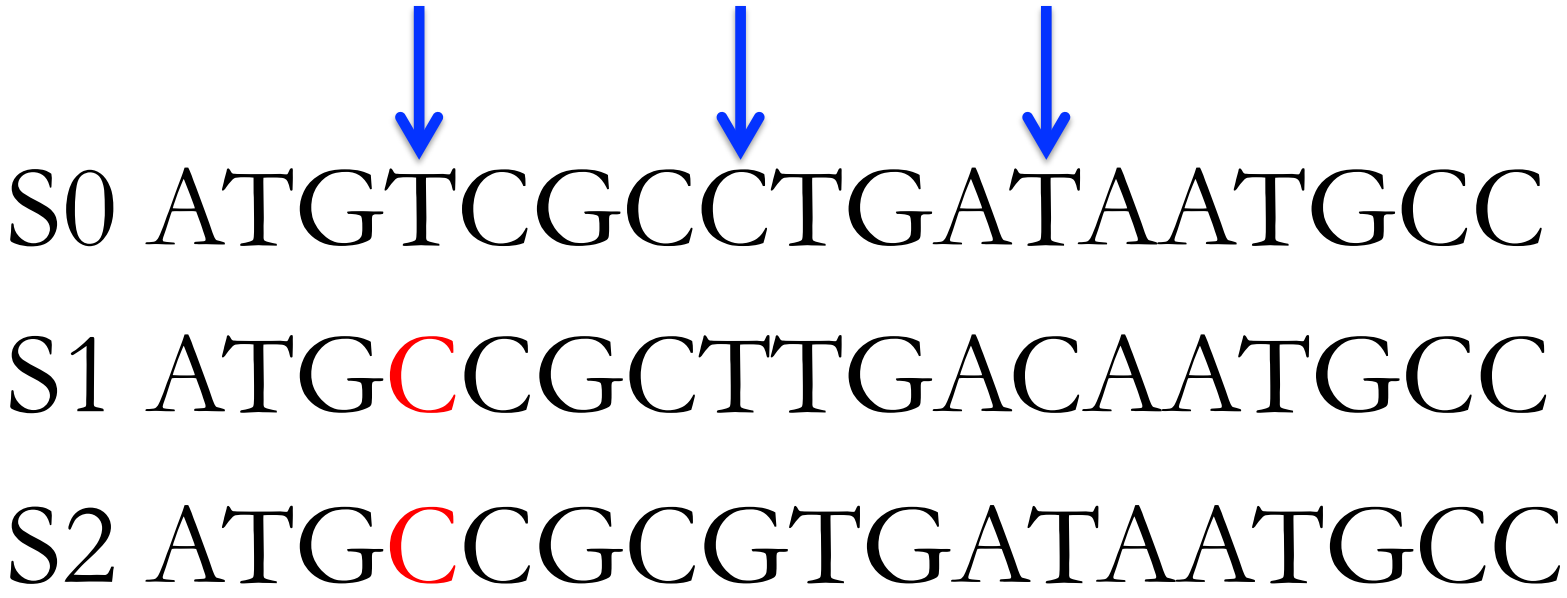
Describing likely descent from a common ancestral sequence of a set of aligned contemporary sequence.

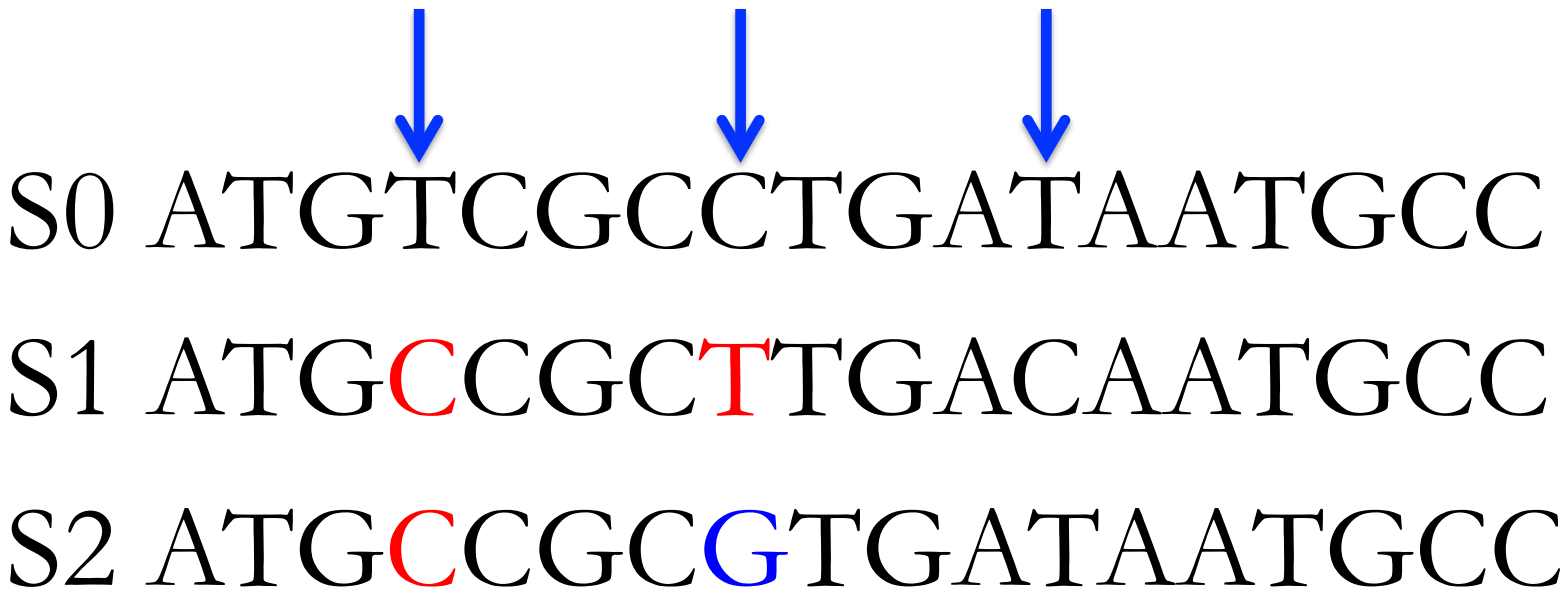

a
c
μ
ν
b
d

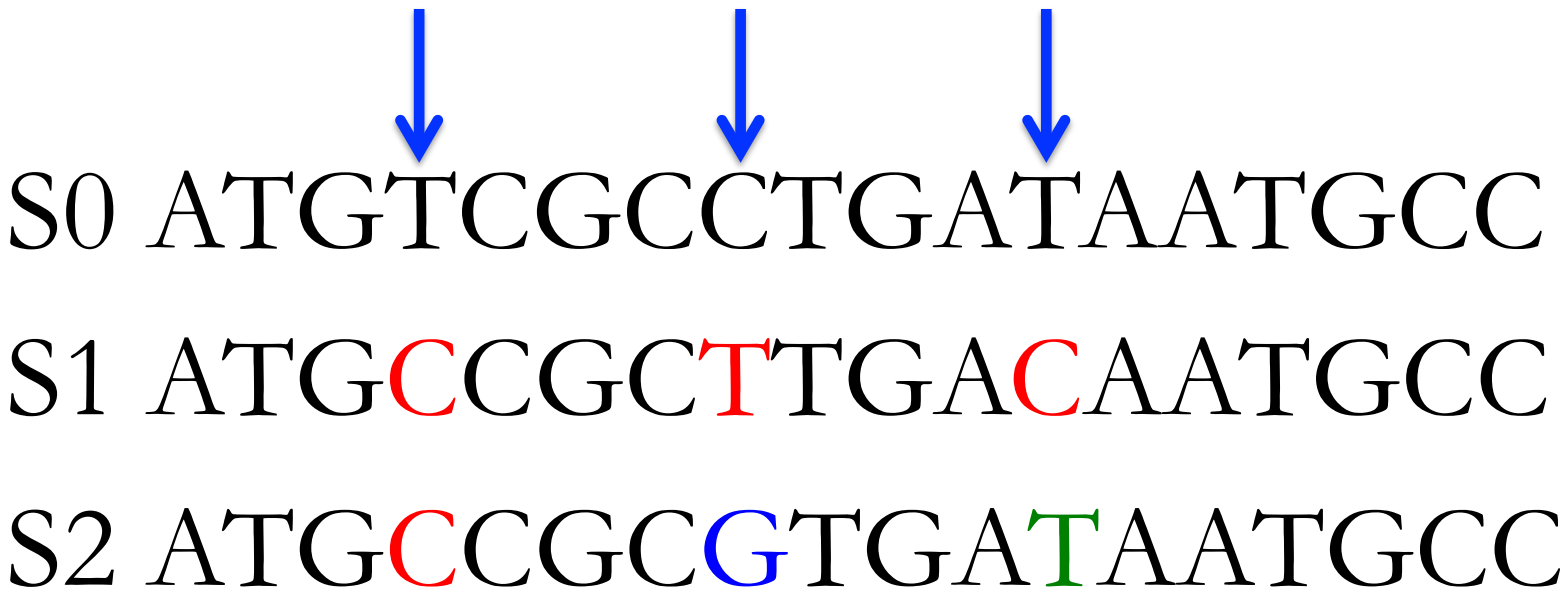

ρ
μ
ν
a
b
c
d

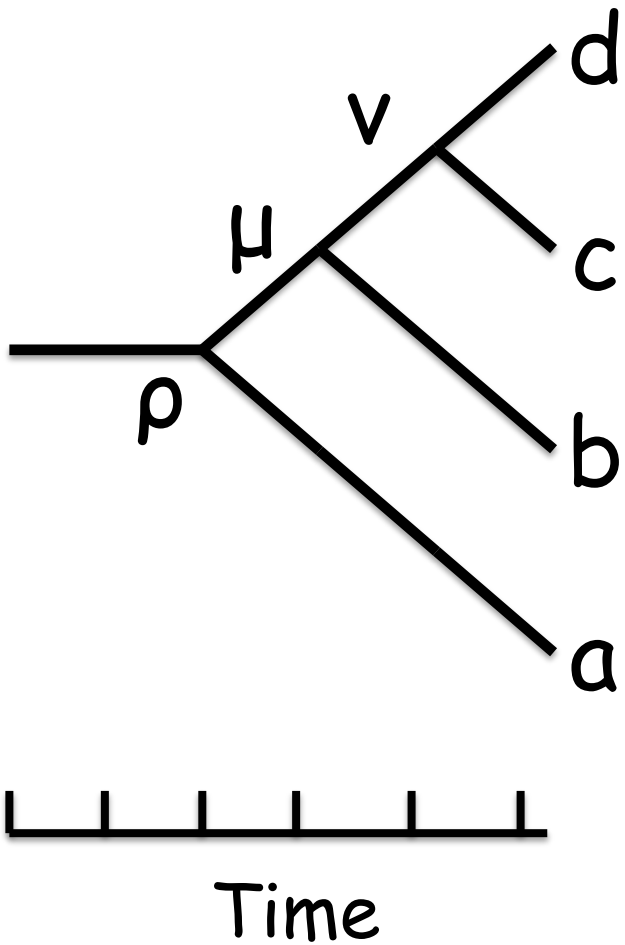

ρ
μ
ν
a
b
c
d

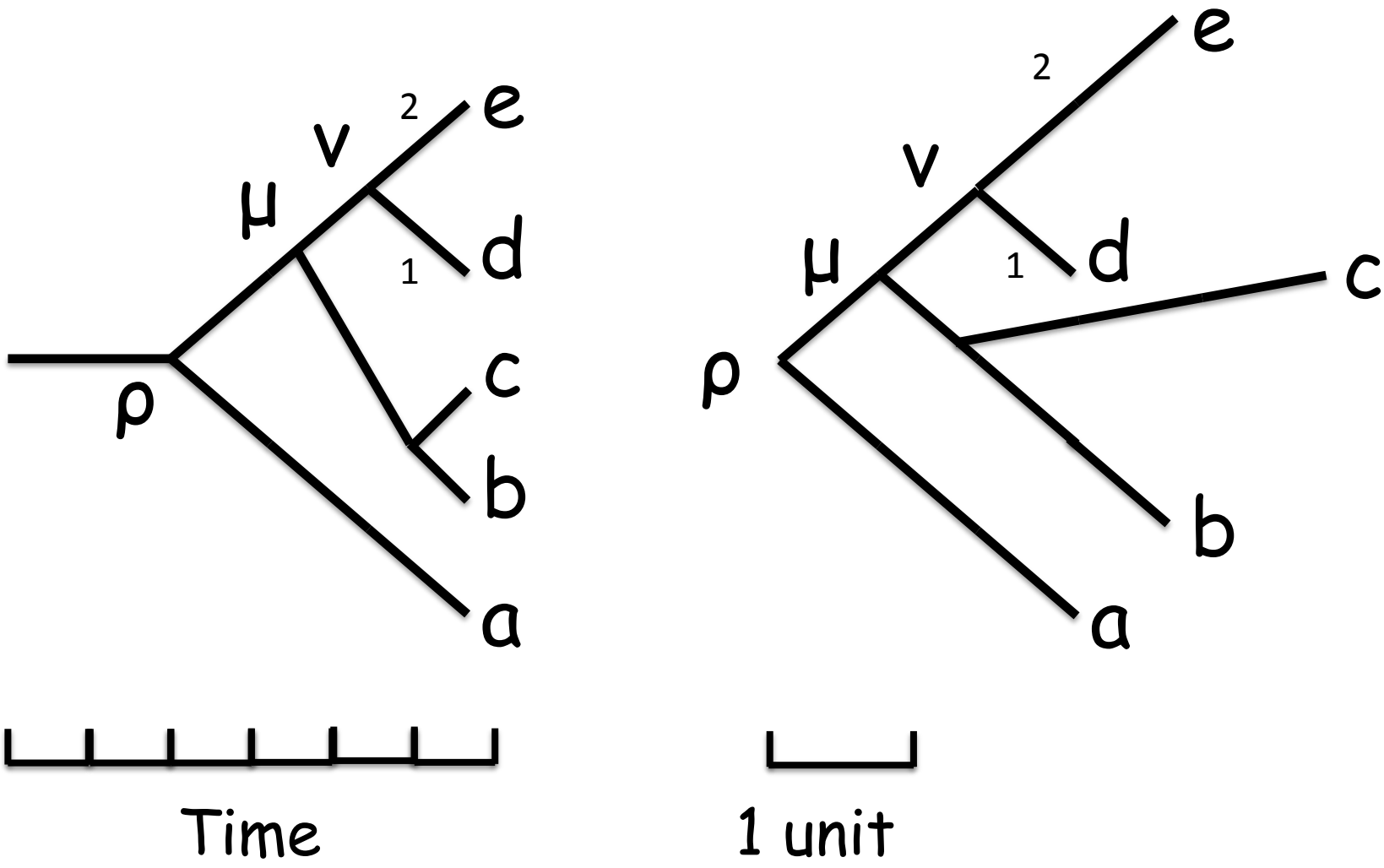

S0 ATGTCGCCTGATAATGCC
S1 ATGCCGCTTGACAATGCC
S2 ATGCCGCGTGATAATGCC


S0 ATGTCGCCTGATAATGCC
S1 ATGCCGCTTGACAATGCC
S2 ATGCCGCGTGATAATGCC


S0 ATGTCGCCTGATAATGCC
S1 ATGCCGGCTTGACAATGCC
S2 ATGCCGGCGTGATAATGCC


S0 ATGTCGCCTGATAATGCC
S1 ATGCCGCTTGACAATGCC
S2 ATGCCGCGTGATAATGCC


d
ν
μ
c
ρ
b
a
Time


e
2
ν
μ
d
1
c
ρ
b
a
Time
e
2
ν
μ
d
1
c
ρ
b
a
1 unit

How many rooted and unrooted possibilities are there?

Number of OTUs	# rooted trees	# unrooted trees
2	1	1
3	3	1
4	15	3
5	105	15
6	954	105
7	10,395	954
8	135,135	10,395
9	2,027,025	135,135
10	34,459,425	2,027,025

Parsimony

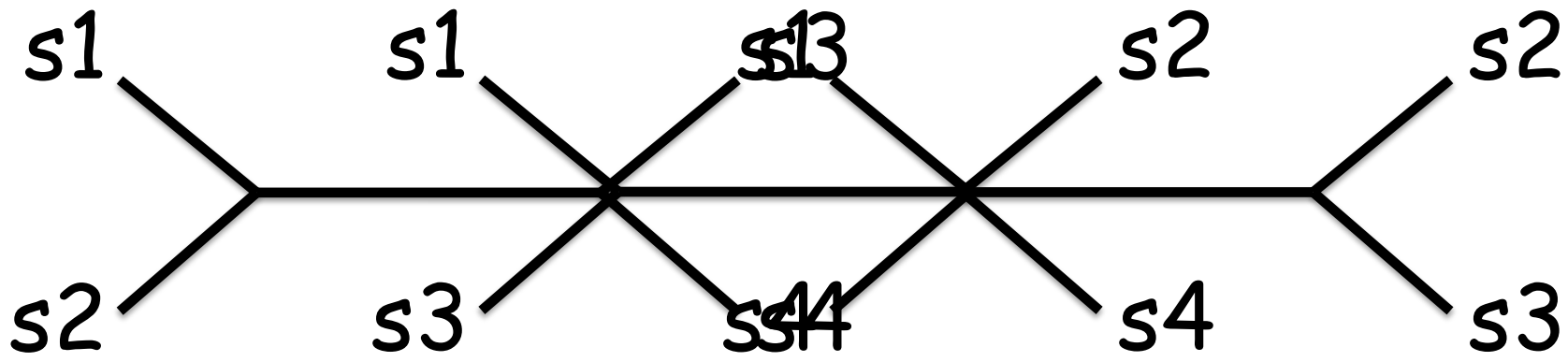
```
          5         10        15
S1 AACTTGCGCATTATC
S2 ATCTTGCGCATCATC
S3 ATCTTGGGCATCATC
S4 AACTTGGGCATTATC
```

Parsimony

```
          5         10        15
S1 AACTTGCGCATTATC
S2 ATCTTGCGCATCATC
S3 ATCTTGGGCATCATC
S4 AACTTGGGCATTATC
```

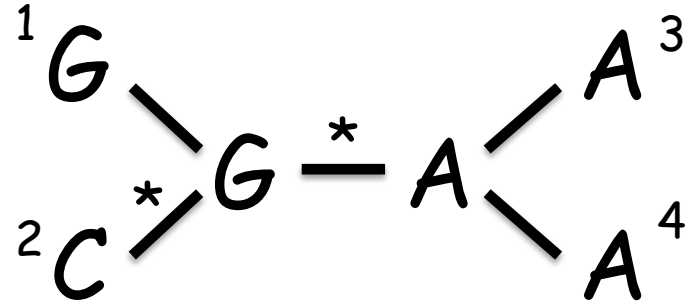
Parsimony

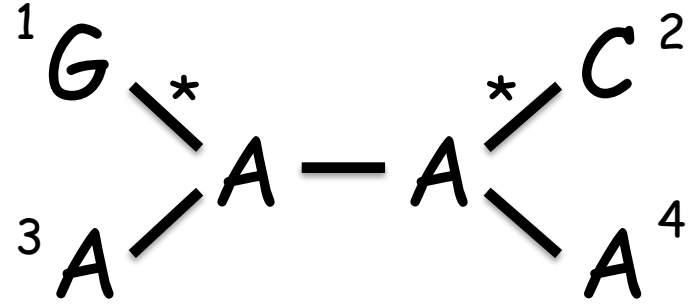
S1	A	A	G	A	G	T	G	C	A
S2	A	G	C	C	G	T	G	C	G
S3	A	G	A	T	A	T	C	C	A
S4	A	G	A	G	A	T	C	C	G

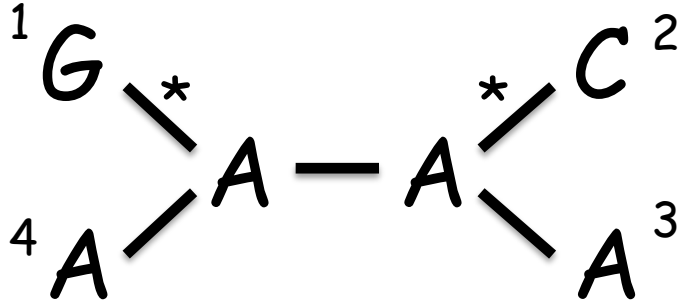


Site 3

G
C
A
A

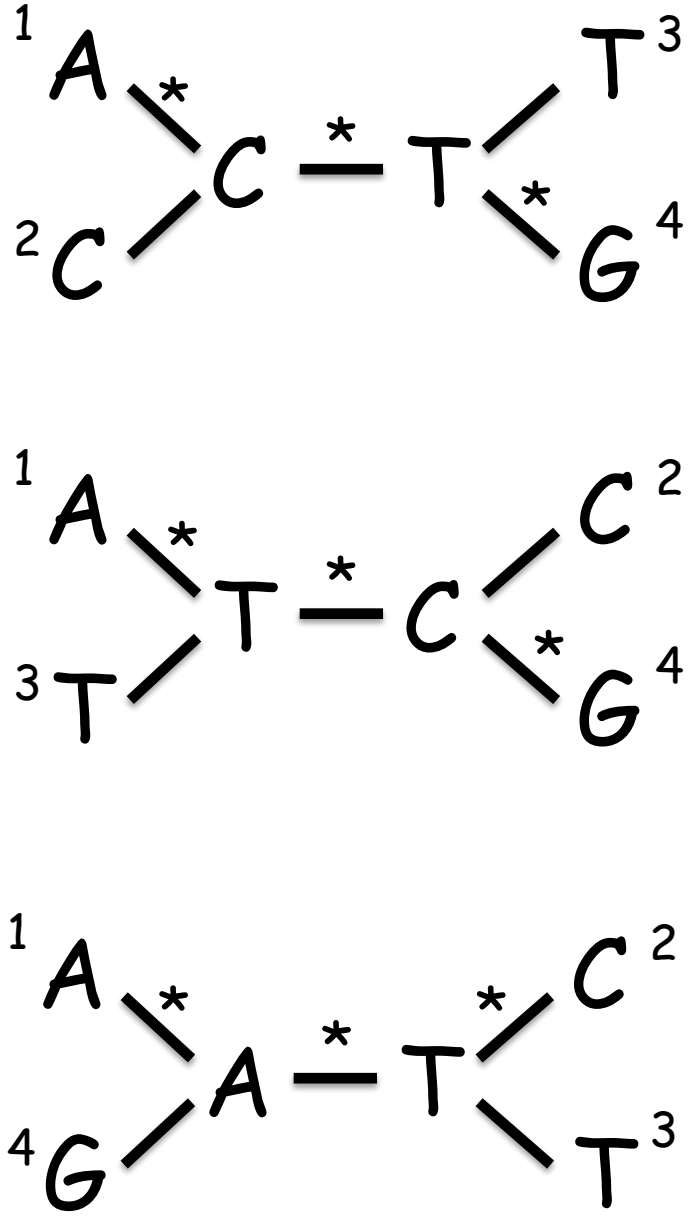






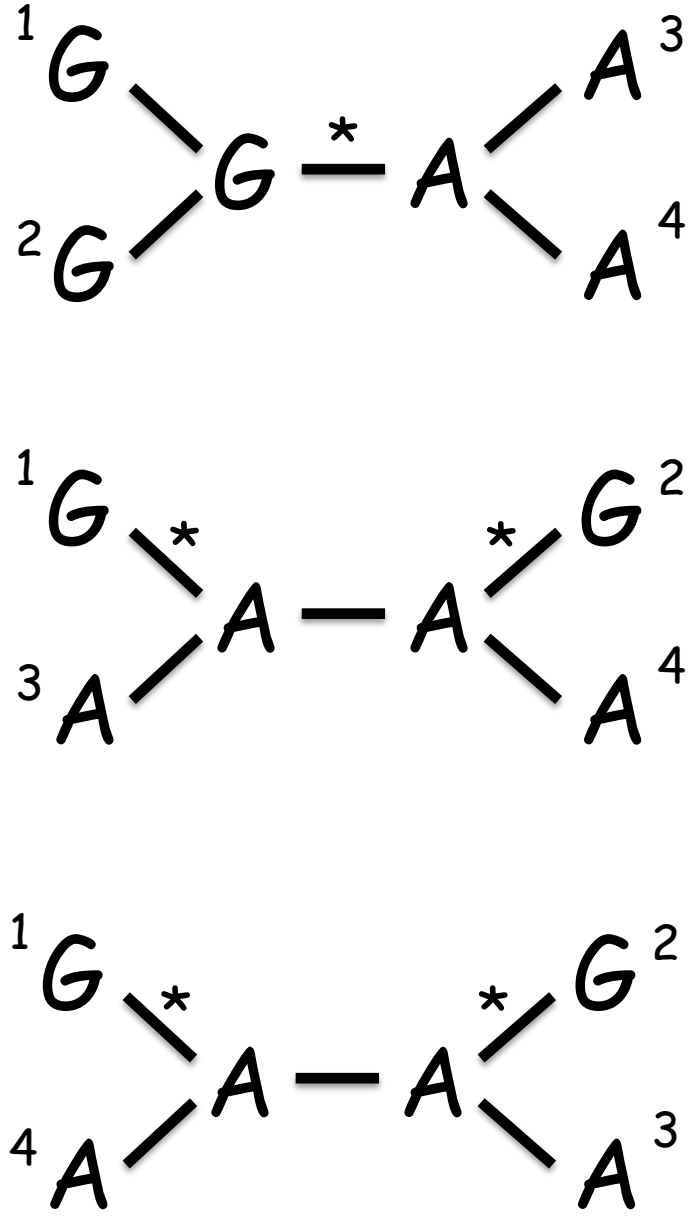
Site 4

A
C
T
G



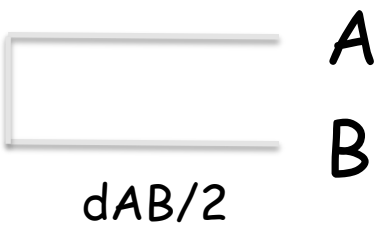
Site 5

G
G
A
A



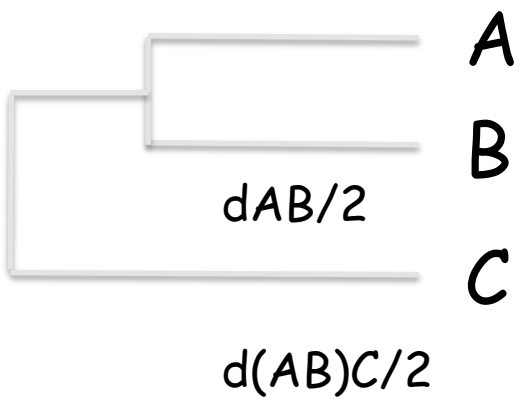
Distance - UPGMA

OTU	A	B	C
B	dAB		
C	dAC	dBC	
D	dAD	dBD	dCD

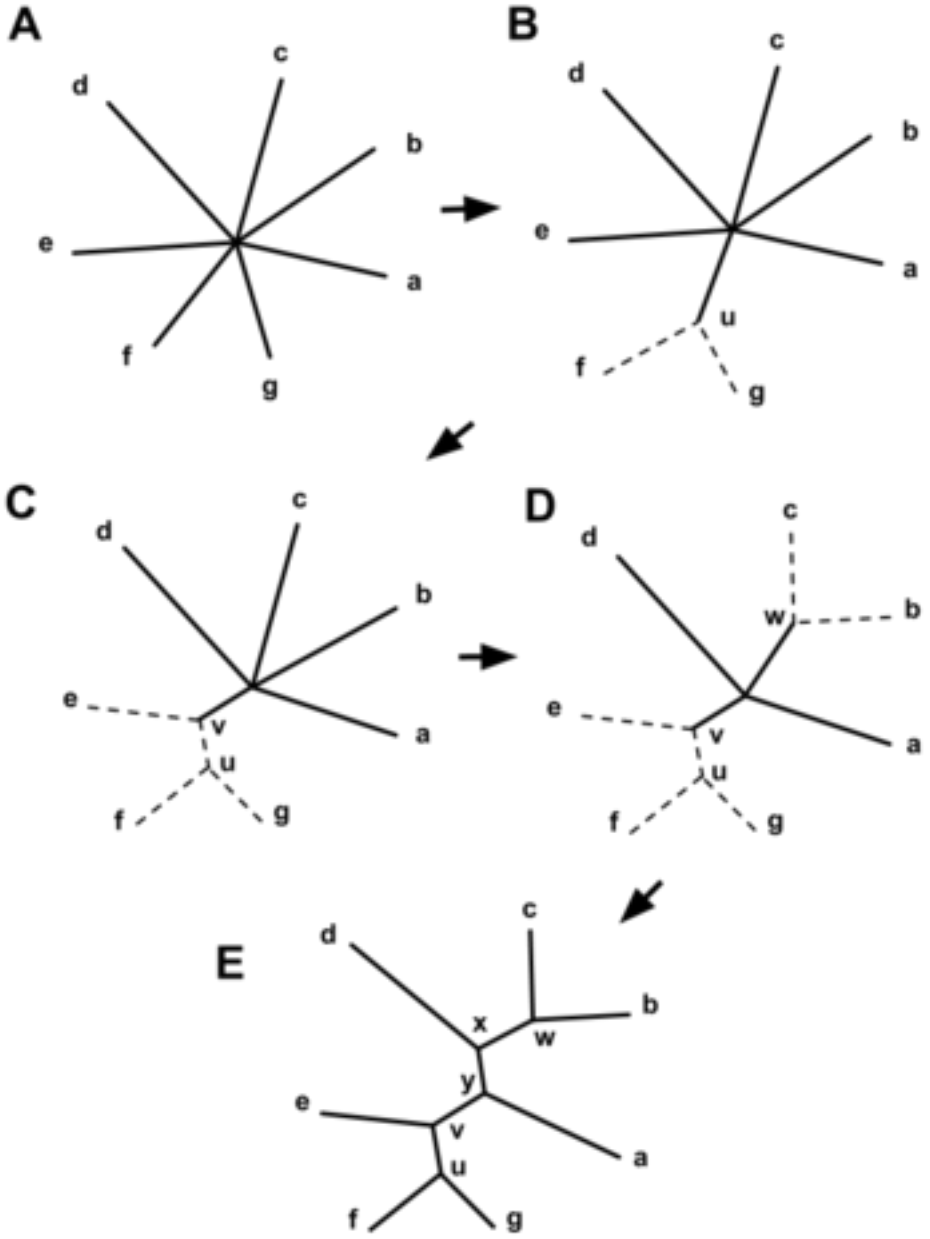


Distance - UPGMA

OTU	(AB)	C
C	d(AB)C	
D	d(AB)D	dCD



NJ

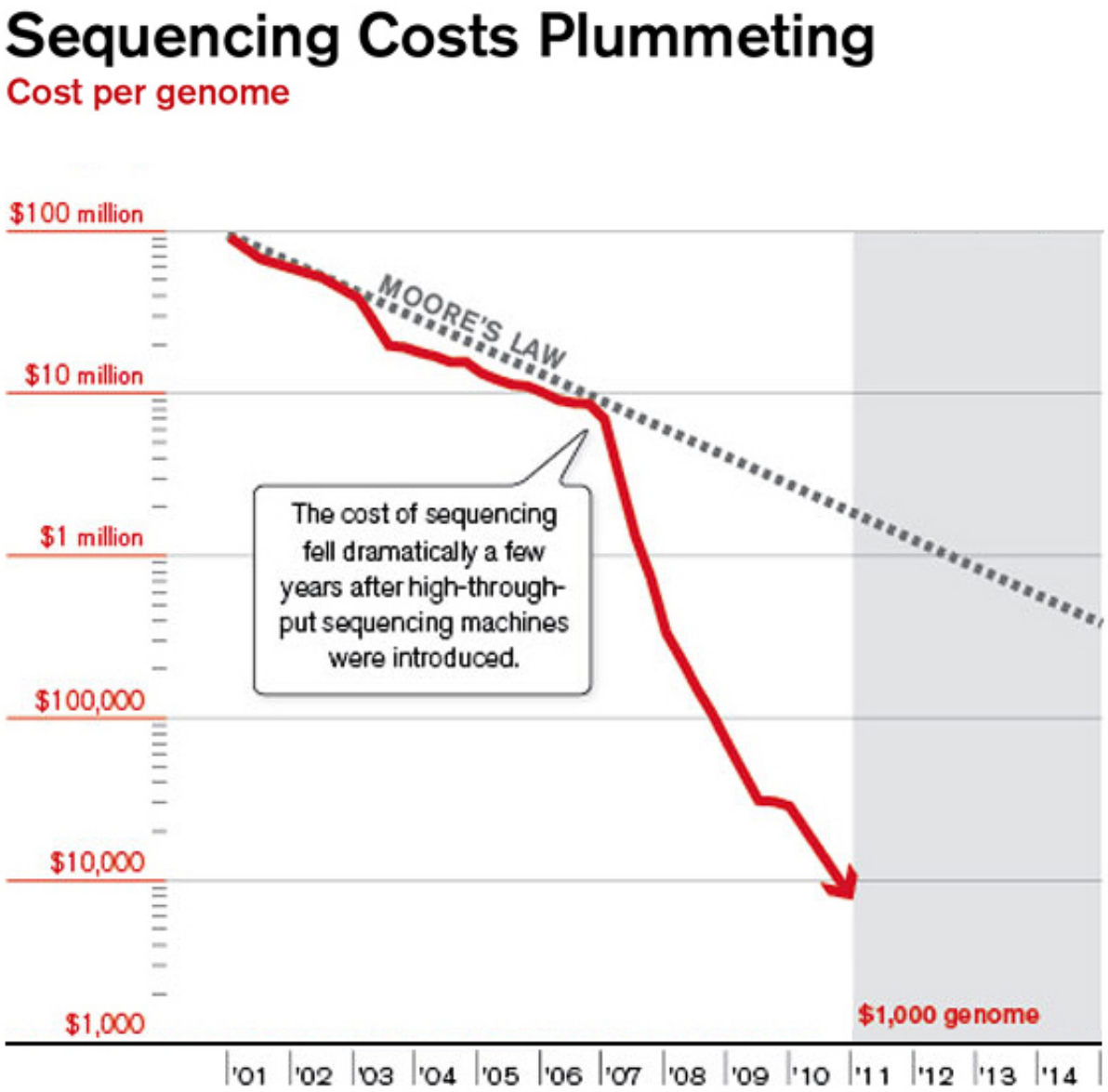


Comparing sequencing platforms in microbiome analysis

Platform	Method	Reads	16S	Metagenome	Notes
Sanger	Dideoxy terminator	750 bp	2-3 reads to cover	Good for database comparisons	Accurate, costly, slow
Pyroseq.	Light emission	400 bp			Good for 16S but not meta
Illumina	Flourescent step-by-step	100-150		More coverage makes up for short reads	High coverage, low cost
3rd generation	Electronic signal	10-100 kb		Great for assembly	Unknown error, usability

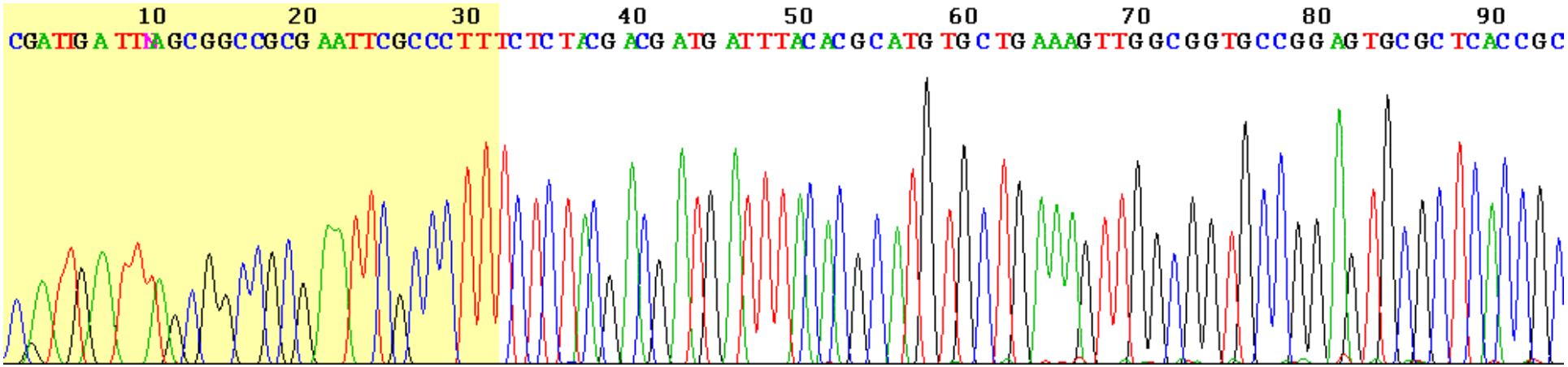
Bases to Bytes (Technology Review April 2012)

Cheap sequencing technology is flooding the world with genomic data.
Can we handle the deluge?

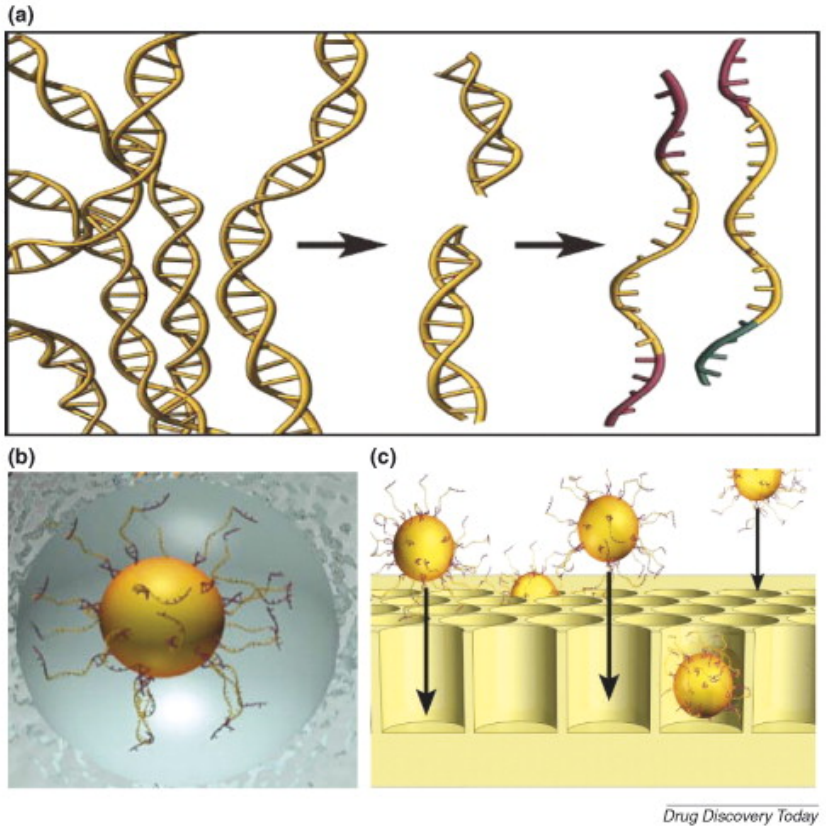


Sanger sequencing

Sanger Sequencing

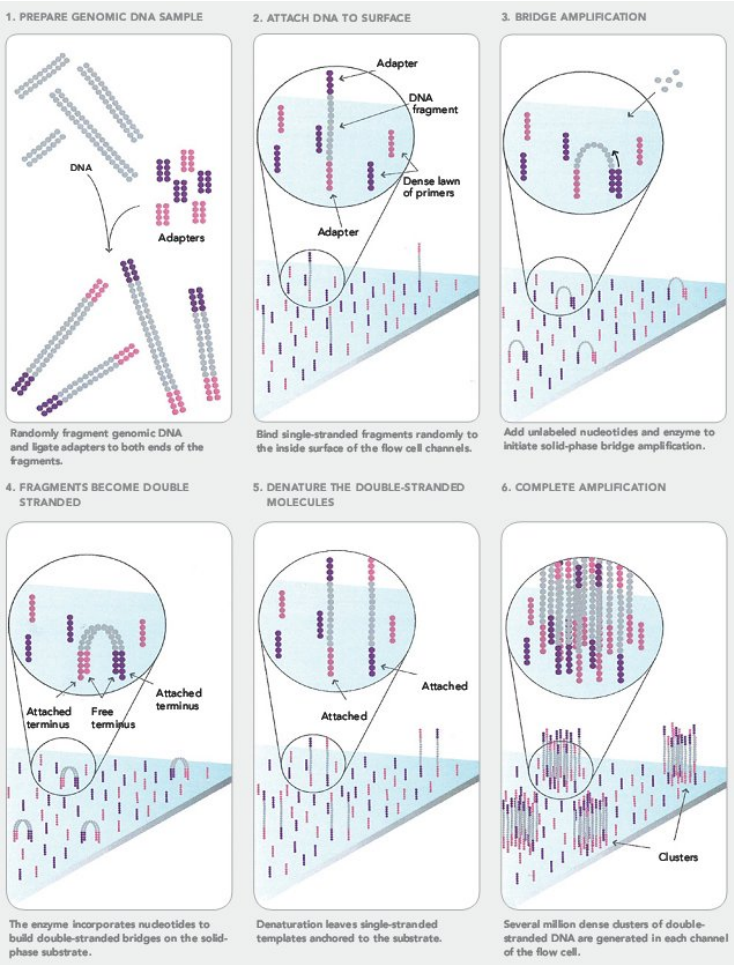


Pyrosequencing



Pyrosequencing

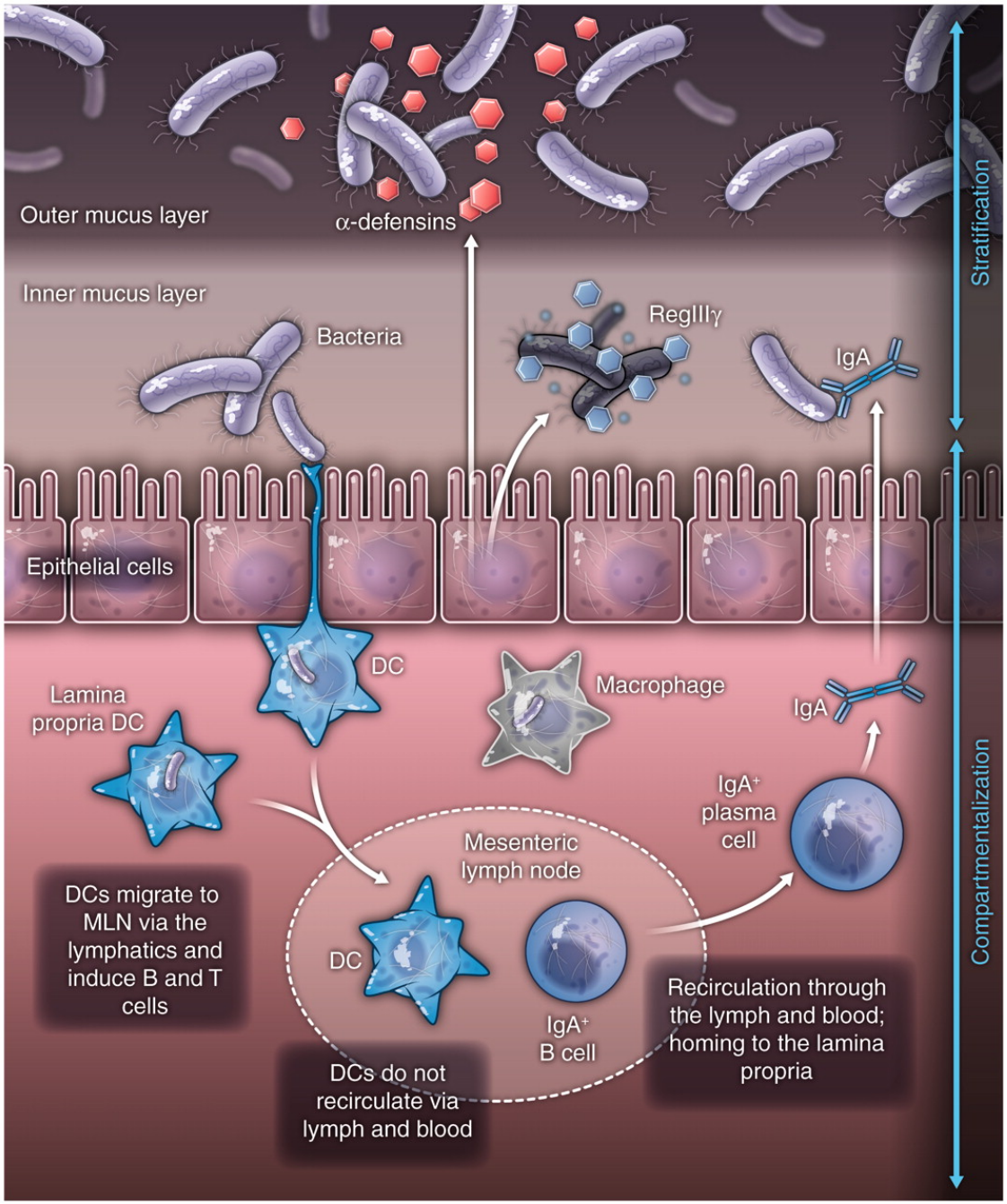
Illumina sequencing



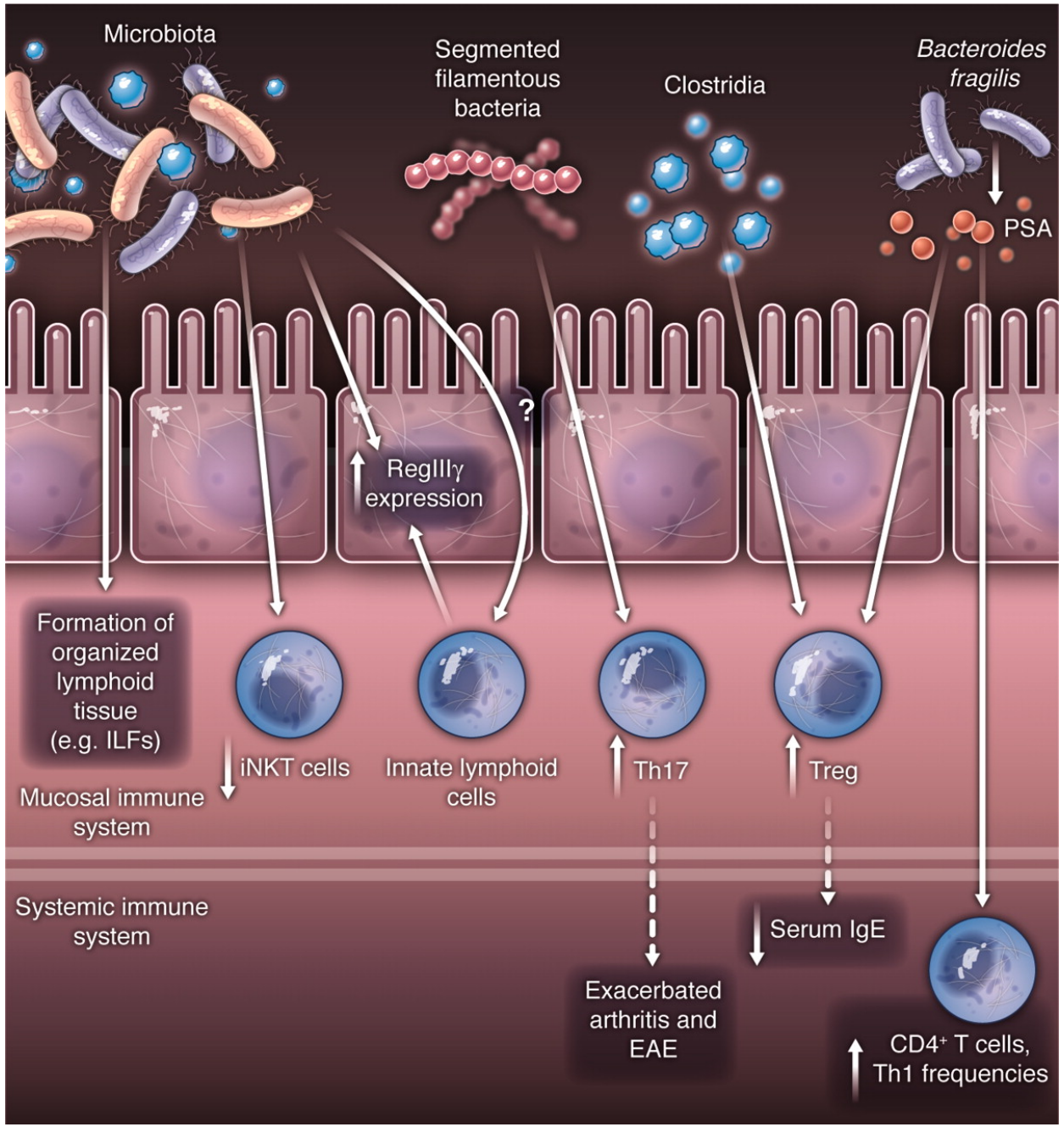
[Illumina Sequencing]

Nanopore Technology

Oxford Nanopore Technology



L V Hooper et al. Science 2012;336:1268-1273



L V Hooper et al. Science 2012;336:1268-1273