

5. Phylogenetic trees

- The tree of life
- The tree, an abstract object
- Building an array of distances
- The UPGMA algorithm
- **Differences are not always what they look like**
- The diversity of bioinformatics algorithms
- The application domains in microbiology

	Sequence1	Sequence2	Number of substitutions being observed	Number of actual substitutions
Unique substitution	C	C → A	1	1
Multiple substitutions	A	A → C → T	1	2
Coincident substitutions at the same position	C → A	G → C	1	2
Parallel substitutions	T → A	T → A	0	2
Convergent substitutions	C → T → A	C → A	0	3
Reverse substitutions	C → T → C	C	0	2

AACATCGTAATCGGTATATACGGTCA
AACATCGTAATCGGTATATACGGTCA

AACATCGTAATCGGTATATACGGTCA
AACATCGTAATAGGTATATACGGTCA

	Sequence1	Sequence2	Number of substitutions being observed	Number of actual substitutions
Unique substitution	C	C → A	1	1
Multiple substitutions	A	A → C → T	1	2
Coincident substitutions at the same position	C → A	G → C	1	2
Parallel substitutions	T → A	T → A	0	2
Convergent substitutions	C → T → A	C → A	0	3
Reverse substitutions	C → T → C	C	0	2

AACATCGTAATCGGTATATACGGTCA
AACATCGTAATCGGTATATACGGTCA

AACATCGTAATCGGTATATACGGTCA
AACATCGTACTCGGTATATACGGTCA

AACATCGTAATCGGTATATACGGTCA
AACATCGTA^TTCGGTATATACGGTCA

	Sequence1	Sequence2	Number of substitutions being observed	Number of actual substitutions
Unique substitution	C	C → A	1	1
Multiple substitutions	A	A → C → T	1	2
Coincident substitutions at the same position	C → A	G → C	1	2
Parallel substitutions	T → A	T → A	0	2
Convergent substitutions	C → T → A	C → A	0	3
Reverse substitutions	C → T → C	C	0	2

AACATCGTAATCGGTATATACGGTCA
AACGGGGTTCGCGGTACGTACGGTCA

AACATAGTAATCGGTATATACGGTCA
AACGGCGTTCGCGGTACGTACGGTCA

	Sequence1	Sequence2	Number of substitutions being observed	Number of actual substitutions
Unique substitution	C	C → A	1	1
Multiple substitutions	A	A → C → T	1	2
Coincident substitutions at the same position	C → A	G → C	1	2
Parallel substitutions	T → A	T → A	0	2
Convergent substitutions	C → T → A	C → A	0	3
Reverse substitutions	C → T → C	C	0	2

AACATCGTAATCGGTATATACGGTCA
AACATCGTAATCGGTATATACGGTCA

AACAACGTAATCGGTATATACGGTCA
AACAACGTAATCGGTATATACGGTCA

	Sequence1	Sequence2	Number of substitutions being observed	Number of actual substitutions
Unique substitution	C	C → A	1	1
Multiple substitutions	A	A → C → T	1	2
Coincident substitutions at the same position	C → A	G → C	1	2
Parallel substitutions	T → A	T → A	0	2
Convergent substitutions	C → T → A	C → A	0	3
Reverse substitutions	C → T → C	C	0	2

AACATAGTAATCGGTATATACGGTCA
AACGGCGTTCGCGGTACGTACGGTCA

AA**T**ATAGTAATCGGTATATACGGTCA
AACGGCGTTCGCGGTACGTACGGTCA

AA**A**ATAGTAATCGGTATATACGGTCA
AA**A**GGCGTTCGCGGTACGTACGGTCA

	Sequence1	Sequence2	Number of substitutions being observed	Number of actual substitutions
Unique substitution	C	C → A	1	1
Multiple substitutions	A	A → C → T	1	2
Coincident substitutions at the same position	C → A	G → C	1	2
Parallel substitutions	T → A	T → A	0	2
Convergent substitutions	C → T → A	C → A	0	3
Reverse substitutions	C → T → C	C	0	2

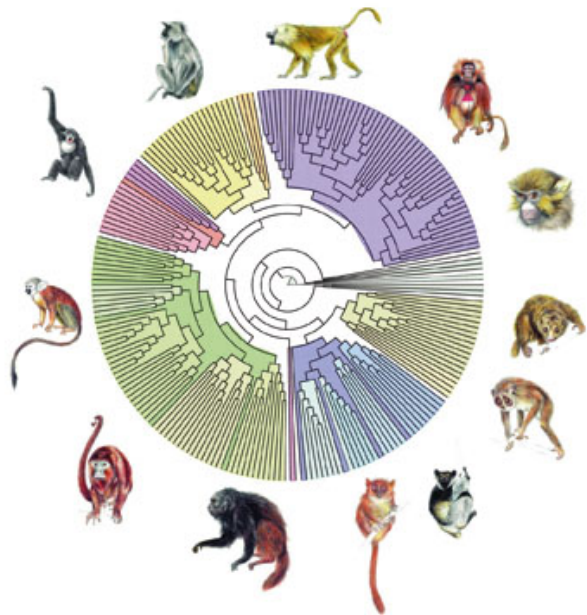
AACATAGTAATCGGTATATACGGTCA
AACGGCGTTCGCGGTACGTACGGTCA

AACATAGTAATCGGTATATA**T**GGTCA
AACGGCGTTCGCGGTACGTACGGTCA

AACATAGTAATCGGTATATA^CGGTCA
AACGGCGTTCGCGGTACGTACGGTCA

More pertinent methods are required

- Other distance methods
 - NJ (Neighbor-Joining)
 - Corrections
- Parsimony
- Maximum likelihood



Pictures & movies : material licensing

p.24 : Rights reserved