An Improved **Method For** Constructing

Acqua Phylogenies

Nomi Hadar & Itay Mayrose

Department of Molecular Biology and Ecology of Plants, Tel Aviv University, Tel Aviv Israel

- Reconstructing the "Tree of Life" is one of the enduring goals of evolutionary
- The massive accumulation of sequence data should provide more accurate phylogenies with the hope to resolve the tree of life.
- Recently, the tree of birds, tree of plants and tree of reptiles were presented.
- But, the reconstructed trees can only be as good as their underlying multiple sequence alignment (MSA).

The Problem

Aligning rapidly evolving loci for highly diverged taxa leads to poor alignments and inaccurate trees.

Slowly evolving regions: Easy to align good for solving ancient divergence events, cannot be used to solve recent events.

Rapidly evolving regions: Hard to align good for solving recent event, cannot be used to solve ancient divergence events.

Background: Mega Phylogenies

biology.

We Compare Mega Phylogenies to **Current Taxonomical Knowledge**

If the tree is correct, we expect the species that are part of the same hierarchy family (or other taxonomic hierarchy) to form a monophyletic group in the tree.

Three measures were used to asses monophyly.

E.g., Family = { , , , , } History?

Description Of The

True Evolutionary

Strict = 0

Relaxed 1 = 4/6

of species in family

of species under the last common ancestor

Divide The Alignment Space To Well-Aligned Regions

Locus 1

For Reconstructing Phylogenies

A New Approach

Locus 2 -> Aligned TATCCCGGG - - -AAGCTAATT -AATGGAGGAATTTCAAGTATATT TAT ----GGGCA -------ATGGA -- AATTTCAA ---TAT TATCCCTGG - - - AAAGCTAAT - CAATGGAGGAATTTCAAGTATAT TAA - - - - GGGCA - - - - - - - - - - ATGGA - - AATTGCAA - - - TAT TAT ---- GCGCA ------ATGCA -- AATTTCAA ---TAT Locus 3 -> Seq1: CATTGATGTTTCCTTAATCC - AGATCGGAGGACGTGTTTC Seq2: -----TCGATG------ACGAAAGCGAA---GTCCTC Seg3: CATTG----- CCCCATTAA----- CGCAC-- GTTTCA Seq4: -----TCGATG------AGGATCCGCGAA---GTTCT-Seq5: CATTG----- CCTACATCC----- CGCGT-- CACCAC

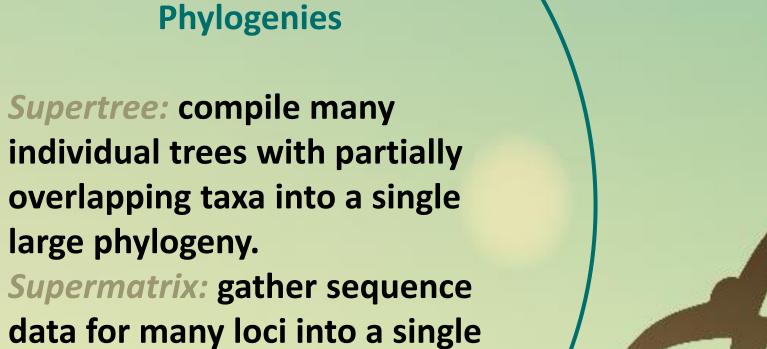
Locus 1 ->

Locus 2a Locus 2b TATCCCGGGAAGCTAATTAATGGAGGAATTTCAAGTATATT TATCCCTGGAAAGCTAATCAATGGAGGAATTTCAAGTATAT---TATGGGCAATGGAAATTTCAATAT --TAAGGGCAATGGAAATTGCAATAT --TATGCGCAATGCAAATTTCAATATT Locus 3a Locus 3b Locus 3c CATTGATGTTTCCTTAATCCAGATCGGAGGACGTGTTTC---TCGATGACGAAAGCGAA-GTCCTC---TCGATGAGGATCCGCGAA-GTTCT--

Using the resulted alignment to reconstruct the tree







Current Approaches For

Reconstructing Mega

multiple sequence alignment,

reconstruct a large phylogeny.

which is then used to



Is family monophyltic?

Relaxed 2 = 3/4

of species in largest monophyletic subgroup

of species in family