Google as a taxonomic engine

Alexey Shipunov

Background: tree

cow

mouse

rat
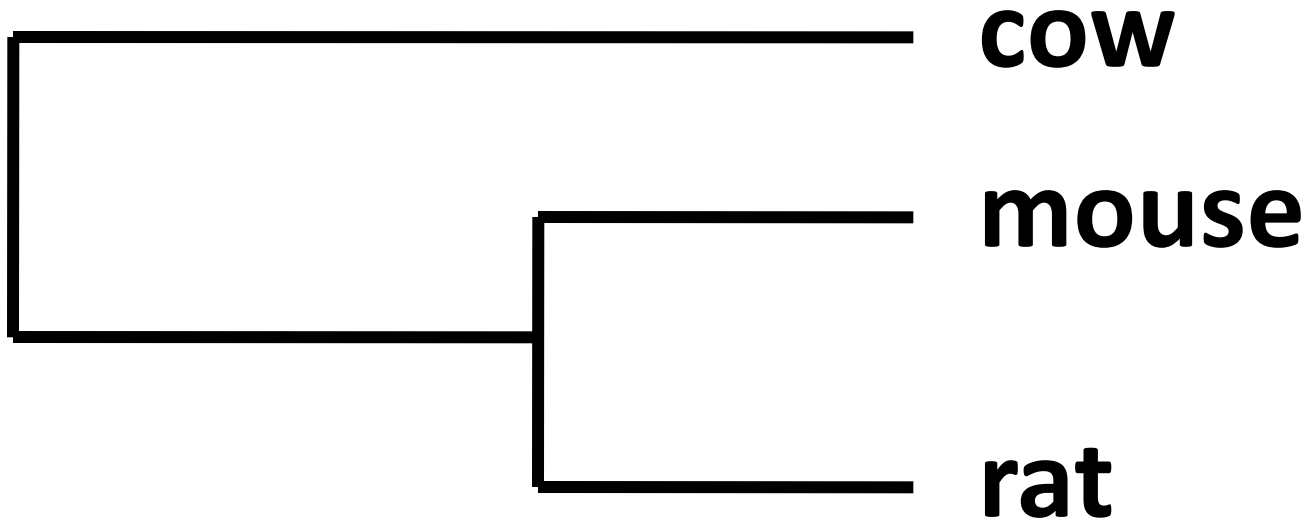
Background: Google Scholar

- Extract phyla names
- Obtain numbers of joint hits
- Do some magic
- Calculate similarity
- Make clusters
- Same, with "-ecology"
- Same, with classes names

Numbers of joint hits

- Phyla names from my "synat" classification
- **R** script to make command-line queries
- **Links** text browser to make textual output
- UNIX text tools (sed, grep) to clean results
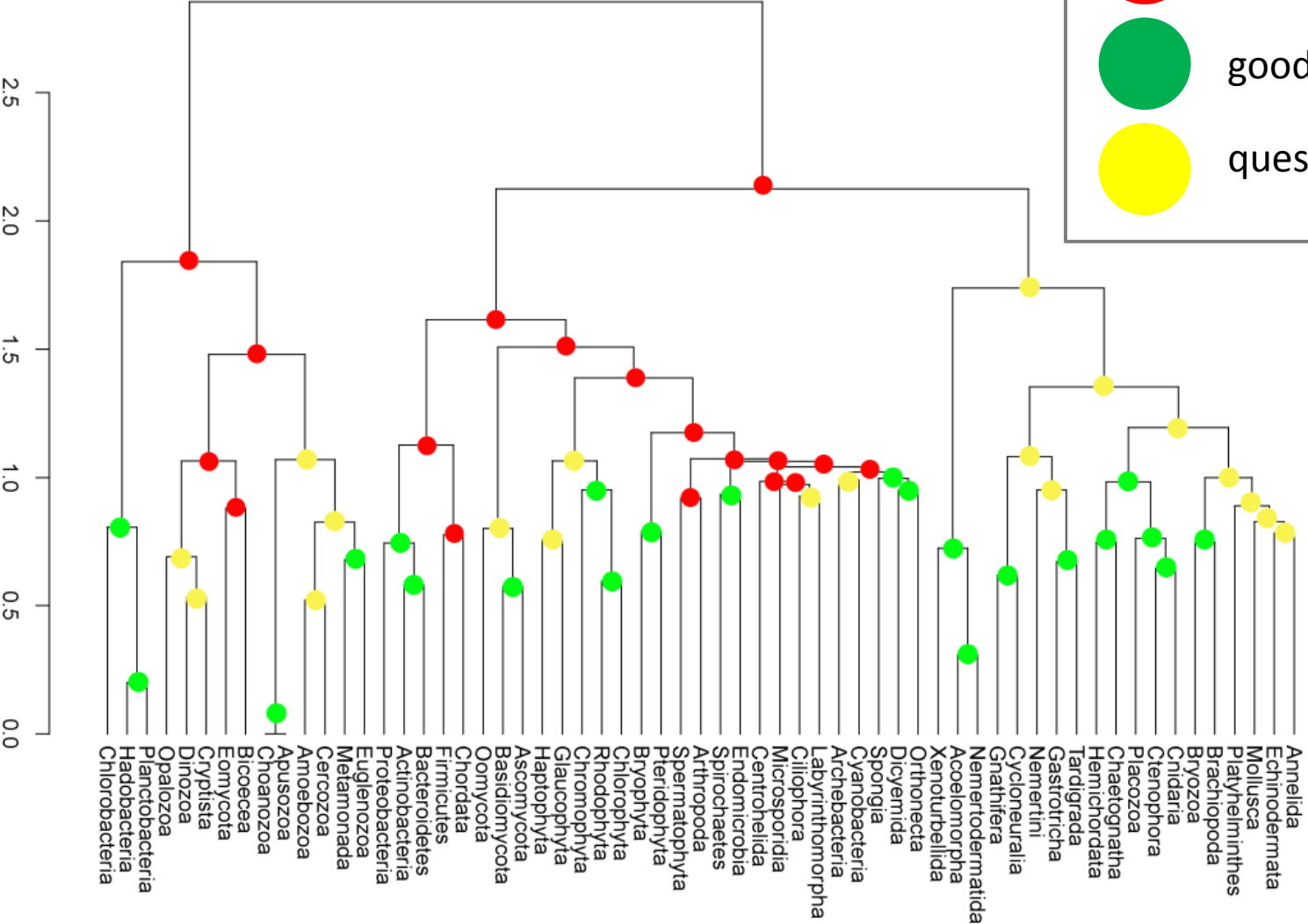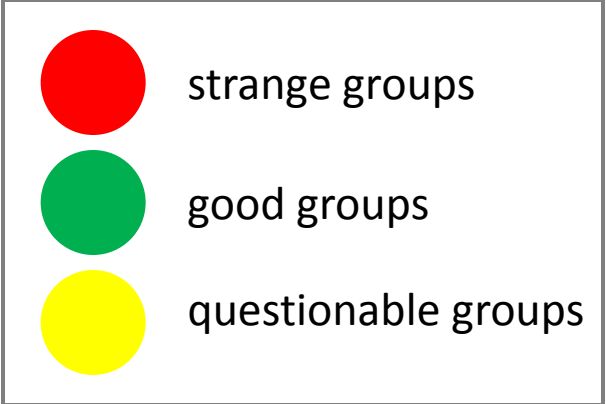- Comma-delimited file for import into **R**

Magic with numbers

- Some names appear much often than others
  - found numbers of individual hits
  - make weights
  - multiple numbers of joint hits by geometric means of individual hits from each taxon in pair
- Convert table of three column into square matrix
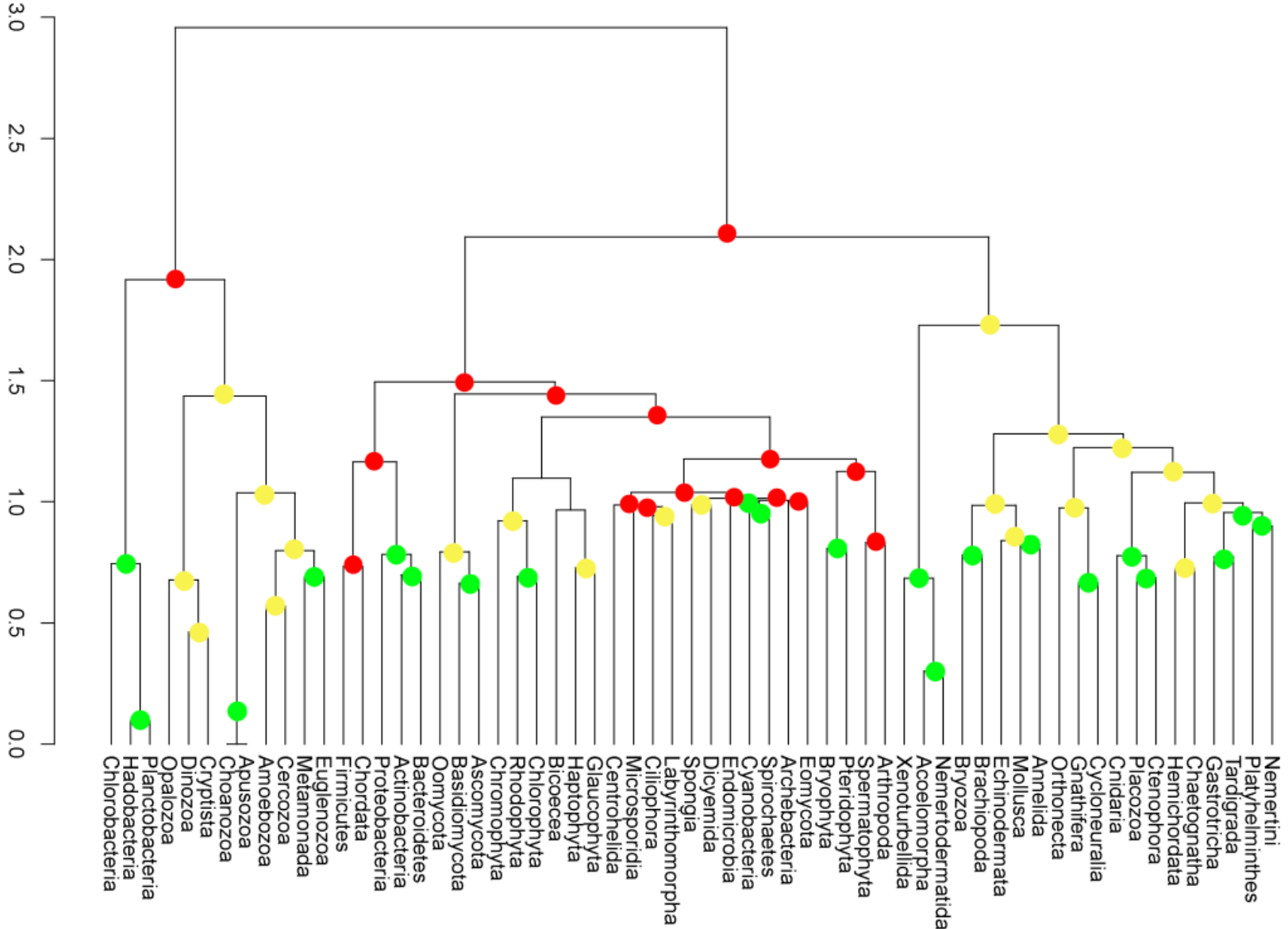- Convert similarities to dissimilarities

Similarity

- Calculate Euclidean distances
- Ward's method hierarchical clustering
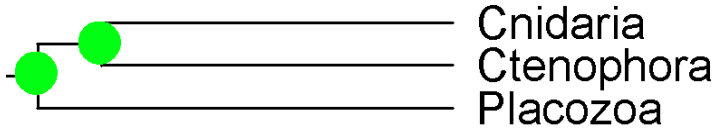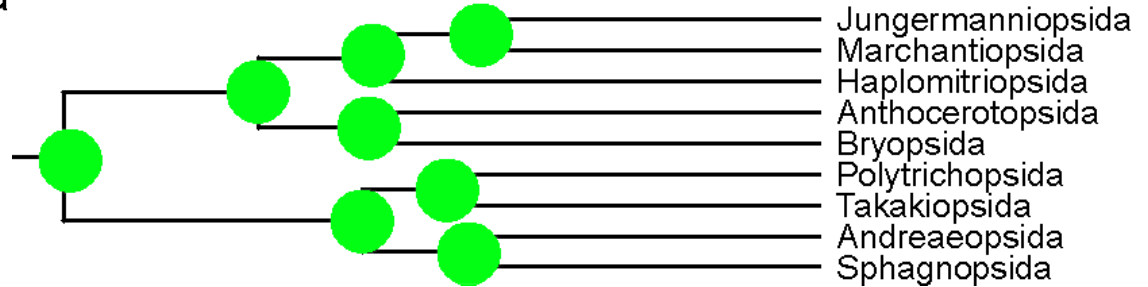- Tree of clusters is NOT a phylogenetic tree

Result

strange groups
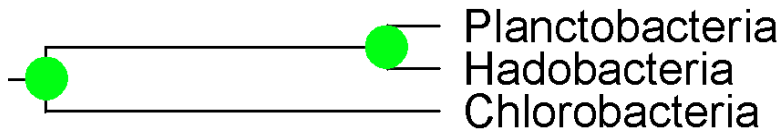
good groups

questionable groups

Same, -ecology

Classes

Some reliable groups

Cnidaria
Ctenophora
Placozoa

Primitive animals

Jungermanniopsida
Marchantiopsida
Haplomitriopsida
Anthocerotopsida
Bryopsida
Polytrichopsida
Takakiopsida
Andreaeopsida
Sphagnopsida

Mosses

Planctobacteria
Hadobacteria
Chlorobacteria

Basal bacterial groups

Lecanoromycetes
Orbiliomycetes
Sordariomycetes
Eurotiomycetes
Dothideomycetes
Leotiomycetes
Pezizomycetes
Neolectomycetes
Taphrinomycetes
Schizosaccharomycetes
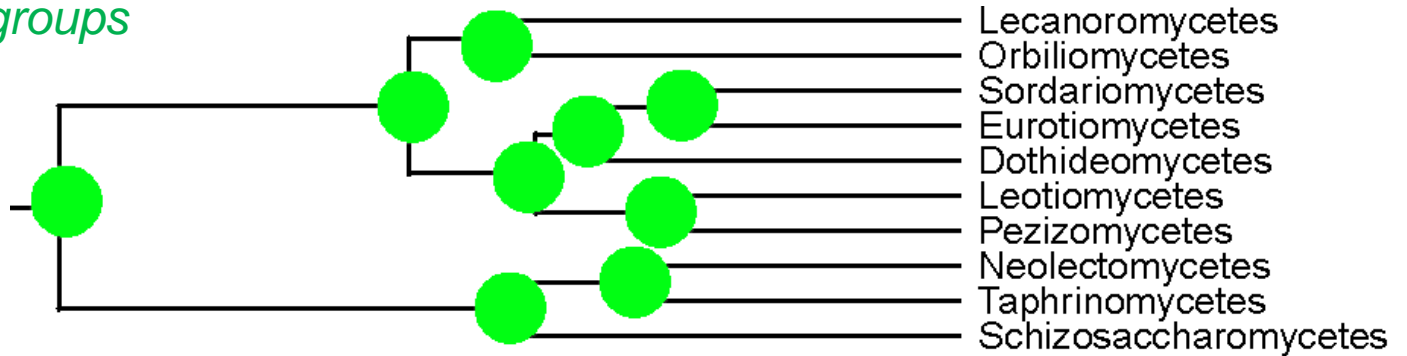
*Algae with primary
chloroplasts*

Chlorophyta
Rhodophyta

Ascomycete fungi

Questionable groups
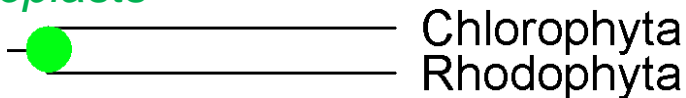
Annelida
Echinodermata
Mollusca
Platyhelminthes
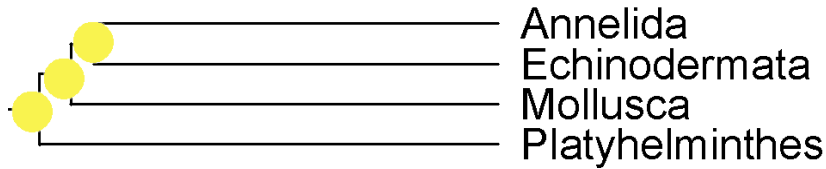
Spiralians + starfishes

Algal mix

Chlorophyceae
Bacillariophyceae
Chrysophyceae
Euglenophyceae
Phaeophyceae
Prasinophyceae
Prymnesiophyceae
Raphidophyceae

Cercozoa
Amoebozoa
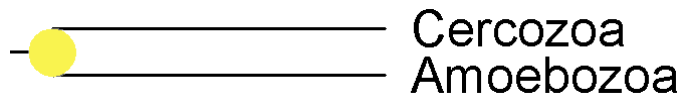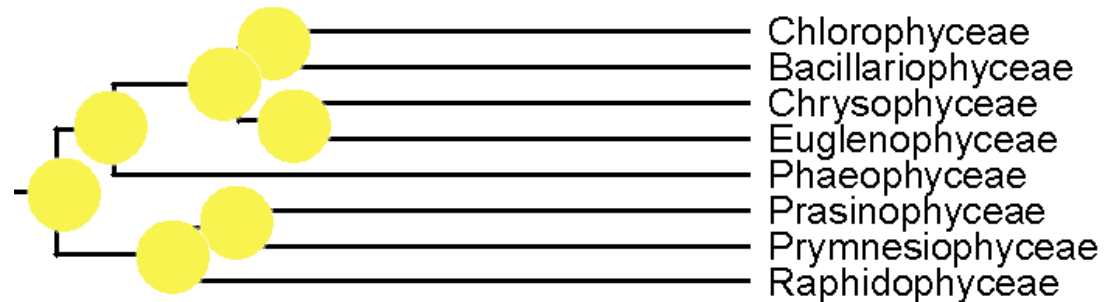
Two distantly related groups, both contain amoebae

Strange groups

Arthropoda
Spermatophyta
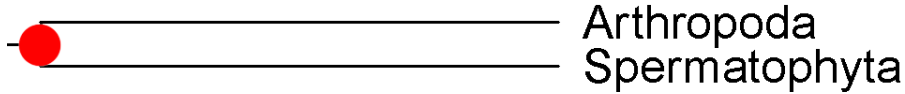
Insects + flowering plants

Chordata
Firmicutes
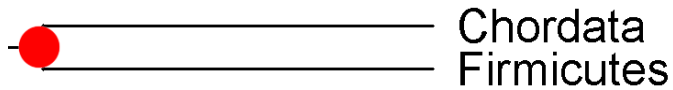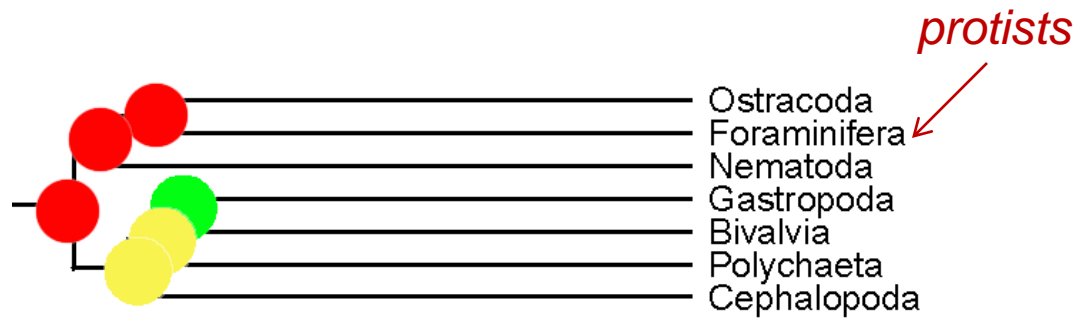
Vertebrates + Gram-positive bacteria

protists

Ostracoda
Foraminifera
Nematoda
Gastropoda
Bivalvia
Polychaeta
Cephalopoda

Creatures with shells

Myxosporea
Malacosporea
Microsporea
Ichthyostraca
Polypodiozoa
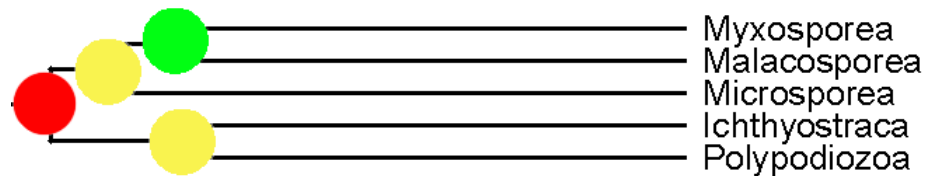
Parasitic creatures, animals and protists

Conclusions

- **It is working!**
- In most cases, only closest taxa were revealed; animals, bacteria and fungi were intermixed
- "-ecology" did not help
- Classes are generally better than phyla

"Best classification is a classification which does not exist,
it is a constantly changing product of processing all reliable data available on-line"