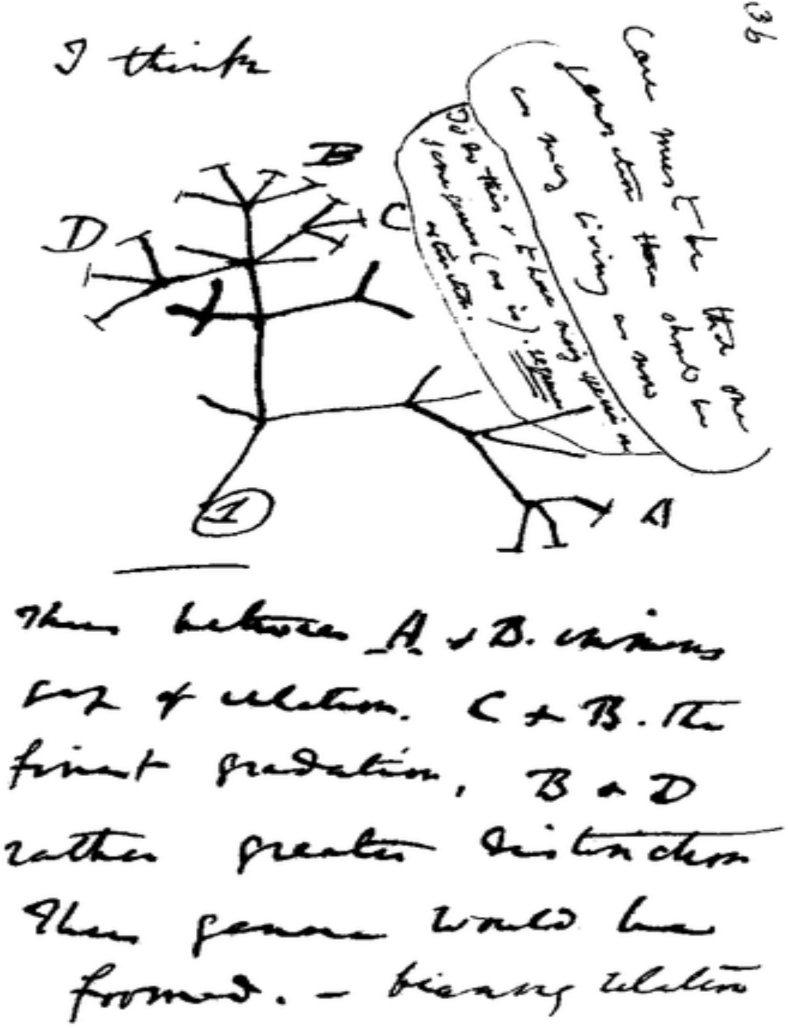
4. Sequence comparison

- How to predict gene/protein functions?
- **Why gene/protein sequences may be similar?**
- Measuring sequence similarity
- Aligning sequences is an optimization problem
- A sequence alignment as a path
- A path is optimal if all its sub-paths are optimal
- Alignment costs
- A recursive algorithm
- Recursion can be avoided: an iterative version
- How efficient is this algorithm?

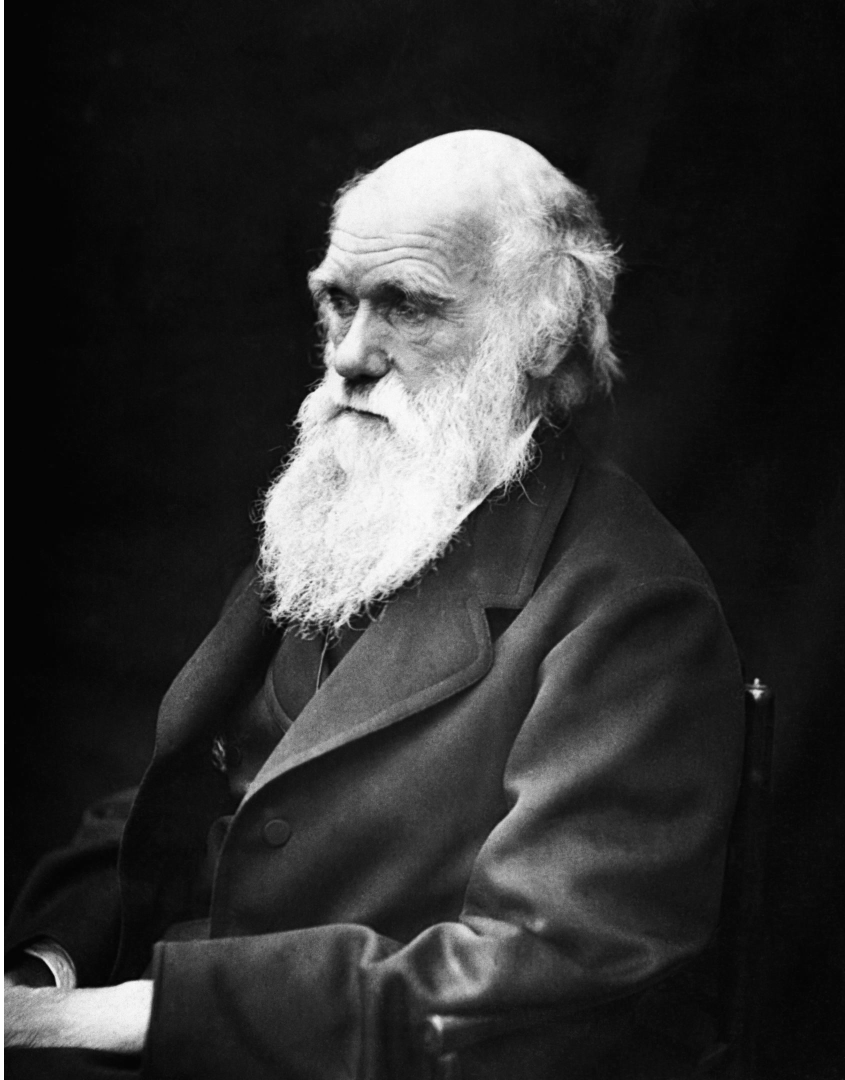


36

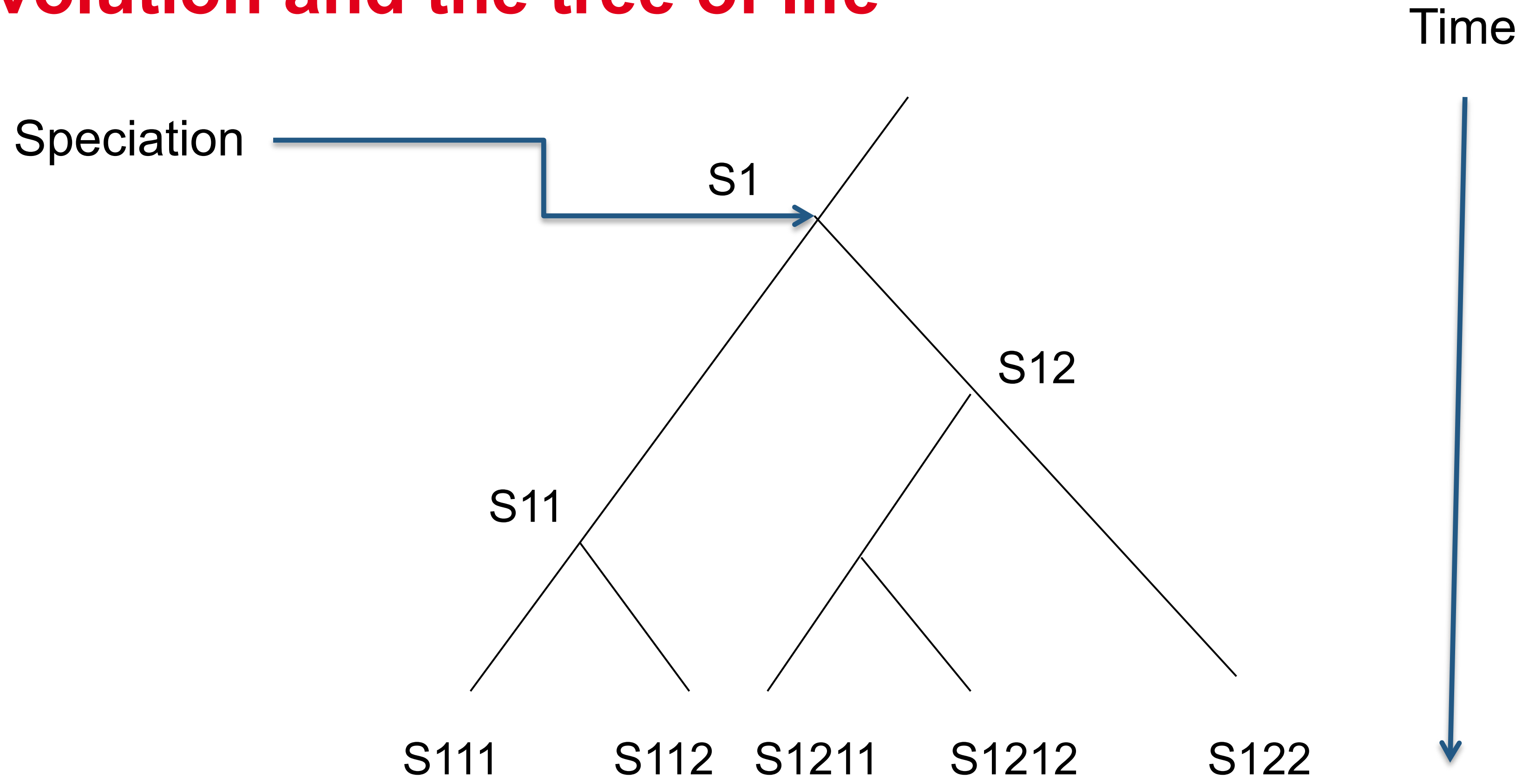
I think

Case must be that one generation then should be as many living as now. To do this & to have many species in same genus (as is). requires extinction.

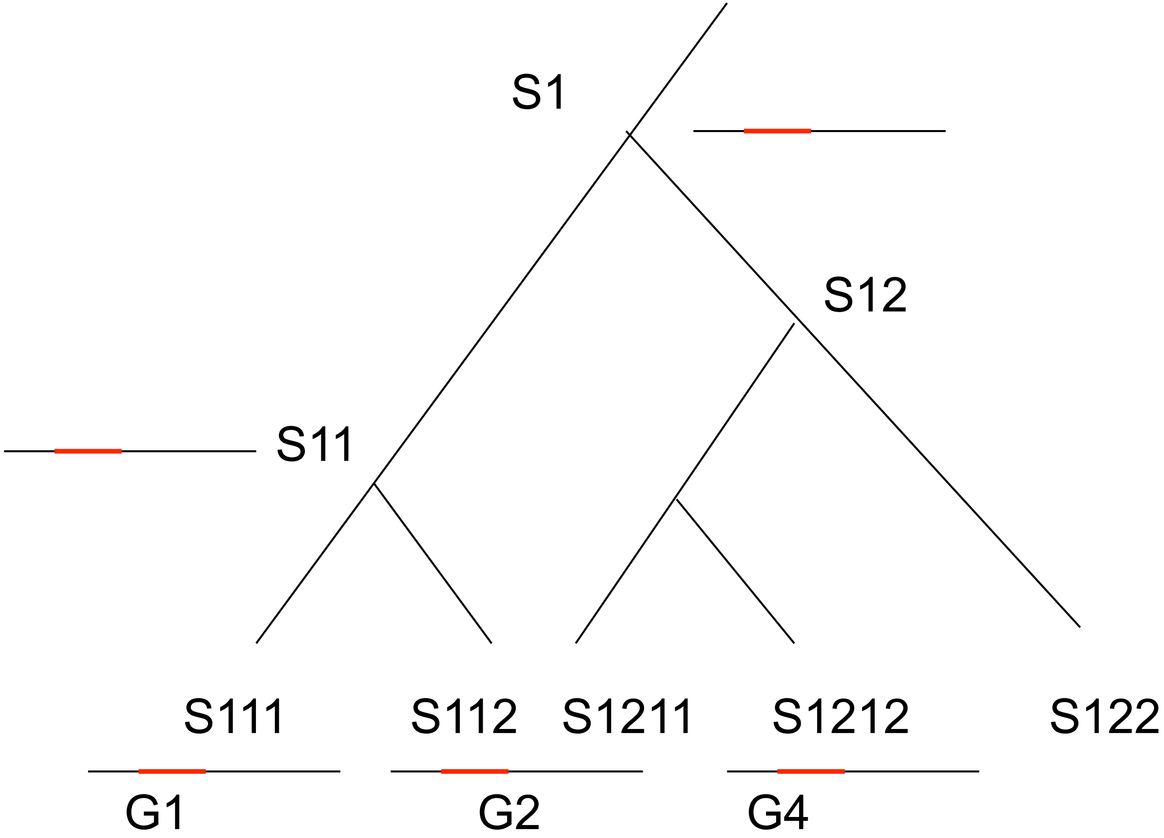
Thus between A & B. immense gap of relation. C & B. the finest gradation, B & D rather greater distinction. Thus genera would be formed. — bearing relation



Evolution and the tree of life



Evolution and the tree of life



Substitutions, deletions, insertions

ACCTCTGTATCTATTCGGGATCATCAT

Substitutions, deletions, insertions

ACCTCTGTATCTATTCGGGATCATCAT

ACCCCTGTATCTATTCGGGATCATCAT

Substitutions, deletions, insertions

```
ACCTCTGTATCTATTCGGGATCATCAT
ACCCCTGTATCTATTCGGGATCATCAT
ACCTCTGTATCT--TCGGGATCATCAT
```

Substitutions, deletions, insertions

ACCTCTGTATCTATTCGGGATCATCAT

ACCCCTGTATCTATTCGGGATCATCAT

ACCTCTGTATCT--TCGGGATCATCAT

ACCTCTGTATCTATTCGTGGATCATCAT

Pictures & movies : material licensing

p. 2 : Public domain, Wikipedia

p. 2 : Public domain, Wikipedia