

The Rarity of Evolution Through Gene Shuffling

Gavin Conant

Institut für Informatik, Universität Leipzig and
Department of Biology, The University of New Mexico



The University of New Mexico



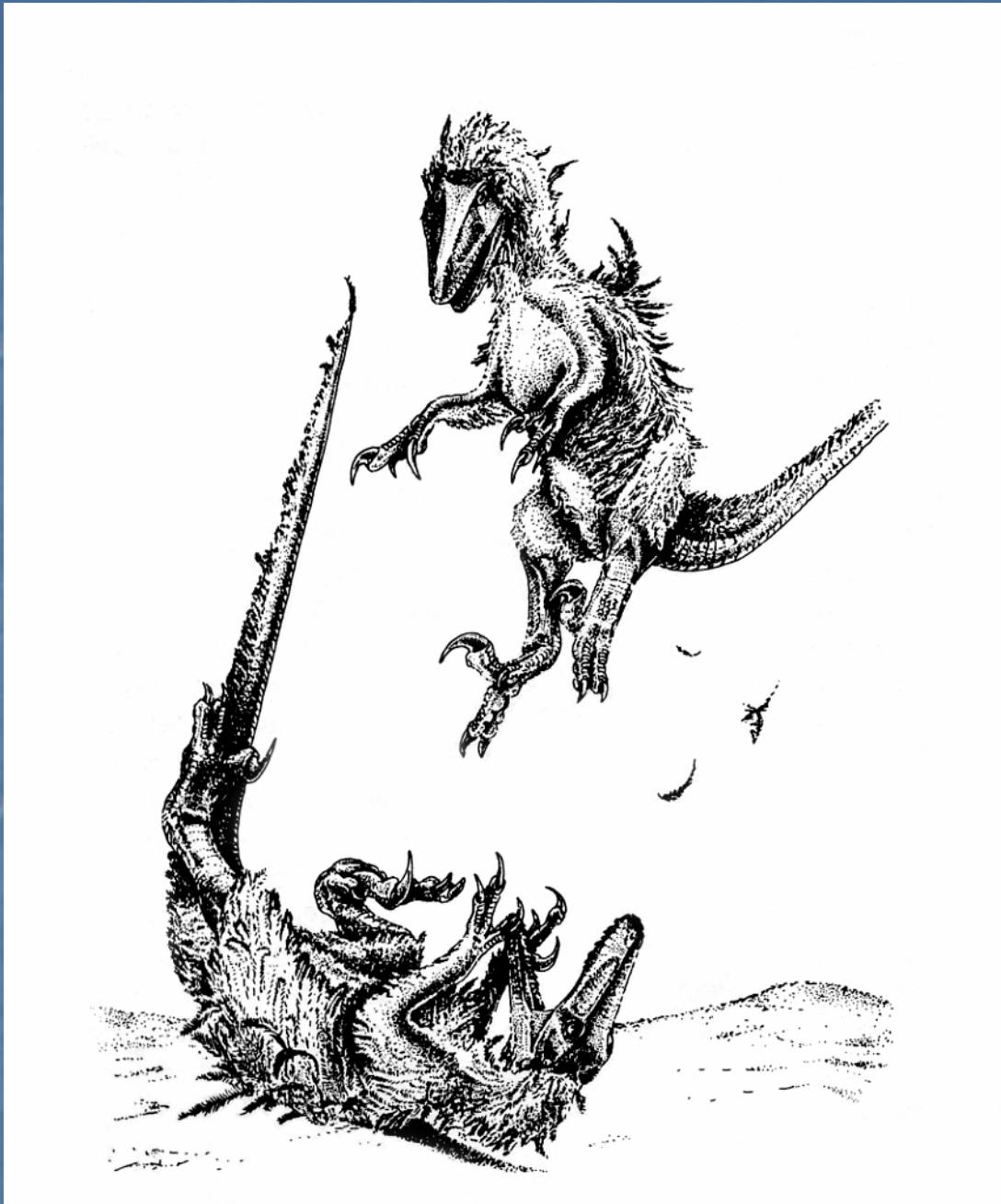
Overview

- Introduction: Origins of evolutionary novelty
- What is “gene shuffling”?
- Frequency of shuffling in genomes
- Parting thoughts
- Acknowledgements

Origins of novel structures

- One of Darwin's most difficult problems was to explain how novel structures could be constructed by evolution





From *The Dinosaur Heresies*, Bakker, 1986

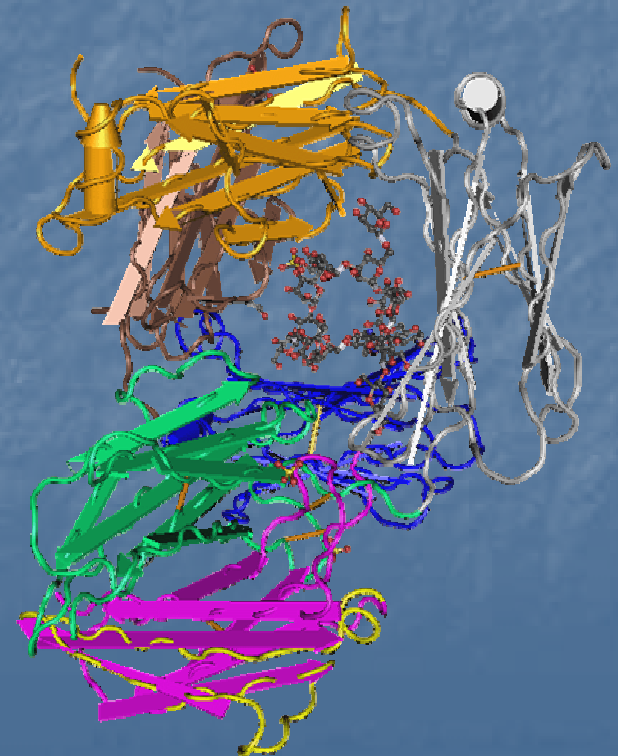


Ji *et. al*, 2001 *Nature* 410:1084

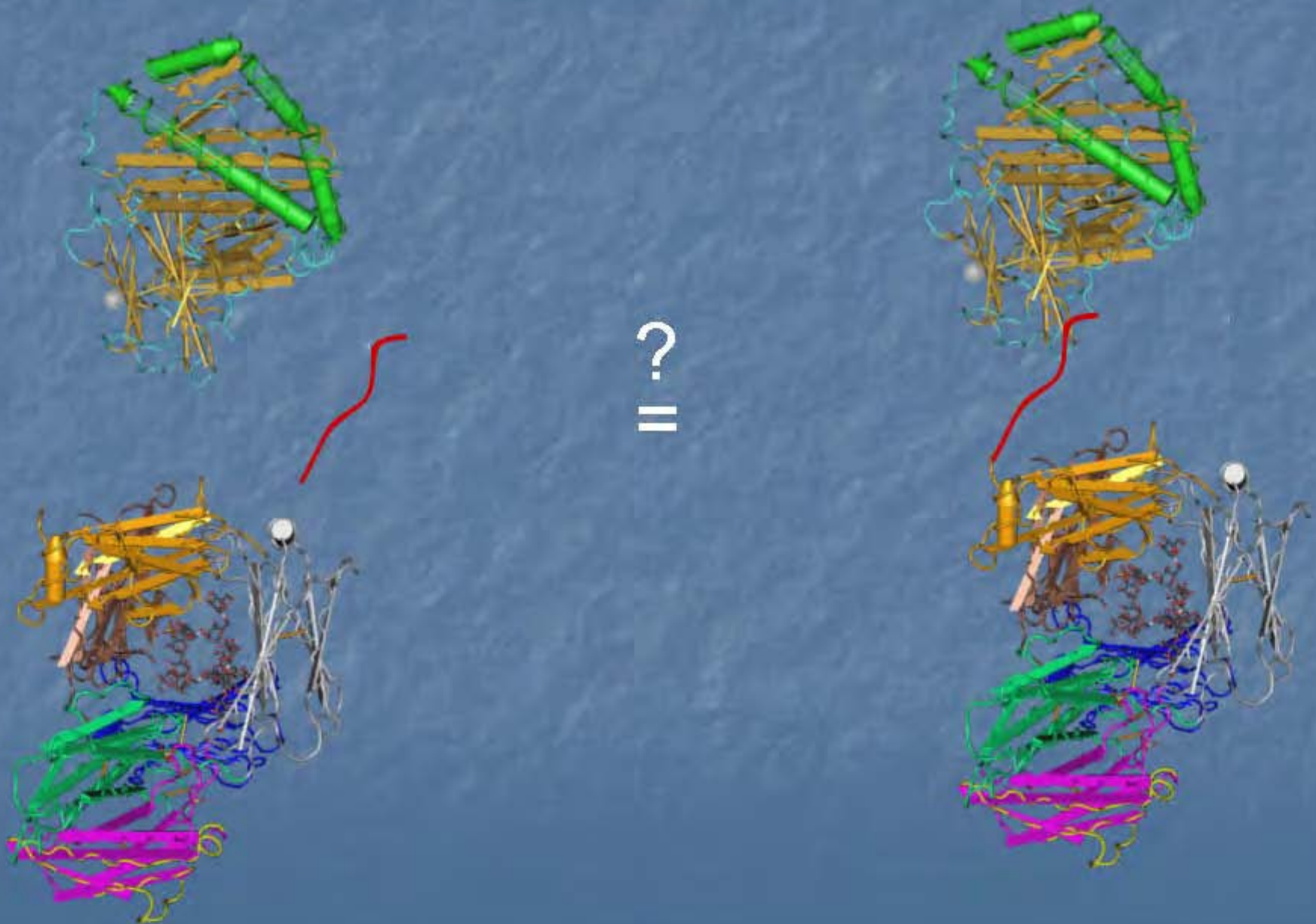
Proteins

- A linear sequence of “residues” is encoded in the genome and converted to the protein “alphabet”
- By a process that is still imperfectly understood, that linear polymer folds into a specific 3-dimensional shape

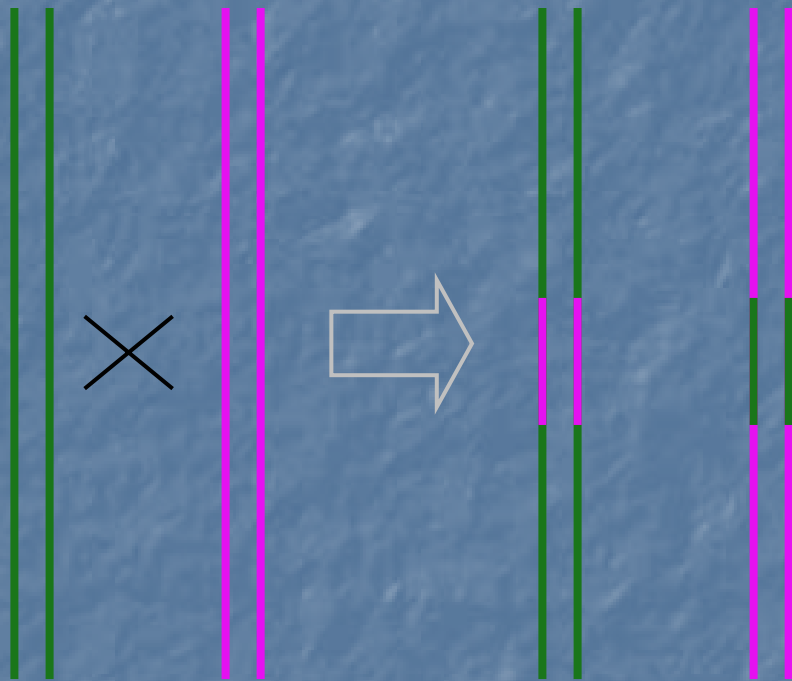
MTTYRSTAAGHKKTE



Are proteins modular?

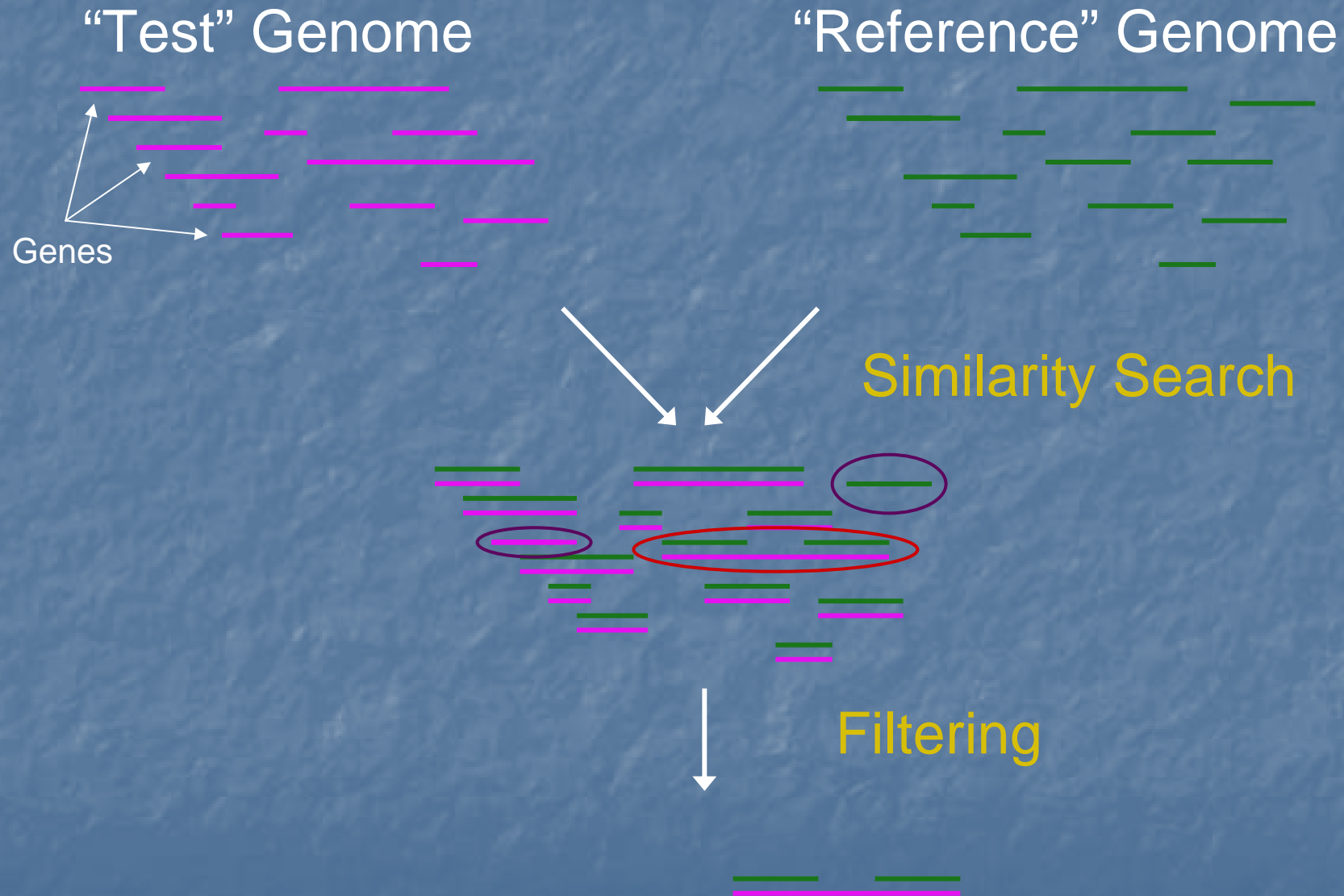


DNA Recombination



- Usually requires some sequence similarity
- Can introduce chimeras composed of parts of 2 genes

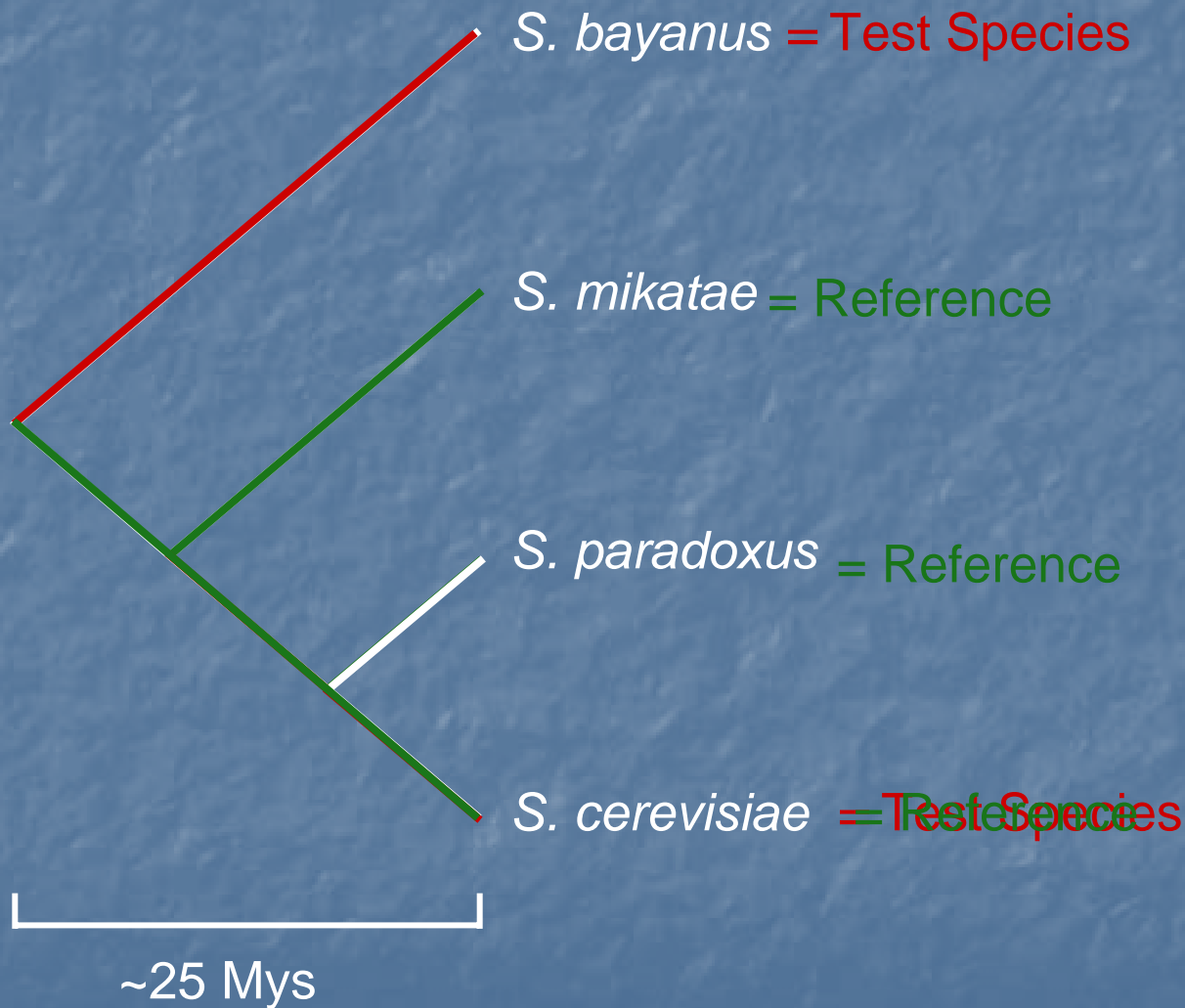
Scanning Genomes for "Shuffled Genes"



Analysis details:

- All protein-encoding genes in the test species were compared to the reference using BLAST
- Significant hits were locally aligned
- Test species proteins with non-overlapping hits to two reference species genes were analyzed with a special purpose algorithm
- The result was a list of putatively shuffled genes

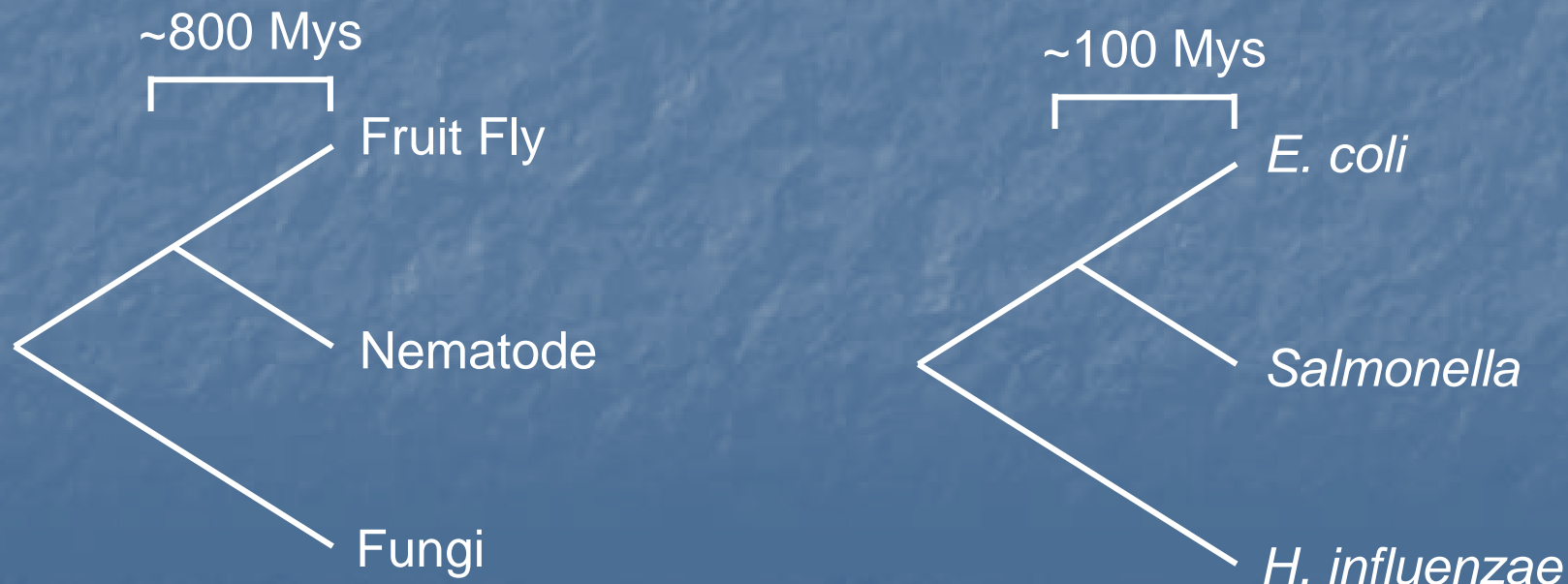
Recent shuffling Events

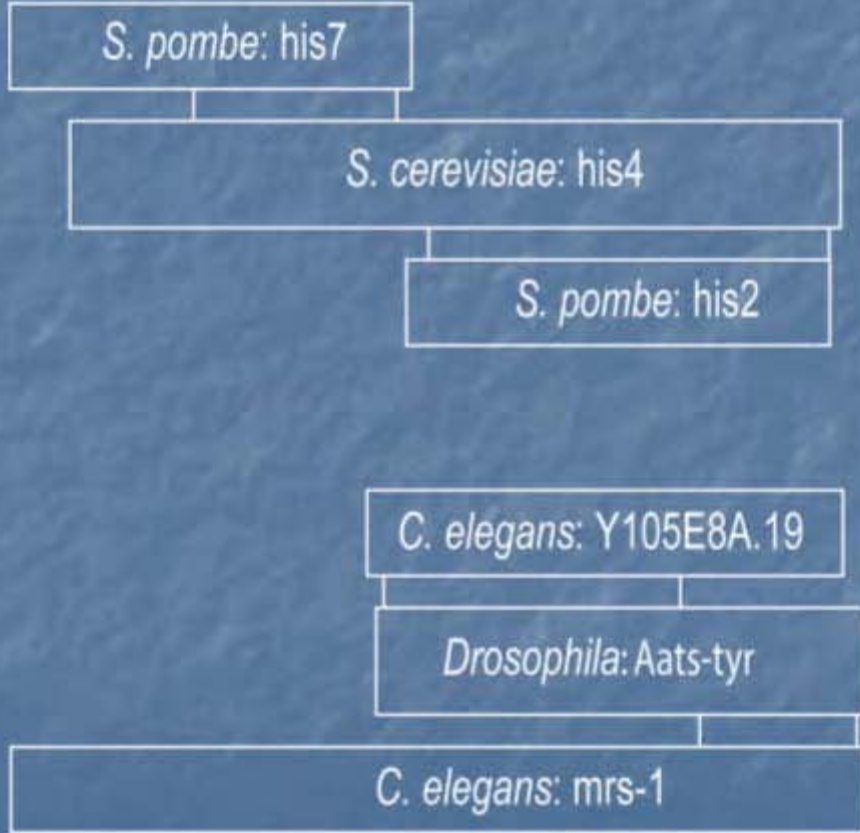
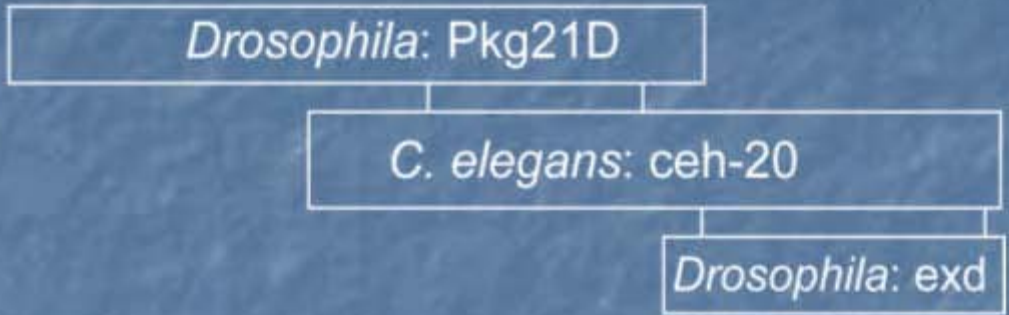
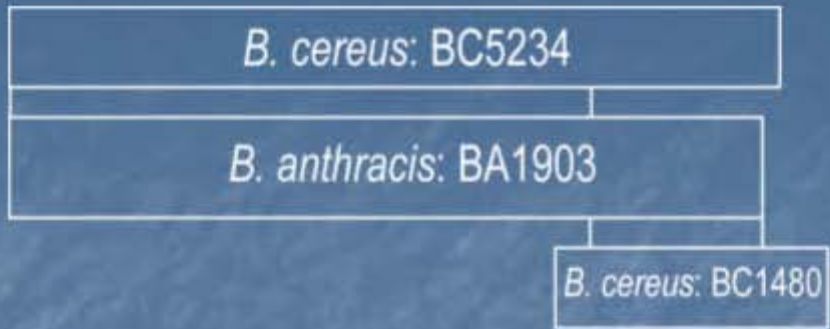


- We hypothesized that gene shuffling would be common....,
- but actually it was rare to non-existent

Large-scale Analysis

- We next took a broad taxonomic approach, surveying species from all three kingdoms of life:
 - 4 Bacteria, including anthrax and *E. coli*
 - 4 Eukaryotes: including yeast (again) and fruit fly
 - 2 Archeans

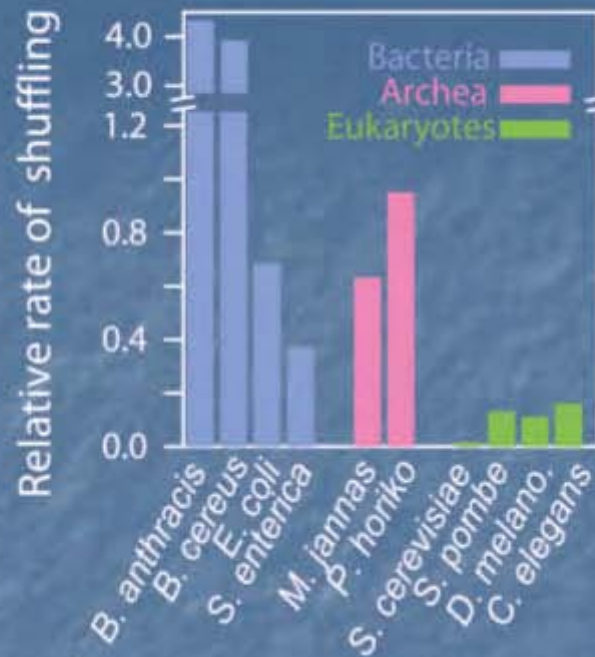




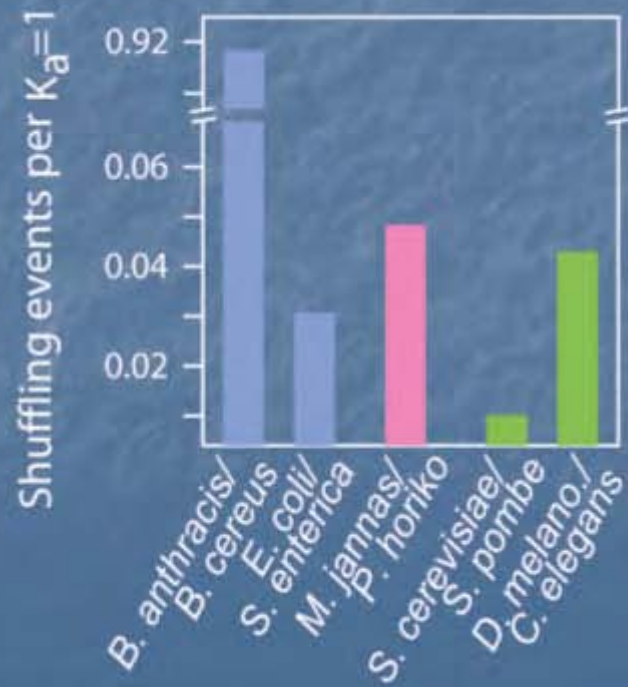
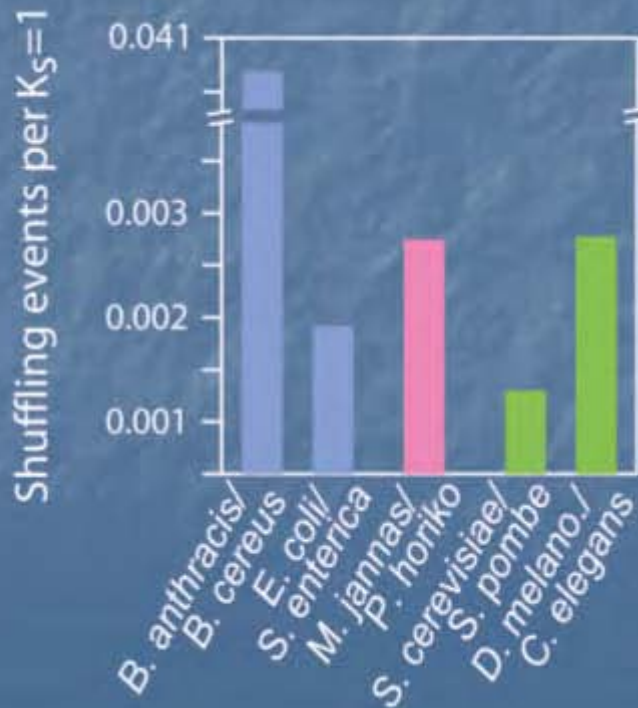
Time Calibration

- We calibrated our observed shuffling events against two other kinds of events:
 - The number of duplication events since the common ancestor
 - The average number of gene sequence changes since the common ancestor
- Absolute divergence times are not known for many test species
- “Molecular clocks” can give approximate time estimates

Duplication Clock



Sequence-change Clocks



Conclusions

- Examples of gene shuffling were found
- Shuffling seems to be quite rare
 - In fruit flies, only about 1 shuffling event is preserved every 4 million years
 - In that same period, at least 40 gene duplications would be expected (shuffling rates was $>10\%$ of duplication rate across the eukaryotes)
- Differences in rates between prokaryotes and eukaryotes may be related to the differential strength of genetic drift and natural selection in these populations

Parting Thoughts

Origins of novelty

- Gene shuffling is likely still an important source of evolutionary novelty, especially in bacteria
- Modularity may not be the primary mode of innovation in protein sequences

Acknowledgements

- Andreas Wagner
- Annette Evangelisti
- Michael Fuller
- Michael Gilchrist
- The Department of Energy and the Krell Institute



The University of New Mexico

