

# Evolutionary Genomics and the Origin of Life: Insights and Constraints

Greg Fournier  
Department of Biological Engineering  
MIT



*in parrington*



# Outline

1. Role of evolutionary biology in Origin of Life research
2. Complex histories in the Tree of Life
3. Reconstructing ancestral sequences

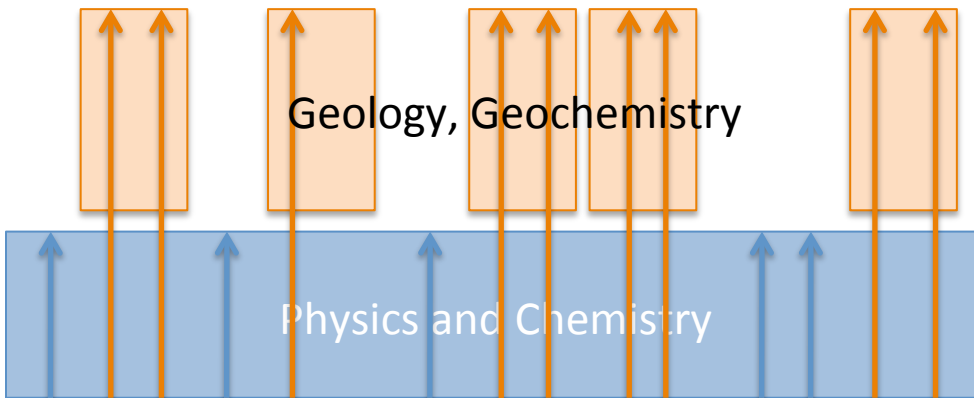
1. How does biological information inform “pre-biotic” investigations?



Theoretical Paths to  
Biocomplexity

Contingency

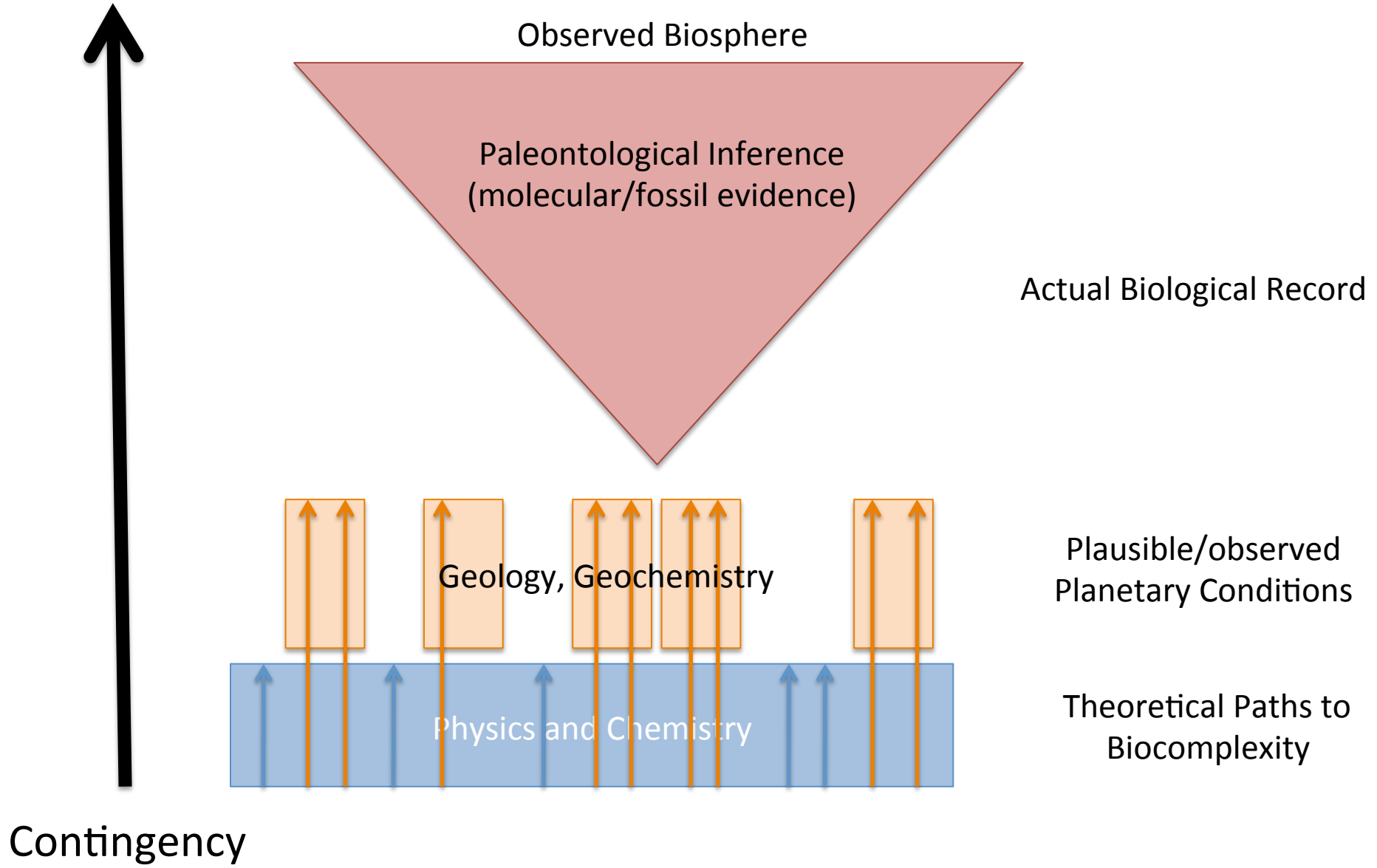


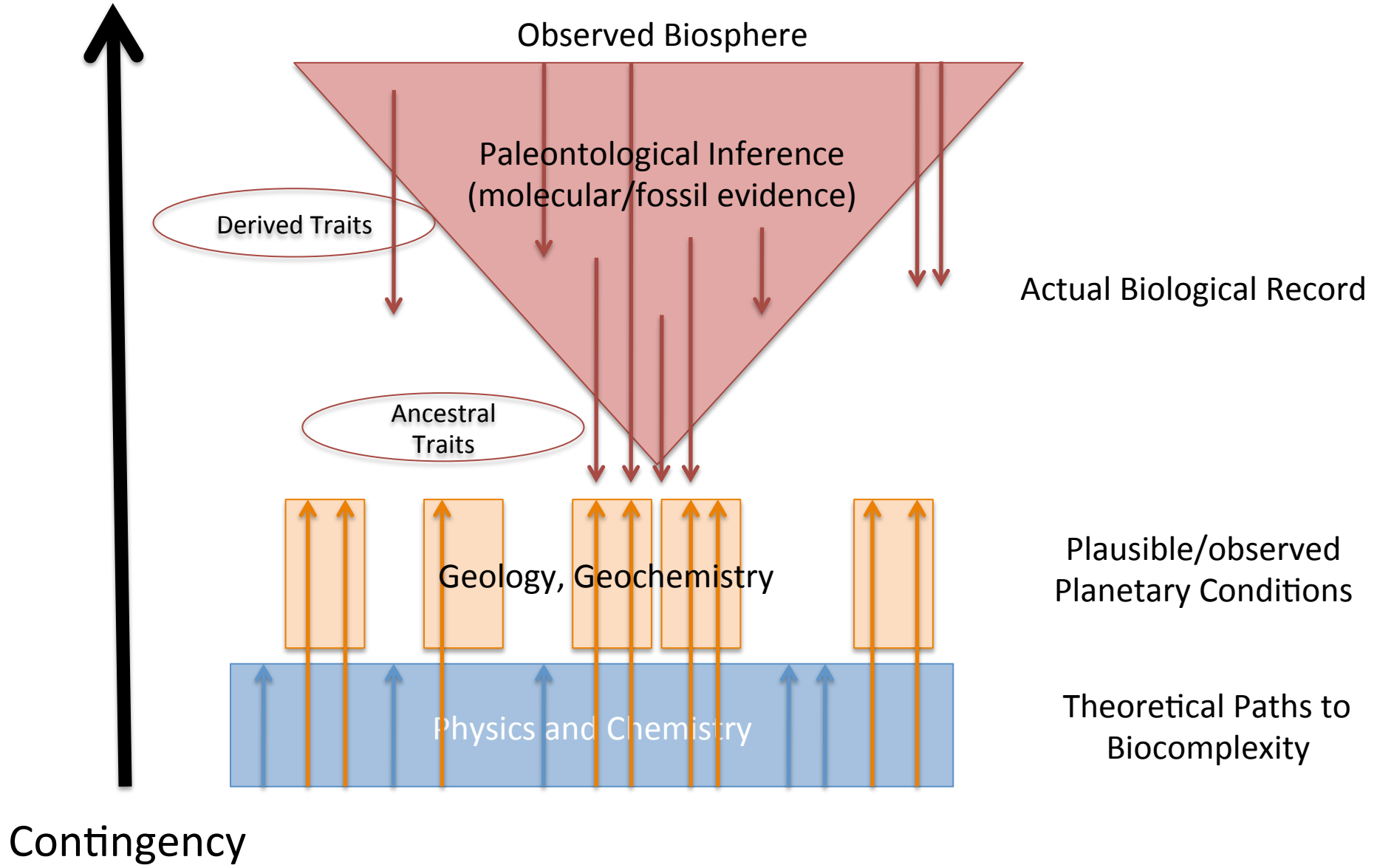


Plausible/observed  
Planetary Conditions

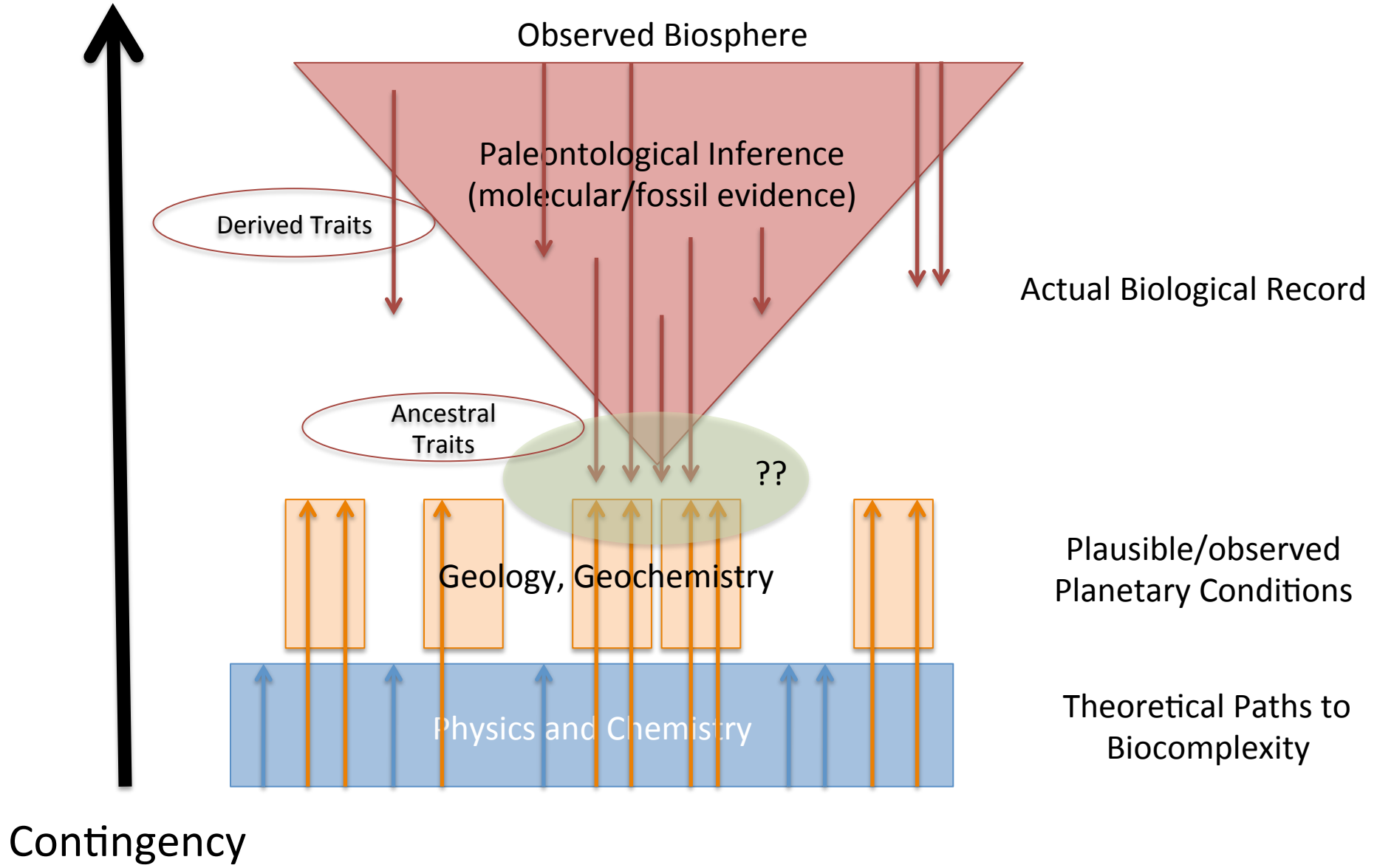
Theoretical Paths to  
Biocomplexity

Contingency





Contingency



Observed Biosphere

Paleontological Inference  
(molecular/fossil evidence)

Derived Traits

Actual Biological Record

Ancestral  
Traits

??

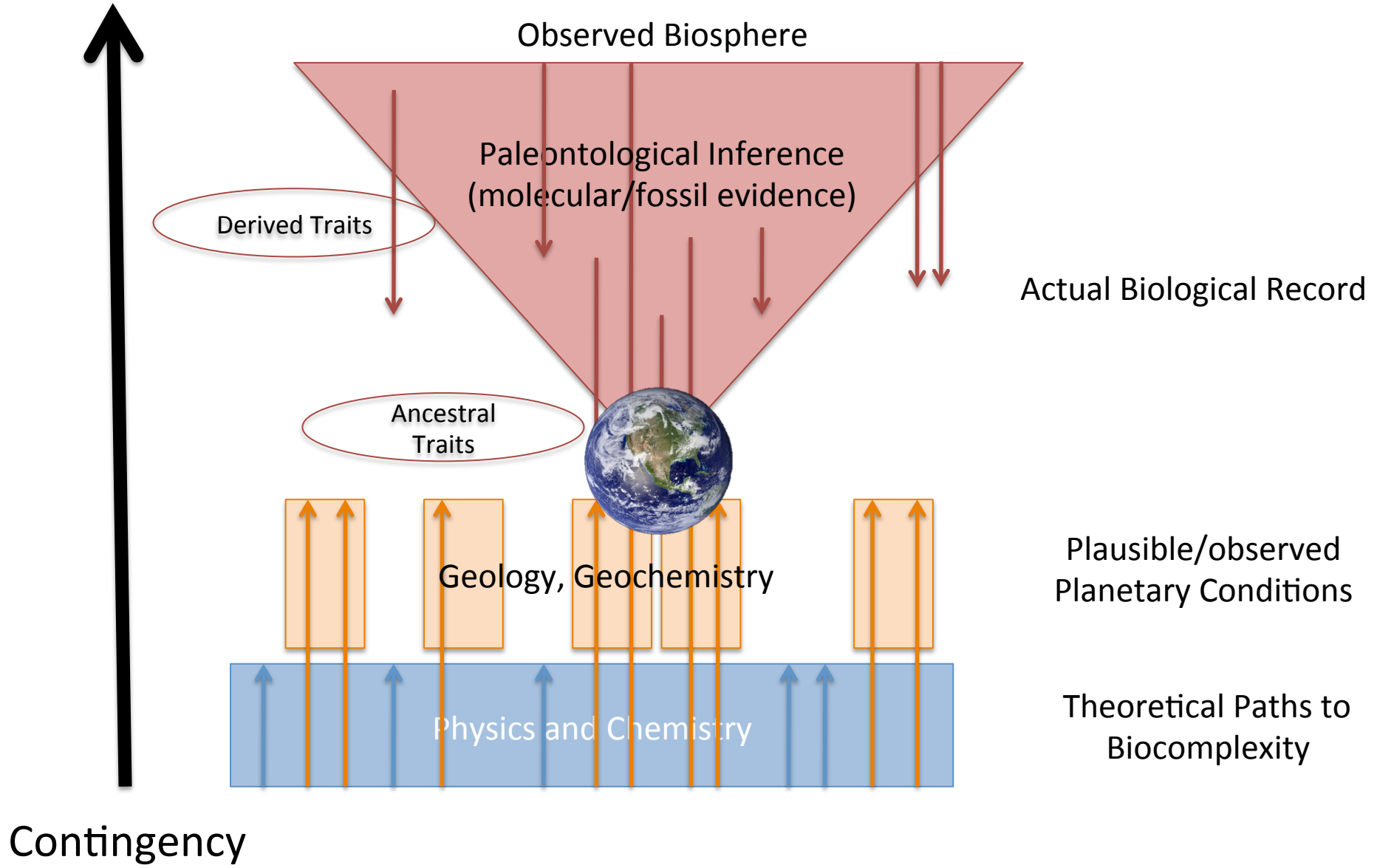
Geology, Geochemistry

Plausible/observed  
Planetary Conditions

Physics and Chemistry

Theoretical Paths to  
Biocomplexity

Contingency



Observed Biosphere

Paleontological Inference  
(molecular/fossil evidence)

Derived Traits

Actual Biological Record

Ancestral  
Traits

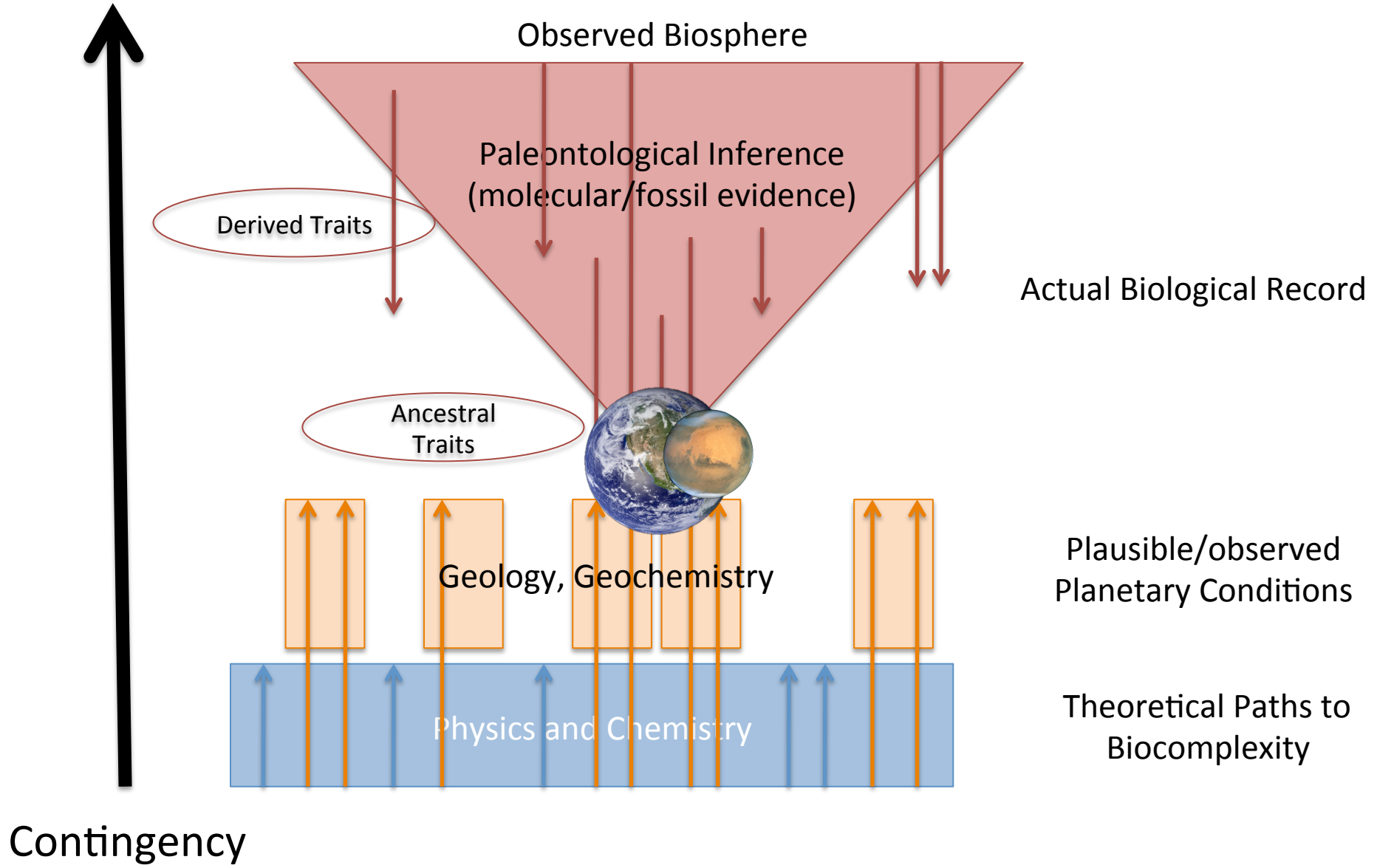
Geology, Geochemistry

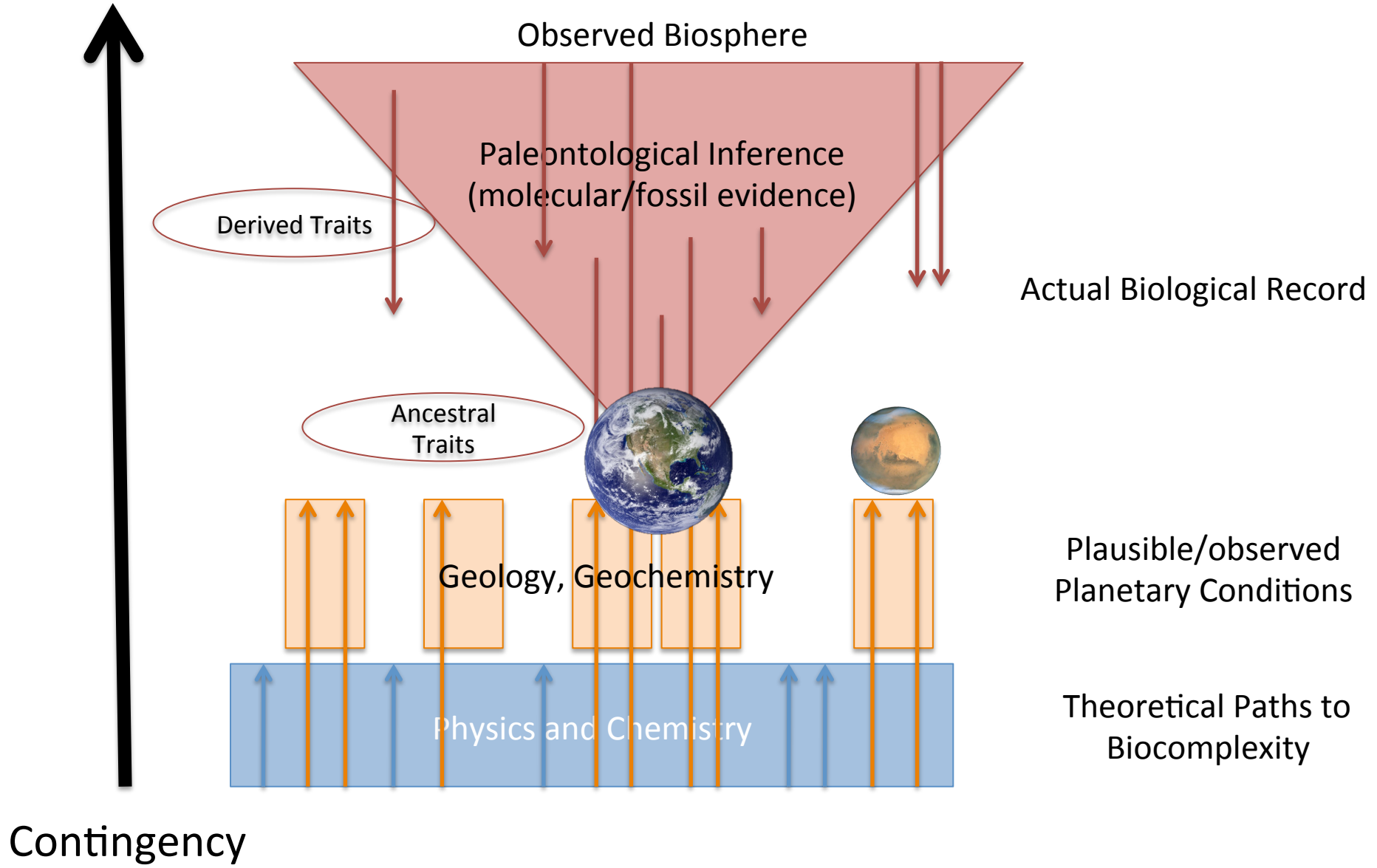
Plausible/observed  
Planetary Conditions

Physics and Chemistry

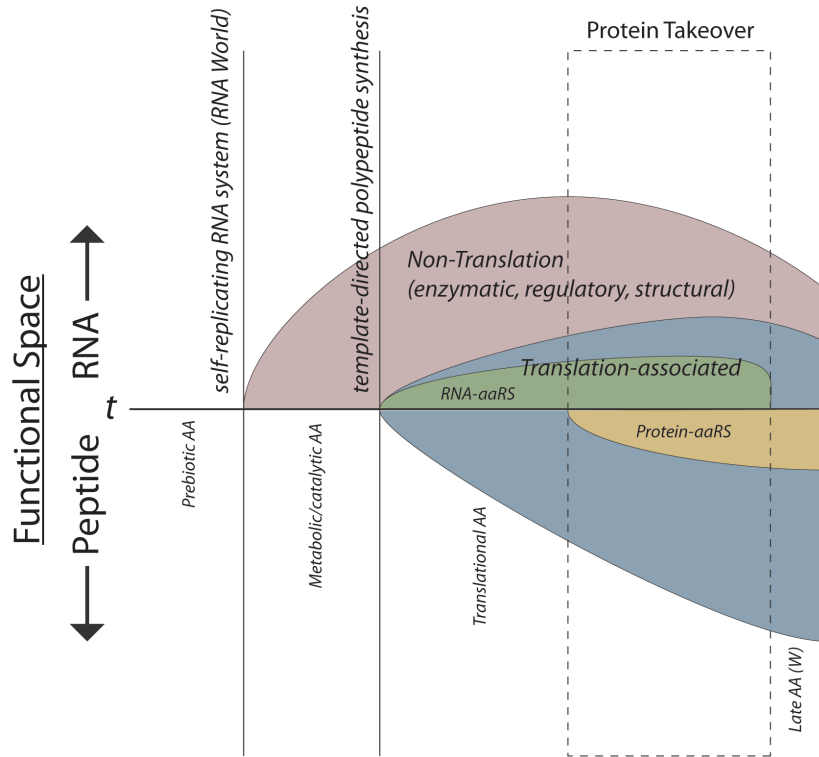
Theoretical Paths to  
Biocomplexity

Contingency

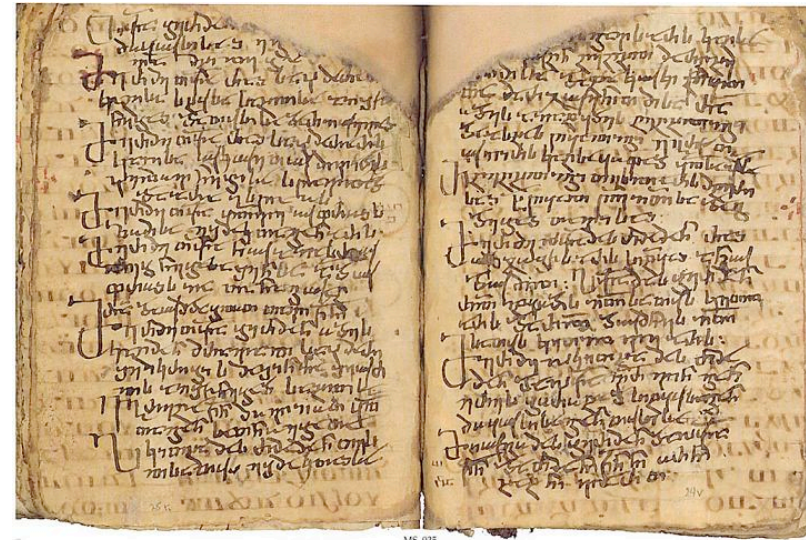




# Is the earliest history of life “Overwritten?”



Modern Cell Biology (Cenancestor)



MS 018  
Codex Sinaiticus Zosimi Rescriptus, Palimpsest, Mt. Sinai, Egypt, 6th c. and 979

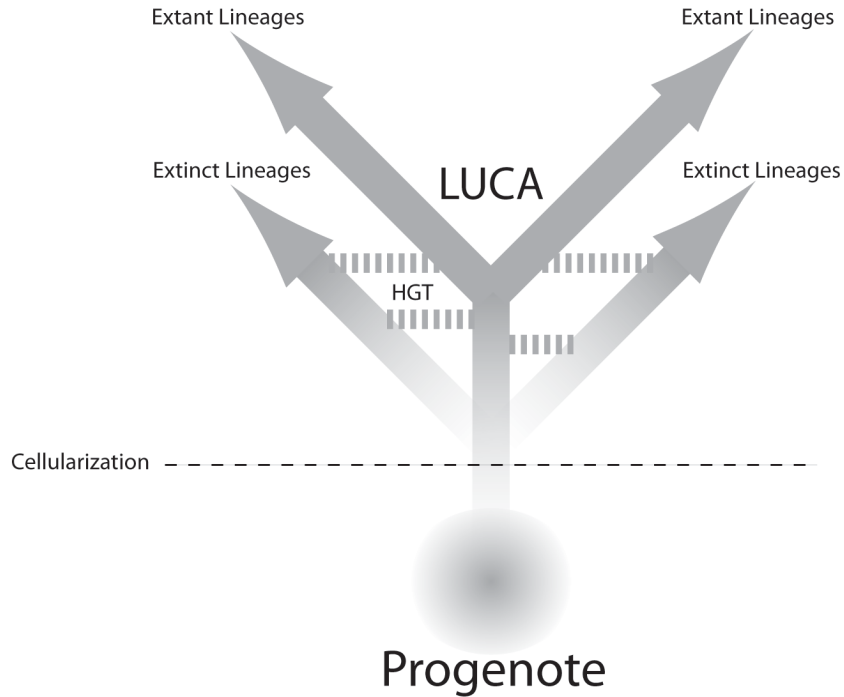
Egyptian codex, 6<sup>th</sup> century and 979 CE

Takeover from RNA-based physiology

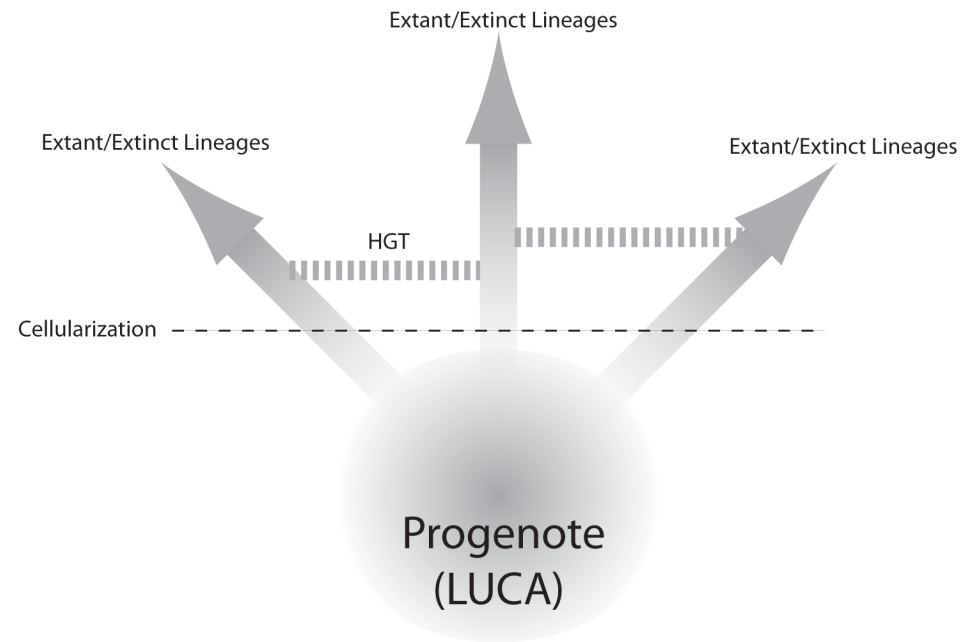


# Nature of the Last Universal Common Ancestor

## LUCA Late

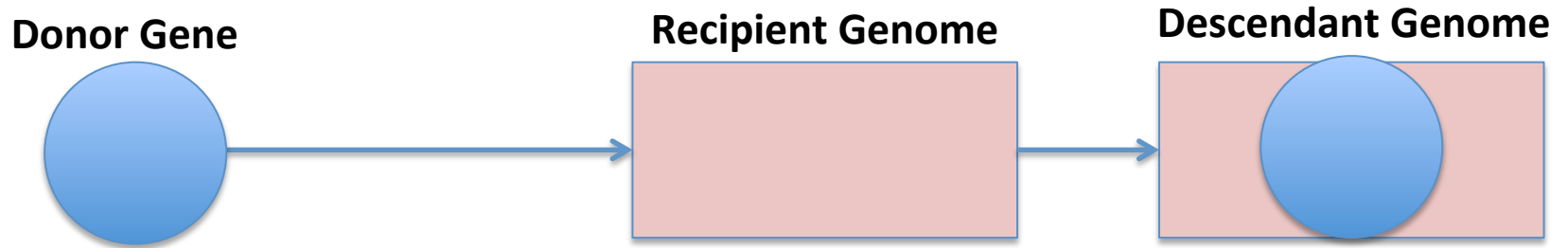


## LUCA Early

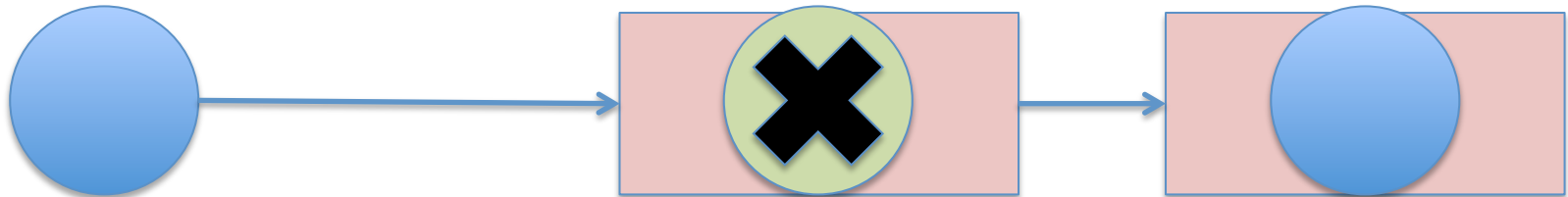


## 2. Horizontal Gene Transfer and Coalescence

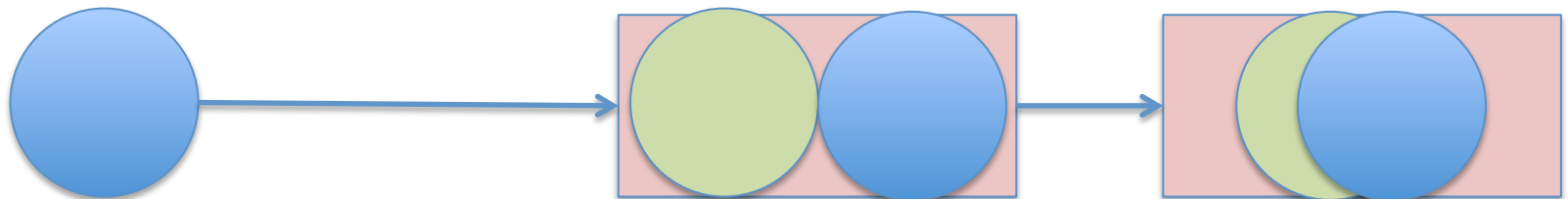
## Novel Acquisition



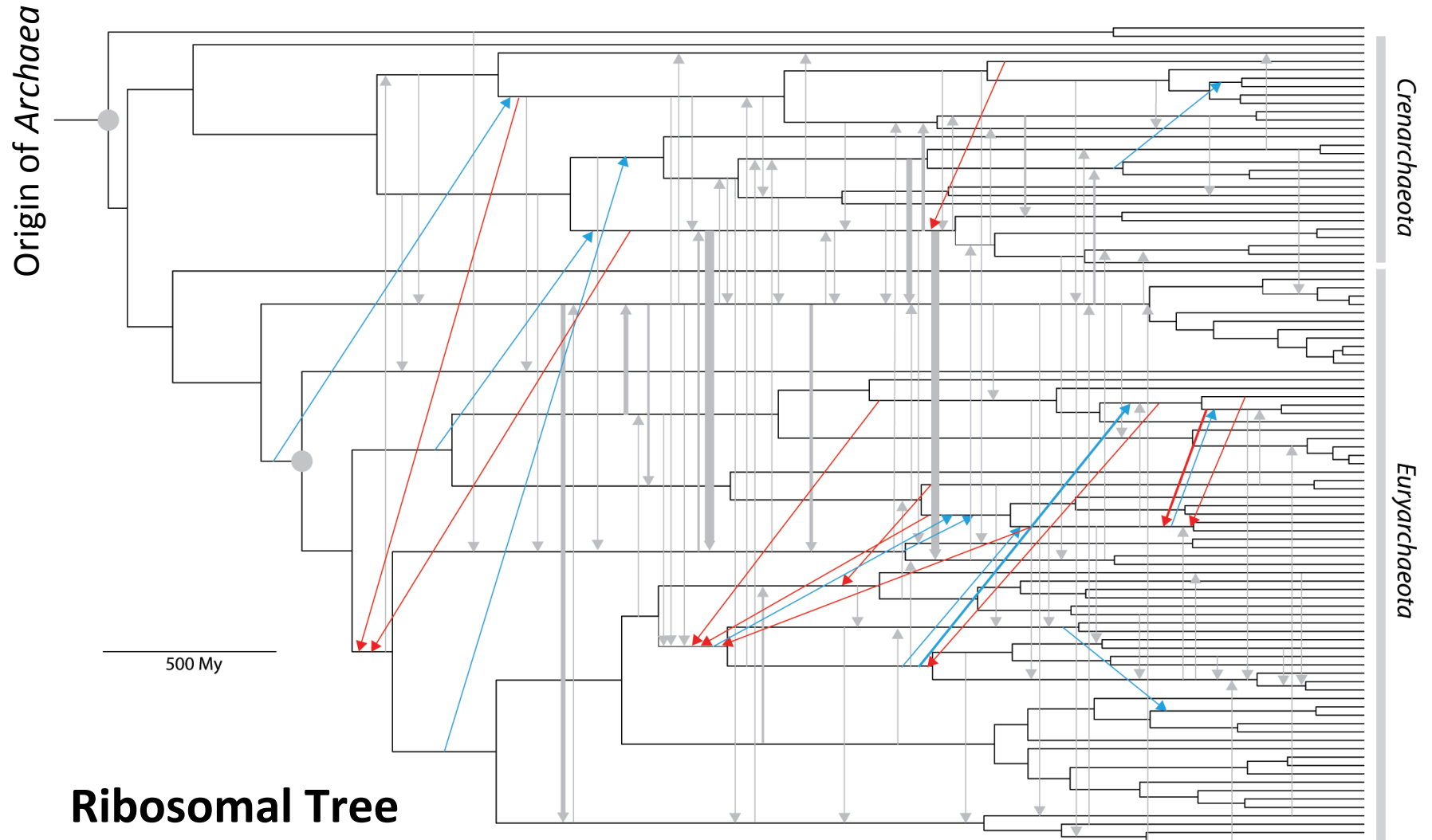
## Orthologous Displacement

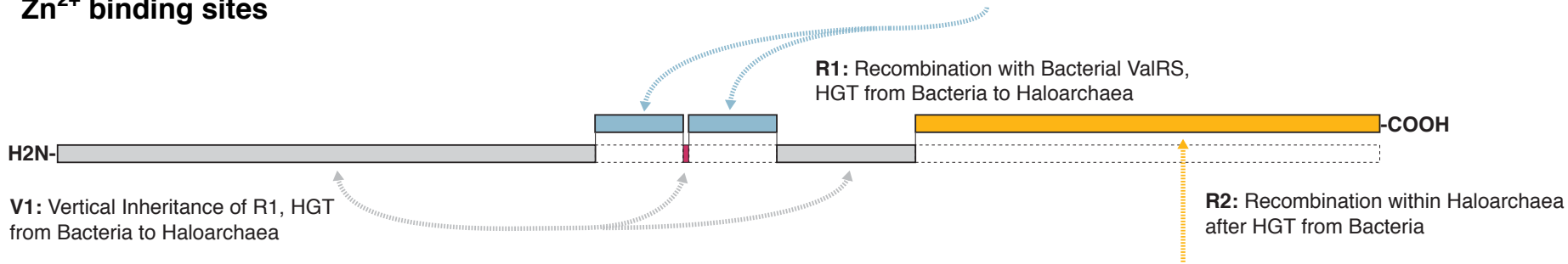
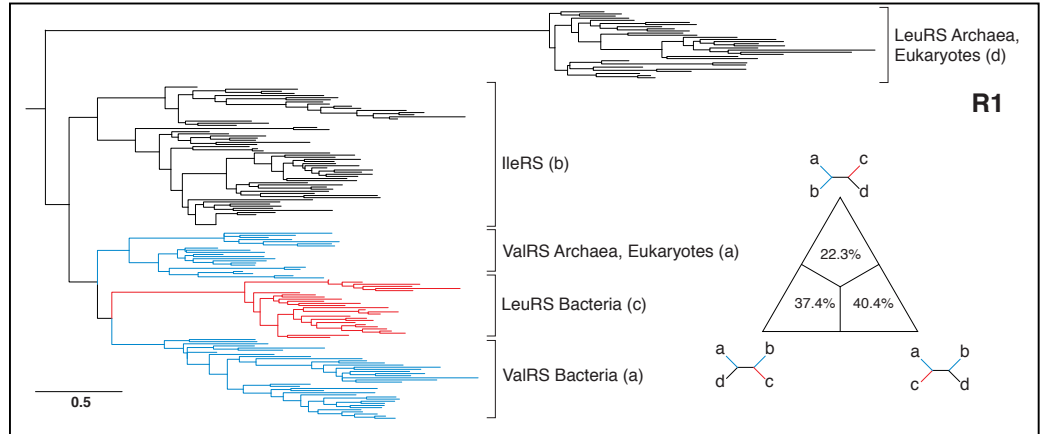
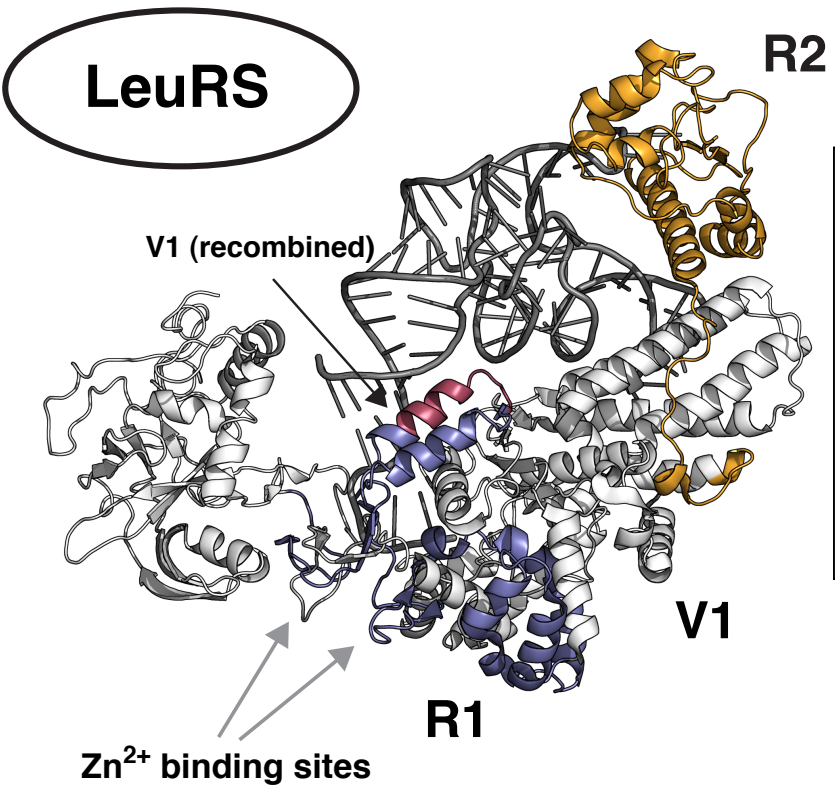


## Orthologous Displacement with Recombination



# HGT is extensive and conflates gene origins



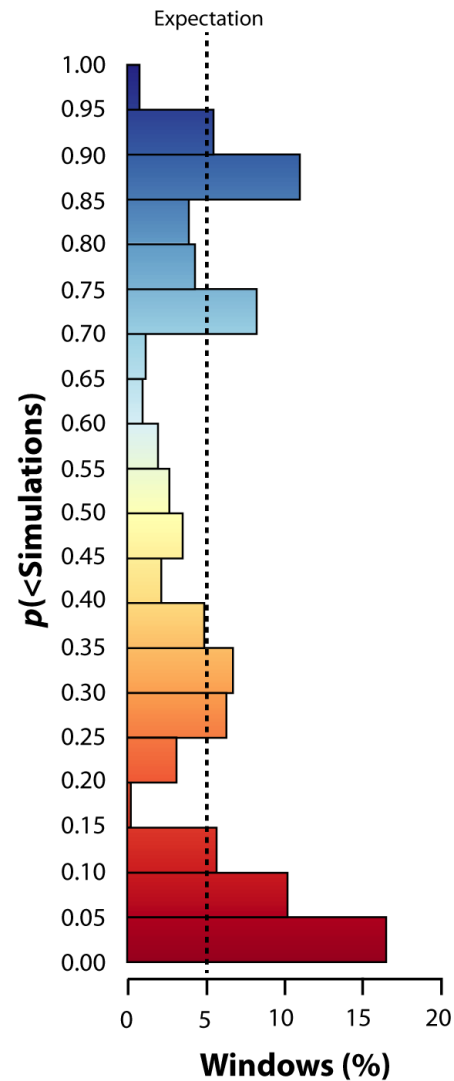
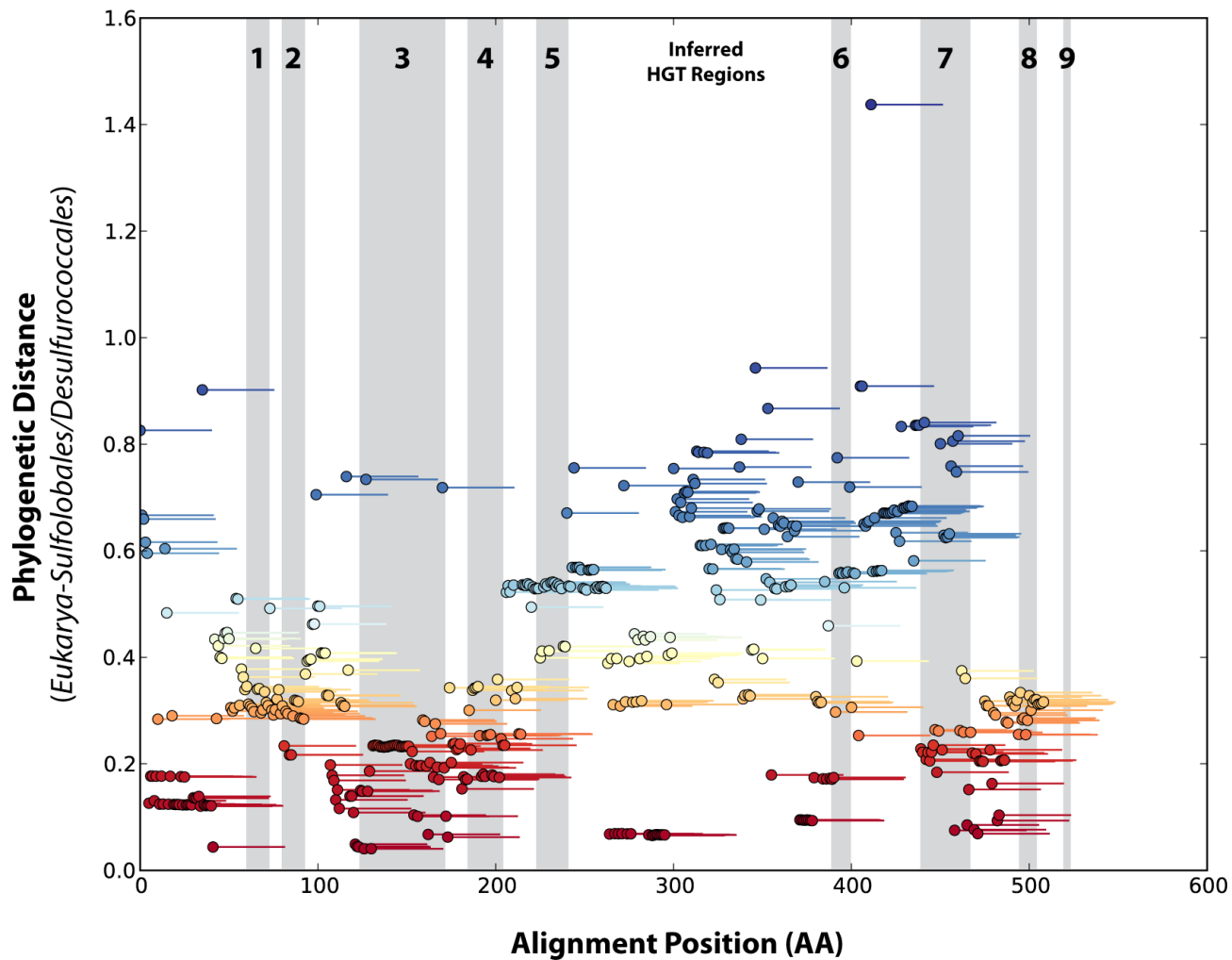


# Recombination in ancient HGT events

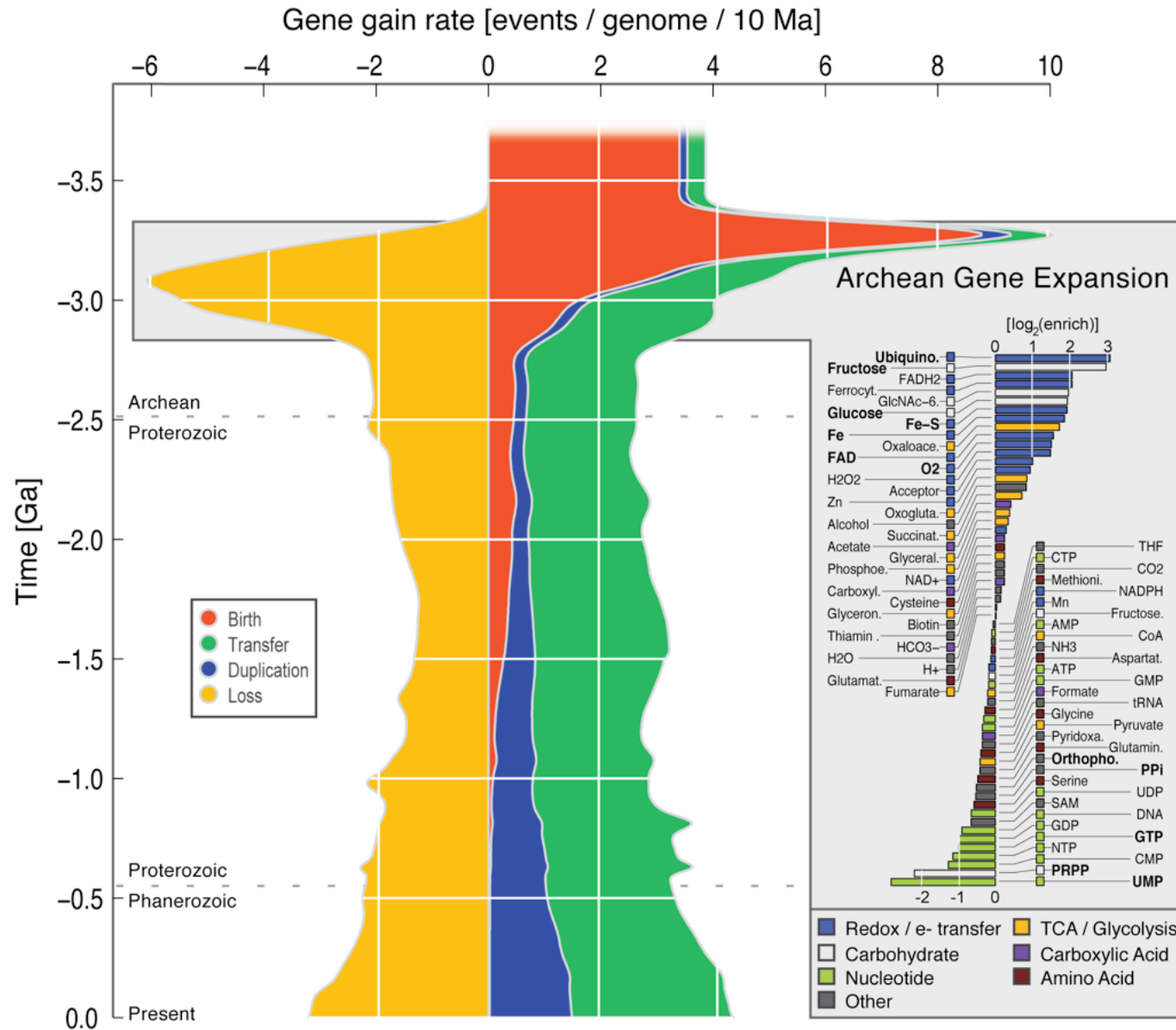
# Detecting Multiple Evolutionary Histories within Single Genes

EF-1 $\alpha$

*Eukarya* Sequence Windows



# Patterns of Gene Family Evolution across the Tree of Life

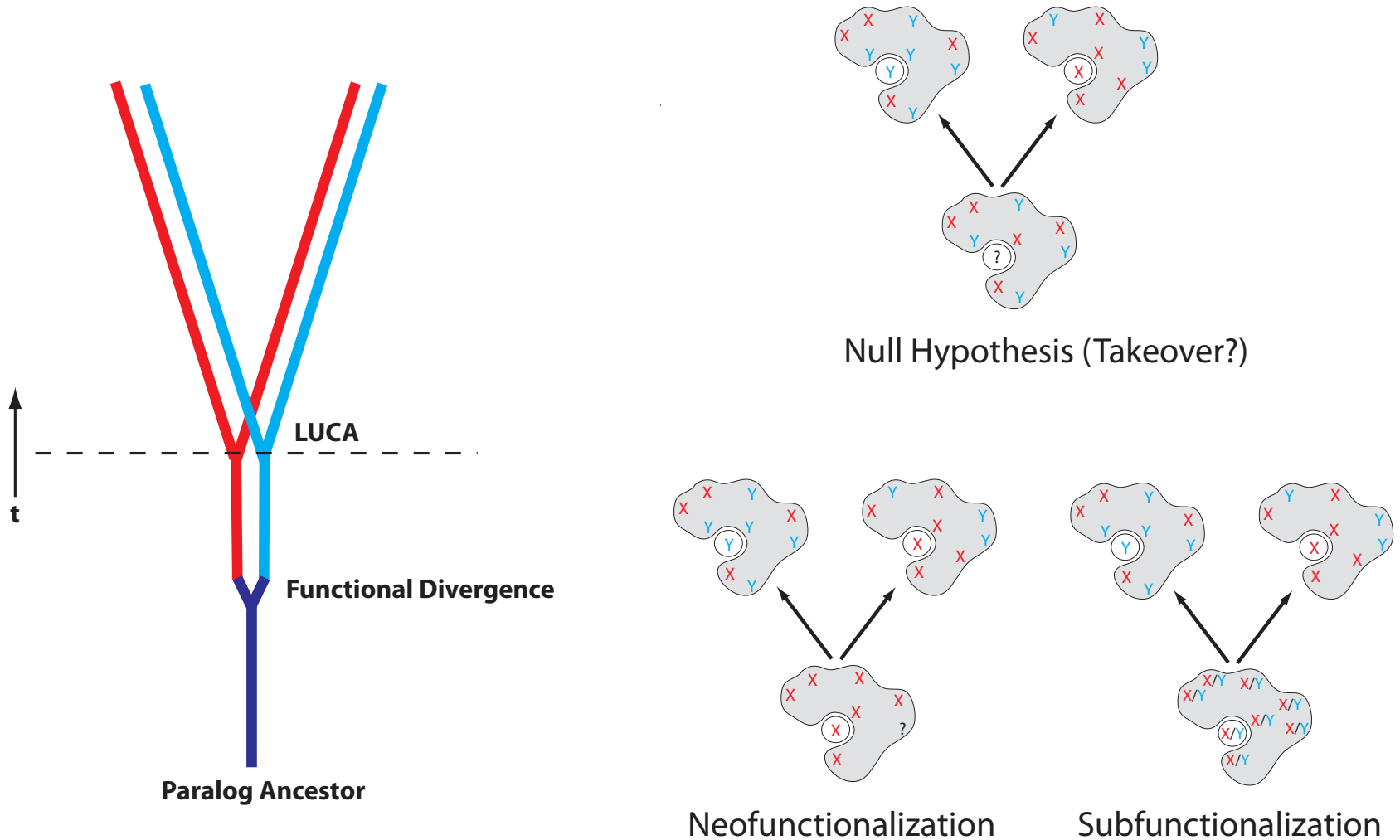


# 3. Reconstruction of Ancestral Sequences

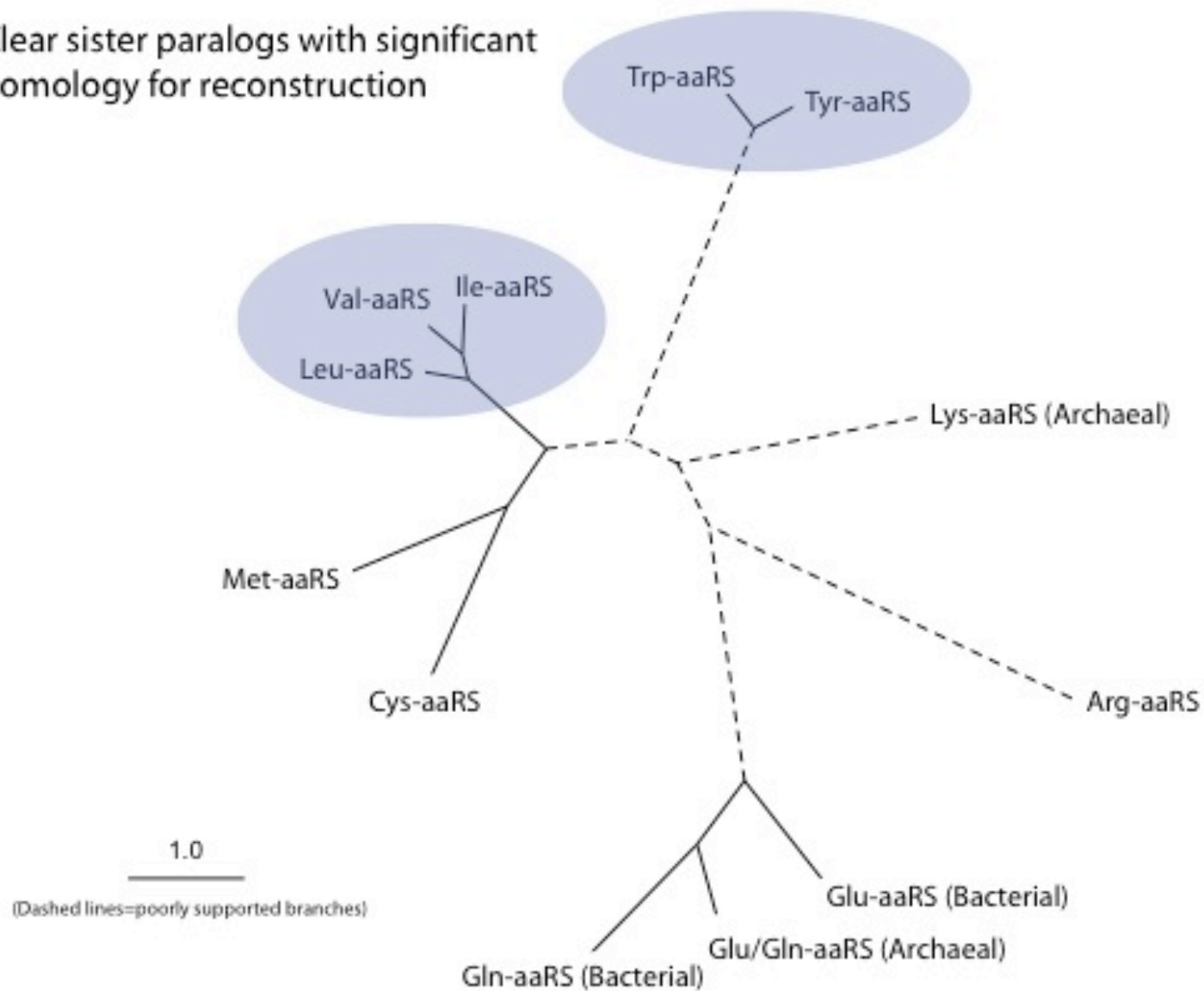




# Pre-LUCA evolution of aminoacyl-tRNA synthetases

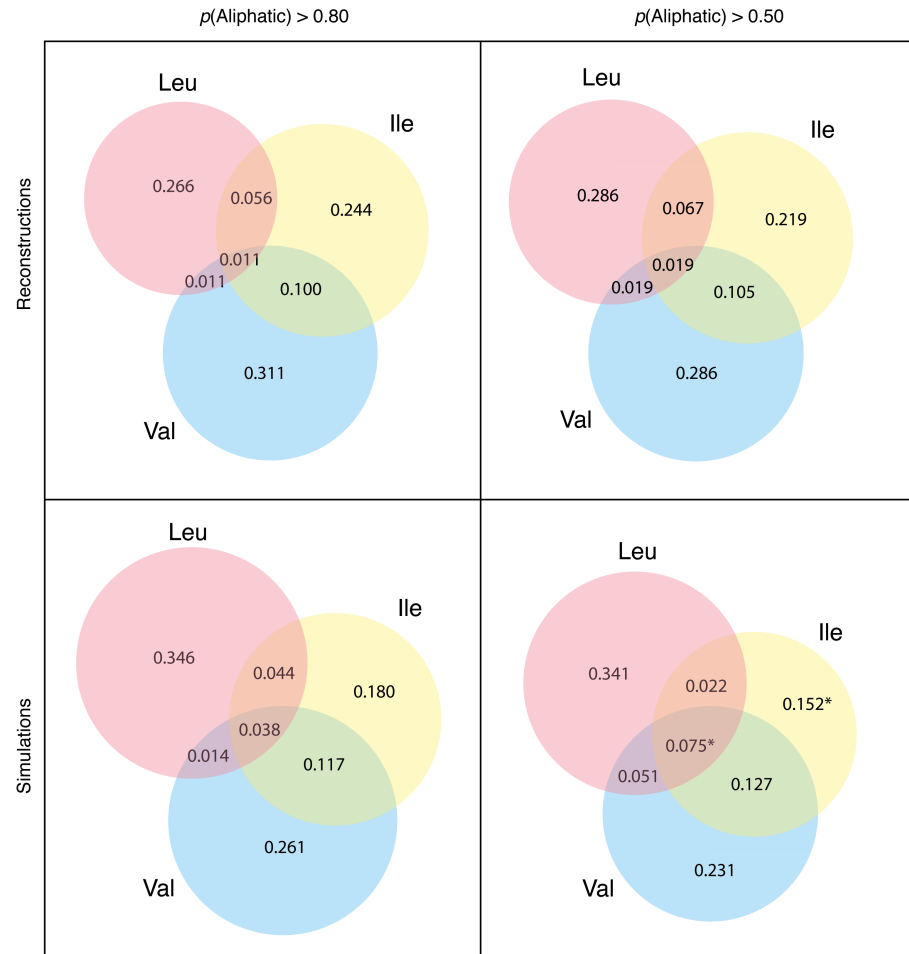


Clear sister paralogs with significant homology for reconstruction



# ValRS-IleRS pre-LUCA Ancestor contained both Ile and Val

Reconstructions



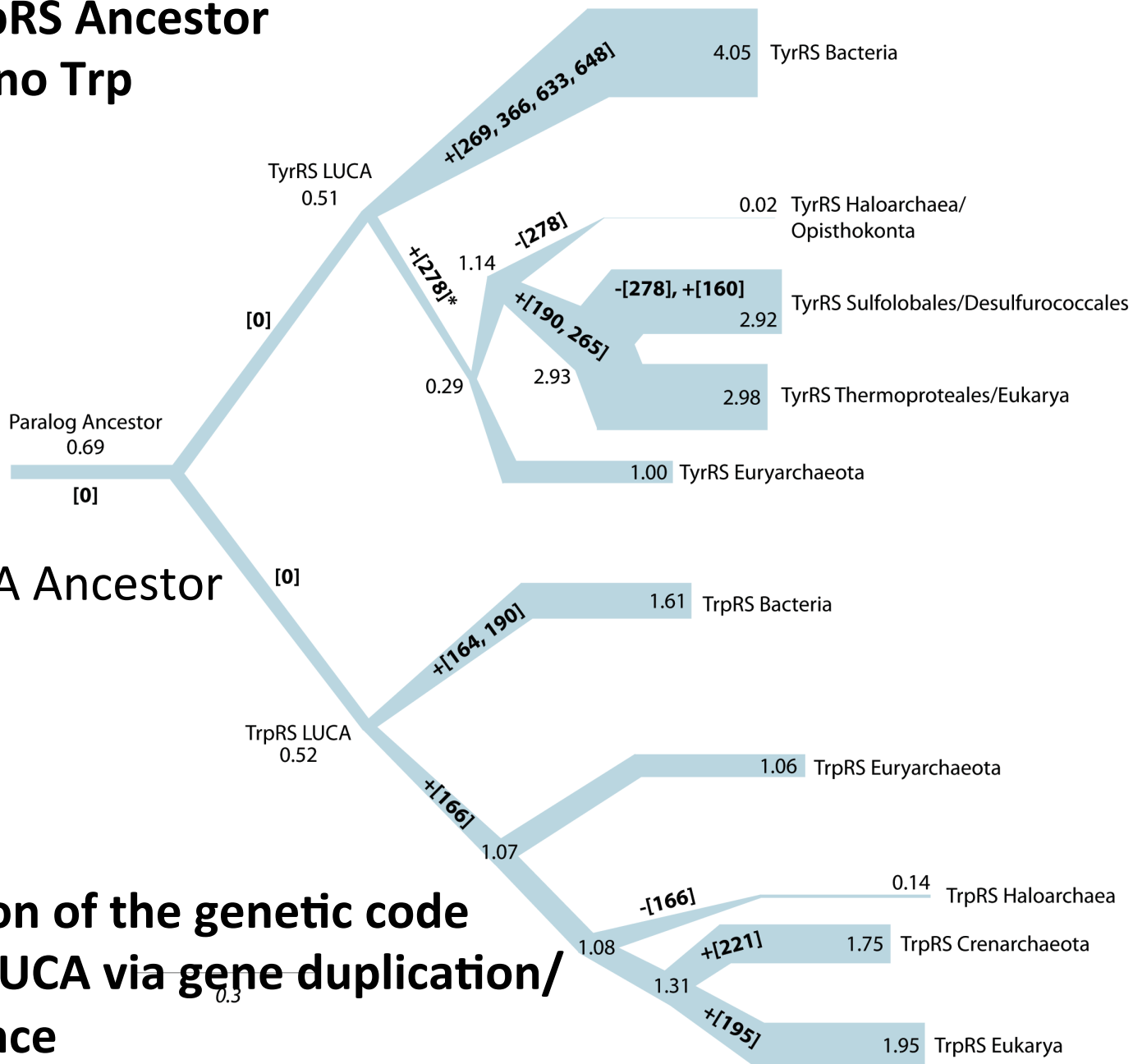
Simulations

Ile and Val use in the genetic code predates cognate proteins:  
**Takeover of more ancient system?**

# TyrRS-TrpRS Ancestor contains no Trp

pre-LUCA Ancestor

Expansion of the genetic code before LUCA via gene duplication/divergence



# Conclusions

- Biological information provides valuable “top-down” perspective and constraint on Origin of Life questions, and is a strong complement to prebiotic chemistry investigations;
- Understanding complex patterns of inheritance within genomes lets us extract more precise and accurate information about the earliest biological systems;
- Improved methods in sequence reconstruction lets us infer physiological characters of early life, even before LUCA.

# Acknowledgements



N A S A  
ASTROBIOLOGY  
INSTITUTE

## Alm Lab

Dr. Eric Alm

Dr. Francisco Camas

Dr. Manu Tamminen

Dr. Sarah Pacocha Preheim

Sean Clarke

Sarah Spencer

Eli Papa

Mark Smith

Chris Smillie

Allison Perrotta

