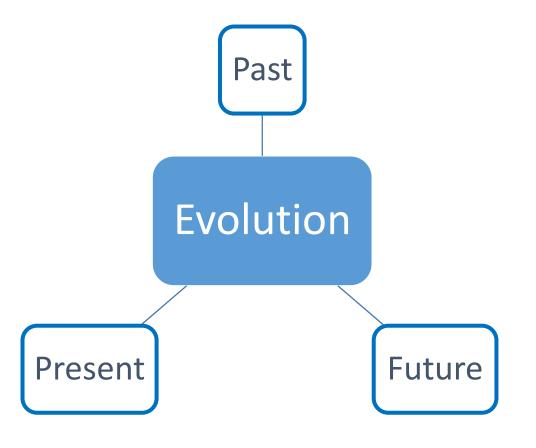


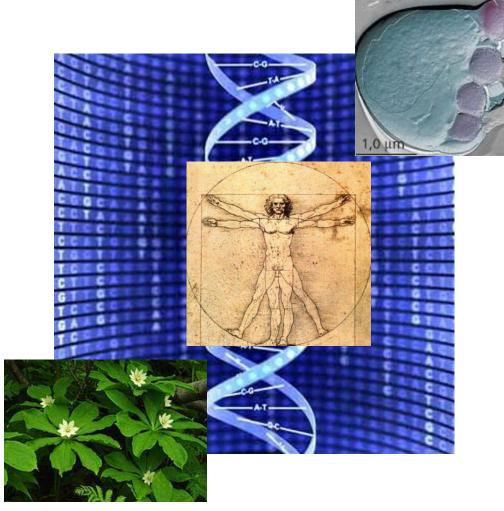
Earth, genomes, and time: a big data approach to integrative evolutionary histories

Fabia U. Battistuzzi
Biological Sciences
battistu@oakland.edu
Dec. 1, 2016

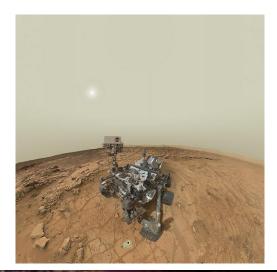
- Past history of life is a predictor of current and future changes
 - Medical field
 - Climate science
 - Astrobiology
 - Conservation biology
 - Sustainable energy
 - •

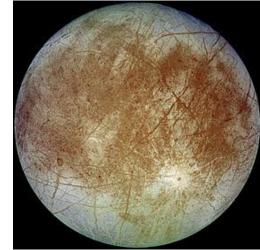


- Genomes are repositories of billions of data points (DNA bases)
 - Human genome: 3 billion DNA base pairs
 - 7 billion individuals on Earth
 - 2.1 e+19 base pairs
 - Species estimates: 10 million to 1 trillion
 - Many will be much smaller than us (< 1 million base pairs)
 - Many are much larger than us (up to ~150 billion base pairs)

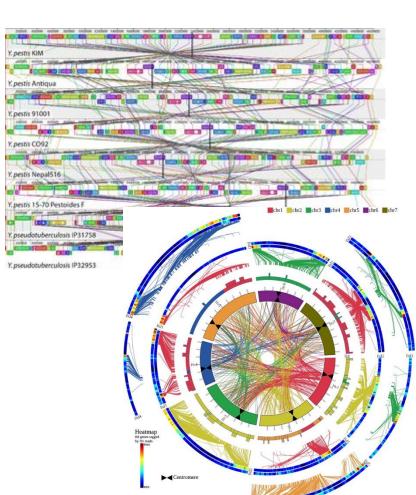


- Comparative genomics
 - How and where did life originate
 - And where should we look for other life (Astrobiology)

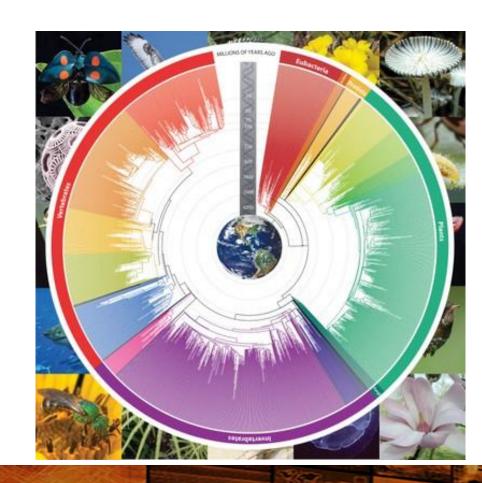








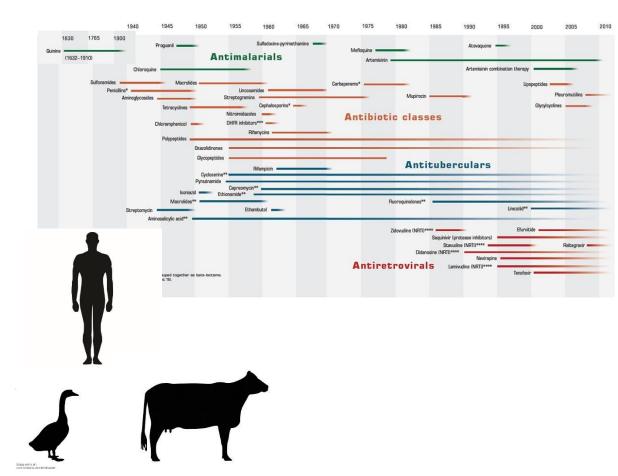
- Comparative genomics
 - How did life survive on Earth through major environmental changes
 - Microbes are the longest living lineages on Earth (~4 billion years)
 - They survived and thrived during planetary-scale climate changes (conservation and sustainability)





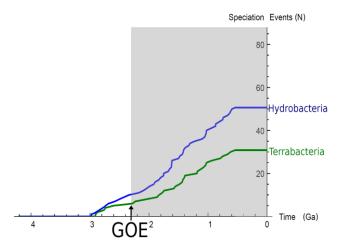
- Comparative genomics
 - How do pathogens escape our immune system and drugs
 - What changes at the genomic level allow them to adapt?
 - How do pathogens arise and spread?

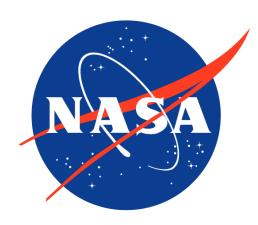


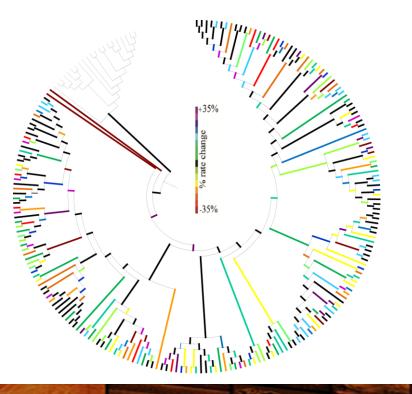


Evolution in the Blab

- Early life evolution
 - How to accurately reconstruct the evolutionary histories of microbes on Earth
 - Conditions for life to thrive
 - Adaptations that sustained microbial life through climate changes
 - Rate of species diversification

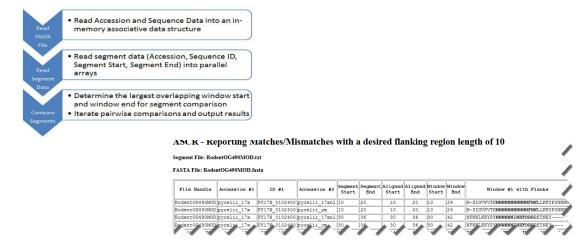




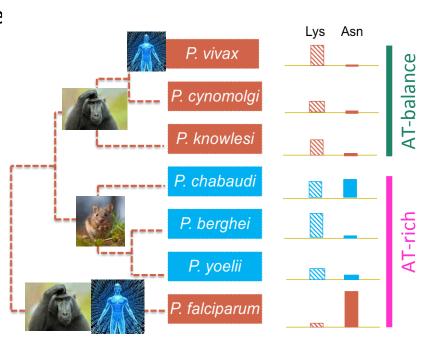


Evolution in the Blab

- Evolution of malaria
 - How does it adapt to humans and other hosts?
 - Are genomes evolving differently depending on the host?
 - Are genes involved in antimalarial drug resistance evolving faster?

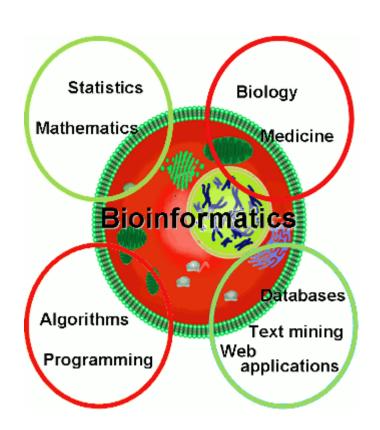






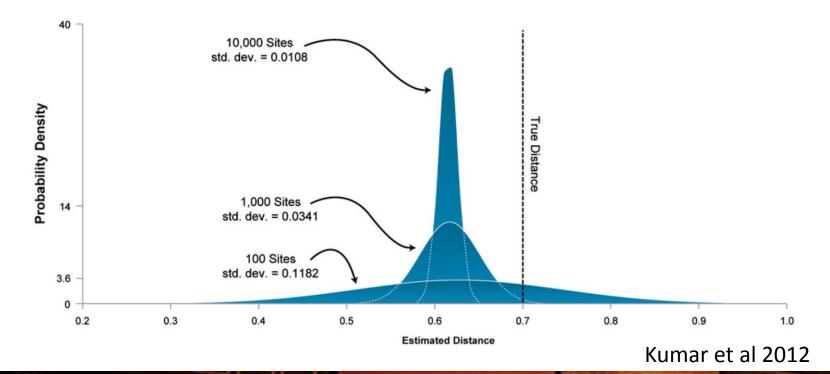
Evolution in the Blab

- Students involvement
 - Undergrads & grads (current size: 11+1 students)
- Skills
 - Scripting/programming
 - Phylogenetics
 - Comparative genomics
 - Data mining
 - Statistics



New opportunities with CDaS

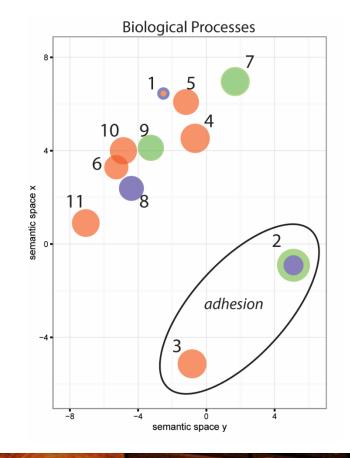
- Explore new statistical applications for Big Data
 - Systematic bias
 - False discovery rates



New opportunities with CDaS

- Connect comparative and functional genomics
 - Text mining of functional databases
 - Integration of multiple databases





New opportunities with CDaS

- Explore new strategies to gain computational support
 - On-site high-performance computational cluster (working on it...)
 - Cloud-based and off-site clusters (exploring options...)
 - Programming, database architecture







Contact info

battistu@oakland.edu

340 Dodge Hall Biological Sciences Oakland University