

# PHYLOGENOMICS

MARIA CHAVEZ AND ANDIE HELD



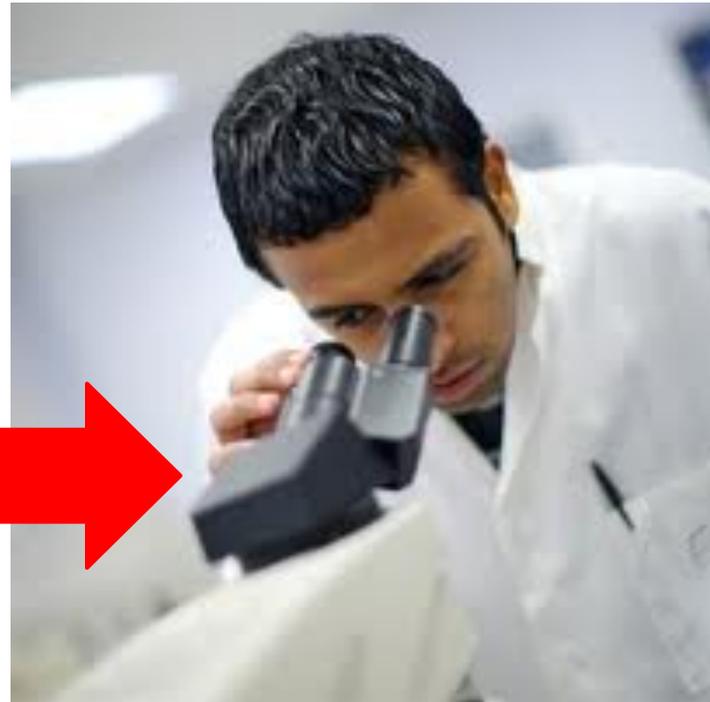
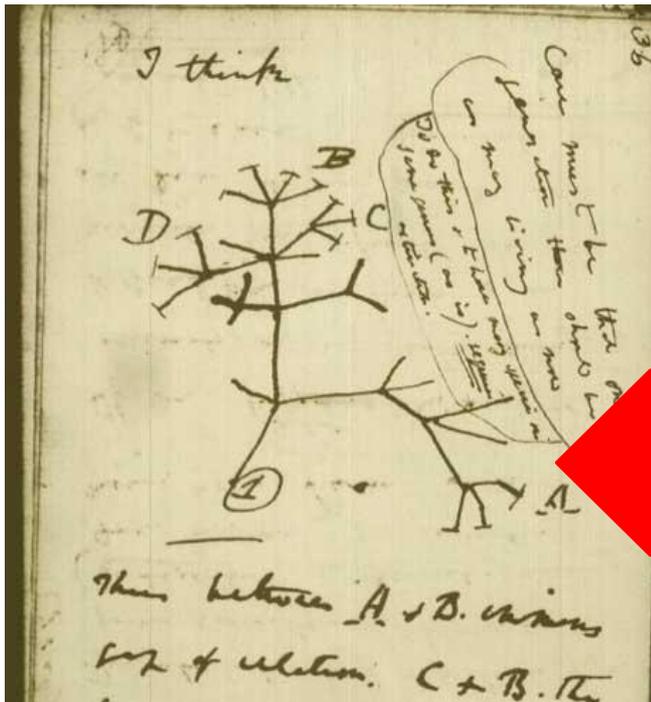
# Evolution + Genomics = Phylogenomics



The construction of evolutionary history

<http://photosofwar.net/rockefeller-center-construction-new-york-new-york-1932/>

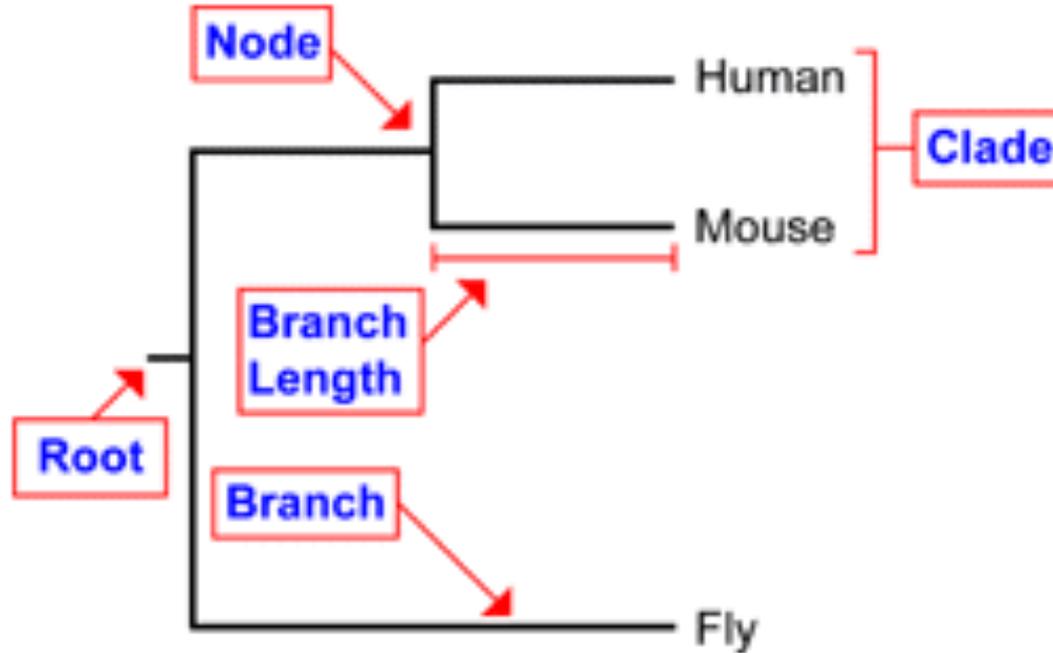
# Why is phylogenomics important?



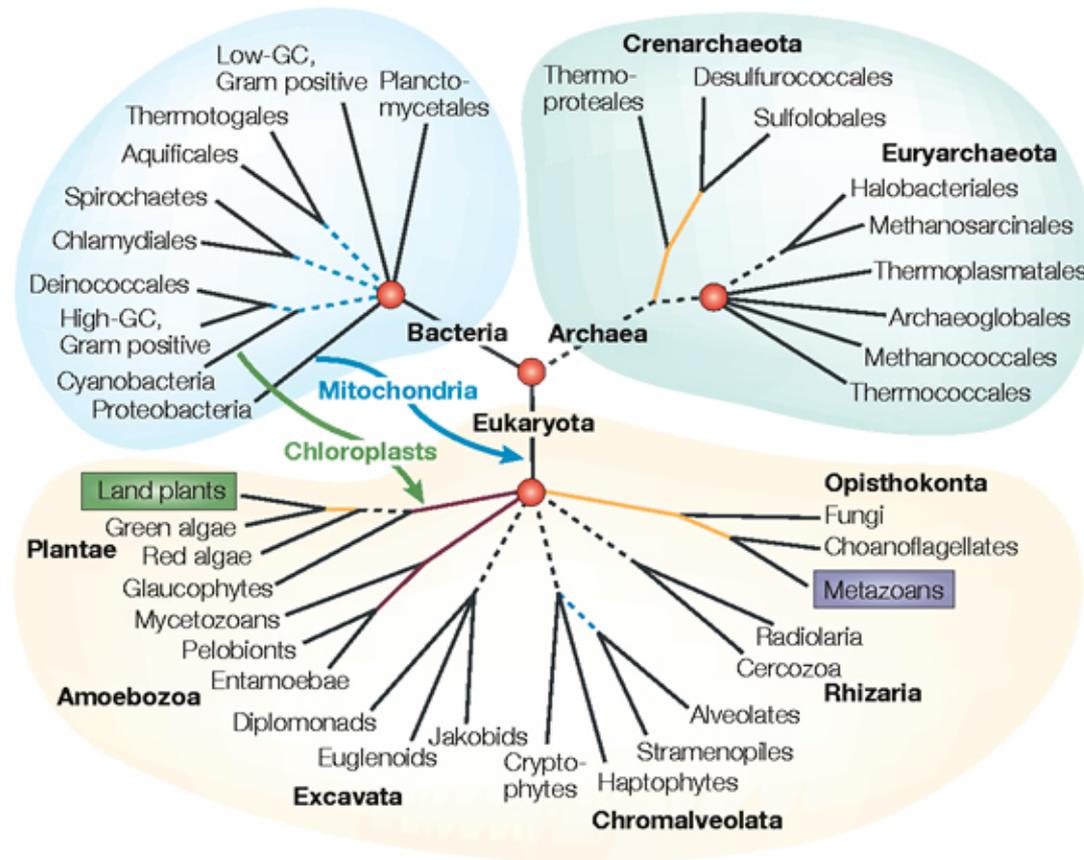
<http://www.theguardian.com/science/2009/jan/21/charles-darwin-evolution-species-tree-life>

<http://mncs.nges-patan.org/subject-botany.aspx>

# Tree Basics



# Tree of Life

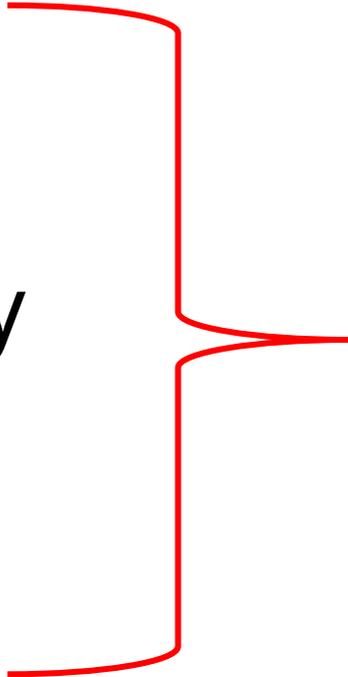


# Reconstruction Methods

1. Distance

2. Maximum Parsimony

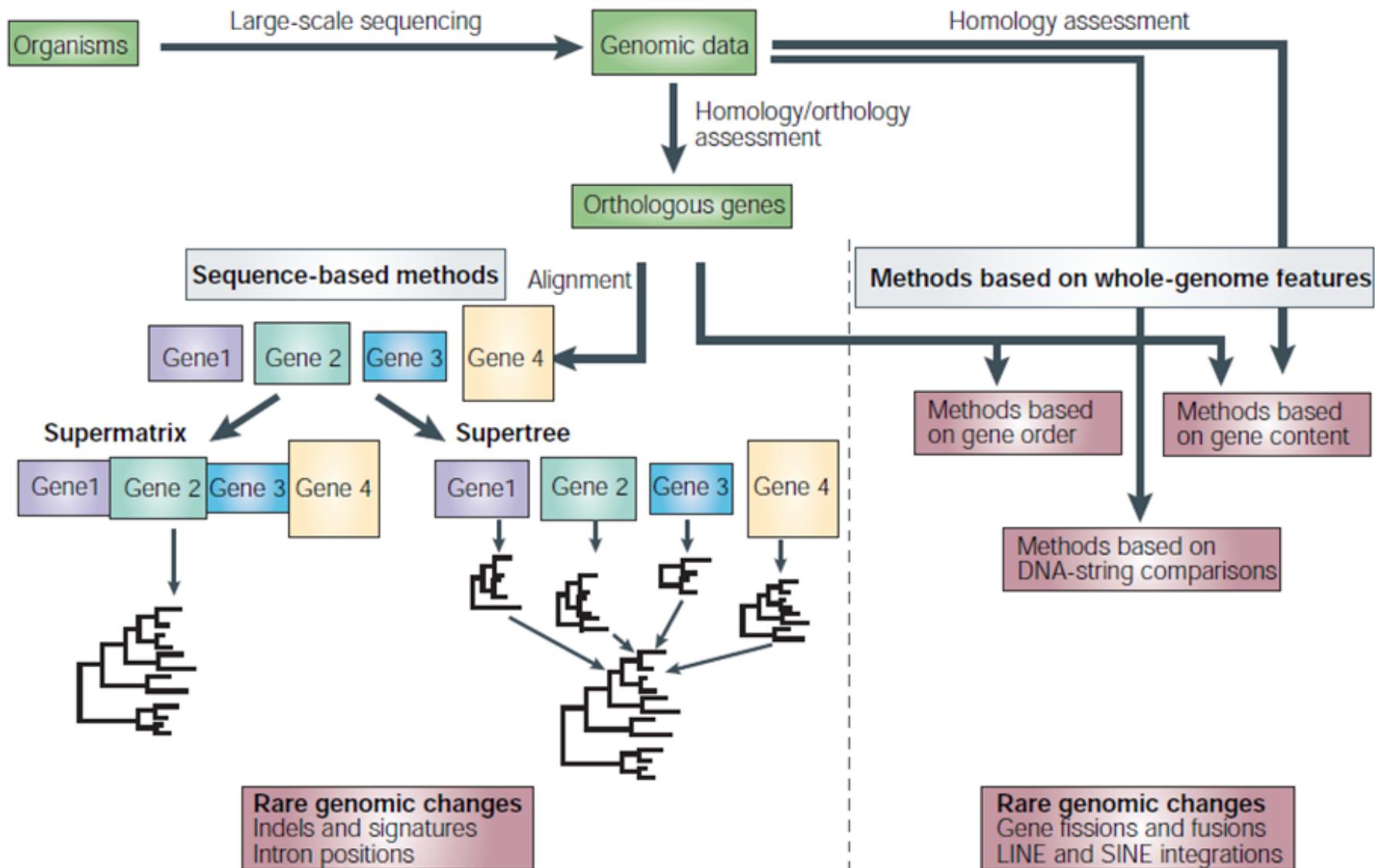
3. Likelihood

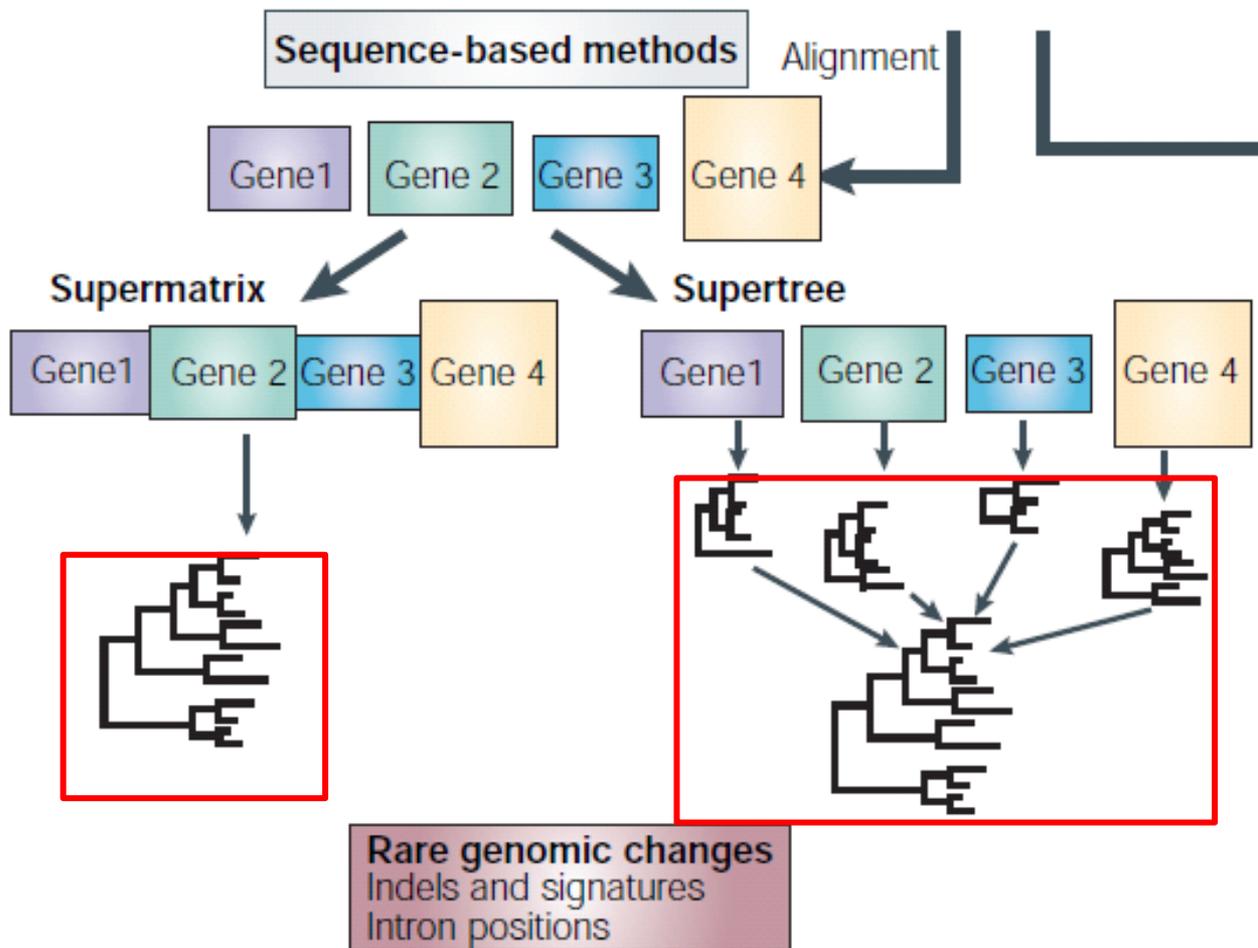


Ways to  
build our  
trees

# How do we do it?

## Box 2 | Methods of phylogenomic inference





# Clustal Omega

[Input form](#) | [Web services](#) | [Help & Documentation](#)

[Tools](#) > [Multiple Sequence Alignment](#) > Clustal Omega

## Multiple Sequence Alignment

Clustal Omega is a new multiple sequence alignment program that uses seeded guide trees and HMM

### STEP 1 - Enter your input sequences

Enter or paste a set of  sequences in any supported format:

Or, upload a file:  No file chosen

### STEP 2 - Set your parameters

OUTPUT FORMAT

DEALIGN INPUT SEQUENCES

MBED-LIKE CLUSTERING GUIDE-TREE

MBE

MAX GUIDE TREE ITERATIONS

MAX HMM ITERATIONS

ORF

### STEP 3 - Submit your job

What we're using  
for our own  
trees!

Sequence-based  
alignment  
program

# Whole Genome approach

## Gene order vs. Gene content vs. DNA string

- Estimates presence/absence of pairs of orthologous genes
- Needs more efficient algorithm
- Uses distance and parsimony methods
- 'Big Genome Attraction' problem
- Based on distribution of oligonucleotides
- Saturation of signal limits approach

# Rare Genomic Changes

## What are they for?

Can be useful for reconstructing phylogenetic trees

Used as “signatures” to support certain nodes

Based on their absence or presence

## Can come in many forms!

Indels

Intron positions

SINES/LINES

Gene fusion and fission events

# Challenges and inconsistencies

**stochastic/sampling errors:** Shorter length=  
more errors

**systematic errors:** depend on data and  
methods



# Phylotype

a similarity in DNA sequence(s) used to classify organisms; a grouping of these similar organisms is also called a phylotype





# A jungle in there

Hulcr, J., et al.

# Why study the belly button?

consistent environment

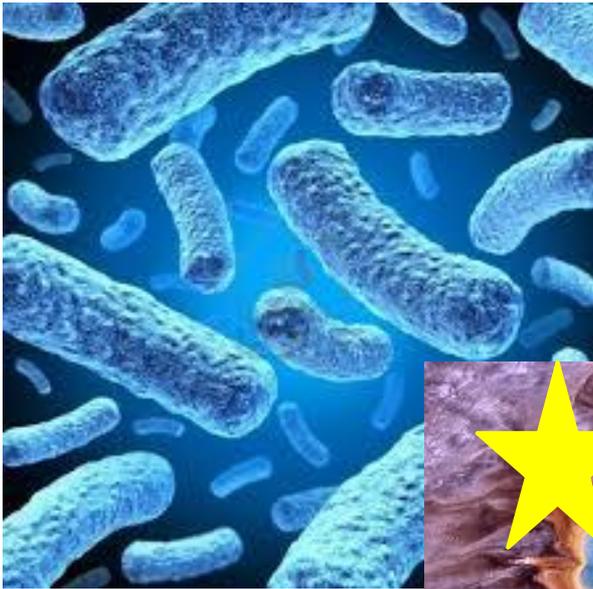
less frequently disturbed  
-no secretions or oils

public interest



[http://www.huffingtonpost.com/2013/09/12/what-your-belly-button-says-about-you\\_n\\_3901527.html](http://www.huffingtonpost.com/2013/09/12/what-your-belly-button-says-about-you_n_3901527.html)

# What lives in the belly button?



photos from [http://www.huffingtonpost.com/2013/09/12/what-your-belly-button-says-about-you\\_n\\_3901527.html](http://www.huffingtonpost.com/2013/09/12/what-your-belly-button-says-about-you_n_3901527.html); <http://www.123rf.com/stock-photo/bacteria.html>; <http://en.wikipedia.org/wiki/Archaea>

# Why is this important?

public outreach and  
science education

personal hygiene



# How was the research conducted?

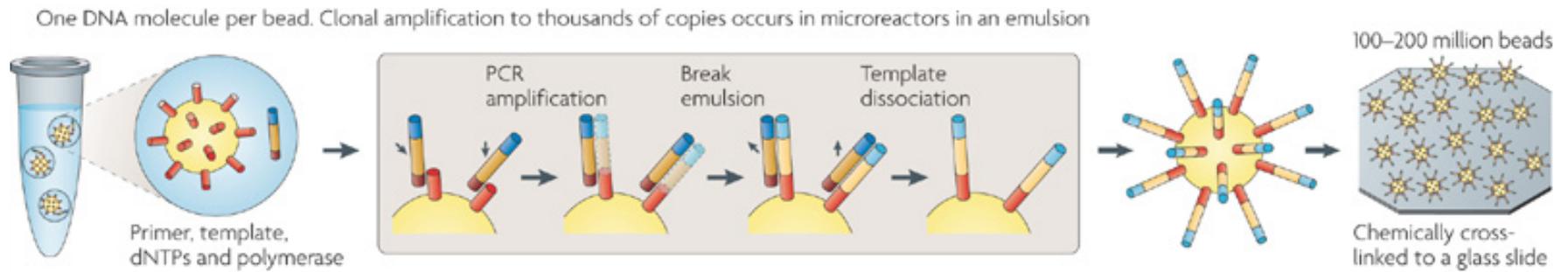


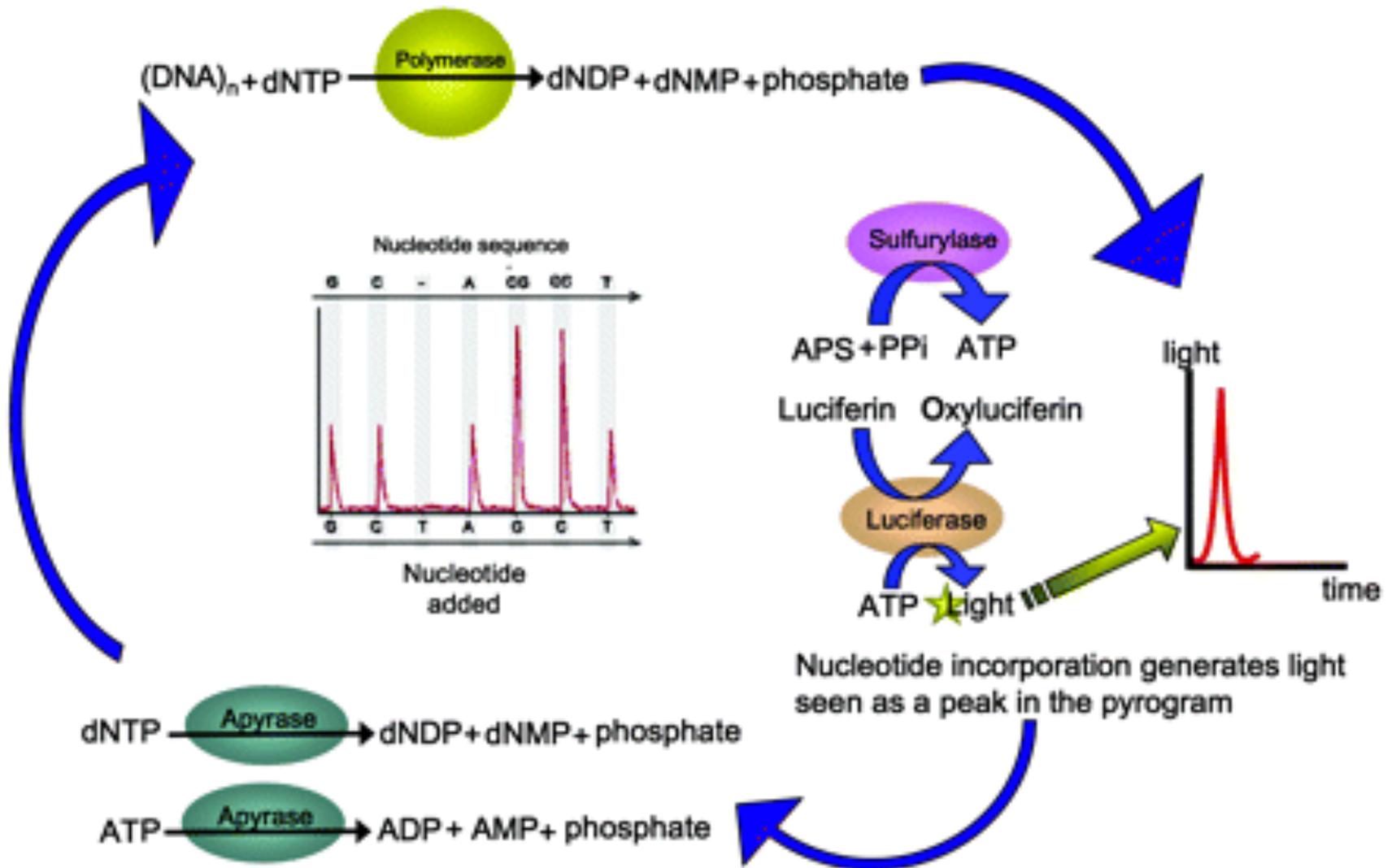


# Methods Outline

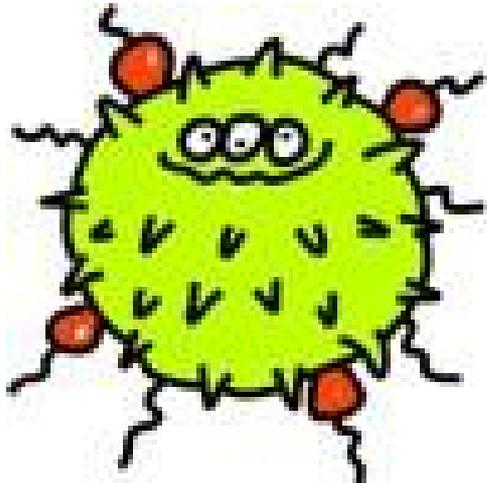
1. Roche 454 Pyrosequencing
2. PyNAST

# Remember this?!





1. Sequences  $\geq 97\%$  were grouped into Operational Taxonomic Units (OTU)
2. OTUs were aligned using PyNAST
  - a. uses 16s rDNA and BLAST



<http://www.mybiosoftware.com/alignment/11775>

# WHY?

# Vocab you should know

**frequent**= occurring on many humans

**abundant**= found many times in the belly button of the humans on which they are found

frequent + abundant = .....

# Oligarch



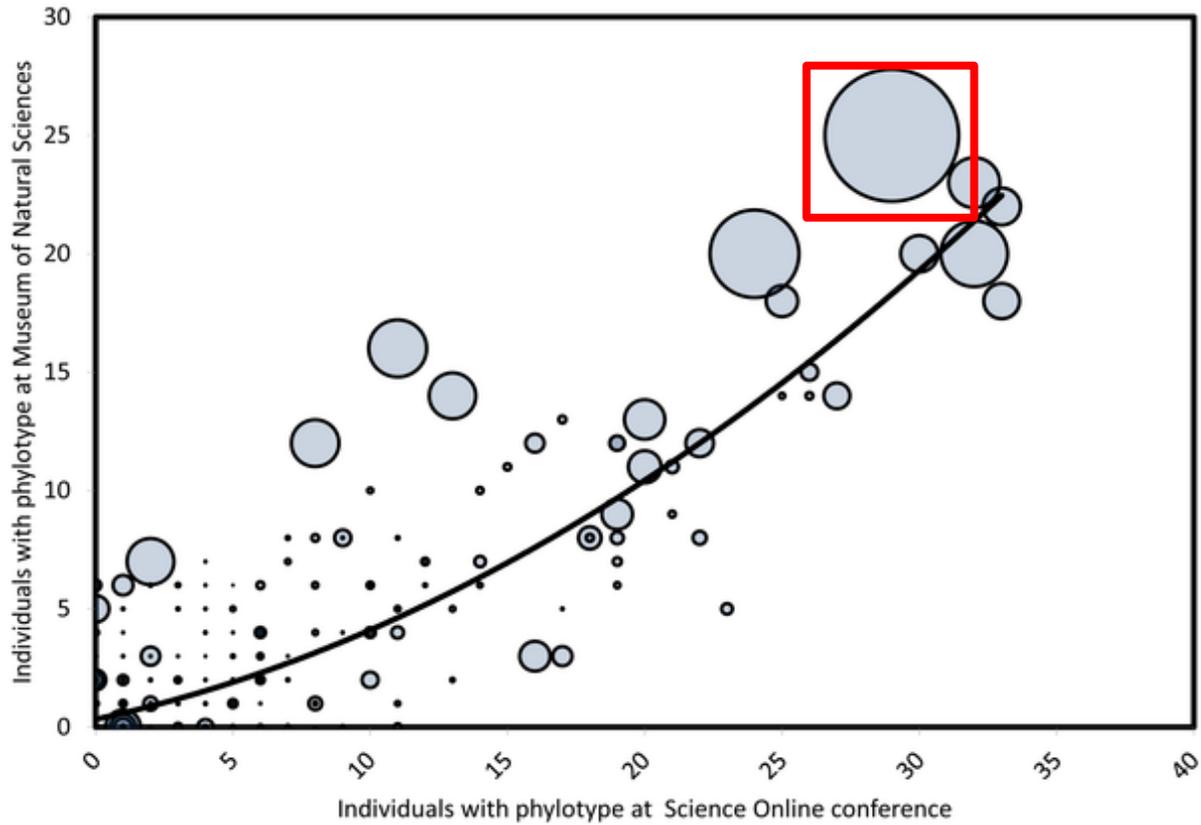
predictably frequent among individuals and often abundant

# Research Question

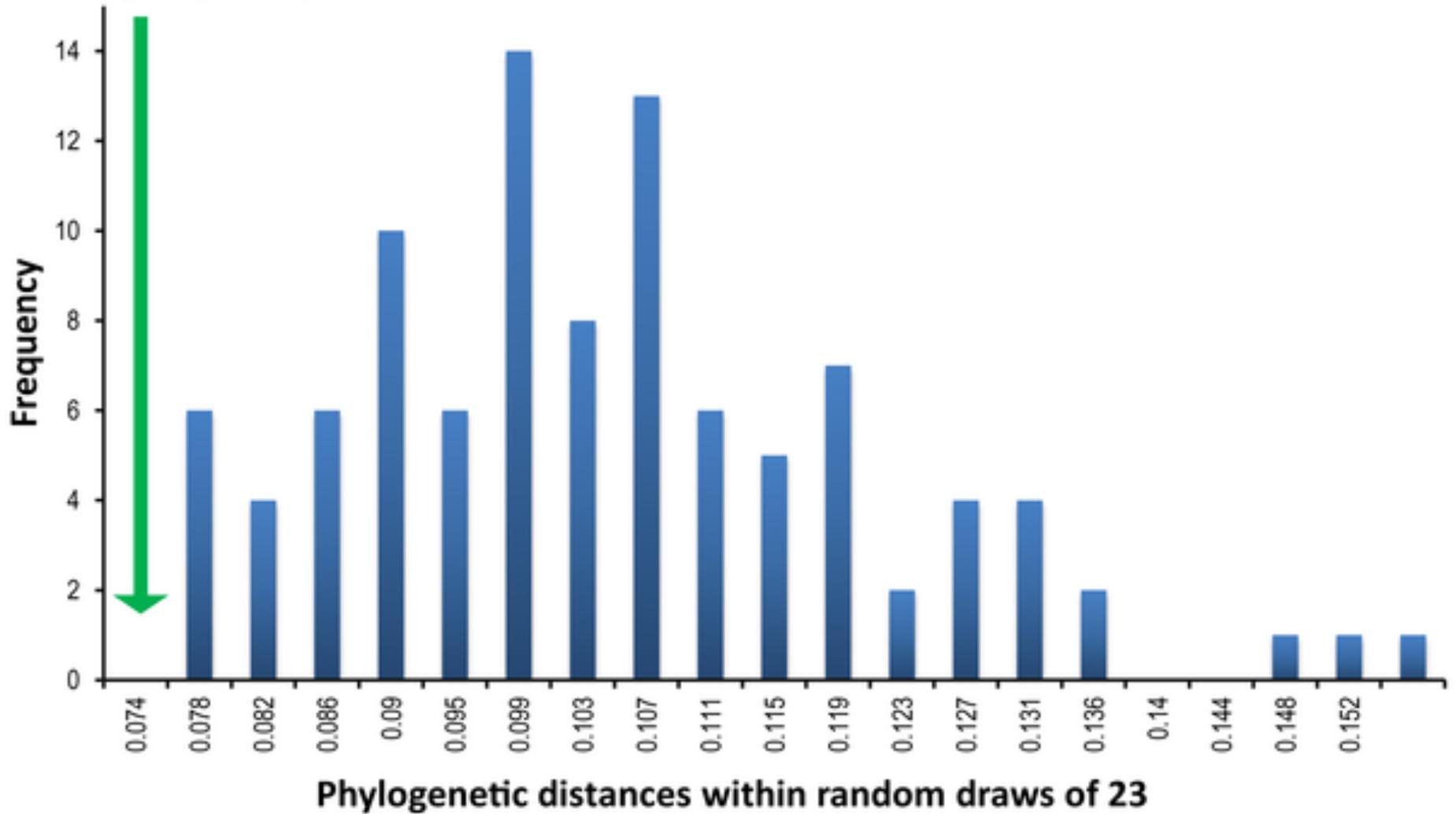
Are the oligarchs predictable? Are they more closely related than other phylotypes selected at random?

|   | <b>phylotypes</b> | <b>% human samples</b> |
|---|-------------------|------------------------|
| phylotypes present on<br><50% people      | 2188              | 1–10                   |
|   | 97                | 11–20                  |
|   | 31                | 21–30                  |
|   | 12                | 31–40                  |
|   | 17                | 41–50                  |
| 23 “oligarchs”, present on<br>>50% people | 11                | 51–60                  |
|   | 4                 | 61–70                  |
|   | 2                 | 71–80                  |
|   | 2                 | 81–90                  |
|   | 4                 | 91–100                 |

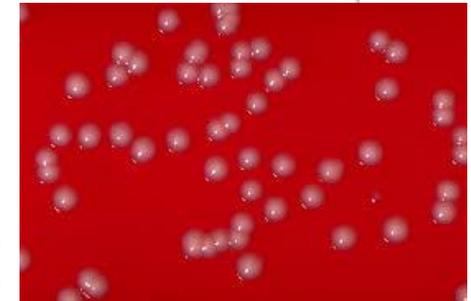
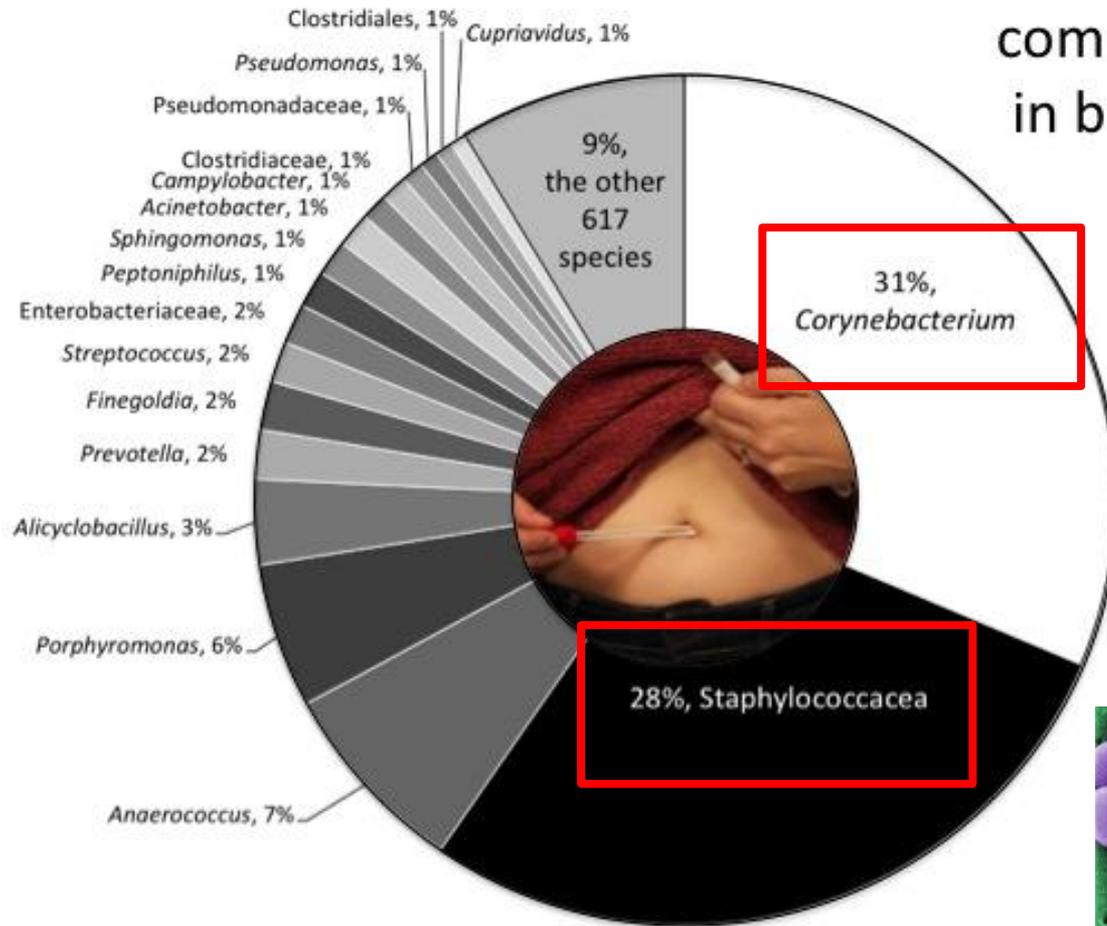
doi:10.1371/journal.pone.0047712.t001



Mean phylogenetic distance among "oligarchs", 0.070



# The most common species in belly buttons





**Micrococcus  
(Actinobacteria)**



**Clostridiales**



**Staphylococcus**



**Bacillus**

# Additional findings



Belly buttons vary in diversity

More diverse, the more rare and infrequent

# Conclusions

frequent and abundant phylotypes are phylogenetically related and similar across populations

belly button taxonomic compositions are predictable due to oligarchs

<http://www.wildlifeofyourbody.org/>

# Future Implications

science outreach

end the war against bacteria

personal hygiene

skin immune function

future studies



# Discussion Questions

Why do you think different people's belly buttons contained different bacteria?

Do you think sampling people in different states or countries would produce different results?

# Questions for us?

