

# Reticulate evolutionary strategies are favoured when pathogens cross host species barriers.

Eric J. Ma, Nichola J. Hill, Kyle Yuan, Justin Zabilansky, Jonathan A. Runstadler  
Department of Biological Engineering & Division of Comparative Medicine, MIT

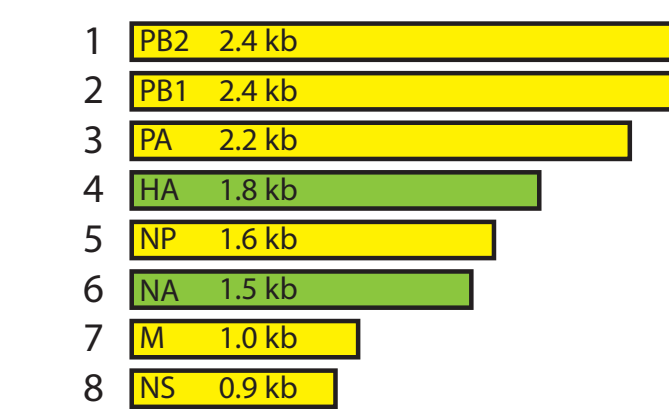
## Introduction

### Research Questions

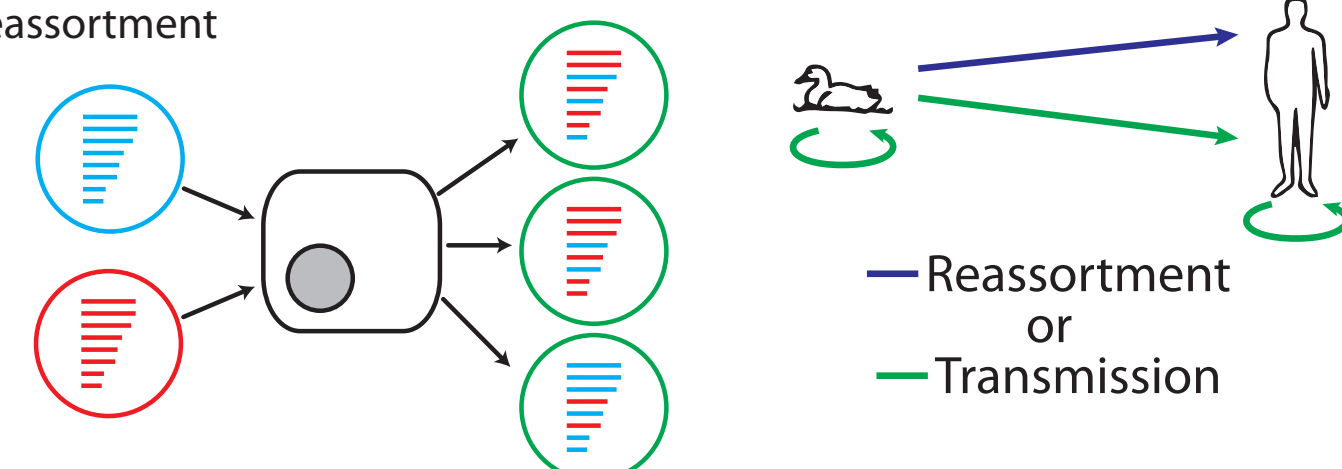
- How important is reticulate evolution as a process for pathogen evolution?
- What is the role of reticulate evolution in pathogen ecology?

### Flu: A Model Pathogen

Influenza Genome Structure



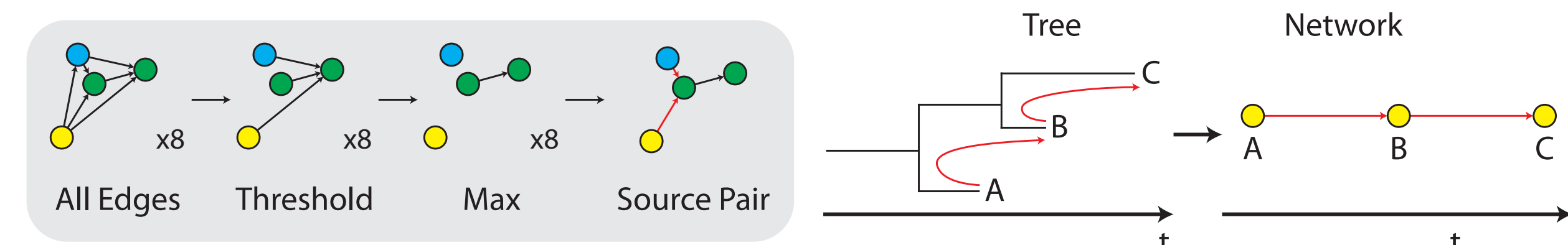
Reassortment



- Segmented genome; can reassort (this is influenza's reticulate evolution mechanism)
- Multiple subtypes, competition for hosts, but also cooperation via gene sharing.
- Abundant sequence data with matched metadata, densely sampled.

## Reassortment Detection

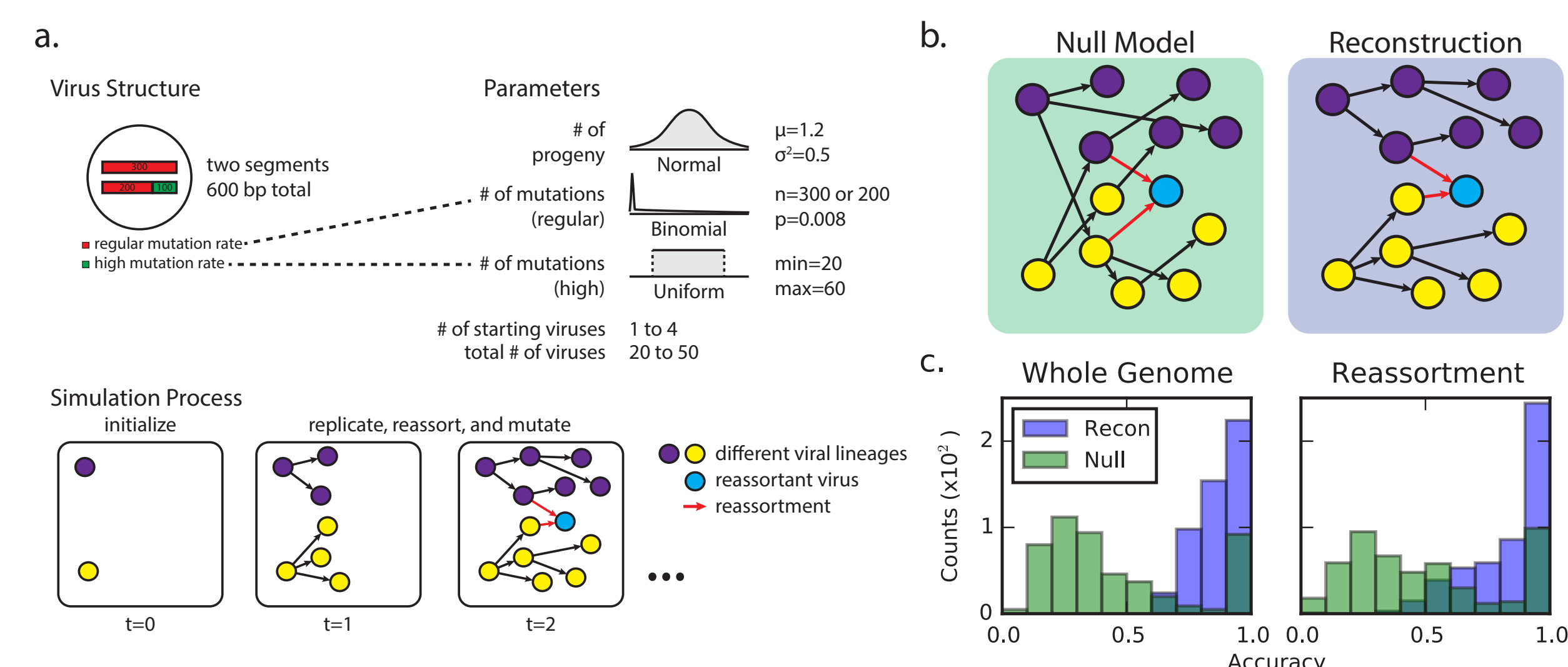
### Algorithm



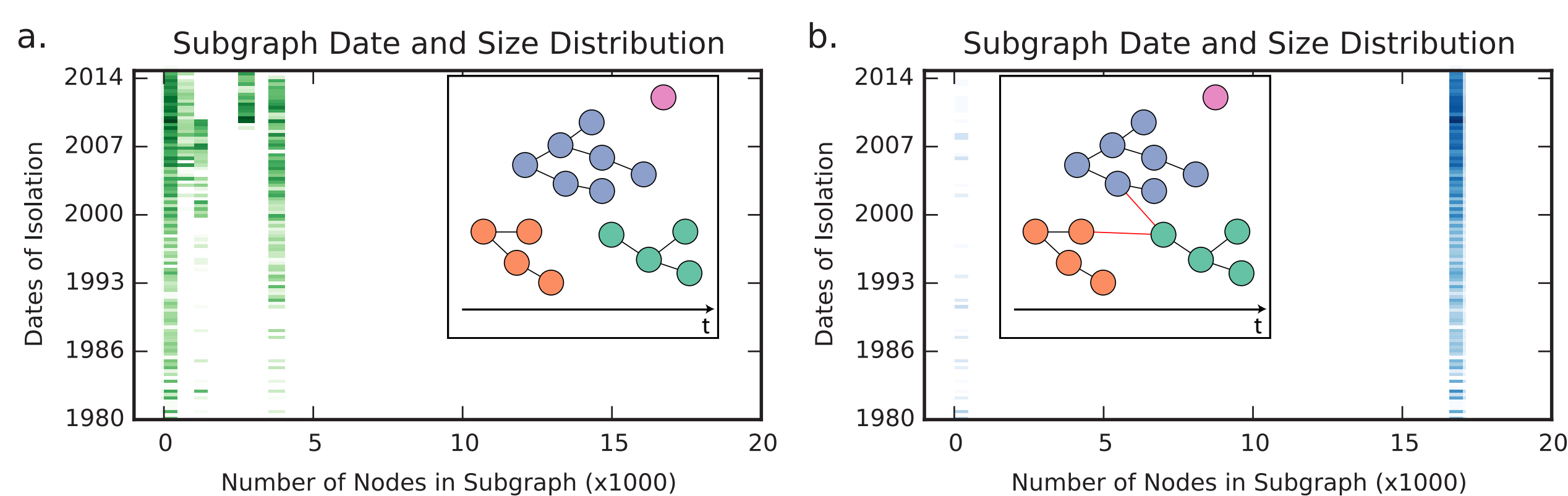
- Phylogenetic heuristic: search for sources of genetic material.
- Maximize genetic similarity, while minimizing number of sources.
- Heuristic method is akin to "flattening" a phylogenetic tree.

### Performance

Accurate reconstruction in simulation studies.



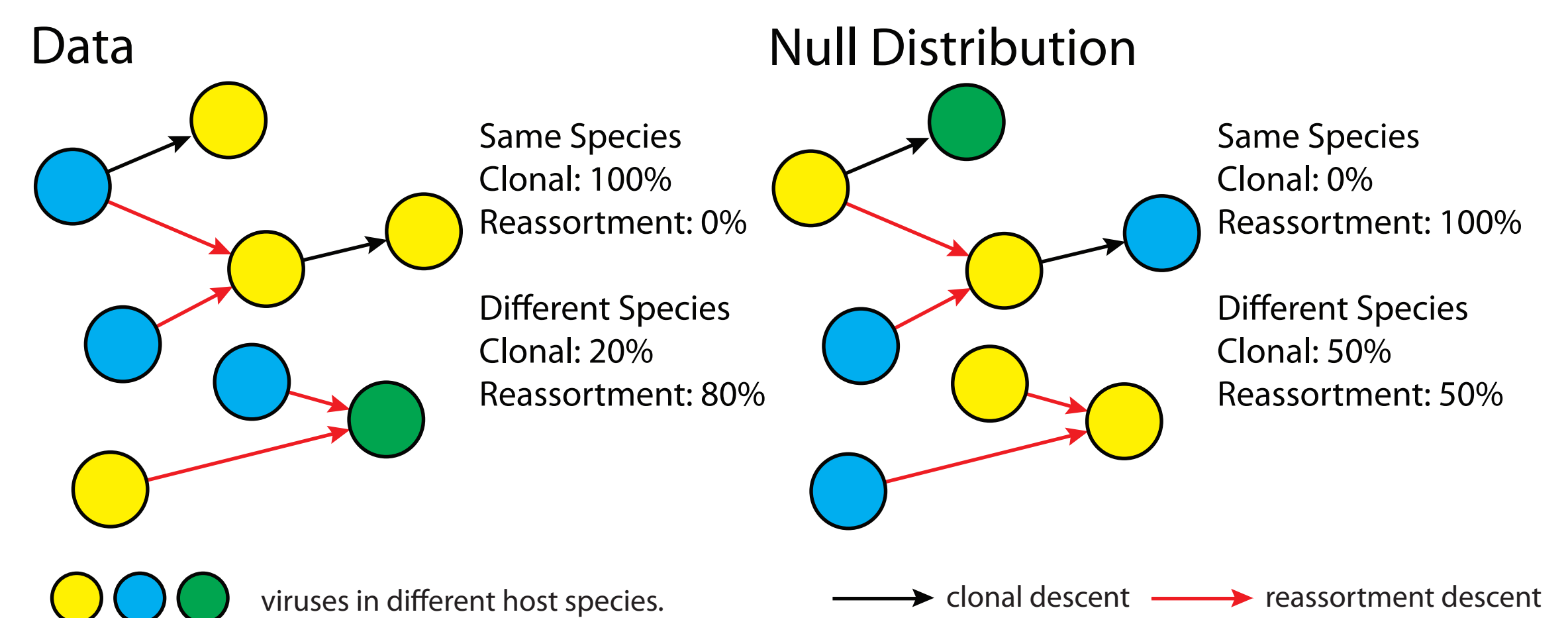
Network captures known viral circulation and reassortants.



- In network (n=18632 viral isolates), known circulation of human and swine viruses captured in the network.
- "Famous" reassortants identified - pandemic H1N1 (2009), H7N9 (2013).
- Reassortment connects viral subtypes together in a global network of gene exchange.

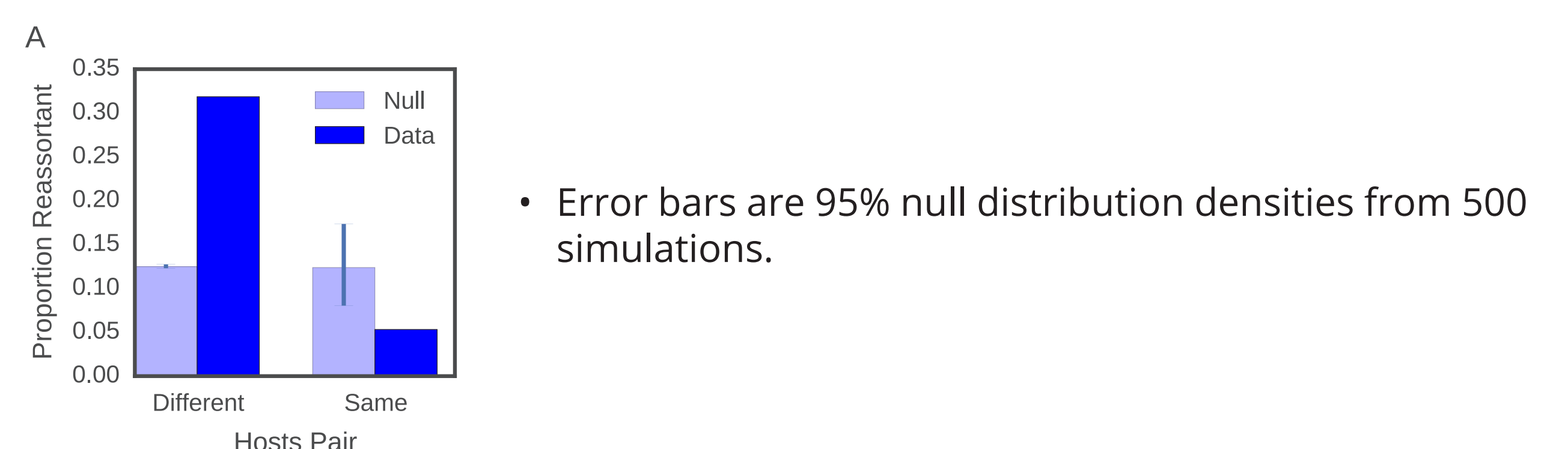
## Reassortment Importance

### Network Statistical Test

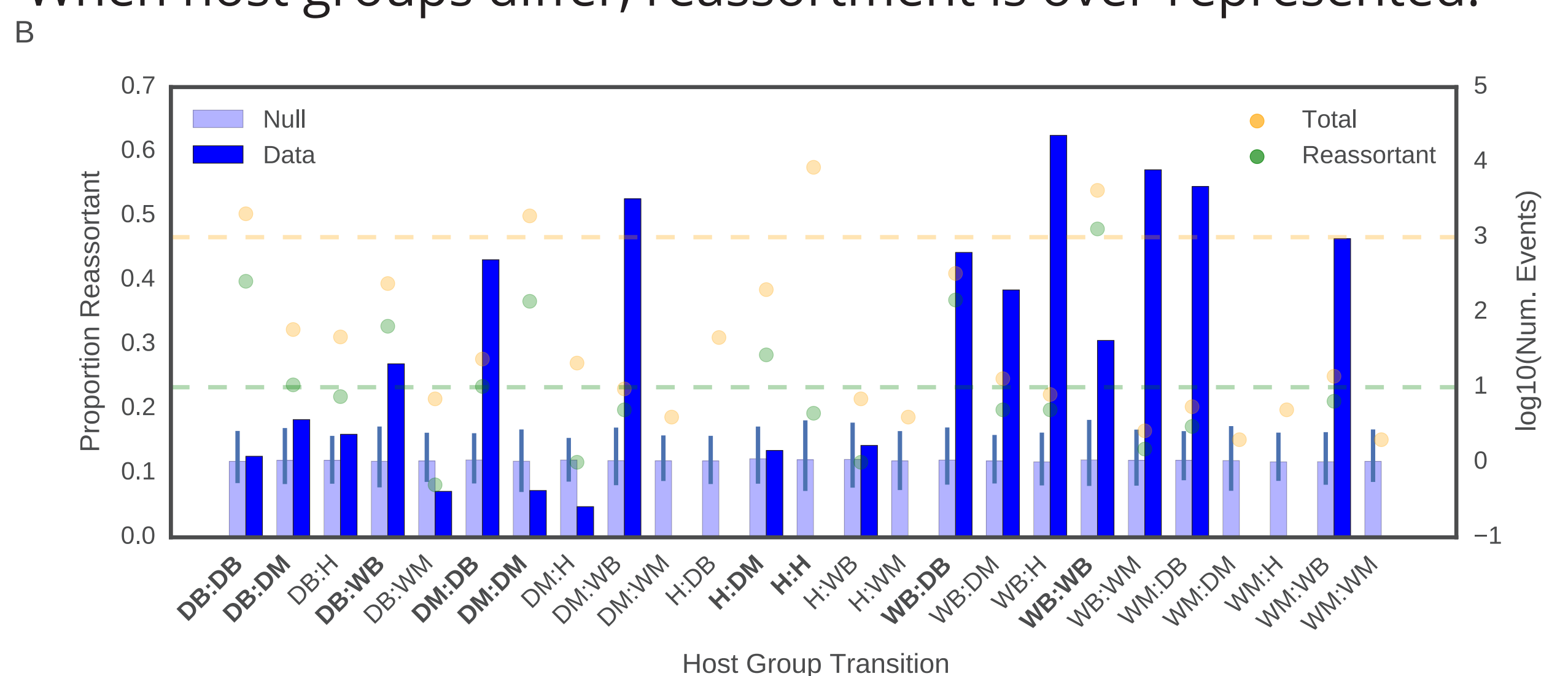


### Results

When hosts are different, reassortment is over-represented.

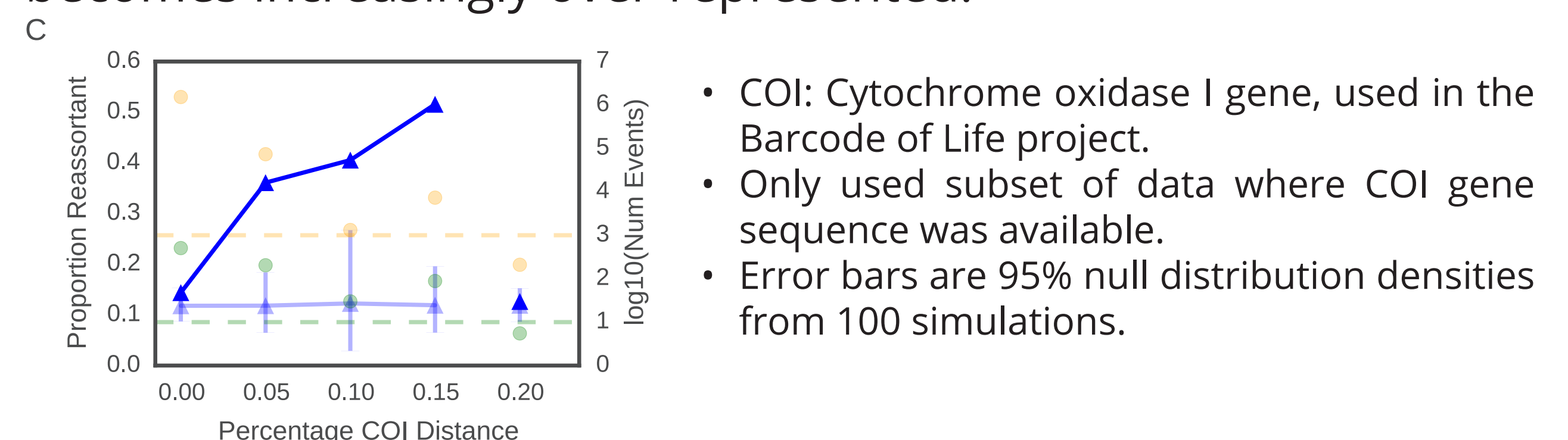


When host groups differ, reassortment is over-represented.



- D & W refer to "Domestic" and "Wild". B, M and H refer to "Bird", "Mammal" and "Human". Example: DB = Domestic Bird, WM = Wild Mammal.
- Dotted lines represent threshold number of reassortants for calculating proportion of reassortment, i.e. either dot has to be above same-colored line.
- Error bars are 95% null distribution densities from 500 simulations.

As hosts are increasingly evolutionarily distant, reassortment becomes increasingly over-represented.



- COI: Cytochrome oxidase I gene, used in the Barcode of Life project.
- Only used subset of data where COI gene sequence was available.
- Error bars are 95% null distribution densities from 100 simulations.

## Conclusions

- Network phylogenetic heuristic accurately captures known clonal and reassortment transitions between viral hosts.
- The greater the difference between ecological niches, the greater the importance of reticulate evolution in enabling niche switches.