1. Introduction
We are interested in the drivers of parasitism in primates...

2. Data
Global Primate Parasite Database (GPPD)

3. Network and Prediction
Phylogeny-based prediction methods identified undocumented host-parasite interactions.

4. PGLS
Coefficients in PGLS Models with Citation as Weight and Covariate Predicting Total Count and Total Imputed Count

Fig. 3 Degree centrality before and after adding predicted edges.

Fig. 4 Models predicting parasite counts show certain predictors have larger influences on count.
Data

Global Primate Parasite Database (GPPD)

A compilation of parasites+pathogens from wild primate hosts

Phylogeny

Some species are studied much more than other species
Phylogenetic Generalized Least Squares (PGLS) models

- A consensus chronogram was obtained from 10K tree, and used in phylogenetic regression
- log(Citation) used as weights and included as covariate to control for sampling efforts
- Possible models were ranked by AIC values, and coefficients were obtained through model averaging

**Important Predictors of Parasitism**

- Diel Activity
- Locomotion
- Mating System
- Realm
Predicting Undocumented Host-Parasite Interactions

- Evolutionary relationships encoded as weighted network.
- Top edges were ranked by relative probabilities of existing \( p \).
- Citation count variable decreased in importance (eased sampling effort problem)
- 5-Fold Cross Validation to test predictive accuracy.
**Proportion of Data**

**Degree Centrality**: the number of edges connected to a node. i.e. the number of parasites connected to a host.