

# GMPD IDEAS Capstone

## How do host and parasite traits predict parasite species richness across host threat levels?

Overall, we predict parasite species richness will decrease when hosts are classified as more threatened.

However, we predict this general pattern might not be the same across all host and parasite traits.

*Host* traits we are interested in studying include:

- host threat status (either binary or grouped IUCN levels)
- population density change
- group size

*Parasite* traits we're interested in include:

- type (bacteria, virus, protozoa, etc.)
- transmission mode (close, non-close, vector)
- host breadth (mpd)

## Data

*Host-parasite pairs:*

- Records of parasitic organisms reported from free-living mammals were obtained from GMPD.

```
##      hostName      dietBreadth      habitatBreadth      maxLongevity
## Length:19587      Min.      :1.000      Min.      :1.000      Min.      : 60.0
## Class :character  1st Qu.:1.000      1st Qu.:1.000      1st Qu.:204.0
## Mode  :character  Median :2.000      Median :1.000      Median :262.0
##                               Mean  :2.891      Mean  :1.197      Mean  :293.2
##                               3rd Qu.:5.000      3rd Qu.:1.000      3rd Qu.:354.0
##                               Max.   :7.000      Max.   :3.000      Max.   :744.0
##                               NA's   :2409      NA's   :3304      NA's   :2254
##      popGrpSize.x      socGrpSize.x      popDenChange      popGrpSize.y
## Min.      :      1.0      Min.      : 1.00      Min.      : -999.000      Min.      :      1.0
## 1st Qu.:      5.5      1st Qu.:  2.25      1st Qu.:  0.040      1st Qu.:      5.5
## Median :     30.0      Median : 12.00      Median :   0.060      Median :     30.0
## Mean   :    8029.4      Mean   : 14.37      Mean   :  -4.293      Mean   :    8029.4
## 3rd Qu.:    500.0      3rd Qu.: 17.50      3rd Qu.:  0.110      3rd Qu.:    500.0
## Max.    :100000.0      Max.    :110.00      Max.    :   0.440      Max.    :100000.0
## NA's     :16457      NA's     :11752      NA's     :2891      NA's     :16457
##      socGrpSize.y      comGrpSize      Order.1.2
## Min.      : 1.00      Min.      : 1.00      Length:19587
## 1st Qu.:  2.25      1st Qu.:  2.25      Class :character
## Median : 12.00      Median : 13.00      Mode  :character
## Mean   : 14.37      Mean   : 2799.08
## 3rd Qu.: 17.50      3rd Qu.: 22.60
## Max.    :110.00      Max.    :100000.00
## NA's     :11752      NA's     :10744
##      Family.1.2      Genus.1.2      IUCN.Status.1.2
## Length:19587      Length:19587      Length:19587
## Class :character      Class :character      Class :character
```

```

## Mode :character Mode :character Mode :character
##
##
##
##
##      Mass.g           Group           parasiteName
## Min.      : 42.5      Length:19587      Length:19587
## 1st Qu.: 6000.0      Class :character      Class :character
## Median : 27939.1      Mode  :character      Mode  :character
## Mean    : 90973.8
## 3rd Qu.: 77874.4
## Max.    :2949986.3
## NA's    :789
##      ParType           ParPhylum           Prevalence           HostsSampled
## Length:19587      Length:19587      Min.    : 0.000      Min.    : 0.0
## Class :character      Class :character      1st Qu.: 0.100      1st Qu.: 9.0
## Mode  :character      Mode  :character      Median : 0.310      Median : 27.0
##                                     Mean    : 0.433      Mean    : 114.7
##                                     3rd Qu.: 0.714      3rd Qu.: 74.0
##                                     Max.    :100.000      Max.    :62650.0
##                                     NA's    :3164      NA's    :2683
##      SamplingType      HostEnvironment      Citation      close
## Length:19587      Length:19587      Length:19587      Min.    :0.0000
## Class :character      Class :character      Class :character      1st Qu.:0.0000
## Mode  :character      Mode  :character      Mode  :character      Median :0.0000
##                                     Mean    :0.3756
##                                     3rd Qu.:1.0000
##                                     Max.    :1.0000
##                                     NA's    :1108
##      nonclose           vector           intermediate           ParasiteTraitsCitation
## Min.    :0.0000      Min.    :0.0000      Min.    :0.0000      Length:19587
## 1st Qu.:0.0000      1st Qu.:0.0000      1st Qu.:0.0000      Class :character
## Median :1.0000      Median :0.0000      Median :0.0000      Mode  :character
## Mean    :0.5603      Mean    :0.1583      Mean    :0.2532
## 3rd Qu.:1.0000      3rd Qu.:0.0000      3rd Qu.:1.0000
## Max.    :1.0000      Max.    :1.0000      Max.    :1.0000
## NA's    :1108      NA's    :1108      NA's    :1108
##      parRich
## Min.    : 1.0
## 1st Qu.: 64.0
## Median : 155.0
## Mean    : 362.8
## 3rd Qu.: 531.0
## Max.    :1623.0
##

```

Group	n
carnivores	152
primates	182
ungulates	116
Total	450

This left us with 19587 rows of data and 448.

*Parasite traits:* Host specificity (whether present in one or more hosts), transmission mode (close, non-close, VBD), and type (virus, protozoan, fungus, arthropod, helminth, or bacterium) were taken from GMPD. Host breadth (mpd across host spp) was taken from Park et al. (2018?) but TBD.

Table 2: Distribution of parasite type in dataset.

Type	n
Arthropod	412
Bacteria	235
Fungus	32
Helminth	931
Prion	1
Protozoa	256
Virus	164

Table 3: Distribution of transmission modes of parasites in dataset.  
Some parasites may have multiple modes of transmission.

close	nonClose	vector	intermediate
472	841	302	443

We have 2031 unique parasites.

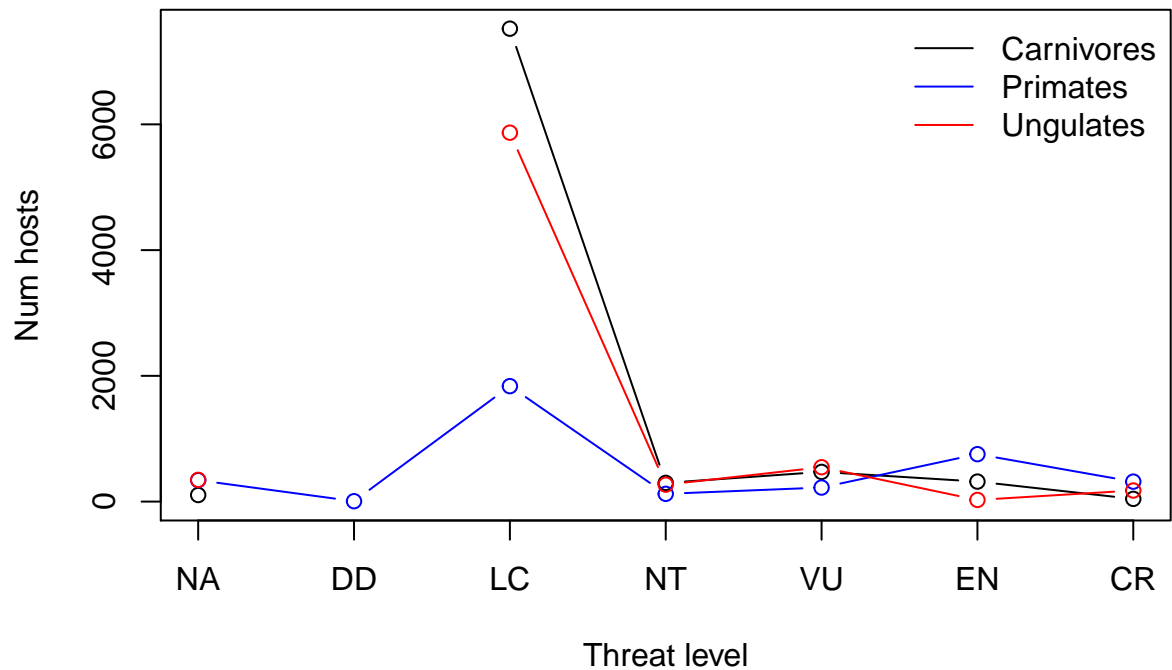
*Host traits:* Host traits were obtained from Pantheria and Phylacine.

- Data were joined by host species names and filtered by hosts with records in GMPD.
- Threat levels:

Extinct (EX) – beyond reasonable doubt that the species is no longer extant. Extinct in the wild (EW) – survives only in captivity, cultivation and/or outside native range, as presumed after exhaustive surveys. Critically endangered (CR) – in a particularly and extremely critical state. Endangered (EN) – very high risk of extinction in the wild, meets any of criteria A to E for Endangered. Vulnerable (VU) – meets one of the 5 red list criteria and thus considered to be at high risk of unnatural (human-caused) extinction without further human intervention. Near threatened (NT) – close to being at high risk of extinction in the near future. Least concern (LC) – unlikely to become extinct in the near future. Data deficient (DD) Not evaluated (NE)

Table 4: Distribution of threat levels from phylacine database shows there are many hosts without IUCN levels characterized.

Group		DD	LC	NT	VU	EN	CR
carnivores	104	NA	7522	297	472	318	43
primates	340	6	1836	123	223	754	317
ungulates	345	NA	5868	269	546	26	176

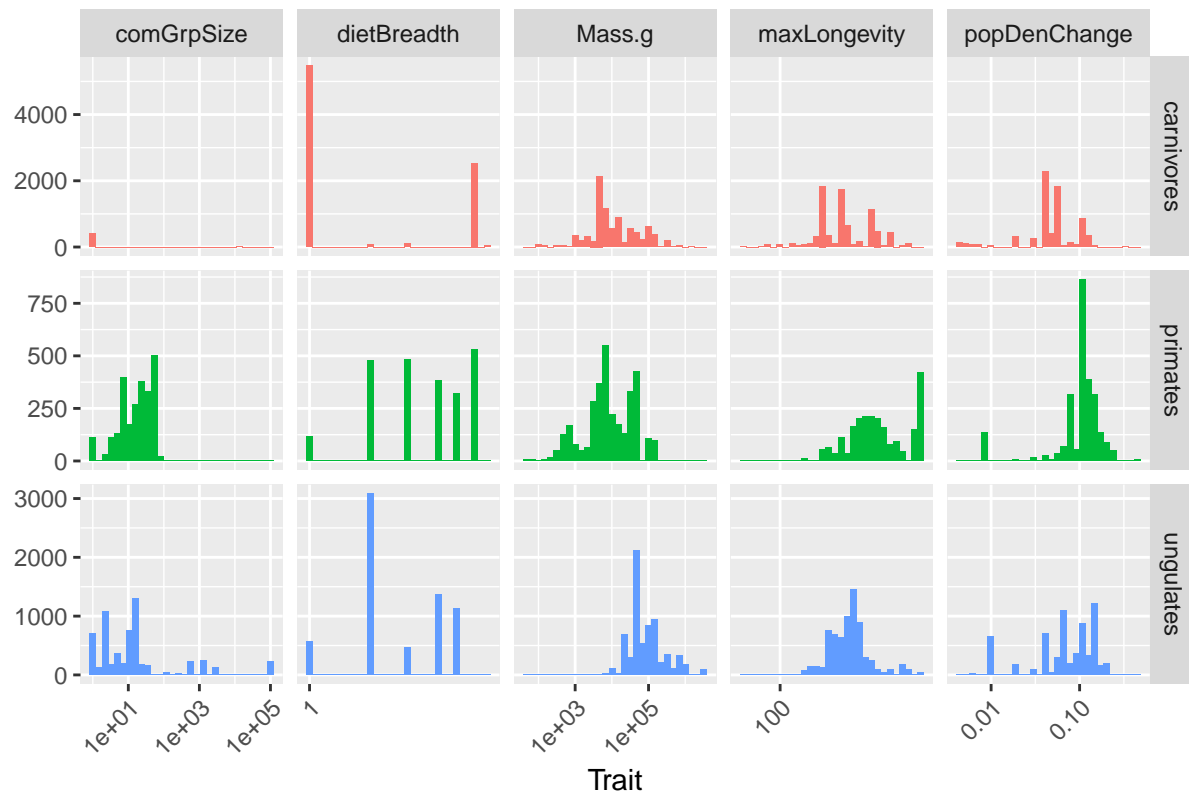


## Warning in self\$trans\$transform(x): NaNs produced

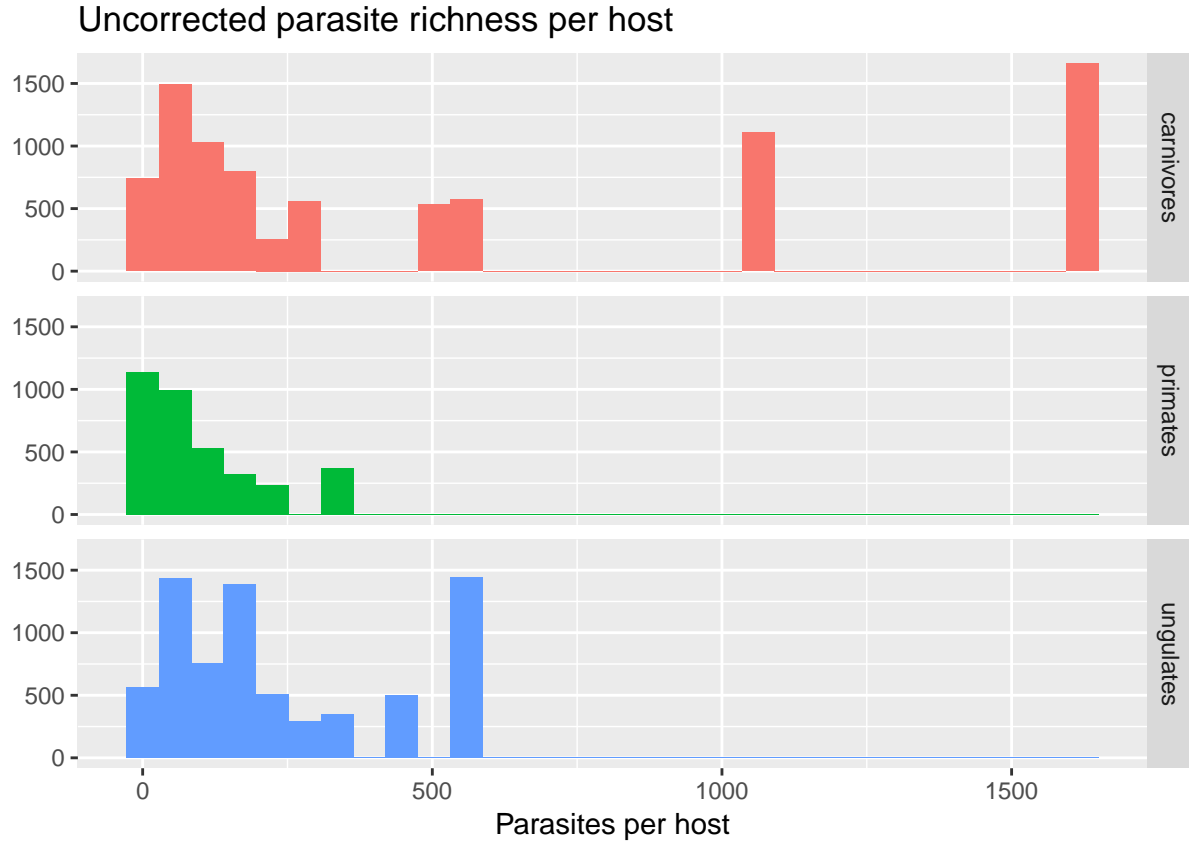
## Warning: Transformation introduced infinite values in continuous x-axis

## Warning: Removed 19802 rows containing non-finite values (stat\_bin).

### Host traits



- Parasite richness: calculated by summing number of unique parasites per host



## Warning: package 'bindrcpp' was built under R version 3.4.4

Table 5: Top 3 host species per mammal group by uncorrected parasite richness.

hostName	Group	parRich
Canis_lupus	carnivores	557
Procyon_lotor	carnivores	1075
Vulpes_vulpes	carnivores	1623
Papio_anubis	primates	163
no_binomial_name	primates	206
Pan_troglodytes	primates	336
Aepyceros_melampus	ungulates	437
Cervus_elaphus	ungulates	542
Odocoileus_virginianus	ungulates	576

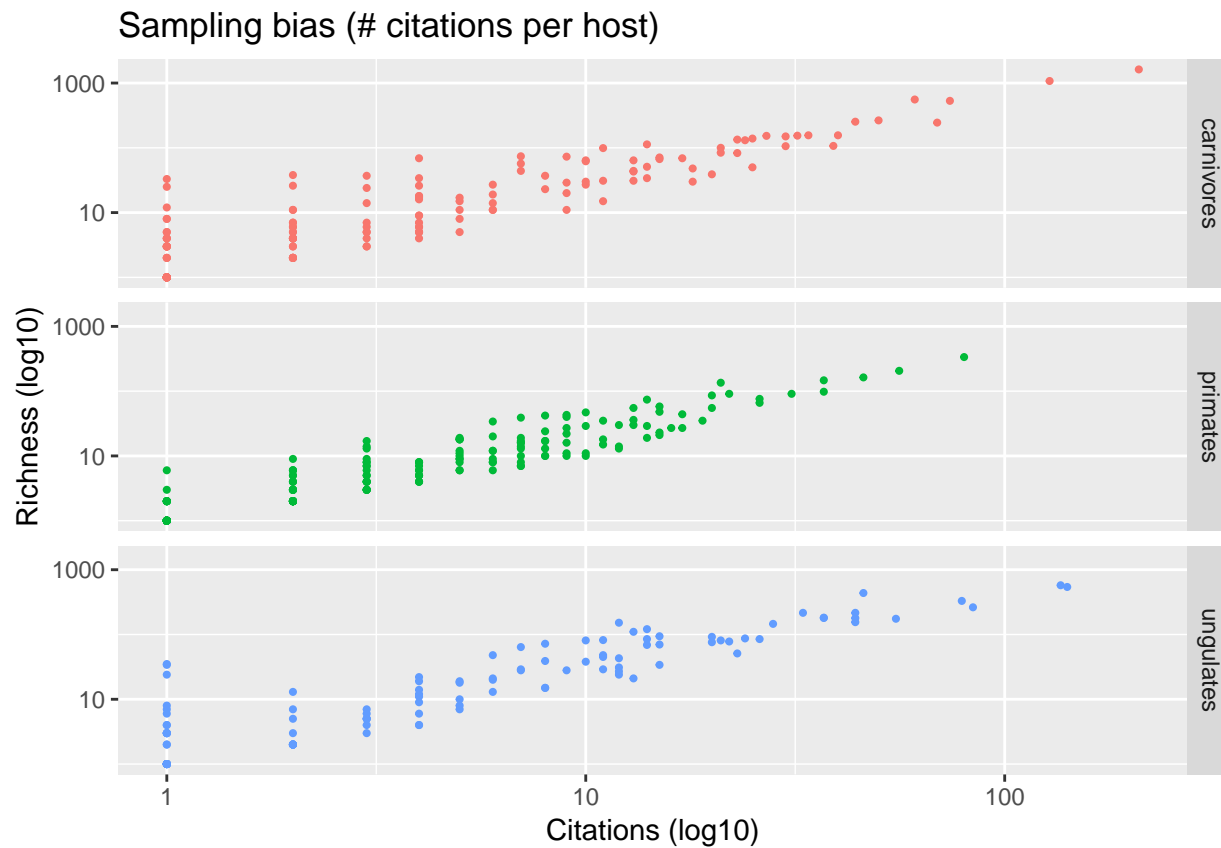
## Statistical Analyses

*Type of model/overall modeling scheme:* TBD

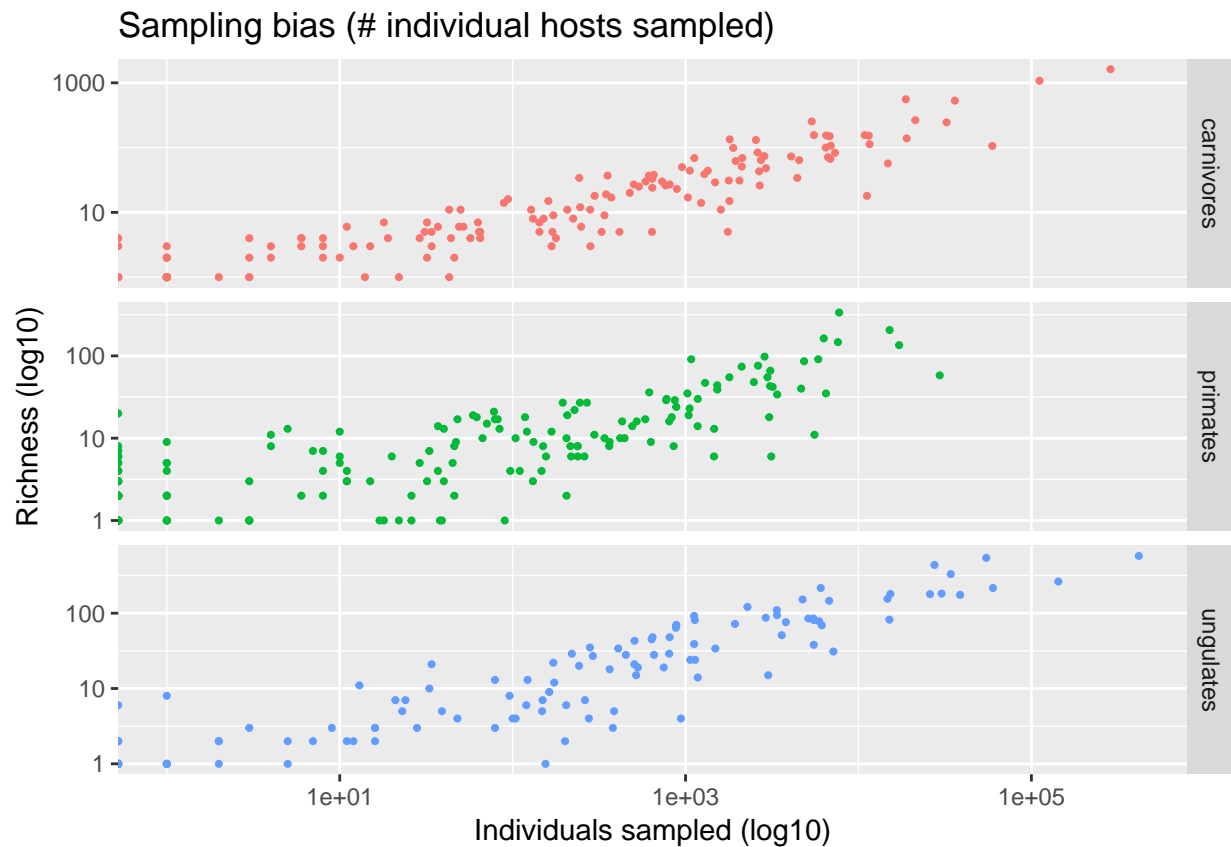
- Because carnivores, primates, and ungulates have different life histories, ecologies, and predictors of extinction risk, we will conduct models separately for each group, though the number of hosts per parasite (specialist/generalist variable) will be taken from merged host-parasite lists across all three groups.

*Controlling for uneven sampling:* To control for uneven sampling of parasites among hosts, we will include as

a covariate the number of unique citations per host (Altizer, Nunn & Lindenfors 2007). Note: could also try # of individuals sampled per host.



## Warning: Transformation introduced infinite values in continuous x-axis



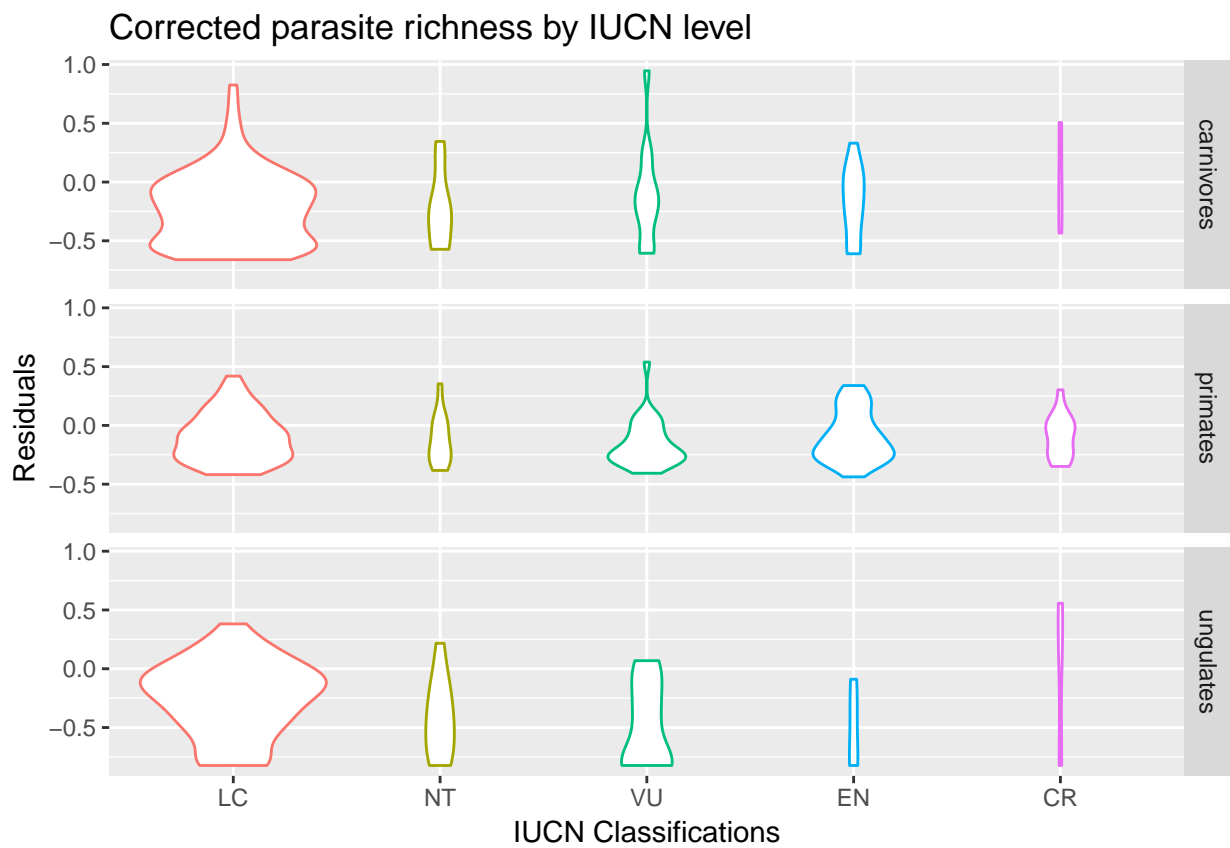
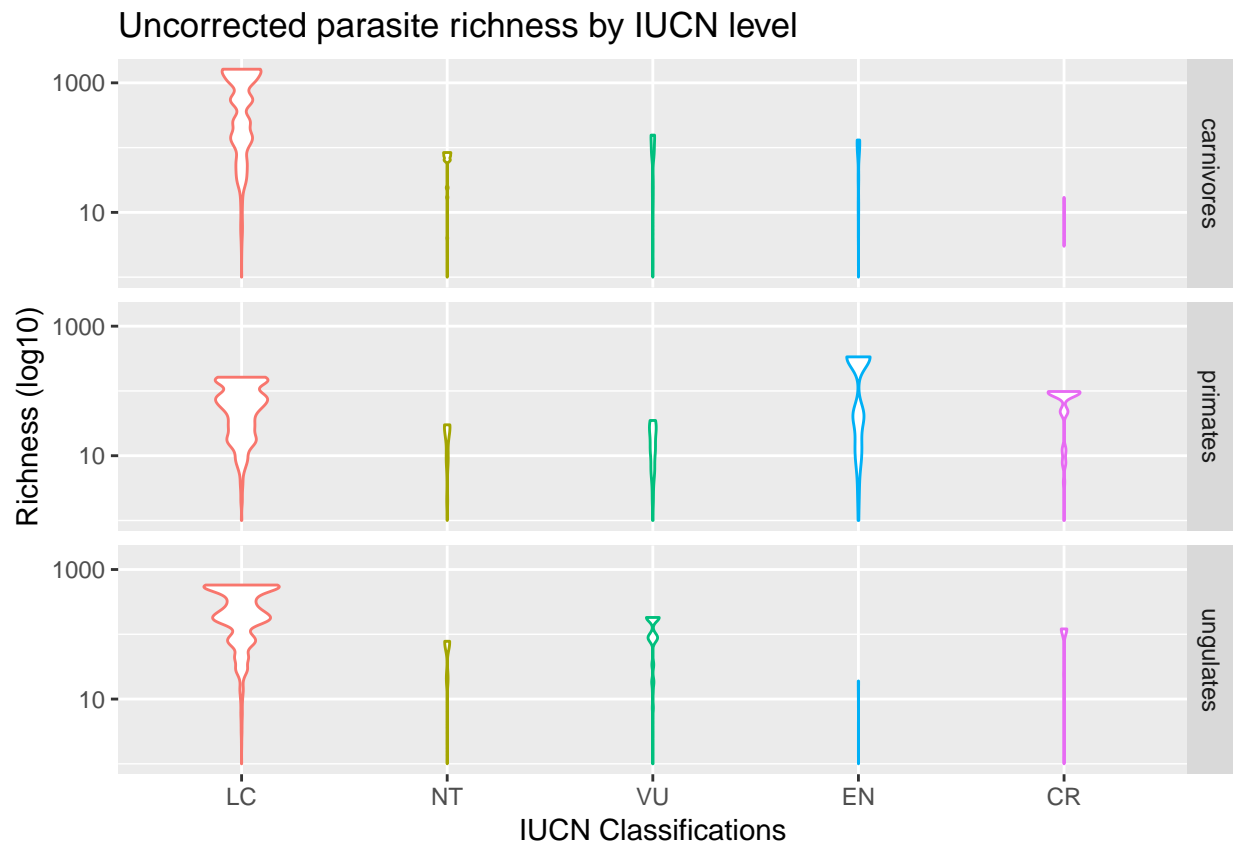
*Details of model:* \* Variables will be checked for colinearity \* Which packages used for which analyses: TBD  
 \* Whether to account for phylogenetic signal in residuals: TBD

#### *Sensitivity analyses*

- Reducing models by taking out predicts with  $p > 0.1$  and refitting with sampling/phylogenetic bias: TBD
- But threat status & citations will be forced into all models
- Could do bivariate model with threat status and citation count to explore effect of threat status in absence of ecological covariates
- Others?

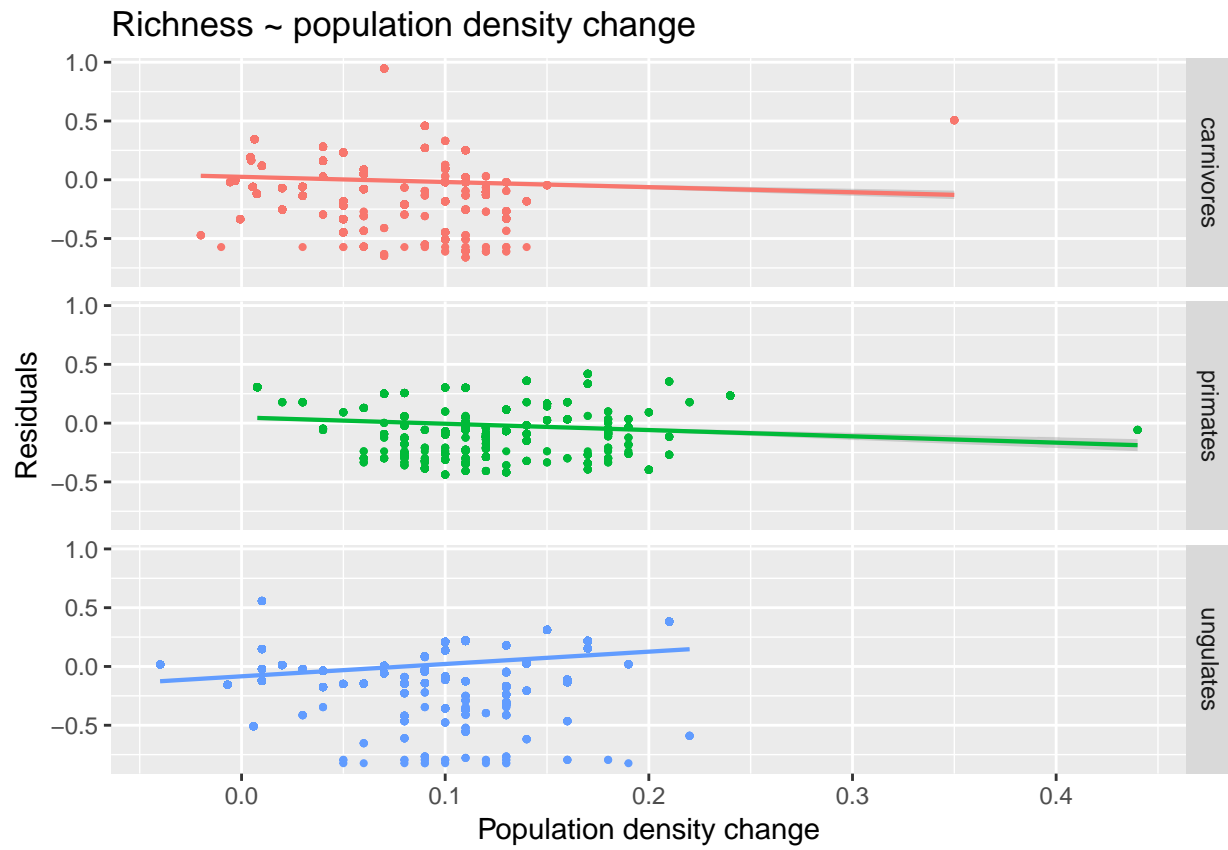
#### **Results:**

- Pattern of parasite spp richness across threat level (violin plot)





- pattern of parasite species richness by amount of population change



- pattern of parasite species richness by amount of population change & parasite type