Generalism data results

Josephine Walker

14 May 2016

## Summary of Metrics Calculated

See Poulin et al 2011 for a description of the facets of host specificity/generalism

|  |  |  |
| --- | --- | --- |
| Metric | Description | Facet |
| degree | number of hosts | basic |
| taxonomic SPD | mean pairwise phylogenetic distance between all hosts | phylogenetic |
| Faith's PD | minimum total length of all the phylogenetic branches required to span all hosts on the phylogenetic tree | phylogenetic |
| sumPD | sum of pairwise phylogenetic distance between all hosts | phylogenetic |

## Methods

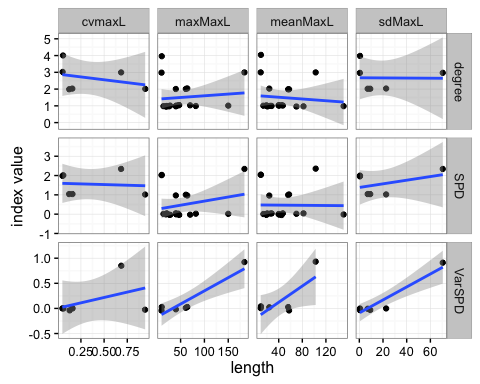
The generalism metrics for each parasite species were compared to measures of the characteristics described in Table 1. Metrics for endoparasites with direct life cycles were compared separately to summary measures (mean, maximum, standard deviation (SD), and coefficient of variation (CV)) of the maximum length reported for each of their hosts using linear models. Note that SD and CV of the host length are only calculated for parasites with more than one host.

Due to the distribution of degree for ectoparasites in this case we used a negative binomial generalized linear model (GLM) with a log link function; standard linear models were used for SPD and VarSPD. The mean maximum length was additionally divided into a categorical variable according to quartiles.

The effect of geographic range on parasites with a direct life cycle was calculated for ectoparasites only, due to the small number of endoparasites with a direct life cycle. As above, a linear model/ANOVA was used to assess the effect of geographical category on SPD and VarSPD, while a negative binomial GLM was used for degree. Regions were assessed as defined in Table 1 and also divided into two groups, where Antarctica, Nearctic, and Palearctic were assumed to be colder than Africa, Australia, Indopacific, and Neotropical regions. Some host-parasite associations were reported in more than one region, so for this analysis the generalism metrics were calculated separately for each region.

The binary parasite characteristics of complex life cycle and trophic transmission were compared to the parasite indices as above – using a linear model/t-test for SPD and VarSPD and a negative binomial GLM for degree. In this case, endo and ectoparasites are considered in the same analysis.

## Endoparasites & Host Body Size



summary(lm(data=hp.endoDirect,cvmaxL~degree))

##   
## Call:  
## lm(formula = cvmaxL ~ degree, data = hp.endoDirect)  
##   
## Residuals:  
## 7 11 16 17 18 20   
## -0.23790 0.50639 -0.27655 -0.31015 0.39853 -0.08031   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 0.7336 0.6118 1.199 0.297  
## degree -0.1486 0.2210 -0.673 0.538  
##   
## Residual standard error: 0.4034 on 4 degrees of freedom  
## (14 observations deleted due to missingness)  
## Multiple R-squared: 0.1016, Adjusted R-squared: -0.123   
## F-statistic: 0.4523 on 1 and 4 DF, p-value: 0.5381

summary(lm(data=hp.endoDirect,cvmaxL~SPD))

##   
## Call:  
## lm(formula = cvmaxL ~ SPD, data = hp.endoDirect)  
##   
## Residuals:  
## 7 11 16 17 18 20   
## -0.2642 0.5764 -0.2065 -0.2401 0.3896 -0.2552   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 0.41857 0.50588 0.827 0.455  
## SPD -0.05221 0.30557 -0.171 0.873  
##   
## Residual standard error: 0.4241 on 4 degrees of freedom  
## (14 observations deleted due to missingness)  
## Multiple R-squared: 0.007245, Adjusted R-squared: -0.2409   
## F-statistic: 0.02919 on 1 and 4 DF, p-value: 0.8726

summary(lm(data=hp.endoDirect,cvmaxL~VarSPD))

##   
## Call:  
## lm(formula = cvmaxL ~ VarSPD, data = hp.endoDirect)  
##   
## Residuals:  
## 7 11 16 17 18 20   
## -2.176e-01 6.753e-01 -1.077e-01 -1.413e-01 2.498e-16 -2.086e-01   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 0.2676 0.1701 1.573 0.191  
## VarSPD 0.4711 0.4686 1.005 0.372  
##   
## Residual standard error: 0.3803 on 4 degrees of freedom  
## (14 observations deleted due to missingness)  
## Multiple R-squared: 0.2017, Adjusted R-squared: 0.00214   
## F-statistic: 1.011 on 1 and 4 DF, p-value: 0.3716

summary(lm(data=hp.endoDirect,maxMaxL~degree))

##   
## Call:  
## lm(formula = maxMaxL ~ degree, data = hp.endoDirect)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -53.147 -27.207 -8.057 8.497 125.550   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 42.860 20.516 2.089 0.0512 .  
## degree 5.197 11.845 0.439 0.6661   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 45.87 on 18 degrees of freedom  
## Multiple R-squared: 0.01058, Adjusted R-squared: -0.04439   
## F-statistic: 0.1925 on 1 and 18 DF, p-value: 0.6661

summary(lm(data=hp.endoDirect,maxMaxL~SPD))

##   
## Call:  
## lm(formula = maxMaxL ~ SPD, data = hp.endoDirect)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -61.273 -23.530 -4.380 3.924 108.245   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 44.38 11.68 3.800 0.00131 \*\*  
## SPD 13.45 12.88 1.044 0.31032   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 44.78 on 18 degrees of freedom  
## Multiple R-squared: 0.05709, Adjusted R-squared: 0.004708   
## F-statistic: 1.09 on 1 and 18 DF, p-value: 0.3103

summary(lm(data=hp.endoDirect,maxMaxL~VarSPD))

##   
## Call:  
## lm(formula = maxMaxL ~ VarSPD, data = hp.endoDirect)  
##   
## Residuals:  
## 7 11 16 17 18 20   
## -2.710e+01 2.900e+00 2.690e+01 2.390e+01 5.329e-15 -2.660e+01   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 37.10 11.72 3.167 0.03397 \*   
## VarSPD 165.26 32.28 5.119 0.00689 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 26.2 on 4 degrees of freedom  
## (14 observations deleted due to missingness)  
## Multiple R-squared: 0.8676, Adjusted R-squared: 0.8345   
## F-statistic: 26.2 on 1 and 4 DF, p-value: 0.006892

summary(lm(data=hp.endoDirect,meanMaxL~degree))

##   
## Call:  
## lm(formula = meanMaxL ~ degree, data = hp.endoDirect)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -33.162 -25.490 -7.162 13.228 102.838   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 51.177 15.933 3.212 0.00483 \*\*  
## degree -4.015 9.199 -0.436 0.66768   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 35.63 on 18 degrees of freedom  
## Multiple R-squared: 0.01047, Adjusted R-squared: -0.0445   
## F-statistic: 0.1905 on 1 and 18 DF, p-value: 0.6677

summary(lm(data=hp.endoDirect,meanMaxL~SPD))

##   
## Call:  
## lm(formula = meanMaxL ~ SPD, data = hp.endoDirect)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -34.958 -24.526 -5.376 11.475 104.624   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 45.3761 9.3397 4.858 0.000126 \*\*\*  
## SPD -0.4756 10.3001 -0.046 0.963676   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 35.81 on 18 degrees of freedom  
## Multiple R-squared: 0.0001185, Adjusted R-squared: -0.05543   
## F-statistic: 0.002132 on 1 and 18 DF, p-value: 0.9637

summary(lm(data=hp.endoDirect,meanMaxL~VarSPD))

##   
## Call:  
## lm(formula = meanMaxL ~ VarSPD, data = hp.endoDirect)  
##   
## Residuals:  
## 7 11 16 17 18 20   
## -2.190e+01 -7.363e+00 2.614e+01 2.464e+01 4.441e-15 -2.151e+01   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 31.36 10.69 2.933 0.0427 \*  
## VarSPD 80.22 29.47 2.722 0.0529 .  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 23.91 on 4 degrees of freedom  
## (14 observations deleted due to missingness)  
## Multiple R-squared: 0.6495, Adjusted R-squared: 0.5618   
## F-statistic: 7.412 on 1 and 4 DF, p-value: 0.05285

summary(lm(data=hp.endoDirect,sdMaxL~degree))

##   
## Call:  
## lm(formula = sdMaxL ~ degree, data = hp.endoDirect)  
##   
## Residuals:  
## 7 11 16 17 18 20   
## -17.778 3.924 -9.511 -11.632 52.214 -17.218   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 19.6076 45.3709 0.432 0.688  
## degree -0.4523 16.3860 -0.028 0.979  
##   
## Residual standard error: 29.92 on 4 degrees of freedom  
## (14 observations deleted due to missingness)  
## Multiple R-squared: 0.0001905, Adjusted R-squared: -0.2498   
## F-statistic: 0.000762 on 1 and 4 DF, p-value: 0.9793

summary(lm(data=hp.endoDirect,sdMaxL~SPD))

##   
## Call:  
## lm(formula = sdMaxL ~ SPD, data = hp.endoDirect)  
##   
## Residuals:  
## 7 11 16 17 18 20   
## -25.6984 13.9379 0.5029 -1.6185 38.4669 -25.5908   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) -8.792 32.627 -0.269 0.801  
## SPD 17.482 19.708 0.887 0.425  
##   
## Residual standard error: 27.35 on 4 degrees of freedom  
## (14 observations deleted due to missingness)  
## Multiple R-squared: 0.1644, Adjusted R-squared: -0.04453   
## F-statistic: 0.7868 on 1 and 4 DF, p-value: 0.4252

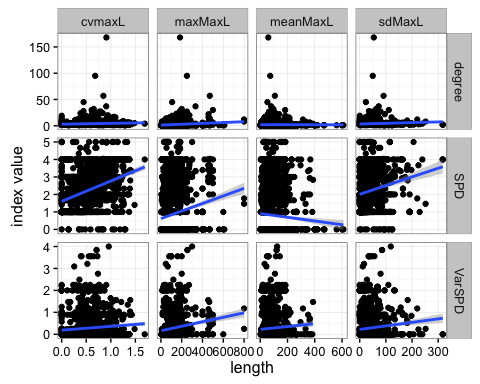
summary(lm(data=hp.endoDirect,sdMaxL~VarSPD))

##   
## Call:  
## lm(formula = sdMaxL ~ VarSPD, data = hp.endoDirect)  
##   
## Residuals:  
## 7 11 16 17 18 20   
## -7.516e+00 1.464e+01 1.204e+00 -9.177e-01 7.550e-15 -7.409e+00   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 7.989 4.049 1.973 0.11979   
## VarSPD 70.286 11.159 6.299 0.00325 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 9.055 on 4 degrees of freedom  
## (14 observations deleted due to missingness)  
## Multiple R-squared: 0.9084, Adjusted R-squared: 0.8855   
## F-statistic: 39.67 on 1 and 4 DF, p-value: 0.003247

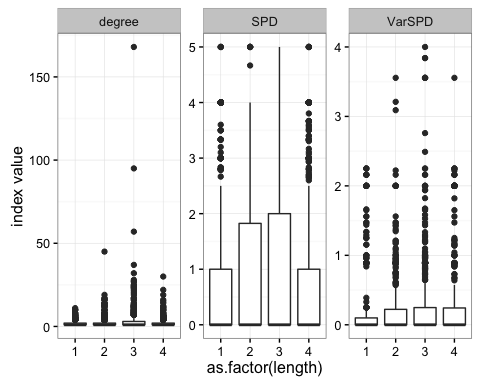
## Ectoparasites & Host Body Size

## Warning: Removed 21278 rows containing non-finite values (stat\_smooth).

## Warning: Removed 21278 rows containing missing values (geom\_point).



## Warning: Removed 2653 rows containing non-finite values (stat\_boxplot).



Negative binomial GLM seems like the best way to compare Ecto degree to quartiles of length categories. The estimates are log values, so exp() to get the estimate. Except, it's strictly positive so should use a zero truncated model?

summary(lm(data=hp.ectoDirect,cvmaxL~SPD))

##   
## Call:  
## lm(formula = cvmaxL ~ SPD, data = hp.ectoDirect)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.8175 -0.2446 -0.0123 0.2220 1.0840   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.236676 0.018157 13.04 <2e-16 \*\*\*  
## SPD 0.116167 0.007408 15.68 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.3327 on 1574 degrees of freedom  
## (2647 observations deleted due to missingness)  
## Multiple R-squared: 0.1351, Adjusted R-squared: 0.1346   
## F-statistic: 245.9 on 1 and 1574 DF, p-value: < 2.2e-16

summary(lm(data=hp.ectoDirect,cvmaxL~VarSPD))

##   
## Call:  
## lm(formula = cvmaxL ~ VarSPD, data = hp.ectoDirect)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.61554 -0.27229 -0.04986 0.21649 1.22153   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.468859 0.009929 47.220 < 2e-16 \*\*\*  
## VarSPD 0.065194 0.015474 4.213 2.66e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.3543 on 1554 degrees of freedom  
## (2667 observations deleted due to missingness)  
## Multiple R-squared: 0.01129, Adjusted R-squared: 0.01066   
## F-statistic: 17.75 on 1 and 1554 DF, p-value: 2.662e-05

summary(lm(data=hp.ectoDirect,maxMaxL~SPD))

##   
## Call:  
## lm(formula = maxMaxL ~ SPD, data = hp.ectoDirect)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -118.84 -61.57 -31.36 62.36 699.55   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 86.359 1.529 56.466 <2e-16 \*\*\*  
## SPD 9.621 1.016 9.466 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 83.25 on 4210 degrees of freedom  
## (11 observations deleted due to missingness)  
## Multiple R-squared: 0.02084, Adjusted R-squared: 0.02061   
## F-statistic: 89.61 on 1 and 4210 DF, p-value: < 2.2e-16

summary(lm(data=hp.ectoDirect,maxMaxL~VarSPD))

##   
## Call:  
## lm(formula = maxMaxL ~ VarSPD, data = hp.ectoDirect)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -142.38 -66.59 -22.83 61.87 690.91   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 102.828 2.511 40.952 < 2e-16 \*\*\*  
## VarSPD 25.177 3.921 6.422 1.78e-10 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 89.89 on 1565 degrees of freedom  
## (2656 observations deleted due to missingness)  
## Multiple R-squared: 0.02567, Adjusted R-squared: 0.02505   
## F-statistic: 41.24 on 1 and 1565 DF, p-value: 1.781e-10

summary(lm(data=hp.ectoDirect,meanMaxL~SPD))

##   
## Call:  
## lm(formula = meanMaxL ~ SPD, data = hp.ectoDirect)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -83.40 -54.40 -26.54 45.60 525.60   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 84.4009 1.3372 63.117 < 2e-16 \*\*\*  
## SPD -3.5445 0.8887 -3.989 6.76e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 72.78 on 4210 degrees of freedom  
## (11 observations deleted due to missingness)  
## Multiple R-squared: 0.003765, Adjusted R-squared: 0.003528   
## F-statistic: 15.91 on 1 and 4210 DF, p-value: 6.759e-05

summary(lm(data=hp.ectoDirect,meanMaxL~VarSPD))

##   
## Call:  
## lm(formula = meanMaxL ~ VarSPD, data = hp.ectoDirect)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -82.65 -46.50 -19.99 27.25 308.00   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 74.002 1.756 42.135 < 2e-16 \*\*\*  
## VarSPD 7.672 2.742 2.798 0.00521 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 62.87 on 1565 degrees of freedom  
## (2656 observations deleted due to missingness)  
## Multiple R-squared: 0.004977, Adjusted R-squared: 0.004341   
## F-statistic: 7.828 on 1 and 1565 DF, p-value: 0.005208

summary(lm(data=hp.ectoDirect,sdMaxL~SPD))

##   
## Call:  
## lm(formula = sdMaxL ~ SPD, data = hp.ectoDirect)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -50.165 -24.441 -12.788 8.708 273.403   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 19.7782 2.1296 9.287 < 2e-16 \*\*\*  
## SPD 6.0773 0.8689 6.994 3.93e-12 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 39.03 on 1574 degrees of freedom  
## (2647 observations deleted due to missingness)  
## Multiple R-squared: 0.03014, Adjusted R-squared: 0.02953   
## F-statistic: 48.92 on 1 and 1574 DF, p-value: 3.929e-12

summary(lm(data=hp.ectoDirect,sdMaxL~VarSPD))

##   
## Call:  
## lm(formula = sdMaxL ~ VarSPD, data = hp.ectoDirect)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -47.06 -25.52 -12.79 10.84 287.03   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 30.465 1.096 27.787 < 2e-16 \*\*\*  
## VarSPD 7.376 1.709 4.317 1.68e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 39.12 on 1554 degrees of freedom  
## (2667 observations deleted due to missingness)  
## Multiple R-squared: 0.01185, Adjusted R-squared: 0.01122   
## F-statistic: 18.64 on 1 and 1554 DF, p-value: 1.681e-05

# these LMs don't have straight Q-Q plots. Need to think about the implications.  
  
library(MASS)

##   
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':  
##   
## select

summary(glm.nb(data=hp.ectoDirect,degree ~ cvmaxL))

##   
## Call:  
## glm.nb(formula = degree ~ cvmaxL, data = hp.ectoDirect, init.theta = 2.764163414,   
## link = log)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.5676 -0.6411 -0.4357 -0.0065 11.7204   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 1.08921 0.03469 31.397 <2e-16 \*\*\*  
## cvmaxL 0.46109 0.05513 8.364 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for Negative Binomial(2.7642) family taken to be 1)  
##   
## Null deviance: 1403.5 on 1575 degrees of freedom  
## Residual deviance: 1341.9 on 1574 degrees of freedom  
## (2647 observations deleted due to missingness)  
## AIC: 7269.7  
##   
## Number of Fisher Scoring iterations: 1  
##   
##   
## Theta: 2.764   
## Std. Err.: 0.143   
##   
## 2 x log-likelihood: -7263.732

summary(glm.nb(data=hp.ectoDirect,degree ~ maxMaxL))

##   
## Call:  
## glm.nb(formula = degree ~ maxMaxL, data = hp.ectoDirect, init.theta = 2.928829563,   
## link = log)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.2294 -0.5662 -0.3934 0.1606 15.0265   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.3657332 0.0216160 16.92 <2e-16 \*\*\*  
## maxMaxL 0.0032613 0.0001509 21.61 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for Negative Binomial(2.9288) family taken to be 1)  
##   
## Null deviance: 3558.3 on 4211 degrees of freedom  
## Residual deviance: 3105.6 on 4210 degrees of freedom  
## (11 observations deleted due to missingness)  
## AIC: 15085  
##   
## Number of Fisher Scoring iterations: 1  
##   
##   
## Theta: 2.929   
## Std. Err.: 0.118   
##   
## 2 x log-likelihood: -15078.712

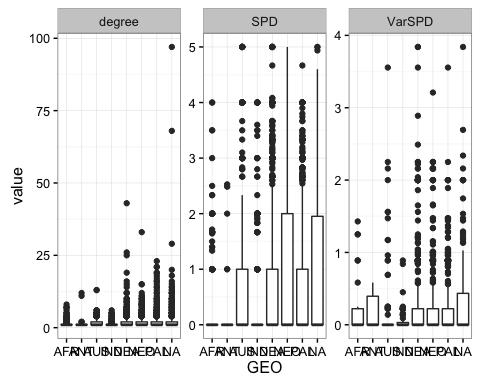
summary(glm.nb(data=hp.ectoDirect,degree ~ meanMaxL))

##   
## Call:  
## glm.nb(formula = degree ~ meanMaxL, data = hp.ectoDirect, init.theta = 2.52344473,   
## link = log)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.6446 -0.6381 -0.6148 -0.0257 15.6915   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.7292920 0.0217387 33.548 <2e-16 \*\*\*  
## meanMaxL -0.0001822 0.0002000 -0.911 0.362   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for Negative Binomial(2.5234) family taken to be 1)  
##   
## Null deviance: 3296.1 on 4211 degrees of freedom  
## Residual deviance: 3295.3 on 4210 degrees of freedom  
## (11 observations deleted due to missingness)  
## AIC: 15522  
##   
## Number of Fisher Scoring iterations: 1  
##   
##   
## Theta: 2.5234   
## Std. Err.: 0.0967   
##   
## 2 x log-likelihood: -15516.0440

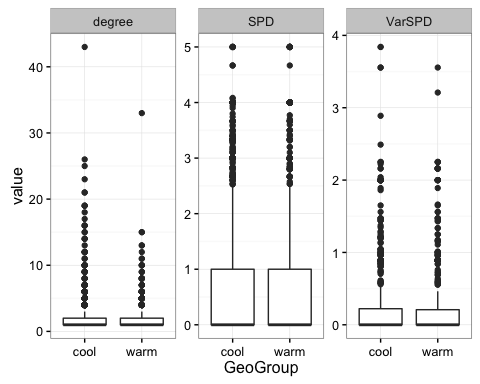
summary(glm.nb(data=hp.ectoDirect,degree ~ as.factor(mean.lcat)))

##   
## Call:  
## glm.nb(formula = degree ~ as.factor(mean.lcat), data = hp.ectoDirect,   
## init.theta = 2.763942524, link = log)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.9668 -0.5946 -0.4342 0.0486 14.2012   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.48940 0.03024 16.184 < 2e-16 \*\*\*  
## as.factor(mean.lcat)2 0.17390 0.04179 4.161 3.17e-05 \*\*\*  
## as.factor(mean.lcat)3 0.56872 0.03987 14.265 < 2e-16 \*\*\*  
## as.factor(mean.lcat)4 0.05284 0.04255 1.242 0.214   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for Negative Binomial(2.7639) family taken to be 1)  
##   
## Null deviance: 3459.3 on 4222 degrees of freedom  
## Residual deviance: 3191.8 on 4219 degrees of freedom  
## AIC: 15297  
##   
## Number of Fisher Scoring iterations: 1  
##   
##   
## Theta: 2.764   
## Std. Err.: 0.109   
##   
## 2 x log-likelihood: -15287.051

## Warning: Removed 3639 rows containing non-finite values (stat\_boxplot).



## Warning: Removed 5018 rows containing non-finite values (stat\_boxplot).



summary(lm(data=GeoDirectEcto,SPD ~ GEO))

##   
## Call:  
## lm(formula = SPD ~ GEO, data = GeoDirectEcto)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.0253 -0.7426 -0.6403 0.3597 4.2574   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.35421 0.06917 5.121 3.16e-07 \*\*\*  
## GEOANT -0.15829 0.18385 -0.861 0.389294   
## GEOAUS 0.45001 0.09124 4.932 8.41e-07 \*\*\*  
## GEOIND -0.06451 0.08562 -0.753 0.451222   
## GEONEA 0.38842 0.07677 5.059 4.36e-07 \*\*\*  
## GEONEO 0.67107 0.07953 8.438 < 2e-16 \*\*\*  
## GEOPAL 0.28606 0.07518 3.805 0.000143 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.155 on 4827 degrees of freedom  
## (581 observations deleted due to missingness)  
## Multiple R-squared: 0.03534, Adjusted R-squared: 0.03414   
## F-statistic: 29.47 on 6 and 4827 DF, p-value: < 2.2e-16

summary(lm(data=GeoDirectEcto,VarSPD ~ GEO))

##   
## Call:  
## lm(formula = VarSPD ~ GEO, data = GeoDirectEcto)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.2775 -0.2603 -0.1959 -0.0381 3.5625   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.14653 0.06485 2.259 0.0240 \*  
## GEOANT 0.04904 0.23915 0.205 0.8376   
## GEOAUS 0.02000 0.07835 0.255 0.7985   
## GEOIND -0.07752 0.08797 -0.881 0.3783   
## GEONEA 0.13095 0.06969 1.879 0.0604 .  
## GEONEO 0.11381 0.07006 1.625 0.1045   
## GEOPAL 0.04933 0.06885 0.716 0.4739   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.5147 on 1552 degrees of freedom  
## (3856 observations deleted due to missingness)  
## Multiple R-squared: 0.01126, Adjusted R-squared: 0.007442   
## F-statistic: 2.947 on 6 and 1552 DF, p-value: 0.007303

summary(glm.nb(data=GeoDirectEcto,degree ~ GEO))

##   
## Call:  
## glm.nb(formula = degree ~ GEO, data = GeoDirectEcto, init.theta = 6.922037142,   
## link = log)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.6612 -0.5933 -0.4841 0.1268 9.9275   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.36772 0.05476 6.715 1.89e-11 \*\*\*  
## GEOANT 0.05213 0.14292 0.365 0.715299   
## GEOAUS 0.12002 0.07079 1.696 0.089974 .   
## GEOIND -0.16152 0.06948 -2.325 0.020093 \*   
## GEONEA 0.22264 0.05983 3.721 0.000198 \*\*\*  
## GEONEO 0.28537 0.06138 4.650 3.33e-06 \*\*\*  
## GEOPAL 0.20177 0.05883 3.430 0.000604 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for Negative Binomial(6.922) family taken to be 1)  
##   
## Null deviance: 3392.6 on 4833 degrees of freedom  
## Residual deviance: 3291.8 on 4827 degrees of freedom  
## (581 observations deleted due to missingness)  
## AIC: 15430  
##   
## Number of Fisher Scoring iterations: 1  
##   
##   
## Theta: 6.922   
## Std. Err.: 0.487   
##   
## 2 x log-likelihood: -15414.026

summary(glm.nb(data=GeoDirectEcto,degree ~ GeoGroup))

##   
## Call:  
## glm.nb(formula = degree ~ GeoGroup, data = GeoDirectEcto, init.theta = 6.649502335,   
## link = log)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.5761 -0.5761 -0.4802 0.1434 9.8901   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.57628 0.01598 36.052 < 2e-16 \*\*\*  
## GeoGroupwarm -0.09070 0.02510 -3.614 0.000301 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for Negative Binomial(6.6495) family taken to be 1)  
##   
## Null deviance: 3355.2 on 4833 degrees of freedom  
## Residual deviance: 3342.1 on 4832 degrees of freedom  
## (581 observations deleted due to missingness)  
## AIC: 15507  
##   
## Number of Fisher Scoring iterations: 1  
##   
##   
## Theta: 6.650   
## Std. Err.: 0.459   
##   
## 2 x log-likelihood: -15501.253

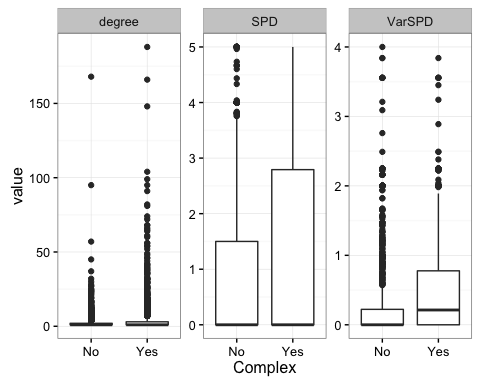
summary(lm(data=GeoDirectEcto,SPD ~ GeoGroup))

##   
## Call:  
## lm(formula = SPD ~ GeoGroup, data = GeoDirectEcto)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.7046 -0.7046 -0.6771 0.3229 4.3229   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.67710 0.02226 30.411 <2e-16 \*\*\*  
## GeoGroupwarm 0.02755 0.03422 0.805 0.421   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.176 on 4832 degrees of freedom  
## (581 observations deleted due to missingness)  
## Multiple R-squared: 0.0001341, Adjusted R-squared: -7.286e-05   
## F-statistic: 0.6479 on 1 and 4832 DF, p-value: 0.4209

summary(lm(data=GeoDirectEcto,VarSPD ~ GeoGroup))

##   
## Call:  
## lm(formula = VarSPD ~ GeoGroup, data = GeoDirectEcto)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.2325 -0.2325 -0.2076 -0.0103 3.6075   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.23248 0.01716 13.551 <2e-16 \*\*\*  
## GeoGroupwarm -0.02486 0.02653 -0.937 0.349   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.5167 on 1557 degrees of freedom  
## (3856 observations deleted due to missingness)  
## Multiple R-squared: 0.0005635, Adjusted R-squared: -7.835e-05   
## F-statistic: 0.8779 on 1 and 1557 DF, p-value: 0.3489

## Warning: Removed 4610 rows containing non-finite values (stat\_boxplot).



summary(glm.nb(data=hp.comp,degree ~ Complex))

##   
## Call:  
## glm.nb(formula = degree ~ Complex, data = hp.comp, init.theta = 1.347301944,   
## link = log)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.9710 -0.9710 -0.5436 -0.0173 12.5390   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.71240 0.01704 41.8 <2e-16 \*\*\*  
## ComplexYes 0.57178 0.02433 23.5 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for Negative Binomial(1.3473) family taken to be 1)  
##   
## Null deviance: 7230.4 on 7621 degrees of freedom  
## Residual deviance: 6673.4 on 7620 degrees of freedom  
## AIC: 32424  
##   
## Number of Fisher Scoring iterations: 1  
##   
##   
## Theta: 1.3473   
## Std. Err.: 0.0277   
##   
## 2 x log-likelihood: -32418.3850

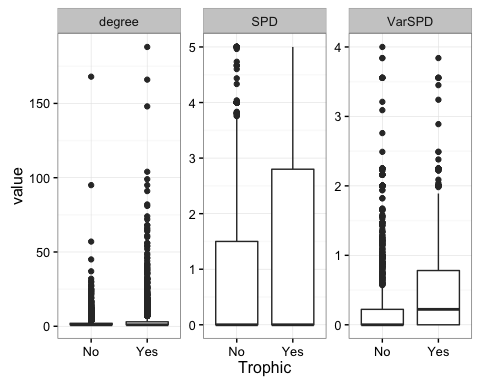
summary(lm(data=hp.comp,SPD ~ Complex))

##   
## Call:  
## lm(formula = SPD ~ Complex, data = hp.comp)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.1924 -1.1924 -0.8166 1.1167 4.1834   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.81663 0.02139 38.17 <2e-16 \*\*\*  
## ComplexYes 0.37577 0.03213 11.70 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.394 on 7620 degrees of freedom  
## Multiple R-squared: 0.01763, Adjusted R-squared: 0.0175   
## F-statistic: 136.8 on 1 and 7620 DF, p-value: < 2.2e-16

summary(lm(data=hp.comp,VarSPD ~ Complex))

##   
## Call:  
## lm(formula = VarSPD ~ Complex, data = hp.comp)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.4343 -0.2725 -0.2725 0.1467 3.7275   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.27253 0.01469 18.546 < 2e-16 \*\*\*  
## ComplexYes 0.16180 0.02128 7.603 3.85e-14 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.5834 on 3010 degrees of freedom  
## (4610 observations deleted due to missingness)  
## Multiple R-squared: 0.01884, Adjusted R-squared: 0.01852   
## F-statistic: 57.8 on 1 and 3010 DF, p-value: 3.851e-14

## Warning: Removed 4457 rows containing non-finite values (stat\_boxplot).



summary(glm.nb(data=hp.trop,degree ~ Trophic))

##   
## Call:  
## glm.nb(formula = degree ~ Trophic, data = hp.trop, init.theta = 1.328485495,   
## link = log)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.9842 -0.5550 -0.5550 -0.0331 12.3777   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.73020 0.01696 43.05 <2e-16 \*\*\*  
## TrophicYes 0.57808 0.02490 23.22 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for Negative Binomial(1.3285) family taken to be 1)  
##   
## Null deviance: 7014.3 on 7369 degrees of freedom  
## Residual deviance: 6468.7 on 7368 degrees of freedom  
## AIC: 31441  
##   
## Number of Fisher Scoring iterations: 1  
##   
##   
## Theta: 1.3285   
## Std. Err.: 0.0276   
##   
## 2 x log-likelihood: -31435.2590

summary(lm(data=hp.trop,SPD ~ Trophic))

##   
## Call:  
## lm(formula = SPD ~ Trophic, data = hp.trop)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.1875 -0.8269 -0.8269 1.1052 4.1731   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.82695 0.02121 38.98 <2e-16 \*\*\*  
## TrophicYes 0.36057 0.03283 10.98 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.39 on 7368 degrees of freedom  
## Multiple R-squared: 0.01611, Adjusted R-squared: 0.01598   
## F-statistic: 120.7 on 1 and 7368 DF, p-value: < 2.2e-16

summary(lm(data=hp.trop,VarSPD ~ Trophic))

##   
## Call:  
## lm(formula = VarSPD ~ Trophic, data = hp.trop)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.4392 -0.2716 -0.2716 0.1322 3.7284   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.27162 0.01455 18.668 < 2e-16 \*\*\*  
## TrophicYes 0.16759 0.02175 7.704 1.8e-14 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.5838 on 2911 degrees of freedom  
## (4457 observations deleted due to missingness)  
## Multiple R-squared: 0.01998, Adjusted R-squared: 0.01964   
## F-statistic: 59.35 on 1 and 2911 DF, p-value: 1.798e-14

with(hp.def,table(Trophic,Complex))

## Complex  
## Trophic No Yes Yes\_No  
## No 4233 54 0  
## Yes 10 3008 0

# Trophic and Complex are almost perfectly correlated in this subset of the dataset.