Mixed models for CS paper 2, all data

2024-01-19

## GLMMs from January 2024

I am coming back to this after several months away. I think I need to analyze each question using GLMM, with phylum\_class as a random effect. I think I understand the syntax here, which is this:

summary(glmer(CSss~Eco\_diff+(1|phylum\_class),data=surveymorpho,family=binomial))

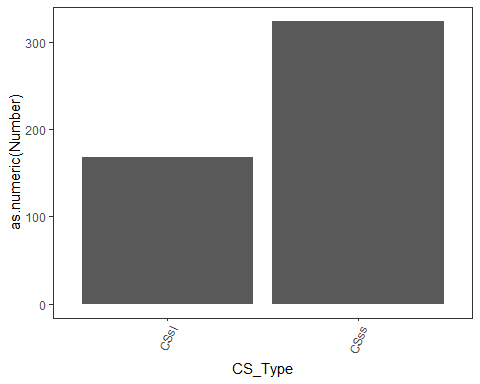
The probability of a species having CSss based on Eco\_diff, once the variance from phylum\_class is accounted for.

So, game plan: build a model and then work through the table with it.

First, code to load libraries, read in files, etc

1. Taxonomic work is needed? Are there more CSss or CSsl overall? This can probably stay the same, chi-square test.

##   
## Pearson's Chi-squared test with Yates' continuity correction  
##   
## data: cbind(expected, observed)  
## X-squared = 24.723, df = 1, p-value = 6.619e-07



More CSss than the 50% benchmark (significant), so reject the idea that most CS are just due to taxonomic oversight. BUT, does this vary with phyla or classes?

## Single term deletions  
##   
## Model:  
## CSss ~ phylum\_wormsV1  
## Df Deviance AIC LRT Pr(>Chi)  
## <none> 612.48 642.48   
## phylum\_wormsV1 14 631.73 633.73 19.249 0.1557

## Single term deletions  
##   
## Model:  
## CSss ~ class\_wormsV1  
## Df Deviance AIC LRT Pr(>Chi)   
## <none> 548.08 628.08   
## class\_wormsV1 39 623.60 625.60 75.517 0.0004055 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## Single term deletions  
##   
## Model:  
## CSss ~ phylum\_class  
## Df Deviance AIC LRT Pr(>Chi)   
## <none> 549.38 635.38   
## phylum\_class 42 631.73 633.73 82.354 0.0001979 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Phylum is not significant, but class is (both class\_wormsV1 and phylum\_class). This is because of a few classes, mostly Ascidea (others are minor classes, ie <10 species). So, from here I will be using mixed models with phylum\_class as a random factor. This will *remove variation due to phylum\_class* and allow us to see if there is a main effect once phylum\_class is accounted for. This is not a proper phylo correction per se but should get the job done.

1. Are there differences in time of description? For this, the independent variable is time (year of description OR decade of description; not sure why we need to bin but I will run both of them). Dependent variable is CSss.

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ decade + (1 | phylum\_class)  
## Data: surveymorpho  
##   
## AIC BIC logLik deviance df.resid   
## 550.1 662.9 -247.0 494.1 387   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.6639 -0.9388 0.4680 0.6979 1.9937   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 0.5113 0.715   
## Number of obs: 415, groups: phylum\_class, 41  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.70146 0.53727 1.306 0.1917   
## decade1770 0.15120 0.81507 0.186 0.8528   
## decade1780 0.82260 0.95374 0.863 0.3884   
## decade1790 -0.86605 0.91533 -0.946 0.3441   
## decade1800 1.97343 1.16978 1.687 0.0916 .  
## decade1810 -1.38065 0.79364 -1.740 0.0819 .  
## decade1820 -0.25268 0.75825 -0.333 0.7390   
## decade1830 -0.57851 0.65601 -0.882 0.3779   
## decade1840 -0.50099 0.66675 -0.751 0.4524   
## decade1850 0.16644 0.67643 0.246 0.8056   
## decade1860 -0.42001 0.76372 -0.550 0.5823   
## decade1870 -0.43422 0.66909 -0.649 0.5164   
## decade1880 -0.03201 0.72735 -0.044 0.9649   
## decade1890 -0.46197 0.80828 -0.572 0.5676   
## decade1900 0.53913 0.86465 0.624 0.5329   
## decade1910 2.15711 1.18805 1.816 0.0694 .  
## decade1920 -0.37760 1.04635 -0.361 0.7182   
## decade1930 -0.39618 0.85466 -0.464 0.6430   
## decade1940 -1.21309 0.83336 -1.456 0.1455   
## decade1950 -0.25444 1.36823 -0.186 0.8525   
## decade1960 1.34971 1.23259 1.095 0.2735   
## decade1970 -1.75899 1.04746 -1.679 0.0931 .  
## decade1980 0.22930 0.90405 0.254 0.7998   
## decade1990 0.89915 0.97110 0.926 0.3545   
## decade2000 -0.12298 0.80820 -0.152 0.8791   
## decade2010 0.39591 0.68820 0.575 0.5651   
## decade2020 -0.11368 0.64664 -0.176 0.8605   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
## optimizer (Nelder\_Mead) convergence code: 0 (OK)  
## Model failed to converge with max|grad| = 0.00490098 (tol = 0.002, component 1)

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ yearb + (1 | phylum\_class)  
## Data: surveymorpho  
##   
## AIC BIC logLik deviance df.resid   
## 538.1 550.2 -266.1 532.1 412   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -1.9167 -1.0852 0.5701 0.7388 1.1237   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 0.4049 0.6363   
## Number of obs: 415, groups: phylum\_class, 41  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)  
## (Intercept) -1.0217218 2.1249649 -0.481 0.631  
## yearb 0.0008689 0.0011261 0.772 0.440  
##   
## Correlation of Fixed Effects:  
## (Intr)  
## yearb -0.996  
## optimizer (Nelder\_Mead) convergence code: 0 (OK)  
## Model failed to converge with max|grad| = 0.0635708 (tol = 0.002, component 1)  
## Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## Model is nearly unidentifiable: large eigenvalue ratio  
## - Rescale variables?

So no effect of time on CSss, either by year (continuous) or decade (bins).

1. Are there differences in the number that are easy or hard to access through time? First, test trend in HARD to access through time:

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: Difficult ~ yearb + (1 | phylum\_class)  
## Data: surveyaccess  
##   
## AIC BIC logLik deviance df.resid   
## 544.5 558.7 -269.3 538.5 830   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -0.8243 -0.3683 -0.2557 -0.1867 6.8326   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 1.379 1.175   
## Number of obs: 833, groups: phylum\_class, 45  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -1.544e+01 1.540e+00 -10.027 <2e-16 \*\*\*  
## yearb 6.727e-03 7.947e-04 8.465 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr)  
## yearb -0.969  
## optimizer (Nelder\_Mead) convergence code: 0 (OK)  
## Model failed to converge with max|grad| = 0.046852 (tol = 0.002, component 1)  
## Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## Model is nearly unidentifiable: large eigenvalue ratio  
## - Rescale variables?

So there IS a trend where Difficult Access CS are described later. How about easy?

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: Easy ~ yearb + (1 | phylum\_class)  
## Data: surveyaccess  
##   
## AIC BIC logLik deviance df.resid   
## 764.2 778.4 -379.1 758.2 830   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -0.9769 -0.5137 -0.3571 -0.2603 3.9045   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 1.013 1.007   
## Number of obs: 833, groups: phylum\_class, 45  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 3.470069 2.076215 1.671 0.0947 .  
## yearb -0.002680 0.001117 -2.400 0.0164 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr)  
## yearb -0.994  
## optimizer (Nelder\_Mead) convergence code: 0 (OK)  
## Model failed to converge with max|grad| = 0.154402 (tol = 0.002, component 1)  
## Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## Model is nearly unidentifiable: large eigenvalue ratio  
## - Rescale variables?

…and where easy ones are described earlier.

1. How about HKK (habitat connectivity) and an association with CSss?

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ HKKv3 + (1 | phylum\_class)  
## Data: surveymorpho  
##   
## AIC BIC logLik deviance df.resid   
## 581.5 593.8 -287.7 575.5 449   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -2.3029 -1.1119 0.6233 0.7640 1.1755   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 0.3137 0.5601   
## Number of obs: 452, groups: phylum\_class, 42  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.5222582 0.1731721 3.016 0.00256 \*\*  
## HKKv3 0.0008152 0.0004124 1.977 0.04806 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr)  
## HKKv3 -0.320  
## optimizer (Nelder\_Mead) convergence code: 0 (OK)  
## Model failed to converge with max|grad| = 0.00249359 (tol = 0.002, component 1)  
## Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?

Slight effect of HKKv3, with a positive effect (large HKK more likely to show more CSss).

1. Is habitat stability associated with CSss?

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ Hstable + (1 | phylum\_class)  
## Data: surveymorpho  
##   
## AIC BIC logLik deviance df.resid   
## 623.2 635.8 -308.6 617.2 489   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -2.3187 -1.1324 0.6098 0.7218 1.0001   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 0.2208 0.4699   
## Number of obs: 492, groups: phylum\_class, 43  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.5515 0.1555 3.548 0.000389 \*\*\*  
## HstableTRUE 0.7342 0.2771 2.649 0.008069 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr)  
## HstableTRUE -0.321

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ Hstable + HKKv3 + (1 | phylum\_class)  
## Data: surveymorpho  
##   
## AIC BIC logLik deviance df.resid   
## 580.0 596.5 -286.0 572.0 448   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -2.2479 -1.1363 0.5843 0.7174 1.2197   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 0.3332 0.5773   
## Number of obs: 452, groups: phylum\_class, 42  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 4.731e-01 1.776e-01 2.664 0.00772 \*\*  
## HstableTRUE 7.306e-01 4.066e-01 1.797 0.07238 .   
## HKKv3 6.723e-05 5.850e-04 0.115 0.90850   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) HsTRUE  
## HstableTRUE -0.136   
## HKKv3 -0.124 -0.714  
## optimizer (Nelder\_Mead) convergence code: 0 (OK)  
## Model failed to converge with max|grad| = 0.00341633 (tol = 0.002, component 1)  
## Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## Model is nearly unidentifiable: large eigenvalue ratio  
## - Rescale variables?

No effect of habitat stability with these habitat classes. <- Note April 4 – NOW THAT I HAVE MADE CORAL NOT STABLE, THIS IS NOT TRUE. MORE CSSS IN STABLE HABITATS.

1. Are CSss more common in sympatry or not?

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ Sympatric + (1 | phylum\_class)  
## Data: surveymorpho  
##   
## AIC BIC logLik deviance df.resid   
## 611.9 624.5 -303.0 605.9 483   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -2.1845 -1.1038 0.5460 0.7155 1.0469   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 0.2389 0.4888   
## Number of obs: 486, groups: phylum\_class, 43  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 1.1725 0.2125 5.517 3.46e-08 \*\*\*  
## SympatricTRUE -0.6869 0.2137 -3.214 0.00131 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr)  
## SymptrcTRUE -0.703

CSss are LESS common in sympatry, which matches predictions (sympatric species have diagnostic diffs).

1. Are species in sympatry more likely to have ecological differentiation?

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: Eco\_diff ~ Sympatric + (1 | phylum\_class)  
## Data: survey  
##   
## AIC BIC logLik deviance df.resid   
## 157.4 166.4 -75.7 151.4 145   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -2.0354 0.4913 0.4913 0.4913 0.5774   
##   
## Random effects:  
## Groups Name Variance Std.Dev.   
## phylum\_class (Intercept) 8.275e-16 2.877e-08  
## Number of obs: 148, groups: phylum\_class, 27  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 1.0986 0.3651 3.009 0.00262 \*\*  
## SympatricTRUE 0.3228 0.4387 0.736 0.46187   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr)  
## SymptrcTRUE -0.832  
## optimizer (Nelder\_Mead) convergence code: 0 (OK)  
## boundary (singular) fit: see help('isSingular')

No, no relation

1. Are CSss more likely in species with eco diff?

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ Eco\_diff + (1 | phylum\_class)  
## Data: surveymorpho  
##   
## AIC BIC logLik deviance df.resid   
## 128.8 136.9 -61.4 122.8 104   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -2.1714 -1.0158 0.4605 0.6387 0.9874   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 0.503 0.7092   
## Number of obs: 107, groups: phylum\_class, 24  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 2.0688 0.6704 3.086 0.00203 \*\*  
## Eco\_diffYes -1.5193 0.6840 -2.221 0.02634 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr)  
## Eco\_diffYes -0.895

Species with ecological differentiation are LESS likely to be CSss (more likely to have diagnostic differences).

1. Biological traits associated with CSss?

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ hard\_skeleton + (1 | phylum\_class)  
## Data: surveymorpho  
##   
## AIC BIC logLik deviance df.resid   
## 597.3 609.7 -295.6 591.3 461   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -1.9032 -1.1096 0.5773 0.7634 0.9510   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 0.2719 0.5215   
## Number of obs: 464, groups: phylum\_class, 39  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.9118 0.2812 3.243 0.00118 \*\*  
## hard\_skeletonY -0.3557 0.3428 -1.038 0.29946   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr)  
## hard\_skltnY -0.822

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ ferti + (1 | phylum\_class)  
## Data: surveymorpho  
##   
## AIC BIC logLik deviance df.resid   
## 598.4 614.9 -295.2 590.4 460   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -1.7939 -1.2097 0.5646 0.7251 1.0472   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 0.1889 0.4346   
## Number of obs: 464, groups: phylum\_class, 39  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)  
## (Intercept) 0.3321 0.3391 0.979 0.327  
## fertiE 0.5330 0.3890 1.370 0.171  
## fertiI 0.1985 0.4729 0.420 0.675  
##   
## Correlation of Fixed Effects:  
## (Intr) fertiE  
## fertiE -0.851   
## fertiI -0.771 0.642

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ genitals + (1 | phylum\_class)  
## Data: surveymorpho  
##   
## AIC BIC logLik deviance df.resid   
## 599.3 615.8 -295.6 591.3 460   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -1.7480 -1.1816 0.5749 0.7428 0.9035   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 0.1864 0.4317   
## Number of obs: 464, groups: phylum\_class, 39  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)  
## (Intercept) 0.39876 0.49083 0.812 0.417  
## genitalsN 0.38638 0.52283 0.739 0.460  
## genitalsY 0.02626 0.58537 0.045 0.964  
##   
## Correlation of Fixed Effects:  
## (Intr) gntlsN  
## genitalsN -0.939   
## genitalsY -0.838 0.779

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ image + (1 | phylum\_class)  
## Data: surveymorpho  
##   
## AIC BIC logLik deviance df.resid   
## 598.2 610.6 -296.1 592.2 461   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -1.7974 -1.1105 0.5738 0.7416 0.9985   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 0.2847 0.5336   
## Number of obs: 464, groups: phylum\_class, 39  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.7171 0.1958 3.663 0.000249 \*\*\*  
## imageY -0.1342 0.3467 -0.387 0.698703   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr)  
## imageY -0.565

1. Ok, now larvae.

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ Repro\_invest + (1 | phylum\_class)  
## Data: surveylarvmorpho  
##   
## AIC BIC logLik deviance df.resid   
## 372.5 383.6 -183.3 366.5 289   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -1.7820 -1.2218 0.5771 0.6963 0.9800   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 0.2533 0.5033   
## Number of obs: 292, groups: phylum\_class, 42  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.70190 0.22014 3.188 0.00143 \*\*  
## Repro\_investLOW 0.05605 0.28962 0.194 0.84654   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr)  
## Rpr\_nvstLOW -0.599

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ Larv\_type + (1 | phylum\_class)  
## Data: surveylarvmorpho  
##   
## AIC BIC logLik deviance df.resid   
## 372.7 387.4 -182.4 364.7 288   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -1.9163 -1.2082 0.5684 0.6936 1.0485   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 0.2241 0.4734   
## Number of obs: 292, groups: phylum\_class, 42  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.9430 0.2875 3.280 0.00104 \*\*  
## Larv\_typeLT -0.5358 0.3977 -1.347 0.17793   
## Larv\_typePT -0.1634 0.3347 -0.488 0.62527   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) Lrv\_LT  
## Larv\_typeLT -0.658   
## Larv\_typePT -0.733 0.525

No effect of larval type (3 categories) or reproductive investment (2 categories).

1. Sediment vs rocky only

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ Hsubstrate + (1 | phylum\_class)  
## Data: surveysed  
##   
## AIC BIC logLik deviance df.resid   
## 281.5 291.8 -137.7 275.5 224   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -1.8099 -1.1580 0.5525 0.8636 0.8636   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 0 0   
## Number of obs: 227, groups: phylum\_class, 31  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.2933 0.1992 1.473 0.14083   
## HsubstrateSediment 0.8932 0.2910 3.069 0.00214 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr)  
## HsbstrtSdmn -0.684  
## optimizer (Nelder\_Mead) convergence code: 0 (OK)  
## boundary (singular) fit: see help('isSingular')

Significant! More likely to have CSss in sediment than on rocky shores.

# Now let’s make some graphs

Fig 1: ss vs sl by major class. I want this to be a barplot with Proportion CSss on the y and major classes on the x, with a horizontal dashed line for the overall proportion.

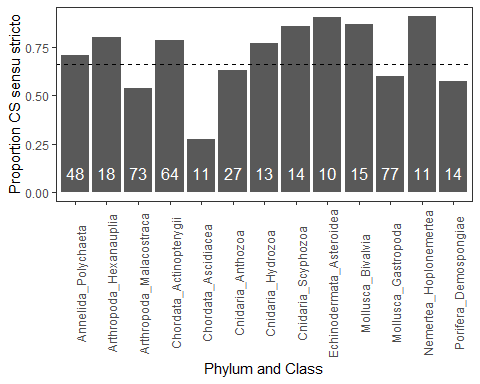


Fig 2: CS through time, with habitat access plotted over. NEED TO THINK THROUGH

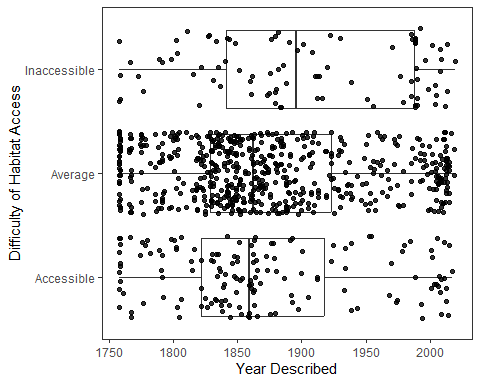


Fig 3: HKK and proportion CSss. This will be a bar graph with the 6 HKK values on the x, proportion ss on the y.

##   
## 1 5 10 25 100 1000  
## Diagnostic 4 37 49 49 9 11  
## No differentiation 1 20 34 37 22 11  
## Statistic 0 33 57 33 13 32

## 1 5 10 25 100 1000  
## stat\_total 0.0 0.3666667 0.4071429 0.2773109 0.2954545 0.5925926  
## diag\_total 0.8 0.4111111 0.3500000 0.4117647 0.2045455 0.2037037  
## nod\_total 0.2 0.2222222 0.2428571 0.3109244 0.5000000 0.2037037

##   
## 1 5 10 25 100 1000   
## 5 90 140 119 44 54

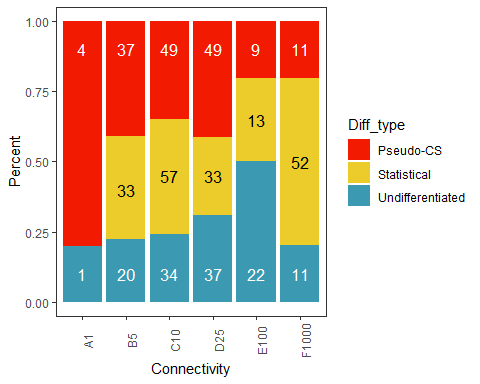
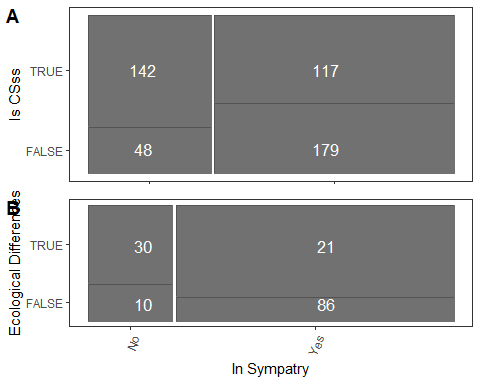
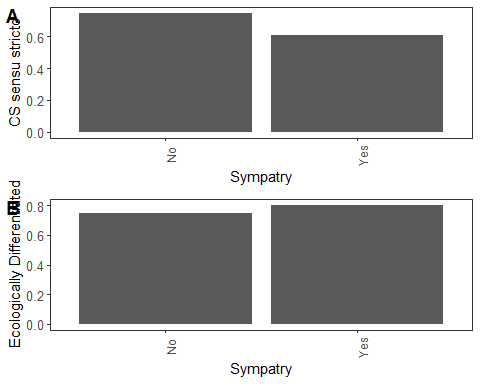


Fig 4: Sympatry vs morpho and eco diffs. This will be a two-panel with Proportion ss on the y, and on the x will be morpho diffs OR eco diffs (two panels).



# OK. Break in code and thoughts.

Here we are, March 7, and I’m trying to use AIC to compare different glmers with variables to see which is ‘most important.’

The variables that are significant in one-by-one analysis are sympatry, HKKv3, and sediment vs rocky habitat. We will therefore test this using the subset of species that are either soft sediment or rocky.

## cll df cAIC Refit  
## CSss ~ HKKv3 + (1 | phylum\_class) -118.07 0 236.15 FALSE  
## CSss ~ Sympatric + (1 | phylum\_class) -138.03 0 276.06 FALSE  
## CSss ~ HKKv3 \* Sympatric + (1 | phylum\_class) -115.65 0 231.30 FALSE

Next try: need to just run regular models to get AIC/BIC values, then make a table - I’m not going to be able to do the comparison functions (habitat data has zero covariance or something, look at error message if needed)

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ HKKv3 + (1 | phylum\_class)  
## Data: surveysed  
## AIC BIC logLik deviance df.resid   
## 254.9878 264.8977 -124.4939 248.9878 198   
## Random effects:  
## Groups Name Std.Dev.  
## phylum\_class (Intercept) 0.4763   
## Number of obs: 201, groups: phylum\_class, 27  
## Fixed Effects:  
## (Intercept) HKKv3   
## 0.28242 0.01899

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ Sympatric + (1 | phylum\_class)  
## Data: surveysed  
## AIC BIC logLik deviance df.resid   
## 286.4186 296.6802 -140.2093 280.4186 223   
## Random effects:  
## Groups Name Std.Dev.  
## phylum\_class (Intercept) 0.2275   
## Number of obs: 226, groups: phylum\_class, 31  
## Fixed Effects:  
## (Intercept) SympatricTRUE   
## 1.1201 -0.5842

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ Hsubstrate + (1 | phylum\_class)  
## Data: surveysed  
## AIC BIC logLik deviance df.resid   
## 281.4861 291.7610 -137.7431 275.4861 224   
## Random effects:  
## Groups Name Std.Dev.  
## phylum\_class (Intercept) 0   
## Number of obs: 227, groups: phylum\_class, 31  
## Fixed Effects:  
## (Intercept) HsubstrateSediment   
## 0.2933 0.8932   
## optimizer (Nelder\_Mead) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warnings

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ HKKv3 \* Sympatric + (1 | phylum\_class)  
## Data: surveysed  
## AIC BIC logLik deviance df.resid   
## 249.5437 266.0353 -119.7719 239.5437 195   
## Random effects:  
## Groups Name Std.Dev.  
## phylum\_class (Intercept) 0.362   
## Number of obs: 200, groups: phylum\_class, 27  
## Fixed Effects:  
## (Intercept) HKKv3 SympatricTRUE   
## 0.973544 0.003611 -1.279043   
## HKKv3:SympatricTRUE   
## 0.038005

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ HKKv3 \* Hsubstrate + (1 | phylum\_class)  
## Data: surveysed  
## AIC BIC logLik deviance df.resid   
## 253.6599 270.1764 -121.8300 243.6599 196   
## Random effects:  
## Groups Name Std.Dev.  
## phylum\_class (Intercept) 0   
## Number of obs: 201, groups: phylum\_class, 27  
## Fixed Effects:  
## (Intercept) HKKv3 HsubstrateSediment   
## 0.033322 0.010469 0.692369   
## HKKv3:HsubstrateSediment   
## 0.004861   
## optimizer (Nelder\_Mead) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warnings

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ Hsubstrate \* Sympatric + (1 | phylum\_class)  
## Data: surveysed  
## AIC BIC logLik deviance df.resid   
## 279.0738 296.1764 -134.5369 269.0738 221   
## Random effects:  
## Groups Name Std.Dev.   
## phylum\_class (Intercept) 1.871e-09  
## Number of obs: 226, groups: phylum\_class, 31  
## Fixed Effects:  
## (Intercept) HsubstrateSediment   
## 0.8690 0.4660   
## SympatricTRUE HsubstrateSediment:SympatricTRUE   
## -0.9708 0.7167   
## optimizer (Nelder\_Mead) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warnings

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ Hsubstrate \* Sympatric \* HKKv3 + (1 | phylum\_class)  
## Data: surveysed  
## AIC BIC logLik deviance df.resid   
## 250.8377 280.5226 -116.4189 232.8377 191   
## Random effects:  
## Groups Name Std.Dev.  
## phylum\_class (Intercept) 0.1126   
## Number of obs: 200, groups: phylum\_class, 27  
## Fixed Effects:  
## (Intercept) HsubstrateSediment   
## 0.8524817 0.3470714   
## SympatricTRUE HKKv3   
## -1.5504044 -0.0023419   
## HsubstrateSediment:SympatricTRUE HsubstrateSediment:HKKv3   
## 0.6090769 0.0045305   
## SympatricTRUE:HKKv3 HsubstrateSediment:SympatricTRUE:HKKv3   
## 0.0358587 -0.0003689

# Here again March 13 to test whether zones are related to CSss.

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ Spol + (1 | phylum\_class)  
## Data: spolmorpho  
##   
## AIC BIC logLik deviance df.resid   
## 474.9 486.6 -234.4 468.9 364   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -1.9138 -1.1305 0.5545 0.7587 1.1584   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 0.394 0.6277   
## Number of obs: 367, groups: phylum\_class, 39  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.5338 0.1897 2.814 0.00489 \*\*  
## SpolTRUE 1.0738 0.6180 1.738 0.08228 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr)  
## SpolTRUE -0.212

# No significant results.

# Consider only shallow species, sediment vs rocky (206 cases)

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ Hsubstrate + (1 | phylum\_class)  
## Data: shallow  
##   
## AIC BIC logLik deviance df.resid   
## 240.8 250.6 -117.4 234.8 190   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -2.1054 -1.0391 0.5245 0.7905 1.0307   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 0.1174 0.3426   
## Number of obs: 193, groups: phylum\_class, 31  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.2159 0.2416 0.894 0.3715   
## HsubstrateSediment 1.0190 0.3371 3.023 0.0025 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr)  
## HsbstrtSdmn -0.638

#Need something about zones?

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ Trop + (1 | phylum\_class)  
## Data: tropmorpho  
##   
## AIC BIC logLik deviance df.resid   
## 478.3 490.0 -236.1 472.3 364   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -1.8211 -1.1342 0.5703 0.7685 1.1084   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 0.3455 0.5878   
## Number of obs: 367, groups: phylum\_class, 39  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.55190 0.26280 2.100 0.0357 \*  
## TropTRUE 0.08883 0.26556 0.334 0.7380   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr)  
## TropTRUE -0.731

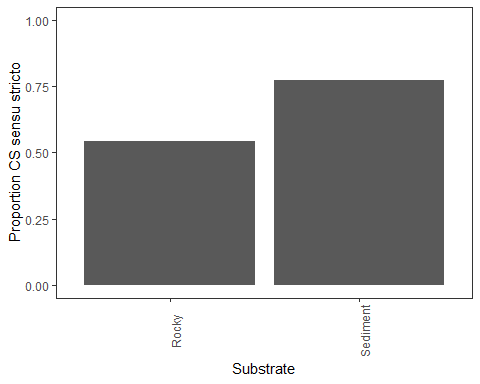
## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ Npol + (1 | phylum\_class)  
## Data: npolmorpho  
##   
## AIC BIC logLik deviance df.resid   
## 477.4 489.1 -235.7 471.4 364   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -1.8272 -1.0541 0.5591 0.7565 1.1449   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 0.347 0.589   
## Number of obs: 367, groups: phylum\_class, 39  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.6593 0.1850 3.565 0.000364 \*\*\*  
## NpolTRUE -0.3550 0.3561 -0.997 0.318822   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr)  
## NpolTRUE -0.242

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ Ntemp + (1 | phylum\_class)  
## Data: ntempmorpho  
##   
## AIC BIC logLik deviance df.resid   
## 477.6 489.3 -235.8 471.6 364   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -1.9937 -1.1428 0.5680 0.7538 1.0957   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 0.3471 0.5891   
## Number of obs: 367, groups: phylum\_class, 39  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.8130 0.2877 2.826 0.00472 \*\*  
## NtempTRUE -0.2488 0.2821 -0.882 0.37787   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr)  
## NtempTRUE -0.781

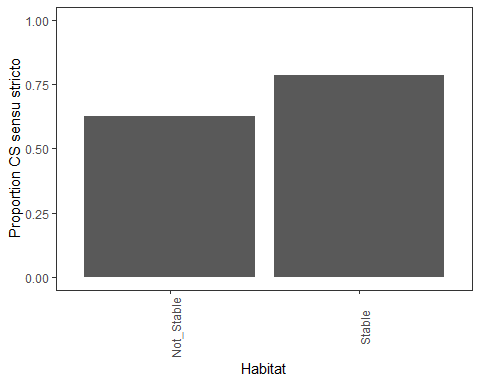
## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ Spol + (1 | phylum\_class)  
## Data: spolmorpho  
##   
## AIC BIC logLik deviance df.resid   
## 474.9 486.6 -234.4 468.9 364   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -1.9138 -1.1305 0.5545 0.7587 1.1584   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 0.394 0.6277   
## Number of obs: 367, groups: phylum\_class, 39  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.5338 0.1897 2.814 0.00489 \*\*  
## SpolTRUE 1.0738 0.6180 1.738 0.08228 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr)  
## SpolTRUE -0.212

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: CSss ~ Stemp + (1 | phylum\_class)  
## Data: stempmorpho  
##   
## AIC BIC logLik deviance df.resid   
## 478.4 490.1 -236.2 472.4 364   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -1.8093 -1.1615 0.5570 0.7484 1.0759   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## phylum\_class (Intercept) 0.3431 0.5857   
## Number of obs: 367, groups: phylum\_class, 39  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.60792 0.21850 2.782 0.0054 \*\*  
## StempTRUE 0.01556 0.23064 0.067 0.9462   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr)  
## StempTRUE -0.574

## Graphs on April 12 2024, prepping for Montpellier seminar

First, a graph looking at the shallow rocky vs sediment habitat 

Now let’s do this for stable habitats overall.



Need graph of the biological factors

