

GLMMs and GAMs summary

Naima Madi

08/03/2022

1. GAMs with polymorphism rate in a focal species as a function of community diversity

1.1. Polymorphism rate as a function of shannon diversity

```
##
## Family: Beta regression(83.127)
## Link function: logit
##
## Formula:
## polymorphism_rate ~ s(species_name, Shannon_alpha_diversity,
##   bs = "fs") + s(sample_id, bs = "re") + s(subject_id, bs = "re") +
##   s(Shannon_alpha_diversity) + s(total_reads_orig)
##
## Parametric coefficients:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.50462    0.06222  -88.47  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                                     edf  Ref.df   Chi.sq  p-value
## s(species_name,Shannon_alpha_diversity)  94.2839 657.000 3045.219 < 2e-16 ***
## s(sample_id)                             0.1071 464.000   0.064 1.000000
## s(subject_id)                            126.8345 248.000  341.220 < 2e-16 ***
## s(Shannon_alpha_diversity)                1.0654   1.110    4.687 0.031412 *
## s(total_reads_orig)                       2.1916   2.622   15.605 0.000883 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) = 0.198  Deviance explained = 33.2%
## -REML = -44506  Scale est. = 1          n = 8324
```

1.2. Polymorphism rate as a function of species richness (all data)

```
##
## Family: Beta regression(82.337)
## Link function: logit
##
## Formula:
```

```

## polymorphism_rate ~ s(species_richness) + s(total_reads_orig) +
##   s(species_name, species_richness, bs = "fs") + s(sample_id,
##   bs = "re") + s(subject_id, bs = "re")
##
## Parametric coefficients:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.50343    0.06384  -86.21  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##           edf Ref.df  Chi.sq p-value
## s(species_richness)      2.3921  2.897  10.596  0.0172 *
## s(total_reads_orig)      2.6499  3.180  29.839 2.32e-06 ***
## s(species_name,species_richness) 74.3303 653.000 2899.471 < 2e-16 ***
## s(sample_id)              0.1199 464.000    0.069  1.0000
## s(subject_id)            129.6311 248.000  355.976 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) = 0.191  Deviance explained = 32.3%
## -REML = -44490  Scale est. = 1          n = 8324

```

1.3. Polymorphism rate as a function of species richness (rarefied data)

```

##
## Family: Beta regression(81.057)
## Link function: logit
##
## Formula:
## polymorphism_rate ~ s(species_richness) + s(species_name, species_richness,
##   bs = "fs") + s(sample_id, bs = "re") + s(subject_id, bs = "re")
##
## Parametric coefficients:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.49214    0.06354  -86.43  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##           edf Ref.df  Chi.sq p-value
## s(species_richness)      2.75961  3.384  26.06 1.93e-05 ***
## s(species_name,species_richness) 65.10056 664.000 2835.25 < 2e-16 ***
## s(sample_id)              0.07985 435.000    0.05    1
## s(subject_id)            121.52785 234.000  317.22 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) = 0.191  Deviance explained = 32%
## -REML = -43612  Scale est. = 1          n = 8153

```

2. GAMs with polymorphism rate in a focal species as a function of shannon diversity at higher taxonomic levels

2.1. Polymorphism rate as a function of the phyla shannon diversity

```
##
## Family: Beta regression(82.102)
## Link function: logit
##
## Formula:
## polymorphism_rate ~ s(species_name, phyla_shannon, bs = "fs") +
##   s(sample_id, bs = "re") + s(subject_id, bs = "re") + s(phyla_shannon) +
##   s(total_reads_orig)
##
## Parametric coefficients:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.50443   0.06413  -85.83  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##           edf Ref.df  Chi.sq p-value
## s(species_name,phyla_shannon) 63.0641 667.000 2845.917 < 2e-16 ***
## s(sample_id)                   0.0525 464.000   0.030 1.000000
## s(subject_id)                  135.0904 248.000  391.928 < 2e-16 ***
## s(phyla_shannon)                1.0008   1.001   0.399 0.528475
## s(total_reads_orig)            2.2930   2.750  20.019 0.000224 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) = 0.189  Deviance explained = 32%
## -REML = -44485  Scale est. = 1          n = 8324
```

2.2. Polymorphism rate as a function of the class shannon diversity

```
##
## Family: Beta regression(82.102)
## Link function: logit
##
## Formula:
## polymorphism_rate ~ s(species_name, class_shannon, bs = "fs") +
##   s(sample_id, bs = "re") + s(subject_id, bs = "re") + s(class_shannon) +
##   s(total_reads_orig)
##
## Parametric coefficients:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.50443   0.06413  -85.83  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##           edf Ref.df  Chi.sq p-value
```

```

## s(species_name,class_shannon) 63.07154 668.000 2846.265 < 2e-16 ***
## s(sample_id) 0.05227 464.000 0.030 1.000000
## s(subject_id) 135.07734 248.000 391.858 < 2e-16 ***
## s(class_shannon) 1.00078 1.001 0.418 0.518425
## s(total_reads_orig) 2.29239 2.749 19.996 0.000227 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) = 0.189 Deviance explained = 32%
## -REML = -44485 Scale est. = 1 n = 8324

```

2.3. Polymorphism rate as a function of the order shannon diversity

```

##
## Family: Beta regression(82.095)
## Link function: logit
##
## Formula:
## polymorphism_rate ~ s(species_name, order_shannon, bs = "fs") +
##   s(sample_id, bs = "re") + s(subject_id, bs = "re") + s(order_shannon) +
##   s(total_reads_orig)
##
## Parametric coefficients:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.50422  0.06409  -85.88  <2e-16 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##           edf Ref.df  Chi.sq p-value
## s(species_name,order_shannon) 63.05687 667.000 2836.488 < 2e-16 ***
## s(sample_id) 0.03421 464.000 0.019 1.000000
## s(subject_id) 134.82575 248.000 389.563 < 2e-16 ***
## s(order_shannon) 1.00290 1.005 0.732 0.394604
## s(total_reads_orig) 2.28820 2.744 19.891 0.000235 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) = 0.189 Deviance explained = 32%
## -REML = -44485 Scale est. = 1 n = 8324

```

2.4. Polymorphism rate as a function of the family shannon diversity

```

##
## Family: Beta regression(82.6)
## Link function: logit
##
## Formula:
## polymorphism_rate ~ s(species_name, family_shannon, bs = "fs") +
##   s(sample_id, bs = "re") + s(subject_id, bs = "re") + s(family_shannon) +
##   s(total_reads_orig)
##

```

```

## Parametric coefficients:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.50114    0.06373  -86.32  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##           edf Ref.df   Chi.sq p-value
## s(species_name,family_shannon) 81.6475 665.000 2929.823 < 2e-16 ***
## s(sample_id)                   0.0644 464.000    0.038 1.000000
## s(subject_id)                  128.9861 248.000  351.108 < 2e-16 ***
## s(family_shannon)               1.0027  1.005    4.676 0.030850 *
## s(total_reads_orig)            2.1864  2.617   15.737 0.000807 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) = 0.192  Deviance explained = 32.6%
## -REML = -44497  Scale est. = 1          n = 8324

```

2.5. Polymorphism rate as a function of the genus shannon diversity

```

##
## Family: Beta regression(82.709)
## Link function: logit
##
## Formula:
## polymorphism_rate ~ s(species_name, genus_shannon, bs = "fs") +
##   s(sample_id, bs = "re") + s(subject_id, bs = "re") + s(genus_shannon) +
##   s(total_reads_orig)
##
## Parametric coefficients:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.5000    0.0636  -86.48  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##           edf Ref.df   Chi.sq p-value
## s(species_name,genus_shannon) 84.64871 665.000 2937.879 < 2e-16 ***
## s(sample_id)                   0.04288 464.000    0.026 1.00000
## s(subject_id)                  125.68401 248.000  330.285 < 2e-16 ***
## s(genus_shannon)               1.00319  1.006    6.506 0.01093 *
## s(total_reads_orig)            2.17477  2.603   15.552 0.00086 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) = 0.193  Deviance explained = 32.7%
## -REML = -44502  Scale est. = 1          n = 8324

```

3. GAMs with polymorphism rate in a focal species as a function of species richness at higher taxonomic levels

3.1. Polymorphism rate as a function of the phyla richness

```
##
## Family: Beta regression(82.85)
## Link function: logit
##
## Formula:
## polymorphism_rate ~ s(species_name, phyla_nb, bs = "fs", k = 5) +
##   s(sample_id, bs = "re", k = 5) + s(subject_id, bs = "re",
##   k = 5) + s(phyla_nb, k = 5) + s(total_reads_orig, k = 5)
##
## Parametric coefficients:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.50599   0.06421  -85.75  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##           edf  Ref.df   Chi.sq  p-value
## s(species_name,phyla_nb)  84.18112 332.000 3039.620 < 2e-16 ***
## s(sample_id)              0.05118 464.000   0.029 1.00000
## s(subject_id)            129.94503 248.000  356.246 < 2e-16 ***
## s(phyla_nb)               1.00193   1.003   7.517 0.00614 **
## s(total_reads_orig)       2.19187   2.550   24.352 2.07e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) = 0.196  Deviance explained = 32.8%
## -REML = -44497  Scale est. = 1          n = 8324
```

3.2. Polymorphism rate as a function of the class richness

```
##
## Family: Beta regression(82.57)
## Link function: logit
##
## Formula:
## polymorphism_rate ~ s(species_name, class_nb, bs = "fs", k = 5) +
##   s(sample_id, bs = "re", k = 5) + s(subject_id, bs = "re",
##   k = 5) + s(class_nb, k = 5) + s(total_reads_orig, k = 5)
##
## Parametric coefficients:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.50526   0.06414  -85.83  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##           edf  Ref.df   Chi.sq  p-value
```

```

## s(species_name,class_nb) 7.888e+01 336.000 2985.657 < 2e-16 ***
## s(sample_id)             3.942e-03 464.000    0.002  1.0000
## s(subject_id)           1.317e+02 248.000  368.973 < 2e-16 ***
## s(class_nb)             1.003e+00  1.005    6.501  0.0109 *
## s(total_reads_orig)     2.254e+00  2.618   26.160 1.23e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.194  Deviance explained = 32.5%
## -REML = -44492  Scale est. = 1          n = 8324

```

3.3. Polymorphism rate as a function of the order richness

```

##
## Family: Beta regression(82.285)
## Link function: logit
##
## Formula:
## polymorphism_rate ~ s(species_name, order_nb, bs = "fs", k = 5) +
##   s(sample_id, bs = "re", k = 5) + s(subject_id, bs = "re",
##   k = 5) + s(order_nb, k = 5) + s(total_reads_orig, k = 5)
##
## Parametric coefficients:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.50518    0.06406  -85.94  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##           edf Ref.df  Chi.sq p-value
## s(species_name,order_nb) 72.524 337.000 2913.982 < 2e-16 ***
## s(sample_id)             0.039 464.000    0.022  1.0000
## s(subject_id)           132.918 248.000  372.550 < 2e-16 ***
## s(order_nb)             1.009  1.015    5.628  0.0181 *
## s(total_reads_orig)     2.214  2.576   25.006 1.01e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.19  Deviance explained = 32.2%
## -REML = -44488  Scale est. = 1          n = 8324

```

3.4. Polymorphism rate as a function of the family richness

```

##
## Family: Beta regression(82.251)
## Link function: logit
##
## Formula:
## polymorphism_rate ~ s(species_name, family_nb, bs = "fs", k = 5) +
##   s(sample_id, bs = "re", k = 5) + s(subject_id, bs = "re",
##   k = 5) + s(family_nb, k = 5) + s(total_reads_orig, k = 5)
##

```

```

## Parametric coefficients:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.50380    0.06389  -86.15  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##           edf Ref.df  Chi.sq p-value
## s(species_name,family_nb) 71.93665 338.000 2902.469 < 2e-16 ***
## s(sample_id)              0.03462 464.000   0.020 1.00000
## s(subject_id)             130.42838 248.000  354.350 < 2e-16 ***
## s(family_nb)              1.00032   1.001   9.753 0.00179 **
## s(total_reads_orig)       2.24867   2.615  27.497 3.37e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) = 0.191  Deviance explained = 32.2%
## -REML = -44490  Scale est. = 1          n = 8324

```

3.5. Polymorphism rate as a function of the genus richness

```

##
## Family: Beta regression(82.453)
## Link function: logit
##
## Formula:
## polymorphism_rate ~ s(species_name, genus_nb, bs = "fs", k = 5) +
##   s(sample_id, bs = "re", k = 5) + s(subject_id, bs = "re",
##   k = 5) + s(genus_nb, k = 5) + s(total_reads_orig, k = 5)
##
## Parametric coefficients:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.50417    0.06381  -86.26  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##           edf Ref.df  Chi.sq p-value
## s(species_name,genus_nb) 78.66673 338.000 2923.286 < 2e-16 ***
## s(sample_id)              0.01746 464.000   0.010 1.00000
## s(subject_id)             129.82405 248.000  353.898 < 2e-16 ***
## s(genus_nb)               1.00718   1.012   8.095 0.00448 **
## s(total_reads_orig)       2.28861   2.659  28.260 2.42e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) = 0.193  Deviance explained = 32.4%
## -REML = -44492  Scale est. = 1          n = 8324

```


4. GLMMs with strain count in a focal species as a function of community diversity

4.1. Strain count in a focal species as a function of Shannon diversity

```
## Family: truncated_nbinom2 ( log )
## Formula:
## strain_nb ~ shannon_diversity + total_reads_orig + (shannon_diversity |
##   species_id) + (1 | sample_id) + (1 | subject_id)
## Data: datsc
##
##      AIC      BIC  logLik deviance df.resid
## 27664.6 27734.4 -13823.3 27646.6   17281
##
## Random effects:
##
## Conditional model:
## Groups      Name              Variance Std.Dev. Corr
## species_id (Intercept)      0.28677  0.53551
##           shannon_diversity 0.00458  0.06768 -0.41
## sample_id (Intercept)      0.02731  0.16527
## subject_id (Intercept)     0.06415  0.25328
## Number of obs: 17290, groups:  species_id, 194; sample_id, 469; subject_id, 249
##
## Overdispersion parameter for truncated_nbinom2 family (): 6.02e+05
##
## Conditional model:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -0.33976    0.05383  -6.312 2.75e-10 ***
## shannon_diversity 0.08056    0.02116   3.806 0.000141 ***
## total_reads_orig 0.08365    0.01993   4.198 2.69e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

4.2. Strain count in a focal species as a function of species richness (all data)

```
## Family: truncated_nbinom1 ( log )
## Formula:
## strain_nb ~ species_richness + total_reads_orig + (species_richness |
##   species_id) + (1 | sample_id) + (1 | subject_id)
## Data: datsc
##
##      AIC      BIC  logLik deviance df.resid
## 27668.1 27737.9 -13825.0 27650.1   17281
##
## Random effects:
##
## Conditional model:
## Groups      Name              Variance Std.Dev. Corr
## species_id (Intercept)      0.2821584 0.53119
##           species_richness 0.0005942 0.02438  0.88
## sample_id (Intercept)      0.0233862 0.15293
```

```

## subject_id (Intercept)      0.0699615 0.26450
## Number of obs: 17290, groups:  species_id, 194; sample_id, 469; subject_id, 249
##
## Overdispersion parameter for truncated_nbinom1 family (): 3.94e-06
##
## Conditional model:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -0.33776    0.05339  -6.326 2.52e-10 ***
## species_richness -0.08880    0.02066  -4.299 1.72e-05 ***
## total_reads_orig  0.11849    0.02145   5.525 3.30e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

4.3. Strain count in a focal species as a function of richness (rarefied data)

```

## Family: truncated_nbinom1 ( log )
## Formula:
## strain_nb ~ species_richness + (species_richness | species_id) +
## (1 | sample_id) + (1 | subject_id)
## Data: datsc
##
##      AIC      BIC  logLik deviance df.resid
## 26835.6 26897.4 -13409.8 26819.6   16624
##
## Random effects:
##
## Conditional model:
## Groups      Name              Variance Std.Dev. Corr
## species_id (Intercept)      2.870e-01 0.535739
##           species_richness 8.102e-05 0.009001 -0.71
## sample_id (Intercept)      3.460e-02 0.186023
## subject_id (Intercept)     5.660e-02 0.237900
## Number of obs: 16632, groups:  species_id, 190; sample_id, 437; subject_id, 235
##
## Overdispersion parameter for truncated_nbinom1 family (): 6.73e-06
##
## Conditional model:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -0.33196    0.05393  -6.155 7.49e-10 ***
## species_richness -0.04108    0.01925  -2.134  0.0328 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

5. Strain count in a focal species as a function of shannon diversity at higher taxonomic levels

5.1. Strain count in a focal species as a function of phyla shannon diversity

```

## Family: truncated_nbinom1 ( log )
## Formula:
## strain_nb ~ phyla_shannon + total_reads_orig + (phyla_shannon |

```

```

##      species_id) + (1 | sample_id) + (1 | subject_id)
## Data: datsc2
##
##      AIC      BIC   logLik deviance df.resid
## 27677.8 27747.6 -13829.9 27659.8   17281
##
## Random effects:
##
## Conditional model:
## Groups      Name          Variance Std.Dev. Corr
## species_id (Intercept) 0.289567 0.53811
##          phyla_shannon 0.001328 0.03645 -0.99
## sample_id (Intercept) 0.026701 0.16340
## subject_id (Intercept) 0.070000 0.26458
## Number of obs: 17290, groups:  species_id, 194; sample_id, 469; subject_id, 249
##
## Overdispersion parameter for truncated_nbinom1 family (): 1.82e-06
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -0.33929   0.05434  -6.244 4.27e-10 ***
## phyla_shannon   0.03711   0.01910   1.943  0.052 .
## total_reads_orig 0.08143   0.02002   4.067 4.76e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

5.2. Strain count in a focal species as a function of classes shannon diversity

```

## Family: truncated_nbinom1 ( log )
## Formula:
## strain_nb ~ class_shannon + total_reads_orig + (class_shannon |
##      species_id) + (1 | sample_id) + (1 | subject_id)
## Data: datsc2
##
##      AIC      BIC   logLik deviance df.resid
## 27677.5 27747.3 -13829.8 27659.5   17281
##
## Random effects:
##
## Conditional model:
## Groups      Name          Variance Std.Dev. Corr
## species_id (Intercept) 0.290706 0.53917
##          class_shannon 0.001383 0.03719 -1.00
## sample_id (Intercept) 0.027240 0.16505
## subject_id (Intercept) 0.069897 0.26438
## Number of obs: 17290, groups:  species_id, 194; sample_id, 469; subject_id, 249
##
## Overdispersion parameter for truncated_nbinom1 family (): 1.35e-06
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -0.33957   0.05443  -6.239 4.40e-10 ***
## class_shannon   0.03789   0.01914   1.980  0.0477 *

```

```
## total_reads_orig 0.08129 0.02006 4.052 5.07e-05 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

5.3. Strain count in a focal species as a function of orders shannon diversity

```
## Family: truncated_nbinom1 ( log )
## Formula:
## strain_nb ~ order_shannon + total_reads_orig + (order_shannon |
##   species_id) + (1 | sample_id) + (1 | subject_id)
## Data: datsc2
##
##      AIC      BIC  logLik deviance df.resid
## 27678.5 27748.4 -13830.3 27660.5 17281
##
## Random effects:
##
## Conditional model:
## Groups      Name          Variance Std.Dev. Corr
## species_id (Intercept) 0.290321 0.53881
## order_shannon 0.001303 0.03609 -1.00
## sample_id (Intercept) 0.027042 0.16444
## subject_id (Intercept) 0.070228 0.26501
## Number of obs: 17290, groups: species_id, 194; sample_id, 469; subject_id, 249
##
## Overdispersion parameter for truncated_nbinom1 family (): 5.22e-07
##
## Conditional model:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.33895 0.05440 -6.230 4.66e-10 ***
## order_shannon 0.03390 0.01911 1.774 0.076 .
## total_reads_orig 0.08095 0.02005 4.038 5.38e-05 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

5.4. Strain count in a focal species as a function of families shannon diversity

```
## Family: truncated_nbinom1 ( log )
## Formula:
## strain_nb ~ family_shannon + total_reads_orig + (family_shannon |
##   species_id) + (1 | sample_id) + (1 | subject_id)
## Data: datsc2
##
##      AIC      BIC  logLik deviance df.resid
## 27679.8 27749.6 -13830.9 27661.8 17281
##
## Random effects:
##
## Conditional model:
## Groups      Name          Variance Std.Dev. Corr
## species_id (Intercept) 0.288718 0.5373
## family_shannon 0.001037 0.0322 -0.95
```

```

## sample_id (Intercept)    0.027718 0.1665
## subject_id (Intercept)  0.068444 0.2616
## Number of obs: 17290, groups:  species_id, 194; sample_id, 469; subject_id, 249
##
## Overdispersion parameter for truncated_nbinom1 family (): 3.93e-06
##
## Conditional model:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -0.33864    0.05412  -6.258 3.91e-10 ***
## family_shannon  0.04001    0.01947   2.055  0.0399 *
## total_reads_orig 0.08174    0.02005   4.077 4.56e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

5.5. Strain count in a focal species as a function of genera shannon diversity

```

## Family: truncated_nbinom1 ( log )
## Formula:
## strain_nb ~ genus_shannon + total_reads_orig + (genus_shannon |
##   species_id) + (1 | sample_id) + (1 | subject_id)
## Data: datsc2
##
##      AIC      BIC  logLik deviance df.resid
## 27665.8 27735.6 -13823.9 27647.8   17281
##
## Random effects:
##
## Conditional model:
## Groups      Name          Variance Std.Dev. Corr
## species_id (Intercept)  0.294181 0.54238
##           genus_shannon 0.002799 0.05291 -0.97
## sample_id (Intercept)  0.028057 0.16750
## subject_id (Intercept) 0.065306 0.25555
## Number of obs: 17290, groups:  species_id, 194; sample_id, 469; subject_id, 249
##
## Overdispersion parameter for truncated_nbinom1 family (): 4.08e-06
##
## Conditional model:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -0.34008    0.05447  -6.243 4.28e-10 ***
## genus_shannon  0.06839    0.01994   3.430 0.000605 ***
## total_reads_orig 0.08094    0.02001   4.045 5.23e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

6. GLMMs with strain count in a focal species as a function of richness at higher taxonomic levels

6.1. Strain count in a focal species as a function of the phyla richness

```

## Family: truncated_nbinom2 ( log )

```

```

## Formula:
## strain_nb ~ phyla_nb + total_reads_orig + (phyla_nb | species_id) +
##   (1 | sample_id) + (1 | subject_id)
## Data: datsc2
##
##      AIC      BIC  logLik deviance df.resid
## 27674.0 27743.8 -13828.0 27656.0   17281
##
## Random effects:
##
## Conditional model:
## Groups      Name          Variance Std.Dev. Corr
## species_id (Intercept) 0.2802192 0.52936
##           phyla_nb     0.0005011 0.02239 0.99
## sample_id (Intercept) 0.0231251 0.15207
## subject_id (Intercept) 0.0744282 0.27282
## Number of obs: 17290, groups:  species_id, 194; sample_id, 469; subject_id, 249
##
## Overdispersion parameter for truncated_nbinom2 family (): 2.09e+04
##
## Conditional model:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -0.33972    0.05354  -6.345 2.22e-10 ***
## phyla_nb       -0.06447    0.01842  -3.500 0.000466 ***
## total_reads_orig 0.09490    0.02007   4.727 2.27e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

6.2. Strain count in a focal species as a function of the classes richness

```

## Family: truncated_nbinom1 ( log )
## Formula:      strain_nb ~ class_nb + total_reads_orig + (1 | species_id) +
##   (1 | sample_id) + +(1 | subject_id)
## Data: datsc2
##
##      AIC      BIC  logLik deviance df.resid
## 27669.4 27723.7 -13827.7 27655.4   17283
##
## Random effects:
##
## Conditional model:
## Groups      Name          Variance Std.Dev.
## species_id (Intercept) 0.28418 0.5331
## sample_id (Intercept) 0.02245 0.1498
## subject_id (Intercept) 0.07361 0.2713
## Number of obs: 17290, groups:  species_id, 194; sample_id, 469; subject_id, 249
##
## Overdispersion parameter for truncated_nbinom1 family (): 2.92e-06
##
## Conditional model:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -0.33925    0.05383  -6.303 2.93e-10 ***
## class_nb      -0.06821    0.01810  -3.768 0.000164 ***

```

```
## total_reads_orig 0.10267 0.02032 5.052 4.38e-07 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

6.3. Strain count in a focal species as a function of the orders richness

```
## Family: truncated_nbinom1 ( log )
## Formula: strain_nb ~ order_nb + total_reads_orig + (1 | species_id) +
## (1 | sample_id) + (1 | subject_id)
## Data: datsc2
##
## AIC BIC logLik deviance df.resid
## 27657.3 27711.6 -13821.7 27643.3 17283
##
## Random effects:
##
## Conditional model:
## Groups Name Variance Std.Dev.
## species_id (Intercept) 0.28583 0.5346
## sample_id (Intercept) 0.02274 0.1508
## subject_id (Intercept) 0.06628 0.2575
## Number of obs: 17290, groups: species_id, 194; sample_id, 469; subject_id, 249
##
## Overdispersion parameter for truncated_nbinom1 family (): 2.91e-06
##
## Conditional model:
## Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.33685 0.05364 -6.279 3.40e-10 ***
## order_nb -0.09959 0.01926 -5.172 2.32e-07 ***
## total_reads_orig 0.12437 0.02106 5.906 3.51e-09 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

6.4. Strain count in a focal species as a function of the families richness

```
## Family: truncated_nbinom1 ( log )
## Formula: strain_nb ~ family_nb + total_reads_orig + (1 | species_id) +
## (1 | sample_id) + (1 | subject_id)
## Data: datsc2
##
## AIC BIC logLik deviance df.resid
## 27647.2 27701.5 -13816.6 27633.2 17283
##
## Random effects:
##
## Conditional model:
## Groups Name Variance Std.Dev.
## species_id (Intercept) 0.28598 0.5348
## sample_id (Intercept) 0.02070 0.1439
## subject_id (Intercept) 0.06614 0.2572
## Number of obs: 17290, groups: species_id, 194; sample_id, 469; subject_id, 249
```

```

##
## Overdispersion parameter for truncated_nbinom1 family (): 2.94e-06
##
## Conditional model:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.33543    0.05361  -6.256 3.94e-10 ***
## family_nb   -0.11393    0.01863  -6.115 9.65e-10 ***
## total_reads_orig 0.12324    0.02023   6.091 1.12e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

6.5. Strain count in a focal species as a function of genera richness

```

## Family: truncated_nbinom1 ( log )
## Formula:
## strain_nb ~ genus_nb + total_reads_orig + (genus_nb | species_id) +
##           (1 | sample_id) + (1 | subject_id)
## Data: datsc2
##
##           AIC          BIC    logLik deviance df.resid
## 27673.9 27743.8 -13828.0 27655.9    17281
##
## Random effects:
##
## Conditional model:
## Groups      Name          Variance Std.Dev. Corr
## species_id (Intercept) 0.2820803 0.53111
##           genus_nb     0.0006509 0.02551 0.70
## sample_id (Intercept) 0.0232048 0.15233
## subject_id (Intercept) 0.0732920 0.27072
## Number of obs: 17290, groups: species_id, 194; sample_id, 469; subject_id, 249
##
## Overdispersion parameter for truncated_nbinom1 family (): 4.05e-06
##
## Conditional model:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.33869    0.05358  -6.322 2.59e-10 ***
## genus_nb    -0.07489    0.02103  -3.560 0.00037 ***
## total_reads_orig 0.11218    0.02159   5.197 2.02e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

7. DBD over time in the gut (HMP time series)

7.1. Polymorphism change in a focal species as a function of community diversity at the earlier time point

7.1.1. Polymorphism change in a focal species as a function of Shannon diversity at the earlier time point

```
##
```



```

## Family: gaussian
## Link function: identity
##
## Formula:
## delta_polymorphism ~ s(total_reads_orig1) + s(alpha_div_tp1) +
##   s(species, alpha_div_tp1, bs = "fs") + s(sample_tp1, bs = "re") +
##   s(subject, bs = "re")
##
## Parametric coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.997e-05 1.690e-04  0.118  0.906
##
## Approximate significance of smooth terms:
##           edf Ref.df      F p-value
## s(total_reads_orig1)  1.000e+00  1.001 1.212 0.271004
## s(alpha_div_tp1)      2.178e+00  2.615 1.306 0.404512
## s(species,alpha_div_tp1) 1.456e-04 591.000 0.000 0.990101
## s(sample_tp1)         5.046e+01 215.000 0.329 0.000324 ***
## s(subject)            3.025e-04 159.000 0.000 0.801293
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) = 0.0243  Deviance explained = 4.14%
## GCV = 6.4579e-05  Scale est. = 6.3427e-05  n = 3063

```

7.1.2. Polymorphism change in a focal species as a function of species richness at the earlier time point

```

##
## Family: gaussian
## Link function: identity
##
## Formula:
## delta_polymorphism ~ s(total_reads_orig1) + s(richness_tp1) +
##   s(species, richness_tp1, bs = "fs") + s(sample_tp1, bs = "re") +
##   s(subject, bs = "re")
##
## Parametric coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept) 7.741e-06 1.707e-04  0.045  0.964
##
## Approximate significance of smooth terms:
##           edf Ref.df      F p-value
## s(total_reads_orig1)  1.000e+00  1.000 1.582 0.208529
## s(richness_tp1)      1.362e+00  1.586 0.220 0.640216
## s(species,richness_tp1) 1.621e-04 583.000 0.000 0.999852
## s(sample_tp1)         5.385e+01 215.000 0.356 0.000185 ***
## s(subject)            1.338e-04 159.000 0.000 0.821320
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) = 0.0249  Deviance explained = 4.28%
## GCV = 6.4594e-05  Scale est. = 6.3388e-05  n = 3063

```

7.1.3. Polymorphism change in a focal species as a function of the rarefied species richness at the earlier time point

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## delta_polymorphism ~ s(richness_rare_tp1) + s(species, richness_rare_tp1,
##   bs = "fs") + s(sample_tp1, bs = "re") + s(subject, bs = "re")
##
## Parametric coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.808e-06  1.713e-04  -0.016   0.987
##
## Approximate significance of smooth terms:
##               edf Ref.df    F  p-value
## s(richness_rare_tp1)      1.000e+00     1 0.462 0.496749
## s(species,richness_rare_tp1) 6.756e+00   512 0.016 0.187483
## s(sample_tp1)             5.544e+01   216 0.368 0.000135 ***
## s(subject)                1.109e-04   159 0.000 0.839634
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.0278  Deviance explained = 4.79%
## GCV = 6.4553e-05  Scale est. = 6.32e-05  n = 3063
```

7.2. Gene gain in a focal species as a function of community diversity at the earlier time point

7.2.1. Gene gain in a focal species as a function of Shannon diversity at the earlier time point

```
## Family: nbinom1 ( log )
## Formula:
## num_gene_gains ~ alpha_div_tp1 + total_reads_orig1 + (alpha_div_tp1 |
##   species) + (1 | subject) + (1 | sample_tp1)
## Data: datsc
##
##      AIC      BIC  logLik deviance df.resid
## 1516.0 1562.5  -749.0  1498.0    1287
##
## Random effects:
##
## Conditional model:
##   Groups      Name      Variance Std.Dev. Corr
## species (Intercept) 0.28885  0.5374
##   alpha_div_tp1 0.08704  0.2950  -0.37
## subject (Intercept) 2.12132  1.4565
## sample_tp1 (Intercept) 1.00077  1.0004
## Number of obs: 1296, groups: species, 54; subject, 154; sample_tp1, 210
##
## Overdispersion parameter for nbinom1 family (): 121
##
```

```

## Conditional model:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -0.36000    0.41245  -0.873   0.383
## alpha_div_tp1    0.07365    0.21687   0.340   0.734
## total_reads_orig1 -0.08165    0.23305  -0.350   0.726

```

7.2.2. Gene gain in a focal species as a function of species richness at the earlier time point

```

## Family: nbinom1 ( log )
## Formula:
## num_gene_gains ~ richness_tp1 + total_reads_orig1 + (1 | species) +
##   (1 | subject) + (1 | sample_tp1)
## Data: datsc
##
##      AIC      BIC   logLik deviance df.resid
## 1513.1 1549.2  -749.5  1499.1    1289
##
## Random effects:
##
## Conditional model:
## Groups      Name          Variance Std.Dev.
## species     (Intercept) 0.2914  0.5398
## subject     (Intercept) 2.1331  1.4605
## sample_tp1 (Intercept) 0.8869  0.9418
## Number of obs: 1296, groups: species, 54; subject, 154; sample_tp1, 210
##
## Overdispersion parameter for nbinom1 family (): 123
##
## Conditional model:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -0.28956    0.39887  -0.726   0.468
## richness_tp1    0.09735    0.19348   0.503   0.615
## total_reads_orig1 -0.14904    0.24677  -0.604   0.546

```

7.2.3. Gene gain in a focal species as a function of rarefied richness at the earlier time point

```

## Family: nbinom1 ( log )
## Formula:
## num_gene_gains ~ richness_rare_tp1 + (1 | species) + (1 | subject) +
##   (1 | sample_tp1)
## Data: datsc
##
##      AIC      BIC   logLik deviance df.resid
## 1511.1 1542.1  -749.5  1499.1    1290
##
## Random effects:
##
## Conditional model:
## Groups      Name          Variance Std.Dev.
## species     (Intercept) 0.2925  0.5408
## subject     (Intercept) 2.1466  1.4651
## sample_tp1 (Intercept) 0.9015  0.9495
## Number of obs: 1296, groups: species, 54; subject, 154; sample_tp1, 210

```

```
##
## Overdispersion parameter for nbinom1 family (): 123
##
## Conditional model:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -0.2833    0.3971  -0.713   0.476
## richness_rare_tp1  0.1127    0.1691   0.666   0.505
```

7.3. Gene loss in a focal species as a function of community diversity at the earlier time point

7.3.1. Gene loss in a focal species as a function of Shannon diversity at the earlier time point

```
## Family: nbinom2 ( log )
## Formula:
## num_gene_losses ~ alpha_div_tp1 + total_reads_orig1 + (alpha_div_tp1 |
##   species) + (1 | subject) + (1 | sample_tp1)
## Data: datsc
##
##      AIC      BIC  logLik deviance df.resid
## 1696.2 1742.7  -839.1  1678.2    1287
##
## Random effects:
##
## Conditional model:
## Groups      Name          Variance Std.Dev. Corr
## species     (Intercept)    1.7511  1.3233
##             alpha_div_tp1  0.2719  0.5214  -0.57
## subject     (Intercept)    2.3094  1.5197
## sample_tp1  (Intercept)   10.3411  3.2158
## Number of obs: 1296, groups: species, 54; subject, 154; sample_tp1, 210
##
## Overdispersion parameter for nbinom2 family (): 0.111
##
## Conditional model:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -3.5455    0.8326  -4.259 2.06e-05 ***
## alpha_div_tp1    0.8360    0.3950   2.117  0.0343 *
## total_reads_orig1  0.8956    0.5299   1.690  0.0910 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

7.3.2. Gene loss in a focal species as a function of species richness at the earlier time point

```
## Family: nbinom2 ( log )
## Formula:
## num_gene_losses ~ richness_tp1 + total_reads_orig1 + (1 | species) +
##   (1 | subject) + (1 | sample_tp1)
## Data: datsc
##
##      AIC      BIC  logLik deviance df.resid
## 1695.1 1731.2  -840.5  1681.1    1289
```

```

##
## Random effects:
##
## Conditional model:
## Groups      Name          Variance Std.Dev.
## species     (Intercept) 1.577    1.256
## subject     (Intercept) 3.432    1.852
## sample_tp1  (Intercept) 8.125    2.850
## Number of obs: 1296, groups: species, 54; subject, 154; sample_tp1, 210
##
## Overdispersion parameter for nbinom2 family (): 0.102
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -3.3062   0.7822  -4.227 2.37e-05 ***
## richness_tp1    0.7851   0.3751   2.093  0.0364 *
## total_reads_orig1 0.3357   0.5010   0.670  0.5028
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

7.3.3. Gene loss in a focal species as a function of rarefied richness at the earlier time point

```

## Family: nbinom2 ( log )
## Formula:
## num_gene_losses ~ richness_rare_tp1 + (1 | species) + (1 | subject)
## Data: datasc
##
##      AIC      BIC   logLik deviance df.resid
## 1695.8 1721.6  -842.9  1685.8    1291
##
## Random effects:
##
## Conditional model:
## Groups      Name          Variance Std.Dev.
## species     (Intercept) 1.655    1.287
## subject     (Intercept) 5.925    2.434
## Number of obs: 1296, groups: species, 54; subject, 154
##
## Overdispersion parameter for nbinom2 family (): 0.0626
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.9154   0.5366  -3.569 0.000358 ***
## richness_rare_tp1 0.6903   0.2652   2.603 0.009240 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

8. Black Queen Hypothesis

8.1. GAM with number of genes in a focal species as a function of community gene richness

```
##
## Family: Negative Binomial(126.741)
## Link function: log
##
## Formula:
## pres_genes ~ s(species_name, community_pres_genes, bs = "fs") +
##   s(sample_id, bs = "re") + s(subject_id, bs = "re") + s(community_pres_genes) +
##   s(total_reads_orig)
##
## Parametric coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  7.88488    0.03015   261.5  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##             edf   Ref.df   Chi.sq  p-value
## s(species_name,community_pres_genes) 250.672 1206.000 216769.89 < 2e-16 ***
## s(sample_id)                          1.034  465.000     1.00 0.633634
## s(subject_id)                          171.340 248.000   630.79 < 2e-16 ***
## s(community_pres_genes)                 2.515   3.004    28.40 2.92e-06 ***
## s(total_reads_orig)                     2.498   3.040    16.43 0.000964 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.936   Deviance explained = 94.4%
## -REML = 92585   Scale est. = 1           n = 13105
```

8.3. GAM with number of pathways in a focal species as a function of community pathway richness

```
##
## Family: poisson
## Link function: log
##
## Formula:
## pathway_richness ~ s(community_pathway_richness_set, species,
##   bs = "fs") + s(sample_id, bs = "re") + s(subject_id, bs = "re") +
##   s(community_pathway_richness_set) + s(total_reads_orig)
##
## Parametric coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   4.186      3.666   1.142   0.254
##
## Approximate significance of smooth terms:
##             edf   Ref.df   Chi.sq  p-value
## s(community_pathway_richness_set,species) 8.601e+02 1711.000 1.123e+11 <2e-16
```

```

## s(sample_id)                4.670e+02  467.000  8.541e+04  <2e-16
## s(subject_id)               1.994e-05  248.000  0.000e+00    1.00
## s(community_pathway_richness_set)  8.452e+00    8.689  4.039e+03  <2e-16
## s(total_reads_orig)         1.000e+00    1.000  8.000e-03    0.93
##
## s(community_pathway_richness_set,species) ***
## s(sample_id)                ***
## s(subject_id)                ***
## s(community_pathway_richness_set) ***
## s(total_reads_orig)
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.954  Deviance explained = 95.5%
## UBRE = 0.011673  Scale est. = 1          n = 22832

```