

# Species and genetic diversity patterns show diverging responses to land use intensity in central European grasslands

*Reisch & Hartig, Supplementary Material S2*

*7/20/2017*

## Abstract

This document contains the results from the data analysis of Reisch & Hartig, Species and genetic diversity patterns show diverging responses to land use intensity in central European grasslands. The complete analysis, code and data is available at <https://github.com/florianhartig/Reisch-Hartig-2020>

## Contents

	1
<b>Analysis of alpha diversity</b>	<b>1</b>
Species alpha . . . . .	1
Genetic alpha . . . . .	3
<b>Analysis of beta diversity</b>	<b>8</b>
Species beta . . . . .	8
Genetic beta . . . . .	11
<b>Analysis of gamma diversity</b>	<b>13</b>

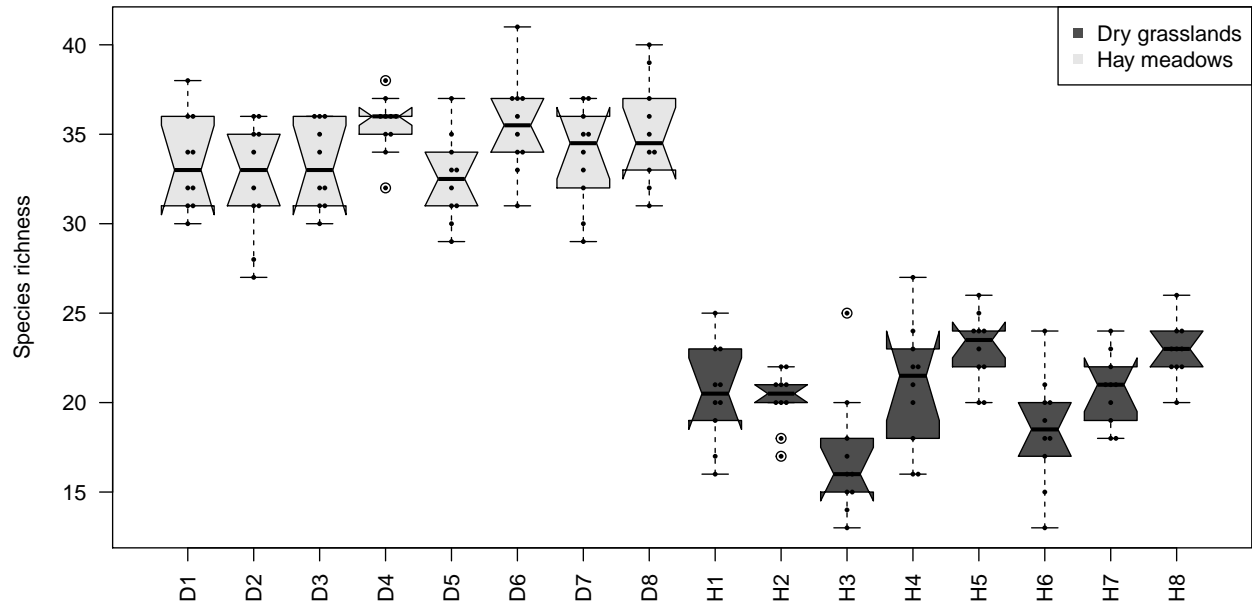
Notes: habitats are occasionally abbreviate with German words MR (Magerrasen) = dry grassland, FW (Fettwiese) = Hay Meadow.

## Analysis of alpha diversity

### Species alpha

Looking at species alpha diversity, we see a clear habitat pattern in which dry grasslands have higher local species richness (this is Fig. 2 in main paper)

```
## Warning in bxp(list(stats = structure(c(30, 31, 33, 36, 38, 27, 31, 33, :  
## some notches went outside hinges ('box'): maybe set notch=FALSE
```



Regression to test for the dependence of alpha on all predictors confirms the effect of habitat

```
## boundary (singular) fit: see ?isSingular

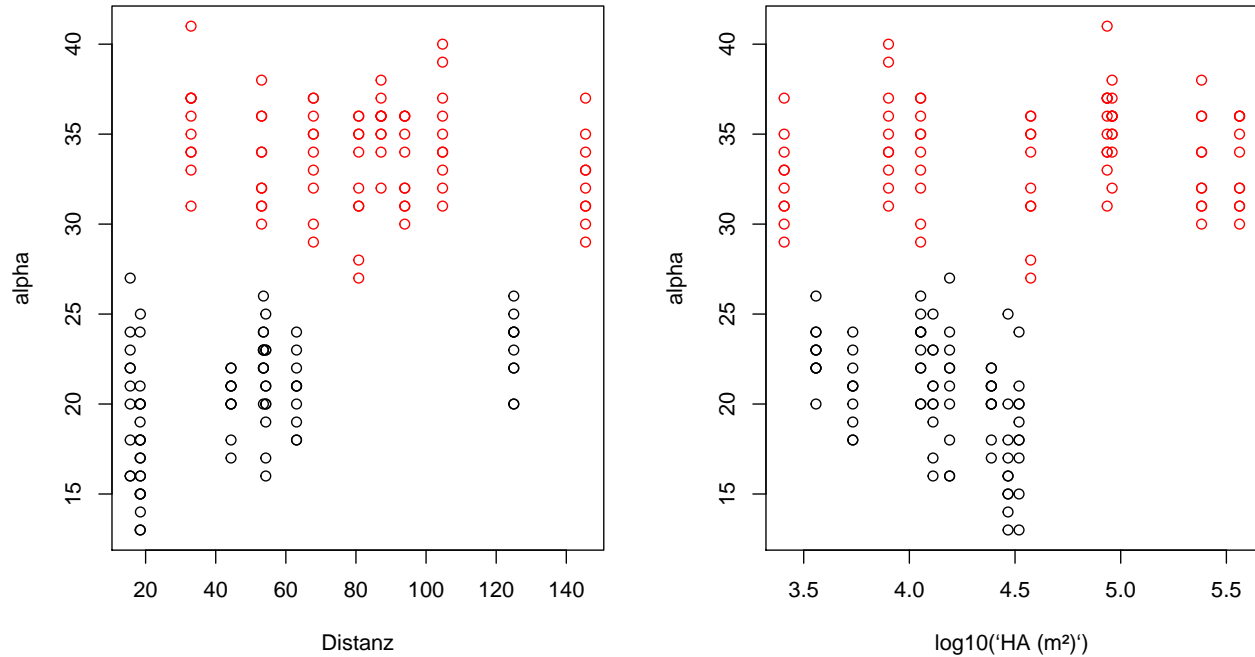
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: poisson ( log )
## Formula: alpha ~ Habitat + scale(Distanz) + scale(`HA (m²)`) + (1 | Plot)
## Data: sites$alpha2
##
##      AIC      BIC   logLik deviance df.resid
##    883.6    899.0   -436.8    873.6     155
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.58893 -0.39694  0.08864  0.37352  1.53683
##
## Random effects:
##  Groups Name         Variance Std.Dev.
##  Plot   (Intercept)  0          0
## Number of obs: 160, groups: Plot, 16
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    3.024937   0.027301 110.800   <2e-16 ***
## HabitatMagerrasen 0.492702   0.039113  12.597   <2e-16 ***
## scale(Distanz)    0.018518   0.017535   1.056    0.291
## scale(`HA (m²)`) -0.003164   0.015641  -0.202    0.840
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Warning in abbreviate(rn, minlength = 11): abbreviate used with non-ASCII
## chars
##
## Correlation of Fixed Effects:
```

```
## Warning in abbreviate(rn, minlength = 6): abbreviate used with non-ASCII
## chars

##          (Intr) HbttMg scl(D)
## HabttMgrrsn -0.827
## scal(Dstnz)  0.340 -0.487
## sc(`HA(²)` )  0.319 -0.444  0.207
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

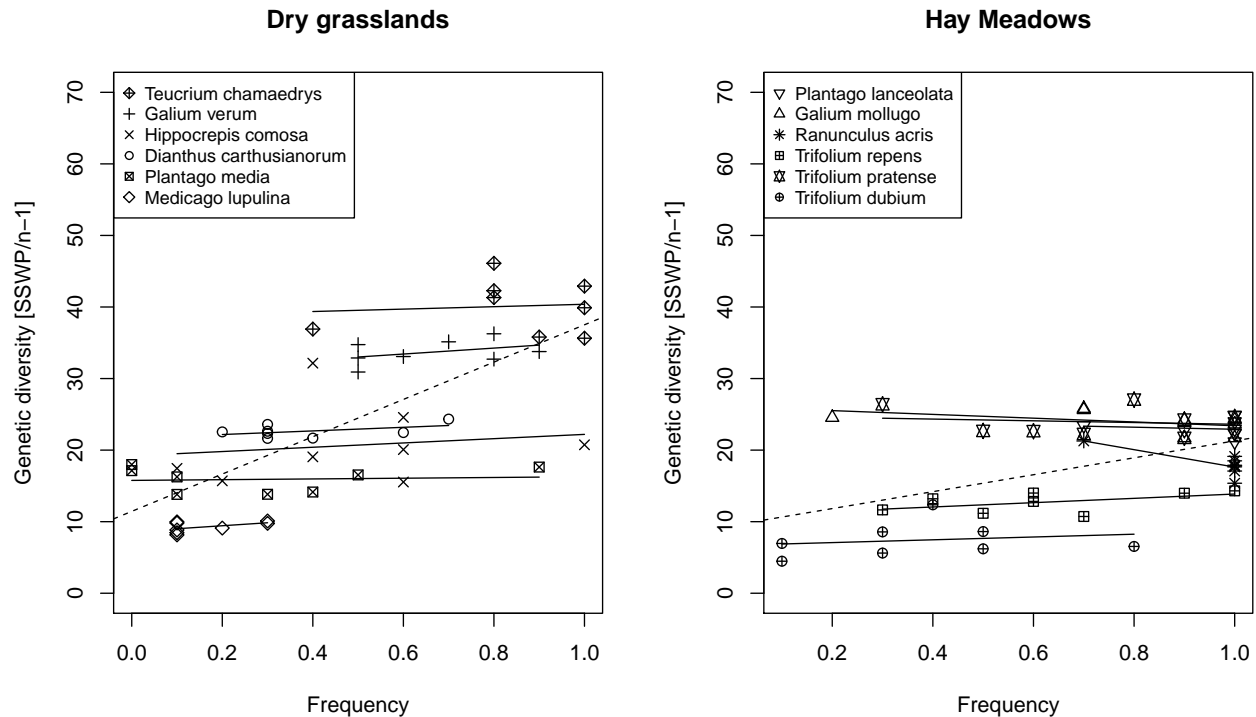
We did some further tests to explore possible nonlinear confounding with other predictors, but didn't see an issue that would question these results



Our conclusion the main factor associated with species alpha is habitat (dry grassland has higher tax alpha).

## Genetic alpha

We consider genetic alpha diversity (measured in SSWP/n-1) for the 6 species each in the dry and mesic grasslands against the frequency (FR) of subplots in which the species was found. The dashed line in each plot shows a linear regression estimate across all species in a given habitat, the solid lines show linear regression estimate per species.



The comparison suggests that genetic diversity differs little within species, and apparent correlations between FR and genetic diversity occur because of a confounding of FR with species. To separate these effects, we calculated linear mixed models, with population and species as random intercepts.

```
## Linear mixed model fit by maximum likelihood . t-tests use
## Satterthwaite's method [lmerModLmerTest]
## Formula:
## `SSWP/n-1` ~ Habitat * scale(FR) + scale(`HA (m²)`) + scale(Distanz) +
## (1 | Population) + (1 | Art)
## Data: sites$genAlpha
##
##      AIC      BIC   logLik deviance df.resid
##    511.2    534.3   -246.6   493.2     87
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.3640 -0.5007 -0.0612  0.4707  4.6340
##
## Random effects:
##  Groups      Name                Variance Std.Dev.
##  Population (Intercept)  0.2302   0.4798
##  Art          (Intercept) 66.9680   8.1834
##  Residual                    5.4261   2.3294
## Number of obs: 96, groups: Population, 16; Art, 12
##
## Fixed effects:
##              Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)    18.4747    3.3823  12.0918   5.462 0.000141
## HabitatMagerrasen    5.4216    4.7906  12.1681   1.132 0.279569
## scale(FR)         -0.1911    0.5982  85.4007  -0.319 0.750224
## scale(`HA (m²)`)    0.2236    0.3066  13.2769   0.729 0.478428
```

```

## scale(Distanz)          0.2373      0.3246 15.7950   0.731 0.475424
## HabitatMagerrasen:scale(FR) 1.1369      0.8206 83.1458   1.386 0.169593
##
## (Intercept)            ***
## HabitatMagerrasen
## scale(FR)
## scale(`HA (m²)`)
## scale(Distanz)
## HabitatMagerrasen:scale(FR)
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Warning in abbreviate(rn, minlength = 11): abbreviate used with non-ASCII
## chars

##
## Correlation of Fixed Effects:

## Warning in abbreviate(rn, minlength = 6): abbreviate used with non-ASCII
## chars

##          (Intr) HbttMg sc(FR) s(`HA( scl(D)
## HbttMgrrsn -0.710
## scale(FR)   -0.089  0.072
## sc(`HA(²)`) 0.054 -0.073 -0.053
## scal(Dstnz) 0.075 -0.085 -0.304  0.200
## HbttMg:(FR) 0.069 -0.027 -0.750  0.042  0.291

Which showed that none of the factors are significant. Habitat is also not significant in a single regression, as
long as species identity is included as a random intercept.

## Linear mixed model fit by maximum likelihood . t-tests use
## Satterthwaite's method [lmerModLmerTest]
## Formula: `SSWP/n-1` ~ Habitat + (1 | Population) + (1 | Art)
## Data: sites$genAlpha
##
##      AIC      BIC    logLik deviance df.resid
##    506.8    519.6   -248.4    496.8      91
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.1299 -0.5121 -0.0549  0.4621  4.7175
##
## Random effects:
## Groups      Name      Variance Std.Dev.
## Population (Intercept) 0.1304  0.3612
## Art      (Intercept) 72.8985  8.5381
## Residual              5.6877  2.3849
## Number of obs: 96, groups: Population, 16; Art, 12
##
## Fixed effects:
##              Estimate Std. Error    df t value Pr(>|t|)
## (Intercept)      18.186      3.505 12.030   5.189 0.000224 ***
## HabitatMagerrasen   5.530      4.957 12.030   1.116 0.286378
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##

```

```
## Correlation of Fixed Effects:
##           (Intr)
## HabttMgrsrn -0.707
```

However, species is obviously a significant predictor for both genetic diversity. The problem is that species is perfectly collinear with habitat, and closely colinear with FR, so that it's difficult to separate these effects.

```
## Type III Analysis of Variance Table with Satterthwaite's method
##      Sum Sq Mean Sq NumDF  DenDF F value    Pr(>F)
## Art 7709.1  700.82    11 75.813  123.25 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Both habitat and frequency would be significant when not controlling for species identity, but given the previous results, and the apparent collinearity between species identity and other factors, this seems statistically not sensible.

```
## boundary (singular) fit: see ?isSingular

## Linear mixed model fit by maximum likelihood . t-tests use
## Satterthwaite's method [lmerModLmerTest]
## Formula:
## `SSWP/n-1` ~ Habitat * scale(FR) + scale(`HA (m²)`) + scale(Distanz) +
##      (1 | Population)
## Data: sites$genAlpha
##
##      AIC      BIC    logLik deviance df.resid
##    647.2    667.7   -315.6    631.2      88
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.66122 -0.70808  0.02347  0.63996  2.32656
##
## Random effects:
## Groups      Name      Variance Std.Dev.
## Population (Intercept)  0.00    0.00
## Residual              41.99    6.48
## Number of obs: 96, groups: Population, 16
##
## Fixed effects:
##              Estimate Std. Error      df t value
## (Intercept)    16.669891   1.212008  96.000000   13.754
## HabitatMagerrasen  10.461055   1.855476  96.000000    5.638
## scale(FR)         3.816992   1.103762  96.000000    3.458
## scale(`HA (m²)`)   0.130509   0.761002  96.000000    0.171
## scale(Distanz)     0.004005   0.787422  96.000000    0.005
## HabitatMagerrasen:scale(FR)  4.598303   1.502450  96.000000    3.061
## Pr(>|t|)
## (Intercept)      < 2e-16 ***
## HabitatMagerrasen  1.73e-07 ***
## scale(FR)         0.000812 ***
## scale(`HA (m²)`)   0.864195
## scale(Distanz)     0.995952
## HabitatMagerrasen:scale(FR) 0.002865 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Warning in abbreviate(rn, minlength = 11): abbreviate used with non-ASCII
## chars

##
## Correlation of Fixed Effects:

## Warning in abbreviate(rn, minlength = 6): abbreviate used with non-ASCII
## chars

##              (Intr) HbttMg sc(FR) s(`HA( scl(D)
## HabttMgrrsn -0.799
## scale(FR)    -0.456  0.346
## sc(`HA(²)`    0.357 -0.456 -0.039
## scal(Dstnz)   0.441 -0.503 -0.230  0.197
## HbttMg:(FR)   0.355 -0.130 -0.746  0.031  0.220
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

In fact, it seems that both the habitat effects and the frequency effects in the latter regression are essentially driven by species identity, in that some species simply have a higher diversity, or at least there are not enough species to separate between an effect of species identity, and true FR / Habitat effects.

To confirm that intuition, we additionally calculated a t-test between the average genetic diversity values per species for each habitat (n=6 for each group)

```
##
## Welch Two Sample t-test
##
## data:  SSWP/n-1 by Habitat
## t = -1.0192, df = 8.2069, p-value = 0.3372
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -17.986717  6.927195
## sample estimates:
## mean in group Fettwiese mean in group Magerrasen
##                18.18584                23.71560
```

The results show that habitat effects are n.s. when comparing only mean genetic diversity per species between habitat (i.e. n=6 per group).

Moreover, the colinearity between FR and Species is created by the fact that the species with higher FR range also seem to have higher genetic diversity, although there is no significant correlation between FR and genetic diversity within species.

To test if this relationship is significant, we correlated mean FR and mean genetic diversity per species

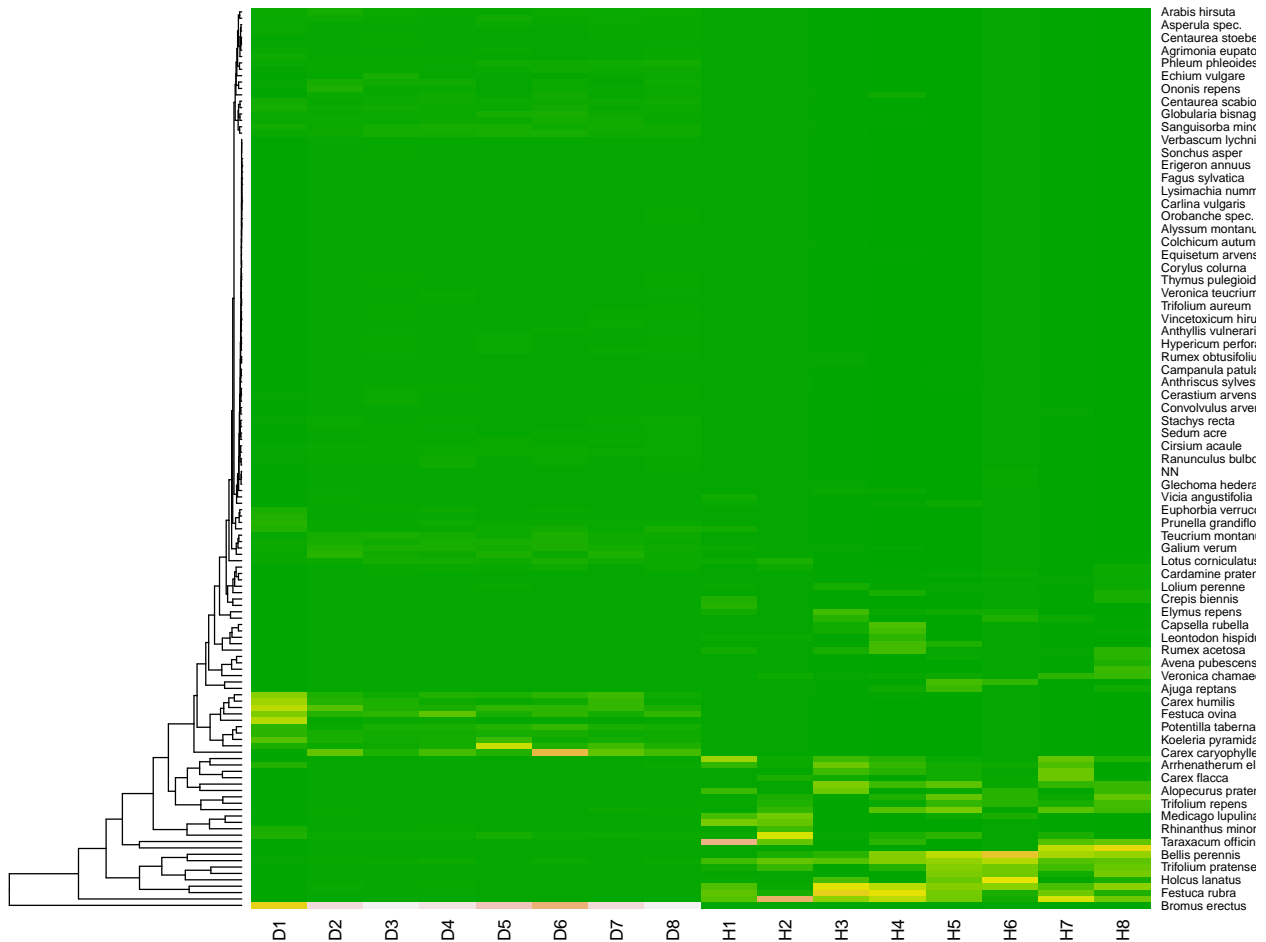
```
##
## Pearson's product-moment correlation
##
## data:  x1$`SSWP/n-1` and x2$FR
## t = 2.1086, df = 10, p-value = 0.06119
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.02807041  0.85610208
## sample estimates:
##      cor
## 0.5547683
```

The relationship was again n.s. However, given the small sample size (n=12), and the general plausibility of such an effect, we would encourage future studies to look again for such an effect.

## Analysis of beta diversity

### Species beta

Visualization of the abundance of site / specioes, clustered by species



Test for genetic beta via multivariate homogeneity of group dispersions (variances)

```
##
## Homogeneity of multivariate dispersions
##
## Call: betadisper(d = distSpecies, group = sites$alpha$Habitat)
##
## No. of Positive Eigenvalues: 15
## No. of Negative Eigenvalues: 0
##
## Average distance to median:
## Fettwiese Magerrasen
##      0.3756      0.2435
##
```



```

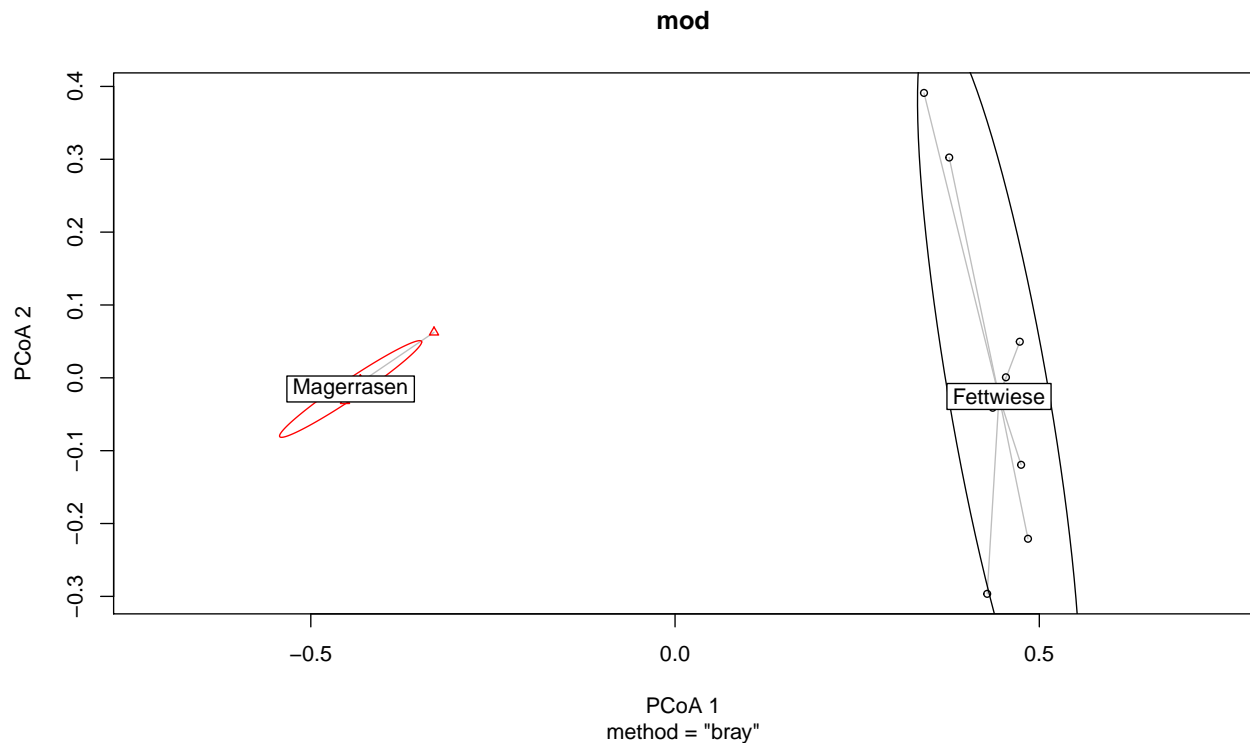
## Eigenvalues for PCoA axes:
## (Showing 8 of 15 eigenvalues)
##   PCoA1   PCoA2   PCoA3   PCoA4   PCoA5   PCoA6   PCoA7   PCoA8
## 3.04065 0.40635 0.31489 0.23082 0.20848 0.14773 0.12665 0.07148

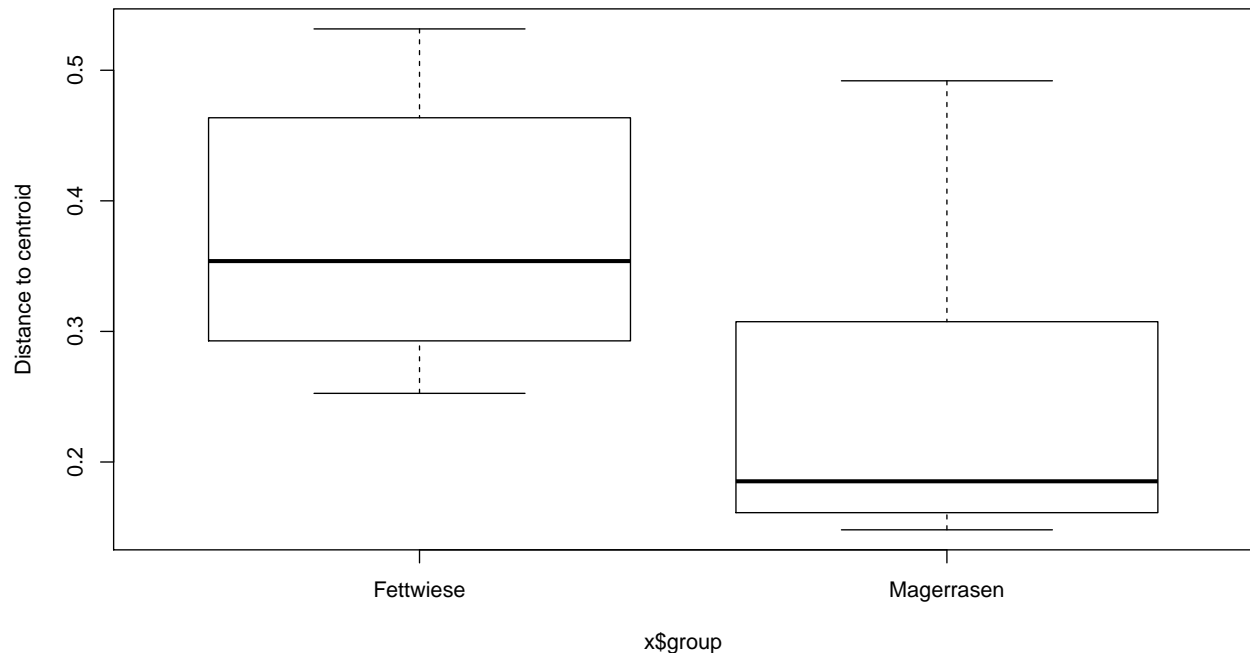
## Analysis of Variance Table
##
## Response: Distances
##           Df   Sum Sq Mean Sq F value Pr(>F)
## Groups      1 0.069824 0.069824  5.515 0.03407 *
## Residuals   14 0.177250 0.012661
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
## Permutation test for homogeneity of multivariate dispersions
## Permutation: free
## Number of permutations: 499
##
## Response: Distances
##           Df   Sum Sq Mean Sq    F N.Perm Pr(>F)
## Groups      1 0.069824 0.069824 5.515   499 0.028 *
## Residuals   14 0.177250 0.012661
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
## Pairwise comparisons:
## (Observed p-value below diagonal, permuted p-value above diagonal)
##           Fettwiese Magerrasen
## Fettwiese              0.024
## Magerrasen 0.034066

```





Additional Mantel test for spatial structure, separate for habitats

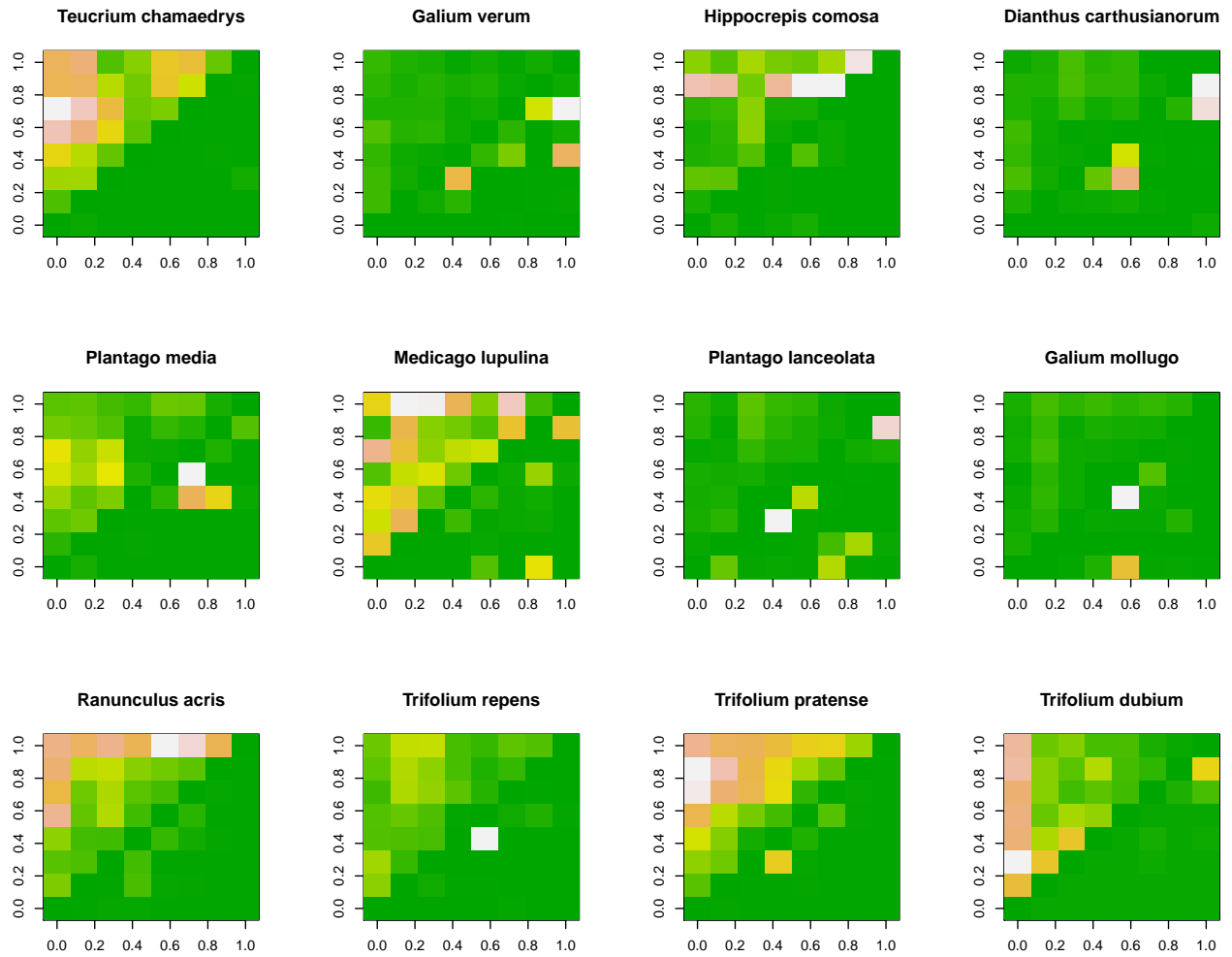
```
##
## Mantel statistic based on Pearson's product-moment correlation
##
## Call:
## mantel(xdis = as.matrix(distSpecies)[1:8, 1:8], ydis = sites$geoDist[1:8, 1:8], method = "pearson")
##
## Mantel statistic r: 0.3382
##      Significance: 0.103
##
## Upper quantiles of permutations (null model):
##   90%   95% 97.5%  99%
## 0.340 0.403 0.456 0.534
## Permutation: free
## Number of permutations: 999

##
## Mantel statistic based on Pearson's product-moment correlation
##
## Call:
## mantel(xdis = as.matrix(distSpecies)[9:16, 9:16], ydis = sites$geoDist[9:16, 9:16], method = "pearson")
##
## Mantel statistic r: 0.5542
##      Significance: 0.01
##
## Upper quantiles of permutations (null model):
##   90%   95% 97.5%  99%
## 0.278 0.365 0.459 0.521
## Permutation: free
## Number of permutations: 999
```

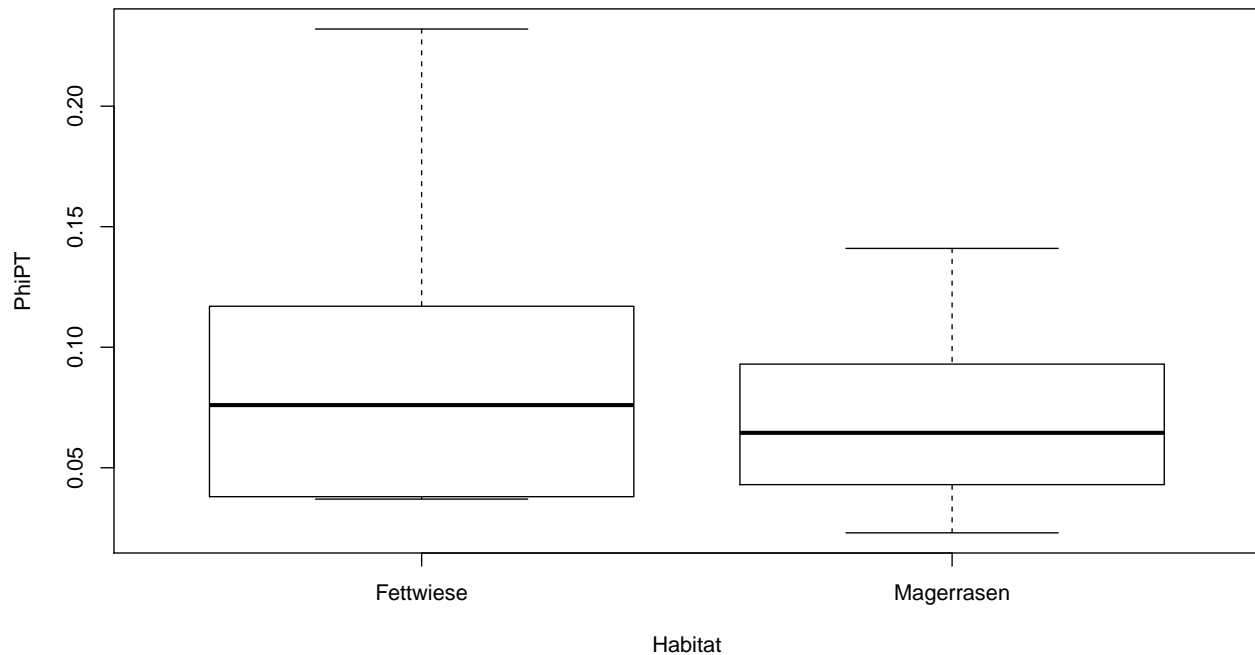
Result: FW not significant with geographic distance, but MR is probably (stronger driven by environment)

## Genetic beta

Genetic turnover for all species



Regression PhiPT gegen Habitat



```
##
## Call:
## lm(formula = PhiPT ~ Habitat, data = sites$genAlpha[c(1:6, 49:54),
##   ])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.05900 -0.03350 -0.01475  0.02113  0.13600
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.09600    0.02426   3.958  0.0027 **
## HabitatMagerrasen -0.02450    0.03430  -0.714  0.4914
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05941 on 10 degrees of freedom
## Multiple R-squared:  0.04854,    Adjusted R-squared:  -0.04661
## F-statistic: 0.5101 on 1 and 10 DF,  p-value: 0.4914
##
## Welch Two Sample t-test
##
## data:  sites$genAlpha$PhiPT[1:6] and sites$genAlpha$PhiPT[49:54]
## t = -0.71424, df = 7.9428, p-value = 0.4955
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.10370052  0.05470052
## sample estimates:
## mean of x mean of y
##    0.0715    0.0960
```

PhiPT slightly higher in MR

Mantel tests to test for spatial signal in genetic diversity

```
## [1] 0.160 0.192 0.030 0.020 0.003 0.041
```

```
## [1] 0.002 0.012 0.468 0.090 0.029 0.095
```

Result: again, FW are all n.s., but so is MR (only one species significant), potentially slightly higher spatial effects in MR

## Analysis of gamma diversity

We can only consider species gamma, there is no measure of genetic gamma diversity.

```
countsFW = apply(sites$community[1:8,],2, sum)
countsMR = apply(sites$community[9:16,],2, sum)
sum(countsFW > 0)
```

```
## [1] 92
```

```
sum(countsMR > 0)
```

```
## [1] 70
```

Result: FW has higher diversity