

# Environmental filtering and habitat (mis)matching of riverine invertebrate communities

NRSA DisEQ-Figures

## Contents

<b>Site Map</b>	<b>4</b>
<b>Filtering &amp; Habitat Matching ANOVAs</b>	<b>6</b>
<b>Habitat Matching &amp; Mismatch Vector Component ANOVAs</b>	<b>8</b>
<b>Functional Trait Diversity ANOVAs</b>	<b>10</b>
<b>DisEQ-by-Traits</b>	<b>12</b>
Filtering-by-Traits Figure . . . . .	12
Mismatch-by-Traits Figure . . . . .	14
<b>Trait-by-Environment</b>	<b>16</b>
<b>DisEQ &amp; Functional Diversity ANCOVAs</b>	<b>19</b>
Filtering by Functional Diversity . . . . .	19
Habitat Matching by Functional Diversity . . . . .	20
<b>Environmental Filtering &amp; Habitat Matching Model</b>	<b>23</b>
<b>Community Composition</b>	<b>25</b>
<b>R Session Information</b>	<b>27</b>

```
## Reorder trait variables for plotting
```

```
trait.axis.limits <- c(  
  "dispersal.low", "dispersal.high",  
  "flying.strength.none", "flying.strength.weak", "flying.strength.strong",  
  "size.small", "size.medium", "size.large",  
  "depositional", "depositional.erosional", "erosional",  
  "cold", "cool.warm", "warm",  
  "sensitive", "medium", "tolerant",  
  "CG", "CF", "HB", "PR"  
)
```

```
## Rename trait variables for plotting
```

```
trait.axis.labels <- c(  
  "dispersal.low" = "Low Dispersal", "dispersal.high" = "High Dispersal",  
  "flying.strength.none" = "Nonflyer", "flying.strength.weak" = "Weak Flyer",  
  "flying.strength.strong" = "Strong Flyer",  
  "size.small" = "Small Size", "size.medium" = "Medium Size", "size.large" = "Large Size",  
  "depositional" = "Depositional", "depositional.erosional" = "Depo-Eros", "erosional" = "Erosional",  
  "cold" = "Cold", "cool.warm" = "Cool-Warm", "warm" = "Warm",  
  "sensitive" = "Sensitive", "medium" = "Intermediate", "tolerant" = "Tolerant",  
  "CG" = "CG", "CF" = "CF", "HB" = "HB", "PR" = "PR"  
)
```

```
## Ecoregion labels (remove acronyms)
```

```
ecoregion.labels <- c(  
  "CPL" = "Coastal Plain", "NAP" = "Northern Appalachians", "NPL" = "Northern Plains",  
  "SAP" = "Southern Appalachians", "SPL" = "Southern Plains", "TPL" = "Temperate Plains",  
  "UMW" = "Upper Midwest", "WMT" = "Western Mountains", "XER" = "Xeric"  
)
```

```

## Ecoregion-----#
## Set colour palette (viridis)
viridis.continuous <- viridis(n = 9)

# View the colour palette
#show_col(viridis.continuous)

# Set the ecoregion palette
viridis.ecoregion <- viridis.continuous[c(9, 7, 3, 8, 4, 5, 6, 1, 2)]

# Alphabetical --> Geographical
# CPL (1-9), NAP (2-7), NPL (3-3), SAP (4-8), SPL (5-4), TPL (6-5), UMW (7-6), WMT (8-1), XER (9-2)

# Reordered
# WMT, XER, NPL, SPL, TPL, UMW, NAP, SAP, CPL

## Functional Traits-----#
## Set colour palette (magma)
magma.continuous <- magma(n = 16)

# View the colour palette
#show_col(magma.continuous)

# Set the functional traits palette
magma.functional.traits <- magma.continuous[c(3, 8, 13)]

## Environmental Predictors-----#
## Set colour palette (plasma)
plasma.continuous <- plasma(n = 16)

# View the colour palette
#show_col(plasma.continuous)

# Set the environ palette
plasma.environmental.predictors <- plasma.continuous[c(1, 9, 14)]

```

## Site Map

```
## Figure was created by evaluating the supplementary R code and data
## provided by King et al. 2019 (Ecological Applications), deposited by
## Katelyn King on Zenodo at:

# https://zenodo.org/record/3246537#.XgDUKRdKhSw

## Reorder ecoregions for plotting
final.data$ecoregion <- fct_relevel(
  final.data$ecoregion,
  levels = c("WMT", "XER", "NPL", "SPL", "TPL", "UMW", "NAP", "SAP", "CPL")
)

## Set base projection and site locations
usa <- map_data("usa") # pull out the USA map
base.map <- ggplot(data = usa) +
  geom_polygon(aes(x = long, y = lat, group = group),
    fill = "white", color = "black") +
  coord_fixed(1.3)
points <- select(final.data, site.long, site.lat, ecoregion)

## Add points to the US map
site.map <- base.map +
  geom_point(data = points, size = 2,
    aes(x = site.long, y = site.lat, colour = ecoregion, shape = ecoregion)) +
  scale_colour_manual(values = viridis.ecoregion, name = "Ecoregion") +
  scale_shape_manual(values = c(0:8), name = "Ecoregion") +
  north(data = usa, symbol = 3, scale = 0.1, location = "bottomright",
    anchor = c(x = -120, y = 27)) +
  ggsn::scalebar(data = usa, dist = 500, dist_unit = "km", transform = TRUE, model = "WGS84",
    st.size = 2, location = "bottomleft") +
  theme(panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(),
    axis.line.x = element_blank(),
    axis.title.x = element_blank(),
    axis.ticks.x = element_blank(),
    axis.text.x = element_blank(),
    axis.line.y = element_blank(),
    axis.title.y = element_blank(),
    axis.ticks.y = element_blank(),
    axis.text.y = element_blank(),
    panel.background = element_blank()) +
  theme(legend.position = c(0.88, 0.20), legend.text = element_text(size = 9)) +
  theme(legend.text = element_text(size = 9))
```

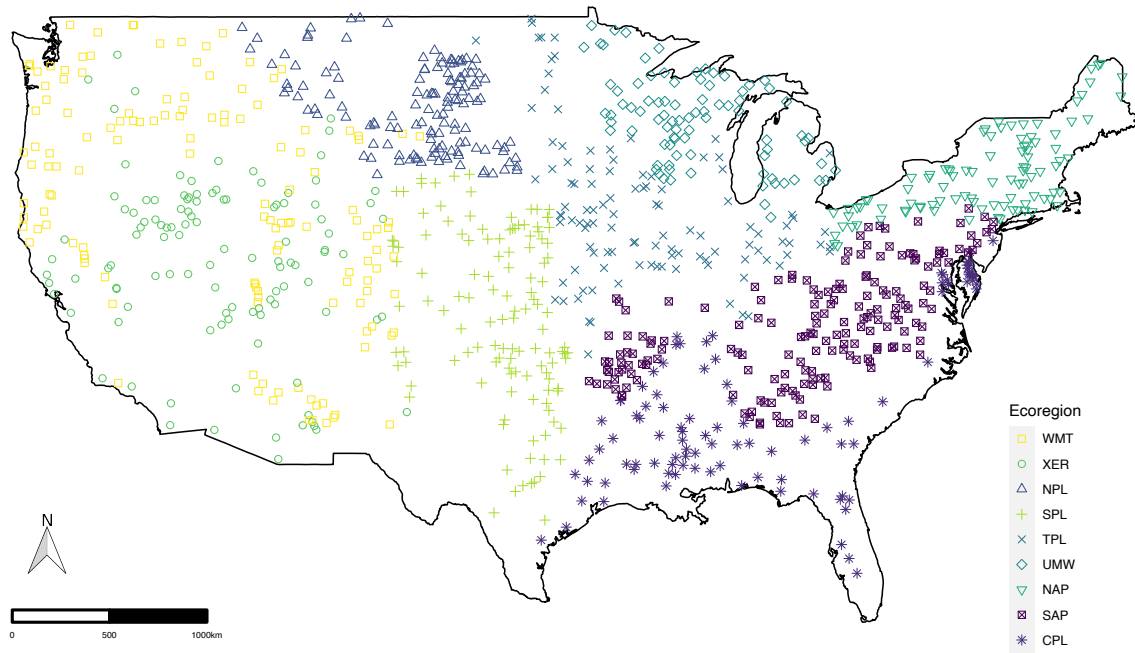


Figure 1: Map of all 1078 sites included in the study.

## Filtering & Habitat Matching ANOVAs

```
## Filtering and habitat matching data management for plotting
```

```
DisEQ.data.wide <- final.DisEQ.data %>%  
  select(ecoregion, filtering.scaled, mismatch.scaled)
```

```
DisEQ.data.long <- DisEQ.data.wide %>%  
  pivot_longer(!ecoregion, names_to = "DisEQ.metric", values_to = "measurement") %>%  
  na.omit()
```

```
DisEQ.figure <- ggerrorplot(  
  data = DisEQ.data.long,  
  x = "ecoregion",  
  y = "measurement",  
  size = 1.5,  
  width = 1.5,  
  desc_stat = "mean",  
  color = "ecoregion",  
  xlab = "Ecoregion",  
  ylab = "Measurement",  
  title = "DisEQ by Ecoregion",  
  palette = viridis.ecoregion,  
  add = "boxplot",  
  ggtheme = theme_pubr(),  
  legend = "right"  
) %>%  
facet(  
  facet.by = "DisEQ.metric",  
  nrow = 2,  
  ncol = 1,  
  scales = "free_y"  
) +  
scale_x_discrete(limits = c("WMT", "XER", "NPL", "SPL", "TPL", "UMW", "NAP", "SAP", "CPL"),  
  labels = ecoregion.labels) +  
rotate_x_text(angle = 45) +  
geom_hline(yintercept = 0, linetype = 3, size = 1, colour = "gray20") +  
font("xlab", size = 16, color = "gray0") +  
font("ylab", size = 16, color = "gray0") +  
font("xy.text", size = 12, color = "gray20")
```

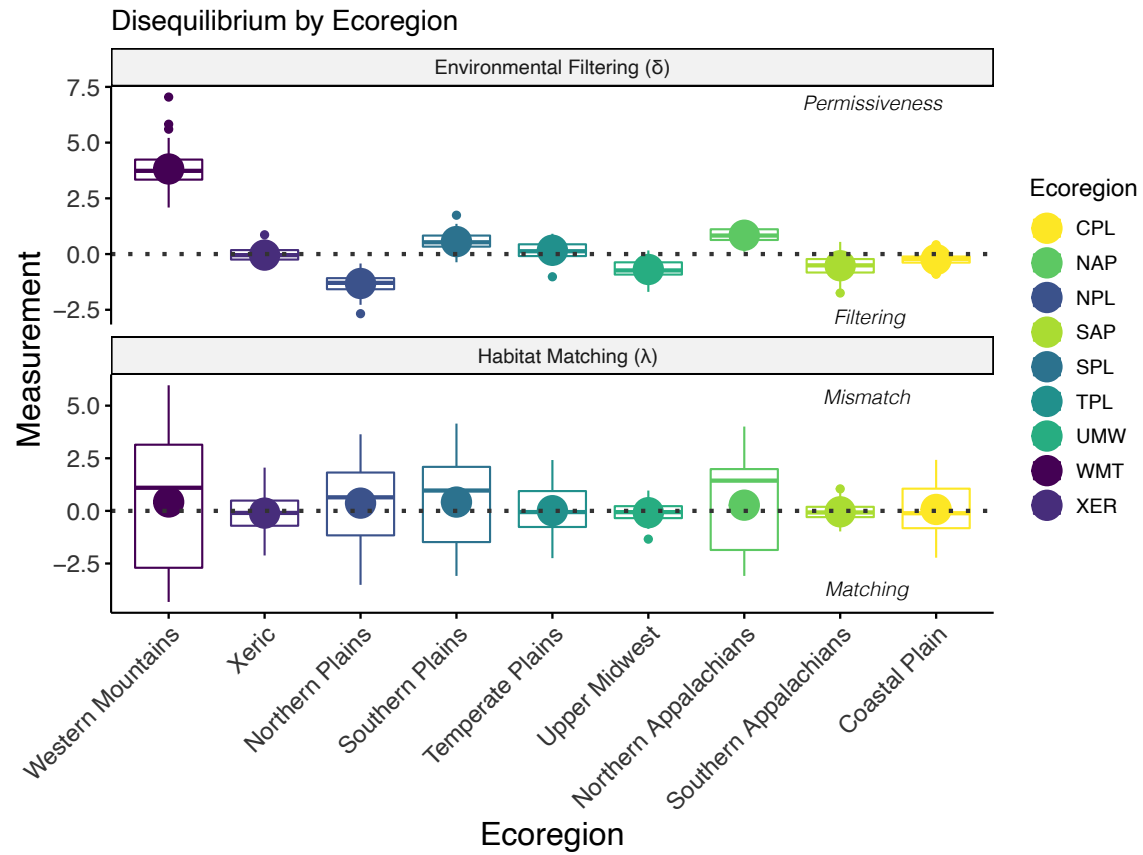


Figure 2: Plots of environmental filtering and habitat matching by ecoregion.

## Habitat Matching & Mismatch Vector Component ANOVAs

```
## Vector component data management for plotting
vector.component.wide <- final.DisEQ.data %>%
  select(ecoregion, Tmax.direction, Tmin.direction, pH.direction, cond.direction)

vector.component.long <- vector.component.wide %>%
  pivot_longer(!ecoregion, names_to = "vector.component", values_to = "measurement") %>%
  na.omit()

## Re-order variables for the facet plots
vector.component.long$vector.component <- fct_relevel(
  vector.component.long$vector.component,
  levels = c("Tmax.direction", "Tmin.direction", "pH.direction", "cond.direction")
)

vector.component.figure <- ggerrorplot(
  data = vector.component.long,
  x = "ecoregion",
  y = "measurement",
  size = 1.5,
  width = 1.5,
  desc_stat = "mean",
  color = "ecoregion",
  xlab = "Ecoregion",
  ylab = "Measurement",
  title = "Vector Components",
  palette = viridis.ecoregion,
  add = "boxplot",
  ggtheme = theme_pubr(),
  legend = "right") %>%
  facet(
    facet.by = "vector.component",
    nrow = 2,
    ncol = 2,
    scales = "free_y"
  ) +
  scale_x_discrete(limits = c("WMT", "XER", "NPL", "SPL", "TPL", "UMW", "NAP", "SAP", "CPL"),
    labels = ecoregion.labels) +
  rotate_x_text(angle = 60) +
  font("xlab", size = 16, color = "gray0") +
  font("ylab", size = 16, color = "gray0") +
  font("xy.text", size = 12, color = "gray20")
```



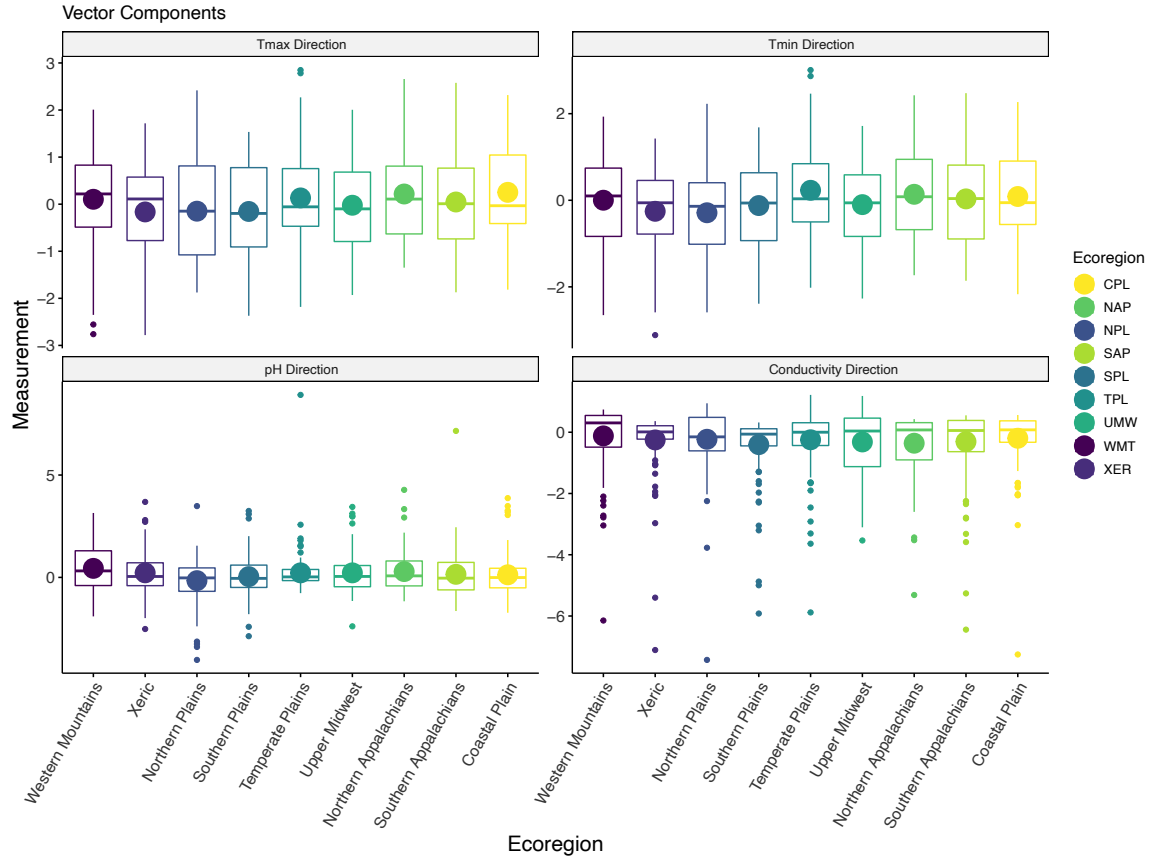


Figure 3: Plots of vector components of habitat matching and mismatch.

## Functional Trait Diversity ANOVAs

```
## Trait diversity data management for plotting
trait.diversity.wide <- final.DisEQ.data %>%
  select(ecoregion, FRic, FEve, FDiv, FDis)

trait.diversity.long <- trait.diversity.wide%>%
  pivot_longer(!ecoregion, names_to = "diversity.measure", values_to = "measurement") %>%
  na.omit()

trait.diversity.figure <- ggerrorplot(
  data = trait.diversity.long,
  x = "ecoregion",
  y = "measurement",
  size = 1.5,
  width = 1.5,
  desc_stat = "mean",
  color = "ecoregion",
  xlab = "Ecoregion",
  ylab = "Measurement",
  title = "Trait Diversity",
  palette = viridis.ecoregion,
  add = "boxplot",
  ggtheme = theme_pubr(),
  legend = "right"
) %>%
facet(
  facet.by = "diversity.measure",
  nrow = 2,
  ncol = 2,
  scales = "free_y"
) +
scale_x_discrete(limits = c("WMT", "XER", "NPL", "SPL", "TPL", "UMW", "NAP", "SAP", "CPL"),
  labels = ecoregion.labels) +
rotate_x_text(angle = 60) +
font("xlab", size = 16, color = "gray0") +
font("ylab", size = 16, color = "gray0") +
font("xy.text", size = 12, color = "gray20")
```

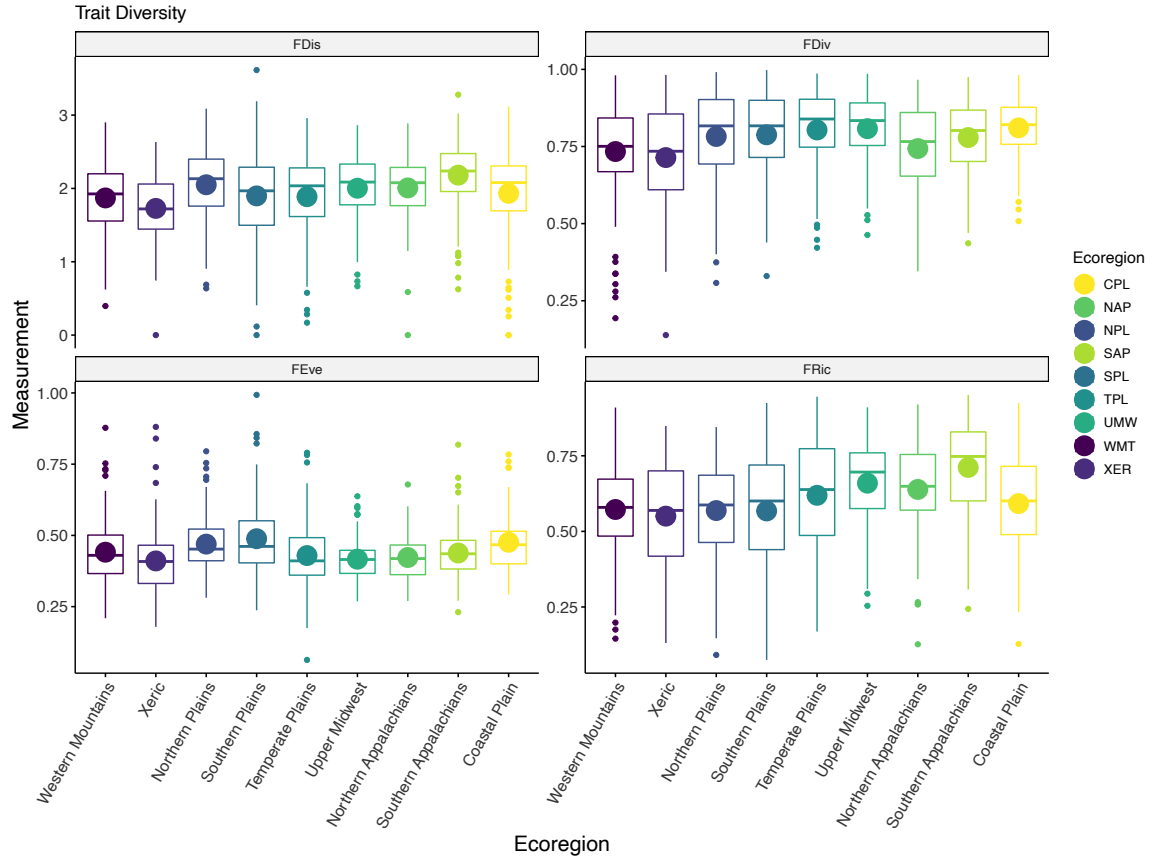


Figure 4: Plots of FRic, FEve, FDiv, and FDis by ecoregion.

## DisEQ-by-Traits

### Filtering-by-Traits Figure

```
## Re-order ecoregion for facet plots
filtering.by.trait.BRT.results$ecoregion <- fct_relevel(
  filtering.by.trait.BRT.results$ecoregion,
  levels = c("WMT", "UMW", "NAP", "XER", "NPL", "SAP", "SPL", "TPL", "CPL")
)

filtering.by.traits.figure <- ggbarplot(
  data = filtering.by.trait.BRT.results,
  x = "var",
  y = "rel.inf",
  fill = "predictor.type",
  xlab = "Trait",
  ylab = "Relative Influence",
  title = "Filtering by Traits",
  facet.by = "ecoregion",
  palette = magma.functional.traits,
  ggtheme = theme_pubr(),
  legend = "right"
) +
  scale_x_discrete(
    limits = trait.axis.limits,
    labels = trait.axis.labels
  ) +
  scale_y_continuous(breaks = c(0, 5, 10), limits = c(0, 11.25)) +
  rotate_x_text(angle = 90) +
  geom_hline(yintercept = 5, linetype = 3, size = 1, colour = "gray20") +
  font("xlab", size = 16, color = "gray0") +
  font("ylab", size = 16, color = "gray0") +
  font("xy.text", size = 12, color = "gray20")
```

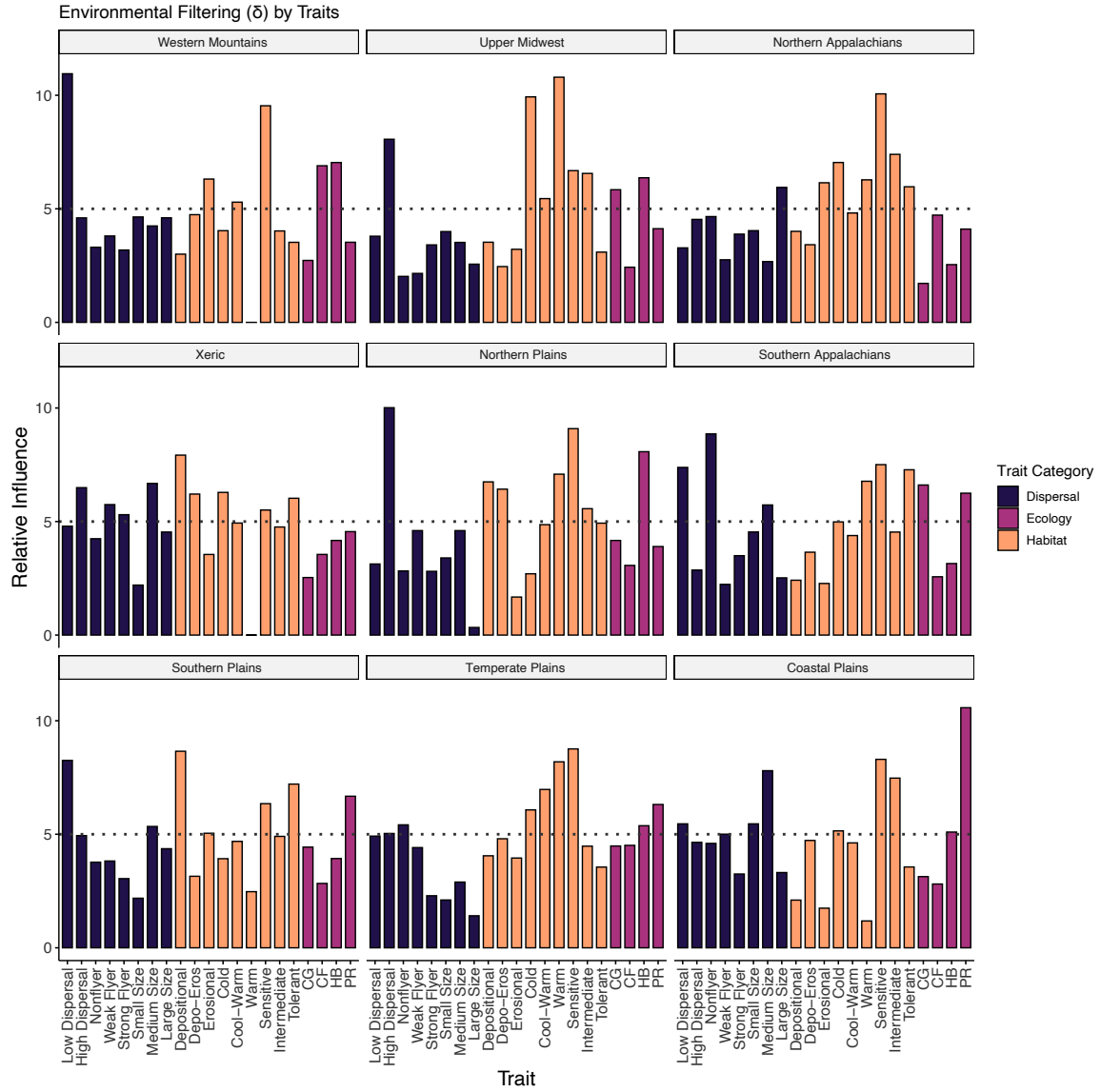


Figure 5: Plots of relative influence for each of 21 traits across ecoregions

## Mismatch-by-Traits Figure

```
## Re-order ecoregion for facet plots
mismatch.by.trait.BRT.results$ecoregion <- fct_relevel(
  mismatch.by.trait.BRT.results$ecoregion,
  levels = c("WMT", "UMW", "NAP", "XER", "NPL", "SAP", "SPL", "TPL", "CPL")
)

mismatch.by.traits.figure <- ggbarplot(
  data = mismatch.by.trait.BRT.results,
  x = "var",
  y = "rel.inf",
  fill = "predictor.type",
  xlab = "Trait",
  ylab = "Relative Influence",
  title = "Habitat Matching by Traits",
  facet.by = "ecoregion",
  palette = magma.functional.traits,
  ggtheme = theme_pubr(),
  legend = "right"
) +
scale_x_discrete(
  limits = trait.axis.limits,
  labels = trait.axis.labels) +
rotate_x_text(angle = 90) +
geom_hline(yintercept = 5, linetype = 3, size = 1, colour = "gray20") +
font("xlab", size = 16, color = "gray0") +
font("ylab", size = 16, color = "gray0") +
font("xy.text", size = 12, color = "gray20")
```

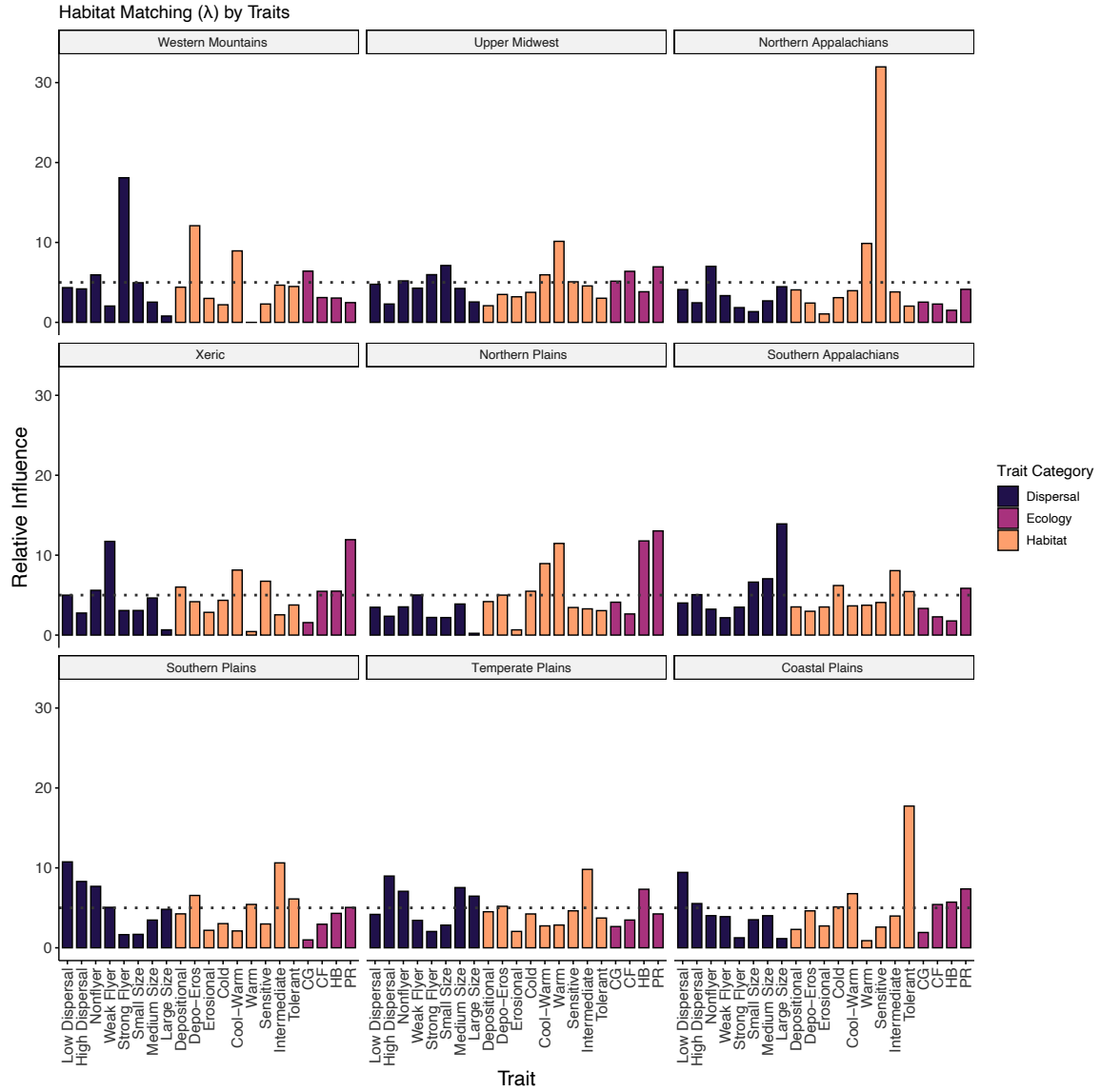


Figure 6: Plots of relative influence by dispersal, ecology, and habitat traits for each of 21 traits across all 9 ecoregions

## Trait-by-Environment

```
## Bind all BRT results into a long dataframe for plotting
trait.by.environment.long.bin.1 <- bind_rows(
  low.dispersal.environment.BRT.results,
  high.dispersal.environment.BRT.results,
  nonflyer.environment.BRT.results,
  weak.flyer.environment.BRT.results,
  strong.flyer.environment.BRT.results,
  small.size.environment.BRT.results,
  medium.size.environment.BRT.results,
  large.size.environment.BRT.results,
  depositional.environment.BRT.results,
  depositional.erosional.environment.BRT.results,
  erosional.environment.BRT.results,
  cold.water.environment.BRT.results,
  cool.warm.water.environment.BRT.results,
  warm.water.environment.BRT.results,
  sensitive.tolerance.environment.BRT.results,
  intermediate.tolerance.environment.BRT.results,
  tolerant.environment.BRT.results,
  CG.environment.BRT.results,
  CF.environment.BRT.results,
  HB.environment.BRT.results,
  PR.environment.BRT.results
)

## Add trait vector to the dataframe
trait.by.environment.long.bin.1$trait <- rep(
  c("Low Dispersal", "High Dispersal",
    "Nonflyer", "Weak Flyer", "Strong Flyer",
    "Small Size", "Medium Size", "Large Size",
    "Depositional", "Depo-Eros", "Erosional",
    "Cold", "Cool-Warm", "Warm",
    "Sensitive", "Intermediate", "Tolerant",
    "CG", "CF", "HB", "PR"),
  each = 162
)

## Filter predictors with relative influence < 5.00
trait.by.environment.long.bin.2 <- trait.by.environment.long.bin.1 %>%
  filter(rel.inf > 5.00)

## Sum relative influence by predictor type and ecoregion for each trait
trait.by.environment.long <- aggregate(
  rel.inf ~ trait + ecoregion + predictor.type,
  data = trait.by.environment.long.bin.2,
  FUN = sum
)

## Re-order ecoregion for facet plots
trait.by.environment.long$ecoregion <- fct_relevel(
  trait.by.environment.long$ecoregion,
  levels = c("WMT", "UMW", "NAP", "XER", "NPL", "SAP", "SPL", "TPL", "CPL")
)
```



)

```
trait.by.environment.figure <- ggbarplot(  
  data = trait.by.environment.long,  
  x = "trait",  
  y = "rel.inf",  
  fill = "predictor.type",  
  color = "predictor.type",  
  xlab = "Trait",  
  ylab = "Relative Influence",  
  title = "Traits by Environment",  
  facet.by = "ecoregion",  
  palette = plasma.environmental.predictors,  
  ggtheme = theme_pubr(),  
  legend = "right"  
) +  
scale_x_discrete(  
  limits = c(  
    "Low Dispersal", "High Dispersal", "Nonflyer", "Weak Flyer", "Strong Flyer",  
    "Small Size", "Medium Size", "Large Size", "Depositional", "Depo-Eros", "Erosional",  
    "Cold", "Cool-Warm", "Warm", "Sensitive", "Intermediate", "Tolerant",  
    "CG", "CF", "HB", "PR"),  
  labels = c(  
    "Low Dispersal", "High Dispersal", "Nonflyer", "Weak Flyer", "Strong Flyer",  
    "Small Size", "Medium Size", "Large Size", "Depositional", "Depo-Eros", "Erosional",  
    "Cold", "Cool-Warm", "Warm", "Sensitive", "Intermediate", "Tolerant",  
    "CG", "CF", "HB", "PR")  
) +  
rotate_x_text(angle = 90) +  
font("xlab", size = 16, color = "gray0") +  
font("ylab", size = 16, color = "gray0") +  
font("xy.text", size = 12, color = "gray20")
```

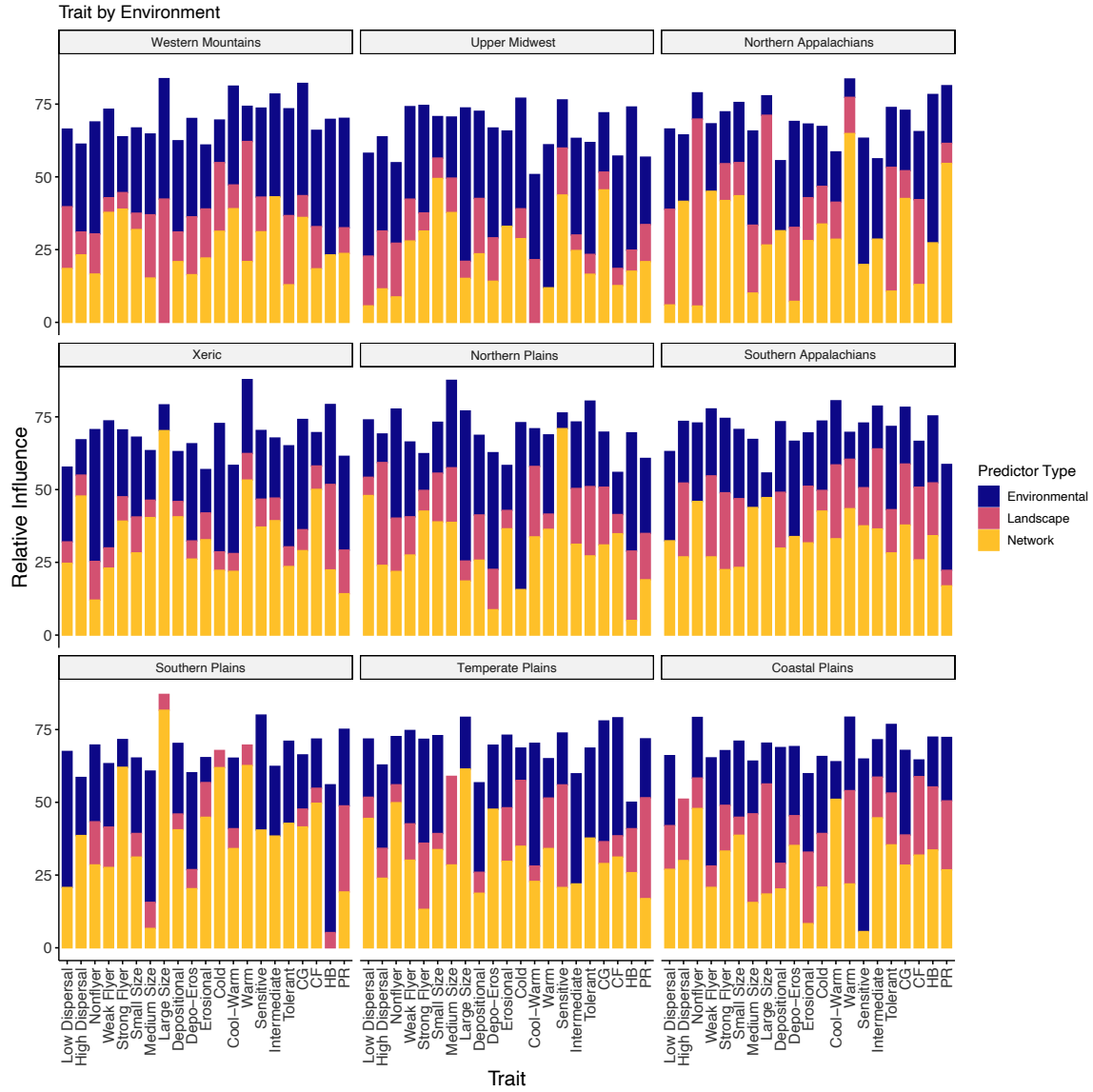


Figure 7: Plots of relative influence by environmental, landscape, and network predictors for each of 21 traits across all 9 ecoregions

## DisEQ & Functional Diversity ANCOVAs

```
## DisEQ and trait diversity data management
# Wide format
DisEQ.trait.diversity.wide <- final.DisEQ.data %>%
  select(filtering.scaled, mismatch.scaled, ecoregion, FRic, FEve, FDiv, FDis)

# Switch from wide to long
DisEQ.trait.diversity.long <- DisEQ.trait.diversity.wide %>%
  pivot_longer(cols = FRic:FDis, names_to = "trait.metric", values_to = "measurement") %>%
  na.omit()

## Reorder trait metrics for plotting
DisEQ.trait.diversity.long$trait.metric <- fct_relevel(
  DisEQ.trait.diversity.long$trait.metric,
  c("FRic", "FEve", "FDiv", "FDis")
)
```

### Filtering by Functional Diversity

```
filtering.by.trait.diversity.figure <- ggscatter(
  data = DisEQ.trait.diversity.long,
  x = "measurement",
  y = "filtering.scaled",
  size = 1,
  color = "ecoregion",
  xlab = "Trait Diversity",
  ylab = "Environmental Filtering",
  title = "Environmental Filtering by Functional Diversity",
  palette = viridis.ecoregion,
  ggtheme = theme_pubr(),
  legend = "right",
  add = "reg.line",
  conf.int = TRUE
) %>%
facet(
  facet.by = "trait.metric",
  nrow = 2,
  ncol = 2,
  scales = "free"
) +
font("xlab", size = 16, color = "gray0") +
font("ylab", size = 16, color = "gray0") +
font("xy.text", size = 12, color = "gray20") +
geom_hline(yintercept = 0, linetype = 3, size = 1, colour = "gray40")
```

## Habitat Matching by Functional Diversity

```
mismatch.by.trait.diversity.figure <- ggscatter(  
  data = DisEQ.trait.diversity.long,  
  x = "measurement",  
  y = "mismatch.scaled",  
  size = 1,  
  color = "ecoregion",  
  xlab = "Trait Diversity",  
  ylab = "Habitat Matching",  
  title = "Habitat Matching by Functional Diversity",  
  palette = viridis.ecoregion,  
  ggtheme = theme_pubr(),  
  legend = "right",  
  add = "reg.line",  
  conf.int = TRUE  
) %>%  
facet(  
  facet.by = "trait.metric",  
  nrow = 2,  
  ncol = 2,  
  scales = "free"  
) +  
font("xlab", size = 16, color = "gray0") +  
font("ylab", size = 16, color = "gray0") +  
font("xy.text", size = 12, color = "gray20") +  
geom_hline(yintercept = 0, linetype = 3, size = 1, colour = "gray40")
```

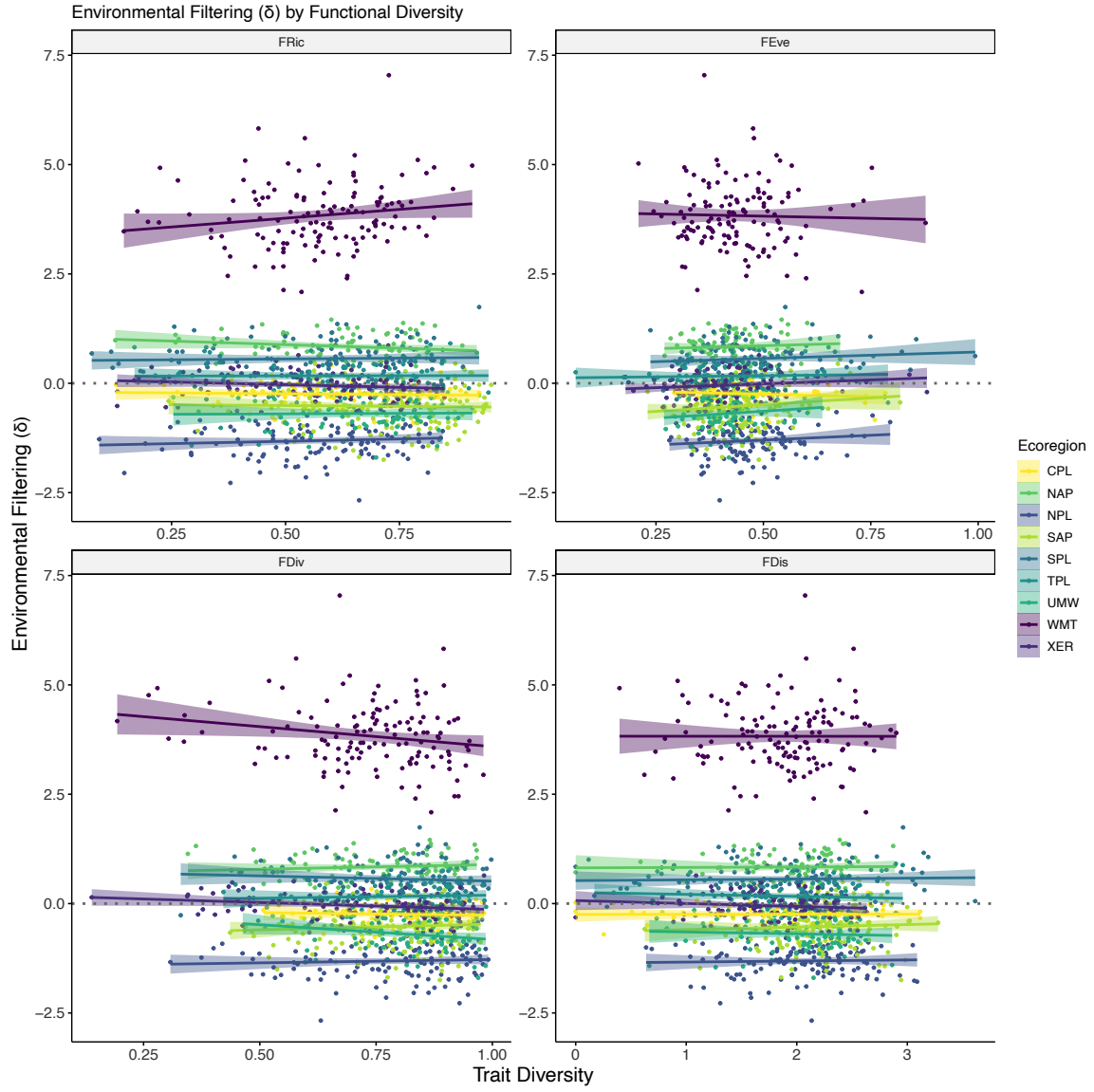


Figure 8: Plots of environmental filtering against measures of functional trait diversity.

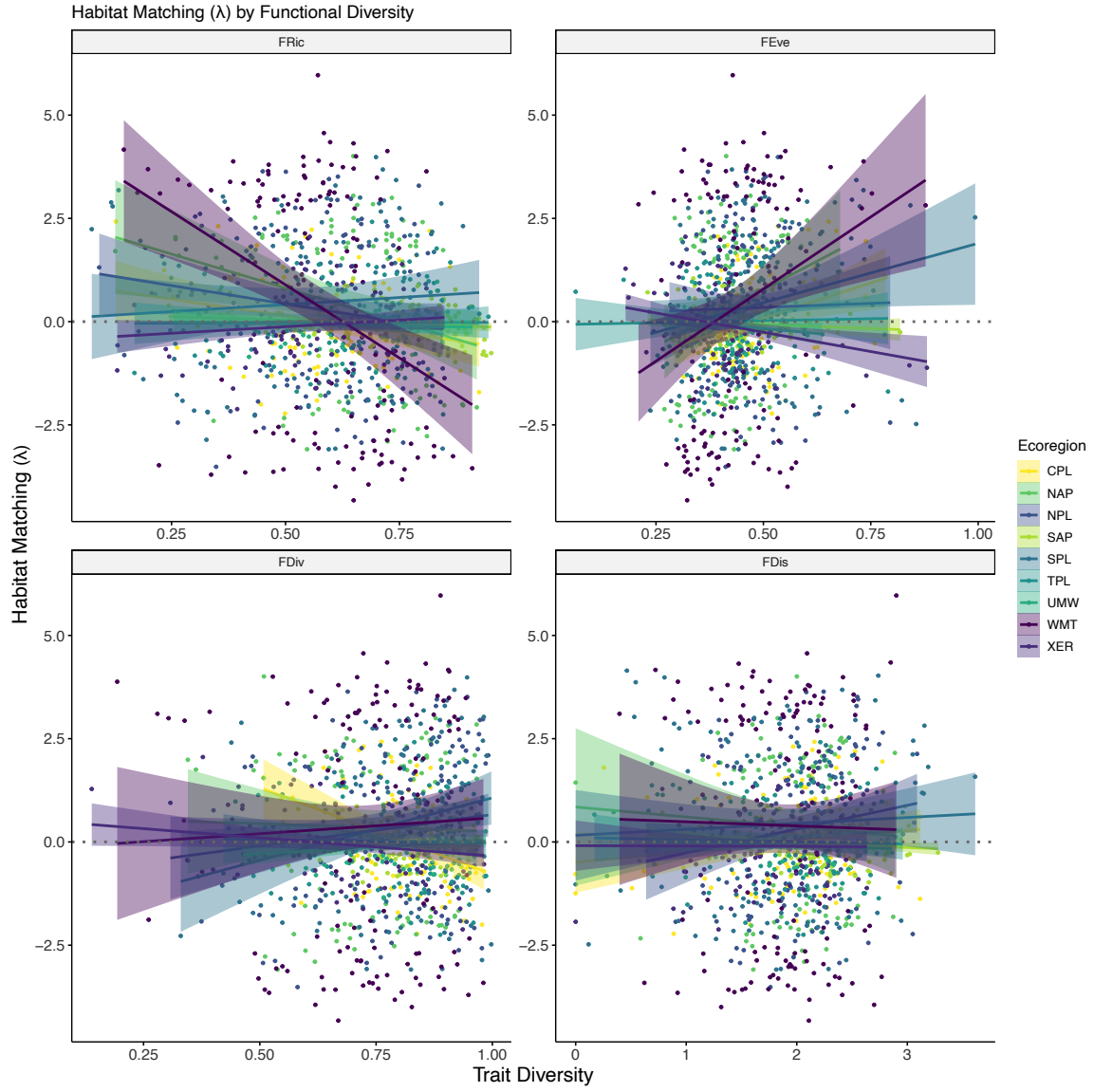


Figure 9: Plots of habitat matching against measures of functional trait diversity.

## Environmental Filtering & Habitat Matching Model

```
DisEQ.LMM.figure <- ggscatter(  
  data = final.DisEQ.data,  
  x = "mismatch.scaled",  
  y = "filtering.scaled",  
  size = 1,  
  color = "ecoregion",  
  xlab = "Habitat Matching",  
  ylab = "Environmental Filtering",  
  palette = viridis.ecoregion,  
  ggtheme = theme_pubr(),  
  legend = "right",  
  add = "reg.line",  
  conf.int = TRUE  
) +  
  geom_hline(yintercept = 0, linetype = 3, size = 1, colour = "gray40") +  
  geom_vline(xintercept = 0, linetype = 3, size = 1, colour = "gray40")
```

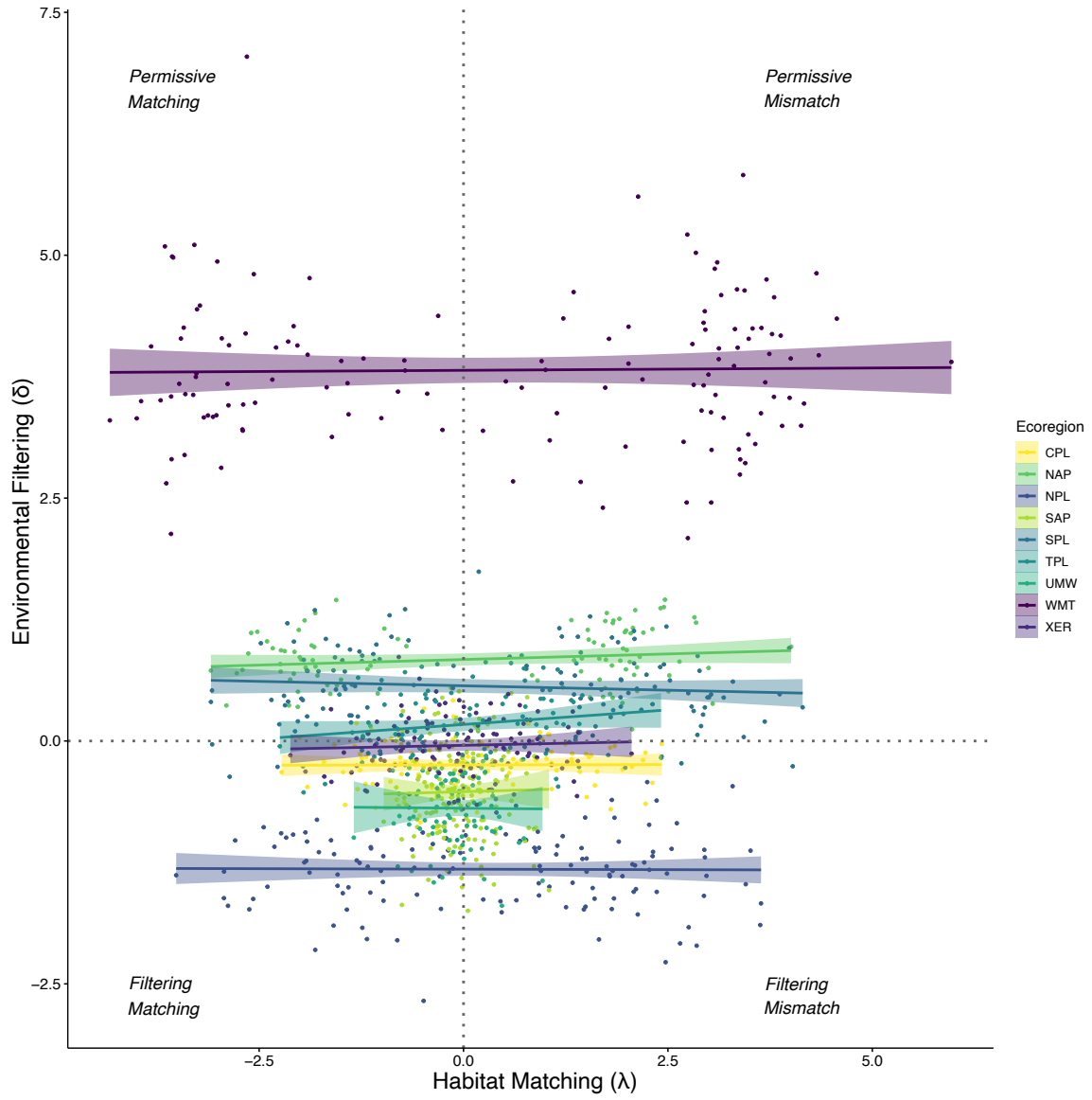


Figure 10: Plot of environmental filtering against habitat matching while controlling for the effect of ecoregion.



## Community Composition

```
## Make panel figure
par(mfrow = c(1, 1), mar = c(5, 5, 5, 5))

## Create identifier
NMDS.groups <- final.data[rowSums(final.data[, 67:142]) > 0, 25]

## Set figure dimensions
pdf("figures/base/figure_S6A-community_composition-base.pdf", width = 8, height = 6)

## NMDS plot (axes 1 and 2)
axes.1_2 <- ordiplot(BC.NMDS, choices = c(1, 2), type = "none", display = "sites",
                    xlim = c(-1.5, 1.5), ylim = c(-1.5, 1.5), xlab = "NMDS 1", ylab = "NMDS 2")
ordiellipse(BC.NMDS, groups = NMDS.groups,
            draw = "lines", col = viridis.ecoregion, label = FALSE, lwd = 5)
mtext(text = "Stress = 0.199", side = 3, line = -1.5, adj = 0.95,
      cex = 1.05)
mtext(text = "(A)", side = 3, line = 1.75, adj = 0.05, cex = 1.25)
legend(1.75, 0.45, c("CPL", "NAP", "NPL", "SAP", "SPL", "TPL", "UMW", "WMT", "XER"),
      cex = 0.75, lty = 1, lwd = 4, x.intersp = 0.95, y.intersp = 0.95,
      col = viridis.ecoregion)

## Export the figure
dev.off()

## Set figure dimensions
pdf("figures/base/figure_S6B-community_composition-base.pdf", width = 8, height = 6)

## NMDS plot (axes 1 and 3)
axes.1_3 <- ordiplot(BC.NMDS, choices = c(1, 3), type = "none", display = "sites",
                    xlim = c(-1.5, 1.5), ylim = c(-1.5, 1.5), xlab = "NMDS 1", ylab = "NMDS 3")
ordiellipse(axes.1_3, groups = NMDS.groups,
            draw = "lines", col = viridis.ecoregion, label = FALSE, lwd = 5)
mtext(text = "(B)", side = 3, line = 1.75, adj = 0.025, cex = 1.25)

## Export the figure
dev.off()
```

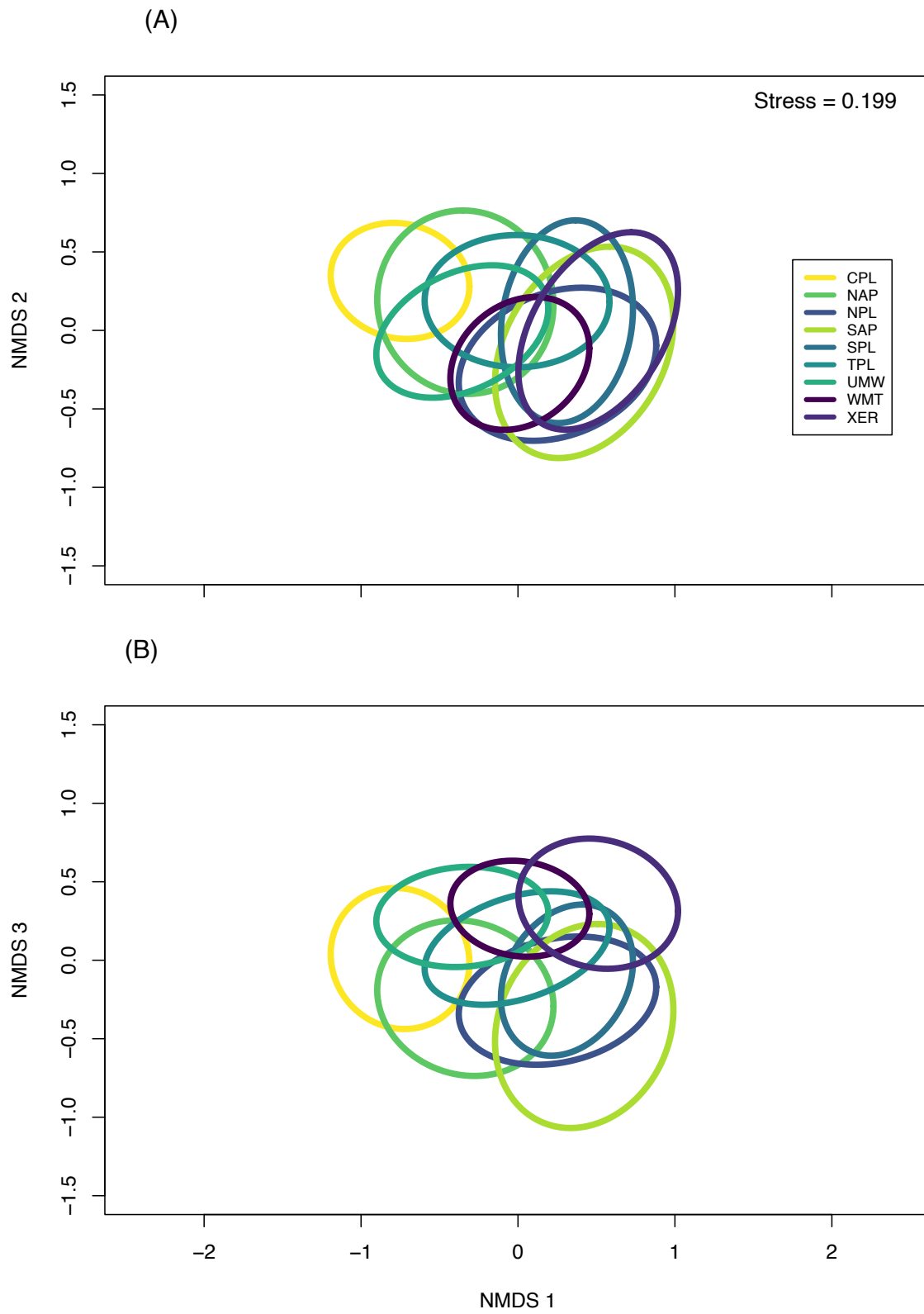


Figure 11: Plot of community composition by ecoregion.

## R Session Information

Table 1: Packages required for data management and figure creation.

Package	Loaded Version	Date
dplyr	1.0.9	2022-04-28
forcats	0.5.1	2021-01-27
ggplot2	3.3.6	2022-05-03
ggpubr	0.4.0	2020-06-27
ggsn	0.5.0	2019-02-18
gridExtra	2.3	2017-09-09
kableExtra	1.3.4	2021-02-20
knitr	1.39	2022-04-26
lattice	0.20-45	2021-09-22
mapdata	2.3.0	2018-03-30
maps	3.4.0	2021-09-25
permute	0.9-7	2022-01-27
purrr	0.3.4	2020-04-17
readr	2.1.2	2022-01-30
rgdal	1.5-30	2022-04-08
scales	1.2.0	2022-04-13
sp	1.5-0	2022-06-05
stringr	1.4.0	2019-02-10
tibble	3.1.7	2022-05-03
tidyr	1.2.0	2022-02-01
tidyverse	1.3.1	2021-04-15
vegan	2.6-2	2022-04-17
viridis	0.6.2	2021-10-13
viridisLite	0.4.0	2021-04-13