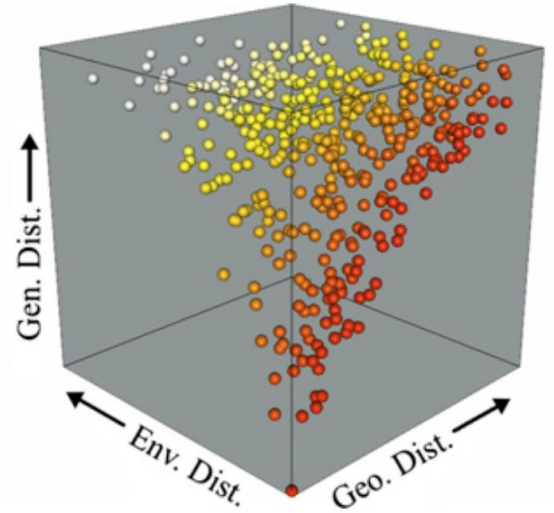
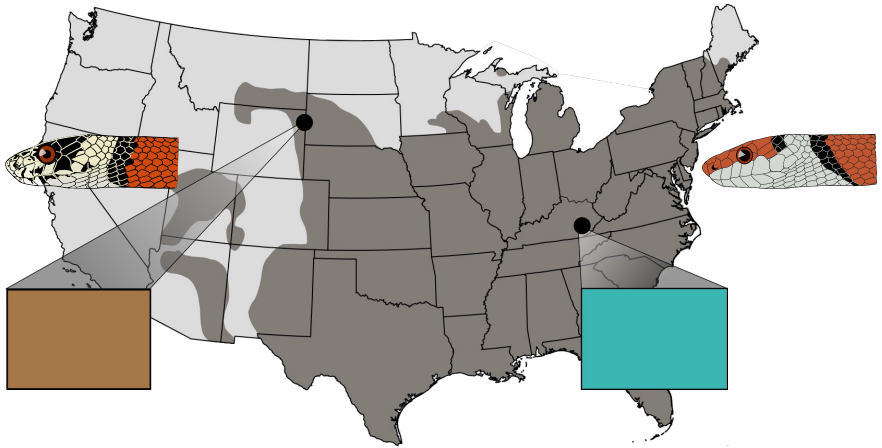
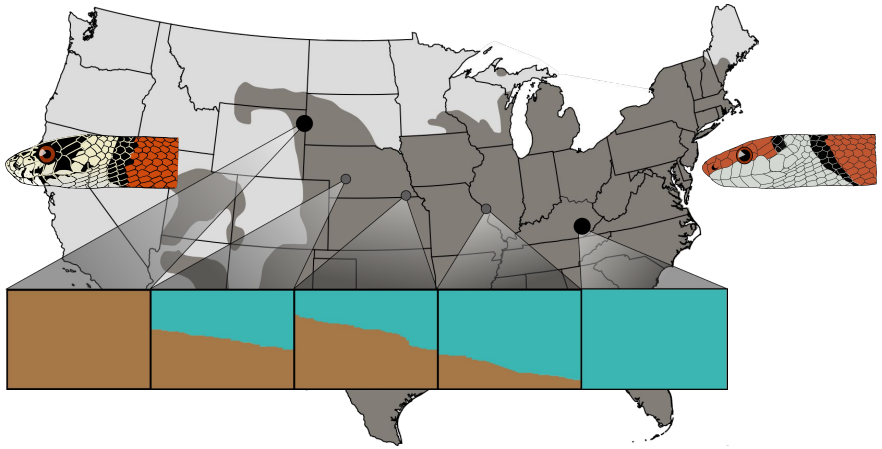


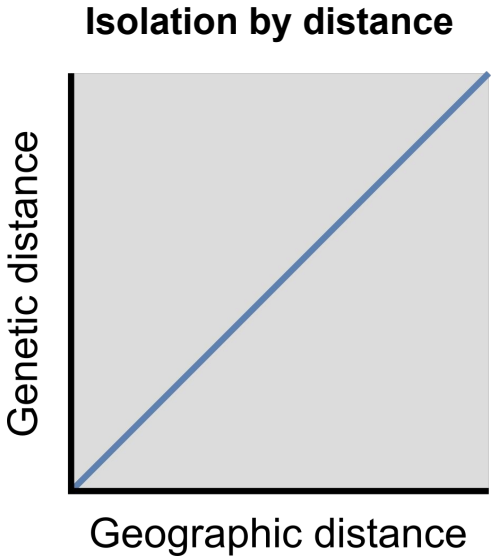

# MMRR:

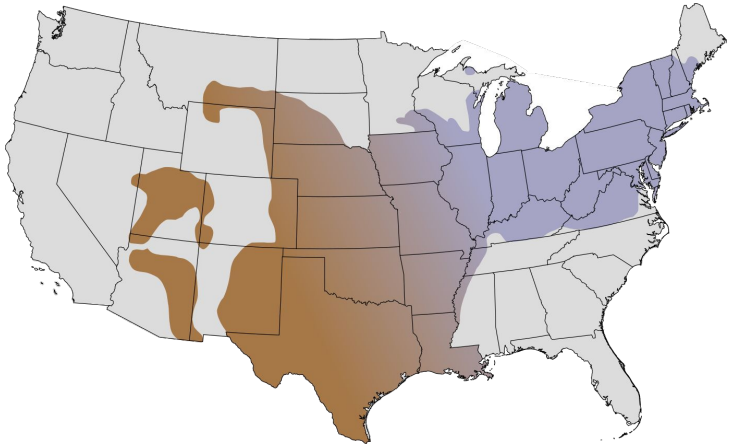
## Multiple Matrix Regression with Randomization

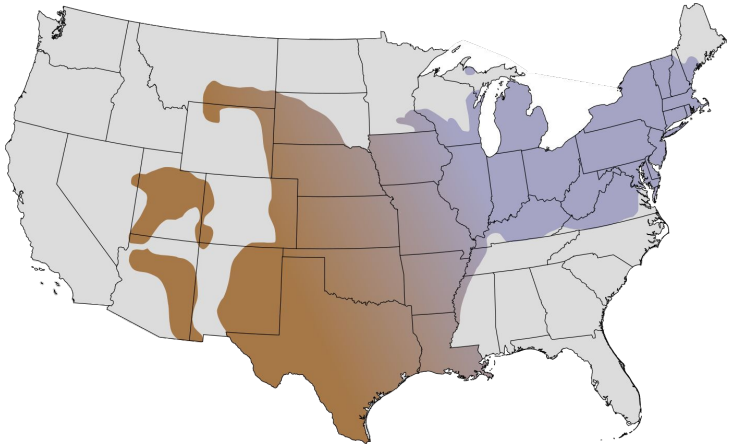
Anne Chambers & Anusha Bishop (2024)

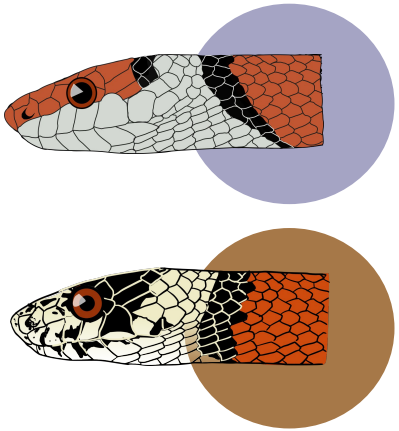


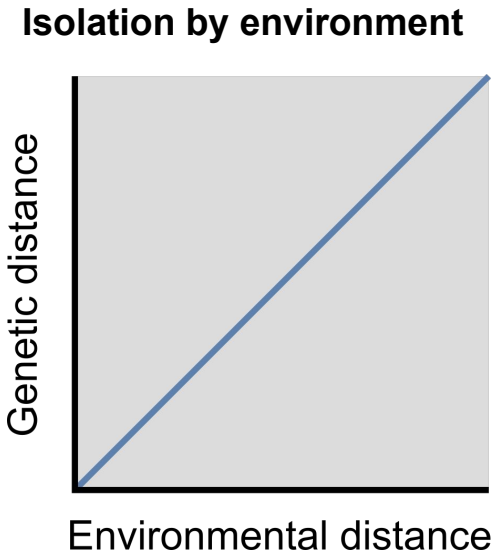



Isolation by distance
Genetic distance
Geographic distance

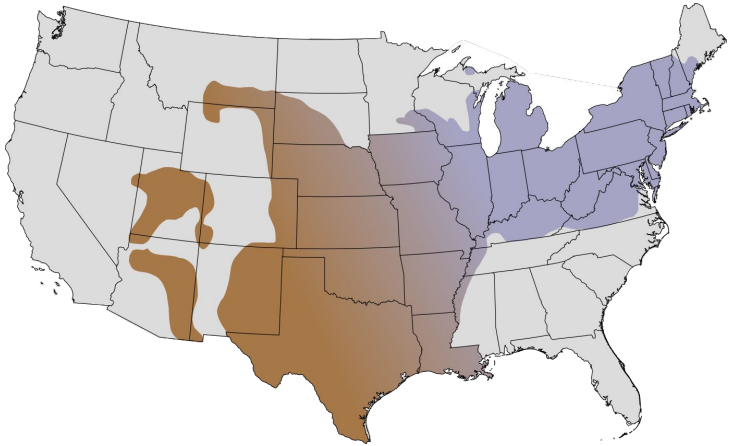


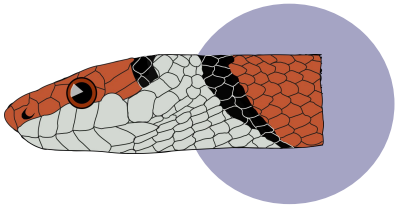


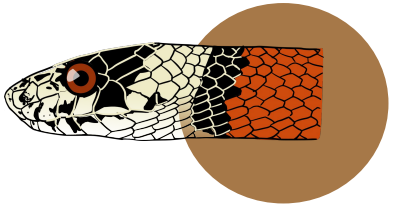



Isolation by environment
Genetic distance
Environmental distance

# Isolation by environment



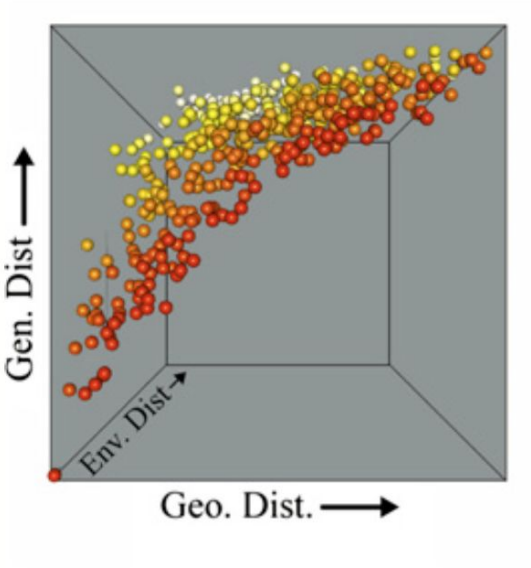




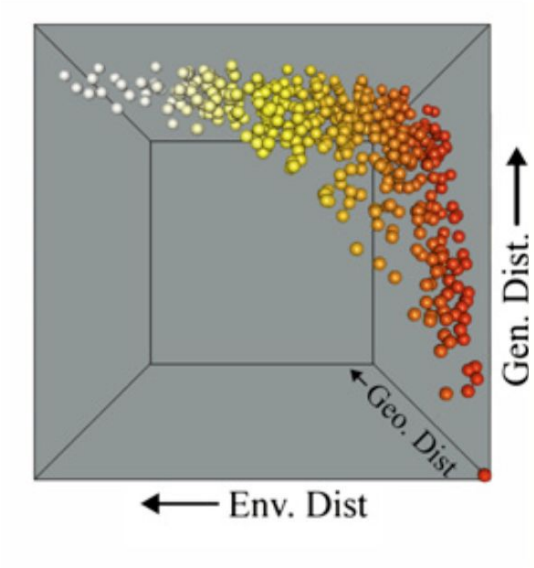
- Natural selection
- Sexual selection
- Hybrids have reduced fitness
- Biased dispersal

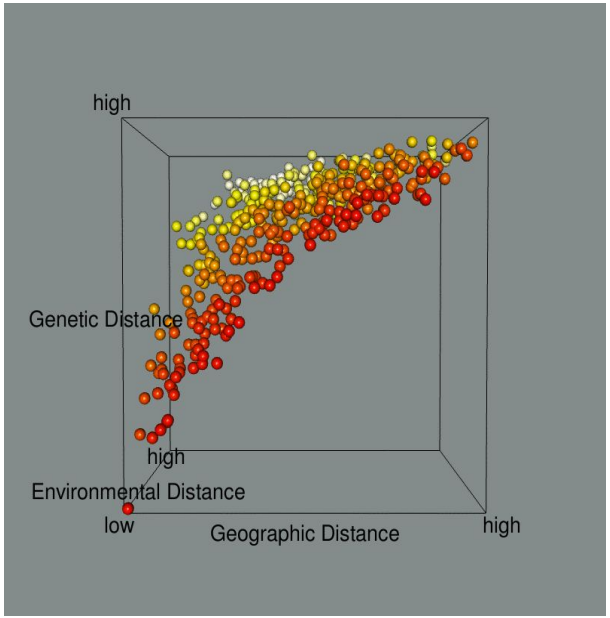
# IBD and IBE can covary with one another

**Isolation by distance**




**Isolation by environment**

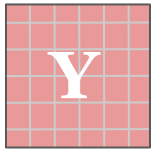


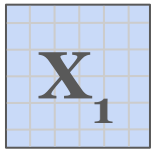

Wang & Bradburd (2014)

# Multiple matrix regression with randomization (MMRR)

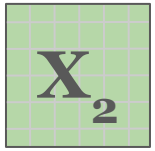
Genetic distance matrix




Geographic distance matrix



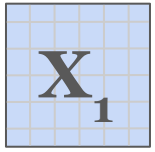

Environmental distance matrix



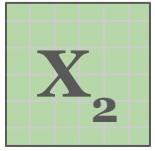
# Multiple matrix regression with randomization (MMRR)

Genetic distance matrix


Geographic distance matrix

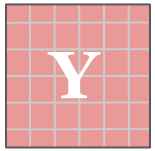


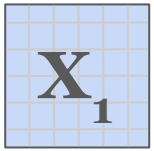

Environmental distance matrix

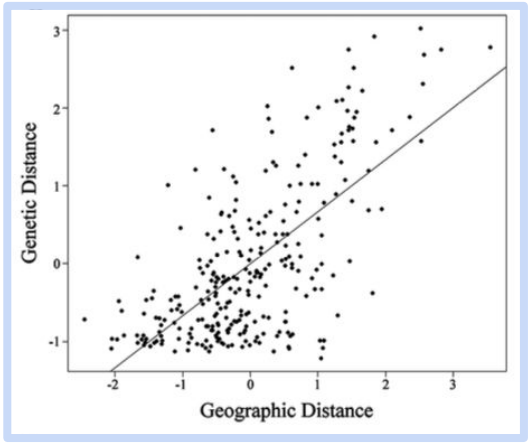



$$y_{ij} = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \varepsilon_{ij}$$

Relationship between $y_{ij}$ and $x_{1ij}$ when $x_{2ij}$ constant



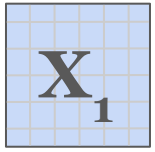



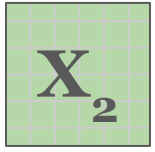
# Multiple matrix regression with randomization (MMRR)

Genetic distance matrix

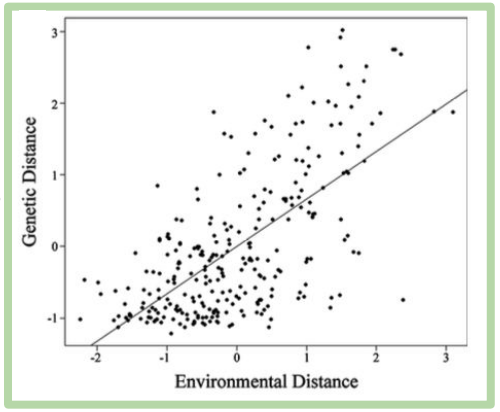

Geographic distance matrix




Environmental distance matrix

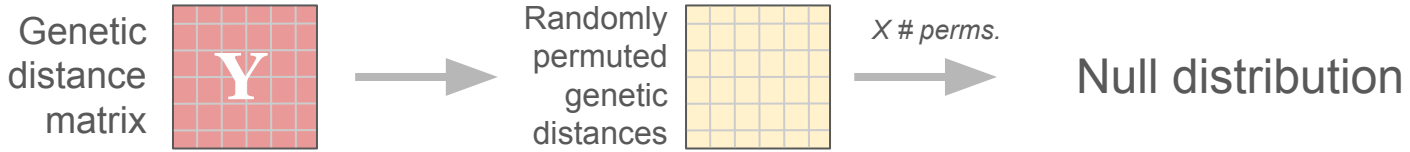



$$y_{ij} = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \varepsilon_{ij}$$




Relationship between $y_{ij}$ and $x_{2ij}$ when $x_{1ij}$ constant


# Significance testing & variable selection in MMRR

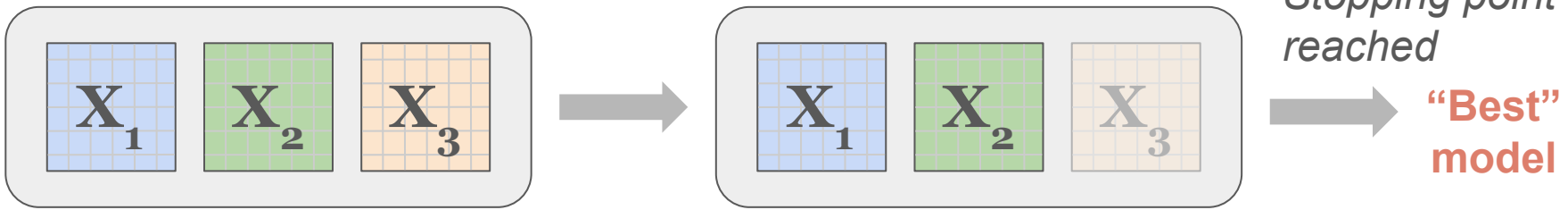



## Backward elimination

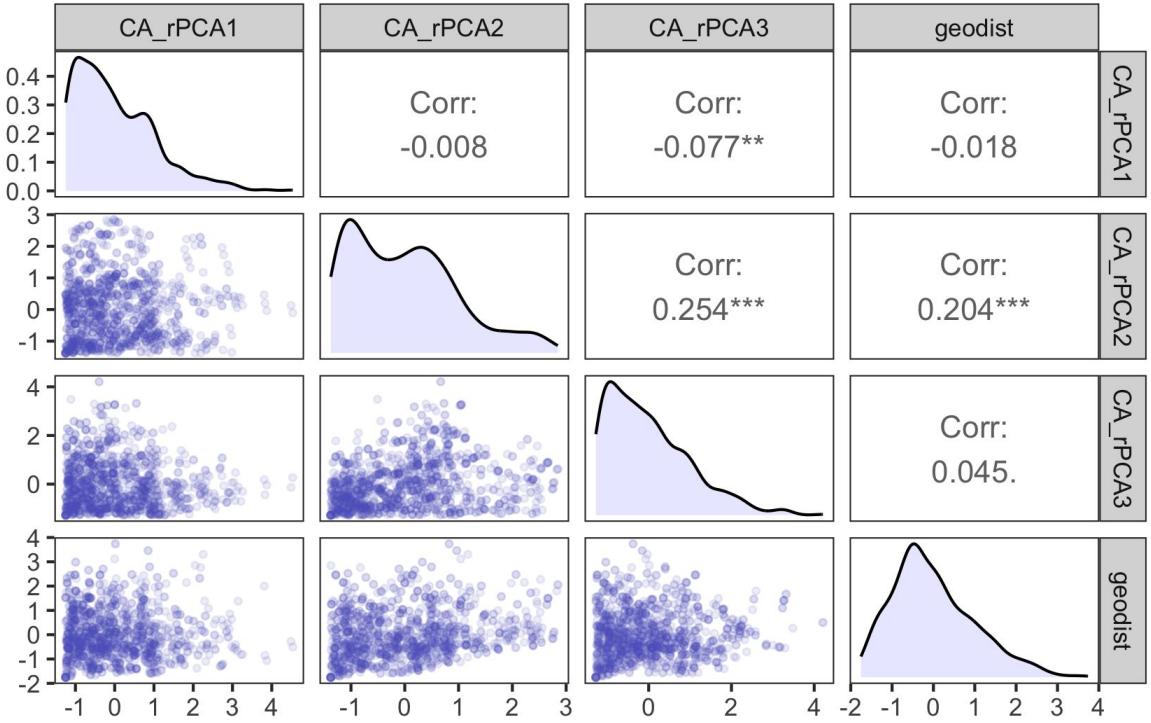
**Full model**

*Remove least significant variable*

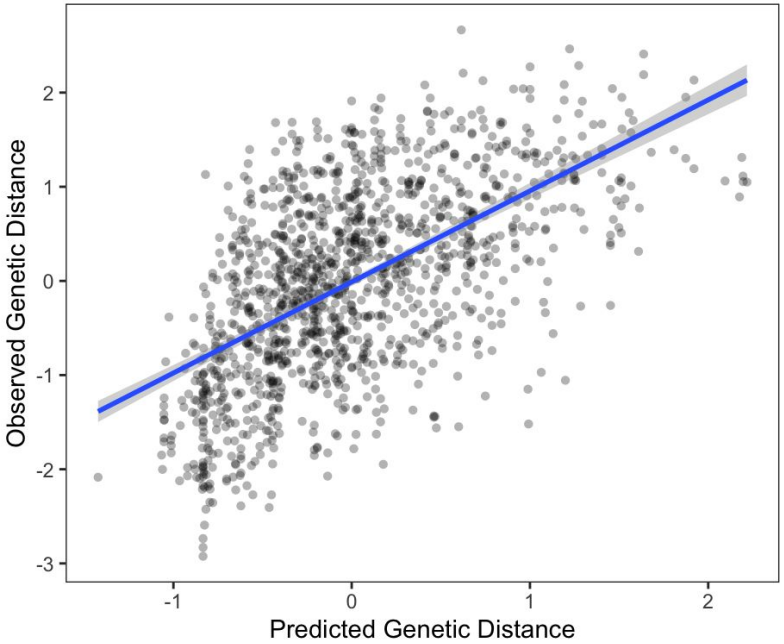
*Stopping point reached*



# Interpreting MMRR results



# Interpreting MMRR results




*Intercept + $\beta$1(geo dist) + $\beta$2(enviro dist) + Error*

Take ratio of IBD:IBE coefficients ($\beta 1/\beta 2$)

# EXERCISE

# Process input data files

```
# Convert genetic data to matrix
Y <- as.matrix(liz_gendist)
```

```
# Extract values from our environmental raster
env <- raster::extract(CA_env,
                       liz_coords)

# Calculate environmental distances
X <- env_dist(env)

# Add geographic distance to X
X[["geodist"]] <- geo_dist(liz_coords)
```

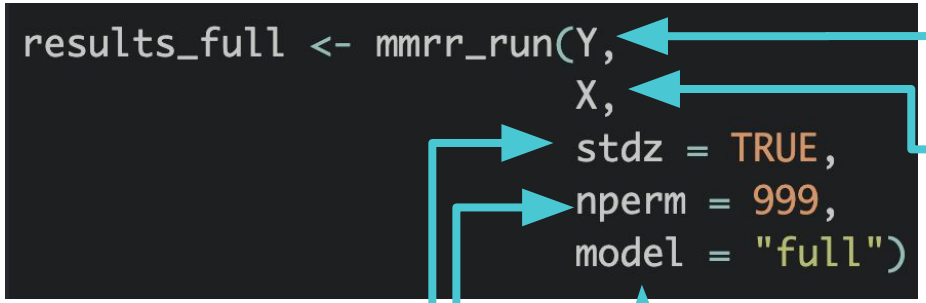
Raster from which to extract values

Data frame with points for which values will be extracted from raster

Data frame of environmental variables for locations

**`coords`** = sampling coordinates for which to calculate geographic distances

# Run MMRR




**Y** = dependent matrix (genetic distances)

**X** = independent matrix (environmental and geographic distances)

**stdz** = whether to standardize input data matrices

**nperm** = number of permutations for significance testing

**model** = run MMRR on all predictor variables with no variable selection (**"full"**) or perform variable selection (**"best"**)

*This syntax is standardized across all relevant algatr functions*

# Interpret and plot MMRR results

# Interpret and plot MMRR results

```
# Single variable plot
mmrr_plot(Y,
          X,
          mod = results_full$mod,
          plot_type = "vars",
          stdz = TRUE)
```

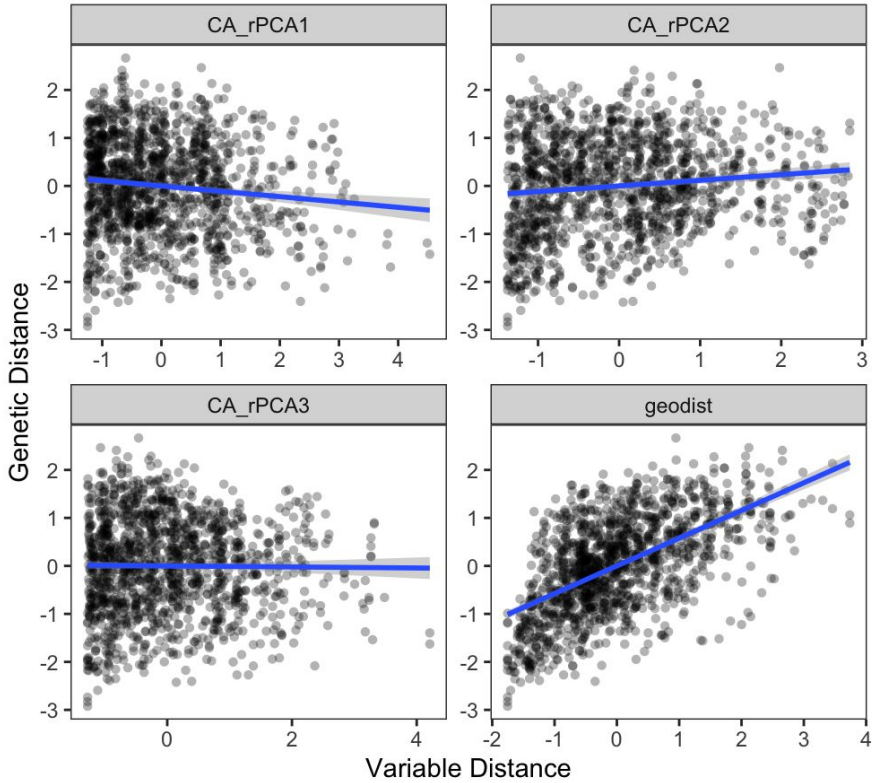
**mod** = MMRR model

**plot_type** = relationship to plot

**stdz** = standardize data matrices

# Interpret and plot MMRR results

```
# Single variable plot
mmrr_plot(Y,
          X,
          mod = results_full$mod,
          plot_type = "vars",
          stdz = TRUE)
```




Try out the other two options for the **`plot_type`** argument. What is plotted?

# Interpret and plot MMRR results

```
mmrr_table(mmrr_results = results_full,
           digits = 2,
           summary_stats = TRUE)
```

| var | estimate | p | 95% Lower | 95% Upper |
|---|---|---|---|---|
| CA_rPCA1 | −0.10 | 0.06 | −0.14 | −0.07 |
| CA_rPCA2 | 0.01 | 0.78 | −0.03 | 0.05 |
| CA_rPCA3 | −0.05 | 0.36 | −0.08 | −0.01 |
| geodist | 0.58 | 0.00 | 0.54 | 0.61 |
| Intercept | 0.00 | 0.30 | −0.04 | 0.04 |

**R-Squared:** **0.35**
**F-Statistic:** **181.28**
**F p-value:** **0.00**

See what other arguments there are within the **`mmrr_table`** function and try them out!

https://thewanglab.github.io/algatr/articles/MMRR_vignette.html

# Exercise

1. Load the example dataset
2. Calculate environmental distances using `env_dist()`
3. Run MMRR using `mmrr_run()` with and without variable selection
4. Visualize MMRR results using `mmrr_plot()` and `mmrr_table()`