Variation in patterns of selection and nonlinear selection estimates

Yedra García

9 April 2024

Variation in patterns of selection

Individuals living in different populations may experience varying selective pressures which can lead to variation in the individuals trait values and/or in the individuals fitness ultimately affecting patterns of selection. For instance, the abundance and composition of pollinators, herbivores or predators often vary locally. Similarly, variation in abiotic selective pressures such as nutrient availability, soil moisture or temperature may also differ in space and time. We thus might be interested in comparing how phenotypic selection is acting on different populations.

In addition to variation in selective pressures, the opportunity for selection I (variance in relative fitness $I = var(\frac{W}{\bar{W}})$) may also differ between populations.

In order to compare how does the direction and strength of selection vary among populations (it could also be among years or experimental treatments), we use the regression approach by Lande and Arnold 1983 to estimate phenotypic linear selection gradients β for each trait in each population followed by analysis of covariance (ANCOVA) to test for differences in selection gradients β across populations (years or treatments).

Notice that we scale the traits for each population separately. In the same way as we do for trait standardization, we estimate the relative fitness $w = W/\bar{W}$ for each population, this is dividing individual fitness by its population's mean fitness.

Exercise 1: Selection on *Ipomoea purpurea* floral traits under water deficit

In this study by Garcia et al. (2023), we set up a common garden study with the common morning glory *Ipomoea purpurea*. We studied phenotypic selection in three different experimental populations or treatments consisting on 1) well-watered plants (watered every other day) exposed to open pollination and which we called *control* treatment 2) plants under water deficit named *drought* treatment and 3) well-watered plants but with restricted access to pollinators, called *pollinator restriction* treatment. The motivation of this study was to address whether two selective agents on floral traits such as pollinators and abiotic resources (water availability) which are predicted to change within a context of global change (pollinator decline and increasing droughts) may have an impact on phenotypic selection on flowers.

During the flowering season, we measured several phenotypic traits that function as floral signals and rewards to pollinators including flower size, nectar volume and concentration and plant size, which was estimated from the diameter of the plant's main stem. We counted the number of fruits and estimated seed set (mean number of seeds per fruit multiplied by fruits) as a reproductive fitness measure.

```
ipo<-read.csv("ipomoea_data.csv")
head(ipo)
```

```
##   Plant_ID Block Source Treatment date_sampling days_setup nectar_vl
## 1      BN1     1     BN         C    2021-08-13          1  2.589844
## 2     BN10     4     BN         C    2021-09-01         20  2.291406
## 3     BN12     1     BN         C    2021-08-14          2  1.092188
## 4     BN13     1     BN         C    2021-08-13          1  2.307031
## 5     BN17     4     BN         C    2021-09-02         21  2.528125
## 6     BN19     1     BN         C    2021-08-13          1  1.374219
##   nectar_conc fl_width fl_lenght mean_size diameter buds flowers buds_2
## 1       42.35   44.215    46.435  45.31141     4.69    0       1      0
## 2       39.65   35.605    45.795  40.37983     4.84    0       0      0
## 3       38.40   36.320    38.285  37.28956     5.16    0       0      0
## 4       36.75   41.250    43.510  42.36493     5.22    0       0      0
## 5       41.65   57.875    53.640  55.71728       NA   NA      NA     NA
## 6       40.90   37.950    40.305  39.10978     4.83    2       5      0
##   flowers_2 init_fruits init_fruits_2 mature_fruits mature_fruits_2 tot_fruits
## 1         0          45            38            87              90        132
## 2         0          32            20            93              93        125
## 3         0         106            76            20              34        126
## 4         0          27            16            79              79        106
## 5        NA          NA            NA            NA              NA         NA
## 6         1          68            52            13              25         81
##   tot_fruits2 fruit_tot mean_seeds seedset
## 1         128     130.0        5.3   678.4
## 2         113     119.0        5.5   621.5
## 3         110     118.0        4.6   506.0
## 4          95     100.5        4.9   465.5
## 5          NA        NA        3.5      NA
## 6          77      79.0        4.8   369.6
```

```
ipo$Treatment<-as.factor(ipo$Treatment)
ipo[1:5, 7:12]# the phenotypic traits
```

```
##   nectar_vl nectar_conc fl_width fl_lenght mean_size diameter
## 1  2.589844       42.35   44.215    46.435  45.31141     4.69
## 2  2.291406       39.65   35.605    45.795  40.37983     4.84
## 3  1.092188       38.40   36.320    38.285  37.28956     5.16
## 4  2.307031       36.75   41.250    43.510  42.36493     5.22
## 5  2.528125       41.65   57.875    53.640  55.71728       NA
```

#Check trait correlations

```
cor.ipo= cor(ipo[,c(7:10, 12)], use="pairwise")
signif(cor.ipo, 2)
```

```
##             nectar_vl nectar_conc fl_width fl_lenght diameter
## nectar_vl       1.000       0.083    0.430    0.4500   0.0870
## nectar_conc     0.083       1.000    0.120    0.0780   0.0220
## fl_width        0.430       0.120    1.000    0.7800   0.0270
## fl_lenght       0.450       0.078    0.780    1.0000  -0.0034
## diameter        0.087       0.022    0.027   -0.0034   1.0000
```

#Explore the data in each experimental population

```r
#In the control treatment (well-watered plants)
control= ipo %>%
  filter(Treatment %in% "C")

par(mfcol= c(1,4))
plot(control$mean_size, control$seedset)
plot(control$nectar_vl, control$seedset)
plot(control$nectar_conc, control$seedset)
plot(control$diameter, control$seedset)
```
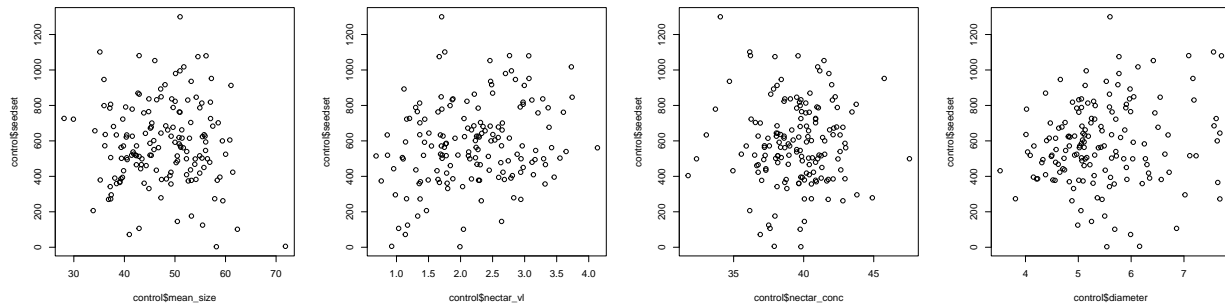


```r
#Estimating relative fitness in the control treatment
control$relfit= control$seedset/mean(control$seedset, na.rm= TRUE)
```

```r
#In the drought treatment (plants in water deficit)
drought= ipo %>%
  filter(Treatment %in% "D")

par(mfcol= c(1,4))
plot(drought$mean_size, drought$seedset)
plot(drought$nectar_vl, drought$seedset)
plot(drought$nectar_conc, drought$seedset)
plot(drought$diameter, drought$seedset)
```
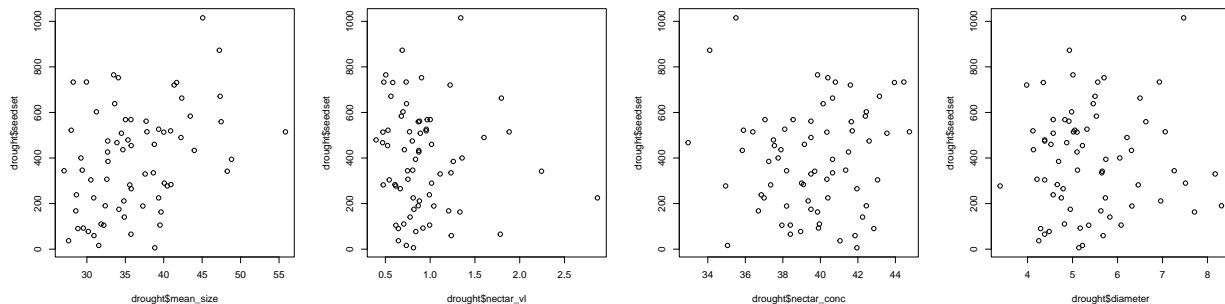


```r
#Estimating relative fitness in the drought (water deficit) treatment
drought$relfit= drought$seedset/mean(drought$seedset, na.rm= TRUE)
```
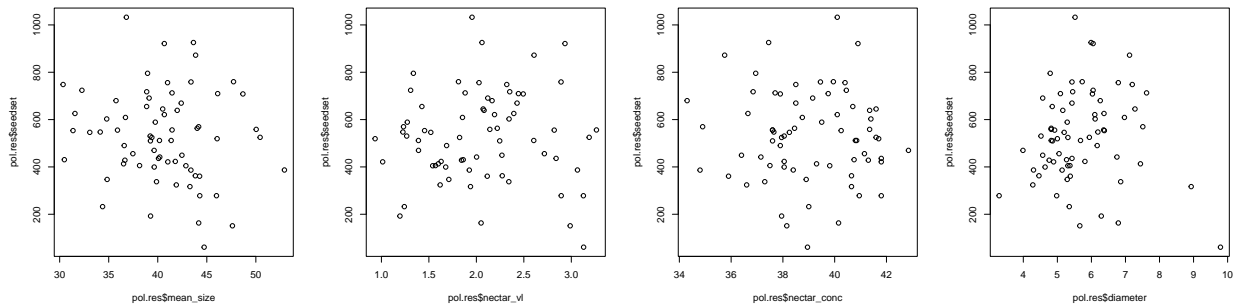
#In the pollinator restriction treatment (well-watered plants without pollinator access)

3

```r
#In the control treatment (well-watered plants)
pol.res= ipo %>%
  filter(Treatment %in% "PR")

par(mfcol= c(1,4))
plot(pol.res$mean_size, pol.res$seedset)
plot(pol.res$nectar_vl, pol.res$seedset)
plot(pol.res$nectar_conc, pol.res$seedset)
plot(pol.res$diameter, pol.res$seedset)
```



```r
#Estimating relative fitness in the pollination restriction treatment
pol.res$relfit= pol.res$seedset/mean(pol.res$seedset, na.rm= TRUE)
```

#Opportunity for selection in each treatment

```r
#In the control treatment
IC = var(control$relfit, na.rm= TRUE)
IC
```

```
## [1] 0.157721
```

```r
#In the drought treatment
ID = var(drought$relfit, na.rm= TRUE)
ID
```

```
## [1] 0.3409269
```

```r
#In the pollinator restriction treatment
IPR = var(pol.res$relfit, na.rm= TRUE)
IPR
```

```
## [1] 0.131094
```

```r
#Opportunity for selection is higher in the water deficit experimental population
```

#Variance-standardized selection gradients:

```
#In the control treatment
m = lm(relfit~ scale(nectar_vl) + scale(nectar_conc)
+ scale(mean_size) + scale(diameter), data= control, na=na.exclude)
summary(m)$coef
```

```
##                       Estimate Std. Error     t value      Pr(>|t|)
## (Intercept)        1.00085029 0.03096119 32.3259627 1.766101e-69
## scale(nectar_vl)   0.09859594 0.03210311  3.0712270 2.531822e-03
## scale(nectar_conc) -0.02682126 0.03224230 -0.8318657 4.068071e-01
## scale(mean_size)   -0.03589359 0.03273290 -1.0965600 2.745911e-01
## scale(diameter)    0.05844165 0.03150725  1.8548637 6.557876e-02
```

We notice that there is linear selection to increase nectar volume (positive sign of the parameter's Estimate) in flowers from control conditions. This is, there is an increase of 9% (nectar vol Estimate= 0.09) in fitness (seed set) per standard deviation increase in nectar volume. In addition, we detect selection to increase plant size (stem diameter), this is, larger plants produce more seeds. In this case seed set increased by 6% per standard deviation increase in plant size (stem diameter Estimate= 0.058)

#Mean-standardized selection gradients:

```
#We center the traits to the mean (subtract the mean)
#and mean-scale the traits (divide by the mean)
control$nectar_vl_m =
  (control$nectar_vl-mean(control$nectar_vl, na.rm=T))/mean(control$nectar_vl, na.rm=T)
control$nectar_conc_m =
  (control$nectar_conc-mean(control$nectar_conc, na.rm=T))/mean(control$nectar_conc, na.rm=T)
control$mean_size_m =
  (control$mean_size-mean(control$mean_size, na.rm=T))/mean(control$mean_size, na.rm=T)
control$diameter_m =
  (control$diameter-mean(control$diameter, na.rm = T))/mean(control$diameter, na.rm = T)

#Then we estimate the multivariate linear selection gradients:
m.mean = lm(relfit~ nectar_vl_m + nectar_conc_m
+ mean_size_m + diameter_m, data= control, na=na.exclude)

summary(m.mean)$coef
```

```
##                  Estimate Std. Error     t value      Pr(>|t|)
## (Intercept)     1.0008503 0.03096119 32.3259627 1.766101e-69
## nectar_vl_m     0.2940615 0.09574722  3.0712270 2.531822e-03
## nectar_conc_m  -0.4356811 0.52373976 -0.8318657 4.068071e-01
## mean_size_m    -0.2195454 0.20021282 -1.0965600 2.745911e-01
## diameter_m      0.3523579 0.18996430  1.8548637 6.557876e-02
```

In these case the mean-scaled linear selection gradients, β_μ, for nectar volume and stem diameter (0.294 vs 0.352) rank differently than variance-scaled linear selection gradients, β_σ (0.098 vs 0.058)

```
#In the water deficit treatment:
#Variance-scaled multivariate selection gradients:
md = lm(relfit~ scale(nectar_vl) + scale(nectar_conc)
+ scale(mean_size) + scale(diameter), data= drought, na=na.exclude)

summary(md)$coef
```

```
##                      Estimate Std. Error    t value     Pr(>|t|)
## (Intercept)       1.002805276 0.06560678 15.2850857 4.873794e-23
## scale(nectar_vl)  -0.133208586 0.06940373 -1.9193289 5.940343e-02
## scale(nectar_conc) 0.029288669 0.06453968  0.4538087 6.515017e-01
## scale(mean_size)   0.250715982 0.06865726  3.6517041 5.263122e-04
## scale(diameter)    0.006141644 0.06733941  0.0912043 9.276151e-01
```

In contrast to control plants, we observe selection favoring plants with lower nectar volumes (negative sign of the parameter's Estimate) suggesting a cost of nectar rewards in stressful conditions (water deficit). Specifically, fitness decreases by 13% (-0.133) per standard deviation increase in nectar volume. Moreover, selection for lower nectar seems to be a 37.4% stronger under water stress than in controlled conditions (β_σ for nectar volume in drought= 0.133 vs 0.098 in control). Interestingly, we find strong positive selection on floral size under water deficit (0.250), with an increase of 25% in fitness by a standard deviation increase in flower size.

#Mean-standardized selection gradients in water deficit:

```
#First we center to the mean and mean-scale the traits:
drought$nectar_vl_m =
  (drought$nectar_vl-mean(drought$nectar_vl, na.rm=T))/mean(drought$nectar_vl, na.rm=T)
drought$nectar_conc_m =
  (drought$nectar_conc-mean(drought$nectar_conc, na.rm=T))/mean(drought$nectar_conc, na.rm=T)
drought$mean_size_m =
  (drought$mean_size-mean(drought$mean_size, na.rm=T))/mean(drought$mean_size, na.rm=T)
drought$diameter_m =
  (drought$diameter-mean(drought$diameter, na.rm=T)) /mean(drought$diameter, na.rm = T)

#Then we estimate the multivariate linear selection gradients:
md.mean = lm(relfit~ nectar_vl_m + nectar_conc_m
+ mean_size_m + diameter_m, data= drought, na=na.exclude)

summary(md.mean)$coef
```

```
##                  Estimate Std. Error    t value     Pr(>|t|)
## (Intercept)     1.0028053 0.06560678 15.2850857 4.873794e-23
## nectar_vl_m    -0.2939973 0.15317715 -1.9193289 5.940343e-02
## nectar_conc_m   0.4655346 1.02583872  0.4538087 6.515017e-01
## mean_size_m     1.5496989 0.42437691  3.6517041 5.263122e-04
## diameter_m      0.0312788 0.34295316  0.0912043 9.276151e-01
```

Again, we observed strong selection on flower size, which is stronger than selection on fitness as a trait($\beta_\mu = 1.549 > 1$).

As we have seen, the direction in patterns of selection on nectar volume differ between treatments, with positive directional selection in plants from control conditions contrasting with negative directional selection in plants under water deficit. Selection to increase flower size in water deficit also differs from what we observed in well-watered conditions (control and PR treatments), where we did not detect selection acting on flower size.

To quantify whether the direction and strength of selection varies between treatments we perform analysis of covariance (ANCOVA) in which we compare the means of the selection gradients β. For this, we need to add the interaction terms between each trait and the variable that represents the different populations or treatments (here "Treatment").

```
#Estimate relative fitness within each treatment:
ipo.treat=ipo %>%
  group_by(Treatment)%>%
  mutate("relfit"= seedset/mean(seedset, na.rm= TRUE))%>%
  ungroup()
```

Here, because of the unbalanced design (large difference in sample size between treatments, with larger size in the control treatment) we performed a paired ANCOVA between the control and the two manipulative treatments (i.e., water deficit and and poll. restriction).

Control vs water deficit:

```
#Fit the ANCOVA model to test for differences in selection between
#control and water deficit:
ipo.treat.cd = ipo.treat %>%
  filter(!Treatment %in% "PR") %>% #exclude pol.res treatment for comparison
  droplevels()

#Notice that the ANCOVA includes the interaction (*) between each trait and the treatment variable
m.ancova.cd= aov(relfit ~ scale(nectar_vl)+ scale(nectar_conc)
+ scale(mean_size)+ scale(diameter)+ scale(nectar_vl)*Treatment+
scale(nectar_conc)*Treatment+ scale(mean_size)*Treatment+
  scale(diameter)*Treatment, na= na.exclude, data = ipo.treat.cd)

summary(m.ancova.cd)
```

```
##                              Df Sum Sq Mean Sq F value  Pr(>F)
## scale(nectar_vl)              1   0.44   0.444   2.304  0.1306
## scale(nectar_conc)            1   0.00   0.000   0.000  0.9920
## scale(mean_size)              1   0.03   0.025   0.130  0.7185
## scale(diameter)               1   0.39   0.388   2.010  0.1577
## Treatment                     1   0.48   0.482   2.498  0.1155
## scale(nectar_vl):Treatment    1   0.76   0.761   3.943  0.0484 *
## scale(nectar_conc):Treatment  1   0.15   0.147   0.762  0.3836
## scale(mean_size):Treatment    1   3.79   3.786  19.628 1.5e-05 ***
## scale(diameter):Treatment     1   0.16   0.158   0.821  0.3659
## Residuals                   214  41.28   0.193
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 12 observations deleted due to missingness
```

We can conclude that the direction and strength of linear selection on nectar volume and floral size differed between the control and the water deficit treatment. This is, water deficit changed patterns of selection on floral traits in our study system.

Control vs pol. restriction:

```
ipo.treat.cpr = ipo.treat %>%
  filter(!Treatment %in% "D") %>% #exclude water deficit treatment for #comparison
  droplevels()

m.ancova.cpr= aov(relfit ~ scale(nectar_vl)+ scale(nectar_conc)+
scale(mean_size)+ scale(diameter)+ scale(nectar_vl)*Treatment+
```

```
    scale(nectar_conc)*Treatment+ scale(mean_size)*Treatment+ scale(diameter)*Treatment,
na= na.exclude, data = ipo.treat.cpr)

summary(m.ancova.cpr)
```

```
##                              Df Sum Sq Mean Sq F value Pr(>F)
## scale(nectar_vl)              1  0.950  0.9503   6.561 0.0111 *
## scale(nectar_conc)            1  0.093  0.0927   0.640 0.4246
## scale(mean_size)              1  0.306  0.3058   2.111 0.1477
## scale(diameter)               1  0.262  0.2618   1.808 0.1802
## Treatment                     1  0.053  0.0531   0.366 0.5456
## scale(nectar_vl):Treatment    1  0.273  0.2732   1.887 0.1710
## scale(nectar_conc):Treatment  1  0.004  0.0044   0.030 0.8626
## scale(mean_size):Treatment    1  0.040  0.0395   0.273 0.6018
## scale(diameter):Treatment     1  0.229  0.2294   1.584 0.2096
## Residuals                   214 30.993  0.1448
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 9 observations deleted due to missingness
```

In contrast, there were no differences in patterns of selection between well-watered plants with varying access to pollinators (i.e control vs. pollinator restriction)

#ANCOVA on mean-standardized multivariate selection gradients:

```
#Center and scale the traits by their mean within each treatment
ipo.treat= ipo.treat %>%
  group_by(Treatment)%>%
  mutate(across(c(nectar_vl, nectar_conc, mean_size, diameter),
              ~(.x-mean(.x, na.rm = TRUE))/mean(.x, na.rm = TRUE),
              .names= "{col}_m"))%>%
  ungroup()

#Fit the ANCOVA model to test for differences in selection between control and water deficit:
ipo.treat.cd= ipo.treat %>%
  filter(!Treatment %in% "PR") %>% #exclude pol.res treatment for
  #comparison
  droplevels()

m.ancova.cdm= aov(relfit ~ nectar_vl_m+ nectar_conc_m+ mean_size_m+
diameter_m+ nectar_vl_m*Treatment+ nectar_conc_m*Treatment+
  mean_size_m*Treatment+ diameter_m*Treatment,
na= na.exclude, data = ipo.treat.cd)
summary(m.ancova.cdm)
```

```
##                        Df Sum Sq Mean Sq F value  Pr(>F)
## nectar_vl_m             1   0.30   0.304   1.578  0.2104
## nectar_conc_m           1   0.00   0.000   0.000  0.9915
## mean_size_m             1   0.45   0.455   2.358  0.1262
## diameter_m              1   0.35   0.355   1.838  0.1766
## Treatment               1   0.00   0.000   0.000  0.9838
## nectar_vl_m:Treatment   1   1.25   1.252   6.488  0.0116 *
## nectar_conc_m:Treatment 1   0.16   0.162   0.840  0.3606
```

```
## mean_size_m:Treatment      1    3.50    3.503  18.158 3.05e-05 ***
## diameter_m:Treatment        1    0.16    0.161   0.836   0.3615
## Residuals                 214   41.28    0.193
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 12 observations deleted due to missingness
```

Nonlinear selection

##Stabilizing and disruptive selection

Up to now we have seen different examples on how to measure linear, also known as directional, phenotypic selection. This mode of selection acts by changing the mean trait values in the population (e.g. plants with larger flowers or lower nectar volumes are selected). However, there are other modes of selection, where selection acts by changing the phenotypic variance in the population rather than the mean. For instance, selection might favor those individuals with trait values closer to the average phenotype (they will have higher fitness) while individuals with trait values close to the extremes of the distribution will show lower fitness values. This mode of selection is known as *stabilizing selection* and it acts decreasing the phenotypic variance in the population. In contrast to stabilizing selection, selection might favor those individuals with extreme trait values (close to the tails of the trait distribution) acting against average individuals. This case of nonlinear selection is called *disruptive selection* and increases the phenotypic variance.

In both cases we will need curve-fitting approaches to describe the selection surface (the relationship between the individuals phenotypic trait values and fitness). The regression approach by Lande and Arnold (1983) also allow us to quantify nonlinear phenotypic selection gradients by applying second-order polynomial regression. For this, we add a quadratic term to the regression model of relative fitness on phenotype, which can be written as:

$w = \alpha + \sum_i z_i\beta_i + 1/2\sum_i z_i^2\gamma_i + \epsilon_i$

where w is the relative fitness of the individual, α is an intercept, β's are partial regression coefficients (linear regression terms) and γ's are the quadratic coefficients of the regression model of relative fitness (w) on the individual's trait zi values. The 1/2 factor is used to make the quadratic terms equivalent to second derivatives (Arnold 2003).

The quadratic regression terms γ_i's are used to measure nonlinear selection. γ_i's <0 indicate stabilizing selection while γ_i's >0 indicate disruptive selection. It is important to notice that to obtain the nonlinear selection gradients, we need to double the resulting quadratic regression coefficients, γ_i, as well as their standard error (Stinchcombe et al. 2008)

We will continue with the example on *Ipomoea purpurea* to quantify nonlinear quadratic selection. We will focus on plants in the control treatment, as sample sizes in the drought and pollinator restriction were too small to quantify nonlinear selection (to include quadratic terms to the regression model).

To add the quadratic terms to the model, we need to square the traits values:

```
#We first variance-scale the traits (here with "scale" function)
#and then we square (^2) the traits:
control= control %>%
mutate(across(c(nectar_vl, nectar_conc, mean_size, diameter),
            ~(scale(.)), .names= "{col}_s"))%>%
mutate(across(c(nectar_vl_s, nectar_conc_s, mean_size_s, diameter_s),
            ~ (.x^2), .names= "{col}q"))
```

#Fit univariate variance-scaled nonlinear selection gradients:

```
##               Estimate Std. Error    t value      Pr(>|t|)
## (Intercept)  1.03166733 0.04143634 24.897648 1.671139e-55
## nectar_vl_s  0.09278596 0.03127233  2.967031 3.493767e-03
## nectar_vl_sq -0.03126944 0.02716306 -1.151175 2.514678e-01
```

```
##               Estimate Std. Error    t value      Pr(>|t|)
## (Intercept)  1.05856485 0.04004726 26.4328929 1.022406e-58
## mean_size_s  -0.01771574 0.03228409 -0.5487451 5.839855e-01
## mean_size_sq -0.06158708 0.02597363 -2.3711389 1.898672e-02
```

```
##               Estimate Std. Error    t value      Pr(>|t|)
## (Intercept)  1.005655052 0.04025354 24.9830209 1.099344e-55
## diameter_s   0.062359090 0.03618735  1.7232291 8.688065e-02
## diameter_sq -0.005154685 0.02503957 -0.2058616 8.371745e-01
```

There is stabilizing selection on flower size (mean_size_sq= -0.06).

Let's see what happens when we account for indirect effects of selection acting on other correlated traits:

#Fit multivariate variance-scaled nonlinear selection gradients:

```
m= lm(relfit~ nectar_vl_s+ nectar_vl_sq +
        nectar_conc_s+ nectar_conc_sq+ mean_size_s+ mean_size_sq
      + diameter_s+ diameter_sq, na= na.exclude, data= control)
summary(m)$coef
```

```
##                 Estimate Std. Error    t value      Pr(>|t|)
## (Intercept)    1.075693940 0.05827238 18.4597557 5.233693e-40
## nectar_vl_s    0.089647487 0.03266674  2.7443051 6.825221e-03
## nectar_vl_sq  -0.028493658 0.02755650 -1.0340086 3.028419e-01
## nectar_conc_s -0.020225446 0.03267405 -0.6190064 5.368768e-01
## nectar_conc_sq 0.004953490 0.01805891  0.2742961 7.842448e-01
## mean_size_s   -0.029598703 0.03285518 -0.9008838 3.691337e-01
## mean_size_sq  -0.047367224 0.02608059 -1.8161871 7.139317e-02
## diameter_s     0.065800573 0.03611433  1.8220072 7.050044e-02
## diameter_sq   -0.006177247 0.02473155 -0.2497720 8.031148e-01
```

There is still (marginally significant) stabilizing (Estimate= -0.047) selection on flower size.

To obtain the nonlinear selection gradient γ_i, we need to double the quadratic regression coefficient and its standard error (Stinchcombe et al. 2008).

```
gamma_meansize= summary(m)$coef[7]*2
gamma_meansize
```

```
## [1] -0.09473445
```

```
SE= 0.026*2
SE
```

```
## [1] 0.052
```

γ for mean flower size is = -0.094 (+/- 0.052)

Correlational selection

In some cases, selection may favor individuals with specific combinations of trait values (Phillips and Arnold 1989). This mode of selection is known as correlational selection and implies that selection on one trait depends on the value of another trait. Once again we can use Lande and Arnold's regression approach to quantify correlational selection. To do that we add the cross-product terms $\gamma_{i,j}$ for each pair of traits to the regression model with the linear and the quadratic regression terms. For two traits z_i and z_j the model can be written as:

$w = \alpha + z_i\beta_i + z_j\beta_j + 1/2z_i^2\gamma_{i,i} + 1/2z_j^2\gamma_{j,j} + z_iz_j\gamma_{i,j} + \epsilon$

where $\gamma_{i,j}$ denotes the correlational selection gradient for traits z_i and z_j

Exercise 2: Selection by pollinators and seed predators on floral traits in two *Ipomopsis* species.

In this work by Dianne Campbell and collaborators (Campbell et al. 2022), they investigated linear and nonlinear phenotypic selection acting on a set of floral traits, including floral size-related traits, nectar, color and scent, in two species of *Ipomopsis* (pollinated by hummingbirds and hawk moths) and their hybrids during two different stages of the plant's life cycle which are: 1) pollination (i.e. seeds initiated) and 2) seed predation (proportion of fruits escaping fly predation).

Previous work on this system has shown that certain traits are under directional selection by pollinators and seed predators. Based on this knowledge, here they tested several predictions in which correlational selection may arise as the effect of combining directional selection by both pollinators and seed predators.

Here, we will focus on *I. aggregata* and three traits (floral width, color and nectar production) for which sample sizes are large enough to test for linear, quadratic and correlational selection through multivariate regression following Lande and Arnold's approach.

```
camp22<-read.csv("Campbell_etal_22.csv")
names(camp22)
```

```
##  [1] "Site"      "Year"      "Type"      "width"     "color"     "nectar"
##  [7] "sepal"     "totpinene" "indoleN"   "seeds"     "seedsinit" "flyescape"
```

```
camp22$Site= as.factor(camp22$Site)
camp22$Type= as.factor(camp22$Type)
```

#Filter to include only I. aggregata sites

```
aggre= camp22 %>%
  filter(Site== "agg") %>%
dplyr::select(c(1:6, 11, 12))%>% #Filter to include petal width, color #and nectar
droplevels()
```
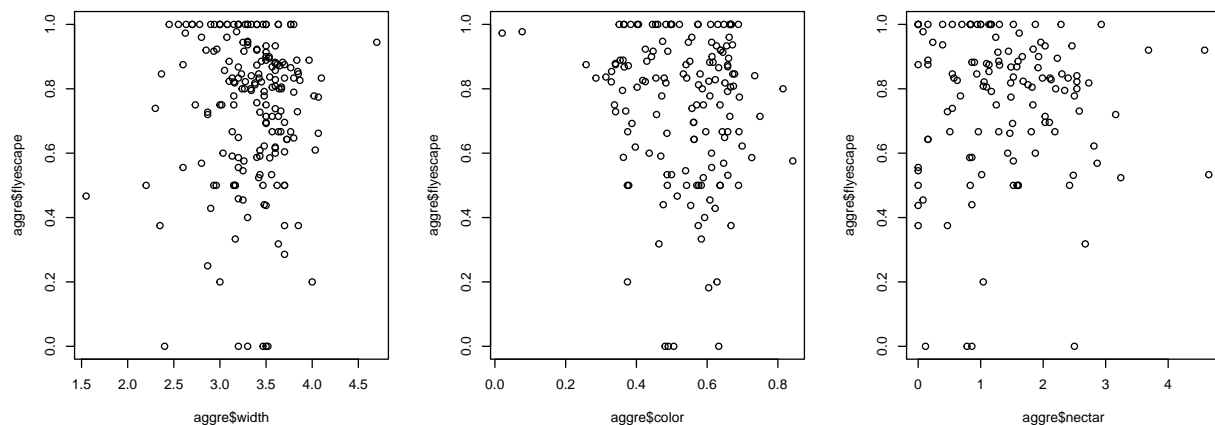
Explore the data: at pollination stage (seeds initiated)

```
par(mfrow= c(1,3))
plot(aggre$width, aggre$seedsinit)
plot(aggre$color, aggre$seedsinit)
plot(aggre$nectar, aggre$seedsinit)
```
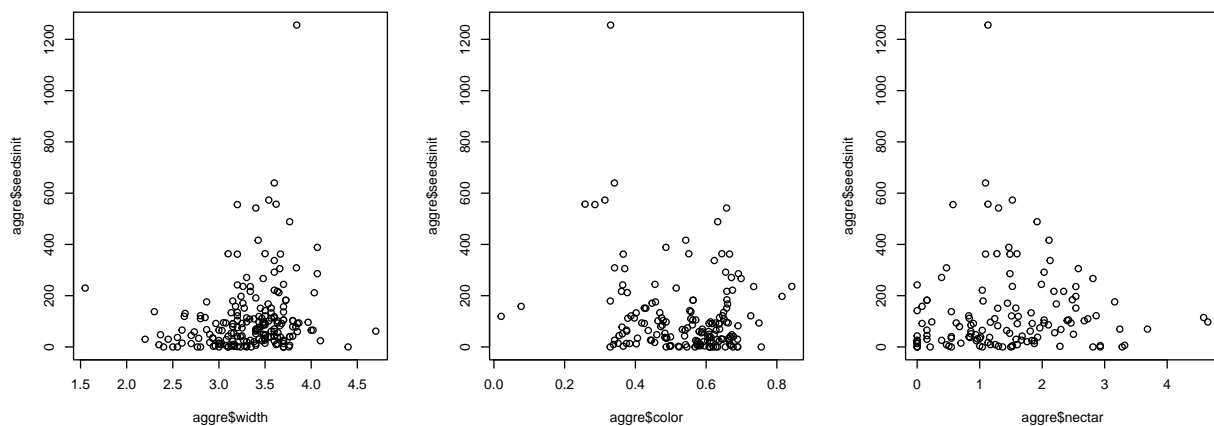

Explore the data: at predation stage (seeds escaping fly predation)

```
par(mfrow= c(1,3))
plot(aggre$width, aggre$flyescape)
plot(aggre$color, aggre$flyescape)
plot(aggre$nectar, aggre$flyescape)
```

Estimate relative fitness at pollination and seed predation stages:

```r
aggre$relfit.pol = aggre$seedsinit/mean(aggre$seedsinit, na.rm = T)
aggre$relfit.pred = aggre$flyescape/mean(aggre$flyescape, na.rm = T)
```

#Variance-scaled univariate linear selection gradients at the pollination stage

```r
summary(lm(relfit.pol~ scale(width), data= aggre, na= na.exclude))$coef
```

```
##                Estimate Std. Error    t value      Pr(>|t|)
## (Intercept)  1.0157628 0.09402000 10.803690 1.097042e-21
## scale(width) 0.2722926 0.09519785  2.860281 4.686646e-03
```

```r
summary(lm(relfit.pol~ scale(color), data= aggre, na= na.exclude))$coef
```

```
##                Estimate Std. Error   t value      Pr(>|t|)
## (Intercept)   1.0998237  0.1134841  9.691433 8.783222e-18
## scale(color) -0.2949649  0.1122830 -2.626977 9.446904e-03
```

```r
summary(lm(relfit.pol~ scale(nectar), data=aggre, na= na.exclude))$coef
```

```
##                 Estimate Std. Error   t value      Pr(>|t|)
## (Intercept)   1.21404288  0.1370054 8.8612749 4.676412e-15
## scale(nectar) 0.07696696  0.1374180 0.5600939 5.763647e-01
```

We observe positive directional selection to increase flower width and negative directional selection to decrease petal color (selection for pale-colored flowers)

Now when we accounting for selection acting on other traits: #Variance-scaled multivariate linear selection gradients at the pollination stage

```r
m = lm(relfit.pol~ scale(width) + scale(color)
+ scale(nectar), data= aggre, na= na.exclude)
summary(m)$coef
```

```
##                   Estimate Std. Error    t value     Pr(>|t|)
## (Intercept)      1.23422786  0.1316850  9.3725764 3.469012e-16
## scale(width)     0.48654477  0.1434723  3.3912115 9.279303e-04
## scale(color)    -0.30200110  0.1262586 -2.3919254 1.822624e-02
## scale(nectar)    0.05930215  0.1346535  0.4404055 6.603921e-01
```

We still detect linear selection through seeds initiated (i.e. pollination stage) acting on flower width (β_σ= 0.48) and petal color (β_σ= -0.30)

#Mean-scaled multivariate linear selection gradients at pollination stage

```
#Mean-scale the traits
aggre= aggre %>%
mutate(across(c(width, color, nectar),~ (.x-mean(.x, na.rm = T))/mean(.x, na.rm = T),
.names= "{col}_m"))


m = lm(relfit.pol~ width_m + color_m+ nectar_m, na= na.exclude, data= aggre)
summary(m)$coef
```

```
##                   Estimate Std. Error    t value     Pr(>|t|)
## (Intercept)      1.23422786  0.1316850  9.3725764 3.469012e-16
## width_m          3.86634780  1.1401081  3.3912115 9.279303e-04
## color_m         -1.25863303  0.5262008 -2.3919254 1.822624e-02
## nectar_m         0.08658552  0.1966041  0.4404055 6.603921e-01
```

Mean-standardized multivariate selection gradients also show strong directional selection on flower width (positive) and floral color (negative), with both β_μ's >1 (3.86 and -1.25) which are stronger than selection on fitness as a trait.

#Estimating multivariate quadratic selection gradients:

```
#Variance-scale and square the traits
aggre= aggre %>%
  mutate(across(c(width, color, nectar),  ~(scale(.)), .names= "{col}_s")) %>%
mutate(across(c(width_s, color_s, nectar_s),  ~ .x^2, .names= "{col}q"))

#Fit the quadratic regression model:
m = lm(relfit.pol~ width_s + color_s+ nectar_s+ width_sq+ color_sq+ nectar_sq,
       na= na.exclude, data= aggre)
summary(m)$coef
```

```
##                  Estimate Std. Error    t value     Pr(>|t|)
## (Intercept)    1.0550669 0.20737844   5.087640 1.305635e-06
## width_s        0.6291963 0.17116970   3.675863 3.514559e-04
## color_s       -0.2097932 0.13999649  -1.498560 1.365292e-01
## nectar_s       0.1820166 0.14683179   1.239627 2.174544e-01
## width_sq       0.1755319 0.11928787   1.471498 1.436897e-01
## color_sq       0.1427399 0.07504765   1.901990 5.949216e-02
## nectar_sq     -0.1313350 0.09120246  -1.440038 1.523765e-01
```

Disruptive selection (color_sq$coeff= 0.14 >0) on floral color. To estimate the quadratic selection gradient γ_i we double the quadratic regression coefficient and the standard error for flower color:

```
gamma_color =summary(m)$coef[6, 1]*2
SE= 0.075*2
```

##Measuring correlational selection:

#Variance-scaled multivariate nonlinear selection gradients Include the two-way interaction terms for each pair of traits to test for correlational selection

```
m = lm(relfit.pol~ width_s + color_s+ nectar_s+
        width_sq + color_sq+ nectar_sq+
  width_s*color_s+ width_s*nectar_s+  color_s*nectar_s,
  na= na.exclude, data= aggre)
summary(m)$coef
```

```
##                    Estimate Std. Error    t value      Pr(>|t|)
## (Intercept)      1.00480974 0.20303951  4.9488384 2.440654e-06
## width_s          0.66668676 0.16838790  3.9592321 1.273946e-04
## color_s         -0.16445462 0.13661907 -1.2037457 2.310376e-01
## nectar_s         0.23777649 0.14370088  1.6546627 1.005846e-01
## width_sq         0.07139757 0.12127719  0.5887139 5.571506e-01
## color_sq         0.27747596 0.08799030  3.1534834 2.034931e-03
## nectar_sq       -0.15437757 0.09101317 -1.6962113 9.241720e-02
## width_s:color_s  -0.51241680 0.15674675 -3.2690745 1.405132e-03
## width_s:nectar_s -0.01008042 0.15178227 -0.0664137 9.471581e-01
## color_s:nectar_s -0.06011296 0.15145515 -0.3969027 6.921384e-01
```

There is negative correlational selection ($\gamma_{i,j}$= -0.512) acting on flower width and color. Selection favors plants with wide flowers when flowers are pale in color.

#Mean-scaled multivariate nonlinear selection gradients:

```
#Square the mean-scaled traits
aggre= aggre %>%
mutate(across(c(width_m, color_m, nectar_m),  ~ .x^2, .names= "{col}q"))

m = lm(relfit.pol~ width_m + color_m+ nectar_m+ width_mq +
        color_mq+ nectar_mq+ width_m*color_m+ width_m*nectar_m+
        color_m*nectar_m, na= na.exclude, data= aggre)
summary(m)$coef
```

```
##                    Estimate Std. Error    t value      Pr(>|t|)
## (Intercept)      1.0048097  0.2030395  4.9488384 2.440654e-06
## width_m          5.2978535  1.3381013  3.9592321 1.273946e-04
## color_m         -0.6853883  0.5693796 -1.2037457 2.310376e-01
## nectar_m         0.3471713  0.2098139  1.6546627 1.005846e-01
## width_mq         4.5085785  7.6583524  0.5887139 5.571506e-01
## color_mq         4.8195508  1.5283260  3.1534834 2.034931e-03
## nectar_mq       -0.3291046  0.1940233 -1.6962113 9.241720e-02
## width_m:color_m  -16.9703997  5.1911940 -3.2690745 1.405132e-03
## width_m:nectar_m  -0.1169585  1.7610594 -0.0664137 9.471581e-01
## color_m:nectar_m  -0.3657915  0.9216150 -0.3969027 6.921384e-01
```