Appendix S3: Statistical Models, Results and Figures

This document lays out R code for all analyses and creates figures for "Local climate and habitat continuity interact to alter contemporary dispersal potential", by:xxxxx (submitted 2022 to *The American Naturalist*). Code was created by xxxxx, please contact with questions.

```
library(plyr)
library(ggplot2)
library(lme4)
library(lmerTest)
library(lsmeans)
library(tidyverse)
library(MuMIn)
library(rootSolve)
library(pracma)
library(gridExtra)
library("cowplot")

##Data
gdat <- read.csv("geum_data.csv", header=TRUE) #trait and climate data
tvdat <- read.csv("terminalvelocity_data.csv", header=TRUE) #terminal velocity data
height <- read.csv("height_data.csv", header=TRUE) #flowering height data from herbaria
wind <- read.csv("wind_data.csv", header=TRUE) #wind data

#data clean
gdat$Replicate <- as.factor(gdat$Replicate) #rep per mom
```

#Data Analysis and Figure Creation

We first analyze all diaspore trait data and create the figures in the manuscript. We examine diaspore mass, morphology, and terminal velocity.

```
#####
#
#  MASS MEASUREMENTS
#
#####

##DIASPORE MASS
masstest <- lmer(sqrt(Weight_Tot) ~ Region * log(DD5_TOT) + (1|Population_Code) +
              (1|Year), data= gdat, REML=FALSE)
summary(masstest)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: sqrt(Weight_Tot) ~ Region * log(DD5_TOT) + (1 | Population_Code) +
##     (1 | Year)
##    Data: gdat
##
##      AIC      BIC   logLik deviance df.resid
```

```
## -13191.0 -13151.9   6602.5 -13205.0     1977
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.1542 -0.6053  0.0175  0.6235  2.5967
##
## Random effects:
##  Groups          Name        Variance  Std.Dev.
##  Population_Code (Intercept) 6.670e-05 0.008167
##  Year            (Intercept) 4.684e-05 0.006844
##  Residual                    6.395e-05 0.007997
## Number of obs: 1984, groups:  Population_Code, 105; Year, 4
##
## Fixed effects:
##                              Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)                   0.15955    0.08574 103.42057   1.861   0.0656 .
## RegionPrairie                -0.23113    0.09916 103.15263  -2.331   0.0217 *
## log(DD5_TOT)                 -0.01146    0.01129 103.35251  -1.015   0.3125
## RegionPrairie:log(DD5_TOT)    0.03073    0.01309 103.34535   2.347   0.0208 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) RgnPrr l(DD5_
## RegionPrair -0.872
## lg(DD5_TOT) -0.999  0.873
## RP:(DD5_TOT  0.873 -1.000 -0.874
```

```
r.squaredGLMM(masstest)
```

```
##             R2m       R2c
## [1,] 0.04636448 0.6563929
```

```
#Set up color palatte for all figures
theme_ba <- function (base_size = 12, base_family = "")
{
  theme_grey(base_size = base_size, base_family = base_family) %+replace%
  theme(axis.text = element_text(size = rel(0.8)),
      axis.ticks = element_line(colour = "black"),
  legend.key = element_rect(colour = "grey80"),
      panel.background = element_rect(fill = "white",
  colour = NA), panel.border = element_rect(fill = NA,
  colour = "black"), panel.grid.major = element_line(colour = "white",
  size = 0.2), panel.grid.minor = element_line(colour = "white",
  size = 0.5), strip.background = element_rect(fill = "light grey",
  colour = "grey50", size = .75))
}

##Create Figure Panel 3a
mass <- ggplot(gdat, aes(x=log(DD5_TOT), y=sqrt(Weight_Tot), linetype = Region))+
  geom_point(aes(col=Region),size=3)+
    geom_point(shape=1,size=3,colour='black')+
    geom_smooth(aes(col=Region, fill = Region), method='lm', alpha = 1)+
    theme_ba()+
    scale_color_manual(values=c('gray20','gray70'))+
```

```
  scale_fill_manual(values=c('gray40','gray40'))+
  scale_linetype_manual(values=c("dashed", "dashed"))+
    ylab('sqrt(Total Mass (mg))')+
  theme (legend.position="none" )+
  labs(tag = "A)")+
  theme(axis.title.x=element_blank())


#####
#
#  MORPHOLOGY MEASUREMENTS
#
#####

##TOTAL DIASPORE LENGTH
lengthtest <- lmer(sqrt(Dispersal_and_seed_mm) ~ Region * log(DD5_TOT) +
          (1|Population_Code/Sample_ID) + (1|Year), data = gdat, REML=FALSE)
summary(lengthtest)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula:
## sqrt(Dispersal_and_seed_mm) ~ Region * log(DD5_TOT) + (1 | Population_Code/Sample_ID) +
##     (1 | Year)
##    Data: gdat
##
##      AIC      BIC   logLik deviance df.resid
##   4435.0   4479.7  -2209.5   4419.0     1975
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.3336 -0.3884  0.0486  0.4242  3.6753
##
## Random effects:
##  Groups                     Name        Variance Std.Dev.
##  Sample_ID:Population_Code (Intercept) 0.3041   0.5515
##  Population_Code           (Intercept) 0.4920   0.7014
##  Year                      (Intercept) 0.2940   0.5422
##  Residual                              0.2849   0.5337
## Number of obs: 1983, groups:
## Sample_ID:Population_Code, 1058; Population_Code, 105; Year, 4
##
## Fixed effects:
##                             Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)                  13.5825     7.5042 100.7620   1.810  0.07328 .
## RegionPrairie               -26.8529     8.6734 100.1710  -3.096  0.00254 **
## log(DD5_TOT)                 -0.9270     0.9884 100.7643  -0.938  0.35056
## RegionPrairie:log(DD5_TOT)    3.5425     1.1452 100.3778   3.093  0.00256 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) RgnPrr l(DD5_
## RegionPrair -0.872
```

```
## lg(DD5_TOT) -0.999  0.872
## RP:(DD5_TOT  0.873 -1.000 -0.874
```

```
r.squaredGLMM(lengthtest)
```

```
##            R2m       R2c
## [1,] 0.0773782 0.8088472
```

```
##DIASPORE AREA
areatest <- lmer(sqrt(SeedArea) ~ Region + log(DD5_TOT) + (1|Population_Code/Sample_ID) +
          (1|Year), data = gdat, REML=FALSE)
#summary(areatest)
#r.squaredGLMM(areatest)


##DIASPORE SHAPE INDEX
shapetest <- lmer(sqrt(SeedShapeIndex) ~ Region + log(DD5_TOT) +
          (1|Population_Code/Sample_ID) + (1|Year), data = gdat, REML=FALSE)
#summary(shapetest)
#r.squaredGLMM(shapetest)


##Create Figure Panel 3b
lnth <- ggplot(gdat, aes(x=log(DD5_TOT), y=sqrt(Dispersal_and_seed_mm),
                         col=Region, linetype = Region))+
    geom_point(aes(col=Region),size=3)+
  geom_point(shape=1,size=3,colour='black')+
    geom_smooth(aes(col=Region, fill = Region),method='lm', alpha = 1)+
    theme_ba()+
    scale_color_manual(values=c('gray20','gray70'))+
  scale_fill_manual(values=c('gray40','gray40'))+
  scale_linetype_manual(values=c("dashed", "dashed"))+
    ylab('sqrt(Total Length (mm))')+
  theme (legend.position="none" )+
  labs(tag = "B)")+
  theme(axis.title.x=element_blank())


#####
#
#  TERMINAL VELOCITY MEASUREMENTS
#
#####

##TERMINAL VELOCITY
tvtest <- lmer(sqrt(tv) ~ Region * log(DD5_TOT) + (1|Population/Sample_ID) + (1|Year),
          data=tvdat, REML=FALSE)
summary(tvtest)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: sqrt(tv) ~ Region * log(DD5_TOT) + (1 | Population/Sample_ID) +
##     (1 | Year)
##    Data: tvdat
##
```

```
##      AIC      BIC   logLik deviance df.resid
##  -1660.3  -1621.2    838.1  -1676.3      966
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.2857 -0.5438  0.0165  0.5834  4.9745
##
## Random effects:
##  Groups               Name        Variance  Std.Dev.
##  Sample_ID:Population (Intercept) 0.0035682 0.05973
##  Population           (Intercept) 0.0036991 0.06082
##  Year                 (Intercept) 0.0007076 0.02660
##  Residual                         0.0076454 0.08744
## Number of obs: 974, groups:  Sample_ID:Population, 210; Population, 43; Year, 2
##
## Fixed effects:
##                             Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)                  0.06491    0.75664 42.94635   0.086  0.93203
## RegionPrairie                2.91763    0.98306 41.14785   2.968  0.00498 **
## log(DD5_TOT)                 0.11313    0.09993 42.96569   1.132  0.26391
## RegionPrairie:log(DD5_TOT) -0.37984    0.12967 41.21750  -2.929  0.00552 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) RgnPrr l(DD5_
## RegionPrair -0.739
## lg(DD5_TOT) -0.999  0.737
## RP:(DD5_TOT  0.740 -1.000 -0.740
```

```
r.squaredGLMM(tvtest)
```

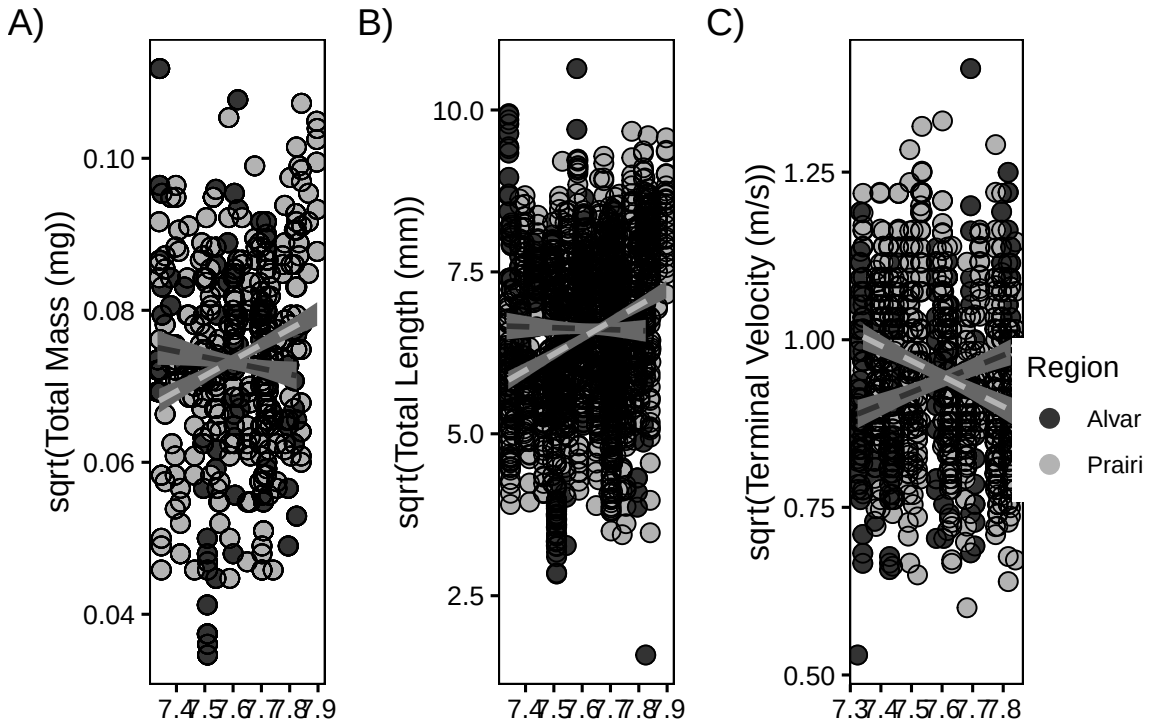
```
##             R2m       R2c
## [1,] 0.08930447 0.5542548
```

```
##Creates Figure Panel 3c
tv <- ggplot(tvdat, aes(x=log(DD5_TOT), y=sqrt(tv),
                        col=Region, linetype = Region))+
    geom_point(aes(col=Region),size=3)+
  geom_point(shape=1,size=3,colour='black')+
    geom_smooth(aes(col=Region, fill = Region),method='lm', alpha = 1)+
    theme_ba()+
    scale_color_manual(values=c('gray20','gray70'))+
  scale_fill_manual(values=c('gray40','gray40'))+
  scale_linetype_manual(values=c("dashed", "dashed"))+
    ylab('sqrt(Terminal Velocity (m/s))')+
  theme (legend.position="none" )+
  labs(tag = "C)")+
  theme(axis.title.x=element_blank())

#to make the legend readable.
leg <- ggplot(tvdat, aes(x=log(DD5_TOT), y=sqrt(tv),col=Region))+
    geom_point(aes(col=Region),size=3, position = position_dodge(width = 0.1))+
    scale_color_manual(values=c('gray20','gray70'))+
  theme_bw()
```

```
g_legend<-function(a.gplot){
  tmp <- ggplot_gtable(ggplot_build(a.gplot))
  leg <- which(sapply(tmp$grobs, function(x) x$name) == "guide-box")
  legend <- tmp$grobs[[leg]]
  return(legend)}

mylegend<-g_legend(leg)

##Creates Figure 3
Fig3_a <- plot_grid(mass, lnth, tv,
                    nrow = 1)
```

```
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
```

```
title_x <- ggdraw()+
                  draw_label(bquote('log(Degree Days above' ~5^o~C*')'),
                             fontface = "bold",
                             x = .35, y = .9,
                             hjust = 0,
                             size = 16)

Fig3_b <- plot_grid(Fig3_a, NULL, title_x,
                       ncol = 1,
                       rel_heights = c(1,0.05, 0.1))

Fig3_all <- plot_grid(Fig3_b, mylegend,
                    rel_widths = c(10,1),
                    nrow = 1)

Fig3_all
```




```
# tiff("Fig3.tiff", units="in", width=8.5, height=4, res=300, compression = 'lzw')
# Fig3_all
# dev.off()
```

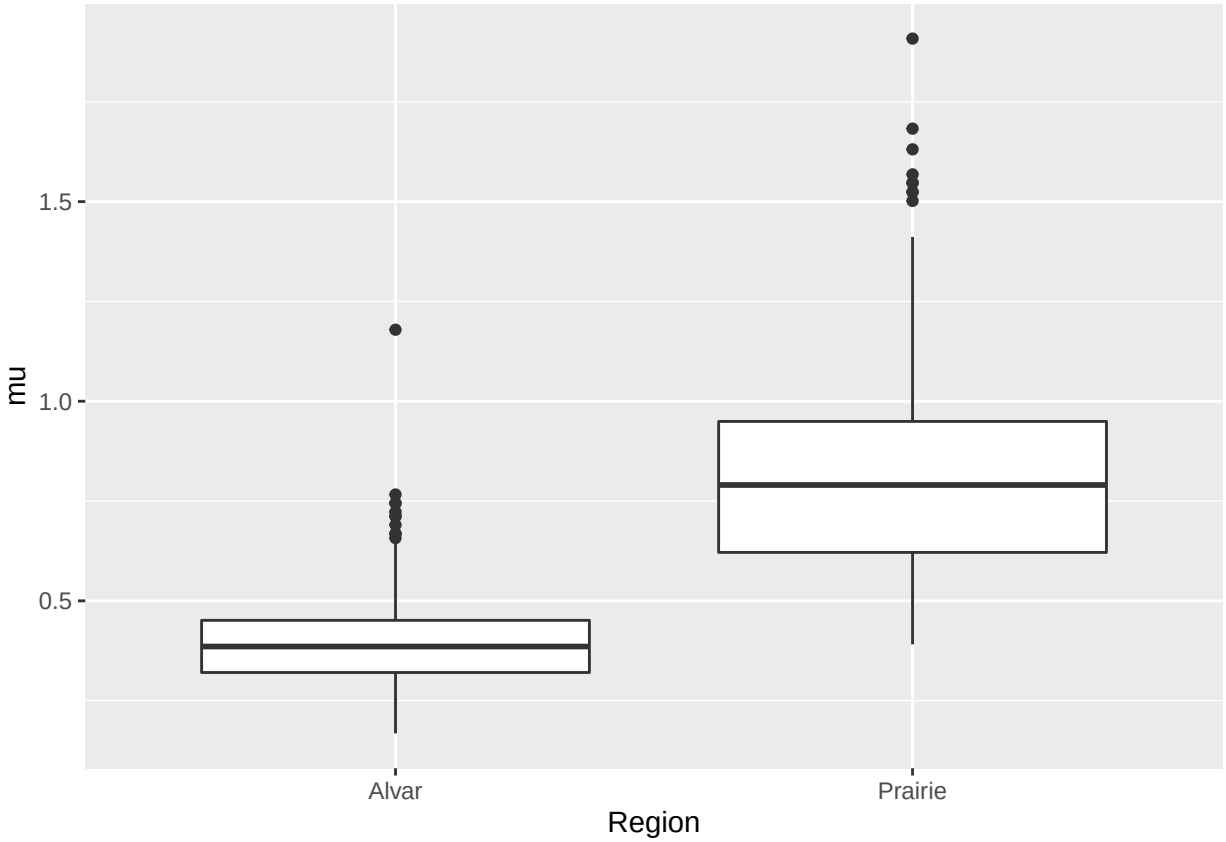
Next we take the terminal velocity data and translate it into a dispersal potential ability at the 99% tail of the dispersal kernel.

```
## Examine height differences by region
summary(aov(height_cm ~ Region, data = height))
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## Region       1   17.4   17.38   0.461  0.499
## Residuals   83 3131.3   37.73
```

```
##Translate TV into max dispersal

height_avg <- ddply(height, .(Region), summarize, avg_flwr_ht = mean(height_cm)/100)
wind_avg <- ddply(wind, .(measurement), summarize, wind_average =
          (mean(value)*1000)/3600)  #km/h into m/s

height_avg$max_wind <- 0
height_avg$equal_wind <- 0
height_avg[2,3] <- wind_avg[2,2]      #add prairie wind - which is avg of June winds at
                                      #   Geum common garden for 2017 and 2018
height_avg[1,3] <- wind_avg[2,2]/2    #assume alvar wind is 50% of prairie wind
height_avg[,4] <- wind_avg[2,2]       #for equal wind between prairie and alvar.

tvdat_all <- merge(tvdat, height_avg, by="Region")
```

```
##Creating the kernels using the WALD model (based on Katul et al. 2005)
#Wind-based, and other non-plant parameters
kappa <- .3      #set between .3 and .4; (.3 = canopies that are very dense)
h <- .2          #Canopy height.
sigma2 <- kappa*h*(2*(1/2)) #sigma2 <- kappa*h*(2*(sigmaw/U)) - sigmaw/U = ratio of 1/2.
g <- 9.8 #m/s
rho <- .001225*1000000 #g/m3, density of air at 15oC at sea level is 0.001225 g/cm3

#Inverse-gaussian params (WALD), calculated from data and wind parameters
tvdat_all$lambda <- (tvdat_all$avg_flwr_ht/sqrt(sigma2))^2  #lambda
tvdat_all$mu <- (tvdat_all$avg_flwr_ht*tvdat_all$max_wind)/tvdat_all$tv    #mu

#plot exploring mean dispersal distance by region
ggplot(tvdat_all, aes(x=Region, y=mu))+
  geom_boxplot()
```




```
#Use CDF to extract tail, or the distance the farthest 1% will travel.
results_dstar <- matrix(nrow=0,ncol=2)
row <- as.vector(1:nrow(tvdat_all))
max_tail <- 10000

for(i in 1:length(row)){
  #i=3
  temp <- tvdat_all[i,]
  lambda <- temp$lambda
  mu <- temp$mu
```

```
  fun_99<- function(x) (.5*erfc((sqrt(lambda/x)*(mu-x))/(sqrt(2)*mu)) +
        .5*(exp((2*lambda)/mu))*erfc((sqrt(lambda/x)*(mu+x))/(sqrt(2)*mu))-.99)
  dstar_99 <- uniroot(fun_99, c(.00001,max_tail))$root

  new <- cbind(temp,  dstar_99)
  results_dstar <- rbind(results_dstar, new)

  #print(i/length(row))  #keeps track of progress
}

#write_csv(results_dstar, "GEUM_dispersal_kernel_integration.csv")

##DISPERSAL POTENTIAL
disp99_test <- lmer((dstar_99) ~ Region * log(DD5_TOT) + (1|Population/Sample_ID) +
          (1|Year), data=results_dstar, REML=FALSE)
summary(disp99_test)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: (dstar_99) ~ Region * log(DD5_TOT) + (1 | Population/Sample_ID) +
##     (1 | Year)
##    Data: results_dstar
##
##      AIC      BIC   logLik deviance df.resid
##   2641.5   2680.5  -1312.7   2625.5      966
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.9823 -0.4352 -0.1124  0.2675  8.0772
##
## Random effects:
##  Groups               Name        Variance Std.Dev.
##  Sample_ID:Population (Intercept) 0.1792   0.4234
##  Population           (Intercept) 0.2424   0.4924
##  Year                 (Intercept) 0.0356   0.1887
##  Residual                         0.6821   0.8259
## Number of obs: 974, groups:  Sample_ID:Population, 210; Population, 43; Year, 2
##
## Fixed effects:
##                            Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)                  3.3730     6.0790  43.3026   0.555   0.5818
## RegionPrairie              -18.9393     7.9174  41.2638  -2.392   0.0214 *
## log(DD5_TOT)                -0.2794     0.8028  43.3296  -0.348   0.7295
## RegionPrairie:log(DD5_TOT)   2.7365     1.0443  41.3439   2.620   0.0122 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) RgnPrr l(DD5_
## RegionPrair -0.739
## lg(DD5_TOT) -0.999  0.738
## RP:(DD5_TOT  0.741 -1.000 -0.740
```

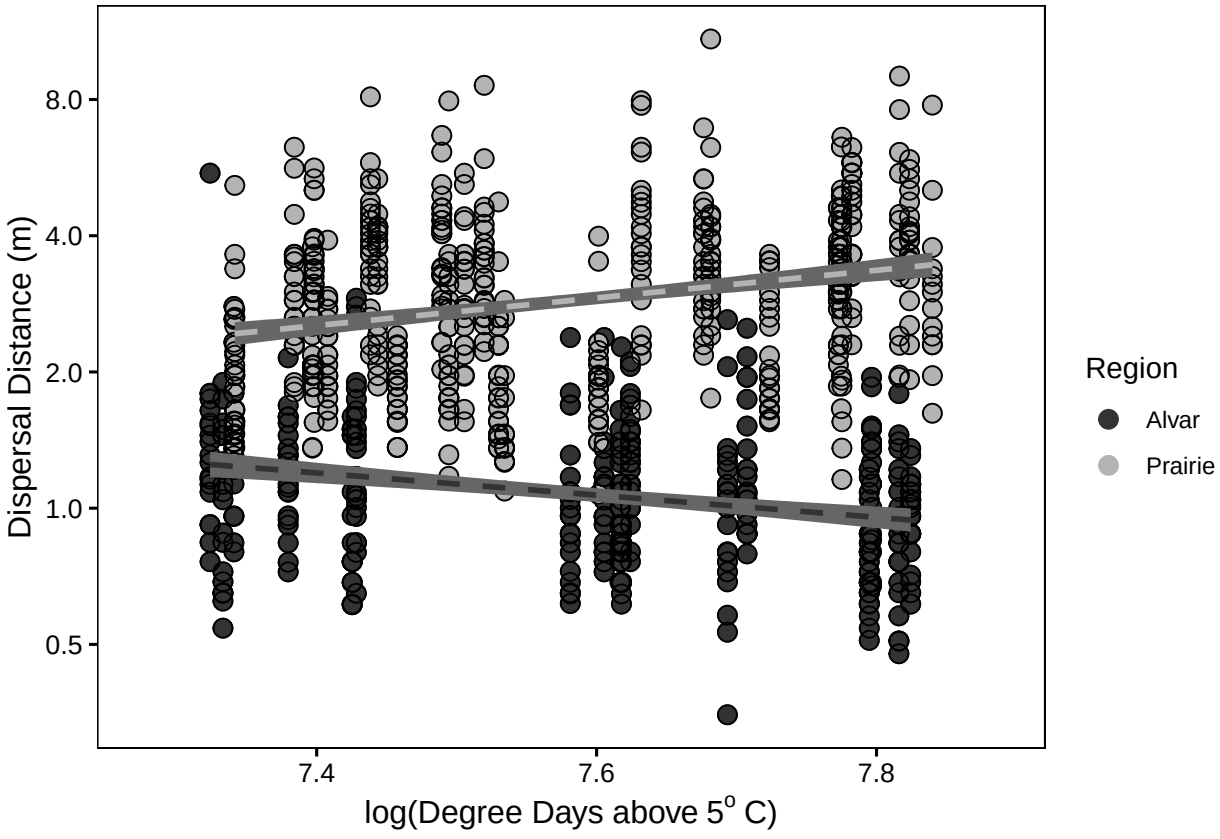
```
r.squaredGLMM(disp99_test)
```

```
##            R2m       R2c
## [1,] 0.4324425 0.6602234
```

```
##Creates Figure 4
Fig4_a <- ggplot(results_dstar, aes(x=log(DD5_TOT), y=dstar_99,col=Region,
                                    linetype = Region))+
    geom_point(aes(col=Region),size=3, position = position_dodge(width = 0.1))+
  geom_point(shape=1,size=3,colour='black', position = position_dodge(width = 0.1))+
    geom_smooth(aes(col=Region, fill = Region),method='lm', alpha = 1)+
    theme_ba()+
    scale_color_manual(values=c('gray20','gray70'))+
  scale_fill_manual(values=c('gray40','gray40'))+
  scale_linetype_manual(values=c("dashed", "dashed"))+
  xlab(bquote('log(Degree Days above' ~5^o~C*')'))+
  ylab('Dispersal Distance (m)')+
  theme (legend.position="none" )+
  scale_y_continuous(trans = 'log2')


Fig4_all <- plot_grid(Fig4_a, mylegend,
                      rel_widths = c(3,0.5),
                      nrow = 1)
```

```
## `geom_smooth()` using formula 'y ~ x'
```

```
Fig4_all
```



```
# tiff("Fig4.tiff", units="in", width=5, height=4, res=300, compression = 'lzw')
# Fig4_all
# dev.off()
```