```r
# The path to the ms directory may need to be changed for different users
# Manuscript files
pathMS <- "~/Google Drive/CardLocalAdaptation/ms"

# Data files for paper
pathDat <- paste(pathMS, "/Data", sep = "")

# Figures for paper
pathFig <- paste(pathMS, "/Figures", sep = "")

# Tables for paper
pathTab <- paste(pathMS, "/Tables", sep = "")

# Small function to make statistical significance asterisks
sigStar <- function(pvalue)
{
  if (pvalue >= 0.05) return("")
  if (pvalue < 0.05 & pvalue >= 0.01) return("*")
  if (pvalue < 0.01 & pvalue >= 0.001) return("**")
  if (pvalue < 0.001) return("***")
}

# Color palette
# Primary color:

 primShade0 <- rgb(1, 0, 0.286)
 primShade1 <- rgb(1, 0.643, 0.745)
 primShade2 <- rgb(1, 0.439, 0.6)
 primShade3 <- rgb(0.765, 0, 0.22)
 primShade4 <- rgb(0.588, 0, 0.169)

# Secondary color (1):

sec1shade0 <- rgb(0.008, 0.957, 1)
sec1shade1 <- rgb(0.643, 0.984, 1)
sec1shade2 <- rgb(0.443, 0.976, 1)
sec1shade3 <- rgb(0, 0.49, 0.514)
sec1shade4 <- rgb(0, 0.376, 0.396)

# Secondary color (2):

sec2shade0 <- rgb(0.988, 1, 0)
sec2shade1 <- rgb(0.996, 1, 0.643)
sec2shade2 <- rgb(0.992, 1, 0.439)
sec2shade3 <- rgb(0.835, 0.847, 0)
sec2shade4 <- rgb(0.643, 0.651, 0)
```

(working title) Local Physiological Adaptation in *Mimulus cardinalis*

Christopher D. Muir and Amy L. Angert

March 26, 2015

List of Figures

List of Tables

Abstract

Introduction

Local adaptation is one of the most ubiquitous observations in nature: organisms perform well in their natal environment, but poorly outside it. Local adaptation within species most often involves genetic differences in ecologically relevant traits caused by selection to different environments, leading to predictable trait-environment correlations that persist in a common garden. Although local adaptation is well-supported by data from reciprocal

transplant experiments, trait-environment associations, and ecological genetics/genomics, predicting which traits respond to selection exceeds our understanding. In particular, the role of physiological traits is poorly understood, even though physiology explains adaptation to the abiotic environment, which we need to understand as a baseline for anticipating how organisms will respond to climate change. A related problem is identifying the exact selective agents, such as abiotic factors like soil water availability or biotic interactions [etc.]... Here, we simultaneously investigate which traits and causal selective agents explain range-wide patterns of local physiological adaptation to climate in *Mimulus*, a model genus for local adaptation studies.

One surprising pattern is that populations adapt to different environments by adjusting their life history to stay within the same fundamental physiological niche rather than have the niche itself evolve. For example, mosquitoes adjust diapause length with latitude rather than evolve altered ... [**CDM:** *what other examples could we include here?*] This suggests that in many situations it is evolutionarily easier to change life history rather than fundamental physiological tolerance.

We also know little about the selective agents, such as climatic factors or biotic interactions, that are most important in driving adaptive divergence. In certain situations, such as pollinator shifts and edaphic adaptation, it is straightforward to identify putative selective agents and test predicted mechanisms. However, it is more challenging for cases in which traits vary with less specific factors like mean annual precipitation or latitude. For example, body size clines (larger size at higher latitude) have evolved repeatedly on three contients in *Drosophila subobscura*, strongly indicating selection. But which one or more of the many climatic factors that covary with latitude actually controls selection on body size? Selection might be even more complex for species that live over multiple seasons, meaning that environmental variation rather than the average environment may be a selective agent.

Linking physiological traits to potentially complex patterns of local adaptation requires integrating multiple lines of evidence from comparative, experimental, genomic studies under both lab and field conditions. Many classic and contemporary studies of local adaptation have been conducted using species from genus *Mimulus* because of its natural history, easy propogation, and genetic/genomic resources. ... Conspicuous absence of phys mechanism

1. 'intrinsic' physiological variation 2. variation in stress response 3. what ecological factors explain divergence (i.e. putative selective agents underlying local adaptation)

Conclude that observations are best explained by stabilizing selection on the fundamental niche, but spatially varying divergent selection on coordinated life history and physiological traits.

Methods

Population Selection

We used 16 populations from throughout the range of *M. cardinalis* (Table 1). Seeds were collected in the field [**CDM:** *Amy, can you explain seed collection methods?*].

Table 1: 16 Focal populations

Name	Region	Demo?	Pop gen?	Latitude	Longtiude	Elevation (mas)
HAU	South Margin	yes	yes			
CTC	South Margin	yes	no			
CUR	South Margin	yes	yes			
GRP	South Margin	no	no			
WWC	Transverse	yes	yes			
MIL	Transverse	yes	no			
WFM	Transverse	yes	no			
NMT	South Sierras	yes	yes?			
PRD	South Sierras	yes	yes			
RWD	South Sierras	yes	yes			
WNA	Central Sierras	yes	yes			
RBW	Central Sierras	yes	no			
MYU	North Sierras	yes	yes			
LIJ	North Sierras	yes	yes			
DPC	North Coast	yes	yes			
RCC	North Margin	yes	yes			

Plant propagation

On 14 April, 2014, 3-5 seeds per family were sown directly on sand (Quikrete Play Sand, Georgia, USA) watered to field capacity in RLC4 Ray Leach cone-tainers placed in RL98 98-well trays (Stuewe & Sons, Inc., Oregon, USA). We used pure sand both to facilitate root-washing and because *M. cardinalis* typically grows in sandy, riparian soils (A. Angert, pers. obs.). Two jumbo-sized cotton balls at the bottom of cone-tainers prevented sand from washing out. Cone-tainers were continuously bottom-watered during germination by placing them in medium-sized flow trays (FLOWTMD, Stuewe & Sons, Inc., Oregon, USA) filled part way with water, placed on benches in greenhouses at the University British Columbia campus in Vancouver, Canada (49°15' N, 123°15' W). Misters thoroughly wetted the top of the sand every two hours during the day. Most seeds germinated between 1

and 2 weeks, but we allowed 3 weeks before transferring seedlings to growth chambers. Germination was recorded daily from one to two weeks after sowing, and every few days thereafter. On 5 May (21 days after sowing), seedlings were transferred to one of two MODEL Growth Chambers (Conviron, Manitoba, Canada). We thinned seedlings to one plant per cone-tainer, leaving the center-most, largest plant. 702 of 768 (91.4%) had plants that could be used in the experiment. We allowed one week at constant, non stressful conditions (day: 20°C, night: 16°C) for plants to acclimate to growth chambers before starting treatments. The initial size of seedlings, measured as the length of the first true leaves, did not differ between populations, families, or treatments (Table S#).

Monitoring environment Cross check with Poorter paper
Light (PAR sensors), temp/humidity (HOBOs), soil moisture

Treatments

We used 2 watering treatments on two populations on two soils. The watering treatments are: well-watered, drought.

Temperature

We simulated typical growing season (June and July) air temperatures at the two most thermally divergent focal sites in our study, Whitewater Canyon (High Temp) and Little Jameson (Low Temp). We downloaded daily interpolated mean, minimum, and maximum air temperature from 13 years (2000-2012) at both sites from ClimateWNA (Wang et al., 2012). Daily temperatures from ClimateWNA are usually highly correlated with the air temperature recorded from data loggers in the field at these sites (A. Angert, unpub. data). Hence, the ClimateWNA temperature profiles are likely to similar to actual thermal regimes experienced by *M. cardinalis* in nature. To create realistic temperature regimes, we calculated the mean temperature trend from June to July using LOESS (**?**). The residuals were highly autocorrelated at both sites (warmer than average days are typically followed by more warm days) and there was strong correlation ($r = 0.65$) between sites (warm days in WWC were also warm in LIJ). The 'VARselect' function in the **vars** package for R (**?**) indicated that a lag two Vector Autoregression (VAR(2)) model best captured the within-site autocorrelation as well as between-site correlation in residuals. We fit and simulated from the VAR(2) using the package **dse** (Gilbert and D., 2014) in R. Simulated data closely resembled the autocorrelation and between-site correlation of the actual data. From simulated mean temperature, we next selected minimum and maximum daily temperatures. Mean, min, and max temperature were highly correlated at both sites. We chose min and max temperatures using site-specific fitted linear models between mean,

max, and min temperature, with additional variation given by normally-distributed random deviates with variance equal to the residual variance of the linear models. For each day, the nighttime (22:00 - 6:00) chamber temperature was set to the simulated minimum temperature. During the middle of the day, chamber temperature was set to the simulated maximum temperature, with a variable period of transition between min and max so that the average temperature was equal the simulated mean temperature.

Water

Wet: daily/constant irrigation with cooled water Dry: start at week #, gradually decrease watering frequency/level

[**CDM:** *Are there any data from California about soil water content through the season?*]

Growth and photosynthesis

Table 2: Traits

Trait	Units
Day of germination	day
Leaf expansion rate	mm day^{-1}
Shoot elongation rate	mm or cm day^{-1}
Harvest dry mass	g
Photosynthetic rate	μmol CO_2 m^{-2} s^{-1}
Mortality	Probability?

Day of germination We tested for population variation in germination rate, measured as Days to Germination, using a lognormal survival model fit using the survreg function in the R package **survival** version 2.38 (Therneau, 2015). The model was fit with Population as a fixed effect and Family as random effect using a Γ frailty function. The signifcance of the Population effect was determined using analysis of deviance.

```
library(survival)
pathMS <- "~/Google Drive/CardLocalAdaptation/ms"
pathDat <- paste(pathMS, "/Data", sep = "")

master <- read.csv(paste(pathDat, "/MasterDatasheet_out.csv", sep = ""),
```

```
    row.names = 1)

  # Create Surv object
        germ <- with(subset(master, !is.na(master$MinGermDay)), Surv(MinGermDay, MaxGermDay
                type = "interval2"))

        # lognormal produced best fit to data
        fitGerm <- survreg(germ ~ PopID + frailty(Family, sparse = F),
                data = subset(master, !is.na(master$MinGermDay)), dist = "lognormal")
        suppressWarnings(anova(fitGerm)) # significant family and population effects
```

```
##                                   Df Deviance Resid. Df    -2*LL
## NULL                              NA       NA  700.0000 2796.417
## PopID                        15.0000 151.9495  685.0000 2644.468
## frailty(Family, sparse = F) 106.9738 254.7233  578.0262 2389.745
##                                Pr(>Chi)
## NULL                                 NA
## PopID                       9.887489e-25
## frailty(Family, sparse = F) 4.530519e-14
```

Growth rate: leaf expansion and shoot elongation We censused longest leaf length > 1 mm on 10 days (twice per week) between 12 May and 12 June (28 to 59 days after sowing). We ceased measuring leaf length once it appeared to asymptote and growth shifted to shoot elongation. We also censused plant height on 7 days (twice per week) between 29 May and 20 June (45 to 67 days after sowing). Both leaf expansion and shoot elongation were modeled as a second-order polynomials of time with individual coefficients (separate for leaf and shoot growth) using empirical Bayes' estimates from linear mixed-effects models fit using the R package **lme4** version 1.1-7 (Bates et al., 2014).

```
  library(lme4)
```

```
## Loading required package:  Matrix
## Loading required package:  Rcpp
```

```
  library(lmerTest)
```

```
##
## Attaching package:  'lmerTest'
##
## The following object is masked from 'package:lme4':
##
##    lmer
```

```
##
## The following object is masked from 'package:stats':
##
##     step
```

```
  # New (blank) datasheet on which to add data
  newdata <- read.csv(paste(pathDat, "/MasterDatasheet_in.csv", sep = ""))

  # Read in LER data
  LERdata <- read.csv(paste(pathDat, "/LERdata.csv", sep = ""))

  #
        # Empirical Bayes Approach: get coefficients for every plant
        #

        # suppressWarnings(mmLER <- lmer(log(LLL) ~ poly(DayN, 2, raw = T) +
    # (poly(DayN, 2, raw = T)|indiv), data = LERdata))
  # save(mmLER, file = paste(pathDat, "/mmLER.RData", sep = ""))
  load(file = paste(pathDat, "/mmLER.RData", sep = ""))

  # Extract coefficients and add to datasheet
        Y <- t(unlist(fixef(mmLER)) + t(ranef(mmLER)$indiv))[match(newdata$indiv,
                rownames(ranef(mmLER)$indiv)), ]
        colnames(Y) <- c("b0_LER", "b1_LER", "b2_LER")
  newdata <- cbind(newdata, Y)

  # Model-based LER
        # Timeframe for Cool treatment: Day 0 - 31
        # Timeframe for Hot treatment: Day 0 - 24
        newdata$LERstart <- exp(newdata$b0_LER)
        newdata$LERend <- exp(newdata$b0_LER +
                                        newdata$b1_LER * ifelse(newdata$TempTrt == "Cool",
                                        newdata$b2_LER * ifelse(newdata$TempTrt == "Cool",
        newdata$LER_AbsGrowth <- (newdata$LERend - newdata$LERstart) /
                ifelse(newdata$TempTrt == "Cool", 31, 24)

  # Mixed model ANOVA with family as random factor fit using step down procedure
        mm <- lmer(LER_AbsGrowth ~ AvgGermDay + PopID * TempTrt * WaterTrt + (1|Family),
                data = master)
        fitLER <- step(mm, reduce.random = F) # includes pop, temp, h2o, population x water
        # fitLER <- step(mm, ddf = "Kenward-Roger", reduce.random = F) # same result using
```

```r
# Read in SER data
SERdata <- read.csv(paste(pathDat, "/SERdata.csv", sep = ""))

#
# Empirical Bayes Approach: get coefficients for every plant
      #

      # suppressWarnings(mmSER <- lmer(log(height + 1) ~ poly(DayN, 2, raw = T) +
  # (poly(DayN, 2, raw = T)|indiv), data = SERdata))
# save(mmSER, file = paste(pathDat, "/mmSER.RData", sep = ""))
load(file = paste(pathDat, "/mmSER.RData", sep = ""))

# Extract coefficients and add to datasheet
Y <- t(unlist(fixef(mmSER)) + t(ranef(mmSER)$indiv))[match(newdata$indiv,
              rownames(ranef(mmSER)$indiv)), ]
      colnames(Y) <- c("b0_SER", "b1_SER", "b2_SER")
newdata <- cbind(newdata, Y)

# Model-based SER
      # Timeframe for Cool treatment: Day 0 - 31
      # Timeframe for Hot treatment: Day 0 - 24
      newdata$SERstart <- exp(newdata$b0_SER)
      newdata$SERend <- exp(newdata$b0_SER +
                                    newdata$b1_SER * ifelse(newdata$TempTrt == "Cool",
                                    newdata$b2_SER * ifelse(newdata$TempTrt == "Cool",
      newdata$SER_AbsGrowth <- (newdata$SERend - newdata$SERstart) /
              ifelse(newdata$TempTrt == "Cool", 31, 24)

mm <- lmer(SER_AbsGrowth ~ AvgGermDay + PopID * TempTrt * WaterTrt +
              (1|Family), data = master)
      fitSER <- step(mm, reduce.random = F) #includes pop, temp, h2o, temp x water
```

Photosynthesis During the week of 10 to 16 June (57 to 63 days after sowing), we measured daytime photosynthetic rate and stomatal conductance on a subset of 329 plants evenly spread between treatments and families within populations. The youngest, fully-expanded leaf acclimated for 3 minutes to reach steady state in a 6 cm^2 chamber of a LI-COR 6400XT Portable Photosynthesis System (LI-COR Biosciences, Lincoln, Nebraska). All measurements were made at ambient light (400 μmol m^{-2} s^{-1}), temperature, and mod-

erate relative humidity. During this period, we suspended normal day-to-day temperature fluctuations and set daytime temperatures to its average for that period (Cool: 26.5°; Hot: 36.1° [**CDM:** *get exact temps used*]) so that all plants within a temperature treatment were measured under the same conditions.

```
  # Photosynthetic rate
  mm <- lmer(Photo ~ PopID * TempTrt * WaterTrt + (1|Family),
    data = subset(master, !is.na(master$Photo)))
  fitPhoto <- step(mm, reduce.random = F)

  # Intrinsic photosynthetic rate
  # = residuals of Conductance - Photosynthesis regression
  # Removed one outlier (may change if data are rearranged!)
       mm <- lmer(resPhoto ~ PopID * TempTrt * WaterTrt + (1|Family),
    data = subset(master, !is.na(master$Photo))[-52, ])
       fitIntrnPhoto <- step(mm, reduce.random = F)
```

Mortality We assayed mortality during twice-weekly growth measurements. We could not get GLMM with Family effects to converge, so we used GLM with a quasibinomial error structure and assessed signifiance using Type 2 Analysis of Deviance with the R package **car**.

```
  library(car)

  fitMort <- glm(DiedFromStress ~ PopID * TempTrt * WaterTrt,
    data = subset(master, !is.na(master$DiedFromStress)), family = "quasibinomial")

  Anova(fitMort, type = 2)
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: DiedFromStress
##                         LR Chisq Df Pr(>Chisq)
## PopID                     38.169 15  0.0008519 ***
## TempTrt                  247.763  1  < 2.2e-16 ***
## WaterTrt                  79.508  1  < 2.2e-16 ***
## PopID:TempTrt             26.188 15  0.0360962 *
## PopID:WaterTrt            11.547 15  0.7129658
## TempTrt:WaterTrt          48.543  1  3.231e-12 ***
## PopID:TempTrt:WaterTrt     8.685 15  0.8934160
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Biomass (show correlation between growth rate and biomass?)

Intrinsic variation and plasticity

[**CDM:** *which traits: germination, growth rate, biomass, biomass allometry, leaf traits, photosynthesis?, mortality?*] We tested for Population, Treatment, and Population × Treatment interactions using mixed model ANOVA with Family included as a random factor. Models were fit by restricted maximum likelihood using lmer from the R package **lme4** (Bates et al., 2014). Significant fixed effect terms were selected using a step-wise backward elimination procedure implemented with the step function in the R package **lmerTest** version 2.0-11 (Kuznetsova et al., 2014). Denominator degrees of freedom for F-tests were estimated using Satterthwaite's approximation. Significant Population effect indicate intrinsic trait differences; significant Population × Treatment effects indicate population differences in plasticity. For growth rate, we also accounted for differences in germination rate by including day of germination as a factor.

Principal components of germination, growth, and photosynthesis

```
popMeans <- data.frame(
  germ = c(coef(fitGerm)[1], coef(fitGerm)[1] + coef(fitGerm)[2:16]),
  LER = fitLER$lsmeans.table[1:16, "Estimate"],
  SER = fitSER$lsmeans.table[1:16, "Estimate"],
  photo = fitIntrnPhoto$lsmeans.table[1:16, "Estimate"])

traitPC <- prcomp(popMeans, center = T, scale. = T)
focClim$PC1 <- traitPC$x[, "PC1"]

# Percent variance explained by first principal component
pve <- paste(round(100 * traitPC$sdev[1] ^ 2 / sum(traitPC$sdev ^ 2), 1), "%")

# Plot latitude versus PC1
pdf(paste(pathFig, "/Figure_PC1vLat.pdf", sep = ""), 5, 5)
par(mar = c(5, 5, 1, 1))
plot(focClim$Lat, focClim$PC1, xlim = c(32, 44), ylim = c(-4, 4), axes = F,
  frame.plot = T, type = "n", xlab = "", ylab = "")
title(xlab = "Latitude", cex.lab = 1.5)
```

```
  title(ylab = substitute(plain(Principal~Component~1)~group("(", x, ")"),
    list(x = pve)))
  axis(1, at = seq(32, 44, 4), labels = c(expression(32*degree), expression(36*degree),
    expression(40*degree), expression(44*degree)), lwd = 0, lwd.ticks = 1)
  axis(2, at = seq(-4, 4, 2), las = 1, lwd = 0, lwd.ticks = 1)
  fitLat <- lm(PC1 ~ Latitude, data = focClim)

  # polygon of confidence intervals
  x <- seq(min(focClim$Latitude), max(focClim$Latitude), l = 1e3)
  u <- predict(fitLat, new = data.frame(Latitude = x))
  s <- sqrt(sapply(x, function(X) t(c(1, X)) %*% vcov(fitLat) %*% c(1, X)))
  confInt <- data.frame(lowerCI = qnorm(0.025, u, s), upperCI = qnorm(0.975, u, s))
  polygon(c(x, rev(x)), c(confInt$lowerCI, rev(confInt$upperCI)),
    col = "grey80", border = NA)
  lines(x, confInt$lowerCI, lwd = 2, lty = 2)
  lines(x, confInt$upperCI, lwd = 2, lty = 2)
  lines(range(focClim$Lat), rev(range(predict(fitLat))), lwd = 2)
  points(focClim$Lat, focClim$PC1, pch = 21, col = "black", bg = primShade0, cex = 1.5,
    lwd = 2)
  dev.off()
```

```
## pdf
##   2
```

```
    # plot axes could be range maps with areas highlighted (x-axis) and short to tall plan
  #plot(focClim£Lat, traitPC£x[, "PC1"])
  #cor.test(focClim£Lat, traitPC£x[, "PC1"])
  #cor.test(popMeans[, 4], traitPC£x[, "PC1"])
  #plot(focClim£bio16_cva, traitPC£x[, "PC1"])
```

Selective agents and environmental correlates

```
library(VSURF)
```

```
## Loading required package: randomForest
## randomForest 4.6-10
## Type rfNews() to see new features/changes/bug fixes.
## Loading required package: rpart
## Loading required package: doParallel
## Loading required package: foreach
## Loading required package: iterators
```

```
## Loading required package:  parallel
```

```
  # Load VSURF results of climatic variables associated with latitude at cardinalis occuren
  # Note that occurences have been randomly thinned to account for variation in sampling ef
  load(paste(pathDat, "/clim2lat_fit_t1.RData", sep = "")) # fit_t1
  vsurfLat <- fit_t1

  # Figure of variable importance for latitude
  pdf(paste(pathFig, "/Figure_clim2lat_imp.pdf", sep = ""), 4, 5)
  par(mar = c(5, 6, 1, 1))
  plot(0, 0, type = "n", xlim = c(0, ceiling(vsurfLat$ord.imp$x)[1]), ylim = c(1,
    length(vsurfLat$varselect.pred)), axes = F, frame.plot = T, xlab = "Importance",
    ylab = "", cex.lab = 1.5)
  for (i in 1:length(vsurfLat$varselect.pred))
  {
     lines(x = c(-1, length(vsurfLat$varselect.pred)), y = c(i, i), col = "grey")
  }
  points(vsurfLat$ord.imp$x[1:length(vsurfLat$varselect.pred)],
    length(vsurfLat$varselect.pred):1, pch = 21, col = "black", bg = "white", cex = 1.5)
        mtext(climVars[vsurfLat$varselect.pred], side = 2,
    at = length(vsurfLat$varselect.pred):1, las = 1, line = 1)
```

```
## Error in as.graphicsAnnot(text):  object 'climVars' not found
```

```
        axis(1, lwd = 0, lwd.ticks = 1)
        dev.off()
```

```
  # Columns with climatic variables to test
  load(paste(pathDat, "/Exploratory_climate_variables.RData", sep = "")) #climVars

  # Variable selection using random forests (VSURF)
  linmod.PC1 <- as.formula(paste("PC1 ~ ", paste(climVars, collapse = " + "), sep = ""))
  set.seed(69646)
  vsurfPC1 <- VSURF.parallel(linmod.PC1, data = focClim[, c("PC1", climVars)],
    ncores = 3)
```

```
## Warning in VSURF.parallel.formula(linmod.PC1, data = focClim[, c("PC1", :
VSURF with a formula-type call outputs selected variables
## which are indices of the input matrix based on the formula:
## you may reorder these to get indices of the original data
```

```r
  # Plot most important climatic predictor versus PC1
  pdf(paste(pathFig, "/Figure_bio16cva_v_Lat.pdf", sep = ""), 5, 5)
  par(mar = c(5, 5, 1, 1))
  plot(focClim$bio16_cva, focClim$PC1, xlim = c(0.25, 0.55), ylim = c(-4, 4), axes = F,
    frame.plot = T, type = "n", xlab = "", ylab = "")
  title(xlab = "Precipitation of Wettest Quarter", cex.lab = 1.5)
  title(xlab = "(Coefficient of variation)", cex.lab = 1, line = 4)
  title(ylab = substitute(plain(Principal~Component~1)~group("(", x, ")"),
    list(x = pve)))
  axis(1, lwd = 0, lwd.ticks = 1)
  axis(2, at = seq(-4, 4, 2), las = 1, lwd = 0, lwd.ticks = 1)
  fitBio16cva <- lm(PC1 ~ bio16_cva, data = focClim)

  # polygon of confidence intervals
  x <- seq(min(focClim$bio16_cva), max(focClim$bio16_cva), l = 1e3)
  u <- predict(fitBio16cva, new = data.frame(bio16_cva = x))
  s <- sqrt(sapply(x, function(X) t(c(1, X)) %*% vcov(fitBio16cva) %*% c(1, X)))
  confInt <- data.frame(lowerCI = qnorm(0.025, u, s), upperCI = qnorm(0.975, u, s))
  polygon(c(x, rev(x)), c(confInt$lowerCI, rev(confInt$upperCI)),
    col = "grey80", border = NA)
  lines(x, confInt$lowerCI, lwd = 2, lty = 2)
  lines(x, confInt$upperCI, lwd = 2, lty = 2)
  lines(range(focClim$bio16_cva), range(predict(fitBio16cva)), lwd = 2)
  points(focClim$bio16_cva, focClim$PC1, pch = 21, col = "black", bg = primShade0,
    cex = 1.5, lwd = 2)
  dev.off()
```

```
## pdf
##   2
```

```r
  # Overlap between latitude predictors and phenotype predictors
  imp2 <- climVars[vsurfPC1$varselect.pred]
  nImpPC1 <- length(imp2)

  vsurfLat$varselect.pred
```

```
##  [1] 45 43 42 20 47 41 46 39 19 13
```

```r
  imp1 <- climVars[vsurfLat$varselect.pred]
  nImpLat <- length(imp1)

  # Figure of variable importance for traits
```

```
        pdf(paste(pathFig, "/Figure_clim2pc1_imp.pdf", sep = ""), 4, 5)
        par(mar = c(5, 1, 1, 6))

        plot(0, 0, type = "n", xlim = c(0, 30), ylim = c(1, nImpLat),
                axes = F, frame.plot = T, xlab = "Importance", ylab = "", cex.lab = 1.5)
        for (i in (nImpLat:(nImpLat - nImpPC1 + 1))) lines(x = c(-100, 100), y = c(i, i),
                col = "grey")
        points(100 * vsurfPC1$ord.imp$x[1:nImpPC1], nImpLat:(nImpLat - nImpPC1 + 1), pch =
                col = "black", bg = "white", cex = 1.5)
        mtext(climVars[vsurfPC1$varselect.pred], side = 4, las = 1, line = 1,
    at = nImpLat:(nImpLat - nImpPC1 + 1),
                font = ifelse(vsurfPC1$varselect.pred %in% vsurfLat$varselect.pred,     2,
        axis(1, lwd = 0, lwd.ticks = 1, labels = seq(0, 3, 0.5), at = seq(0, 30, 5))
  dev.off()
```

```
## pdf
##   2
```

(this also goes into the goal of trying to figure out ecological selective agent driving trait variation). Why does latitude predict trait variation, often better than any of the climatic variables that vary with latitude? Some hypotheses:

- H1: Latitude encompasses multivariate climatic dimension.
- H2: Latitude is substitute for average selective environment. i.e. populations are connected by gene flow, such that alleles at a given site must do well there and in surrounding areas.

Predictions:

- P1: Climatic axis in high dimensional climate space should correlate with latitude (across card sites especially, not just any site [using pseudoabsences]). Focal site position in high-dimensional climatic space should correlated with traits as well or better than latitude.
- P2: Climate (averaged over nearby sites) varies with latitude and the nearby average climate predicts trait values as well or better than latitude.
- P1+2: These are not mutually exclusive and both could be imortant. i.e. average selective environment in high-dimensional climate space predicts trait values as well or better than latitude.

Methods specific to latitude vs. trait correlation: We used Random Forest regression (Liaw and Wiener, 2002) to identify putative climatic factors underlying trait-latitude associations in *M. cardinalis*. We used a comprehensive set of [**CDM:** *358 in total before*

thinning] recent (since 2000) known *M. cardinalis* occurences recorded from herbarium records and an exhaustive field survey in 2010-11 (Angert, 2015). For each occurence, we used a 90m digital elevation model from HydroSHEDS (Lehner et al., 2006) to extract elevation. Monthly interpolated climate layers were calculated using ClimateWNA (Wang et al., 2012), which accurately downscales climate data specifically for the rugged topography of western North America. For each occurence, we calculated bioclimatic variables using the biovars function in the R package **dismo** (Hijmans et al., 2014). In total, we included X environmental variables from ClimateWNA and 17? bioclimatic variables (Table S#). For each variable, we calculated both a 30-year normal by averging annual values between 1981 and 2010 and 30-year coefficient of variation, a standardized metric of interannual climatic variation). Temperatures were converted to Kelvin to be on an interval scale appropriate for calculating the coefficient of variation.

Selection at given site might not reflect simply the local environment, but rather the average environment in a surrounding region, since alleles will experience selection across populations linked by migration. We therefore calculated spatially averaged [**CDM:** *maybe spatially-integrated is better term?*] climate means and coefficients of variation. We sampled climate at 1000 random points (at 90-m resolition) within a 10 km buffer around focal populations. Since *M. cardinalis* is found exclusively in riparian areas, we only selected points along streams using the National Hydrogeoraphy Dataset (United States Geological Survey, 2015). Climatic means and CVs were weighted by the predicted suitability of each point, using ENM models from Angert (2015).

To identify climatic variables that were closely associated with latitude in *M. cardinalis*, we used a "Variable Selection Using Random Forest" (VSURF) algorithm in the R **VSURF** version 0.8.2 (Genuer et al., 2014). In this algorithm, climatic variables are regressed against latitude, and the most important variables were selected after a three-stage selection process that progressively excludes unimportant variables.

Predictions: looked for climatic variables (means or CVs) that were strong predictors of both population latitudes (from all card pops) and trait variation in 16 focal popultions. If spatial average environment is important, then expect greater predictive power and closer association with latitude-predictors than point averages.

Evaluated fit on test set (correlation between predicted versus actual latitude) and on pseudoabsence set. , Despite this common, and long-standing pattern, only certain traits In particular, adaptation to the local abiotic environment is generally thought to be the major cause of natural selection. Physiology, the study of organismal function, connects fitness to the abiotic environment, yet we actually know little about the physiological basis of adaptation to different environments within a species. To address this question, we looked at physiological variation across a broad latitudual gradient within *Mimulus cardinalis*...

For 16 focal populations, I took the average climate within a 10 km buffer, clipped to streams based on the National Hydrological Database (United States Geological Survey, 2015) weighted by climatic suitability determined [multimodel average/ensemble?] from ecological niche modeling (Angert, 2015).

Table 3: Climatic variables

Climate variable	Abbreviation
DD_0	degree-days below 0°C(chilling degree-days)
DD5	degree-days above 5°C(growing degree-days)
DD_18	degree-days below 18°C(heating degree-days)
DD18	degree-days above 18°C(cooling degree-days)
NFFD	number of frost-free days
PAS	precipitation as snow (mm) between August in previous year and July in current
Eref	Hargreaves reference evaporation (mm)
CMD	Hargreaves climativ moisture deficit (mm)
MAR	mean annual solar radiation (NOTE: removing because too many missing values)
RH	mean annual relative humidity
bio1	Annual Mean Temperature
bio2	Mean Diurnal Range (Mean of monthly (max temp - min temp))
bio3	Isothermality (bio2/bio7) (* 100)
bio4	Temperature Seasonality (standard deviation *100)
bio5	Max Temperature of Warmest Month
bio6	Min Temperature of Coldest Month
bio7	Temperature Annual Range (bio5-bio6)
bio8	Mean Temperature of Wettest Quarter
bio9	Mean Temperature of Driest Quarter
bio10	Mean Temperature of Warmest Quarter
bio11	Mean Temperature of Coldest Quarter
bio12	Annual Precipitation
bio15	Precipitation Seasonality (Coefficient of Variation)
bio16	Precipitation of Wettest Quarter
bio17	Precipitation of Driest Quarter
bio18	Precipitation of Warmest Quarter
bio19	Precipitation of Coldest Quarter

Table 4: Population, treatment effects

	Germination survreg	Leaf expansion lmer	Stem elongation lmer	Photosynthesis lmer	Intrin. Photo lmer
Population	***	***	***	***	***
Temperature	NA	***	***	***	
Water	NA	***	***	***	
Pop × Temp	NA			*	
Pop × Water	NA	*			
Temp × Water	NA		***		
Pop × Temp × Water	NA				

Results

Coordinated latitudinal clines in germination, growth, and photosynthesis

Little evidence population by treatment interactions

Latitude and phenotypic clines are associated with more variable precipitation

Table 5: CAPTION

	SS	MS	df1	df2	*F*	P
Day of Germination	18.2	18.2	1	633.1	35.2	4.78×10^{-9}
Population	22.3	1.5	15	121.5	4.3	2.57×10^{-6}
Temperature	114.4	114.4	1	567.7	352.3	1.64×10^{-61}
Water	4.2	4.2	1	567.7	12.6	4.12×10^{-4}
Population × Temperature	3.1	0.2	15	553.3	0.6	0.86
Population × Water	8.9	0.6	15	567.8	1.8	0.03
Temperature × Water	0	0	1	550.5	0	0.99
Population × Temperature × Water	4	0.3	15	535.4	0.8	0.69

Table 6: CAPTION

	SS	MS	df1	df2	*F*	P
Day of Germination	6.2	6.2	1	656.2	20.9	5.9×10^{-6}
Population	10	0.7	15	109.9	4.3	3.69×10^{-6}
Temperature	70.3	70.3	1	568	445.3	2.05×10^{-73}
Water	2.9	2.9	1	567.7	17.8	2.92×10^{-5}
Population $\times$ Temperature	2.7	0.2	15	552.7	1.2	0.25
Population $\times$ Water	2.6	0.2	15	539.8	1.1	0.33
Temperature $\times$ Water	2.3	2.3	1	565.2	14.7	1.37×10^{-4}
Population $\times$ Temperature $\times$ Water	1.5	0.1	15	521.4	0.6	0.86

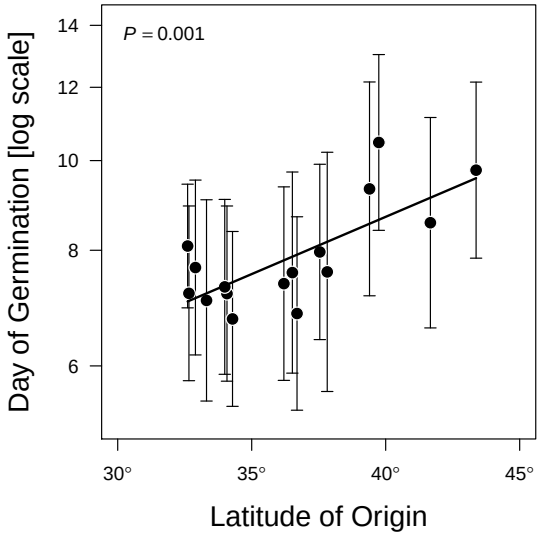

Figure 1: CAPTION

Latitude versus growth rate

Discussion

How organisms adapt depends on the grain of the environment, ..., and gene flow between populations

Possibilities: 1. no trait variation: no environmental variation, high gene flow, generalist (convex) 2. variation in mean: 3. variation in plasticity: convex?,

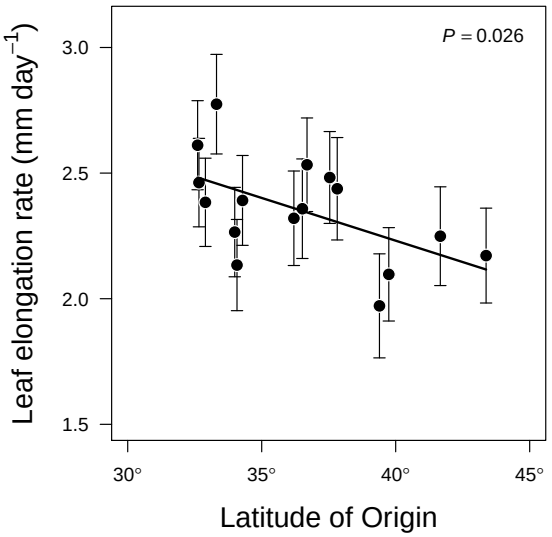

Figure 2: CAPTION

our data are most consistent with: - concave tradeoffs, fine-grained (coarse+concave would predict alt morphs), high gene flow

References

Angert, A. L., 2015. Cardinalis enm. TBD 1:1–1.

Bates, D., M. Maechler, B. Bolker, and S. Walker, 2014. lme4: Linear mixed-effects models using Eigen and S4. URL `http://CRAN.R-project.org/package=lme4`. R package version 1.1-7.

Genuer, R., J.-M. Poggi, and C. Tuleau-Malot, 2014. VSURF: Variable Selection Using Random Forests. URL `http://CRAN.R-project.org/package=VSURF`. R package version 0.8.2.

Gilbert and P. D., 2014. Brief User's Guide: Dynamic Systems Estimation. URL `http://cran.r-project.org/web/packages/dse/vignettes/dse-guide.pdf`.

Hadfield, J. D., 2010. Mcmc methods for multi-response generalized linear mixed models: The MCMCglmm R package. Journal of Statistical Software 33:1–22. URL `http://www.jstatsoft.org/v33/i02/`.

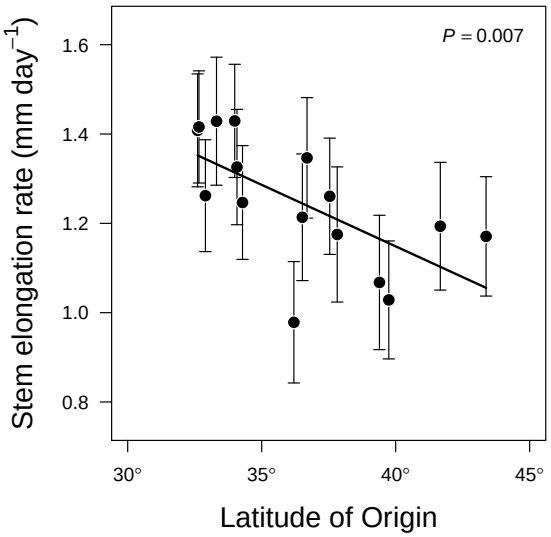

Figure 3: CAPTION

Hijmans, R. J., S. Phillips, J. Leathwick, and J. Elith, 2014. dismo: Species distribution modeling. URL `http://CRAN.R-project.org/package=dismo`. R package version 1.0-5.

Kuznetsova, A., P. Bruun Brockhoff, and R. Haubo Bojesen Christensen, 2014. lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). URL `http://CRAN.R-project.org/package=lmerTest`. R package version 2.0-11.

Lehner, B., K. Verdin, and A. Jarvis, 2006. HydroSHEDS technical documentation. World Wildlife Fund, Washington, D.C. URL `www.worldwildlife.org/hydrosheds`.

Liaw, A. and M. Wiener, 2002. Classification and regression by randomforest. R News 2:18–22. URL `http://CRAN.R-project.org/doc/Rnews/`.

Therneau, T. M., 2015. survival: A Package for Survival Analysis in S. URL `http://CRAN.R-project.org/package=survival`. R package version 2.38.

United States Geological Survey, 2015. National Hydrogeoraphy Dataset. United States Geological Survey, Washington, D.C. URL `http://nhd.usgs.gov/index.html`.

Wang, T., A. Hamann, D. L. Spittlehouse, and T. Q. Murdock, 2012. ClimateWNA – high-resolution spatial climate data for western North America. Journal of Applied Meteorology and Climatology 51:16–29.

Table 7: Initial size of seedlings did not vary among Populations, Families, or Treatments. We used a censored Gaussian model of initial size at the outset of the experiment (longest leaf length of the first true leaves). The model was censored because we could not accurately measure leaves less than 0.25 mm with digital callipers (217 of 702, 30.9%, were too small). We fit models using a Bayesian MCMC method implemented using the MCMCglmm function with default priors in the R package **MCMCglmm** version 2.17 (Hadfield, 2010). We estimated the posterior distribution from 1000 samples of an MCMC chain run for 10^5 steps after a 10^4 step burn-in. We step-wise backward elimination procedure to find the best-supported model according to Deviance Information Criterion (DIC).

Model	Random	DIC
Population + Water + Temperature + Population:Water + Population:Temperature + Water:Temperature + Population:Water:Temperature	Family	1638
Population + Water + Temperature + Population:Water + Population:Temperature + Water:Temperature	Family	1605.2
Population + Water + Temperature + Population:Water + Population:Temperature	Family	1603.4
Population + Water + Temperature + Population:Water + Water:Temperature	Family	1577.5
Population + Water + Temperature + Population:Temperature + Water:Temperature	Family	1579.9
Population + Water + Temperature + Population:Water	Family	1577.3
Population + Water + Temperature + Water:Temperature	Family	1550.5
Population + Water + Temperature	Family	1549.3
Population + Water	Family	1541.7
Population + Temperature	Family	1546.8
Water + Temperature	Family	1551.1
Population	Family	1541.9
Water	Family	1543.9
-	Family	1541.7
-	-	1538.3