Evaluating Showy Milkweed (*Asclepias speciosa*) ecotypic variation in adaptive traits using a genecological approach

Part 2- Testing associations among climate variables and plant characteristics.

Background

Climate variables have a broad influence on plant energy balance, productivity, and reproduction, and thus serve as important baseline selection pressures for modeling plant fitness about the landscape. Because plants are sessile (stationary) and unable to effectively relocate away from intense selection pressures, many species have evolved a broad range of adaptive traits that help them survive in the prevailing and sometimes variable environmental conditions. Central to climate-based selection pressures are the seasonal and annual patterns of temperature and precipitation and how they interact (Table 1). Because areas of the western US vary greatly in topography, elevation and latitude, and contentiality, there is wide variation in climate patterns and thus putative selection pressures. These variations undoubtedly result in the development of abrupt changes in plant traits across the landscape as specialized populations (or ecotypes) display adaptations that enhance fitness in unique local environments.

Secondary factors that shape variation in plant characteristics on the landscape include variations in population genetics (diversity, population size, drift, gene flow, and inbreeding), as well as the fact that that individual plants can display phenotypic plasticity (where the same genotype can produce a range of phenotypes depending on environmental inputs and plant epigenetic responses). In phenotypic plasticity, a single genotype may produce different phenotypes depending on environmental inputs. Phenotypic plasticity most often manifests through epigenetic controls that influence the expression of genes that regulate growth, development, or physiological processes. While individuals experience changes in characteristics due to phenotypic plasticity that ultimately enhances their survival and fitness, phenotypic plasticity is not considered adaptive, because selection pressures do not induce changes in the genetic composition of populations.

Scientific Objective

In **part two** of this lab, you will construct and test hypotheses about how climate and other abiotic factors shape putative adaptive traits in western showy milkweed populations. In efforts to identify the extent of how climate variables impact the phenotypic variation of populations of milkweed, we will test correlations among remotely sensed climate data (Climate WNA; Wang et al. 2012) in conjunction with measured traits from the CWI milkweed common garden (Table 1, 2). Along the way, you'll be exposed to some additional methods that researchers use to analyze, simplify, and interpret multivariate data sets. Your inquiry and analyses will further help shed light on potential explanations for how climate and elevation drive adaptation in local milkweed ecotypes of this species.

Student Learning Objectives

- Use foundational knowledge of plant physiology and critical thinking to construct hypotheses and predictions about relationships among climate variables and plant traits
- Gain exposure to:
    - conducting data analysis in R and RStudio
    - Principle Components analyses (PCA) conceptual framework and interpreting results of PCAs
    - Interpreting Pearson correlation coefficients
- Evaluate hypotheses based on statistical results
- Discuss results in the context of plant adaptations to prevailing local climates
- Discuss potential sources of variation and error as well as opportunities to improve the resolution of this work.

Statistical background- principal components methods and results.- Dusty will go over this in class- look over one time, but don't stress about these details-

Because relating plant characters to climate variables is often complex and requires multivariate modeling, as a first step, we will use Pearson correlation coefficients to test for general correlations that can inform future, more detailed analyses (e.g. this methods is a first step to help reduce non- explanatory variables to streamline the multivariate modeling process. One method that scientists may use to simplify the complexity of datasets with multiple variables is principal components analysis (PCA). PCA is a mathematical procedure that transforms a number of (possibly) correlated variables into a (smaller) number of uncorrelated variables called principal components.

The section below summarizes the results of a PCA that was conducted for all of the variables in (table 1; raw data is in: Ecology_Morph_Climate_Corr.csv).

Table 1: Definitions and units of plant characters measured among 354 showy milkweed individuals from 26 populations of showy milkweed.

Measured Characteristics	**Descriptions (units)**
Biomass	Measured weight of dried plant (g)
Total plant height	Height measurement from the emergence from soil (cm)
Stem length	Length of the 3 longest stems from the emergence from soil to the petiole of the last leaf (cm)
Total stem length	Sum length of three longest stems (cm)
Longest stem length	The measurement of the longest stem out of the three measured (cm)
Canopy area	Product of the widest point of canopy and a perpendicular widest point (cm^2)
Stem growth	CM/day from min stem length to max
Est. leaf area	Product of leaf length x leaf width (cm^2)
Leaf L/W ratio	Quotient of leaf length/leaf width
Dimension 1	Multivariate principal component: Stem growth (17%), Stem length (17%), Height (16%), Biomass (15%)
Dimension 2	Multivariate principal component: Length/Width ratio (69%), Leaf area (14%)

Results

- Single-factor ANOVAS calculated for each variable among 36 populations indicate significance differences in all measured traits between populations ($p \leq 0.0001$ (exception: number of stems and leaf area p= 0.013 and 0.019 respectively).
- 80% of variation among populations is explained by the first two PCs (PC1= 66%, PC2= 14%
    - Dimension 1: Stem growth (17%), Stem length (17%), Height (16%), Biomass (15%)
        - Higher PC1 loadings = plants with larger size and higher growth rate (Fig. 2).
    - Dimension 2: Length/Width ratio (69%), Leaf area (14%).
        - Higher PC2loadings = plants with longer/thinner leaves and smaller leaf areas (Fig. 2).

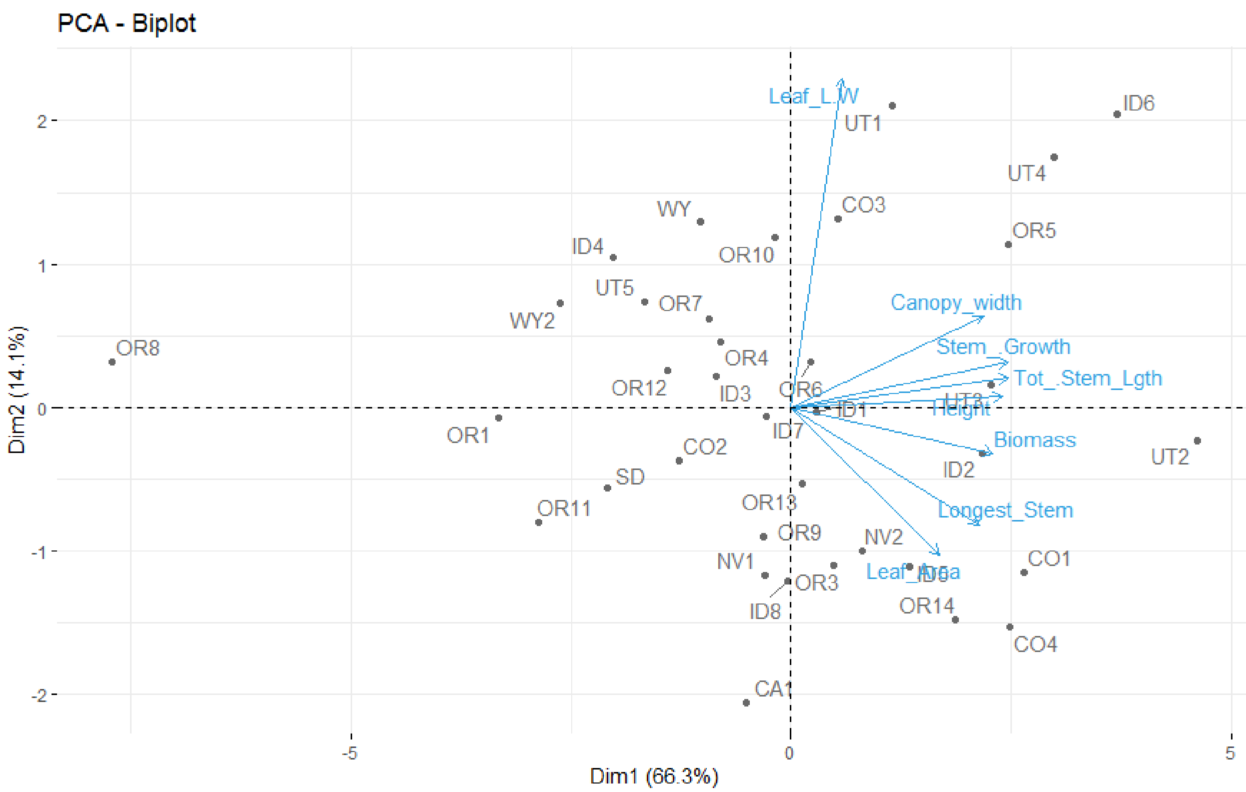

Figure 2. Biplot with principal components analysis separations with dimensional loading factors among 36 populations of Showy Milkweed. Heads of vectors indicate increasing values for each denoted variable.

Interpreting Results

- PCA results provide support that showy milkweed displays genecological differences among growth form traits, growth rates, and leaf characters among populations in the western US.
- Significant difference among populations (ANOVAs) are attributable to differences driven by a few unique populations (OR8, CA1, ID6, UT2, CA1).

<u>Lab Procedure</u>

<u>Hypothesis testing</u>

1. Acquire and conduct a cursory review of the following resources.
   a. Download:
      i. Milkweed genecolgy worksheet.- (this will be the thing you turn in for the lab notebook)
      ii. Ecology_Morph_Climate_Corr.csv
      iii. Ecology_Morph_Climate_Corr - R file
      iv. OPTIONAL: climate WNA help document (for expanded descriptions of climate variables if needed)
   b. Familiarize yourself with basic premises regarding plant adaptations:
      i. Chapter 6 of the Smith and Smith Ecology textbook (sections 6.7-6.10; beginning on pp 100).
      ii. You may also find the discussion of ecotypic variation (section 5.8 pp 79-86) helpful.
2. At the prompts (**Denoted with ♦**) complete actions and answer questions in the Milkweed genecolgy worksheet.

3. Carefully review the plant traits and climate descriptions in Tables 1 and 2.
4. variables (these climate variables may be different for each plant trait)

♦ Use **Tables 1 and 2** to identify at least two plant traits that you wish to evaluate.

Table 2: Definitions and units of the climate variables derived from climate WNA (Climate WNA; Wang et al. 2012).

Climate Variable	Definition (unit)
Latitude	Angular distance north or south of Earth's equator (°)
Longitude	Angular distance east or west of Prime Meridian (°)
Elevation	Height above sea level (m)
MAT	Mean average temperature (°C)
MWMT	Mean warmest month temperature (°C)
MCMT	Mean coldest month temperature (°C)
TD	Temperature difference between MWMT and MCMT (°C)- higher values of this variable correspond to greater continentality
MAP	Mean annual precipitation (mm)
MSP	Mean annual summer precipitation (mm)
AHM	Annual heat-moisture index (°C mm^{-1})
SHM	Summer heat-moisture index (°C mm^{-1})
DD<0	degree-days below 0°C, (Number of days below 0°C)
DD>5 (days)	degree-days above 5°C, (Number of days above 5°C)
DD>18 (days)	degree-days below 18°C, (Number of days below 18°C)
DD>18 (days)	degree-days above 18°C, (Number of days above 18°C)
NFFD	the number of frost-free days
FFP (days)	frost-free period
bFFP (Julian date)	the day of the year on which FFP begins
eFFP (Julian date)	the day of the year on which FFP ends
EMT (°C)	extreme minimum temperature over 30 years
EXT (°C)	extreme maximum temperature over 30 years
MAR (MJ m^{-2} d^{-1})	mean annual solar radiation
CMD (mm)	Hargreaves climatic moisture deficit; the amount of water by which potential evapotranspiration (PET) exceeds actual evapotranspiration (AET)

♦ For each of the plant traits you selected, identify at least three climate variables that you think are important drivers of differences we may observe in plant characteristics

♦List your Hypothesis for how each group of climate conditions relate to your selected plant traits below.

Be certain that your hypothesis meets the following criteria:

1. It relates the broad phenomenon of our research question in terms of a cause and effect

2. It is directional, stating a specific response for a specific stimulus (i.e. greater, larger, smaller, earlier, later, etc.)
3. It offers a justification for why you believe it is plausible (based on previous knowledge or understanding of the phenomena at hand).
4. It is supported by a statistical null and alternative hypothesis

It is okay to list related or additive environmental conditions together such as the example below does:

Example:

Hypothesis: plants originating from areas with higher average annual temperatures, lower summer precipitation, and more continental climates should have smaller leaf surface areas and greater leaf length/width ratios.

Justification:

a) Plants lose water through leaf stomata via evapotranspiration. Leaves with smaller surface areas absorb less sunlight and heat up less that larger leaves, thus reducing leaf water losses.
b) Plants with greater leaf length/width ratios will have smaller boundary layers and subsequently lower leaf temperatures, thus reducing leaf water losses

Statistical null and alternative hypothesis:

a) Null= there is no significant relationship ($p>0.05$) among average annual temperature and leaf surface areas of milkweeds
b) Null= there is no significant relationship ($p>0.05$) among average annual temperature and leaf length/width ratios of milkweeds

♦ **Produce descriptive statistics (in excel) for each dependent and independent variable you plant to test,**

♦ **Evaluate the descriptive statistics of the dependent and independent variable you analyzed**

a. **Which dependent variables display the greatest variability? The least? How do you know?**
b. **Which independent variables do you think will be the most informative in predicting outcomes of dependent variables? Support your argument with data from the descriptive statistics.**

Preparation for work in R:

1. Create a folder on your computer or removable storage that will house all data analysis files
2. Download and save:
    a. the milkweed morphology data set (Ecology_Morph_Climate_Corr.xls)
    b. The R studio analysis script (Ecology_Morph_Climate_Corr - R file

Data file preparation:

1. Begin by creating a single folder (on desktop) that will store all input and output files.B
2. Add the following files to the folder you created:
    a. Ecology_Morph_Climate_Corr.xls
    b. Ecology_Morph_Climate_Corr - R file
3. open Ecology_Morph_Climate_Corr.xls *excel file)
4. Delete all climate variables that you are not interested (leave only the variables you want to test)
5. IMPORTANT! Keep:
    a. Column A (Names) and Row 1 (column headers)
    b. Columns AG and AH (these are the principal components dimensions that
6. When you've deleted what's not needed
    a. Select save as

    i. Select the folder you created above
    ii. Under <u>Save as type</u> **select CSV (comma separated values)**

<u>Data analysis in R</u>

Orientation and important notes:

- Hash tags (#) precede notes
- to run a command, place the cursor on that line and press CTRL+ENTER.
- You may also highlight entire strings and press CTRL+ENTER.

<u>Data analysis procedure</u>

1. # Go to Session tab--> Set working Directory --> to source file location
2. **Follow instructions in notations within the R code**. Most of this process is loading the necessary packages and reading in/formatting the data set.
    a. For each command, place cursor in the command line and press CTRL+ENTER.

♦ Produce frequency distribution histograms (in R) for each dependent and independent variable you plant to test (lines 56-66 in the R code)

c. Export these histograms as .pdfs (you'll need to give them file names) then copy paste images into the space below, or attach images as .pdfs to the final exercise.

d. Answer prompts in the worksheet

♦ In R, produce a correlation matrix and p-value matrix among the dependent variables and independent variables you are testing.

e. Highlight all significant p-values among all variables (second table R produces)

f. Highlight all corresponding R-values (correlations) among all variables (first table R produces)

g. Answer prompts in the worksheet

2. Open the modified data file (with just your variables) that you created in the data file preparation step 4.

♦ conduct regression tests that evaluate the strength of the **two most significant relationships** you identified among the **each of the sets of independent and dependent variables** that you used to test your hypotheses.

    a. Put each set of tests on a separate tab in the excel worksheet.
    b. Produce a complete figure that shows the relationship
        i. A trend line, r^2 value and equation of the relationship.

3. Assess each set of original hypotheses you constructed **based on the Pearson significance (p) of relationships among the dependent and independent variables**.
    a. In your assessment, be certain to:
        i. Assess the validity of your null and alternative hypothesis
            1. Let the significance be the primary determinant of your assessment and use the correlation to suggest the strength of the relationship/correlation.
        ii. Support your statements with evidence
    b. For each set of results, use critical thinking to:
            1. Offer potential explanations as to why you got the results that you did (whether results were expected or unexpected).
            2. Consider biological/ecological phenomena that might explain the results.
        ii. Provide potential suggestions to improve the analyses or methods to make stronger inferences.