**Activity 4: Evaluating Essential Questions in Landscape Genetics**

**Learning Outcomes:**

1. Evaluate genetic diversity using the genetic PCA
2. Evaluate evidence of Evolutionary Significant Units on PCA
3. Evaluate evidence of barriers to gene flow using the genetic PCA

**Background:** Introduction to Conservation Management

**Outline:** 1. A guide to your genetic PCA ………………………….. 1

2. Explore your PCA ……………………………….......... 3

3. Evaluate your hypothesis ……………………………… 4

4. Future study and conservation management…………… 5

**1. A guide to your genetic PCA:**

Congratulations! You have a PCA representing the population genetics of a California vertebrate! Now, what does it mean?

*Interpreting colors and shapes:*

First, let’s dissect the color and shape scheme used in the PCA, to see what it tells us about the locality of a sample. Individuals in the PCA plot are colored based on latitude and longitude, and the shape of the point corresponds to the California ecoregion in which the sample was originally located (Fig. 1).

For example, specimens originally collected in the Klamath Mountains, in the northwestern corner of California, near the border with Oregon, will be colored purple/lavender and be represented as a small, filled diamond. A specimen collected from the Mojave or Sonoran Deserts in the southeastern part of CA, on the border with Arizona and Nevada, will be colored a green-blue and be represented as an asterisk. If you return to Activity 1, Fig. 2, you will notice two things. First, here we use 12 ecoregions instead of 19; this reduction is just to simplify the interpretation of your data. Second, you’ll notice that some ecoregions have a large latitudinal range, like the Sierra Nevada, or a large longitudinal range, like the Mojave Desert. Because of this, you may observe large color differences in individuals in your PCA with the same symbol; the color differences indicate the corner of the ecoregion where the individual was collected. To help with your interpretation of CA geography on the PCA, you should have your sampling map available from Activity 2.

*Axes:*

Your PCA axes are labeled PC1 and PC2. PC1 is principal component 1; this is the variable that explains the largest amount of variation in your original sequence dataset. How much variation? On the x-axis, “PC1” is followed by a percentage. A plot with an x-axis label reading “PC1 80%” means that 80% of all the variation in the genetic data can be explained by PC1. Alternatively, PC1 may only describe a small amount of the variation in the genetic data; if the plot reads “PC1 15%”, PC1 only describes 15% of the genetic variation. Similarly, the percentage of variation accounted for by PC2 is labeled on the y-axis. In general, if the amount of variation explained by a principal component is high, this could indicate large differences between populations, but the study design can also influence it, so take this with a grain of salt.

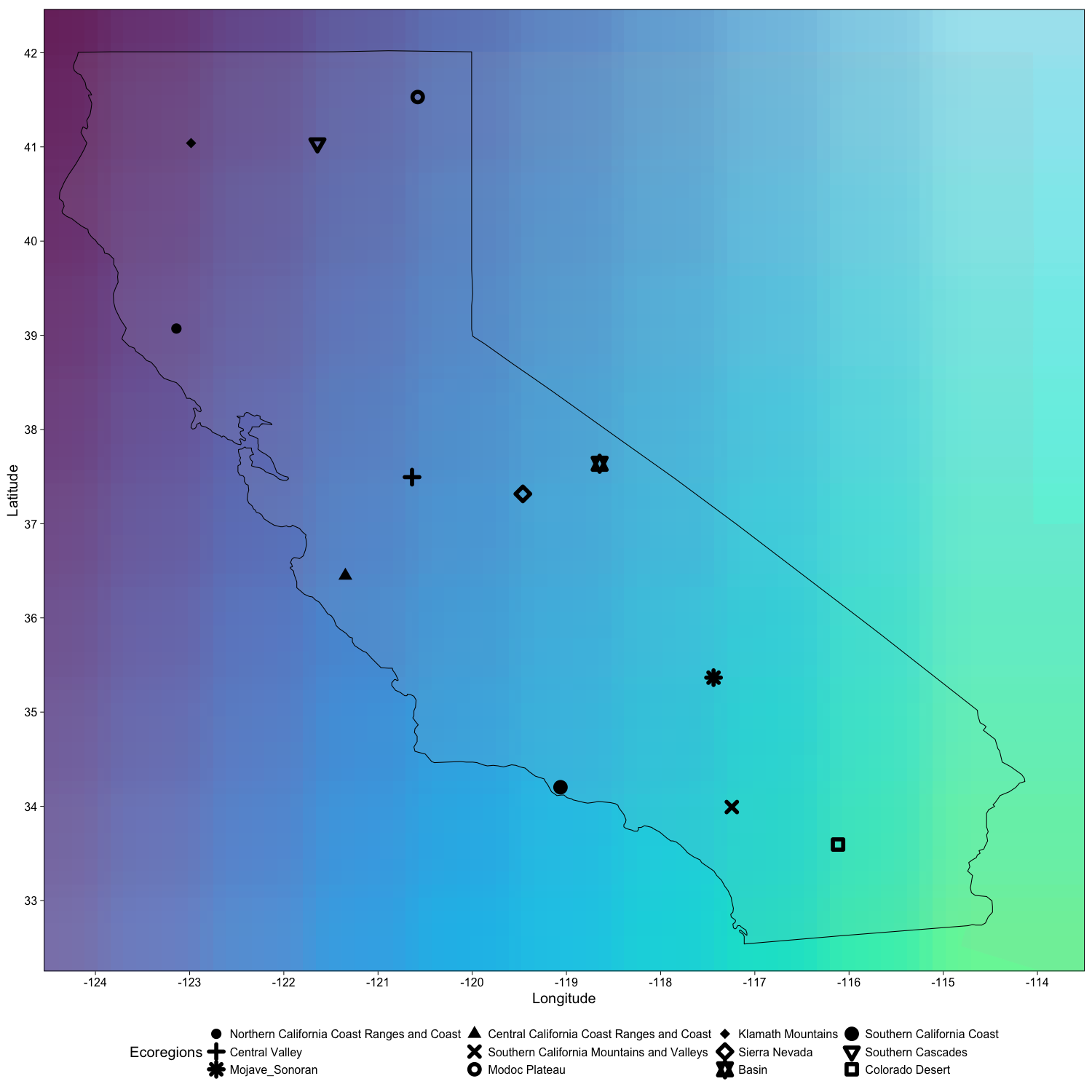


Fig. 1. Key to color and shape scheme used in the BCEEnet genetic PCA function. The color of shapes on the PCA correspond to their latitude and longitude as seen here. The shape of the point in the PCA corresponds to one of 12 ecoregions. In this figure, the shape on the map represents the average latitude and longitude of the ecoregion; the shape legend is at the bottom of the map.

*Genetic diversity:*

Recall that genetic diversity is a measure of the magnitude of genetic variation in a population or species; there are numerous ways to quantify genetic diversity. We won’t do a formal calculation of genetic diversity. However, the PCA illustrates the relative genetic diversity of individual ecoregions with respect to each other and the entire species or species complex; with adequate sampling, these ecoregions can be compared with one another to identify places with high or low genetic diversity. For instance, look at the distribution of individuals from the Sierra Nevada (open diamonds) on your PCA. If individuals from the Sierras are spread widely across the PCA, then the Sierras are a region of high genetic diversity. Alternatively, if the individuals from the Sierras form a tight cluster, the Sierras may contain low genetic diversity. Of course, Sierra samples may be somewhere between these two extremes and be of moderate diversity. You can also use your sample map from Activity 2 and the colors in the PCA to help your interpretation of genetic diversity. Did you sample widely across the Sierras or only in a few nearby locations? If the former, then an observation of high diversity in the Sierras could be the result of numerous separate populations being lumped into one umbrella; on the PCA, you would observe Sierra points with similar colors clumping together. If instead you see high genetic diversity in the Sierras and you only sampled a few localities, then those locations harbor a lot of genetic variation.

*Finding Evolutionary Significant Units (ESUs):*

An evolutionary significant unit or ESU represents a highly distinct population genetically and ecologically. With genetic data alone, you can identify populations that appear genetically unique for further study. You could follow up this discovery with field studies and an examination of morphology in museum specimens. A distinct population or set of populations will occupy their own space, to the exclusion of others, on the PCA plot. Put another way, members of an ESU will be each other’s closest relatives.

*Patterns of landscape connectivity:*

When migration is high among populations, individuals sampled far apart from one another may share genetic variants and thus appear near each other in a genetic PCA. When a barrier impedes migration, populations on either side of the barrier over time will look less and less similar genetically; as a result, in a PCA, they will occupy different parts of the PCA. To test the importance of a physical barrier to migration, you can examine the distribution of individuals from one side of the barrier on the PCA, and compare it to the individuals collected on the other side of the barrier. If you see a natural separation between the regions on the PCA, this could indicate a barrier to gene flow.

**2. Explore the genetic PCA:**

**Q4.1)** Is there genetic variation in your dataset? Give a qualitative description of your PCA. Are there clumps of individuals and if so, how many? How much variation is explained on each axis?Are the colors of individuals structured across the PCA? Are the shapes? [Hint: consider one shape at a time]

**Q4.2)** Do individuals from one ecoregion share similar genetic variation? Give a specific example from your data to illustrate your answer. [Hint: use the shape!]

**Q4.3)** Are individuals from any ecoregion widespread across the PCA? Give a specific example from your data to illustrate your answer. If there IS an ecoregion with high genetic diversity, do you think this reflects wide sampling and some separate populations, or is there high diversity within a single population? Give the reasoning behind your answer, referring to color scheme and your sample map as necessary.

**Q4.4)** If someone asked you to split the individuals on the PCA into two groups, how would you do it? Draw a picture to explain if necessary. Are individuals from the two groups from nearby regions of California? What about three groups? Four?

**Q4.5)** Following the logic of Q4.4,how many natural genetic clusters do you think are represented in your dataset? Does your answer make sense when you consider geography; that is, do clumps contain adjacent regions of CA or not? Justify your answer.

**3. Evaluate essential questions and your hypothesis**

**Q4.6)** Where in California is genetic diversity highest for your species? Justify your answer.

**Q4.7)** Is there evidence of one or more ESU in your species? Consider your sampling in your answer.

**Q4.8)** Is there evidence that any barrier is disrupting migration in your species? Justify your answer.

**Q4.9)** What organismal traits of your species could be impacting your results?

**Q4.10)** Do you accept or reject your hypothesis? Justify your answer.

**4. Future study and conservation management**

Today, you used existing data from a single mitochondrial locus sequenced from vouchered specimens to evaluate the landscape genetics of a vertebrate across a region of the topographically complex state of California.

**Q4.11)** What would you like to know next about this biological system?

**Q4.12)** How you would go about answering your question to Q4.11? You may find the “Recent Advances in Landscape Genetics” section of Part 1:Background helpful, but there is no need to restrict yourself to those methods!

**Q4.13)** What additional evidence would you like to inform your any conservation management recommendation?

**Q4.14)** Do you want to make a management recommendation with respect to your species or species complex? For example, do you think certain genetic clusters should be managed separately from one another?