

Experimental Design

Project 2

Due Tuesday, May 24

The structure of this homework assignment is very similar to that of the first one. You will analyze a set of data and write it up in the form of the Results section of a journal article (with the appropriate Statistical Methods section). Please see me if you have questions about the last assignment.

Problem.—Using a mapping approach, you think that you have identified a novel gene that explains differences in developmental rate among populations of tree frogs (*Hyla regilla*) that live at high and low altitudes. Crosses between frogs are pretty time consuming and you would like to be able to see if the genetic variation you have discovered corresponds to any functional differences among the frog populations, so you decide to conduct a new experiment in which you look at variation in gene expression of this new gene (which we will call *altitudin*). Since one of the major correlates of the change in altitude is a change in temperature, you decide to take a somewhat physiological approach and compare levels of gene expression of lowland and highland tadpoles raised at different temperatures. You sample frogs from four different highland areas and another four different lowland areas. You hope that by sampling the areas within each altitude group you will be able to get some idea of the variation among local frog populations.

You bring eggs collected from these eight different areas back into the lab and use real-time PCR to measure the amount of *altitudin* RNA per gram of tissue collected from metamorphosing tadpoles raised at three different temperatures: 18°, 24°, and 30°. The gene expression for each individual is standardized relative to the level of *actin* RNA within each sample. You measure the responses of 6 tadpoles from each area at each temperature level (no tadpole is used in more than one temperature treatment), yielding 144 total individuals in the experiment. Note that by functional genomic standards, this is a large experiment.

Your job is to analyze the data and tell me which factors influence gene expression in these frogs. The analysis should contain the following (at least):

- Statistics and graphs you deem necessary to describe the data
- The ANOVA model (written out) used to analyze the data (be sure to specify which effects are random and which are fixed). (Note: this would usually be in the *Methods* section.)
- Tests of the ANOVA assumptions.
- Results presented in an ANOVA table, including a symbolic representation of the expected mean squares.

The data are available on the web site. The first column gives the altitude group in which the tadpole was collected (hi/low); the second, the area within that region in which the tadpole was collected (1-4); the third, the temperature at which the experiment was performed (18/24/30); and the fourth, the relative level of gene expression for *altitudin*.