**Instructor Guide: Amylase Copy Number and Diet, modified HHMI activity for Higher Ed**

Written by:

Ruth Buskirk, Ph.D., University of Texas at Austin

Kristin Harvey, Ph.D., University of Texas at Austin

Rebecca Orr, Ph. D., Collin College

**Relevance and goals for the module**

General biology students learn about hydrolysis reactions, role of starch, enzymes, and genes when studying macromolecules. They do not, however, often make connections between the enzymes responsible for hydrolysis and the genes responsible for the production of hydrolytic enzymes. Furthermore, students tend not to appreciate the impact of specific environments on the presence and/or copy number of specific nucleic acids and proteins. This module seeks help students to make these connections while addressing two distinct learning goals, one that is content-based and one that is develops quantitative competency in freshman biology students.

*Learning Goals for "Amylase Copy Number and Diet;*

1. To understand amylase enzyme function and the role the amylase gene copy number might play in conferring a selective advantage for an organism in high-starch and low-starch environments.
2. To develop students' science process skills by teaching students how to utilize descriptive statistics such as mean, median, and standard deviation and how to evaluate histograms.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Content Learning Objectives*** | ***Quantitative Learning Objectives*** | ***Process of Science Learning Objectives*** | ***Formative Assessment Ideas*** | ***Summative Assessment Ideas*** |
| 1. Explain the function of salivary amylase in the hydrolysis of starch. 2. Describe how genetic variation can provide a selective advantage to an organism in a particular environment. | 1) Create and/or analyze numerical summaries and graphical displays of data.  For two groups on one continuous variable:  2) Explain the meaning of key terms used in descriptive statistics such as: sample sizes, means, standard deviations, medians, minimums, and maximums.  3) Analyze a histogram for two groups on a single continuous variable to identify critical components such as graph type, axis labels, and key to symbols.  4) Compare and contrast measures of center and spread based on the shape of the histograms to compare and contrast high vs. low starch diet on the variable of interest, amylase gene copy number. | 1. Propose testable hypotheses to investigate correlation between diet and amylase gene copy number. | 1) Use the embedded questions in the PowerPoint presentation.  2) Have students/teams present answers to questions in each handout, and share/discuss their reflections. | 1) Multiple- choice questions are provided.  2) Present a histogram on an exam and ask student to make conclusions from the data presented. |

**Background for instructor**

Starch, a plant polysaccharide composed of many building blocks of glucose, is a high-energy component of foods found in nature. As early humans transitioned from

hunting and gathering to more agrarian lifestyles, their diets changed to include more high starch foods. Some cultures incorporated more starch into their diets than others, and those cultural differences in human populations are still present in some cultures today.

The gene that encodes salivary amylase (AMY1) is somewhat unique, as most humans have more than one diploid copy of the gene; in fact the number of copies ranges from 2 to 15. The authors of the paper from which this data was taken investigated whether there is a correlation between the number of AMY1 gene copies and the type of diet (high-starch or low-starch) of a population. A correlation would indicate that having more copies of the AMY1 gene provides a selective advantage, allowing individuals to break down starch more efficiently.

In this activity, students analyze amylase gene copy number data from populations that have a high-starch diet and from populations that have a low-starch diet. Each data set provided in this module is a subset taken from the original sample in order to make the data more manageable for students. A random number generator in EXCEL was used to select the subsamples from the larger data set that was available in the supplement of the Perry et al. (2007) paper. Data is provided in accompanying Excel files. Students will need devices with Excel or Excel compatible programs loaded to view the data files.

The data provided can be used in two ways:

1. Ask students to use the Amylase Student Handout (Analysis) handout and the Salivary\_Amylase\_DataFilled Excel file where the following descriptive statistics have been calculated for the student:

* + n (sample size)
  + Mean
  + Median
  + Minimum
  + Maximum
  + Variance and standard deviation

In this handout, students are asked to evaluate the mean of each data set to form a hypothesis. They explore the difference between mean and median, and use calculations of standard deviation for each population to evaluate the variability in each data set. Students are then shown how a histogram can be generated in Excel, are asked to analyze the histogram, and make conclusions the possible effect of diet on amylase gene copy number.

2. Ask students to use the Amylase Student Handout (Generate) handout and the Salivary\_Amylase\_DataBlank Excel file to calculate the following descriptive statistics in Excel:

* + n (sample size)
  + Mean
  + Median
  + Minimum
  + Maximum
  + Variance and standard deviation: Note that space has been allocated for students to calculate standard deviation using the manual method (see manual method tab in Spreadsheet tutorial 5) to ensure students understand the steps required to calculate standard deviation. Students can confirm these calculations by using Excel functions to calculate variance and standard deviation (see Spreadsheet Tutorial 2).

If students are using the Salivary\_Amylase\_DataBlank Excel file, the following HHMI Spreadsheet Tutorials are recommended to teach students how to use Excel to make these calculations:

[Spreadsheet Tutorial 1 -- Formulae, Functions, and Averages](http://www.hhmi.org/biointeractive/spreadsheet-tutorial-1-formulae-functions-and-averages)

[Spreadsheet Tutorial 2 -- Autofill Data, Cell References, and Standard Deviation](http://www.hhmi.org/biointeractive/spreadsheet-tutorial-2-autofill-data-cell-references-and-standard-deviation)

[Spreadsheet Tutorial 5: Histogram](http://www.hhmi.org/biointeractive/spreadsheet-tutorial-5-histogram)

In addition to calculating these descriptive statistics, students are asked to look at the mean of each data set to form a hypothesis. They explore the difference between mean and median, and use calculations of standard deviation for each population to evaluate the variability in each data set. In this exercise students will generate a histogram in Excel and are then asked to analyze their histogram and make conclusions the possible effect of diet on amylase gene copy number.

**Prerequisite knowledge for students**

Before studying this topic, students should already be able to:

|  |
| --- |
| 1. Describe the difference between carbohydrates, proteins, and nucleic acids. 2. Define selective advantage and give an example of how the environment in which an organism lives can lead to natural selection. 3. Explain the role of genes in the production of protein. |

**Scenarios for use of this activity include:**

1. Utilize the PowerPoint presentation of the case study to facilitate class discussion and students' completion of the Amylase Student Handout (Analysis). Ask students to share their answers to generate discussion about generating a testable hypothesis and their evaluation of the data.
2. Have students complete the Amylase Student Handout (Generate) as homework. Use class time to allow students to work in small groups to discuss their answers. Ask students to share their answers to generate discussion about generating a testable hypothesis and their evaluation of the data.

**Sample assessment questions**

1. A histogram would be most useful in showing which of the following types of data?

1. A series of data points recorded over time that measures a change or response to a specific stimuli
2. *A set of data that measures the frequency in which an event or condition occurs*
3. A series of data that measures the response of one variable to another
4. Data measuring changes over time for more than one group

2. Which of the following conditions has been associated with an increase in the number of gene copies for salivary amylase?

1. Populations that have a high-protein diet.
2. Populations that consume a high amount of glycogen.
3. *Populations that have a high-starch diet.*
4. Populations that have a low-protein, low-starch diet.

3. When analyzing data, it can be helpful to understand how much how spread out the data is from the central tendency. Which of the following descriptive statistics would be most helpful in measuring this variability?

1. mean
2. median
3. standard deviation
4. "n"

**References**

1. Amylase Copy Number and Diet, http://www.hhmi.org/biointeractive/amylase-copy-number-and-diet
2. Diet and the Evolution of Salivary Amylase, http://www.hhmi.org/biointeractive/diet-and-evolution-salivary-amylase
3. Perry, G. H. et al. 2007. Diet and the evolution of human amylase gene copy number variation. Nature Genetics39:1256–1260.
4. Spreadsheet Data Analysis Tutorials, http://www.hhmi.org/biointeractive/spreadsheet-data-analysis-tutorials
5. Teacher Guide, Math and Statistics, http://www.hhmi.org/biointeractive/teacher-guide-math-and-statistics