CBE—Life Sciences Education
Vol. 11, 142–151, Summer 2012

Article

Are Africans, Europeans, and Asians Different "Races"? A Guided-Inquiry Lab for Introducing Undergraduate Students to Genetic Diversity and Preparing Them to Study Natural Selection

Steven T. Kalinowski,* Tessa M. Andrews,* Mary J. Leonard,† and Meagan Snodgrass‡

*Department of Ecology, †Department of Education, and ‡Department of Cell Biology and Neuroscience, Montana State University, Bozeman, MT 59717

Submitted September 22, 2011; Revised March 23, 2012; Accepted March 23, 2012
Monitoring Editor: Diane K. O'Dowd

Many students do not recognize that individual organisms within populations vary, and this may make it difficult for them to recognize the essential role variation plays in natural selection. Also, many students have weak scientific reasoning skills, and this makes it difficult for them to recognize misconceptions they might have. This paper describes a 2-h laboratory for college students that introduces them to genetic diversity and gives them practice using hypothetico-deductive reasoning. In brief, the lab presents students with DNA sequences from Africans, Europeans, and Asians, and asks students to determine whether people from each continent qualify as distinct "races." Comparison of the DNA sequences shows that people on each continent are not more similar to one another than to people on other continents, and therefore do not qualify as distinct races. Ninety-four percent of our students reported that the laboratory was interesting, and 79% reported that it was a valuable learning experience. We developed and used a survey to measure the extent to which students recognized variation and its significance within populations and showed that the lab increased student awareness of variation. We also showed that the lab improved the ability of students to construct hypothetico-deductive arguments.

INTRODUCTION

Evolution is the unifying theory of biology (Dobzhansky, 1973) but is poorly understood by many biology students (e.g., Gregory 2009). It is tempting to attribute this misunderstanding to religious beliefs, but a growing body of research shows that personal beliefs have very little effect upon how well students learn evolution (Bishop and Anderson, 1990; Demastes *et al.*, 1995; Ingram and Nelson, 2006). Several facets of evolutionary theory appear intrinsically difficult for students to grasp (e.g., Brumby, 1984; Nehm and Reilly, 2007; Catley and Novick, 2009). For example, natural selection, the principle cause of evolution, is frequently described as a simple idea—even a "staggeringly" simple idea (Coyne, 2009, p. xvi)—but can be remarkably difficult for students to understand. The reasons for this are diverse; three include: typological thinking (Shtulman and Schulz, 2008), misconceptions regarding genetic inheritance (Gregory, 2009), and poorly developed scientific thinking skills (Lawson, 2002).

A potential obstacle for students studying natural selection is a lack of appreciation for the extent and significance of variation within populations. As Darwin documented in the first two chapters of *On the Origin of Species* (Darwin, 1859), individuals of all species exhibit within-species variation of almost all traits. Without such variation, evolution by natural selection would be impossible. Despite the

DOI: 10.1187/cbe.11-09-0087
Address correspondence to: Steven T. Kalinowski (skalinowski@montana.edu).

importance and near ubiquity of such variation, many students have a tendency to dismiss variation within species as unimportant and focus their attention on traits that individuals in a species share (for a review, see Solomon and Zaitchik, 2012). We shall call this tendency "typological thinking." Other authors use the term "essentialism," and if there is clear distinction between the two, it is not commonly made in the evolution literature. Typological thinking arises in childhood and is useful for predicting the characteristics of individuals from their species identity but is an obstacle for understanding natural selection (Shtulman, 2006; Shtulman and Schulz, 2008). The antidote for typological thinking is population thinking, which Mayr (2004) defined as "the realization that in biological populations. . . every individual is unique." Educators therefore frequently stress that evolution instruction should emphasize the importance of individual variation (e.g., Greene, 1990; Gregory, 2009; Shtulman and Schulz, 2008)—just as Darwin did in *On the Origin of Species*.

Students often have misconceptions that impede learning. One of the most common misconceptions held by introductory biology students is that evolution proceeds through the transformation of individuals. Many students believe, for example, that individuals evolve because they need to, because they use or do not use specific body parts, or because the environment directly induces changes (e.g., Brumby, 1979; Bishop and Anderson, 1990; Jensen and Finley, 1996; Nehm and Schonfeld, 2008). These misconceptions reflect a lack of appreciation for the role of DNA in inheritance. To correct these misconceptions, instructors need to help students understand how DNA sequences affect the ability of organisms to survive and reproduce, and how variation in these abilities changes the frequency of DNA sequences in populations (Smith *et al.*, 2009; Kalinowski *et al.*, 2010). This may require some initiative on the part of instructors, because most popular biology textbooks do not contain examples of DNA sequences (Kalinowski *et al.*, 2010).

Misconceptions regarding inheritance and a lack of attention to variation within populations are two potential impediments for students studying natural selection. Poorly developed scientific thinking skills present a third challenge to many students. Evolution is difficult to observe—both because macroevolutionary changes are slow and because microevolutionary genetic changes are invisible to the naked eye—and this makes the entire process difficult for students to understand (Lawson, 2002). Hypothetico-deductive reasoning, one of the most common forms of scientific thinking, is useful for making sense of such processes (e.g., Platt, 1964; Feynman, 1984; Lawson, 2003). In brief, hypothetico-deductive reasoning tests hypotheses by determining what the hypothesis predicts should be observed in a test (Figure 1), and then performing the test to see whether the predicted result is observed. This method for testing hypotheses is often called "the scientific method," but is, of course, only one way in which scientists answer questions. Empirical evidence suggests that improving students' hypothetico-deductive reasoning improves student understanding of evolution (Lavoie, 1999) and decreases student misconceptions regarding natural selection (Lawson and Thompson, 1988). Not surprisingly, recommendations for designing evolution curricula have frequently stressed improving students' scientific thinking skills in general (e.g., Nelson, 2008) and hypothetico-deductive reasoning in particular (e.g., National Research Council, 1998).

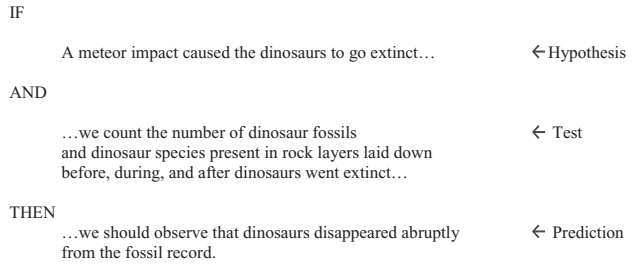

Figure 1. An example of how hypothetico-deductive reasoning can be used to test the hypothesis that a meteor impact caused the dinosaurs to go extinct. The elements of hypothetico-deductive reasoning are labeled on the right.

To understand natural selection, students need to overcome typological thinking, genetic misconceptions, and poor scientific thinking skills. These are relatively long-term goals that are not likely to be achieved in a single lecture, lab, or homework assignment. Most students will need multiple lessons to achieve these learning outcomes. Students learn best when they actively participate in their learning (e.g., Hake, 1998), so these lessons will be most effective if they actively engage students. Unfortunately, there is a shortage of such learning activities available from which instructors can select. In this paper, we describe a 2-h laboratory exercise designed to prepare students to learn about natural selection by introducing them to genetic diversity in human populations. The specific learning objectives of the lab were to:

- Reduce the extent to which students used typological thinking.
- Improve students' ability to use hypothetico-deductive reasoning.
- Engage students in a laboratory exercise they enjoyed and believed was useful.

As the learning objectives reveal, this lab was designed to help *prepare* students to learn natural selection but not to actually teach students how natural selection works. Instructors will need to use this lab in conjunction with other learning activities (e.g., Kalinowski *et al.*, 2006a; Andrews *et al.*, 2011) to teach natural selection.

METHODS

Participants and Context

We used this lab as the first lab of the semester in an introductory biology course on evolution and ecology for biology majors. We collected data on the lab during two semesters. In 2010, 59 students were enrolled in the course: 3% were freshman, 53% sophomores, 34% juniors, and 9% seniors. In 2011, 41 students were enrolled in the course: 85% were freshman, 12% sophomores, and 3% seniors. Virtually all reported (by raised hands) that they were preparing for careers in medicine.

The class met three times a week for a 50-min lecture and once per week for a 3-h lab. Lectures made extensive use of active learning (e.g., Andrews *et al.*, 2011). Labs were guided-inquiry exercises (Herron, 1971; Douglas, 2002) that students

worked on in groups of three or four. At the beginning of each lab, students were given a question to answer and the materials they needed to answer the question, and they were then set loose to answer the question on their own, with assistance (as needed) from the instructor. (See Kalinowski *et al.*, 2006a,b, for descriptions of the second and third labs used in this course.)

Lab Description

The human race lab requires approximately 2 h and, ideally, a computer for every group of three or four students in the course. The lab was divided into four parts: introduction, project design, data analysis, and presentation of results. The introduction to the lab consisted of a 10-min lecture that introduced students to the focal question of the lab and the resources students had available to answer it. We used the following lecture, accompanied by PowerPoint slides, to introduce the lab:

> AIDS is a disease of the immune system caused by the human immunodeficiency virus (HIV). Almost forty million people from all around the world are currently infected with HIV, and each year over two million people die from AIDS. This includes hundreds of thousands of children. Antiviral drugs have been developed that increase the life expectancy of patients with AIDS, but there currently is no cure.
>
> Recent research has shown some people become infected by HIV very quickly after being exposed to HIV but that others do not become infected despite repeated exposure. In addition, some people infected with HIV rapidly develop symptoms of AIDS, while others do not. Not all of these differences are understood, but it is clear that some of this variability is caused by genetic differences among people. For example, people who have a 32–base pair deletion in their *CCR5* gene are highly resistant to HIV infection.
>
> In the future, doctors may tailor their treatments to the specific genetic characteristics of their patients. At this point, however, the genetic basis of most diseases, or the way genes affect treatment, is poorly understood. Therefore, it is usually impossible to use the specific genetic characteristics of patients to select effective treatments.
>
> Some medical researchers have proposed that until we are able to use the specific genetic characteristics of individual patients to select treatments, a patient's race may be useful for selecting treatments. The logic behind this suggestion is that there may be predictable genetic differences between different races of people, and that people of the same race may respond to treatments in a similar manner.
>
> For the purpose of this investigation, let us assume that you are working for a pharmaceutical company developing new drugs to treat AIDS. Your specific assignment is to help design drug trials. The company you work for is concerned that the drugs it develops will work differently in different races. You have been assigned to a research team to study genetic differences among humans from different parts of the world.
>
> The entire concept of race is immensely controversial. This includes scientific controversy regarding what race is and political and social controversy regarding how people should be treated. Racial theories have been used to justify discrimination, exploitation, and genocide. Because of this cultural and historical context, we need to proceed with great caution and respect.
>
> Many different racial classifications have been proposed, and there is no consensus regarding which one, if any, is most appropriate. One of the most commonly proposed classifications is that peoples on different continents belong to different races. For example, it has frequently been proposed that Africans, Europeans, and Asians are three different races. The specific goal of your laboratory investigation is to evaluate this hypothesis. By the end of today's lab period, you will be expected to answer the question "Are Africans, Europeans, and Asians different races?"
>
> To answer this question, you will analyze genetic data from indigenous peoples sampled from Africa, Europe, and Asia [Table 1; Figure 2]. Specifically, eight individuals were sampled from each continent. The mitochondrial genome of each individual was sequenced. This includes over 15,000 nucleotides of DNA sequence. You may have forgotten that mitochondria have DNA and may be wondering why mitochondrial DNA was selected for this study. The answer is that mitochondrial DNA has a high mutation rate and no recombination, but this is not something you need to worry about. For our purpose, mitochondrial DNA sequences can be considered a representative sample of the genome of each individual. For example, if two individuals have genetically similar mitochondrial DNA sequences, we will assume that their entire genomes are similar.
>
> As I just mentioned, you will be analyzing DNA sequences that are 15,000 nucleotides long. Long DNA sequences are difficult to analyze without the assistance of specialized software. Therefore, you will have the computer program Sequence Viewer for Students to assist you.

At this point, the instructor showed the class the computer program Sequence Viewer for Students (available at www.montana.edu/kalinowski/Software/SequenceViewer.htm) and explained how it could be used to view and compare DNA sequences (available in the Supplemental Material). This completed the introduction phase of the lab.

The class was then instructed to open up the DNA sequences in the program, take a look at them, and develop a strategy for answering the focal question of the lab. This initiated the project design phase of the lab. The goal of this phase of the lab was to figure out *how* to answer the focal question of the lab. Many students had very little idea how to get started, either because they do not recognize that the question required hypothetico-deductive reasoning or because they do not know how to apply hypothetico-deductive reasoning. Therefore, the instructor sat down with each group, asked them how they were going to answer the question, and helped guide their thinking as necessary. The following dialogue was typical:

> INSTRUCTOR: How is it going? Have you figured out how to answer the question?
> STUDENT: No, we're lost. I don't get it.
> INSTRUCTOR: OK. Let's talk about this. I see that you are looking at the DNA sequences. Have you observed anything that might be useful for answering our question?
> STUDENT: Well, we can see there are differences between people, but I don't know if this means that these people are different races are not. I don't see any patterns.
> INSTRUCTOR: The purpose of this lab is to test a hypothesis. . . the hypothesis that Africans, Europeans, and Asians are different races. Do you remember from lecture how scientists test hypotheses?

Figure 2. Map of the ethnic groups sampled for this lab. Locations are approximate or representative.

STUDENT: Yeah, we need to make a prediction. My prediction is that Africans are a different race.
INSTRUCTOR: In this context, a prediction isn't what you personally think is true, but is something the hypothesis predicts will be observed. Therefore, you need to figure out what the hypothesis you are testing predicts you will observe in these DNA sequences. If Africans, Europeans, and Asians are different races, what should you observe in the DNA sequences that we have?
STUDENT: You mean, something like. . . Africans should be genetically similar to each other.
INSTRUCTOR: Yes, that is right. Now tell me, when you say that our hypothesis predicts that Africans should be genetically similar to each other, what to do you mean? Genetically similar compared to what?
STUDENT: If Africans are a distinct race, they should be genetically more similar to each other than to Europeans and Asians. The same would be true for people in Europe and Asia.
INSTRUCTOR: Exactly. Now take a look at the DNA sequences to see if you find this pattern or not.

Notice that testing the hypothesis that Africans, Europeans, and Asians are different races requires classic hypothetico-deductive reasoning (Figure 3). In our experience, the hardest part of this lab for students was realizing that hypothetico-deductive reasoning is useful and figuring out how to use it. Once students realized they could use hypothetico-deductive to test their hypothesis, most rapidly figured out how to use it. If the comments made by the instructor in the dialog above are not sufficient for getting students thinking along the right track, it can be useful to focus student attention on the DNA sequence of a single individual in the data and to ask, "What does the hypothesis predict about the DNA sequence of this individual?" Students should recognize that the chosen individual should be more similar to other individuals from the same continent than to individuals from other continents.

Although figuring out how to answer the focal question of the lab is the hardest part of this investigation, the data analysis also forces students to deal with some observations that appear ambiguous. The computer program students use to analyze DNA sequences can count the number of nucleotide differences between the individuals included in the analysis (Table 2). Some of the results of this simple analysis clearly

IF		
	Africans, Europeans, and Asians are different races…	← Hypothesis
AND		
	…we count the number of nucleotide differences between each pair of individuals, and compare the number of differences between people living on different continents…	← Test
THEN		
	…we should observed there are fewer nucleotide differences between people living on the same continent than between people living on different continents.	← Prediction

Figure 3. The hypothetico-deductive reasoning used in the lab described in this paper. The elements of hypothetico-deductive reasoning are labeled on the right.

Table 1. Summary information for mitochondrial DNA sequence used in the lab described in this paper[a]

Continent	Populations	Country	Latitude/longitude	GenBank
Africa	Biaka	Republic of Congo	0, 16 E	AF346968
	Ewondo	Cameroon	2 N, 11 E	AF346980
	Hausa	Nigeria	12 N, 8 E	AF346985
	Mandenka	Senegal	13 N, 13 W	AF346995
	Mbuti	Democratic Republic of Congo	2 N, 29 E	AF346998
	Mkamba	Kenya	1 S, 38 E	AF347000
	San	Namibia	21 S, 20 E	AF347008
	Yoruba	Nigeria	6 N,6 E	AF347014
Europe	Dutch	Netherlands	52 N, 5 E	AF346975
	English	United Kingdom	52 N, 2 W	AF346978
	French	France	46 N, 3 E	AF346981
	Georgian	Georgia	42 N, 44 E	AF346982
	German	Germany	51 N, 10 E	AF346983
	Italian	Italy	42 N, 13 E	AF346988
	Saami	Norway	70 N, 23 E	AF347006
	Tatar	Ukraine	48 N, 33 E	AF346974
Asia	Buriat	Russia (near Mongolian border)	52 N, 107 E	AF346970
	Chinese	China	33 N, 102E	AF346972
	Chuckchi	Russia (extreme northeast)	66 N, 172 E	AF346971
	Evenki	China (extreme northeast)	49 N, 120 E	AF346979
	Indian	India	22 N, 78 E	AF382013
	Inuit	Russia (Siberia)	71 N, 142 E	AF347010
	Japanese	Japan	36 N, 138 E	AF346989
	Uzbek	Uzbekistan	40 N, 65 E	AF347011

[a]The countries, latitudes, and longitudes listed for each population are representative.

Table 2. Number of mitochondrial DNA differences between individuals belonging to 24 populations arranged by continent

Population[a]	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1. Biaka	—	62	62	65	71	74	77	65	60	57	59	66	62	62	57	59	67	69	59	71	58	66	65	57
2. Ewondo	62	—	60	41	67	34	75	29	26	23	25	32	26	28	25	25	29	33	27	33	20	28	27	25
3. Hausa	62	60	—	63	14	72	45	61	60	57	59	66	60	62	59	59	63	67	59	67	52	64	63	59
4. Mandenka	65	41	63	—	70	51	74	46	43	38	42	47	43	45	42	42	46	48	44	50	37	45	44	42
5. Mbuti	71	67	14	70	—	79	54	66	67	64	66	73	67	69	66	66	70	74	66	74	59	71	70	66
6. Mkamba	74	34	72	51	79	—	87	41	38	35	37	44	38	40	37	37	41	47	39	45	32	40	39	37
7. San	77	75	45	74	54	87	—	76	75	72	74	81	75	75	74	74	78	80	74	82	69	79	78	74
8. Yoruba	65	29	61	46	66	41	76	—	31	28	30	37	31	33	30	30	34	38	32	38	25	33	32	30
9. Dutch	60	26	60	43	67	38	75	31	—	9	7	22	18	18	13	7	31	25	19	35	22	30	29	15
10. English	57	23	57	38	64	35	72	28	9	—	8	17	15	15	10	8	28	22	16	32	19	27	26	12
11. French	59	25	59	42	66	37	74	30	7	8	—	21	15	17	12	6	30	24	18	34	21	27	26	14
12. Georgian	66	32	66	47	73	44	81	37	22	17	21	—	18	24	21	21	37	31	25	41	28	36	35	21
13. German	62	26	60	43	67	38	75	31	18	15	15	18	—	20	17	15	31	27	21	35	22	28	27	17
14. Italian	62	28	62	45	69	40	75	33	18	15	17	24	20	—	17	17	33	27	21	37	24	32	31	17
15. Saami	57	25	59	42	66	37	74	30	13	10	12	21	17	17	—	12	30	24	18	34	21	29	28	14
16. Tatar	59	25	59	42	66	37	74	30	7	8	6	21	15	17	12	—	30	24	18	34	21	29	28	14
17. Buriat	67	29	63	46	70	41	78	34	31	28	30	37	31	33	30	30	—	40	32	8	19	27	26	30
18. Chinese	69	33	67	48	74	47	80	38	25	22	24	31	27	27	24	24	40	—	28	44	31	37	38	24
19. Chuckchi	59	27	59	44	66	39	74	32	19	16	18	25	21	21	18	18	32	28	—	36	21	31	30	18
20. Evenki	71	33	67	50	74	45	82	38	35	32	34	41	35	37	34	34	8	44	36	—	23	31	30	34
21. Indian	58	20	52	37	59	32	69	25	22	19	21	28	22	24	21	21	19	31	21	23	—	18	17	21
22. Inuit	66	28	64	45	71	40	79	33	30	27	27	36	28	32	29	29	27	37	31	31	18	—	15	29
23. Japanese	65	27	63	44	70	39	78	32	29	26	26	35	27	31	28	28	26	38	30	30	17	15	—	28
24. Uzbek	57	25	59	42	66	37	74	30	15	12	14	21	17	17	14	14	30	24	18	34	21	29	28	—

[a]Populations 1–8 are Africans, populations 9–16 are European, and populations 17–24 are Asian.

Table 3. Sixteen traits used in a multiple-choice survey to assess the extent to which students were thinking typologically[a]

Trait	Pretest average	Posttest average
Length of femur bone in male lions	2.31	2.44
Have a tail	0.86	0.92
Strength of jaw muscles in male lions	2.61	2.75
Size of the mane in male lions	2.42	2.50
The DNA sequence of *MRC1*, a gene that affects fur color in lions and other mammals	2.39	2.58*
Hunting skill of female lions	2.92	2.92
Have canine teeth	0.75	0.97*
Mating behavior in male lions	2.22	2.33
DNA sequence of *IGF*, a gene that affects the growth rate of mammals	2.58	2.81*
Length of tail	2.22	2.19
Hunt other animals for food	0.97	1.08
Number of cubs a female lioness has in a litter	2.58	2.75*
Volume of roar in male lions	2.53	2.61
Length of canine teeth	2.83	2.64
Testosterone levels in adult male lions	2.61	2.81*
Amino acid sequence of hemoglobin protein	1.54	1.80*

[a]Students were asked whether each trait varied and were given four possible choices: cannot vary (0), could vary but does not (1), varies but differences do not affect individuals (2), and varies and differences affect individuals (3). The pre- and posttest columns show the average response (on the 0 to 3 scale) before and after the lab described in this investigation. The traits are listed in the order given to students.

*$P < 0.05$ (one-tailed, paired t test).

differ from what the African/European/Asian race hypothesis predicts. For example, several Africans in the sample are genetically more similar to Europeans than to other Africans (Table 2). Likewise, some of the Asians are genetically more similar to Europeans than they are to other Asians. Neither of these observations would be predicted by the three-races hypothesis. However, some of the other results do match what the three-races hypothesis predicts. For example, Europeans are genetically more similar to one another than they are to Africans.

Many students will wonder what to do with results that they perceive to be ambiguous or contradictory. This provides a good opportunity to remind students that a wrong hypothesis can make some "correct" predictions. For example, the geocentric model of the solar system correctly predicts that the sun will rise in the east, travel across the sky during the day, and set in the west. However, the geocentric hypothesis does not correctly predict the path that planets follow through the night sky over the course of a year and, therefore, was rejected by Copernicus. What this shows is that it is not sufficient for some of the genetic data examined by students to match the prediction of the three-races hypothesis predicts. For a hypothesis to be supported, *all* data must agree with the predictions made by the hypothesis. Therefore, students should conclusively reject the hypothesis that Africans, Europeans, and Asians are different races.

The final part of the lab was data presentation, in which students reported to the full class their results and discussed the implications of their results for medical trials. Our students had no difficulty concluding that Africans, Europeans, and Asians do not appear to be distinct races. The primary implication of this is that the geographic origin of patients is unlikely to be a good predictor of how patients will respond to medical treatments. Some students noticed that there is much more genetic diversity within Africans than within Europeans or Asians and suggested that if a pharmaceutical company was conducting drug trials, extra effort might be made to sample this diversity.

Assessment

We performed three types of assessment to measure the lab's outcomes. First, we used a survey to assess students' views and reactions to the lab. Second, we developed a multiple-choice survey to quantify the degree to which students were typological thinkers and used this to look for changes after the class had completed the lab. Finally, we created and used two short-answer questions to measure how well students could construct a hypothetico-deductive argument.

Students learn more from tasks they find interesting (Renniger *et al.*, 1992), useful, and important (Zusho *et al.*, 2003). Therefore, we gave students an anonymous survey to get their views about the lab after they completed it. Students were asked to indicate their level of agreement with a series of statements, using a five-point Likert scale (strongly agree, agree, neither agree nor disagree, disagree, strongly disagree). In 2010, 47 students out of 59 students completed the survey, and we calculated the proportion of responses in each category.

We used a multiple-choice survey (see Table 3) immediately before and immediately after the lab in 2011 to measure the degree to which students recognized variation within populations and the significance that this variation has to survival and reproduction. The survey listed 16 morphological, behavioral, and genetic traits that could potentially vary in lions, and students were asked to indicate whether each of the traits: cannot vary, could vary but does not, varies but differences do not affect individuals, or varies and differences affect individuals. We scored the extent to which students perceived variation in populations on a scale of 0 to 3, with 0 being assigned to the "cannot vary" response and 3 being assigned to the "varies and differences affect individuals," etc.

We analyzed student responses in a few ways. In our primary analysis, we averaged each student's response (on the 0 to 3 scale) for all of the questions on each survey, used these averages as an estimate of how strongly each student recognized variation within populations, and performed a paired sample *t* test to determine whether scores increased after the lab. We also calculated the effect size for the lab using Cohen's *d*. While calculating Cohen's *d*, we calculated the pooled standard deviation from the average score of all of the students on the pre- and posttests. Finally, we used paired *t* tests to conduct a post hoc analysis that compared results for individual questions and for subsets of the data. These subsets included: genetic traits, presence/absence traits (have a tail, have canine teeth, hunt other animals for food), quantitative morphological traits (length of femur bone, size of mane in male lions, length of tail, length of canine teeth, and behavioral traits (hunting skill of female lions, mating behavior in male lions).

In addition to assessing how much this lab affected the degree to which students recognized variation within populations, we tested in 2011 for an improvement in hypothetico-deductive reasoning skills. We used a pre/posttest study design to do this, in which all students in the class answered a question requiring hypothetico-deductive reasoning immediately before instruction and another question immediately after instruction. Each student answered the same two questions (Supplemental Material) but the order in which students answered the two questions varied. Approximately half the class answered the "triceratops" question (Supplemental Material) first and approximately half answered the "asteroid" question first. These questions were intended to require the same type of hypothetico-deductive reasoning. We graded student responses to these questions on a scale of 0 to 3. We assigned 3 points to answers that contained a complete hypothetico-deductive argument (e.g., "If an asteroid impact caused dinosaurs to go extinct, and I counted the number of dinosaur fossils in each rock layer, then I should find that dinosaur fossils abruptly end at a single rock layer."). We assigned 2 points to answers that identified specific observations that could be used to test the hypotheses and 1 point to answers that vaguely described some type of observations that could be informative for testing the hypotheses. Finally, we assigned no points to answers that did not begin to answer the question.

RESULTS

Results from the survey of student views showed almost universal positive attitudes toward the lab. Ninety-four percent of the class agreed or strongly agreed with the statement that "The lab was interesting" and 79% agreed or strongly agreed with the statement "The lab was a valuable learning experience."

The survey of typological thinking showed students were more likely to recognize variation within populations and view it as important after completing the lab (Table 3). The average response on the entire survey before instruction was 2.14 (SD = 0.32, $n = 36$). This corresponds to "Varies, but differences do not affect individuals." After instruction, the class average increased to 2.26 (SD = 0.37, $n = 36$). This increase was modest in magnitude (Cohen's $d = 0.35$) but statistically highly significant ($P = 0.007$).

Our assessment of typological thinking was not designed to study how students viewed different types of variation within populations. However, there has been very little research on how university students view variation within species, so we will report results from a few post hoc analyses that may be useful for guiding future research. First, we noted there was a wide variation in the degree to which students perceived variation in lions (Table 3). The three traits with the lowest scores were "Have canine teeth," "Have a tail," and "Hunt other animals for food." Note that these are presence/absence traits (for a body part or a behavior). These were the only presence/absence traits in the study. We included these traits in our survey because they are similar to the traits that Shtulman and Schulz (2008) used in their analysis. The average student score on these traits was 0.86 on pretest, which means that students generally believed these traits could potentially vary in lion populations, but did not. The average pretest student response on the other 13 traits was 2.44, which is much higher. This illustrates that students applied typological thinking much more strongly for some traits than others.

We also noted in our post hoc analysis that the lab changed student views more for some traits than others. The average response increased for 13 out of the 16 traits on the survey, and six of these increases were statistically significant at the 0.05 level (Table 3). These six traits included the three genetic traits on the survey (*MRC1*, *IGF*, hemoglobin). An analysis of subsets of the questions showed students suggested that responses to the genetic traits increased most strongly. The average student response on the three genetic traits on the pretest was 2.15. After the lab, the average was 2.42 ($P = 0.0004$). None of the other subsets of the data showed a statistically significant increase. However, when all of the nongenetic traits were considered together, the increase was statistically significant at the 0.05 level (the average value increased from 2.10 to 2.19, $P = 0.03$), but not if we account for the number of tests that we conducted.

Before instruction, the average score on the hypothetico-deductive reasoning questions was 2.05 ($n = 38$). The average scores for the "triceratops" question and the "asteroid" question before instruction were exactly the same. After instruction, the averages for these two questions were not the same, but the difference was not statistically significant (*t* test: $P = 0.49$). We interpreted this as evidence that the questions were equally difficult, and therefore made no distinction between questions in subsequent analysis.

Our analysis of learning gains for hypothetico-deductive reasoning was complicated by a ceiling effect in the questions we used. Twelve of our 38 students scored a 3 on their hypothetico-deductive reasoning question before instruction. This was the highest score possible. These 12 students, therefore, could not show any improvement after instruction. Including these students in our analysis would decrease the apparent effectiveness of this aspect of the lab, so we dropped these test scores from subsequent consideration.

When we excluded the twelve students with perfect pre-instruction scores, the average on the pretest was 1.62 (SD = 0.62, $n = 26$). After instruction, the average was 2.15 (SD = 0.66, $n = 26$). This increase was highly statistically significant

(one-tailed, paired t test: $P = 0.001$). We interpret this as strong evidence that the 2-h laboratory exercise improved students' ability to construct a hypothetico-deductive argument.

DISCUSSION

We designed this lab to reduce the extent to which students were thinking typologically, to improve the ability of students to construct hypothetico-deductive arguments, and to engage students in an activity they believed was useful and interesting. Our assessment suggests that we accomplished all three goals. This is the first learning activity that we are aware of that reduces typological thinking.

Integrating Lab and Lecture

Any instructional activity is likely to be most effective when it is thoughtfully integrated into the rest of the course. This lab is easy to incorporate into introductory biology courses, because students entering the lab require very little preparation. The only prerequisite knowledge for this lab is a basic understanding of DNA sequences and familiarity with how scientists use hypothetico-deductive reasoning to test hypotheses. Hypothetico-deductive reasoning is explained in the first chapter of many introductory biology tests, but many students may need a review of how hypotheses are tested by making predictions. As part of this lesson, we try to convince students that hypotheses cannot be proven correct but that they can be proven wrong. We use hypothetico-deductive reasoning throughout our course, so the practice students get during this lab is useful in future lessons.

We used this lab as the first part of an instructional unit on natural selection that consisted of two labs (including this one) and six lectures. After students completed this lab, we used two lectures to introduce natural selection and to define the requirements for evolution by natural selection. After these two lectures, students began a second instructional lab (Kalinowski *et al.*, 2006a) that was focused on natural selection. Our third lecture was a 50-min discussion of human evolution designed to elicit and refute common misconceptions regarding selection (Andrews *et al.*, 2011). In these first three lectures, we emphasized the genetic basis of selection (Kalinowski *et al.*, 2010) and repeatedly used the DNA sequences presented in this lab as examples. Our fourth, fifth, and six lectures explored sexual selection, the evolution of complex traits, and the evolution of altruistic behavior, respectively.

Genetic Diversity in Humans

The genetic analysis that students perform in this lab contradicts popular conceptions of race, and although our students were able to answer the focal question of this lab, they frequently had other questions regarding genetic diversity in humans. Therefore, we will briefly review the current state of knowledge of human genetic diversity.

The mitochondrial DNA sequences that students analyzed in this lab are from the seminal study of Ingman *et al.* (2000). Their data included 53 DNA sequences from all over the world; we selected 24 representative sequences from three continents to make the analysis easier. Ingman and colleagues used the 53 DNA sequences to conclude that all humans share a common mitochondrial ancestor who lived approximately 171,000 yr ago in Africa. They noted that some Africans are genetically more similar to Europeans and other non-Africans than to other Africans. Their explanation for this was that the Africans who are genetically similar to non-Africans are living descendants of the Africans who left Africa and settled the rest of the world in the past 100,000 yr.

One of the most common questions students asked after they analyzed their data was "If Africans, Europeans, and Asians are not different races, why do people from each continent look so different?" This question has not been definitively answered in science, but a likely answer is that highly conspicuous traits like skin color are regulated by a small number of genes (Valenzuela *et al.*, 2010) and, therefore, are not a good measure of overall genetic similarity. Differences in skin pigmentation in human populations appear to have evolved through natural selection. Dark skin may be an adaptation in the tropics to prevent cellular damage from UV radiation, and light skin may be an adaptation in higher latitudes to allow the formation of vitamin D (Yuen and Jablonski, 2010). Sexual selection and genetic drift could also explain some of the genetic and morphological differences among human populations.

The mitochondrial DNA sequences the students analyzed in this lab showed that people on each continent are not more similar to one another than to people on other continents. Several students have asked us, "So how is genetic diversity in humans distributed?" This is another complex question. Rosenberg *et al.* (2002) published the first description of human genetic diversity based on a large number of genetic markers and concluded—in contrast with the data presented in this lab—that genetic diversity in humans clusters by continent. Subsequent work challenged this conclusion and has shown human genetic diversity is distributed clinally, that is, genetic differences between human populations tend to be proportional to the geographic distance between populations (Manica *et al.*, 2005; Rosenberg *et al.*, 2005). For example, Rosenberg *et al.* (2005) showed that geographic distance explained 69% of genetic differentiation among human populations and that barriers between continents explained only an additional 5%.

While developing this lab, Kalinowski (2011) reanalyzed the data of Rosenberg *et al.* (2002) that originally had been interpreted as showing human genetic diversity clustered by continent and showed this conclusion may be an artifact of using an inappropriate statistical algorithm to analyze the data. Kalinowski (2011) showed that the genetic similarities among Africans and Europeans that are so evident in the mitochondrial data of Ingman *et al.* (2000) are also present in the largest databases of nuclear genetic data. These relationships had apparently not been noticed before.

Directions for Future Research

The lab described in this paper was designed to help prepare students to learn natural selection. Unfortunately, we were not able to quantify the impact the lab had on how students learned natural selection. The best way to do that would be to have a control group that did not do the lab, and we were not able to do this. Previous work has shown that typological thinking (Shtulman and Shulz, 2008) and poor hypothetico-deductive reasoning skills (Lawson and Thompson, 1988) are obstacles to student learning. We have shown that this lab increases student awareness of variation in populations

and improves their ability to construct hypothetico-deductive arguments, so it is reasonable to conclude that this lab should be helpful to students learning natural selection, but, we were not able to evaluate whether it did.[1] This would be a logical question to address in future research.

A higher priority for future research would be additional study on why students have such a hard time understanding natural selection. Experts view natural selection as a relatively simple process, but scores of studies in the past three decades (for a review, see Gregory, 2009) have shown that students often have a very hard time understanding it. These studies have documented that many students have misconceptions regarding how evolution works and that misconceptions regarding inheritance are common, but almost no research has examined why students have these misconceptions, which students are likely to have them, and what other challenges students face while learning natural selection. It would be easier to design activities to help students understand natural selection if we better understood why natural selection is so difficult for students to understand.

There are two notable exceptions to this lack of research on why students have difficulty learning natural selection (or why they have misconceptions that cause difficulty learning natural selection). As discussed in the *Introduction*, typological thinking and weak hypothetico-deductive reasoning skills have been implicated in difficulty understanding natural selection (Lawson and Thompson, 1988; Lavoie, 1999; Shtulman, 2006; Shtulman and Schulz, 2008). These are important findings that deserve more attention, because if they are true—and further study is warranted—they have obvious and immediate implications for curriculum design. For example, if typological thinking really is a common obstacle to student learning, instructors should target such thinking with activities similar to the lab described in this investigation.

We can only speculate as to why so little work has been done to understand why students have such a hard time understanding natural selection, but a lack of instruments to assess student thinking has probably been a contributing factor. We are not aware, for example, of a validated instrument to assess typological thinking that is appropriate for students in an introductory biology course. This makes understanding how typological thinking affects student learning and how instruction affects typological thinking difficult—and was why we developed our own instrument for this lab. There has been a growing movement to create better instruments to assess student thinking in the biological sciences (e.g., Anderson *et al.*, 2002; Smith *et al.*, 2008; Nadelson and Southerland, 2010), and we hope these efforts will include developing better tools to assess typological thinking and hypothetico-deductive reasoning. Having better tools to identify and quantify these ways of thinking will make it easier to more clearly define specific learning goals and to construct learning activities to achieve them. We hope the lab described here furthers these efforts.

[1]Although we did not evaluate how well the lab described in this paper helped students understand natural selection, we did evaluate the effectiveness of the entire natural selection unit encompassing this lab. As discussed above, this unit was composed of two labs and six lectures. In 2011, we used a 10-question abbreviated version of the Conceptual Inventory of Natural Selection (Anderson *et al.*, 2002; Andrews *et al.*, 2011) to assess student understanding of natural selection before and after our instructional unit on natural selection. The normalized gain (Hake, 1998) for the unit was 0.72. The average normalized gain for this test among randomly selected introductory biology classrooms in the United States is 0.26 (Andrews *et al.*, 2011), so 0.72 is an excellent result.

ACKNOWLEDGMENTS

Thoughtful comments from three anonymous reviewers improved the manuscript. We thank the National Science Foundation for funding (CCLI 0942109).

REFERENCES

Anderson DL, Fisher KM, Norman GJ (2002). Development and evaluation of the conceptual inventory of natural selection. J Res Sci Teach *39*, 952–978.

Andrews TM, Kalinowski ST, Leonard MJ (2011). "Are humans evolving?" A classroom discussion to change student misconceptions regarding natural selection. Evol Educ Outreach *4*, 456–466.

Bishop BC, Anderson C (1990). Student conceptions of natural selection and its role in evolution. J Res Sci Teach *27*, 415–427.

Brumby M (1979). Problems in learning the concept of natural selection. J Biol Educ *13*, 119–122.

Brumby MN (1984). Misconceptions about the concept of natural selection by medical biology students. Sci Educ *68*, 493–503.

Catley KM, Novick LR (2009). Digging deep: exploring college students' knowledge of macroevolutionary time. J Res Sci Teach *46*, 311–332.

Coyne J (2009). Why Evolution Is True, New York: Viking.

Darwin, C. (1859). On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life, 1st ed., London: John Murray.

Demastes SS, Settlage J, Good R (1995). Students' conceptions of natural selection and its role in evolution: cases of replication and comparison. J Res Sci Teach *32*, 535–550.

Dobzhansky T (1973). Nothing in biology makes sense except in the light of evolution. Am Biol Teach *35*, 125–129.

Douglas L (2002). Inquire Within: Implementing Inquiry-Based Science Standards, Thousand Oaks, CA: Corwin Press.

Feynman RP (1984). "Surely You're Joking, Mr. Feynman": Adventures of a Curious Character, New York: WW Norton.

Greene ED (1990). The logic of university students' misunderstanding of natural selection. J Res Sci Teach *27*, 875–885.

Gregory TR (2009). Understanding natural selection: essential concepts and common misconceptions. Evol Educ Outreach *2*, 156–175.

Hake RR (1998). Interactive-engagement versus traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses. Am J Phys *66*, 64–74.

Herron MD (1971). The nature of scientific enquiry. School Rev *79*, 171–212.

Ingman M, Kaessmann H, Paabo S, Gyllensten U (2000). Mitochondrial genome variation and the origin of modern humans. Nature *408*, 708–713.

Ingram EL, Nelson CE (2006). Relationship between achievement and students' acceptance of evolution or creation in an upper-level evolution course. J Res Sci Teach *43*, 7–24.

Jensen MS, Finley FN (1996). Changes in students' understanding of evolution resulting from different curricular and instructional strategies. J Res Sci Teach *33*, 879–900.

Kalinowski ST (2011). The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. Heredity *106*, 625–632.

Kalinowski ST, Leonard MJ, Andrews TM (2010). Nothing in evolution makes sense except in the light of DNA. CBE Life Sci Educ *9*, 87–97.

Kalinowski ST, Taper ML, Metz AM (2006a). Can mutation and natural selection mimic design? A guided inquiry laboratory for undergraduate students. Genetics *174*, 1073–1079.

Kalinowski ST, Taper ML, Metz AM (2006b). How are humans related to other primates? A guided inquiry laboratory for undergraduate students. Genetics *172*, 1379–1383.

Lavoie DR (1999). Effects of emphasizing hypothetico-predictive reasoning within the science learning cycle on high school student's process skills and conceptual understandings in biology. J Res Sci Teach *36*, 1127–1147.

Lawson AE (2002). Sound and faulty arguments generated by preservice biology teachers when testing hypotheses involving unobservable entities. J Res Sci Teach *39*, 237–252.

Lawson AE (2003). The nature and development of hypothetico-predictive argumentation with implications for science teaching. Int J Sci Educ *25*, 1387–1408.

Lawson AE, Thompson LD (1988). Formal reasoning ability and misconceptions concerning genetics and natural selection. J Res Sci Teach *9*, 733–746.

Manica M, Prugnolle F, Balloux F (2005). Geography is a better determinant of human genetic differentiation than ethnicity. Hum Genet *118*, 366–371.

Mayr E (2004). What Makes Biology Unique?, Cambridge, UK: Cambridge University Press.

Nadelson LS, Southerland SA (2010). Development and preliminary evaluation of the measure of understanding of macroevolution: introducing the MUM. J Exp Educ *78*, 151–190.

National Research Council (1998). Teaching about Evolution and the Nature of Science, Washington, DC: National Academies Press.

Nehm RH, Reilly L (2007). Biology majors' knowledge and misconceptions of natural selection. Bioscience *57*, 263–272.

Nehm RH, Schonfeld I (2008). Measuring knowledge of natural selection: a comparison of the CINS, an open-response instrument, and an oral interview. J Res Sci Teach *45*, 1131–1160.

Nelson CE (2008). Teaching evolution (and all of biology) more effectively: strategies for engagement, critical reasoning, and confronting misconceptions. Integr Comp Biol *48*, 213–225.

Platt JR (1964). Strong inference: certain systematic methods of scientific thinking may produce much more rapid progress than others. Science *146*, 347–353.

Renniger KA, Hidi S, Krapp A (eds.) (1992). The Role of Interest in Learning and Development, Mahwah, NJ: Erlbaum.

Rosenberg NA *et al.* (2002). Genetic structure of human populations. Science *298*, 2381–2385.

Rosenberg NA, Mahajan S, Ramachandran S, Zhao C, Pritchard JK, Feldman MW (2005). Clines, clusters, and the effect of study design on the inference of human population structure. PLoS Genet *1*, e70.

Shtulman A (2006). Qualitative differences between naïve and scientific theories of evolution. Cogn Psychol *52*, 170–194.

Shtulman A, Schulz L (2008). The relation between essentialist beliefs and evolutionary reasoning. Cogn Sci *32*, 1049–1062.

Smith JJ, Baum DA, Moore A (2009). The need for molecular genetic perspectives in evolutionary education (and vice versa). Trends Genet *25*, 427–429.

Smith MK, Wood WB, Knight JK (2008). The genetics concept assessment: a new concept inventory for gauging student understanding of genetics. CBE Life Sci Educ *7*, 422–430.

Solomon GE. A, Zaitchik D (2012). Folkbiology. WIRESs Cogn Sci *3*, 105–115.

Valenzuela RK *et al.* (2010). Predicting phenotype from genotype: normal pigmentation. J Forensic Sci *55*, 315–321.

Yuen AWC, Jablonski NG (2010). Vitamin D: In the evolution of human skin colour. Med Hypoth *74*, 39–44.

Zusho A, Pintrich PR, Coppola B (2003). Skill and will: the role of motivation and cognition in the learning of college chemistry. Int J Sci Educ *25*, 1081–1094.