Mapping a Coherent Learning Progression
for the Molecular Basis of Heredity

Jo Ellen Roseman, Ann Caldwell, Arhonda Gogos, and Lori Kurth
Project 2061, American Association for the Advancement of Science

Introduction

Understanding the molecular basis of heredity requires a coherent understanding of two main functions of DNA: (1) determining the characteristics of organisms and (2) passing information from one generation to the next. The expectation is that students know and are able to relate ideas about information transfer and use across several levels of biological organization and over time. Focusing on a learning progression that leads to a coherent understanding of these two functions of DNA presents a potentially useful contrast to the typical treatment of the topic in textbooks, where Mendelian inheritance is presented before DNA, and DNA is unconnected to the actions of proteins (Kurth & Roseman, 2001).

As part of its evaluation of high school biology textbooks, Project 2061 mapped the ideas that students were expected to know for each of the topics being used as a basis for the study. For the topic of the molecular basis of heredity, the textbook evaluation map (http://www.project2061.org/publications/textbook/hsbio/summary/mbh_com.pdf) suggests a learning progression leading to a coherent understanding of the two main functions of DNA. Text statements in boxes on the map specify what ideas students should know and arrows indicate which ideas contribute to which others. Statements are drawn from learning goals in *Benchmarks for Science Literacy* (*Benchmarks*) (American Association for the Advancement of Science [AAAS], 1993) and *National Science Education Standards* (*NSES*) (National Research Council [NRC], 1996). The proposed learning progression differs from current practice in two significant ways by emphasizing proteins before DNA and focusing on DNA before genes and chromosomes. We argue that this reordering of the presentation of ideas makes both logical and psychological sense and is a ripe area for empirical testing.

For the first function of DNA, the map shows that the idea that proteins do the work of cells is precursor to ideas about DNA and heritable characteristics. For students to understand the role of DNA in determining the characteristics of an organism, they need to know that an organism's characteristics reflect the actions of its proteins, that the sequence of (nucleotide) subunits in DNA determines the sequence of amino acid subunits in proteins, and that the sequence of amino acid subunits affects a protein's 3-dimensional structure, and, hence, its function. However, unless students also know that proteins carry out nearly all of the work of cells (and organs/organ systems), they are unlikely to make sense out of the other ideas.

For the second function of DNA, the map shows that ideas about DNA are precursor to more abstract ideas about genes. Ideas about the passage of genetic information—that is, the information for specifying the heritable characteristics of organisms—from one generation to the next are first understood in terms of DNA molecules. Then (and only then) can genes and chromosomes be defined in terms of DNA (Deadman & Kelly, 1978; Lewis, 2000; Longden, 1982; Lucas, 1987; Marbach-Ad, 2001). Given the documented learning difficulties students have with Mendelian inheritance (Banet & Ayuso, 2000; Moll & Allen, 1987; Smith, 1988; Stewart, 1982, 1983), it makes sense to postpone discussion of genes, chromosomes, and patterns

of inheritance until students have a firm grasp of the role of DNA in determining the characteristics of organisms and in passing information from one generation to the next.

In developing a map to serve the needs of the high school biology textbook evaluation, we included learning goals from four different conceptual strand maps in *Atlas of Science Literacy* (AAAS, 2001): DNA and Inherited Characteristics (p. 69), Variation in Inherited Characteristics (p. 71), Cell Functions (p. 73), and Atoms and Molecules (p. 55). Most of the learning goals are from grades 9-12, with a few middle school learning goals included as prerequisites. However, to provide context for discussing the proposed learning progression, we have also included precursors from grades K-2 and 3-5 on the map in Figure 1.

Learning Progressions
The map in Figure 1 displays ideas in a learning progression that leads to an understanding of the role of DNA in determining the characteristics of an organism (the relevant ideas are indicated on the map by blue triangles: ◢). With one exception, the ideas are drawn from the section on Cells in Chapter 5: The Living Environment of *Benchmarks for Science Literacy*. The learning progression proposes a logical and developmentally appropriate sequence of ideas from the primary grades through high school. In grades K-2 students learn that (most) living things need food, water, and air. In grades 3-5 students' notion of living things is expanded to include the idea that single-celled organisms are likely to have the same needs as macroscopic organisms, that some organisms are made of a collection of similar cells that benefit from cooperating, and that some organisms' cells vary greatly in function and perform very different roles in the organism. In middle school students learn that cells carry out basic functions of the organism. In high school students learn that it is the molecules within cells—mainly proteins—that do the work. Another high school benchmark—an organism's characteristics reflect the actions of its proteins—arose out of both logical reasoning (how to help students make the leap from protein molecules to phenotype) and empirical evidence from Project 2061's student interviews, described below.

The same map (Figure 1) also displays ideas in a learning progression that leads to an understanding of the role of DNA in passing information from one generation to the next (the relevant ideas are indicated on the map by red triangles: ◤). Drawing mostly from benchmarks in the Heredity section of *Benchmarks'* Chapter 5: The Living Environment, the learning progression expects students in grades K-2 to learn that offspring resemble their parents (rather than other kinds of organisms). In grades 3-5 students learn that for offspring to resemble their parents, there must be a reliable way to transfer information from one generation to the next. In high school, with their prior knowledge of cells and protein molecules, students are ready to learn about the link between proteins and DNA and, hence, between DNA and traits.

To test the validity of these proposed learning progressions, we need to identify learning activities to help students understand the ideas and assessment items to monitor their progress. As will be discussed below, the careful thinking we engage in to design activities and assessments that are aligned to the specific key ideas often leads to refinements in the learning progression itself.

Assessing and Promoting Student Progress
Project 2061 is currently working to develop assessment items to monitor students' progress along these learning trajectories and to identify phenomena that are likely to help them learn the ideas on the map. In doing so, we are finding it essential to work at a grain size of precision that is finer than individual learning goals (benchmarks or standards). This involves unpacking the learning goal into its component ideas and clarifying the meaning of each one.

Unpacking learning goals. *Science for All Americans* (*SFAA*) (AAAS, 1989) attempted to tell a coherent story involving the ideas that constitute science literacy. Authors of *Benchmarks* and

NSES tried to maintain some of that coherence as they specified what was appropriate for students to learn in earlier grades. As a result, individual benchmarks and standards often include multiple ideas, as illustrated by the following high school benchmark from our map in Figure 1:

> The work of the cell is carried out by the many different types of molecules it assembles, mostly proteins. Protein molecules are long, usually folded chains made from 20 different kinds of amino-acid molecules. The function of each protein molecule depends on its specific sequence of amino acids and the shape the chain takes is a consequence of attractions between the chain's parts (5C/H3).

After considerable discussion, Project 2061 staff unpacked this benchmark into three discrete ideas that would each require its own specific assessment items and instructional activities:

> Idea a: The work of the cell is carried out by the many different types of molecules it assembles, mostly proteins (5C/H3a).

> Idea b: Protein molecules are long, folded chains made from 20 different kinds of smaller (amino-acid) molecules. The shape of a protein molecule depends on interactions among the amino acids and between them and their environment (based on 5C/H3bc).

> Idea c: The function of a protein molecule depends on its shape and, therefore, on its specific sequence of amino acids (based on 5C/H3c).

It is important to note that as we clarify the ideas and design assessment items and activities that are aligned to them, we sometimes decide that an idea is too big or too small to be useful. Both the grain size and the language of the ideas are revised accordingly.

Clarifying ideas. Each idea in the learning progression is then clarified and its boundaries described in terms of what specific knowledge students are and are not expected to know. The clarifications also specify technical terms students should know and terms that go beyond expectations. In setting boundaries for an idea, we look to information provided in *SFAA*, along with other benchmarks and standards. For example, the following clarification of the idea that "the work of the cell is carried out by the many different types of molecules it assembles, mostly proteins" (see Idea a above) limits the cell functions that students should know to those explicitly stated in *SFAA*:

> Students should know that (1) the work of the cell is carried out mainly by proteins and (2) the cell assembles the proteins to do its work.

> By work, we mean the many critical functions a cell needs to carry out for itself, such as repairing cell structures, helping other molecules to get in or out of the cell, replicating genetic information, and generally catalyzing and regulating molecular interactions (*SFAA p.63)*. This idea does not refer to proteins that provide structural support (organization and shape) for the cell, which are discussed under another benchmark, 5C/M1.

> Students should know that proteins carry out critical functions for the organism such as carrying oxygen, effecting contraction, responding to outside stimuli, or providing material for hair, nails, and other body structures (*SFAA p.63)*.

> Students should know that cells and organisms have thousands of different protein molecules and that different proteins do different work (i.e. contribute to different functions). Students are not expected to know how proteins do the work (e.g., mechanisms of enzyme catalysis or ion channels) or the names of specific proteins.

> Students should know that a cell's ability (or an organism's ability) to carry out a particular function depends on the cell's ability to make the particular proteins that do that work. Students should know that (a) if the particular protein is missing or is not functional, then the corresponding function will not be carried out and (b) if the particular molecule is only partially functional then the corresponding function will be carried out less efficiently. Students are not expected to know

that other types of molecules may contribute to the work of proteins (e.g., RNA's role in catalysis or the role of carbohydrates as signaling molecules).

Identifying commonly held student ideas. The next step in developing assessments and activities related to the proposed learning progression is to study and summarize the learning research, identifying common student misconceptions relevant to each key idea. The common misconceptions provide the basis for the design of distractors in our assessment items and inform the selection of phenomena likely to help students make progress in their learning.

If research is not available, we draft questions, interview students, summarize our findings, use the findings to revise questions, interview more students, and so on, until we have identified the most common student ideas. For example, after several rounds of interviews, we have learned that some students do not understand that proteins do the work of the cell. When asked whether proteins are essential, students refer to their importance to the body as a whole rather than to their role in cell function. When asked specifically about the work of proteins in cells, most students have been unable to give examples. Even students who can describe the "lock in key" mechanism of protein function do not relate this to the work proteins do in cells. In probing students' understanding of the relationship between DNA and protein, students who were able to describe adequately the sequence of steps from DNA to mRNA to tRNA to proteins still did not see the big picture—that it is the sequence of nucleotides in DNA that determines the sequence of amino acids in proteins made from that sequence, and that the amino acid sequence affects the shape and, hence, the function of proteins.

What we learn about students' thinking from the available research and from our piloting and student interviews also helps us to revise and fine-tune the clarifications of the key benchmark ideas. For example, the statement in the clarification that different proteins do different work—i.e., have different functions—was added because of interview data showing that students often (a) cannot give a single example of work that proteins do (as opposed to making up hair or nails) or of how different proteins could do different work and (b) have no idea how many proteins would be needed to carry out the many functions needed to sustain life.

Developing assessment items. When we have identified a set of ideas that represents a range of incorrect or incomplete conceptions that many students are likely to have, we test them as plausible distractors for multiple choice assessment items. The items are then piloted with students who are also interviewed or asked to respond to written questions that probe the effectiveness of all of the answer choices—both correct and incorrect—in appealing to students who hold those particular ideas. The items are then revised again based on our findings (DeBoer, et al., 2006). Eventually, these assessment items will become part of an online collection of standards-based assessment resources for middle and early high school science. In addition to the items themselves, the collection will include clarifications of benchmarks and standards and concept maps that depict learning progressions and conceptual connections for more than a dozen science and mathematics topics. For more information on Project 2061's assessment work, go to http://www.project2061.org/research/assessment.htm.

Identifying aligned phenomena. At the same time that we are developing assessment items, we also use the clarification and learning research to draft specifications for phenomena that are likely to help students understand the key ideas in the learning progression. Ideally, phenomena should illustrate or provide examples for each general principle (as articulated in the clarification) that students are expected to know and provide evidence to counteract common student misconceptions. Whenever possible, phenomena should be directly observable by students and,

when not possible, should require few inferences from data. Based on the clarification of Idea a, we have specified the following kinds of phenomena and have begun to identify specific examples of each:

1. Instances in which cells are able to carry out a particular function in the presence of a particular protein but not in its absence. Functions could include repairing cell structures, helping other molecules to get in or out of the cell, replicating genetic information, catalyzing and regulating molecular interactions.
   a. When the action of a protein (myosin) is inhibited (by giving cells blebbistatin), the cells fail to complete the final stages of cell division. Without the inhibitor, cells divide normally.
   http://archives.focus.hms.harvard.edu/2003/March21_2003/cell_biology.html

2. Instances in which cells are able to carry out a particular function less efficiently in the presence of a partially functioning protein than in the presence of a "fully" functioning protein. Functions could include repairing cell structures, helping other molecules to get in or out of the cell, replicating genetic information, or catalyzing and regulating molecular interactions.

3. Instances in which organisms (humans or experimental models) are able to carry out a particular function in the presence of a particular protein but not in its absence. Functions could include carrying oxygen; effecting contraction; responding to outside stimuli;or providing material for hair, nails, and other body structures.

4. Instances in which organisms (humans or experimental models) are able to carry out a particular functions less efficiently in the presence of a partially functioning protein than in the presence of a "fully" functioning protein. Functions could include carrying oxygen; effecting contraction; responding to outside stimuli; or providing material for hair, nails, and other body structures.

Our proposed learning progression for understanding the molecular basis of heredity rests on the assumption that a sufficient number of instances of these types can be found and that they can be made comprehensible to students at the appropriate grade levels. The set of instances should be sufficiently large and diverse to give students a sense of the wide range of functions carried out by proteins and the large number of different proteins that would be needed to carry out those functions.   Identifying such instances is one of our greatest challenges but well worth the effort.

Conclusion
Project 2061's *Atlas of Science Literacy* (AAAS, 2001) presents conceptual strand maps on various science literacy topics that are based on the logic of the discipline and existing learning research. Strand maps represent testable learning trajectories for curriculum studies (Roseman, Kurth, & Stern, 2004). The research and development studies described in this session are a good start. To develop curriculum materials that can help all students achieve science literacy goals, developers need to build instructional activities around phenomena that are aligned with the goals and construct assessment items that are designed to monitor students' progress in learning the precise ideas that they specify. We hope that the assessment and curriculum resources Project 2061 is developing will support curriculum research and development and lead to empirical testing of the learning progressions.

Heritable characteristics ultimately produced in the development of an organism can be observed at molecular and whole-organism levels--in structure, chemistry, or behavior. 5F/H4...

When mutations occur in sex cells, they can be passed on to all cells in the resulting offspring; if mutations occur in other cells, they can be passed on to descendant cells only. ...5B/H5...

An altered gene may be passed on to every cell that develops from it (that cell). ...5B/H4...

A mutation of a DNA segment may not make much difference in the operation of the cell, may fatally disrupt it, or may change it in a significant way. *SFAA*, p. 64

Some faulty operations of body processes are known to be caused by altered genes. They may have a direct, obvious effect, such as causing easy bleeding, or they may only increase the body's susceptibility to developing particular diseases, such as clogged arteries or mental depression. *SFAA*, p. 81

A change in even a single atom in the DNA molecule... can...change the protein that is produced. *SFAA*, p. 64

Insertions, deletions, or substitutions in DNA can alter genes. ...5B/H4...

Each DNA molecule in a cell forms a single chromosome. *NSES* 9-12, p. 185

Genes are segments of DNA molecules. Each DNA molecule contains thousands of discrete genes. *SFAA*, pp. 61-62

The genetic information stored in DNA is used to direct the synthesis of the thousands of proteins that each cell requires. *NSES*, 9-12, p. 184, 5C/H4...

Changes in DNA (mutations) occur... *NSES* 9-12, p. 185

The information passed from parents to offspring is coded in DNA molecules. 5B/H3

DNA molecules are long chains linking just four kinds of smaller molecules, whose precise sequence encodes genetic information. *SFAA*, p. 61

An organism's traits reflect the actions of its proteins. NEW.

Protein molecules are long, often elaborately folded chains made from 20 different kinds of smaller (amino-acid) molecules. The function of each protein molecule depends on its shape. The shape depends on interactions among the amino acids and between them and their environment. 5C/H3bc

The work of the cell is carried out by the many different types of molecules it assembles, mostly proteins. 5C/H3a

All matter is made up of atoms... Atoms may stick together in well-defined molecules or may be packed together in large arrays. Different arrangements of atoms into groups compose all substances. 4D/M1

Within cells, many of the basic functions of organisms—such as extracting energy from food and getting rid of waste—are carried out. The way in which cells function is similar in all living organisms. 5C/M3

Cells repeatedly divide to make more cells for growth and repair. 5C/M2a

For offspring to resemble their parents, there must be a reliable way to transfer information from one generation to the next. 5B/E2

All living things are composed of cells, from just one to many millions, whose details usually are visible only through a microscope. Different body tissues and organs are made up of different kinds of cells. 5C/M1ab

Offspring are very much, but not exactly, like their parents and like one another. 5B/P2

Some living things consist of a single cell. Like familiar organisms, they need food, water, and air; a way to dispose of waste; and an environment they can live in. 5C/E1

Molecular Basis of Heredity

Adapted March 2006 from map used in AAAS Project 2061's Biology Textbook Evaluation
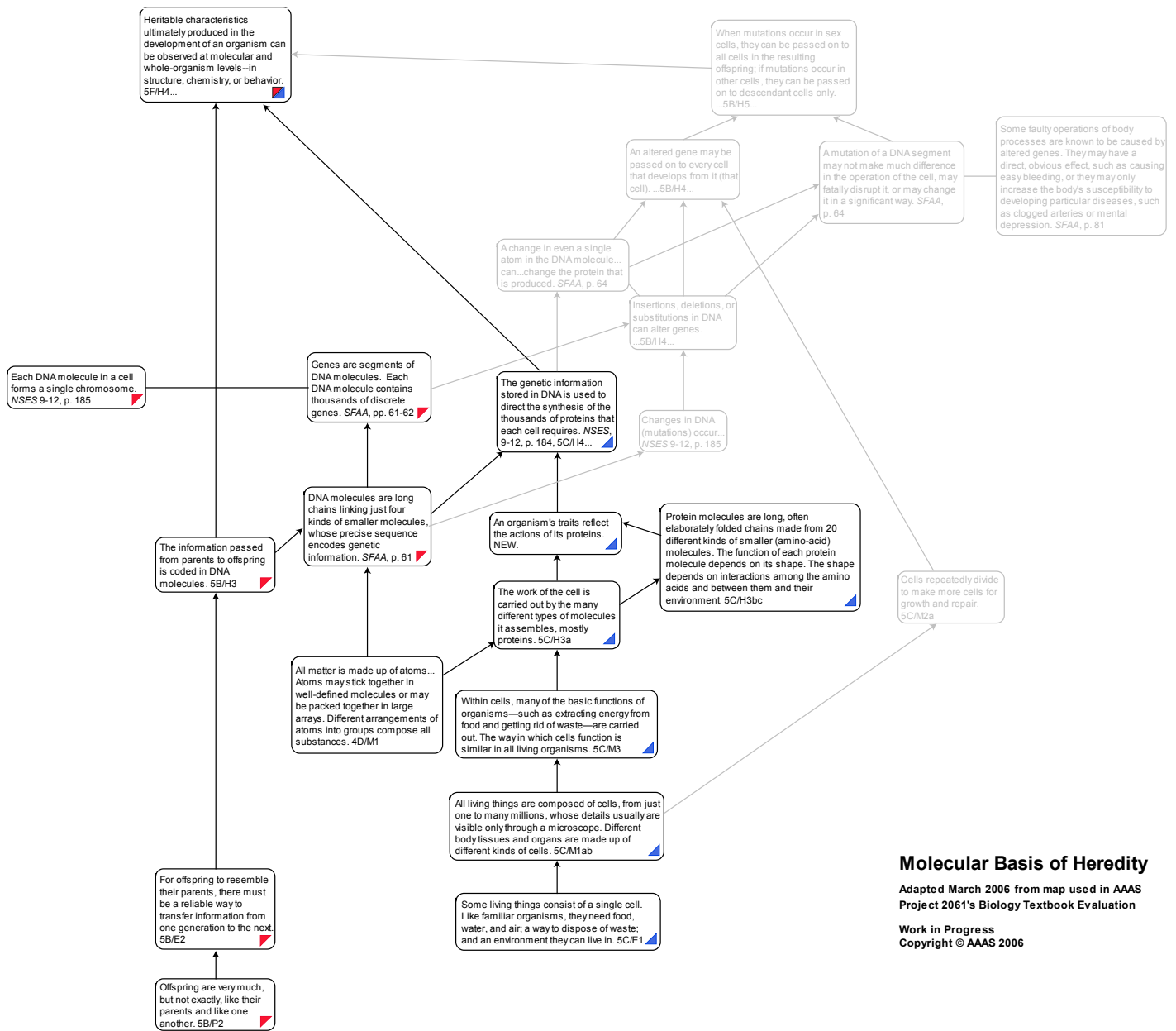
Work in Progress
Copyright © AAAS 2006

Figure 1. Map showing the progression of ideas that leads to an understanding of the role of DNA in determining the characteristics of an organism (▲) and its role in passing information from one generation to the next (▼).

References

American Association for the Advancement of Science. (1989). *Science for all Americans*. New York: Oxford University Press.

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. New York: Oxford University Press.

American Association for the Advancement of Science. (2001). *Atlas of science literacy*. Washington, DC: Author.

Banet, E., & Ayuso, E. (2000). Teaching genetics at secondary school: A strategy for teaching about the location of inheritance information. *Science Education, 84,* 313-351.

Deadman, J., & Kelly, P. (1978). What do secondary school boys understand about evolution and heredity before they are taught the topics? *Journal of Biological Education, 12*, 7-15.

DeBoer, G. E., Roseman, J. E., Gogos, A., Herrmann-Abell, C., Michiels, A., Regan, T., Willard, T., & Wilson, P. (2006, February). *Pilot testing assessment items: How student feedback informs the development of assessment items aligned to content standards.* Poster session presented at the Principal Investigators' Meeting of the National Science Foundation's Instructional Materials Development Program. Alexandria, VA.

Kurth, L. A., & Roseman, J. (2001). *Findings from the high school biology curriculum study: Molecular basis of heredity*. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, St. Louis, MO.

Lewis, J. (2000). Genes, chromosomes, cell division and inheritance—Do students see any relationship? *International Journal of Science Education, 22*(2), 177-195.

Longden, B. (1982). Genetics—Are there inherent learning difficulties? *Journal of Biological Education, 16*(2), 135-140.

Lucas, A. (1987). Public knowledge of biology. *Journal of Biological Education, 21*(1), 41-45.

Marbach-Ad, G. (2001). Attempting to break the code in student comprehension of genetic concepts. *Journal of Biological Education*, *35*(4), 184-189.

Moll, M. B., & Allen, R. D. (1987). Student difficulties with Mendelian genetics problems. *The American Biology Teacher, 49,* 229-233.

National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.

Roseman, J., Kesidou, S., Kurth, L., & Stern, L. (2004). *Mapping for curriculum coherence*. Symposium presented at the Annual Meeting of the National Association for Research in Science Teaching, Vancouver, Canada.

Smith, M. U. (1988). Successful and unsuccessful problem solving in classical genetic pedigrees. *Journal of Research in Science Teaching, 25,* 411-433.

Stewart, J. (1982). Difficulties experienced by high school students when learning basic Mendelian genetics. *American Biology Teacher, 44*, 80-84.

Stewart, J. (1983). Student problem solving in high school genetics. *Science Education, 67*, 523-540.