JOURNAL OF RESEARCH IN SCIENCE TEACHING VOL. 44, NO. 7, PP. 938–959 (2007)

Reasoning Across Ontologically Distinct Levels: Students' Understandings of Molecular Genetics

Ravit Golan Duncan,[1] Brian J. Reiser[2]

[1]*Graduate School of Education, Rutgers, The State University of New Jersey, 10 Seminary Place, New Brunswick, New Jersey 08901*

[2]*School of Education and Social Policy, Northwestern University, 2120 Campus Drive, Evanston, Illinois 60208*

Received 14 December 2005; Accepted 19 October 2006

Abstract: In this article we apply a novel analytical framework to explore students' difficulties in understanding molecular genetics—a domain that is particularly challenging to learn. Our analytical framework posits that reasoning in molecular genetics entails mapping across ontologically distinct levels—an information level containing the genetic information, and a physical level containing hierarchically organized biophysical entities such as proteins, cells, tissues, etc. This mapping requires an understanding of what the genetic information specifies, and how the physical entities in the system mediate the effects of this information. We therefore examined, through interview and written assessments, 10th grade students' understandings of molecular genetics phenomena to uncover the conceptual obstacles involved in reasoning across these ontologically distinct levels. We found that students' described the genetic instructions as containing information about both the structure and function of biological entities across multiple organization levels; a view that is far less constrained than the scientific understandings of the genetic information. In addition, students were often unaware of the different functions of proteins, their relationship to genes, and the role proteins have in mediating the effects of the genetic information. Students' ideas about genes and proteins hindered their ability to reason across the ontologically distinct levels of genetic phenomena, and to provide causal mechanistic explanations of how the genetic information brings about effects of a physical nature. © 2007 Wiley Periodicals, Inc. J Res Sci Teach 44: 938–959, 2007

Keywords: molecular biology; genetics; conceptual change; science literary

Reasoning about scientific phenomena is challenging for many reasons. In some cases it is the very nature of the scientific ideas underlying these phenomena that is intrinsically difficult. One such case is molecular genetics. The phenomena in molecular genetics that are commonly studied at the high school and college level, such as the genetic code, expression of inherited traits, and

Contract grant sponsor: National Science Foundation; Contract grant number: REC-9980055.
Correspondence to: R.G. Duncan; E-mail: rgduncan@rci.rutgers.edu
DOI 10.1002/tea.20186
Published online 16 February 2007 in Wiley InterScience (www.interscience.wiley.com).

genetic disorders [American Association for the Advancement of Science (AAAS), 1993; National Research Council (NRC), 1996], are complicated multileveled phenomena that are not trivial to understand. Prior research on genetics learning has shown that students often fail to understand the interactions between the different organization levels inherent to genetic phenomena (Hill, O'Sullivan, Stanisstreet, & Boyes, 1998; Horwitz, 1996; Lewis & Wood-Robinson, 2000; Marbach-Ad & Stavy, 2000; Stewart, 1990), and are confused about the roles of key entities and mechanisms in genetics (Bahar, Johnstone, & Sutcliffe, 1999; Fisher, 1985; Friedrichsen & Stone, 2004; Kindfield, 1994; Marbach-Ad, 2001; Venville & Treagust, 1998; Wynne, Stewart, & Passmore, 2001). The difficulties students encounter in genetics are also worrisome from a practical perspective due to the growing importance and relevance of this field to our everyday lives (DNA-based evidence, genetic screening, etc.). Consequently, understanding concepts in this domain is, and will continue to be, a critical aspect of scientific literacy (AAAS, 1993: NRC, 1996; Garton, 1992; Lewis & Wood-Robinson, 2000).

Theoretical Framework

What makes understanding molecular genetics phenomena so challenging for learners? Two prevalent explanations exist in the literature. The first explanation attributes students' difficulties to the invisibility and inaccessibility of genetic phenomena. The claim is that it is challenging for students to understand phenomena involving small (and often hidden) entities and processes that they cannot directly experience (Gilbert, Osborne, & Fensham, 1982; Kapteijn, 1990). Moreover, when students attempt to understand these unfamiliar molecular entities by drawing inferences from more familiar macro entities they often err as a result. For example, students tend to assume that the function of the nucleus inside a cell is analogous to the function of the brain in the human body; this is an inadequate and misleading idea (Dreyfus & Jungwirth, 1990).

The second type of explanation is anchored in the complicated structure of genetic phenomena, namely, the multiplicity of levels involved. Phenomena in genetics span multiple biological organization levels—genes, proteins, cells, tissues, organs, etc. The levels are organized hierarchically such that elements at one organization level constitute the elements of progressively higher organization levels; for example, cells constitute tissues, which in turn constitute organs. The dynamics of genetic phenomena are typical to hierarchically organized systems in that interactions at lower organization levels bring about effects at higher organization levels (Casti, 1994; Simon, 1996). Therefore, understanding genetic phenomena entails understanding how mechanisms and interactions at the molecular (genes, proteins) and microlevels (cells) bring about effects at the macrolevel (organism, population). Many researchers have shown that students have difficulty reasoning across different organization levels in genetics (Horwitz, 1996; Marbach-Ad & Stavy, 2000; Stewart & Van Kirk, 1990), and in other contexts such as ecology, evolution, the respiratory system, and chemistry (Ferrari & Chi, 1998; Hmelo, Holton, & Kolodner, 2000; Penner, 2000; Resnick, 1996; Wilensky & Resnick, 1999).

We believe that although the explanations provided in the literature contribute to our understanding of what makes reasoning in this domain challenging, they do not fully account for the difficulties students encounter. We propose that there may be an important third factor contributing to students' difficulties: the ontological differences between the levels of genetic phenomena. These phenomena are brought about by interactions between informational entities (genes) existing at one organizational level and biophysical entities (proteins, cells, tissues, etc.) that exist at other organizational levels. We therefore term genetic phenomena *hybrid hierarchical*, as they are a hybrid of ontologically distinct levels that are organized hierarchically. This characterization of genetic phenomena provides us with a theoretical lens that brings to focus

a new type of conceptual obstacle—the need to coordinate and integrate knowledge about ontologically distinct entities. To reason about hybrid hierarchical phenomena one needs to understand how information brings about effects of a physical nature. In this study we refer to reasoning as the active process of constructing explanations, inferences, and predictions about genetic phenomena. We refer to understanding as the comprehension of the structures, functions, and interactions of entities involved in these phenomena (reasoning about a phenomenon thus employs understandings of it).

To further clarify what we are referring to as hybrid hierarchical phenomena and to illustrate the reasoning involved, let us consider a simpler example of such a phenomenon: the player piano. The information level of the player piano system contains information regarding the type, order, and timing of musical notes. This informational content is coded in the paper scroll of the piano. The physical level of the player piano is composed of the scroll (which physically stores the information by way of the holes punched across the length of the scroll) and the mechanical parts that operate on the scroll to generate the musical tune. An important feature of the player piano, which is also relevant to other hybrid hierarchical phenomena such as genetics, is that the information level does not contain information about how the player piano should work to generate the tune. Rather, the scroll contains information regarding only a small fraction of the entire ensemble of structures and functions that are involved in generating the tune. Specifically, the holes in the scroll allow the penetration of specific pegs at specific points causing certain piano keys to be activated such that particular notes are played. The tune itself emerges from the interaction of those pegs with other physical components of the piano that results in the playing of notes in a sequence that we recognize as a musical tune.

To explain how a simple player piano works, and be able to diagnose and solve problems with it, one needs to understand what exactly is encoded in the scroll and what are the mechanisms that link the information that is embedded in the scroll and the musical tune that we hear. Stated more generally, it is necessary to understand the nature of the informational content in the information level, the way in which this information is manifested in specific physical elements in the system, and the array of physical mechanisms involved in bringing about the information's ultimate effects.

Similar reasoning is needed in genetics. In this case, too, there is an information level (containing the genetic information) and a hierarchically organized physical level (see Figure 1). The informational content embodied in the information level specifies the type and order (sequence) of amino acids—the building blocks of proteins. Thus, the information in our genes determines the structure of proteins by specifying the sequence of amino acids that comprise the protein. The amino acid chain then folds into the functional three-dimensional form of the protein. The structure of the protein (its amino acid sequence) thus determines its function, and the structure itself is determined by the genetic information. The physical level includes the DNA molecule, in which the informational content is stored (similar to the player piano scroll), proteins that are directly coded for by the information, and other structural elements that make up progressively higher organizational levels: cells, tissues, and organs.

We wish to point out that the term *gene* can be used to denote both a unit of information (the gene for a particular trait), and a physical segment on the DNA molecule. The physical DNA sequence is important for the replication of the genetic material (during cell division) and regulation of gene expression (turning the gene on and off) as the molecules that control these aspects of gene function need to interact with the DNA itself. Thus, the role of genes as physical entities is clearly an important one; however, in the context of our theoretical framework we are primarily concerned with the role of genes as comprising the information level of the system.

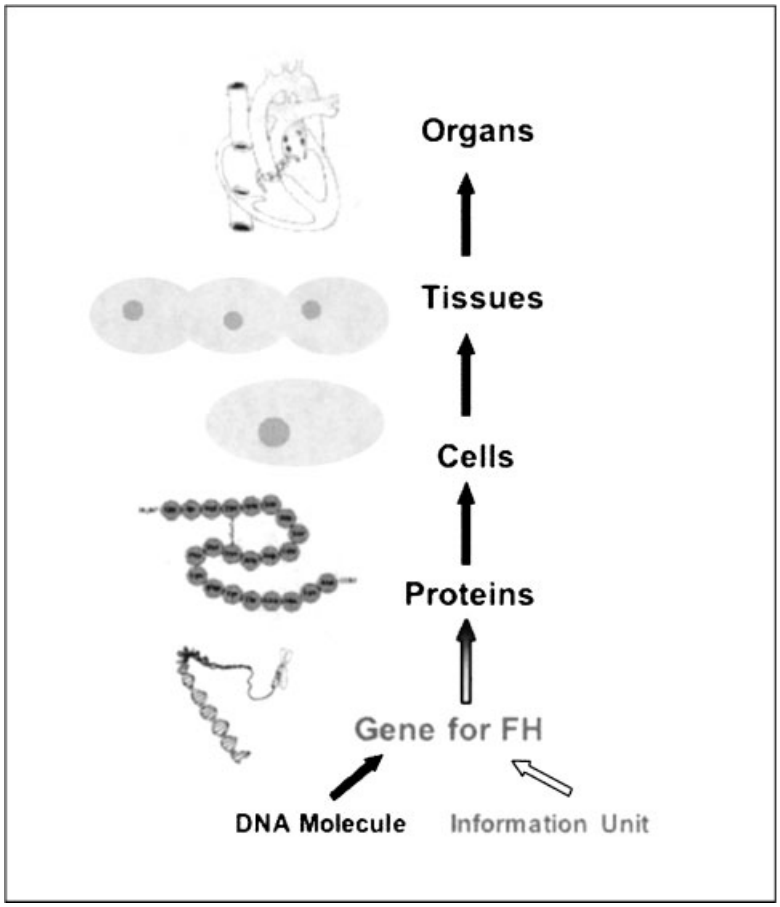

Figure 1. The hybrid hierarchical structure of molecular genetics phenomena.

Therefore, in the remainder of the article we will use the term *gene* to denote a unit of genetic information rather than a physical segment of the DNA molecule.

To explain how genes can bring about an observable physical effect, such as brown eye color, one needs to account for the mechanisms that link the genetic information to the physical outcome (Lewis & Kattmann, 2004). In the case of eye color, the gene encodes a protein that is involved in a cellular process that generates a chemical compound (pigment) that makes the cells appear darker in color. Together, the cells color the tissue that makes up the iris of the eye, resulting in the perceived darker coloration of the eye. This explanation bridges across ontologically distinct levels and specifies: the nature of the informational content (what the information specifies) and the physical mechanisms that bring about the effects of that information. We hypothesize that the understandings involved in generating explanations that bridge the information and physical levels are at the crux of students' difficulties in learning molecular genetics.

In this article we describe an exploratory study in which we examined 10th-grade students' ability to reason across the ontologically distinct levels of genetic phenomena and to explain the mechanisms by which our genes bring about their effects. Specifically, we wanted to know:

1. What are students' understandings of the content of the genetic information?
2. What are students' understandings of the mapping between the information and physical levels, and the physical mechanisms that mediate genetic effects?

Methods

Study Context

The study was conducted during the fall of 2002, in a 10th-grade biology classroom in an urban public high school (a college prep academy) in the Midwest. Students from three class periods, instructed by the same teacher, were included ($n = 64$). We examined students' understandings before and after instruction of relevant concepts. The instruction was carried out over a 5-week period during which students were taught about the discovery of DNA and the deciphering of the genetic code, DNA structure, the processes by which the genetic code is translated into proteins (central dogma), structure–function correlation in proteins (how a change to the amino acid sequence might affect function), and the consequence of mutations to the genetic code.

The main focus of the instruction was on the processes embodied in what biologists refer to as *the central dogma of molecular biology* (Crick, 1970). The central dogma posits that the genetic information is encoded in the DNA molecule. The DNA molecule can replicate itself, segments of the DNA molecule that correspond to genes can also be copied as RNA, and RNA is then translated into a protein. The process of translation matches triplets of nucleotide bases, as they appear in the RNA copy of the DNA, with specific amino acids.

In the study classroom students spent a considerable amount of time building bead models of the DNA molecule and using interactive computer animations to simulate, in detail, the processes of the central dogma. Modeling and simulations have been shown to be effective in promoting understanding of the processes embodied in the central dogma (Rotbain, Marbach-Ad, & Stavy, 2006). Students learned how to "interpret" the genetic code, and had ample experience translating this code into proteins. They also learned about the biological basis of a few genetic phenomena: sickle cell anemia and antibiotic resistance in bacteria. Overall, the progression of instruction was relatively typical for this domain in terms of the topics and their sequence of presentation—based on the district and state curriculum guidelines and the presentation of this subject matter in high school textbooks (such as McLaren, Rotundo, & Gurley-Dilger, 1991; Miller & Levine, 1998). However, there were likely more hands-on activities, in which students manipulated models of the relevant entities and processes, compared to more traditional text-based instruction.

Data Collected

We collected two main sources of data: written assessments and interview data (administered before and after instruction). We also collected video of classroom instruction to document which concepts were taught and the ways in which these concepts were presented to the students.

We conducted clinical interviews (Ginsberg, 1997) before and after instruction with 16 students who were selected based on the teacher's recommendations and who represented all three class periods, both genders, several ethnic backgrounds, and a range of abilities. Interviews took place during a 2-week period before and after the 5-week instruction period. They lasted about 30–45 minutes and were composed of three tasks. There were two comparable versions for the first and second interview task, and the pre- and postinterviews were counterbalanced for these tasks. The interview protocol is provided in Appendix A.

The goal in all three tasks was to elicit students' reasoning, across various genetic phenomena, about how the genetic information brings about effects of a physical nature. In the first task, we asked students to speculate a mechanism at the cellular and molecular organization levels that could account for a phenomenon described in a brief scenario. Both versions of the scenario dealt with genetically determined behaviors of bacterial cells: their ability to sense compounds in their environment (version A), or their ability to move in a particular direction (version B).

In the second task, we asked students to compile an explanation of a genetic phenomenon using nine separate statements we provided about the phenomenon. The two versions of the task were both about genetic disorders: Cystic Fibrosis (version A) and Familial Hypercholesterolemia (version B). Separate statements described the causes of these disorders at the molecular (genetic mutation and altered proteins), cellular (malfunctioning of the relevant cells), tissue/organ level (malfunctioning organ or tissue), and organism level (observable symptoms of the disorder). Although much of the biological details were provided in these statements, students still had to compile a complete mechanistic explanation by inferring the causal connections between the different levels.

The third task was similar to the first task in that students had to provide mechanistic explanations of a genetic phenomenon; however, in this task students could choose the phenomenon they wished to discuss (most students chose a genetic disorder). During this last task we also made sure to ask students to describe the role of genes (and mutations in genes) and the role of proteins in our body. This was to ensure that we discussed these entities even if students did not mention them in previous tasks.

In addition to conducting interviews with a subset of students, we administered pre- and postassessments to students from all three class periods ($n = 64$). The written assessments included nine short answer and multiple-choice questions. The written assessment is provided in Appendix B. The questions were about specific entities in genetic phenomena (mainly proteins and genes), their composition, functions, location in the body, and the relationships between these entities. Some of the questions in the written assessment were adapted from written questionnaires used in prior research (Marbach-Ad, 2001). We did not find it necessary to have two versions of the written assessment because many of the questions were simple and open-ended; moreover, it would have been difficult to find comparable versions of these simple questions. We collected the written assessment data to get a sense of students' understandings of genetic phenomena in the class as a whole, and to triangulate findings from the interviews. Due to the complexity of the interview tasks we could not replicate them in the written assessments. Rather, the assessments focused on students understandings of the structure, function, and relationships of specific entities in the system, whereas the interviews focused on the latter as well as more complex understandings of the mechanisms that link genes to their physical effects.

Data Analysis

In our analysis we attempted to reveal the ways in which students reasoned about: (a) the content of the genetic information, and (b) the physical mechanisms that mediate genetic effects, and the mapping between the information and physical levels. Toward this end we conducted a content analysis of the data, using different parts of the data to ascertain different aspects of students' reasoning.

The interviews were transcribed verbatim, and the transcripts were then coded and analyzed. Classroom videos were viewed once and a narrative was constructed that described the ideas that were mentioned in the lesson and the overall flow of the instructional sequence. This narrative provided a story-board of which concepts were presented to the students, how they were taught,

and in which sequence. Overall, the analyses were carried out through an iterative process in which we constructed and refined the coding schemes through several cycles of analysis (Chi, 1997; Miles & Huberman, 1994). In the Results section we provide details of the analysis and data sources used. Despite the pre–post design of the study we did not conduct any statistical analyses of the data due to the small sample size, and because we were primarily interested in understanding what students find difficult in learning genetics and not whether the instruction was effective in producing significant learning gains.

Results

The findings described herein emerged from the analysis of the pre- and postassessments and the postinterviews. We chose to focus on the postinterviews because we found virtually no substantial differences in students' reasoning between the pre- and postinterviews. In this regard, the postinterviews helped us uncover the aspects of reasoning that remained problematic even after instruction of relevant content. In hindsight, it is not surprising that there was no substantial difference between the pre- and postinterviews because the interview tasks focused on aspects of reasoning—a multileveled and mechanistic view of genetic phenomena that links the genetic information to physical effects—that were not emphasized in the classroom instruction. However, being able to explain how genes bring about their effects was an expected learning outcome in the classroom (expected by us as well as the teacher); an expectation that is consonant with the learning objectives stated in national standards (Benchmark 5C in AAAS, 1993; content Standard C in NRC, 1996). The point we wish to make here is that although these expectations of what students should learn existed, the instruction did not foster such learning. We discuss our speculations as to why that was the case in the Discussion section.

Research Question 1: Students' Understandings of the Content of the Genetic Information

A key aspect of reasoning about hybrid hierarchical phenomena is conceiving of them as containing information and understanding how that information is mapped onto the physical level. We suspected that students' understanding of the interrelation between the information and physical levels would depend to a great degree on their conceptions of the information level itself and what that information specifies. The latter aspect gets at students' understanding of the mapping between the information and physical level because understanding what the information specifies amounts to understanding which physical entity in the system it is mapped onto.

We therefore analyzed students' responses to one of the written assessment questions that specifically asked students to explain what genes are and what they do in the body. We coded students' responses using categories of conceptions of genes described by Venville and Treagust (1998). These researchers reported ontological shifts in tenth grade students' conceptions of genes—from a view of genes as passive particles that are inherited and are correlated with traits, to thinking of genes as active particles that determine traits, to thinking of them as sequences of instructions, and finally to viewing them as productive sequences of instructions that code for proteins. We categorized students' responses to the assessment question as suggesting a passive particle, active particle, instructions, or productive instructions view of genes.

We coded statements such as "genes carry our traits" or "genes are our features" as indicating a view of genes as passive particles associated with traits, because there was no mention of information in genes or an account of genes as actively exerting their effects in some way. Statements such as "genes control/determine our traits" or "genes decide who we are and what we will look like" were coded as indicating a view of genes as active particles, because genes were

presented as playing a more active role (analogous to a commander) in determining our traits but not necessarily as having information. Responses such as ''genes tell us what we will look like,'' ''genes are like maps of our traits,'' or ''genes carry information about our traits'' were coded as indicating a view of genes as sequences of instructions, because there was either an explicit mention of informational content (genes say/tell, or have information), or an implicit notion of informational content (maps/blueprints). If a particular response was consistent with more than one category we coded the response according to the most sophisticated view of genes expressed in it. Thus, categories were mutually exclusive and each response received a single code.

We found that none of the responses we coded (in the pre- or postassessments) fell into the category of genes as productive sequences of instructions that code for proteins. Venville and Treagust (1998) noted similar results in their analysis of post instruction worksheets (none of the 79 responses they coded indicated such an understanding). However, our analysis did show that, following instruction, there was a moderate increase in the proportion of responses indicating an instructions view of genes (see Table 1). This suggests that students' conceptions of genes were positively affected by instruction, and that a greater proportion of students demonstrated a more sophisticated view of genes after instruction.

The reader may note that the proportion of students' responses in the ''genes as passive particles'' category remained the same. We believe that this may be somewhat misleading, and that the proportion of students who held a more sophisticated view of genes (active particle or instructions) is, in fact, higher than our analysis indicates. The assessment question asked about two facets of genes: (a) what genes are, and (b) what genes do. Students' responses varied in terms of which part(s) of the question they addressed. This may have generated a bias in our analysis because responses that addressed what genes *are* (e.g., ''genes are DNA'') were categorized as indicative of a ''genes as passive particles'' view. It is conceivable that students who responded in this way held a more sophisticated view of genes, but because they did not articulate this in their response we have no way of knowing. There were slightly more responses of this ilk in the postassessments than in the preassessments. During the instruction students were repeatedly told that the genetic material is DNA, and there was much focus on the structure of DNA. It may be that students felt more confident in their knowledge about what genes are, rather than what they do, and that their responses reflected this. We realize that students' responses to one question may not

Table 1
Students' conceptions of genes

Gene Conception Categories	Examples of Typical Responses	Number of Responses per Category	
		Pre	Post
Passive particle	They are the genetic makeup; carry our unique hereditary traits; they are passed down to us and give us our traits.	15	16
Active particle	They decide/control everything about the body.	33	22
Instructions	They are a map/blueprint of you; it tells whether you have green or blue eyes.	12	24
No response		3	2
Total number of written responses		63	64
Number of student assessments included in the analysis		$N = 63$	$N = 64$

Note: Analytical categories were mutually exclusive; each student response was assigned to one category only.

reveal all they know about genes; we therefore turned to the richer interview data to further analyze students' understandings of the information level.

Understanding genes as containing information is a prerequisite to reasoning about hybrid hierarchical genetic phenomena. However, knowing that genes carry information is only half the battle; one also needs to understand the content of the genetic information—that genes specify the amino acid sequence of proteins. To uncover what students thought about the content of the genetic instructions we analyzed interview episodes (across all tasks) in which students discussed their ideas about genes, DNA, and the genetic information. A few of the interviewed students (four) did not see genes as containing information at all (held a passive particle or active particle view of genes); we therefore omitted these four interviews from the following analysis. We coded the remaining 12 students' ideas in terms of several themes that emerged: (a) the level of organization in which the coded entities exist (e.g., if a student claimed genes code for our eye color we coded this as genes coding at the organ level), (b) whether the information was about entities at one or more than one organization level, and (c) whether the information was about biological structures, biological functions, or both structures and functions.

Our analysis revealed that students conceived of the informational content of genes to be far richer than the scientific understandings of the genetic information. Students' understandings of the genetic information varied in two ways from the canonical notions. First, students described genes as containing information about both structure and function (rather than structure alone). The following statements illustrate such ideas about the genetic information.

> Dick: . . . parts of your genes talk about the channel proteins and tell them what to do, and if they mutate the channel protein won't know what to do.
> John: Well, maybe that code is transferred to determine what brain cells [pause] how you are going to act, your skin cells your eye cells and what color. That code is transferred to those cells and they just do it.

Dick and John claimed that genes have information about the function or behavior of cells, proteins, and even organs. Seven of the 12 students expressed similar notions of genes as coding for the behavior or function of biological entities. We hypothesize that students were not able to fully explain the functions of cells and organs merely in terms of the interacting structures within them, and thus assumed the existence of instructions that direct the behavior of these biological entities. The genetic information can lead to effects that are behavioral but these are indirect (mediated by structures at lower organization levels). For the purpose of generating mechanistic explanations of how genes bring about their effects it is most accurate and productive to think of genes as specifying structures rather than the functions they afford.

The second dimension along which students' understandings of the genetic information differed from scientific ones was in regard to the hierarchical level that the genetic information specified. All 12 students expressed understandings of genes as containing information regarding entities at multiple biological organization levels—cells, tissues, organs, and traits. The following quotes illustrate this trend:

> Mary: The DNA sequence is the structure that codes for what the cell is suppose to be, what the cell's purpose, and what it is suppose to look like.
> Lisa: I mean the DNA, the information that you need, like it would tell you to, Ok, grow hair here or grow a nose or something.
> Amanda: They [genes] make up the chromosomes and the chromosomes tell you what your eye color are going to be and it has all that information on them.

Mary claimed that the information is about cell function and structure, Lisa referred to the information as coding for a whole organ (nose), and Amanda also referred to the information as specifying whole traits. Many of the students did seem to understand that genes code for or have information about the sequence of amino acids that will make up the proteins. However, they did not think of genes as coding *exclusively* for proteins. Students' views of the genetic information allowed for the inappropriate mapping of the information onto both structures and functions of entities at multiple organization levels (rather than only the amino acid sequence of proteins).

Students' understandings of genes reflect inadequate understanding of the information level of the genetic hybrid hierarchical system. Such understandings are problematic because they hinder students' ability to construct causal mechanistic explanations of how the genetic information brings about physical effects (feature or trait). This results in truncated explanations that link the gene directly to the observable outcome without providing the mediating mechanisms. Truncated explanations of this sort were common in the post interview data. The following exchange, which occurred during the third task of the post interview with Anne, illustrates such an explanation. Anne was asked to explain how the gene for the genetic disorder Sickle Cell Anemia (an example of a genetic disorder she suggested) causes the low red blood cell count associated with this disease (a symptom she described).

Anne:	Um, because if you have a low count, if your mom has a low count and if your dad has a low count and both of them are dominant for Sickle Cell, then you should inherit it for sure. And because of them, they are not giving you the correct amount of red blood cell gene, so the gene isn't carrying the right information, or like good information, to a point where you should have an alright blood count.
Interviewer:	OK, so the gene with the information for the blood cells is not there?
Anne:	I mean it is there but, it decides how many like [pause] It is like a little, I mean I am not saying that it is, but I am just saying it is like a little piece of paper that says like ''blue eyes'' or ''blonde hair,'' and it is given to you, that is what you get through the intercourse of your parents.
Interviewer:	So in this case it would say however many red blood cells?
Anne:	Right, or how many to produce.

Anne's explanation directly links the gene (information level) to the physical level of tissue (amount of blood cells), and is missing several key aspects of the causal mechanism at lower physical levels—there is no reference to what is happening at either the protein or the cell level. Anne's understanding of the genetic information results in a tautological explanation that lacks a mechanism; she is unable to explain *how* the gene brings about a lower blood cell count beyond stating the genes tell the body to produce less cells.

Students' ability to provide causal mechanistic explanations of genetic phenomena was also influenced by their knowledge of the molecular entities (proteins) and mechanisms that mediate genetic effects. We next present our analysis of students' understandings of the physical mechanisms that mediate genetic effects; specifically, their understandings of the relationship between genes and proteins, the functions of proteins, and the role of proteins in genetic phenomena.

Research Question 2: Students' Understandings of the Mapping Between the Information and Physical Levels, and the Physical Mechanisms that Mediate Genetic Effects

We first analyzed students' responses to a written assessment question in which they were asked to explain the connection between genes and proteins. We coded the responses to this

question in terms of the different types of relationships students expressed; for example: proteins make genes, both proteins and genes are in our bodies, proteins help genes function, etc. Some responses referred to more than one type of relationship, and were thus coded in multiple categories. Our analysis showed that only a minority (25%) of the responses (even after instruction) indicated an understanding of the relationship as one in which the genes code for proteins. Students' ideas about the gene–protein connection are depicted in Table 2. Most of the ideas expressed in students' responses were either incorrect (genes are made of proteins, proteins are made of genes, proteins store information, etc.), vague (both genes and proteins are important), or unfruitful (both are found in cells). Although the latter two types of responses are not incorrect per se, they are not particularly useful descriptions of the relationship.

The changes in the distribution of students' responses between the pre- and postassessments were rather small, and only a minority of students' responses in the postassessment expressed the scientifically appropriate understanding that genes determine the amino acid sequence of proteins (genes code for proteins). Without the understanding that genes specify protein structure students are less likely to invoke proteins as playing a role in genetically mediated phenomena because they do not realize that genetic phenomena are inevitably mediated by proteins. To examine this issue we turned to the interview data, specifically, the task in which students were asked to speculate about the molecular mechanisms that mediate genetic phenomena.

This task required students to provide a hypothetical mechanism for a molecular genetics phenomenon termed chemotaxis (see Appendix A). Students read a short passage stating that bacterial cells are able to sense and react to substances in their aquatic environment, and that some bacteria are mutated and can no longer sense their environment. We suspected that students would not be familiar with the chemotaxis phenomena or its causal mechanism. Nonetheless, we wanted

Table 2
Students' conceptions of the gene–protein connection

Students Ideas about the Gene–Protein Connection	Number of Responses That Referred to These Ideas	
	Pre	Post
Genes determine the structure/behavior of proteins	6	16
Proteins build/make genes/DNA	3	3
Composition:		
Proteins are made up of genes/	0	1
genes are made of proteins	9	13
Proteins carry/store genetic information	0	2
Proteins transport information (to or from genes)	1	2
Commonalties:		
Common localization: both are in our cells/bodies	5	2
Common purpose: both are important, both help the body, both control our lives	6	5
Common composition: both contain DNA/amino acids	1	2
Both are inherited	1	0
Proteins help genes function/proteins feed genes	6	3
No connection	1	0
Total number of ideas mentioned in the written responses	39	49
No response	28	15
Number of student assessments included in the analysis	$N = 63$	$N = 64$

Note: In the preassessments the total number of ideas referenced in students' responses (39) in addition to the number of nonresponses (28) is 67, which is larger than the overall number of assessments included in the analysis (63) because a few of the students made reference to more than one idea in their response.

to find out what sorts of hypothetical explanations students could provide, and whether these explanations would include proteins as part of the mechanism.

We found that students were able to suggest physical mechanisms to explain the ''sensing'' behavior of the bacteria. The following excerpts illustrate some of the mechanisms students provided in their explanations:

John: . . . or like us, we can smell things, for example [pause] this could be their [pause] maybe they can't exactly smell, but they have some other way of doing it. They may have an organ like cilia, but there is some distinguishing factor about the food or poison. I don't know what it is, but maybe it's like the vibrations or something like that.

Interviewer: Would something from the outside have to get in (to the cell)?

John: No, we have nerve cells so they might have something similar and maybe if something touched the cilia something could transport a message to it. Like how messenger RNA does and kind of like that. It just takes the message to the organelle.

Dick: It might have a little sensor maybe.

Interviewer: How do you think that sensor will work, what is it made of?

Dick: Maybe it is some sort of chemical or material that reacts to food or sugar. Like it makes some sort of reaction with sugar. Maybe it's kind of like a radar and senses food around; that kind of smells sort of, not really smelling, but sensing.

Ron: Probably reactions or senses. Lets say something comes in contact with a cell wall or, for example, or the hair like structure that senses the movement from the outside, or a nerve and it would send a signal like the human brain. Like you know how if you would poke or pinch us we would feel it immediately, instantly. So the message would go to the nucleus and signal the flagella to move.

Many of the students speculated that the bacteria have some sort of sensory structure on the outer membrane that can detect substances in the environment and relay a message to other organelles in the cell (mainly the nucleus). These students seemed to be drawing on knowledge of other detection mechanisms (at the macro level) that they were more familiar with such as radar (Dick), nerve cells (John), hairs that detect vibrations or movement (Ron and John), and taste buds (examples not shown). Very few of these explanations, only 6 out of the 16 students interviewed, referred to proteins as playing a role. Many of the students did understand this to be a genetic phenomenon (from the statement referring to the mutation), yet they did not further reason that a protein must be involved because it was genetic.

We next analyzed students' responses to the written assessment prompt that directly asked them to explain what proteins do in our bodies, as well as the interview episodes in which students discussed their ideas about proteins' biological role and functions in the body. This analysis indicated that students did have a clear sense that proteins were important biological entities (several of the students recalled the teacher noting that ''proteins run our lives''). However, students knew little about the specific roles proteins play in our bodies, and their descriptions of protein functions were broad and vague. The following statements from the postinterviews illustrate students' understandings of proteins and their biological role:

Beth: I know carbs give us energy and proteins I think they [pause] they don't give us energy, they make you strong probably. I don't know.

Lisa: I know they (proteins) are important and they are part of our lives. But I don't know [pause] like I said, they are in our hair, they make our hair grow.

Kate: Well, you need proteins in your blood for nutrition, but you also need it to run your body to also produce enzymes because that enables you to move. Without proteins you would become weak because you are not eating the right amount of proteins and you will stop functioning properly because you will need proteins to help produce the enzymes to keep you moving.

Although none of these ideas are incorrect, they do not refer to specific functions or explain how proteins (or enzymes) make us strong or enable us to move. It seems that students understand that proteins are an important element in the physical level—vital for survival—but are unclear on how proteins promote this vitality.

Analysis of the written assessment data revealed similar trends in students' responses. Table 3 illustrates students' ideas about protein functions as expressed in responses to the assessment question "What are proteins and what do they do?" Some of the students' responses made reference to several ideas, both in terms of what proteins are and in terms of what they do, and were thus coded in multiple categories.

A large proportion of students' written responses referred to proteins' important vitalistic role in maintaining life, but it is not clear that students had any idea of how proteins promote the body's well being. Even responses that further specified the function of proteins, such as regulating cellular function, did not refer to any specific cellular process (such as cellular respiration, a topic the students in this study learned prior to the genetics unit). Students did learn about proteins as macromolecules and various cellular processes in which proteins are involved in the months prior to this unit, and we expected that they would be able to provide more specific functions for proteins.

Table 3
Students' conception of protein functions

Students Ideas about Proteins	Number of Responses That Referred to These Ideas	
	Pre	Post
Proteins are made up of amino acids	10	30
Proteins capture (bind) glucose molecules	0	6
Proteins hold chromatin	0	2
Proteins are enzymes/catalysts	10	11
Proteins gives us energy/proteins are used as nutrients	23	7
Proteins transport things in the cell/body	1	2
Proteins control/regulate functions in cells	1	5
Proteins make up DNA/genes	2	5
Proteins get rid of waste	1	0
Proteins build and repair tissues (muscles and bones)	9	1
Proteins give instructions to body	0	1
Proteins run our lives, important for survival, help the body	15	18
Proteins make up structures in our body (hair, fingernails)	4	9
Total number of ideas mentioned in the written responses	76	97
No response	8	4
Number of student assessments included in the analysis	$N = 63$	$N = 64$

Note: In the pre- and postassessments the total numbers of ideas referenced by students' responses (76 and 97, respectively) are larger than the overall number of assessment included in the analysis (63 and 64, respectively) because some of the students made reference to more than one idea in their responses.

Many of the students also stated that proteins were needed for energy or as nutrients. This statement is correct if students were referring to the breaking down of proteins for energy in extreme starvation, or proteins' role as enzymes in the cellular processes that generate energy (cellular respiration). However, our sense was that students were not necessarily referring to these uses of proteins but rather conceived of proteins as an immediate source of energy. For example, they would note that you need to eat proteins when you work out, and that proteins give you energy (as in commonly sold protein energy drinks). Such responses were more prevalent in the preassessments but were still prominent even after instruction.

Another point to note is that in the postassessments a larger proportion of students' responses referred to proteins as made up of amino acids, and several responses referred to proteins' ability to "capture glucose." We believe that both of these trends are due to a modeling activity in which students were given a list of amino acids and their properties (such as size, charge, and affinity to water) and asked to determine the amino acid sequence of a protein that would be able to transverse the cell membrane and bind to a glucose molecule on the outside. The goal of this activity was to help students construct an understanding of how a protein's structure (amino acid sequence) affects its function (ability to transverse the cell membrane and bind glucose). This activity may have reinforced the idea that proteins are made of amino acids, and that a plausible protein function is capturing glucose.

In summary, to reason about the genetic hybrid hierarchical system one needs to understand how the physical entities and mechanisms (that are part of the physical level) mediate the effects of genes (part of the information level). Our analysis revealed that students in this study did not have a clear understanding of a key physical entity, the protein, and its biological functions in mediating genetic effects (its relationship to genes). This lack of understanding of the centrality of proteins in genetic phenomena was reflected in students' inadequate and truncated explanations of these phenomena. Without an understanding of proteins the link between the genetic information and its physical effects remains a black box. Despite targeted instructional experiences (modeling the processes of the central dogma, and the design-a-protein activity) students did not seem to develop an appreciation for universal connection between genes and proteins and the consequent involvement of proteins in genetic phenomena.

Discussion

In prior research the difficulties in reasoning about genetic phenomena were mainly attributed to two aspects: the need to reason across multiple organization levels from the molecular to the macro scale (Horwitz, 1996; Marbach-Ad & Stavy, 2000), and the inaccessibility of these phenomena, especially at the molecular and micro levels, to our senses (Gilbert et al., 1982; Kapteijn, 1990). In our theoretical framework we highlighted a third aspect that likely contributes to students' difficulties in this domain—genetic phenomena as a hybrid hierarchical system containing ontologically distinct levels—information and physical. Reasoning about this hybrid system entails understanding how the information is mapped onto the physical entities in the system, and how those physical entities mediate the outcomes of the information across organization levels (the molecular mechanisms involved). In this section we discuss our findings as they relate to our theoretical framework, assertions of prior research, and related research about ontological differences involved in conceptual change.

A critical step to reasoning about hybrid hierarchical systems is recognizing that there are two different types of levels in the system to begin with. We found that after instruction, a significant proportion of the students in the study understood that the genetic material contained information (see Table 1). Thus, they did come to perceive genetic phenomena as composed of two distinct levels. However, students' understandings of the content of the information were nonrestricted

and all encompassing, and therefore, they tended to map the information inappropriately onto multiple organization levels. These understandings led to the construction of incomplete causal mechanistic explanation of genetic phenomena that left out important causal events at the molecular and cellular levels.

Findings from research by Lewis and Kattmann (2004) also suggest that mapping between the information and physical levels is conceptually challenging for students. These researchers noted that students' tend not to distinguish between genotype (gene) and phenotype (traits), and as a result, do not consider the microscopic and molecular causal mechanisms of genetic phenomena (Lewis & Kattmann, 2004). Their findings resonate with what we have found in this study, and support the assertion presented in our theoretical framework that students' difficulties are likely due to an inability to appropriately map between the ontologically distinct information and physical levels of the hybrid genetic system. Students in both studies seemed to map the information onto higher organization levels rather than the lower protein level.

Reasoning about hybrid hierarchical systems also entails understanding the physical level and how entities (and mechanisms) within it mediate the effects of the information. We found that many of the students were not familiar with proteins and their central role in mediating genetic effects. Our findings suggest that students' knowledge of physical level features a large gap—a superficial and limited understandings of proteins and their relationship with higher organization levels such as cells, tissues, etc. This is problematic because proteins are at the interface between the information and physical level; misunderstanding the interactions at this interface point impacts the ability to reason about biological events at subsequent levels and to provide a coherent mechanistic explanation of genetic phenomena. We suspect that students' lack of knowledge about proteins is due in part to the limited instructional time devoted to the discussion of proteins' role in mediating genetic phenomena, as well as the limited experience students have with these invisible entities. Prior research has shown that reasoning about imperceptible entities, such as proteins and genes, is difficult for students (Gilbert et al., 1982; Kapteijn, 1990). We believe that the lack of perceptual experience only partly accounts for the difficulties students experience because students had less trouble reasoning about the structure and function of DNA (also an imperceptible entity) compared to proteins.

We conjecture that there is somewhat of a bootstrapping issue underlying students' difficulties in reasoning about the genetic hybrid hierarchical system. Due to their broad conceptions of the nature of the genetic information, students are not constrained to presume that proteins must be involved because they do not conceive of genes as coding only for proteins. Conversely, because students are not aware of the central roles proteins play in biological processes, they do not presume that any biological phenomena, that has genetic origins, is likely mediated by proteins. Thus, students have a poor understanding of the content of the information level on the one hand, and a lack of knowledge of the mechanisms (and entities) in the physical level that bring about the effects of the information on the other. This limits their ability to understand and reason about the genetic system as translating information programs into physical phenomena.

Before we turn to a discussion of the implications of our work we wish to discuss how our work intersects with prior research on conceptual change across ontological categories (Chi, 2005; Chi, Slotta, & de Leeuw, 1994; Ferrari & Chi, 1998; Slotta & Chi, 1996; Slotta, Chi, & Joram, 1995) on the conceptual changes involved in understand. Chi et al. (1994) argues that conceptual change in physics (understanding phenomena like heat transfer, diffusion, current) involves an ontological shift in the classification of phenomena from one ontological kind (things) to another (processes). A similar conceptual change seems to occur with students' understandings of genes (shift from ''passive particle'' to ''productive sequence of instructions'' view of genes) as described by Venville & Treagust (1998). We are arguing that in addition to the conceptual change

involved in coming to understand genes as instructions that code for proteins, students need to understand how proteins function to bring about the effects of the genetic information. Thus, conceptual change in genetics likely involves not only an ontological shift in concept categorization but also the development of cognitive structures that facilitate the mapping between the ontologically distinct levels in genetics. Such cognitive structures might embody understandings about the physical entities that mediate the effects of the genetic information.

Implications

We discuss three instructional implications that follow from our findings. First, our findings suggest that traditional instruction in this domain, while helping students understand that genes are essentially informational (sequences of instructions), does not engender an understanding of the content of that information. Students in this study spent a considerable amount of time modeling the translation of DNA sequences into sequences of amino acids, yet they did not conclude from these experiences that genes code solely for the building blocks of proteins. There seems to be an instructional assumption that by focusing on the details of the processes involved in translating the genetic code and procedurally practicing such decoding students will develop the general understanding that essentially all genes code only for proteins. However, the students in this study did not seem to connect these discrete activities to the overall process by which genes bring about their effects. It is therefore important to help students see the connection between genes and proteins as a general principle in the domain to help students construct more generalized and generative understandings of the genetic information.

Second, we also believe that there should be more instructional emphasis on the functions of proteins and their role in mediating genetic effects. The instruction we observed did not focus on the specific functions proteins have and their role in genetic phenomena. This resulted in students understanding that proteins were very important elements in the body, yet lacking functional knowledge of how those elements bring about their beneficial effects. Proteins are a central component of biological systems at large; they are involved in all biological processes normal and pathological (and often serve as the target of pharmaceutical drugs). However, proteins are entirely short-changed in the high school biology curriculum (and textbooks); proteins are introduced to students mainly as a type of macromolecule that is consumed (part of the food pyramid) with a focus on the molecular structure of the amino acids. Their role in the body is relegated to four main aspects: enzymes, and the building of muscle, nails, and hair. Although these are important functions of proteins, they are the very tip of the iceberg and disregard other central roles of these molecules (such as signaling, regulation, transport etc).

If we want students to productively reason about molecular genetics phenomena, they need to become much more familiar with these important molecules. A recent analysis (AAAS, 2005; Kurth & Roseman, 2001) of the ways in which major biology textbooks address important ideas in genetics as defined by the AAAS benchmarks (AAAS, 1993) showed that although textbooks do provide some examples of protein function (mainly as enzymes) they do not really make salient the generalization that proteins do almost all the work of the cell. Moreover, many of the books discussed the DNA directed synthesis of proteins but did not clearly link this process to the role proteins have in mediating genetic effects.

Third, dovetailing with our previous implication we suggest that instruction in molecular genetics needs to shift in focus from the procedural to the conceptual. In the classroom we studied, much of the instructional time was devoted detailed descriptions of the structure of DNA and procedural decoding of the genetic code. This focus did not engender an understanding of the underlying mechanisms and key entities in genetic phenomena. Although we have no evidence

that the instruction we observed is representative of the instructional practices of other teachers in other schools, we can assert that this instruction followed the district and state content standards and thus reflects those goals.

It is clear that understanding the big ideas in genetics is the intent of the standards and benchmarks for scientific literacy (AAAS, 1993; NRC, 1996), and the teacher we worked with understood these important ideas and wanted her students to learn them. However, the activities used in the instruction did not foster such learning, nor it is clear that the students in her classroom understood these learning goals. Further research is needed to examine teachers' instructional moves and selection of activities, as well as their students' understanding of the intended purpose of the activities they engage in.

Shifting instruction to emphasize the big ideas and conceptual underpinnings of explanations in the domain of molecular genetics is not a simple task. Instruction needs to engage students with genetic phenomena in their entirety and not just the processes involved in the central dogma. This instruction should be based on a cognitive model of reasoning in the domain. In future work we plan to extend our understanding of student reasoning in molecular genetics by developing such a cognitive model that can inform the design of more effective instruction.

Appendix A

Interview Protocol

Task I. Have students read the following passage about bacteria:

> *Version A*: Most bacterial cells can sense and respond to substances in their surroundings. For example if you put bacteria near food substances (sugar) they will sense the food and move towards the food; if you put them near poisonous substances they will sense the poison and move farther away from it. Some bacteria have a mutation and can no longer sense substances in their environment. These bacteria do not move towards food or away from poisons and therefore are more likely to die from starvation or poisoning.
>
> *Version B:* Many types of bacterial cells live in watery environments and most of them can swim and move around in the water. These bacterial cells can move because they have hairBlike structures protruding from the cell. These ''hairs'' can rotate and propel the bacteria forward. Some bacterial cells have a mutation and can't swim or move around. These mutated bacterial just stay still even if they are in danger, and therefore are more likely to die or get damaged.

Ask students to tell you what the passage was about in their own words. Ask follow-up questions:

- How do you think ''normal'' bacteria sense stuff/move in their surroundings? What is going on inside the bacteria cell?
- What does ''mutation'' mean?
- How do you think a mutation can cause a bacteria to lose its ability to sense substances/ move in its surroundings?
- Do you think there could possibly be a mutation that made bacteria more sensitive to substances/ move better? How would that work?

Task II. Give students the following nine statements. Ask them to build a detailed story about the genetic disease using the given statements, tell them that they do not have to use all statements and that they may find some that are not relevant to the story they want to tell.

Once students have complied their story have them explain the story to you in their own words. Then go over the story with them, statement by statement, to determine what students understood from each statement and how each statement relates to the rest of the story.

Version A:

1. Cystic fibrosis is a lung disease. The lungs of people with CF get clogged up with thick and sticky mucous which makes it hard for CF patients to breathe (picture)
2. CF patients have a mutation in one of their genes.
3. The channel protein in CF patients (is mutated it) has a different amino acid sequence in comparison to the channel protein of normal people.
4. There are many different types of channel proteins in the body; different channel proteins allow different substances to flow through them.
5. In healthy people the normal channel protein is found in the cell membranes, and it functions as a channel through which water can pass from inside the cell to the lung cavity. Water in the lung cavity dilutes the mucous inside the lungs and makes it less thick.
6. Both normal people and CF patients have 46 chromosomes that make up 23 pairs.
7. The amount of mucous in the lungs does not affect how much blood gets into the lungs or how fast it flows through.
8. In CF patients the abnormal channel protein cannot get into the membrane, and therefore cannot function as a channel.
9. Without the channel protein water can't get from the cell to the lung cavity.

Version B:

1. Patients with Hyprecholesterolemia (FHC) have too much cholesterol in their blood. Too much cholesterol in the blood can clog up and block blood vessels and this may lead to a heart attack. These patients tend to die at a young age from heart attacks.
2. The cholesterol-binding protein of FHC patients has a different amino acid sequence in comparison to the cholesterol-binding protein of normal people.
3. FHC patients have a mutation in one of their genes.
4. In healthy people the normal cholesterol-binding protein is found in the cell membrane. This protein binds to cholesterol in the blood (which is on the outside of the cell) and brings it into the cell. This is the way in which most of the cholesterol in the blood gets into cells.
5. In FHC patients the abnormal cholesterol-binding protein cannot get into the membrane, and therefore, cannot bind cholesterol.
6. Without the cholesterol-binding protein, cholesterol cannot get from the blood into the cell.
7. There are many different types of binding proteins in the body; different binding proteins bind to different substances.
8. The amount of cholesterol in the blood does not affect the amount of nutrients the blood can carry.
9. Both normal people and FHC patients have 46 chromosomes that make up 23 pairs.

Task III. Ask students to pick a genetic disease they know a bit about.

Ask students the following questions about the genetic disease they chose:

1. What are the symptoms of this disease?
2. Why do we say that it is genetic?
3. How is it caused? (press students to give detailed explanations). If students mentions genes ask:
   - How can a gene cause this disease?
   - Do normal people have this gene?
   - How is this different from being sick with Anthrax or the Flu?

Appendix B

Written Assessment

1. What are proteins and what do they do in the body?
2. What are genes and what do they do in the body?
3. Proteins are made of:
   a) Carbohydrates
   b) Lipids (fat molecules)
   c) Amino acids
   d) Other substances: _____
   e) A combination of:_____
4. Proteins are found:
   a) In all the cells and fluids of the body.
   b) In some of the cells, but not in all of them.
   c) Mainly in the blood and some in the cells.
   d) None of the above, they are found in:_____
5. Proteins are:
   a) Bigger than cells.
   b) Bigger than the nucleus.
   c) Bigger than genes
   d) Bigger than amino acids
   e) Other: _____
6. What is the connection between proteins and genes?
7. Genes are made of:
   a) Proteins.
   b) DNA
   c) Amino acids
   d) Carbohydrates
   e) Chromosomes
   d) Cells
   f) Other: _____
8. What is a genetic disease? Give an example if you know one.
9. Is there a difference between a genetic disease and a disease like the flu? If yes, How are they different?

This research was conducted as part of the KDI/ASSESS project. The authors gratefully acknowledge Bruce Sherin and Daniel Edelson, School of Education and Social Policy at Northwestern University, for their insightful comments and valuable discussions.

References

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

American Association for the Advancement of Science. (2005). High school biology textbooks: A Benchmarks-based evaluation. New York: Oxford University Press.

Bahar, M., Johnstone, A.H., & Sutcliffe, R.G. (1999). Investigation of students' cognitive structure in elementary genetics through word association tests. Journal of Biological Education, 33, 134–141.

Casti, J.L. (1994). Complexification: Explaining a paradoxical world through the science of surprise. New York: Harper Collins.

Chi, M.T.H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. Journal of the learning sciences, 6, 271–315.

Chi, M.T.H. (2005). Commonsense conceptions of emergent processes: Why some misconceptions are robust. Journal of the Learning Sciences, 14, 161–199.

Chi, M.T.H., Slotta, J.D., & de Leeuw, N. (1994). From things to processes: A theory of conceptual change for learning science concepts. Learning and Instruction, 4, 27–43.

Crick, F. (1970). Central dogma of molecular biology. Nature, 227, 561–563.

Dreyfus, A., & Jungwirth, E. (1990). Macro and micro about the living cell: Which explains what? In P.L. Lijnse, P. Licht, W. de Vos, & A.J. Vaarlo (Eds.), Relating macroscopic phenomena to microscopic particles (pp. 108–118). Utrecht, The Netherlands: CD-Press.

Ferrari, M., & Chi, M.T.H. (1998). The nature of naïve explanations of natural selection. International Journal of Science Education, 20, 1231–1256.

Fisher, K.M. (1985). A misconception in biology: Amino acids and translation. Journal of Research in Science Teaching, 22, 53–62.

Friedrichsen, P.M., & Stone, B. (2004, April). Examining students' conceptions of molecular genetics in an introductory biology course for non-science majors. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, Vancouver, BC, Canada.

Garton, G.L. (1992). Teaching genetics in the high school classroom. In M.U. Smith & P.E. Simmons (Eds.), Teaching genetics: Recommendations and research proceedings of a national conference (pp. 20–23). Cambridge, MA: National Science Foundation.

Gilbert, J.K., Osborne, R. J., & Fensham, P. J. (1982). Children's science and its consequences for teaching. Science Education, 66, 623–633.

Ginsberg, H.P. (1997). Entering the child's mind: The clinical interview in psychological research and practice. New York: Cambridge University Press.

Hill, R., O'Sullivan, H., Stanisstreet, M., & Boyes, E. (1998). Biology students' understanding of cystic fibrosis, gene therapy and gene screening. Journal of Biological Education, 32, 103–110.

Hmelo, C.E., Holton, D.L., & Kolodner, J.L. (2000). Designing to learn about complex systems. The Journal of the Learning Sciences, 9, 247–298.

Horwitz, P. (1996). Teaching science at multiple space time scales. Communications of the ACM, 39, 100–102.

Kapteijn, M. (1990). The function of organizational levels in biology for describing and planning biology education. In P.L. Lijnse, P. Licht, W. de Vos, & A.J. Vaarlo (Eds.), Relating macroscopic phenomena to microscopic particles (pp. 139–150). Utrecht, The Netherlands: CD-Press.

Kindfield, A.C.H. (1994). Understanding a basic biological process: Expert and novice models of meiosis. Science Education, 78, 255–283.

Kurth, L.A., & Roseman, J. (2001, April). 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., & Kattmann, U. 2004. Traits, genes, particles and information: Revisiting students' understanding of genetics. The International Journal of Science Education, 26, 195–206.

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

Marbach-Ad, G. (2001). Attempting to break the code in student comprehension of genetic concepts. Journal of Biological Education, 35, 183–189.

Marbach-Ad, G., & Stavy, R. (2000). Students cellular and molecular explanations of genetic phenomena. Journal of Biological Education, 34, 200–210.

McLaren, J.E., Rotundo, L., & Gurley-Dilger, L. (1991). Heath biology. Lexington, MA: D.C. Heath and Company.

Miles, M.B., & Huberman, A.M. (1994). Qualitative data analysis: An expanded sourcebook (2nd ed.). Thousand Oaks, CA: Sage Publications.

Miller, K.R., & Levine, J. (1998). Biology (4th ed.). Upper Saddle River, NJ: Prentice-Hall.

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

Penner, D.E. (2000). Explaining processes: Investigating middle school students' understanding of emergent phenomena. Journal of Research in Science Teaching, 37, 784–806.

Resnick, M. (1996). Beyond the centralized mindset. Journal of the Learning Sciences, 5, 1–22.

Rotbain, Y., Marbach-Ad, G., & Stavy, R. (2006). Effect of bead and illustration models on high school students' achievement in molecular genetics. Journal of Research in Science Teaching, 43, 500–529.

Simon, H.A. (1996). The sciences of the artificial (3rd ed.). Cambridge, MA: MIT Press.

Slotta, J.D., & Chi, M.T.H. (1996). Understanding constraint-based processes: A precursor to conceptual change in physics. In G.W. Cottrell (Ed.), Proceedings of the eighteenth annual conference of the cognitive science society (pp. 306–311). Mahwah, NJ: Erlbaum.

Slotta, J.D., Chi, M.T.H., & Joram, E. (1995). Assessing students' misclassifications of physics concepts: An ontological basis for conceptual change. Cognition and Instruction, 13, 373–400.

Stewart, J. (1990). Biology education research: A view from the field. In M. Gardner, J.G. Greeno, F. Reif, A.H. Schoenfeld, A. diSessa, & E. Stage (Eds.), Toward a scientific practice of science education (pp. 55–68). Hillsdale, NJ: Erlbaum.

Stewart, J., & Van Kirk, J. (1990). Understanding and problem-solving in classical genetics. International Journal of Science Education, 12, 575–588.

Venville, G.J., & Treagust, D.F. (1998). Exploring conceptual change in genetics using a multidimensional interpretive framework. Journal of Research in Science Teaching, 35, 1031–1055.

Wilensky, U., & Resnick, M. (1999). Thinking in levels: A dynamic systems approach to making sense of the world. Journal of Science Education and Technology, 8, 3–19.

Wynne, C.F., Stewart, J., & Passmore, C. (2001). High school students' use of meiosis when solving genetics problems. International Journal of Science Education, 23, 501–515.