Chapter 5
How Can We Help Students Reason About the Mechanisms by Which Genes Affect Traits?

Michal Haskel-Ittah

5.1 The Problem

We have entered an era in which genetic tools are constantly being developed and used by researchers and industry. In this era, every individual can obtain information about his own DNA sequences and consume genetically engineered products. This rapid invasion of genetic issues into our everyday lives has not escaped the media's attention, where these issues are frequently raised, albeit not always in a scientifically correct manner (Stern & Kampourakis, 2017). Hence, today, knowledge about genetics is essential for the ability to make informed everyday decisions and to engage in discussions about ethical issues and the health benefits of applying these genetic tools. Mere familiarity with genetic concepts is not enough. A deeper understanding of genetic mechanisms, the limits of our understanding of them and the boundaries of their predictive power is required (Boerwinkel et al., 2017; Gelbart, 2012; Stern & Kampourakis, 2017).

Almost 20 years ago, it was already being acknowledged that education that wishes to prepare the public for the genetic revolution should itself undergo a change in the oversimplified description of genetics in school (McInerney, 2002). Haga (2006) suggested that educational resources in genetics must be continuously updated and revised to narrow the gap between concurrent scientific findings and public knowledge. The purpose of narrowing this gap is not merely to be up-to-date with current scientific knowledge, but also to comprehend how scientists' understanding of genetics may affect citizens' everyday lives.

For the last 20 years, the understanding of the mechanisms by which genes affect traits has been continually growing. Scientists have revealed the complexity of these mechanisms which involve multiple genes, regulatory pathways, and interactions with the environment. These mechanisms are the basis for technologies and genetic

M. Haskel-Ittah (✉)
Department of Science Teaching, Weizmann Institute of Science, Rehovot, Israel
e-mail: Michal.Haskel@weizmann.ac.il

M. Haskel-Ittah, A. Yarden (eds.), *Genetics Education*, Contributions from Biology Education Research, https://doi.org/10.1007/978-3-030-86051-6_5

tools, and they provide the limits for what can and cannot be learned from the DNA sequence. However, when examining how the public today acknowledges the complexity of the relationship between genes and traits, many people seem to hold an oversimplified conception of this relationship (Chapman et al., 2019; Gericke et al., 2017). For example, people suggest that complex traits such as autism are the product of a single genetic variant, and that a person's behavior can be predicted from looking at their DNA sequence (Chapman et al., 2019).

If our aim is to help students reason about gene-to-trait mechanisms to support everyday decision making and discussion, we should assist them in understanding these mechanisms and their complexities. This is also important for the goal of developing a mechanistic understanding in general, an issue that has been acknowledged for its contribution to critical thinking (Ahn & Kalish, 2000; Koslowski, 1996), as well as to a deep understanding of science (Chin & Brown, 2000; Grove et al., 2012). The question is, how can we support students' mechanistic understanding and mechanistic reasoning in these genetic contexts? To answer this question, we must review what we know about teaching genetics from a mechanistic perspective.

5.2 Current Knowledge About the Problem

All natural phenomena can potentially be explained by underlying mechanisms. Such mechanisms consist of entities with specific properties that enable them to conduct certain activities (Craver & Darden, 2013; Machamer et al., 2000). These activities cause changes which eventually lead to the observed phenomena (Craver, 2001; Darden, 2008). Hence, to understand a mechanism and to reason about similar mechanisms, one should be able to identify and characterize the entities involved in the mechanism, and recognize the activities and functions executed by those entities (Russ et al., 2008). Those entities and interactions should be at least one organizational level lower than the phenomenon itself, and they should be able to explain how interactions between entities at this level lead to the final outcome—the phenomenon (Krist et al., 2018).

In biology, mechanisms typically extend across multiple levels (Craver & Darden, 2013). In the case of the effect of genes on traits, the characterization of the phenomenon (gene to trait) suggests that multiple levels should be addressed in the mechanism: from the molecular gene level to the organismal trait level. In multicellular organisms, this requires a deep understanding of the central entities that act at each level (e.g. Haskel-Ittah & Yarden, 2017; Marbach-Ad & Stavy, 2000; van Mil et al., 2013). However, in the case of genetics teaching, the representation and description of the entity—the gene—is not consistent among textbooks, making it hard for students to thoroughly understand its properties (Gericke et al., 2014; Livni-Alcasid et al., 2018). For example, an analysis of eight textbooks from three different countries revealed inconsistencies in allelic symbols throughout all of them with no explicit explanation for the symbol alterations (Livni-Alcasid et al.,

2018). Gericke and Hagberg (2007) identified five historical models that included different descriptions of the concept of gene. In each model, the properties and function of genes differed. For example, in one model, the gene's function was to hold information for traits (Mendelian model) and in another, it was to manufacture specific proteins (neoclassical model) or code for RNA (modern model). In one model, the gene was an abstract idea with no structure or composition (Mendelian model) and in another, it had the property of being located at a specific locus on the chromosome (classical model). An analysis of science textbooks from six different countries revealed the sporadic appearance of the five different models among textbooks (Gericke et al., 2014). From a mechanistic perspective, this is highly problematic because an entity's description should include its properties and function. When genes are described inconsistently in each historical model, with different properties and functions, it may be confusing in terms of which properties and activities should be used for the purpose of mechanistic reasoning. Such a description may lead to confusion regarding genes as entities in the mechanism.

Proteins, which are central entities in genetic mechanisms, are also difficult for students to understand and use in genetic explanations (e.g. Duncan & Tseng, 2011; Haskel-Ittah & Yarden, 2017; Marbach-Ad & Stavy, 2000). Teachers often refer to proteins as merely gene products without further denoting their active involvement in the mechanism (Thörne & Gericke, 2014). Aside from the inaccuracy of describing proteins as if they were side products of genetic mechanisms, this description offers students a fragmented mechanism that does not explain most aspects of genetics. These problems might explain why students often regard genetic phenomena as cause-and-effect relationships between genes and traits, and not as the result of a complex causal mechanism (Duncan & Tseng, 2011; Haskel-Ittah & Yarden, 2017; Haskel-Ittah et al., 2019).

No less important than these central entities—genes and proteins—is their connection. A comparison between how biology and chemistry textbooks describe the connection between genes and proteins revealed that in the former, the mechanistic connection is not emphasized (Wahlberg & Gericke, 2018). Wahlberg and Gericke (2018) characterized the description of protein synthesis in biology textbooks as more focused around the concepts in isolation, rather than on how their interactions lead to a final outcome. Although characterizing the entities involved in a mechanism is an important part of mechanistic reasoning, it is not enough to provide a mechanistic explanation. With no description of the dynamics between entities and how this leads to the emergence of a certain phenomenon, there is no connection between set up and final conditions, and therefore, no actual mechanism (Krist et al., 2018; Russ et al., 2008).

Another problem emerges from the erroneous descriptions of the relationships between genes and traits: the complexity of the gene-to-trait mechanism is often ignored in the classroom, suggesting that there is a one-to-one relationship between genes and traits. A study with preservice biology teachers showed that they tend to ignore the complexities of these mechanisms. These teachers' explanations about the origin of traits in individuals mostly referred to genes as causes, but not to a process or mechanism (Kampourakis et al., 2016). Moreover, textbooks often

provide examples of classical genetic disorders or traits and rarely deal with multifactorial disorders and traits, leading to the misconception that multifactorial traits are the exception rather than the rule (Dougherty, 2009; Hicks et al., 2014). Such examples outline a simple one-to-one relationship that will, later on, be discordant with knowledge about other components in the mechanism (e.g. proteins). This again may lead students to abandon a mechanistic explanation in favor of a simple cause-and-effect one.

While not prevalent in high-school education, it seems that at a young age, children can already think of genetic phenomena as the result of a mechanism. Although preschoolers may provide inconsistent and sometimes teleological explanations for kinship (Williams, 2012), when they are asked to choose between a mechanistic explanation which involves physical entities and a teleological one that involves intentions, they display a preference for the mechanistic one (Ergazaki et al., 2014; Springer & Keil, 1991). A study conducted with high-school students showed that a mechanistic explanation for genetic phenomena, but not memorization of genetic concepts, was retained in the students' memory even 18 months after instruction (Todd & Romine, 2018). In addition, studies show that causal knowledge plays a major role in the ability to provide a mechanistic explanation. In the absence of causal knowledge, other intuitive explanations (e.g., teleological) emerge (Kampourakis & Zogza, 2009; Trommler et al., 2018). This implies that at least some of the difficulties encountered by students in reasoning about genetics are not because they do not recognize the superiority of a causal mechanism over other explanations in this context, but because they lack the domain-specific knowledge to build one. Indeed, domain-specific knowledge has been shown to play a major role in students' ability to reason in genetics (Duncan, 2007).

5.2.1 A Possible Solution

If we acknowledge that the ability to reason mechanistically in genetics is based on conceptual and mechanistic knowledge in genetics, then conceptual knowledge and mechanistic understanding in genetics should be intertwined and mutually developed. By reviewing some of the literature in the field of genetics education, I came up with three milestones in the development toward a mechanistic understanding of the effect of genes on traits. These three milestones represent levels in the progression from a more naïve phenomenal description of the relationship between genes and traits to a complex understanding of the entities and activities that are involved in these mechanisms.

I. **Establishing a Correct Causal Connection between Genes and Traits.** The first step in the search for mechanisms is a description of the target phenomenon. This description outlines the borders for the space of the mechanistic explanations (Craver & Darden, 2013; Darden, 2008). In the gene-to-trait mechanism, this space is defined between genes and traits. Thus, the first milestone in under-

standing genetics (prior to learning about the mechanism) is an acknowledgment of this space.

Studies in genetics education have mapped numerous alternative conceptions of genetic concepts. Among others, they found a misconception regarding the relationship between genes and traits: the notion that genes are trait-bearing particles or genes and traits are synonyms (Lewis & Kattmann, 2004; Venville & Treagust, 1998). This alternative conception is an incorrect description of the phenomenon of causal relationship between genes and traits. More than a decade later, we decided to explore both the prevalence of this alternative conception, and its effect on mechanistic understanding in genetics. In our study (Haskel-Ittah & Yarden, 2018), we used concept maps to explore the manner in which students describe the relationship between genes and traits. We found that this alternative conception was prevalent among 9th graders. Namely, 28% of the 152 students who participated in the study described genes as containing traits or as being synonyms for traits. Thus, almost a third of the students described genetic phenomena as resulting from a direct connection, which might not provide a space for a mediating mechanism.

According to the knowledge-integration perspective, the identification of a gap in one's knowledge is a crucial step toward constructing connections between new knowledge and prior knowledge (Linn et al., 2004). Describing genes as affecting traits suggests causal relationships with a mediating gap that may be further explained by searching for the mediating mechanisms. In contrast, suggesting that genes are traits or trait-bearing particles means that the gene–trait relationship need not to be explained by a mechanism because there is no gap to be filled. Knowledge about entities and activities in the mechanism would thus be fruitful, and can be linked to prior knowledge in the first case, but is irrelevant for the latter. Indeed, we found that when students described a causal relationship between genes and traits prior to instruction, they used the taught gene-to-trait mechanism as the missing link between genes and traits. These students' knowledge of the involvement of proteins improved and they were successful in reasoning mechanistically about new genetic phenomena. On the other hand, students who did not describe the gene–trait relationship as causal prior to instruction were significantly less successful in learning about proteins and in reasoning about the mechanism.

As a phenomenal description, "genes affect traits" is correct, and seems to be an important stage prior to learning about gene-to-trait mechanisms. Nonetheless, this description holds the risk of imposing a gene-centered deterministic view, i.e., that genes are the only entities determining traits. Such a view can create an educational problem because it may hinder the learning of other ideas in genetics (Jiménez-Aleixandre, 2014), and it can pose a social problem because it may lead to racism and prejudice (Donovan, 2016; Keller, 2005). Thus, teachers and educators should be aware of this risk and possible solutions for avoiding it (Stern & Kampourakis, 2017).

One suggested solution for the problem of genetic determinism is presenting multifactorial genetic traits before the description of Mendelian traits (Dougherty, 2009). Another suggestion is to focus on the developmental aspect of the mutual

effect of genes and environment on traits (Jamieson & Radick, 2017). My suggestion is to implement these two ideas at the first stage of the phenomenon description suggested here. For example, students can sort effectors on traits and conclude that genes affect traits—sometimes this involves more than one gene and sometimes genes are not the only effectors. This, of course, does not provide the mechanisms for this multifactorial effect, but it describes a multifactorial phenomenon in a manner that sets the placeholders for a mechanism (see Fig. 5.1A).

II. **Establishing the Understanding of Genes–Proteins–Traits Mechanisms.** Following the description of genetic phenomena, the mechanistic space can be filled with entities and activities. This might include entities such as RNA-polymerase, mRNA, ribosome, amino acids, various proteins, etc. As more entities and functions are incorporated into the mechanism, the cognitive load in learning this mechanism increases. Thus, in most cases, not all entities and functions will be presented. An important question is what are the key components for allowing mechanistic understanding and mechanistic reasoning in genetics? To answer this question, more empirical studies are needed. Nevertheless, in this chapter, I describe the importance of focusing on entities that are largely overlooked by teachers: the proteins (Thörne & Gericke, 2014).

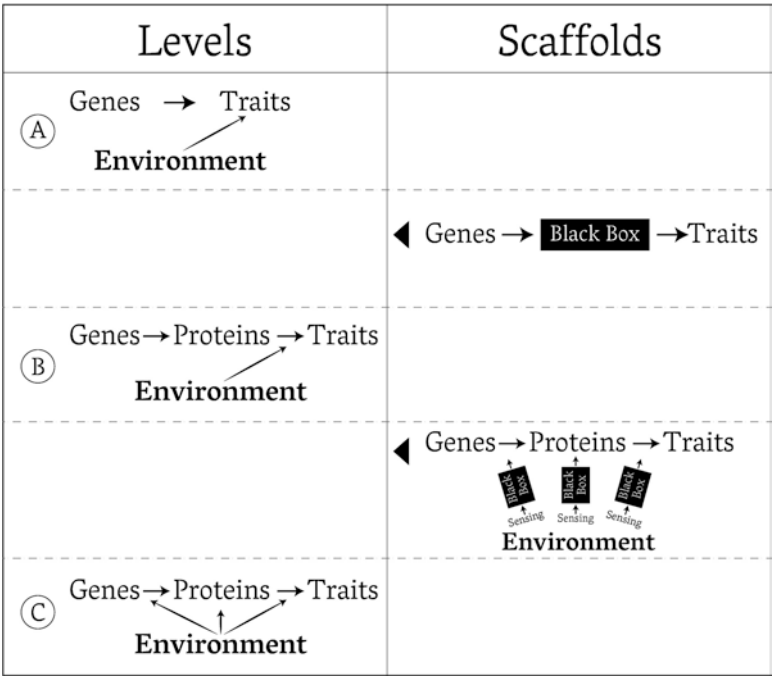

Fig. 5.1 A progression scale toward mechanistic understanding in genetics. Three milestones (A, B, C) in the development of mechanistic understanding in genetics are shown on the left, and suggested scaffolds for facilitating the progression along the scale are shown on the right

Several studies have reported that instructions that focus on protein function lead to gains in the mechanistic understanding of genetics (Freidenreich et al., 2011; van Mil et al., 2013). For example, Duncan and Tseng (2011) developed an instructional module in which 10th-grade students explored multiple phenomena involving proteins with different functions (e.g., enzymes, transporters, and channels). Following instruction, students proposed proteins as key players in their accounts of molecular genetics phenomena and were able to describe the roles of those proteins in the suggested mechanism. Van mil (2016) pointed out that learning about three general protein activities (colliding, binding and changing shape) can bridge an explanatory gap between the molecular and cellular levels. An implementation of this approach with high-school students showed improvement of students' reasoning about cell biology. In a study conducted with 7th graders, we examined how students who had learned about different protein functions (e.g., receptor, enzyme) reasoned using those functions (Haskel-Ittah et al., 2019). Our findings revealed that although the students acknowledged that genes code for proteins, their use of proteins when asked to explain the result of a genetic mutation on a specific trait was inconsistent: they used proteins to explain the given genetic phenomenon in cases that could be explained by a protein function they had learned. However, when the case could not be explained by a previously learned function, students turned back to merely describing a causal relationship between genes and traits without specifying a mechanism.

Altogether, these studies suggest that students are more likely to establish the idea of proteins as mediators between genes and traits via learning about the involvement of proteins as functioning entities and their contribution to trait formation, rather than via learning about how the DNA code is transcribed and translated into proteins. This notion is in accordance with the mechanistic perspective, whereby the presence of an entity in an explanation is important if it changes the predicted behavior of the mechanism (Craver & Darden, 2013). A change in a gene sequence may result in a change in the predicted trait if the protein's function, quantity or stability is affected by the sequence change. The RNA is also changed in this scenario, but it is impossible to explain the trait change only by the RNA change. In the case of the protein, we may describe the gene as coding for a protein whose function was changed without mentioning the RNA. Hence, proteins are central for explaining genetic phenomena.

This does not mean that translation and transcription mechanisms are unimportant, but it implies that these mechanisms may be less fruitful for the purpose of allowing students to reason about the effect of genes on traits using proteins. These mechanisms may be added later to provide a more detailed linkage between genes and proteins (see Fig. 5.1B).

III. Identifying Regulation Points and Understanding How Environmental Signals May Lead to the Modulation of Gene-to-Trait Mechanisms. The final milestone toward a mechanistic understanding in genetics includes a more complex understanding of biological mechanisms that interact with, affect and regulate genetic mechanisms. Genetic mechanisms are subject to multilevel regulation,

> from regulation of gene expression to regulation of RNA translation, protein activity and its stability. Such regulation is the outcome of mechanisms that involve multiple inputs from the internal and external environments, which are sensed and transmitted as signals that affect one or more regulation points. Understanding that gene-to-trait mechanisms are regulated by other mechanisms is important for understanding the complexity of genetics and to demarcate the limits of genetically based predictions.

In a Delphi-like study conducted to ascertain the knowledge needed for genetics literacy in the twenty-first century, the interaction between environmental factors and genetic mechanisms was considered one of the core ideas in modern genetics (Boerwinkel et al., 2017). However, in that study, similar to recommendations in other studies and reports, the interaction between genetic mechanisms and the environment was only vaguely described. This vague description included mentioning that it exists or suggesting that the environment may affect gene expression, while overlooking all other regulation points, such as protein stability (Boerwinkel et al., 2017; Dougherty, 2009; Duncan et al., 2009; National Research Council, 2012).

The mechanism by which a change in the environment may result in a change in traits is also regarded. For example, in the genetics learning progression (LP; Duncan et al., 2009), the environment is first described as influencing our traits; at higher levels, it is described as influencing cell function, and then as causing mutations or altering gene expression. There are no descriptions of any mechanisms that might link the environment to the alteration in the gene's expression. A disregard for the question of how environment might affect genetic mechanisms is also apparent in textbooks. An analysis of five Spanish textbooks showed that, although four of them defined phenotype as a result of gene–environment interactions, the texts provided only a few examples of the mechanisms for environmental influences (Puig & Jiménez-Aleixandre, 2011).

Overall, such a description provides limited knowledge about possible regulation points (only the point of gene expression) and fragmented knowledge about the mechanisms by which environmental signals can lead to the modulation of traits (only an input–output description of the environment as influencing traits). This may form a serious obstacle to understanding the effects of the internal and external environment on the regulation of gene-to-trait mechanisms.

Puig and Jiménez-Aleixandre (2011) analyzed 10th-grade students' explanations for gene–environment interactions when involved in a group discussion about a physical trait. Those students studied with different teachers about how the environment may affect genetic traits. Four out of five groups suggested that either the environment has no effect at all, or that genes and environment affect the phenotype separately, with the former having a greater influence. The two groups that suggested genes as the only influential factor studied with the same teacher, and the authors suggested that they had not had the opportunity to develop their ideas regarding possible mechanisms of gene–environment interactions in class. This indicates the importance of providing at least a general description of the possible mechanisms for gene–environment interactions.

In a study conducted with undergraduate students (Haskel-Ittah, Duncan and Yarden (2020), we provided students with four phenomena in which an environmental change led to a change in trait (e.g., skin color change as a result of UV light, low growth rate as a result of undernutrition). We interviewed these students and asked them to reason about these cases. We found that students with presumably less knowledge about biological mechanisms (most 1st-year students and some 2nd-year students) suggested that the environment affects the trait independently from the genetic mechanism or by causing mutations. Students with a higher level of knowledge were able to explain how sensing the environment might activate signals that can modulate genetic mechanisms via regulation of gene expression or other regulatory points. In addition, we found that knowledge about mechanisms of gene expression, where they were not connected to environmental signals, did not assist students in reasoning about gene–environment interactions.

From these studies, we can conclude that it is not enough to mention that the environment might affect genetic mechanisms. Without establishing a mechanistic connection between the environment and regulation of genetic mechanisms, the possibility of an environmental effect on genetic traits is not considered. For this reason, knowledge about gene regulation that is not described as the output of regulatory signals may not be fruitful in terms of reasoning about genetic phenomena that involve gene–environment interactions. Accordingly, students should first identify possible points of regulation in gene-to-trait mechanisms and acknowledge that they include mechanisms of sensing the environment, which can result in activating/inhibiting these regulation points. Only then can specific mechanisms be taught as an example of regulatory mechanisms, such as regulation of gene expression (e.g., transcription factors, epigenetic modifications) or protein modifications (e.g., ubiquitination, phosphorylation) (see Fig. 5.1C).

How to scaffold movement between the milestones of mechanistic understanding in genetics? Moving between the three suggested milestones of mechanistic understanding might be challenging for students. In this section, I suggest theoretical scaffolds for moving along this progressive scale of mechanistic understanding in genetics. These scaffolds are based on a concept that has been characterized and studied in computer science education: the black box. The term "black box" refers to knowledge which is unknown to the person receiving the explanation and/or to the one who is giving the explanation (Haberman et al., 2002). Thinking in terms of black boxes includes both the process of removing details with the aim of simplification, and the process of identifying the core essence with the aim of generalization (Kramer, 2007). Hence, thinking in terms of black boxes in explanations enables thinking in terms of conceptual ideas, rather than details (Armoni, 2009).

Black boxes are also central to mechanistic explanations in biology, because "every description of a mechanism bottoms out at some point where the gain in detail makes no difference to the researcher" (Craver & Darden, 2013, p. 90). For example, one can provide an explanation involving an enzyme catalyzing a reaction without explaining precisely how the chemical interaction between the enzyme and its substrate leads to the catalysis. In other cases, black boxes may exist simply because their contents, namely the mechanisms, are not yet understood by scientists.

Drawing on its use in computer science education, teaching using black boxes does not mean skipping over or ignoring parts of a mechanism, but characterizing them. This characterization should include an input/output description without any details on its internal function (Statter & Armoni, 2017). Such a description focuses on the functionality of the black box and its contribution to the mechanism. For example, we can describe a "regulation" black box in the gene-to-trait mechanism. The input to this black box is signals from the internal or external environment, and the outputs are inhibitors/activators of the gene-to-trait mechanism. A different "regulation" box may act on different sections of the mechanism (i.e., gene expression, translation or protein function). In each section, the internal mechanism is different but the conceptual idea of modulating the mechanism, as a result of environmental signals, is the same.

Black boxes are effective scaffolds in computer science education for moving toward a deeper conceptual understanding (Ben-David Kolikant & Haberman, 2001). This is because by using black boxes in teaching, novices can first understand the functionality of a process and its contribution to the whole mechanism, before coping with the complicated details of the process itself. In the field of biology, a study was conducted with middle-school students who learned biochemistry by first describing a black box and then learning about the mechanism inside the black box. The authors reported that these students asked deeper questions and acquired a more thorough understanding than the control group (Olsher & Dreyfus, 1999). These studies are in line with several related findings in science teaching: first, that centering on the functionality of processes prior to teaching the process itself can establish a deeper understanding of the latter (Liu & Hmelo-Silver, 2009); second, that describing a process as a black box is, in essence, acknowledging a knowledge gap. Such acknowledgement has been shown to be an important step for learning how to fill this gap (Linn et al., 2004). Lastly, it has been claimed that after learning a mechanism and forgetting its details, the relearning of the mechanism is much faster due to "cognitive traces" in the shape of multiple black boxes between central entities (Keil, 2019). Building on this claim, we can assume that having these black boxes prior to actually learning about the details of the mechanism may facilitate learning.

If we wish to use these black-box scaffolds, we should first identify them along the progression. The first milestone in the aforedescribed progression of mechanistic understanding in genetics is acknowledging that there is a causal connection between genes and traits and that this causal connection may be detailed, meaning acknowledging that there is a black box between genes and traits (Fig. 5.1, right). Genes–black box–traits is not a mechanistic explanation but it is a phenomenal model that describes an indirect relationship mediated by a mechanism. One can use this phenomenal model to make predictions (such as in Mendelian genetics) while still recognizing that this is not a full explanation. Hence, the black box may also act as a placeholder and scaffold the progress toward a more detailed mechanistic explanation. In addition, understanding the limits of reasoning using this black box is important in terms of recognizing the limits of this phenomenal model and avoiding conflation with other, more detailed models, as described by Gericke and

Hagberg (2010). For the same reason, acknowledging that there is a black box that links genes to traits might also inhibit views of genetic determinism.

The second milestone in the progression is opening the black box and revealing the mechanism that mediates the causal relationship, i.e., unpacking the mediating entities and activities. As previously mentioned, not all entities and activities should be unpacked. The use of black boxes at this stage may assist in building a whole mechanism without the risk of forming an erroneous direct connection between indirect entities. For example, moving forward from genes–black box–traits to genes–black box–proteins–traits may allow focusing on the central role of proteins without risking the formation of a direct link between genes and proteins, and without the need for a detailed explanation of transcription and translation. In this example, the input for the black box may be a DNA sequence and the output a specific protein structure.

The move from the second mechanistic milestone to a more complex understanding of regulatory mechanisms may be facilitated by introducing a "regulatory black box"—a recognition that sensing mechanisms can regulate each step in the gene-to-trait mechanism without providing the details of these regulatory mechanisms. This should provide the ability to think about regulation by the internal/external environment in terms of conceptual ideas, without the need for a detailed explanation (Fig. 5.1, right). Later on in the progression, since the placeholders already exist, students may more easily understand the details of a specific regulatory mechanism, such as regulation of gene expression, including transcription factors or chromatin remodeling.

Use of these black boxes may facilitate the move between milestones in the progression scale of mechanistic reasoning in genetics, and they may act as placeholders, inhibiting conceptions of genetic determinism.

5.3 Remaining Issues

The black box scaffolds suggested in this chapter are based on findings and theories from science education but are nonetheless theoretical. Understanding the affordances and constraints of the use of black boxes as scaffolds in genetics teaching warrants further empirical testing. Such testing should first analyze how students grasp the gap between the input and output (the black box) and how we should help them understand this gap. This is because if students do not perceive this input–output as a gap, or they perceive it as a gap that cannot be explained (even in the future), this black box cannot act as a scaffold. One possible indication that students do perceive this black box as one that can be opened is their ability to ask questions about how it may function or raise assumptions regarding its possible content (similar to what was done in Olsher & Dreyfus, 1999). These questions and assumptions lay the groundwork for understanding what knowledge is needed to open the black box. Research should also characterize the difficulties students may encounter when reasoning in biology using a black box. Some of these difficulties have been characterized in the field of computer science, but not in biology (Statter & Armoni, 2017).

5.4 Implications for Teaching

The current LP in genetics is wide, targeting more than gene-to-trait mechanisms. It includes many other important ideas, such as mechanisms and probabilities in sexual reproduction (Duncan et al., 2009). This LP was empirically tested and revised accordingly (Shea & Duncan, 2013; Todd & Kenyon, 2015). In the revised version, the LP includes several constructs that are relevant to the discussion in this chapter (see Table 5.1): genes code for proteins (construct B), proteins do the work of cells (construct C1), proteins as mediators between genes and traits (construct C2), and environmental factors interact with genetic mechanisms (construct H).

This description of the constructs from the genetic LP was adapted from Todd and Romine (2018, p. 7). Each construct (B, C1, C2, H) is a description of a "big idea" in genetics and the levels correspond to the levels of understanding this idea. The progression occurs along these levels. Although construct B deals with the connection between genes and proteins, the first progression level describes genes as instructions for the development of traits. This is essentially a black box description which scaffolds the movement from students' prior knowledge about a link between genes and traits toward a more mechanistic description of how genes affect the appearance and development of an organism. The higher levels focus on the mechanisms by which genes code for proteins and do not deal with how proteins affect

Table 5.1 Condensed description of several constructs from the genetic LP which are discussed in the implications

Construct	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
B—Genes code for proteins	Genes noninformational	Genes are instructions (at the body level)	Genes code for cell entities	Genes code for proteins	Genes translated into proteins	----
C1—Proteins do work of cell	Cells perform function	Proteins are good for your body	Proteins do work of cell	Protein function depends on structure	Protein function and structure depends on amino acids in the protein	----
C2—Proteins connect genes and traits	Change to genes changes traits	Change to genes changes cells	Change to genes changes proteins	Change to genes changes proteins to change traits	Change to genes changes amino acids in proteins	Change to genes changes protein function to change traits
H—The environment interacts with genetic information	Environment can affect organisms	Environment can affect traits or functions	Environment can affect our cells, organs or tissues	Environment can change or mutate things inside cells	Environment can change type and amount of proteins that influence cell function	Environment can change genes which change proteins or gene expression

traits, meaning that only the black box between genes and proteins is opened (explaining how genes code for proteins), but not the one between proteins and traits (no explanation of how protein function affects traits). Construct C2, which deals with proteins as mediators between genes and traits, also begins by establishing a causal relationship between a change in genes and a change in traits. It is only at higher levels that the construct further describes the change in genes as affecting protein structure and consequently, its function. Again, a black box is established between genes and traits, then proteins are described as mediators and only the black box between genes and proteins (and not proteins and traits) is opened at the higher levels.

The construct that opens the black box between proteins and traits is construct C1, dealing with how protein function may lead to traits. As already suggested, the centrality of proteins as mediators arrives mainly via an understanding of their contribution to the formation of traits (namely their function) and less via knowledge about genes as coding for proteins. Thus, construct C1 is essential for gaining an understanding of constructs B and C2. In addition, the relevance of construct C1 to genetics is understood only when higher levels are reached in constructs B and C2. And indeed, such contingencies were empirically detected (Shea & Duncan, 2013). Hence it is recommended not to teach all constructs in parallel but to begin with lower levels in B and C2 which describe the black box between genes and traits (genes affect traits via a mechanism that is not taught for now) and adding C1 before moving to higher levels in these constructs. In addition, the connection between these constructs should be made explicit.

Another idea mentioned in the LP, which is relevant for the discussion here, is the construct dealing with environmental factors that interact with genetic mechanisms (construct H). According to the recommendations here, this construct should be elaborated to include the idea of sensing the environment and should be tightly connected to regulation mechanisms, such as regulation of gene expression (which is suggested in a completely different construct).

In conclusion, this chapter identifies three milestones along the development of mechanistic understanding and mechanistic reasoning in genetics, and suggests black boxes which may be used as scaffolds for supporting this development. These black boxes seem to already exist, in some parts of the current genetics LP, but not explicitly or consistently.

References

Ahn, W.-K., & Kalish, C. W. (2000). The role of mechanism beliefs in causal reasoning. In F. C. Keil & R. A. Wilson (Eds.), *Explanation and cognition* (pp. 199–225). The MIT Press.

Armoni, M. (2009). Reduction in CS: A (mostly) quantitative analysis of reductive solutions to algorithmic problems. *Journal on Educational Resources in Computing, 8*(4), 11-1–11-30.

Ben-David Kolikant, Y., & Haberman, B. (2001). *Activating "black boxes" instead of opening "zipper": A method of teaching novices.* Paper presented at the ITiCSE'01: Proceedings of the Sixth Annual Conference on Innovation and Technology in Computer Science Education. https://doi.org/10.1145/377435.377464.

Boerwinkel, D. J., Yarden, A., & Waarlo, A. J. (2017). Reaching a consensus on the definition of genetic literacy that is required from a twenty-first-century citizen. *Science & Education, 26*(10), 1087–1114.

Chapman, R., Likhanov, M., Selita, F., Zakharov, I., Smith-Woolley, E., & Kovas, Y. (2019). New literacy challenge for the twenty-first century: Genetic knowledge is poor even among well educated. *Journal of Community Genetics, 10*(1), 73–84.

Chin, C., & Brown, D. E. (2000). Learning in science: A comparison of deep and surface approaches. *Journal of Research in Science Teaching, 37*(2), 109–138.

Craver, C. F. (2001). Role functions, mechanisms, and hierarchy. *Philosophy of Science, 68*(1), 53–74.

Craver, C. F., & Darden, L. (2013). *In search of mechanisms: Discoveries across the life sciences*. University of Chicago Press.

Darden, L. (2008). Thinking again about biological mechanisms. *Philosophy of Science, 75*(5), 958–969.

Donovan, B. M. (2016). Framing the genetics curriculum for social justice: An experimental exploration of how the biology curriculum influences beliefs about racial difference. *Science Education, 100*(3), 586–616.

Dougherty, M. J. (2009). Closing the gap: Inverting the genetics curriculum to ensure an informed public. *The American Journal of Human Genetics, 85*(1), 6–12.

Duncan, R. G. (2007). The role of domain-specific knowledge in generative reasoning about complicated multileveled phenomena. *Cognition and Instruction, 25*(4), 271–336.

Duncan, R. G., & Tseng, K. A. (2011). Designing project-based instruction to foster generative and mechanistic understandings in genetics. *Science Education, 95*(1), 21–56.

Duncan, R. G., Rogat, A. D., & Yarden, A. (2009). A learning progression for deepening students' understandings of modern genetics across the 5th–10th grades. *Journal of Research in Science Teaching, 46*(6), 655–674.

Ergazaki, M., Alexaki, A., Papadopoulou, C., & Kalpakiori, M. (2014). Young children's reasoning about physical & behavioural family resemblance: Is there a place for a precursor model of inheritance? *Science & Education, 23*(2), 303–323.

Freidenreich, H. B., Duncan, R. G., & Shea, N. (2011). Exploring middle school students' understanding of three conceptual models in genetics. *International Journal of Science Education, 33*(17), 2323–2349.

Gelbart, M. E. (2012). Catching education up with technology: Preparing the public to make informed choices about personal genetics. *CBE Life Sciences Education, 11*(1), 1–2.

Gericke, N., & Hagberg, M. (2007). Definition of historical models of gene function and their relation to students' understanding of genetics. *Science & Education, 16*(7–8), 849–881.

Gericke, N., & Hagberg, M. (2010). Conceptual incoherence as a result of the use of multiple historical models in school textbooks. *Research in Science Education, 40*(4), 605–623.

Gericke, N., Hagberg, M., dos Santos, V. C., Joaquim, L. M., & El-Hani, C. N. (2014). Conceptual variation or incoherence? Textbook discourse on genes in six countries. *Science & Education, 23*(2), 381–416.

Gericke, N., Carver, R., Castéra, J., Evangelista, N. A. M., Marre, C. C., & El-Hani, C. N. (2017). Exploring relationships among belief in genetic determinism, genetics knowledge, and social factors. *Science & Education, 26*(10), 1223–1259.

Grove, N. P., Cooper, M. M., & Cox, E. L. (2012). Does mechanistic thinking improve student success in organic chemistry? *Journal of Chemical Education, 89*(7), 850–853.

Haberman, B., Shapiro, E., & Scherz, Z. (2002). Are black boxes transparent?—High school students' strategies of using abstract data types. *Journal of Educational Computing Research, 27*(4), 411–436.

Haga, S. B. (2006). Teaching resources for genetics. *Nature Reviews Genetics, 7*(3), 223–229.

Haskel-Ittah, M., & Yarden, A. (2017). Toward bridging the mechanistic gap between genes and traits by emphasizing the role of proteins in a computational environment. *Science & Education, 26*(10), 1143–1160.

Haskel-Ittah, M., & Yarden, A. (2018). Students' conception of genetic phenomena and its effect on their ability to understand the underlying mechanism. *CBE Life Sciences Education, 17*(3). https://doi.org/10.1187/cbe.18-01-0014

Haskel-Ittah, M., Duncan, R. G., Yarden, A., & Gouvea, J. (2020). Students' understanding of the dynamic nature of genetics: Characterizing undergraduates' explanations for interaction between genetics and environment. *CBE—Life Sciences Education, 19*(3), ar37. https://doi.org/10.1187/cbe.19-11-0221

Haskel-Ittah, M., Duncan, R. G., Vázquez-Ben, L., & Yarden, A. (2019). Reasoning about genetic mechanisms: Affordances and constraints for learning. *Journal of Research in Science Teaching, 57*(3), 342–367.

Hicks, M. A., Cline, R. J., & Trepanier, A. M. (2014). Reaching future scientists, consumers, & citizens: What do secondary school textbooks say about genomics & its impact on health? *The American Biology Teacher, 76*(6), 379–383.

Jamieson, A., & Radick, G. (2017). Genetic determinism in the genetics curriculum. *Science & Education, 26*(10), 1261–1290.

Jiménez-Aleixandre, M. P. (2014). Determinism and underdetermination in genetics: Implications for students' engagement in argumentation and epistemic practices. *Science & Education, 23*(2), 465–484.

Kampourakis, K., & Zogza, V. (2009). Preliminary evolutionary explanations: A basic framework for conceptual change and explanatory coherence in evolution. *Science & Education, 18*(10), 1313–1340.

Kampourakis, K., Silveira, P., & Strasser, B. J. (2016). How do preservice biology teachers explain the origin of biological traits?: A philosophical analysis. *Science Education, 100*(6), 1124–1149.

Keil, F. (2019). How do partial understandings work? In S. R. Grimm (Ed.), *Varieties of understanding: New perspectives from philosophy, psychology, and theology* (pp. 191–208). Oxford University Press.

Keller, J. (2005). In genes we trust: The biological component of psychological essentialism and its relationship to mechanisms of motivated social cognition. *Journal of Personality and Social Psychology, 88*(4), 686–702.

Koslowski, B. (1996). *Theory and evidence: The development of scientific reasoning*. MIT Press.

Kramer, J. (2007). Is abstraction the key to computing? *Communications of the ACM, 50*(4), 36–42.

Krist, C., Schwarz, C. V., & Reiser, B. J. (2018). Identifying essential epistemic heuristics for guiding mechanistic reasoning in science learning. *Journal of the Learning Sciences, 28*(2), 160–205.

Lewis, J., & Kattmann, U. (2004). Traits, genes, particles and information: Re-visiting students' understandings of genetics. *International Journal of Science Education, 26*(2), 195–206.

Linn, M. C., Eylon, B.-S., & Davis, E. A. (2004). The knowledge integration perspective on learning. In M. C. Linn, E. A. Davis, & P. Bell (Eds.), *Internet environments for science education* (pp. 29–46). Lawrence Erlbaum Associates.

Liu, L., & Hmelo-Silver, C. E. (2009). Promoting complex systems learning through the use of conceptual representations in hypermedia. *Journal of Research in Science Teaching, 46*(9), 1023–1040.

Livni-Alcasid, G., Haskel-Ittah, M., & Yarden, A. (2018). As symbol as that: Inconsistencies in symbol systems of alleles in textbooks, and students' justifications for them. *Education in Science, 8*(3). https://doi.org/10.3390/educsci8030110

Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science, 67*(1), 1–25.

Marbach-Ad, G., & Stavy, R. (2000). Students' cellular and molecular explanations of genetic phenomena. *Journal of Biological Education, 34*(4), 200–205.

McInerney, J. D. (2002). Education in a genomic world. *The Journal of Medicine and Philosophy, 27*(3), 369–390.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. The National Academies Press.

Olsher, G., & Dreyfus, A. (1999). Biotechnologies as a context for enhancing junior high-school students' ability to ask meaningful questions about abstract biological processes. *International Journal of Science Education, 21*(2), 137–153.

Puig, B., & Jiménez-Aleixandre, M. P. (2011). Different music to the same score: Teaching about genes, environment, and human performances. In T. D. Sadler (Ed.), *Socio-scientific issues in the classroom* (pp. 201–238). Springer.

Russ, R. S., Scherr, R. E., Hammer, D., & Mikeska, J. (2008). Recognizing mechanistic reasoning in student scientific inquiry: A framework for discourse analysis developed from philosophy of science. *Science Education, 92*(3), 499–525.

Shea, N. A., & Duncan, R. G. (2013). From theory to data: The process of refining learning progressions. *Journal of the Learning Sciences, 22*(1), 7–32.

Springer, K., & Keil, F. C. (1991). Early differentiation of causal mechanisms appropriate to biological and nonbiological kinds. *Child Development, 62*(4), 767–781.

Statter, D., & Armoni, M. (2017). *Learning abstraction in computer science: A gender perspective.* Paper presented at the Proceedings of the 12th Workshop on Primary and Secondary Computing Education. https://doi.org/10.1145/3137065.3137081.

Stern, F., & Kampourakis, K. (2017). Teaching for genetics literacy in the post-genomic era. *Studies in Science Education, 53*(2), 193–225.

Thörne, K., & Gericke, N. (2014). Teaching genetics in secondary classrooms: A linguistic analysis of teachers' talk about proteins. *Research in Science Education, 44*(1), 81–108.

Todd, A., & Kenyon, L. (2015). Empirical refinements of a molecular genetics learning progression: The molecular constructs. *Journal of Research in Science Teaching, 53*(9), 1385–1418.

Todd, A., & Romine, W. (2018). The learning loss effect in genetics: What ideas do students retain or lose after instruction? *CBE Life Sciences Education, 17*(4). https://doi.org/10.1187/cbe.16-10-0310

Trommler, F., Gresch, H., & Hammann, M. (2018). Students' reasons for preferring teleological explanations. *International Journal of Science Education, 40*(2), 159–187.

van Mil, M. H. W., Boerwinkel, D. J., & Waarlo, A. J. (2013). Modelling molecular mechanisms: A framework of scientific reasoning to construct molecular-level explanations for cellular behaviour. *Science & Education, 22*(1), 93–118.

van Mil, M. H., Postma, P. A., Boerwinkel, D. J., Klaassen, K., & Waarlo, A. J. (2016). Molecular mechanistic reasoning: Toward bridging the gap between the molecular and cellular levels in life science education. *Science Education, 100*(3), 517–585.

Venville, G. J., & Treagust, D. F. (1998). Exploring conceptual change in genetics using a multidimensional interpretive framework. *Journal of Research in Science Teaching, 35*(9), 1031–1055.

Wahlberg, S. J., & Gericke, N. M. (2018). Conceptual demography in upper secondary chemistry and biology textbooks' descriptions of protein synthesis: A matter of context? *CBE Life Sciences Education, 17*(3). https://doi.org/10.1187/cbe.17-12-0274

Williams, J. M. (2012). Children and adolescents' understandings of family resemblance: A study of naïve inheritance concepts. *British Journal of Developmental Psychology, 30*(2), 225–252.