International Society of the Learning Sciences

Cross-Sectional Study of Students' Molecular Explanations of Inheritance Patterns

Moraima Castro-Faix, Rutgers University, moraimac@rutgers.edu
Ravit Golan Duncan, Rutgers University, ravit.duncan@gse.rutgers.edu

Abstract: Genetics is an important topic in the biology curriculum of many countries. Learning genetics is difficult because it involves the need to reason about complex ideas and to link between mechanisms that occur at different levels of organization. Previous research in genetics education has studied either students' understanding of inheritance patterns or their understandings of molecular genetics; very few studies have examined students' understanding of the connection between these ideas. To address this gap, we conducted a cross-sectional interview study with students from middle school to high school, as well as undergraduate and graduate students, after they had experienced "status-quo" instruction in genetics. We analyzed student responses and described the molecular mechanistic explanations students used to explain inheritance patterns. Our findings allowed us to propose a progression describing the gradual use of genetics ideas. We discuss the implications of the tentative progression for instruction aimed at supporting reasoning in genetics.

Introduction

Discoveries originating from genetics research are increasingly becoming part of our daily lives. These include the discovery of gene editing, the availability of genetic testing (e.g., 23andMe), GMO-based products in grocery stores, increased prenatal testing, and the use of novel therapies for cancer treatment (Boerwinkel, Yarden & Waarlo, 2017; Gericke & Smith, 2014). For instance, the company 23andMe provides genetic testing to the public without consultation with a genetic counselor to help the consumer interpret the results of the tests (23andMe, 2020). This may be problematic for individuals who may not have a full grasp of genetics or have difficulties understanding the probability of developing a genetic disorder. These interpretations of genetic screens may cause anxiety in the consumer or possibly influence important medical decisions, like deciding whether to have breast cancer screenings or prenatal testing (Bellcross, Page, & Meaney-Delman, 2012; Hawkins & Ho, 2012). Additionally, the discovery of CRISPR by Emmanuelle Charpentier and Jennifer Doudna, and their recent winning of the Nobel Prize, has led to important ethical discussions about the role these discoveries have in society and the possibility of genetically modifying human embryos to prevent inherited diseases (Weisbverg, Badagio & Chatterjee, 2017).

Yet, high school graduates and college students often have difficulties comprehending these scientific advances, which can make it difficult to engage with the civic, personal and ethical issues they may encounter stemming from genetics advances (Stern & Kampurorakis, 2017). Stewart, Cartier and Passmore (2005) proposed that genetics literacy consists of understanding three models of genetics. These models include: (a) the inheritance model, which explains patterns of inheritance (e.g. recessive, dominant, sex-linked) and the understanding of how alleles (gene variants) are transmitted across generations; (b) the meiotic model, which involves the physical transfer of genes from one generation to the next through sex cells; and (c) the molecular model, which includes the cellular and molecular mechanisms by which genes bring about traits. This model involves the understanding that genes code for proteins and that changes (mutations) in the recipe (DNA) may lead to changes in the structure and function of proteins. Moreover, genetics literacy involves the linkages among these models along with understanding the role of the environment (Duncan, Rogat & Yarden, 2009; Stewart et al., 2005).

To help students become literate in genetics, it is important to facilitate student understanding of the connection between molecular and the inheritance models as making connections can help students understand the underlying molecular mechanisms involved in the expression of traits (Haskel-Ittah & Yarden, 2019; Marbach-Ad & Stavy 2000). For example, in muscular dystrophy there are two main types of mutations in the dystrophin protein a complete loss-of- function of the dystrophin protein or the production of a misshapen protein that is still partially functional. Duchenne muscular dystrophy (DMD) is the most common type of this disorder and consists of a loss-of-function mutation in the dystrophin gene (Falzarano et al., 2015). The second most common type for this disorder is Becker muscular dystrophy. This type of muscular dystrophy has a milder effect than Duchenne because in Becker's muscular dystrophy the dystrophin protein has structural problems that cause it to have an altered form (Falzarano et al., 2015). However, the protein can partially function to allow for the formation of muscle cells. These disorders may originate through mutations in early development or be inherited.

Regardless of the type of inheritance the problem with the muscle cells are caused by mutations in the genes that code for proteins involved in making muscle cells. Understanding how it is possible that a disorder like muscular dystrophy can have different forms, and origins involves understanding the link between classical and molecular genetics. However, the linking of these models is one of the areas of most difficulty for students because it involves understanding the mechanisms that occur at different levels of organization consisting of understanding what happens at the molecular, cellular and organismal level (Duncan, 2007; Todd & Romine, 2017). Few studies have examined how students form these connections and none have taken a cross-sectional approach (Todd & Romine, 2017; Wolyak, 2013).Therefore, we know little about how such understandings develop over the secondary and undergraduate levels. To address this gap, we conducted a cross-sectional interview study with students from middle school and high school, as well as undergraduate and graduate students, after they had experienced "status-quo" instruction in genetics. Our aims for this work were to (a) characterize the kinds of explanations students use to provide the link between inheritance patterns (recessive or dominant) with the underlying molecular mechanisms, and (b) characterize how these explanations change over the course of schooling.

We developed qualitative analyses of cross-sectional data with the purpose of identifying and characterizing the ways the students in each grade developed molecular explanations of inheritance patterns. The following research questions guided our analyses: 1) What kinds of molecular explanations can students provide for inheritance patterns? , and 2) How does the use of these explanations change over the course of schooling? We next describe the theoretical frameworks that informed our research questions.

Theoretical framework: Framework of genetics literacy and a normative model of genetics

To characterize the explanations that students used, we drew on the framework of genetics literacy by Stewart, Cartier and Passmore (2005) to identify core genetics ideas, and on a normative model of genetics that consisted of Muller's 1932 definition of mutation as the *loss or gain* of function of a protein (Wilkie, 1994). This definition is still widely used by scientists and textbooks to describe the effects of mutation in protein function (Griffiths et al., 2005; Zhou et al, 2020). Together these two frameworks allowed us to identify and characterize the ideas students were using to reason about molecular mechanisms involved in the expression of phenotypes, such as how a gene for albinism brings about observed physical differences in skin and hair color.

The framework of genetics literacy allowed us to classify and identify the ideas that belonged to the inheritance and the molecular model (Stewart et al., 2005). For example, ideas around inheritance patterns and the connection genes and proteins and we then analyzed the ways the students developed molecular explanations for inheritance patterns.

The normative model of genetics allowed is to identify and classify the mechanisms that the students developed to reason about inheritance patterns. Muller's 1932 model describes how scientists reason about mutations and their effects in protein function. Mutations can be classified as *loss of function* or *gain of function* (Muller, 1932). In *loss of function* mutations there may be either a complete loss of the protein or a reduction of the protein function, which can result in disease. For instance, in sickle-cell anemia the hemoglobin protein is affected by a mutation that causes a change in structure (Pauling, Itano, Singer & Wells, 1949). Normally, hemoglobin is a protein that transports oxygen that is located on the cytoplasm of red-blood cells. In the mutated form, the hemoglobin has a different shape and is less efficient at transporting oxygen. As a consequence, the lack of oxygen causes the red blood cell to become "sickled". This causes the cells to stick together and have difficulties moving through the blood vessels (Hopkins, 2020; Pauling et al., 1949). Loss of function mutation is usually the mechanism involved in recessive disorders and in some dominant disorders depending on how the mutation is affecting the individual. In contrast, gain of function mutations can lead to an increase in protein expression (more protein is being produced than what is needed), a protein that interferes with the function of another or the gain of a new function. For example, in Hungtinton's disease there is a mutation in the sequence of DNA that produces a protein that is much longer than normal (Bates, 2005). This protein is then cut into fragments ("toxic proteins") that bind to each other and accumulate in neurons, affecting the normal functions of the cells. Gain of function mutations are more often related to dominant disorders (Jones & Hughes, 2011). Therefore, the normative model allowed us to identify if students were able to develop loss and gain of function explanations of inheritance patterns. Our analysis characterizes the different ways students reasoned about classical genetics and molecular genetics and about problems that involved integrating knowledge from both models. We describe a progression of mechanism exhibited across several grades (middle school to graduate students). Findings from

this study provide evidence about how students make links among the genetics models and contributes to the limited research base on genetics literacy.

Methods

Study context and participants

The aim of this study was to understand the kinds of explanations that students provided to connect the classical and molecular models of inheritance. We gathered cross-sectional interview data from students at different levels and expertise in genetics. We interviewed students from the following groups: 15 students in middle school (MS), 81 11th-grade biology students, 15 high school biology advance placement students (AP), 31 undergraduates and 12 graduate students that had completed coursework in genetics. Prior to the interview we made sure that the students in each group completed status-quo instruction in genetics and participated in the genetics units or courses that were appropriate to their grade level. For instance, the high school students completed lessons on the structure and function of proteins and their relationship to genes, and they also learned about inheritance patterns. The high school (AP and 11th grade students) and middle school students that participated were from two districts. District 1 consisted of a diverse suburban school with the following demographics: 40% African American, 30% Hispanic, 17% Asian and 13% Caucasian with 50% of the students eligible for free or reduced-fee lunch. District 2 consisted of a suburban school with the following demographics: 55% Caucasian, 35% Asian, 4% African American and 6% Hispanic with 15% of the students eligible for free or reduced-fee lunch. The university students (graduate and undergraduate students) were from a large diverse north-eastern university with the following demographic: 39.8% White, 23.3% Asian, 12.2% Hispanic or Latino, 7.82% Black or African American, 2.79% Two or More Races, 0.212% Native Hawaiian or Other Pacific Islanders, and 0.0605% American Indian or Alaska Native.

Data collection

We collected individual interview data from each student. The interviews lasted 45-60 minutes and consisted of open-ended tasks modified from a previous studies (Castro-Faix, Duncan & Choi, 2020; Duncan, Choi, Castro-Faix & Cavera, 2017). In the first task, the students were asked to reason about the molecular mechanisms for the inheritance of a hypothetical recessive disorder that we called blood-clotting disorder (BCD). The students were given a pedigree that showed the recessive trait "skipping" a generation and they were asked about the genes and traits of th0e individuals shown in the pedigree. The goal of this first task was to evaluate student understanding of inheritance patterns, the role of sex cells in inheritance, and the molecular basis of genetic traits. In the second task, the students were asked to reason about two hypothetical disorders, one dominant and one recessive. In the case of the dominant disorder, the patient is heterozygous for the disorder (Bb). The individual inherited a 'b' copy that codes for the "normal" protein and a 'B' mutated allele that codes for a mutated protein that does not work properly. In the case of the heterozygous recessive disorder (Aa), the individual inherited an 'A' copy that codes for the "normal" protein and an 'a' allele that codes for a mutated protein that does not work properly. We then asked students to explain how a heterozygous genotype can result in a healthy individual (Aa – recessive disorder) and a sick individual (Bb – dominant disorder). The purpose of this task was to elicit molecular explanations for the disease and to figure out how they would explain each disorder. In the third task, students were given 16 cards listing genetics terms like: cell, alleles, chromosomes, DNA, proteins, recessive, *etc*. The students were then asked to arrange the cards in a concept web and draw arrows that linked the terms (as many as they could). The concept web was used to encourage the participants to organize their ideas and explain the terms and the relationships they noted among the terms (Butler-Kisber & Poldma, 2010). In this task, our goal was to emphasize the relationship between alleles and proteins, and to observe how the students elaborated on this relationship.

Data Analysis

Each task was then analyzed qualitatively to determine what kinds of molecular explanations students provided for the inheritance patterns of trait; that is, all three tasks were analyzed using a directed content analysis approach with the purpose of identifying and examining how students integrated ideas from the molecular and inheritance model of genetics ideas (Chi, 1997; Hsieh & Shannon, 2005; Mayring, 2000) and used these ideas to develop molecular mechanistic explanations of inheritance patterns. According to Hsieh and Shannon (2005), the goal of this approach is to inform an existing conceptual framework or theory. The existing theory focuses the research questions that are going to be studied and provides direction about the variables to be studied (Hseih & Shannon, 2005). We used Stewart et al.'s (2005) model of genetics literacy to focus our research questions. The framework provided the initial coding scheme and the relationships between the codes that were used to identify the molecular

mechanisms that students used. Moreover, we aimed to characterize their ideas using a normative model of genetics that we previously described (Muller, 1934). This model allowed us to classify students responses into gain and loss of function mutations and to figure out if the students were using the molecular mechanisms previously discussed.

Results

Our results consist of two main findings: (a) we identified several mechanisms students used to connect the molecular and inheritance models, and (b) we characterized a progression (based on cross-sectional data) that describes the kinds of mechanisms students provided to connect the inheritance and molecular models.

What kinds of explanations do students provide to connect these two models?

We first analyzed the molecular mechanisms students were using to explain inheritance patterns. Table 1 shows the mechanisms students used to explain dominant and recessive disorders. The first column illustrates the molecular mechanisms we identified and how they map onto the normative model of genetics. The second column provides an example of students' statements for each context (e.g. dominant or recessive). Our findings indicate that students across all grades (see Table 1, Table 2) developed *loss and gain* of function explanations to explain why the heterozygote (Aa) is sick in a dominant disorder and healthy in a recessive disorder. We will discuss these in greater detail below. In the next section we will discuss each molecular mechanism and how it relates to the respective inheritance pattern.

Recessive disorder

In the context of the recessive disorder, the students were given a pedigree of a recessive disorder and were asked to explain how this disorder was inherited and to explain the molecular mechanisms that caused the heterozygote individual to be healthy. We found that the students invoked the mechanisms of c*ompensation* and *silencing* to explain why a heterozygote (Aa) individual is healthy in a recessive disorder. The following, quote from a middle-school student illustrates the use of the *compensation* mechanism. "The recessive genes have mutations, but the dominant genes make up for it because they make enough protein and that's why the carrier (Aa) is healthy". Students from all the groups, middle school (50%) to graduate school (100%), explained how the dominant allele was *"compensating*" or was producing enough protein to ensure that the heterozygote individual was healthy. That is, they seemed to understand that the dominant allele as being able to complete the function without the recessive allele. Similarly, 10% of the high schoolers, 15% of the AP students and 30% of the undergraduate students operationalized this explanation of compensation with a mechanism of *silencing* that involved the assumption that the recessive allele is non-functional. In this case, the recessive allele is being turned off or being *silenced* by mutations, therefore they assumed that a mutation can cause cells to turn off the expression of the recessive allele. For example, in the following quote, a high school student describes the recessive allele being silenced. "It could be that the gene is somehow mutated, and that gene expression of that gene is turned off. The recessive allele cannot be transcribed. You can't make any proteins from that gene". This is interesting, while not entirely accurate, because it shows that students at the high school level can understand that the silencing of genes can occur. In the next section we discuss how the students reasoned about dominant disorder.

Table 1: Molecular mechanisms that students used to explain the dominant and recessive patterns of inheritance

Molecular Mechanism	Student Example
1.*Loss-of-function* ***Compensation*** -the proteins made by the dominant allele is sufficient to allow the cell to function normally therefore the allele can compensate for the recessive. Inheritance Pattern: Recesive	"The "normal" proteins compensate for the protein that isn't working the DNA that they inherited from the parent that has the big "A" is "enough" to complete the function of the protein. The protein will be enough to carry on the function even if half of the protein is missing"

2. *Loss-of-function* ***Silencing*** Mutations can turn genes off. Inheritance Pattern: Recessive and Dominant	*Recessive:* "In that case just the dominant one, the amino acid that codes for whatever this has, it's going to exhibit only the dominant ones whereas the ones with the (r) are going to be kind of quiet. They're not going to be as active. I wouldn't say the (r) is making anything. I would say the (R) is making everything. I think the (r) is just there to pass down to the offspring." *Dominant:* "Well, sometimes in dominant disorders you might have proteins not working properly. I think in this case, is because he doesn't have enough healthy protein, and this can be caused by low transcription like problems with a promoter or a receptor not being able to initiate the particular protein to be made."
3.*Loss-of-function* ***Haploinsufficiency*** The dominant allele doesn't make enough. Inheritance Pattern: Dominant	"The big R is directions to make that protein, but the little r doesn't make anything, hypothetically. Then somebody with two big Rs would make ... they would have that instruction from both of their big Rs but somebody with a big R and a little r they would ... they would make the protein because of their big R, but maybe not to the same amount because they don't have two big Rs. The little "r" is not making enough protein, but the dominant allele makes enough."
4. *Gain-of-function* ***Dominant-Negative Mutations*** The protein interferes with the function. Inheritance Pattern: Dominant	"The heterozygote is sick because the mutated protein has something wrong in the structure that allows it to it bind to the healthy protein forming dimers and interferes with the function possibly by blocking it's active site."

Dominant disorder

In this context the students were given two multiple-choice questions that asked students to explain why a heterozygote (Aa) individual is sick in a dominant disorder. We found that students used the mechanisms of *haploinsufficiency*, *dominant-negative mutations* and *silencing* to explain why the heterozygote (Aa) individual is affected by the disorder (Table 2, row 3). *Haploinsufficiency* occurs when one allele in the pair of alleles is mutated causing a decrease in the total production of a protein. The use of this explanation suggests that students understood that the "healthy" allele is not able to produce enough protein for the individual to have a "normal" phenotype. For example, in the following quote a high school student states that "Sam" is not healthy because having one healthy allele "a" is not enough for him to be healthy. "Sam is not making enough protein to be healthy. He has one dominant(A) and one recessive allele(a). The recessive is the correct one , this allows him to make the Tropin but he doesn't make it in the quantities that he needs. He is missing one healthy allele therefore he doesn't make enough of the protein because of the mutation in the A allele."

Dominant negative mutation explanations were used to explain dominant disorders by 10% of the high school students, 20% of the AP students, 35% of the undergraduates and of the 90% of the graduate students. In this explanation, the students suggest that the mutated protein is somehow inhibiting the function of the healthy allele, therefore the heterozygote individual (Aa) is sick in a dominant disorder (Table 2, row 4). In the following quote, an undergraduate student explains that there may be a mutated protein interfering with the function of the normal protein. "The heterozygote is sick because the mutated protein has something wrong in the structure that makes it bind to the healthy protein and interferes with the function". The use of this mechanism suggests that students understood that proteins can interact with one another, sometimes leading to inhibition of their function.

Silencing in the context of the dominant disorder was used mainly by undergraduate and graduate students and it consisted of mutations that affect the activation of genes, therefore there was a lack of protein expression. For example, in the following quote a graduate student suggests that a gene is being silenced due to mutations. "Sam may be sick for several reasons. However, it is possible that regions that are related to gene activation were affected by the mutation therefore Sam is unable to make the protein he needs to be healthy."

In the next section we discuss how these explanations were developed by students across several levels of expertise.

Development of molecular explanations over time

Table 2 shows the frequencies of types of mechanistic explanations used by each group of students to explain recessive and dominant disorders. The data seems to indicate several interesting patterns. First, some mechanisms

are observed with an increase in their use corresponding to education level (e.g., Table 2, rows 1-2). For example, the *compensation* explanation was used correctly by students across several grades to explain the inheritance of recessive and dominant disorders. Similarly, constructing *dominant-negative* explanations was challenging, but within reach, for high school students. All of the students used these in the correct context and situation. Second, the mechanism of "s*ilencing*" in the context of the recessive disorder was not used by middle school students, but it was used by high school students, AP students and very few graduate students. This mechanism was used in a non-normative way by the younger students while the graduate students provided more detail about how a gene can be "silenced. Additionally, silencing was used in the context of the dominant disorder by undergraduate and graduate students, including more detail on how the gene may have been inactivated and how this inactivation can lead to disease. Third, the *haploinsufficiency* explanation, in terms of the dominant allele and the dominant-negative mutations, were used in an appropriate context for the dominant disorder - mostly by graduate students (Table 2, rows 4).

Table 2: Progression of molecular explanations over time

Molecular Mechanism	Inheritance Pattern	MS	HS	AP	College	Grad
1. Compensation--the proteins made by the dominant allele is sufficient to allow the cell to function normally therefore the allele can compensate for the recessive.	Recessive	50%	70%	80%	100%	100%
2. Silencing-Mutations can turn genes off.	Recessive	0%	10%	15%	30%	50%
	Dominant	0%	0%	0%	20%	60%
3. Haploinsufficiency-the proteins made by the dominant allele is not sufficient to allow the cell to function normally.	Dominant	20%	35%	40%	50%	100%
4. Dominant Negative- the protein produced by the mutated allele interferes with the function of the "normal" protein.	Dominant	0%	10%	20%	35%	90%

Discussion and instructional implications

Genetics is an important, but challenging, subject for students to learn and understand. According to Stewart et al. (2005), a literate individual understands the links among the classical, molecular and meiotic models of genetic inheritance. There is little research to show how students integrate these models. Our findings characterize the different mechanisms that students used to link classical and molecular patterns of inheritance post status-quo instruction. The molecular explanations that students developed were, for the most part, consistent with the normative model of *loss and gain of function*. For example, the mechanism of *haploinsufficiency* involves the loss of part of the protein function leading to a decrease in the protein's expression, and the mechanism of *compensation* involves the understanding that although there may be a decrease in the protein's expression the other allele is able to compensate. Our findings indicate that middle school students were able to reason about these ideas of *compensation* and *haploinsufficiency*. These ideas involve an understanding that alleles can code for *enough* protein to complete a function and may suggest a molecular understanding of *dosage* (Duncan, 2017). Understanding *dosage* may allow allows students to traverse several levels of organization because it allows them to understand the relationships among alleles, protein and their physical effects. Helping students understand the relationships between genes and how they bring about their effects is conductive to understanding the mechanistic linkages between molecular and classical genetics (Haskel-Ittah & Yarden, 2019).

Our findings also show that students at the high school level can understand that genes can be turned *on or off.* This idea of *silencing* is critical in the understanding that genes are dynamic entities that are influenced by the environment. Authors, 2020 studied how undergraduates understood phenotypic plasticity, that is how organisms respond to changes in the environment. They found two types of mechanistic accounts: (a) the organism is passively responding to environmental changes, and (b) the organism senses and responds to the environment. Our findings seem to support the idea that understanding how organisms respond is difficult and we propose that understanding that genes can be turned on and off is an important part of conceptualizing how the

environment influences gene expression. This is an important idea because understanding that genes are dynamic entities can decrease genetic determinism (Donovan, 2016). Moreover, Duncan, 2007 studied how students understood how plants responded (by ripening) to an environmental change (exposure to ethylene). She found that students that were not able to operationalize a gene expression (or silencing) schema were not able to understand or explain how the organism reacted to the environment or further explain how genes are regulated. Therefore, the understanding that genes can be turned on and off is an important idea for the understanding of gene expression (Duncan, 2007; Haskel-Ittah et al., 2020).

To conclude, we wish to discuss the progression of mechanism use. From these findings we have begun to identify a progression that captures the mechanisms that students used when reasoning about classical and molecular genetics. Duncan, Rogat and Yarden (2009) proposed a learning progression for the integration of molecular and classical ideas. Our findings support the progression and provide details about how students' reasoning about the linkage between molecular and classical genetics may develop over time. Duncan et al.'s (2009) learning progression describes the development of students understanding of eight genetics ideas (constructs) over time (middle school to high school).

Our findings, can begin to specify where instruction can focus to help students develop the mechanistic explanations for the link between molecular and classical genetics, for instance we have shown that instruction should focus on emphasizing that alleles code for proteins and that recessive and dominant are simply categorical to indicate the phenotypic effect of the proteins they code for and that these proteins they are structurally different due to changes in the sequence (recipe). Moreover, instruction should emphasize that proteins have distinct functions that have a direct result in the organism's physical traits. Understanding the connections among genes, proteins, and traits entails reasoning across multiple levels of organization (e.g., sub-cellular, cellular, and organismal levels), a conceptually challenging task since students need to understand the role each level of organization plays in inheritance phenomena and how biological mechanisms at each level relate to each other (Marbach-Ad & Stavy, 2000). Duncan and Tseng, 2011 argue that understanding the link between genes and traits can help students develop a framework for reasoning about complex systems (such as the interaction between classical and molecular genetics) that can be further developed when students understand how physical and cellular phenomena emerge from protein function. Therefore, helping students understand the relationship between genes and proteins can help elucidate the link between the two models.

Similarly, we have shown that helping students understand that genes can be regulated is important but difficult idea, therefore instruction should aim to help students understand that genes are dynamic entities that respond to the environmental changes the organism is interacting with. More studies need to be done to further understand how students reason about these ideas. This article begins to provide evidence about the ways that students link classical and molecular genetics.

References

23andMe. (2020). Retrieved from: https://www.23andme.com/

Bates, G. P. (2005). The molecular genetics of Huntington disease—a history. *Nature Reviews Genetics*, 6(10), 766-773.

Bellcross, C.A., Page, P., & Meaney-Delman, D. (2012). Direct-to-consumer personal genome testing and cancer risk prediction. *The Cancer Journal, 18*(4), 293 - 302.

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.

Butler-Kisber, L., & Poldma, T. (2010). The power of visual approaches in qualitative inquiry: The use of collage making and concept mapping in experiential research. *Journal of Research Practice*, 6(2), M18-M18.

Chi, M. T. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The Journal of the Learning Sciences*, 6(3), 271-315.

Castro-Faix, M., Duncan, R. G., & Choi, J. (2020). Data-driven refinements of a genetics learning progression. *Journal of Research in Science Teaching*.

Donovan, B. M. (2017). Learned inequality: Racial labels in the biology curriculum can affect the development of racial prejudice. *Journal of Research in Science Teaching*, 54(3), 379-411.

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., Castro-Faix, M., & Choi, J. (2017). Informing a learning progression in genetics: Which should be taught first, Mendelian inheritance or the central dogma of molecular biology?. *International Journal of Science and Mathematics Education*, *14*(3), 445-472.

Duncan, R. G., Rogat, A., & Yarden, A. (2009). A learning progression for deepening students'understanding of modern genetics across the 5th-12th grades. Special issue on Learning Progressions for *the Journal of Research in Science Teaching*, 46(6), 644-674.

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

Falzarano, M. S., Scotton, C., Passarelli, C., & Ferlini, A. (2015). Duchenne muscular dystrophy: from diagnosis to therapy. *Molecules*, *20*(10), 18168-18184.

Gericke, N. M., & Smith, M. U. (2014). Twenty-first-century genetics and genomics: Contributions of HPS-informed research and pedagogy. In *International handbook of research in history, philosophy and science teaching* (pp. 423-467). Springer, Dordrecht.

Griffiths, A. J., Wessler, S. R., Lewontin, R. C., Gelbart, W. M., Suzuki, D. T., & Miller, J. H. (2005). *An introduction to genetic analysis*. Macmillan.

Haskel-Ittah, M., Duncan, R. G., & Yarden, A. (2020). Students' Understanding of the Dynamic Nature of Genetics: Characterizing Undergraduates' Explanations for Interaction between Genetics and Environment. *CBE—Life Sciences Educatio*n, 19(3), ar37.

Haskel-Ittah, M., & Yarden, A. (2019). Students' conception of genetic phenomena and its effect on their ability to understand the underlying mechanism. *CBE—Life Sciences Education*, 17(3), ar36.

Hawkins, A.K. & Ho, A. (2012). Genetic counseling and the ethical issues around direct to consumer genetic testing. *Journal of Genetic Counseling*, 21(3), 367 - 373.

Hopkins. (2020). Retrieved from: https://www.hopkinsmedicine.org/health/conditions-and-diseases/sickle-cell-disease

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research*, *15*(9), 1277-1288.

Jones, L., & Hughes, A. (2011). Pathogenic mechanisms in Huntington's disease. In *International review of neurobiology* (Vol. 98, pp. 373-418). Academic Press.

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

Mayring, P. (2000, June). Qualitative content analysis forum qualitative sozialforschung. In *Forum: qualitative social research* (Vol. 1, No. 2, pp. 2-00).

Muller, H. J. (1932). Further studies on the nature and causes of gene mutations. *Proc. Sixth Int. Cong. Genet., Ithaca, New York, USA*, *1*, 213-255.

Shea, N. A., Duncan, R. G., & Stephenson, C. (2015). A tri-part model for genetics literacy: Exploring undergraduate student reasoning about authentic genetics dilemmas. *Research in Science Education*, *45*(4), 485-507.

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

Stewart, J., Cartier, J. L., & Passmore, C. M. (2005). Developing understanding through model-based inquiry. *How Students Learn*, 515-565.

Todd, A.N., & Romine, W. (2017) Difficulties in integrating the genetic-meiotic and genetic molecular models of genetics. A paper presented at the 2017 Annual Meeting of the National Association for Research in Science Teaching, San Antonio, TX.

Weisberg, S. M., Badgio, D., & Chatterjee, A. (2017). A CRISPR new world: attitudes in the public toward innovations in human genetic modification. *Frontiers in public health*, 5, 117.

Wilkie, A. O. (1994). The molecular basis of genetic dominance. *Journal of medical genetics*, 31(2), 89-98.

Wolyniak, M. J. (2013). Improved student linkage of mendelian and molecular genetic concepts through a yeast-based laboratory module. *Biochemistry and Molecular Biology Education*, 41(3), 163-172

Zhou, Y., Borchelt, D., Bauson, J. C., Fazio, S., Miles, J. R., Tavori, H., & Notterpek, L. (2020). Subcellular diversion of cholesterol by gain-and loss-of-function mutations in PMP22. *Glia*.