International Journal of Science Education

ISSN: 0950-0693 (Print) 1464-5289 (Online) Journal homepage: http://www.tandfonline.com/loi/tsed20

Validation of the Learning Progression-based Assessment of Modern Genetics in a college context

Amber Todd & William L. Romine

To cite this article: Amber Todd & William L. Romine (2016) Validation of the Learning Progression-based Assessment of Modern Genetics in a college context, International Journal of Science Education, 38:10, 1673-1698, DOI: 10.1080/09500693.2016.1212425

To link to this article: http://dx.doi.org/10.1080/09500693.2016.1212425

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION, 2016
VOL. 38, NO. 10, 1673–1698
http://dx.doi.org/10.1080/09500693.2016.1212425

Routledge
Taylor & Francis Group

Validation of the Learning Progression-based Assessment of Modern Genetics in a college context

Amber Todd and William L. Romine

Biological Sciences, Wright State University, Dayton, OH, USA

ABSTRACT

Building upon a methodologically diverse research foundation, we adapted and validated the *Learning Progression-based Assessment of Modern Genetics* (LPA-MG) for college students' knowledge of the domain. Toward collecting valid learning progression-based measures in a college majors context, we redeveloped and content validated a majority of a previous version of the LPA-MG which was developed for high school students. Using a Rasch model calibrated on 316 students from 2 sections of majors introductory biology, we demonstrate the validity of this version and describe how college students' ideas of modern genetics are likely to change as the students progress from low to high understanding. We then utilize these findings to build theory around the connections college students at different levels of understanding make within and across the many ideas within the domain.

ARTICLE HISTORY
Received 4 January 2016
Accepted 10 July 2016

KEYWORDS
Assessment; genetics; learning progression; Rasch model

An important element of scientific literacy for college students is understanding modern genetics. Although this domain is relatively new compared to other scientific domains such as physics or astronomy, it is becoming increasingly important for people to understand genetics at a time when issues around technologies such as genetic screening, genetically modified organisms (GMOs), and stem cell therapies are encountered during everyday living. Modern genetics is poised to make a huge impact on detection of diseases, evaluation of risk of diseases, and advancement of personalized medicine, but understanding and taking advantage of this information while understanding its limitations is extremely complex and requires adequate knowledge of the domain (Gollust, Wilfond, & Hull, 2003; Hull & Prasad, 2001; Lewis & Wood-Robinson, 2000). In this vein, *Vision and Change* (AAAS, 2011) identified five core concepts for college undergraduate biological literacy and modern genetics is included in three: evolution; structure and function; and information flow, exchange, and storage.

Although modern genetics is important for students to understand, research consistently demonstrates that this content is difficult to teach and learn (e.g. Fisher, 1992;

CONTACT Amber Todd amber.todd@wright.edu Wright State University, 3640 Colonel Glenn Hwy, Dayton, OH 45435, USA

Supplemental data for this article can be accessed here http://dx.doi.org/10.1080/09500693.2016.1212425.

Gericke & Smith, 2014; Lewis & Kattmann, 2004; Marbach-Ad & Stavy, 2000; Stewart, Cartier, & Passmore, 2005; Wynne, Stewart, & Passmore, 2001). Assessments for genetics do exist at the college level; however, none are tied to current learning progression theory and none have been rigorously evaluated using modern statistical methods such as item response theory or Rasch modeling. We have previously demonstrated that our *Learning Progression-based Assessment for Modern Genetics* (LPA-MG) is a valid, reliable, and unidimensional instrument for measuring high school (10th grade) students' conceptions of modern genetics (Todd, Romine, & Cook Whitt, in press). In this study, we describe the revision of the LPA-MG instrument for university students and subsequent revalidation for this population. We proceed to utilize this new version to create learning progression profiles for how university biology students at different levels of understanding conceptualize modern genetics concepts, and propose a hypothetical concept map showing connections between concepts in the progression. We address the following research questions:

(1) Can quantitative learning progression-based measures for modern genetics deliver a valid, reliable, and unidimensional measure for college students?
(2) How are students' profiles of understanding of modern genetics likely to change as they progress from low to high understanding?

Background and theoretical framework

Modern genetics literacy

Literacy in genetics is extremely important for people in modern society – the general public is encountering molecular genetics during their everyday lives with technologies such as genetic screening, gene therapy, GMOs, DNA sequencing, and stem cell research and therapies now commonplace (Gericke & Smith, 2014; Gollust et al., 2003; Hull & Prasad, 2001; Lewis & Wood-Robinson, 2000). Ideas in modern genetics are included in the *Next Generation Science Standards* (NGSS Lead States, 2013) for students in the K-12 arena, and while there are no set standards for what college students should learn in biology, *Vision and Change* (AAAS, 2011) identifies five core concepts for college undergraduate biological sciences literacy. Modern genetics ideas are included in three of the five: evolution (diversity of life evolved over time by process of mutation, selection, and genetic change); structure and function (basic units of structure such as DNA, proteins, and cells define the function of all living things); and information flow, exchange, and storage (growth and behavior of organisms are activated through expression of genetic information).

Along with standards and policy documents that describe performance expectations of students, research groups have also described what they interpret as modern genetics literacy and the core concepts students need to understand. Stewart et al. (2005) describe literacy as understanding and integration of three inter-related models of genetics: *genetic* (inheritance, Mendelian, classical, or transmission genetics), *meiotic* (the process of meiosis), and *molecular* (genes code for proteins, protein structure function, and gene expression). Duncan (2007) describes heuristics (*genes-code-for-proteins*, *proteins-as-central*, *effects-through-interaction*) and explanatory schemas (*inhibit*, *activate*, *translation*, *regulation-of-gene-expression*, *catalyze*, *transport*, *receptor*, *structural*, *structure-function*) critical for modern genetics reasoning. Shea,

Duncan, and Stephenson (2015) describe that literacy consists of content knowledge plus situational features of the task and argumentation quality.

Modern genetics is an important and relevant part of biology curricula around the world, but it is notoriously difficult for students and teachers alike (e.g. Fisher, 1992; Gericke & Smith, 2014; Lewis & Kattmann, 2004; Marbach-Ad & Stavy, 2000; Stewart et al., 2005; Wynne et al., 2001). Students and teachers have difficulties understanding the concept of a gene (Dikmenli, Cardak, & Kiray, 2011; Gericke & Wahlberg, 2013; Lewis & Kattmann, 2004; Smith & Williams, 2007; Venville & Treagust, 1998), how entities give rise to things seen at higher organizational levels (Duncan & Reiser, 2007; Hallden, 1990; Marbach-Ad & Stavy, 2000), that genes do not directly code for traits but instead code for proteins (Donovan & Venville, 2014; Duncan & Reiser, 2007; Lewis, Leach, & Wood-Robinson, 2000; Thörne & Gericke, 2014), and that topics in genetics are actually related to each other (Knippels, 2002).

Most of the more qualitative research in genetics understanding has focused on the primary and secondary academic levels. Understanding the challenges and struggles students and teachers have with the content has improved with this research. While there are studies that look at teachers' knowledge of genetics content (i.e. Thörne & Gericke, 2014), there are few qualitative studies that assess college students' understandings and none are tied to a learning progression for genetics. Our study is the first to use a learning progression framework for modern genetics to assess college introductory biology students' understandings and provide some qualitative descriptions of their conceptions.

Learning progressions and modern genetics

Modern genetics literacy can also be described in terms of learning progressions. Learning progressions are models of student learning that describe increasingly sophisticated ways of reasoning within a content area or scientific practice. They contain upper and lower bounds and intermediate ideas in the format of learning performances that increase in sophistication (Corcoran, Mosher, & Rogat, 2009). The ordered learning performances do not necessarily describe a singular pathway that all students follow toward an expert-level understanding; they instead describe increasingly sophisticated and productive ways of understanding the domain. Students may move forward or backward over time as their understanding changes and may even transition through certain levels extremely quickly (Corcoran et al., 2009; Rogat et al., 2011; Stevens, Delgado, & Krajcik, 2010).

A learning progression is a useful cognitive model that describes increasingly sophisticated understandings that individuals may have as they progress toward an expert-level understanding of the domain. Though the ultimate goal of instruction in a domain would be the expert-level understanding (upper bounds) and instructors may have targeted curriculum and instruction designed to help students achieve the upper learning performance, achievement of the upper bounds is not guaranteed (Duncan & Hmelo-Silver, 2009).

The modern genetics domain has four published learning or conceptual progressions (Dougherty, 2009; Duncan, Rogat, & Yarden, 2009; Elmesky, 2013; Roseman, Caldwell, Gogos, & Kurth, 2006). Roseman's and Duncan's progressions target grades 5–10, Dougherty's progression focuses on concepts for middle school and high school students, and

Elmesky's progression targets grades K-12; none of the progressions extend into college understandings. A learning progression for a domain is considered a hypothetical model, though based on substantial research, until the progression itself is empirically tested, revised, and validated through an iterative process using classroom data (Rogat et al., 2011; Shea & Duncan, 2013). As progressions are refined, curricula and assessments tied to the progression (Duncan & Hmelo-Silver, 2009) and connections between constructs in a progression (Corcoran et al., 2009; Wilson, 2009) are typically added. Duncan's progression is the only of the four progressions to begin to be empirically tested (Shea & Duncan, 2013; Todd & Kenyon, 2015) and is the most well-defined. As such, we tied our *Learning Progression-based Assessment of Modern Genetics* (LPA-MG) to the recent revisions of Duncan's progression.

Our revisions to Duncan's progression and development of the first version of the LPA-MG (Todd, & Kenyon, 2015; Todd, Romine & Cook Whitt, in press) posit 12 inter-related constructs within the progression, each construct containing five to seven different levels of learning performances. A condensed version of these constructs and their levels is described in Table S1. The original Duncan progression builds on Stewart et al.'s (2005) idea that literacy in this area consists of being able to understand and integrate the three models of genetics (*molecular*, *genetic*, and *meiotic*). While certain constructs in the progression focus on one of the models, several constructs contain learning performances that integrate models; thus, Duncan et al. (2009) explain that the three models should be taught concurrently and in relation to one another and not as separate concepts. Though Duncan's progression is targeted to students in grades 5–10, we used the LPA-MG to assess student understanding of the 12 different constructs within the revised progression in a college context. Several assessments in genetics have been written for college-level students, but none are tied to learning progression theory.

Assessments of modern genetics

The genetics domain has a wealth of assessments, but few have been psychometrically validated (McElhinny, Dougherty, Bowling, & Libarkin, 2014). The Genetics Literacy Assessment Instrument targets college students, including non-science majors (Bowling et al., 2008). It probes 6 broad concepts and 17 sub-concepts, including some modern genetics ideas. It has been evaluated for validity and reliability. The Genetics Concept Assessment targets college genetics majors and non-majors (Smith, Wood, & Knight, 2008). It probes nine concepts and has also been validated. Other college-level assessments that contain topics in genetics include the Genetics Concept Inventory, under development (Elrod, 2007); the Molecular Life Sciences Concept Inventory, under development (Howitt, Anderson, Costa, Hamilton, & Wright, 2008; Wright & Hamilton, 2008); the Biology Concept Inventory, assessing general biology knowledge but validated through interviews and reliability determined by Cronbach's alpha (Klymkowsky, Underwood, & Garvin-Doxas, 2010); the Introductory Molecular and Cell Biology Assessment, assessing molecular and cellular biology and validated for validity and reliability using classical test theory (Shi et al., 2010); and the Molecular Biology Capstone Assessment, targeting molecular biology and genetics for college students completing undergraduate work and evaluated for reliability using Cronbach's alpha, test–retest stability, item difficulty, and question discrimination (Couch, Wood, & Knight, 2015).

At the high school level, Tsui and Treagust (2010) developed a two-tiered genetics assessment and determined reliability using Cronbach's alpha; Zohar and Nemet (2002) developed an instrument to assess genetics in the context of argumentation; and Sadler and Zeidler (2005) developed an instrument to assess genetics in the context of socioscientific reasoning. Our LPA-MG is an instrument aligned with revisions to Duncan's genetics learning progression that has been validated through the Rasch model at the high school level (Todd, Romine, & Cook Whitt, in press).

Here, we describe Version 2 of the LPA-MG and its validation in a college introductory biology context. We use our findings to describe introductory biology students' understandings of modern genetics and propose a hypothetical concept map of an expert-level understanding of genetics and discuss the implications of college students' understandings for instructional purposes.

Methods

Context

Our sample contained 316 students (138 biology majors, 174 non-biology majors, and 4 unknown) from a Midwestern open-enrollment research university. University-wide, the total minority student enrollment is 19.7% (10.4% African-American, 3.3% two or more races, 2.8% Asian American, 2.9% Hispanic American, 0.2% American Indian or Alaskan Native, and 0.1% Native Hawaiian or Pacific Islander) and the total international student enrollment is 11% (65 different countries). Our sample was consistent with these demographics. Students from two sections (fall [$n = 226$] and spring [$n = 90$]) of a college introductory biology course intended for majors were given the assessment. Topics during this course include genetics and the molecular and cellular basis for the unity of life, including the concepts assessed by our instrument. Students completed the assessment during the last few weeks of the course after these concepts were discussed. The course is a traditional introductory biology general education course with three hours of lecture and two hours of lab for four semester credit hours. The course was taught in a lecture format and there was no science prerequisite.

The LPA-MG instrument and revisions

The LPA-MG was designed to assess 12 different concepts (Table 1), with each construct aligning with Duncan's learning progression and its revisions (Duncan et al., 2009; Shea & Duncan, 2013; Todd, 2013; Todd & Kenyon, 2015; Todd, Romine, & Cook Whitt, in press). We constructed three assessment items for each construct using the ordered multiple choice (OMC) framework (Briggs, Alonzo, Schwab, & Wilson, 2006). Using the revised progression constructs (Todd, Romine, & Cook Whitt, in press) for the test's within-assessment structure (Wilson, 2009), the responses for each item corresponded to the different levels for that construct. Generally, the number of levels for each construct corresponded to the number of responses for items probing that construct. More detailed explanations about the item design are discussed in Todd, Romine, & Cook Whitt (in press); outline of progression levels can be found in Table S1.

Table 1. Modern genetics constructs.

Construct	Concept	Assessment Items	Levels
A	Genetic information is hierarchically organized	A (V1, V2, and V3 combined)	0–6
B	Genes code for proteins	V4, V5, V6	0–6
C1	Proteins do the work of the cell	V7, V8, V9	0–5
C2	Proteins connect genes and traits	V10, V11, V12	0–6
D	Cells express different genes	V13, V14, V15	0–6
E	Genetic information is passed on to offspring	V16, V17, V18	0–5
F	There are patterns of correlation between genes and traits	V19, V20, V21	0–5
G1	DNA varies between and within species	V22, V23, V24	0–6
G2	Changes to genetic information result in increased variation and can drive evolution	V25, V26, V27	0–5
H	The environment interacts with genetic information	V28, V29, V30	0–6
I	Only mutations in gametes can be passed to offspring	V31, V32, V33	0–4
J	Gene expression can change at any point during an organism's lifespan	V34, V35, V36	0–4

We previously used this assessment to probe understandings of high school (10th grade) students (Todd, Romine, & Cook Whitt, in press). While the assessment was found to be a valid, reliable, and unidimensional measure for modern genetics understanding in high school students, we chose to revise many items on the assessment based on our data and feedback from four genetics content experts at the university to better tailor the instrument to university students. We will now briefly discuss these changes from Version 1 (Todd, Romine, & Cook Whitt, in press) to Version 2 used in this study. Version 2 of the assessment is included in supplementary materials.

The majority of changes to items entailed small wording changes to make the question stems and distractors clearer, adding details to questions, and/or rearranging wording to keep lengths and answer styles within an item similar. Twenty of the items were modified in this fashion. The context of four questions was modified for simplicity and accuracy and a protein model was changed in one question from a ribbon model to a space-filling model. We also changed all of the items that asked students to select incorrect statements in Version 1 of the assessment. Our reviewers suggested that negatively worded assessment items can be problematic for students, which is also stated in research (Weems, Onwuegbuzie, Schreiber, & Eggers, 2003); hence we changed four items in this manner and asked students to select the correct answers. Students could choose any, all, or none of the answers. A detailed description containing the items changed and the differences between Version 1 and Version 2 are included in supplementary materials as Table S2.

Assessment administration and scoring

The assessment was administered using Qualtrics survey software and was given to students as an extra credit assignment to be completed outside class. The population of students completing the assessment was representative of the course as a whole. Assessment items and item responses were both randomized using Qualtrics. As with Version 1 of the assessment, we chose to retain the certainty of response index (CRI) for each item to control for the confounding effects of guessing. CRI asks students to denote their confidence after each question using a 1–4 rating scale (Romine, Schaffer, & Barrow, 2015). The scale used was 1: guessing, 2: uncertain, 3: certain, and 4: very confident. Each assessment item and its CRI question were placed together on individual pages within the online

survey so students could answer the item and denote their level of confidence on that specific item before moving on to the next item on the next page. The assessment typically took students 20–30 minutes to complete.

Scoring structures for specific types of items were detailed in Todd, Romine, & Cook Whitt (in press), to which we will refer the interested reader. Briefly, since the items were written based on a learning progression framework using the OMC format, each response was targeted to a specific level of that particular construct. Students received a score for this item equal to the level of response selected, unless they indicated in the accompanying CRI question that they were guessing (indicated by selection of the '1: guessing' response). In a multiple choice assessment, students must choose something even if they do not know the answer. Since each response corresponded to a particular level of the construct, guessing could be a serious threat to the validity of this OMC assessment. Students who selected a 'guessing' CRI response were thus given a score of '0' for this item to account for the fact that guessing occurs; our previous work suggests that using the CRI in this manner gives a more accurate measure of the students' true knowledge level (Romine, Schaffer, & Barrow, 2015; Todd, Romine, & Cook Whitt, in press).

Items as measures of the progression

A great deal of qualitative data have been used to revise and refine Duncan's progression; however, this work has only been done with middle school and high school students (Duncan et al., 2009; Shea & Duncan, 2013; Todd, 2013; Todd & Kenyon, 2015) as the progression was intended for grades 5–10. Given this, there have been no studies that have examined understandings of college students in relation to the progression. Tenth grade students are expected to be able to reason at the highest level of each construct; so it is reasonable to deduce that college students should be able to show high levels of modern genetics understanding relating to the progression.

To evaluate construct validity with respect to the progression in a college context, we used a mixture of Rasch partial credit (Masters, 1982) and rating scale (Andrich, 1978) models estimated with BIGSTEPS. Version 2 of the LPA-MG contained 34 items (36 individual items with V1–3 condensed into Item A) linked to the 12 constructs described in Todd, Romine, & Cook Whitt (in press). For items measuring the same construct (i.e. V4–6, measuring construct B, Table 1), the same rating scale was used for fit, making 12 different rating scales for the 34 items. As with our analysis of the data from Version 1 of the LPA-MG, this simpler model of 12 rating scales as opposed to 34 different rating scales, provided it fits the data, reveals the important progression trend in the data while not modeling response error. Satisfactory fit with this model provides the extent to which the items in Version 2 of the assessment conform to the hypothesized progression constructs in a college context.

The Rasch model provides a data-independent criterion for validity of each item and progression structure in its implicit assumption that the probability of a student identifying with a higher level of the progression should increase with his/her knowledge of modern genetics. The Rasch model offers many benefits for validation of assessments (Boone, Townsend, & Staver, 2011), including those which are learning progression based. While the many benefits of Rasch are described elsewhere, notable advantages for this study are twofold: (1) it provides a data-independent criterion by which to test

and possibly falsify the hypothesis that data produced by the instrument are valid and (2) it puts item and person measures on the same scale, allowing straightforward prediction of a particular student's location along the progression based on his/her measure. Infit and outfit statistics with respect to the Rasch model were used to evaluate construct validity of items and rating scales. We obtained mean squares fit indices by comparing observed response patterns to those expected by the Rasch model. Infit is information-weighted, making it less affected by outliers than outfit. We used mean squares fit indices below 1.3 as indicative that an item's response pattern along the progression fits adequately with what would be expected from the Rasch model.

The scale as a measure of the progression

Validity of the scale was evaluated based on the reliability of person and item measures along the scale, unidimensionality, and overlap of student and item measure distributions. We calculated reliability of student and item measures with respect to the 12-rating scale partial credit model.

Since high reliability does not necessarily imply that the test measures a single dimension (Schmitt, 1996); we used principal components analysis (PCA) on standardized residuals as a tool for detecting left-over dimensions in the residuals, if they existed, which would imply that a single scale is not explaining the data completely. Simulation work on polytomous Rasch models suggests that a first eigenvalue around or below 2 indicates that the residuals are random (Linacre & Tennant, 2009), meaning that a single scale is sufficient to capture the data.

Results

Can quantitative learning progression-based measures for modern genetics deliver a valid, reliable, and unidimensional measure for college students?

We found that Version 2 of the LPA-MG provides measures for college students' knowledge of modern genetics related to a learning progression intended for grades 5–10 that are both valid and reliable. We discuss validity evidence for the scale, progression, and items in the following sections.

Scale-level validity

Though the LPA-MG probes 12 different constructs (concepts) within the modern genetics domain (Table 1), we found that Version 2 of the instrument provided measures that are valid, reliable, and measure a single dimension. We determined that items measured a single idea and exhibit low dependency after accounting for the underlying construct, knowledge of modern genetics, which indicates that measures meet the basic assumptions of Rasch analysis. The first eigenvalue from PCA on residuals was 1.20 items of variance, which is well below the cutoff of 2. The largest item residual dependency was 0.06 between V9 (construct C1) and V14 (construct D). That well under 1% of the systematic variation between these items is not accounted for by the Rasch model illustrates negligible item dependency on constructs other than modern genetics knowledge. This gives a first look at item consistency, supporting the important validity argument

that all items measure knowledge of modern genetics, and that the measures they provide are not biased or corrupted by outside factors. This is further supported by the fact that the LPA-MG provides highly reliable measures for college students. Total reliability for student measurement was 0.86 (2.47 separation). Item locations along the Rasch scale were estimated with a reliability of 0.98 (6.45 separation), which is more than sufficient for establishing construct validity of items.

Using a person–item map (Figure 1), we can begin to get a qualitative look at construct validity. Though the participants were college students and the items were written within a progression framework targeted to grade 5–10 students, we see that the distribution of students and measures for items are comparable. Indeed, we see near-perfect overlap between the person measure and item measure distributions. This implies that Version 2 of the LPA-MG is not too easy or too difficult for this population of college students. Similar to our findings with high school students (Todd, Romine, & Cook Whitt, in press), items aligning with construct A: genetic organization, J: gene expression can change at any point (V35), and C1: proteins do the work of the cell (V7) were the easiest items on the instrument. A majority of

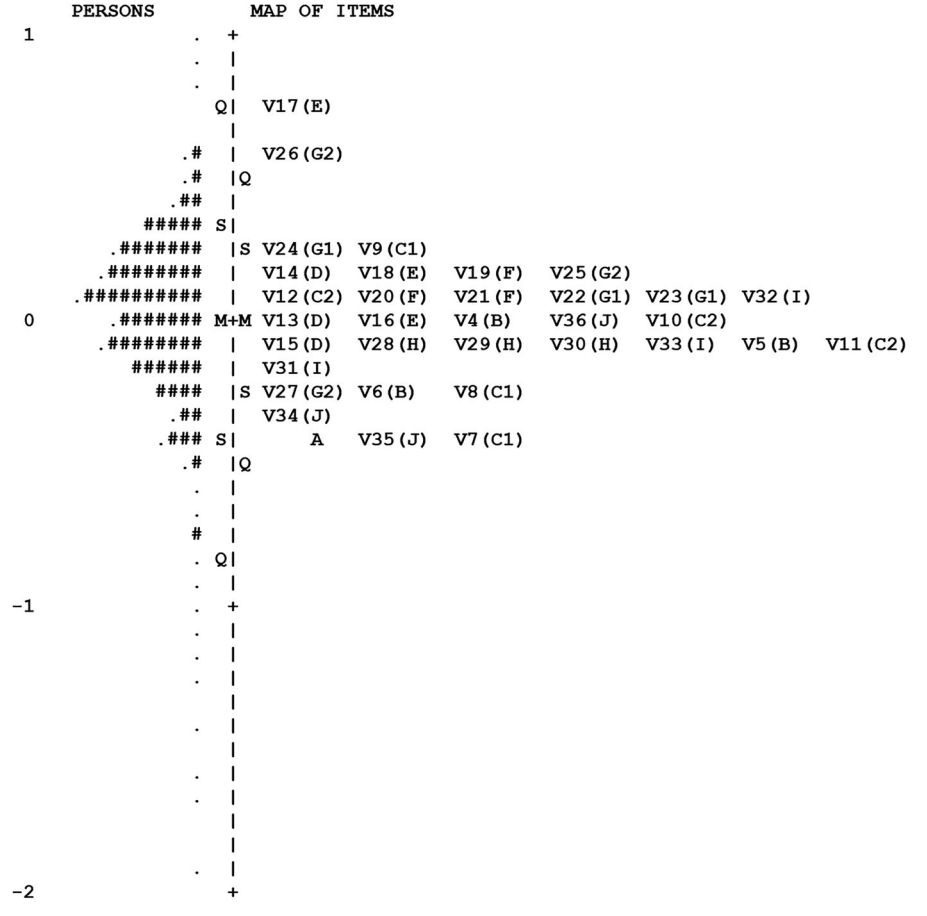

Figure 1. Wright map of person and item measures along the Rasch scale.

students' measures sit above these items, indicating that a majority of students understood how genetic information is organized, that gene expression can change at any point during an organism's lifespan, and that proteins perform the main functions in a cell. Items aligning with construct E: genetic information is passed on to offspring (V17) and G2: changes to genetic information result in increased variation and can drive evolution (V26) were the most difficult items on this version with this student population. The majority of students in this study had measures below these items, indicating that a majority of students did not understand the details of meiosis or that DNA mutations introduce variations which are a driving force of evolution.

Progression-level validity

The tendency to find students at a certain level of the progression generally increased with their level of knowledge of modern genetics as measured by Version 2 of the LPA-MG instrument – item-measure correlations took values between 0.22 and 0.56. Analysis of threshold infit and outfit indices for each of the 12 rating scales showed small potential issues with the middle categories on 5 of the 12 construct ideas. Level 1 was found to be misfitting (infit = 1.55, outfit = 1.40) for construct category C1 (Proteins do the work of the cell). Level 4 was found to misfit students' responses on construct categories D (cells express different genes) (infit = 1.18, outfit = 2.93), F (there are patterns of correlation between genes and traits) (infit = 1.39, outfit = 1.83), and G1 (DNA varies between and within species (infit = 1.15, outfit = 1.62). Level 5 was found to be misfitting (infit = 1.24, outfit = 2.15) for construct category C2 (proteins connect genes and traits). Misfitting of the lower progression levels with the Rasch model revealed the tendency to see higher ability students at these levels. Along a similar line, misfitting of the higher progression levels attests to a tendency to see lower ability students at the higher levels. From a measurement perspective, we do not consider this a significant cause for concern given that the 'messy middle' is a common phenomenon observed in learning progressions research (Alonzo, 2012; Gotwals & Songer, 2010). In this sense, it is noteworthy that a majority of the middle progression levels displayed satisfactory fit with the Rasch model. Further, the highest and lowest progression levels along each concept construct fit well with the Rasch model, indicating valid and well-defined high- and low-progression boundaries.

Item-level validity

As the pre-defined progression structures generally conformed well to the data, so did the items within each construct. Three of the 34 items were found misfitting (V14, V18, and V23), with mean squares fit >1.30 (Figure 2). The misfit of item V14 (construct D) resulted from a tendency for students in the upper third of the distribution to identify with the middle progression levels. The misfit of V18 (construct E) resulted from a lack of knowledge of the details of meiosis in students at the middle of the scale as well as many of the top students identifying with levels 2 and 3 of this construct. The misfit of V23 (construct G1) resulted from a tendency for some of the highest ability students to identify with the lowest progression levels. On all of these items, the higher ability students had some tendency to lack confidence in their responses (resulting in a score of '0').

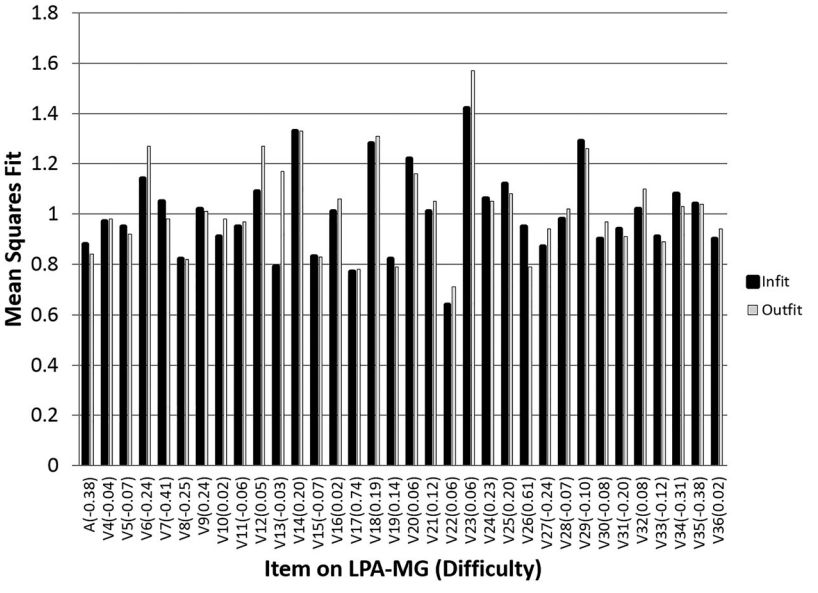

Figure 2. Mean squares fit indices for items on the LPA-MG. Misfit with the Rasch model is indicated by means squares fit >1.30.

How are students' profiles of understanding of modern genetics likely to change as they progress from low to high understanding?

Upon obtaining validity evidence for Version 2 of the LPA-MG instrument, we further analyzed the data to determine how understandings change as the students progress from low to high understanding. The median logit measure for students was 0.03, the first quartile logit measure was −0.19, the third quartile logit measure was 0.18, the minimum logit measure was −3.63, the maximum logit measure was 0.98, and the mean logit measure was −0.08. We made a map of the most probable progression levels in each construct for a student's particular logit measure (Figure 3) to get a more qualitative progression-based description of how college students' understandings change as logit measures by themselves tell little about what a student understands. The box-whisker plot at the top of the figure illustrates the logit measures (under the box-whisker plot from −3 to 3) of the median, quartiles, and some of the outliers in relation to the most probable progression level for each construct. The vertical lines in the map correspond to the logit measures of the minimum and maximum non-outliers and median. The numbers within each construct (between 0 and the maximum level for that construct) indicate the predicted level on the progression associated with a particular logit measure (Figure 3). For example, on construct A, a student with a logit measure of −1 would be predicted to be at a level 0 since this is below the 2 level. Similarly, our Rasch model predicts that a student with a logit measure of 0 would be at a level 5 in this construct since this measure sits above the 5 level but below the 6 level. A student with a logit measure of 1 is predicted to reside at the 6 level. The fact that the 1 and 4 levels are missing illustrates the Rasch model's prediction that these levels are not the most probable responses at any location

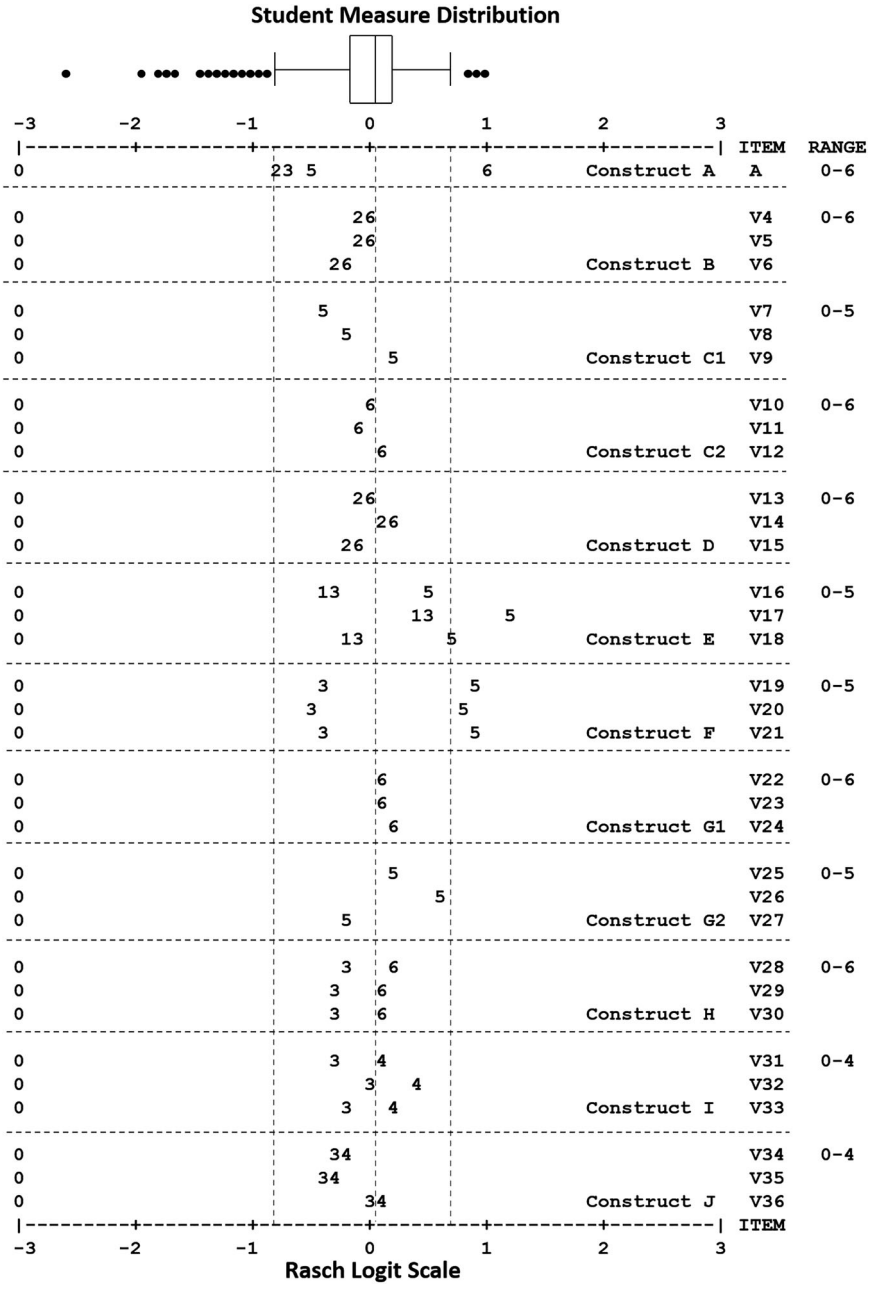

Figure 3. LPA-MG Version 2 Rasch measure for distributions and map of most probable progression levels with respect to students' logit measures. The box-whisker plot represents the range of the data, and the dots are outliers in the data.

along the logit scale in this sample of college-level students. In other words, while there is always a possibility of observing a student at the 1 and 4 levels, it is always more likely that the student will be observed at one of the other progression levels (0, 2, 3, 5, and 6).

Since we fit items measuring the same construct with the same rating scale model, we see 12 different rating scales for the 34 items (Figure 3), each grouped according to construct order (i.e. A (construct A), V4–6 (construct B), V7–9 (construct C1), etc.). The same rating scale model also means that we observed the same most probable progression levels for the items in the same construct; for example, construct B items (V4–6) show most probable progression levels of 0, 2, and 6. We did observe small differences in levels relative to logit measures among the items within the same construct despite using the same rating scale model for all the times within the construct, indicating that items probing the same idea (construct) tend to vary in difficulty. This variation is not surprising as individual items probing the same concept do function differently, but it is interesting that most items within a construct tended to have similar predicted response patterns relative to students' logit measures.

The minimum non-outlier logit measure (Figure 3, left whisker and left line) corresponded to a student with essentially no understanding of modern genetics ideas, as the most probable response for all items except for construct A at this logit measure was 0, which indicates no understanding of that idea. It was most probable for a student at the minimum logit measure to only understand how two of the following six concepts were related to each other: *gene*, *genome*, *chromosome*, *cell*, *nucleotide/base*, *DNA* (construct A, level 2). Interestingly, the maximum non-outlier logit measure (Figure 3, right whisker and right line) did not correspond to a expert-level understanding despite the instrument being linked to a progression intended for students in grades 5–10. The most probable response for construct A (genetic organization) was level 5, the most probable response for item V17 (construct E: genetic information is passed on to offspring) was level 3, and the most probable response for all three items in construct F (there are patterns of correlation between genes and traits (V19–21)) was a level 3 (Table S1). The most probable responses for all other items were at the maximum level for that construct. The median logit measure (Figure 3, line inside the box and middle line) showed varied levels of understanding: it was most probable for students at this logit measure to have the highest level understanding for constructs B, C1, C2, D, and J; the middle progression-level understanding for constructs A, E, F, H, and I; and the lowest progression-level understanding for constructs G1 and G2. This more descriptive analysis tied to the progression levels shows that college students do have a wide range of understandings consistent with the progression from complete novice to near expert, despite the learning progression being targeted to grades 5–10.

Discussion

The assessment and its progression

Although the student population is different from previous work with the LPA-MG (college students versus high school students) and was not included in the grade range of the progression on which the assessment was based, our data indicate that this measure and its learning progression framework is useful for understanding how undergraduate students think about concepts in genetics. Version 2 of the LPA-MG instrument contained revisions to nearly all of the items from Version 1 based on thorough reviews by university biology faculty (Table S2). Our data show that the instrument provides valid

and reliable measures for introductory college students. However, the psychometric numbers, while satisfactory, are not as strong as those from the use of Version 1 on our sample of high school students (Todd, Romine, & Cook Whitt, in press). For example, the reliability for Version 1 was 0.91 for measuring high school students while the reliability for Version 2 was 0.86 for measuring the college students in this study. All items on Version 1 fit well with the Rasch model when used to measure high school students. However, we found three items (V14, V18, and V23) on Version 2 that misfit with the college student data. Are these minor differences primarily due to the difference between tests or the difference between student populations? Given the thorough content review, we contend that Version 2 is indeed an improvement over Version 1 for the college environment as it is reasonable to hypothesize that the slightly deflated psychometric numbers are due to the relative heterogeneity of the introductory college biology students in comparison to students in a STEM school biology course undertaking a 23-week genetics intervention (Todd, Romine, & Cook Whitt in press). The slight deflation of reliability and the misfit of certain items with the Rasch model is reflective of the fact that responses from college students attending a large lecture course offered by an open-enrollment university are expectedly less predictable than responses from high school biology students measured at various time points across extended instruction targeted to a number of the modern genetics constructs measured in the assessment.

To the end of improving both measurement precision and validity of measures, we would like to take this opportunity to emphasize the importance of integrating confidence into OMC learning progression scales. In this version, integration of the confidence tier, or CRI (Romine, Schaffer, & Barrow, 2015), improved reliability of college students' measures from 0.61 to 0.86; we demonstrated a similar increase in reliability from CRI integration on Version 1 of the LPA-MG (Todd, Romine, & Cook Whitt, in press). Due to the inflated probability of obtaining apparently high levels along the progression from guessing, it makes sense that eliminating this unpredictability through the requirement of confidence leads to dramatic improvement in measurement reliability. We recommend that other studies which integrate learning progressions within assessment items using an OMC format consider doing the same.

This study is the first to quantitatively measure college students' knowledge of modern genetics with respect to a learning progression. Several college-level assessments exist, but none are tied to current learning progression theory. We used a population of students in an introductory biology course at a research-intensive university in the Midwest and found the assessment was appropriate for this population of college students. Future studies could extend this research to use the LPA-MG to assess college students' understandings at various types of institutions (i.e. community colleges, small liberal arts colleges, higher institutions that serve specific populations such as women's colleges or historically black universities) with varying student populations in a variety of locations (U.S. and non-U.S.) to explore differential functioning for these populations of students. Our study is limited in that we did not use the assessment as a pre/post to be able to determine potential learning gains as we did with Version 1 (Todd, Romine, & Cook Whitt, in press); rather we captured a snapshot in time. Now that we have determined that the LPA-MG instrument can reliably measure introductory biology college students' understandings, future research could be longitudinal in nature to measure learning gains and assess difficulties that students have with the content. To this end, the Rasch Rating

Scale Model used in this study demonstrates potential utility in extracting interesting qualitative information from quantitative measures of student growth in longitudinal contexts.

College students' understandings of the genetics learning progression

Despite the progression on which the instrument was based being targeted to students in grades 5–10, we chose to use our instrument to assess modern genetics knowledge in college introductory biology students. It is reasonable to speculate that a vast majority of these college students had taken a biology course previously in high school as a life science course is often required for high school graduation. These students were also currently enrolled in an introductory biology course addressing the topics of genetics and the molecular and cellular basis for the unity of life, including concepts assessed by our instrument, and completed the assessment toward the end of their time in the course. Hence we can deduce that introductory college students have been exposed to many of the concepts assessed by the LPA-MG at some point in their schooling. One important aspect of learning progression theory is that achievement of upper levels is not guaranteed even with targeted instruction and curriculum (Duncan & Hmelo-Silver, 2009). Our data provide further evidence that though the upper bounds of the progression may be aligned to performance expectations of students in grades 9–10, it is unreasonable to assume that students in higher grades (even college) successfully demonstrate this level of knowledge. Grade bands for performance expectations in learning progressions can be a useful tool for educators and instructors in designing the appropriate level of instruction for a particular topic, but they should be strongly cautioned not to assume that students at a particular grade band have mastered the performance expectations of the lower grade bands.

Our population of college students represented the gamut of modern genetics understanding from essentially no knowledge (Figure 3, left whisker and left line) to a nearly expert-level of knowledge (Figure 3, right whisker and right line). Expert-level knowledge of modern genetics entails not only understanding the three models of genetics (*molecular*, *genetic*, and *meiotic*) but how they integrate (Stewart et al., 2005); as such, the upper levels of some of the progression constructs represent learning performances that integrate multiple models. For example, the learning performance for level 5 of construct F is the understanding that alleles differ in sequence which affects proteins, leading to trait variations and that dominant/recessive relationships can be explained by protein interactions (Table S1). This construct level integrates ideas from the *genetic* and *molecular* models, which means to successfully achieve this level, students must not only understand the *genetic* and *molecular* models, but how they relate. Similarly with multiple constructs in a learning progression for a domain, there are contingencies and relationships between constructs (Plummer & Krajcik, 2010; Schwarz et al., 2009; Shea & Duncan, 2013; Wilson, 2009).

Expert-level understanding

Based on our previous qualitative work (Todd, 2013; Todd & Kenyon, 2015), knowledge of the domain and its literature, and data from this study and Todd, Romine, & Cook Whitt (in press), we now present a concept map which can be used tentatively as a framework showing the theoretical relationships and contingencies between constructs B and J of

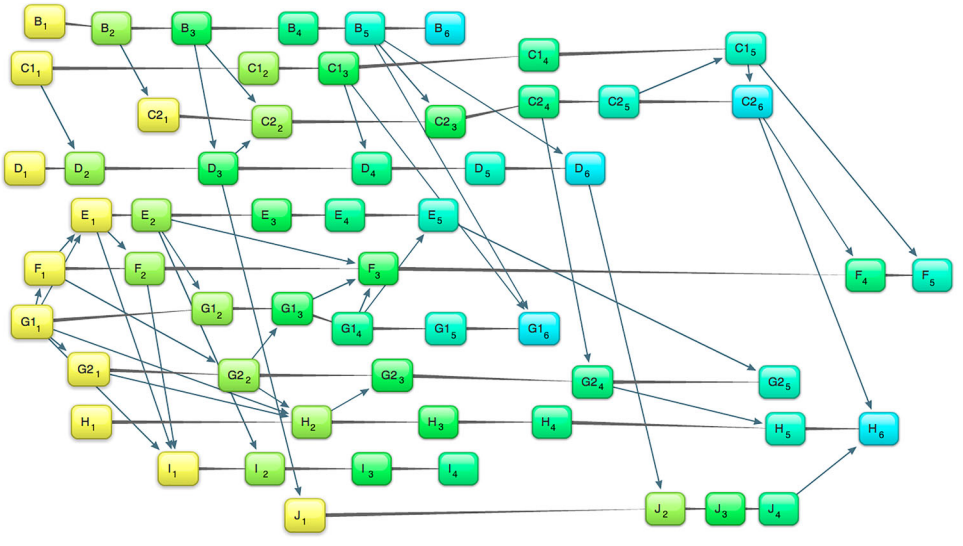

Figure 4. Hypothetical expert-level concept map of constructs B–J in the genetics learning progression. Each bubble has a label corresponding to its level in the progression (i.e. B_1 is level 1 of construct B; $G1_4$ is level 4 of construct G1). Condensed descriptions of connections between construct ideas are outlined in Table 2; a condensed description of each level is outlined in Table S1. Constructs are generally arranged vertically in order from B to J and horizontally from minimum level (left) to maximum level (right). Lines between bubbles indicate relationships between levels in the same construct where the line narrows going toward the more complex idea/higher level. Arrows between bubbles indicate a hypothetical relationship between ideas where understanding the concept at the tail of the arrow influences understanding the concept at the head of the arrow.

the genetics learning progression (Figure 4, Table 2). Though these relationships are based on prior research, the relationships themselves have not yet been tested; therefore they remain hypothetical until empirical testing can be done. We chose to exclude construct A from our concept map because current revisions to this construct changed learning performances to identifying increasingly more connections between the concepts of *gene*, *DNA*, *chromosome*, *nucleotide/base*, *cell*, and *genome*. A level 2 understanding of this construct indicates a correct relationship between two of any of these concepts – a relationship between *DNA* and *gene* would be at the same level as a relationship between *genome* and *cell*; however, these are two very different conceptual understandings. Since there could be multiple different conceptual understandings at each different level of construct A, we did not include it in our concept map.

Our hypothetical concept map of an expert-level understanding of modern genetics (Figure 4) is roughly organized vertically in construct order from top to bottom and horizontally in level order from left to right similarly to Figure 3. Lines between the bubbles indicate relationships between the different levels within the same construct (i.e. line between B_1 [construct B, level 1] and B_2 [construct B, level 2]) where the line narrows toward the higher level or more complex learning performance. We wish to emphasize that distances between concepts are not indicative of amount of 'conceptual leap' or difficulty; instead spacing was chosen with the goal of illustrating the connections between

Table 2. Description of hypothetical links between construct ideas.

Link Between Construct Ideas	Description
$B_2 \rightarrow C2_1$	'Genes are informational' supports 'changes to genes change traits'
$B_3 \rightarrow C2_2$	'Genes instruct the body at different levels' supports 'changes to genes change cells'
$B_3 \rightarrow D_3$	'Genes instruct the body at different levels' supports 'DNA tells cells to be different'
$B_5 \rightarrow C2_3$	'Genes code for proteins' supports 'changes to genes change proteins'
$B_5 \rightarrow D_6$	'Genes code for proteins' supports 'somatic cells have the same DNA to express different proteins'
$B_5 \rightarrow G1_6$	'Genes code for proteins' supports 'the more conserved DNA is between species, the more important the gene product'
$C1_1 \rightarrow D_2$	'Cells perform functions' supports 'cells are different because they have different functions'
$C1_3 \rightarrow D_4$	'Proteins do the cell's work' supports 'different cells have different proteins for their functions'
$C1_3 \rightarrow G1_6$	'Proteins do the cell's work' supports 'the more conserved DNA is between species, the more important the gene product'
$C1_5 \rightarrow C2_6$	'Protein structure and function depends on amino acids in the protein' supports 'changes to genes change protein functions to change traits'
$C1_5 \rightarrow F_5$	'Protein structure and function depends on amino acids in the protein' supports 'alleles differ in sequence which affects proteins to give trait variations, dominant and recessive relationships can be explained by protein interactions'
$C2_4 \rightarrow G2_4$	'Changes to genes change proteins to change traits' supports 'DNA changes can be beneficial, neutral, or harmful, and can change protein structure/function'
$C2_5 \rightarrow C1_5$	'Changes to genes change amino acids in proteins' supports 'protein structure and function depends on amino acids in the protein'
$C2_6 \rightarrow F_4$	'Changes to genes change protein functions to change traits' supports 'alleles differ in sequence which affects proteins to give trait variations'
$C2_6 \rightarrow H_6$	'Changes to genes change protein functions to change traits' supports 'environment can change genes which change proteins, or change gene expression of proteins'
$D_3 \rightarrow C2_2$	'DNA tells cells to be different' supports 'changes to genes change cells'
$D_3 \rightarrow J_1$	'DNA tells cells to be different' supports 'gene expression is not regulated or controlled, or does not change'
$D_6 \rightarrow J_2$	'Somatic cells have the same DNA to express different proteins' supports 'genes can be turned on during development'
$E_1 \rightarrow F_2$	'Organisms can only get traits of their parents' supports 'traits from parents can mix or compete to give offspring traits'
$E_1 \rightarrow I_1$	'Organisms can only get traits of their parents' supports 'a change of traits can be passed down to offspring'
$E_2 \rightarrow F_3$	'Offspring get half of their DNA from each parent' supports 'organisms get one allele per parent, and traits can be predicted'
$E_2 \rightarrow G1_2$	'Offspring get half of their DNA from each parent' supports 'organisms have different DNA'
$E_2 \rightarrow I_2$	'Offspring get half of their DNA from each parent' supports 'DNA mutations can be passed down to offspring'
$E_5 \rightarrow G2_5$	'Chromosomes can swap sections increasing genetic variation' supports 'DNA changes lead to increased genetic variation and evolution'
$F_1 \rightarrow E_1$	'Organisms have different versions of traits' supports 'organisms can only get traits of their parents'
$F_1 \rightarrow G2_2$	'Organisms have different versions of traits' supports 'organisms within a species look and function differently'
$F_2 \rightarrow I_1$	'Traits from parents can mix or compete to give offspring traits' supports 'a change of traits can be passed down to offspring'
$G1_1 \rightarrow E_1$	'Organisms have different traits or functions' supports 'organisms can only get traits of their parents'
$G1_1 \rightarrow F_1$	'Organisms have different traits or functions' supports 'organisms have different versions of traits'
$G1_1 \rightarrow G2_1$	'Organisms have different traits or functions' supports 'species look and function differently'
$G1_1 \rightarrow H_2$	'Organisms have different traits or functions' supports 'environment can affect traits or functions'
$G1_1 \rightarrow I_1$	'Organisms have different traits or functions' supports 'a change of traits can be passed down to offspring'
$G1_3 \rightarrow F_3$	'Organisms have different DNA even within a species' supports 'organisms get one allele per parent, and traits can be predicted'
$G1_4 \rightarrow E_5$	'Organisms within a species have both similar and different DNA' supports 'chromosomes can swap sections increasing genetic variation'

(*Continued*)

Table 2. Continued.

Link Between Construct Ideas	Description
$G1_4 \rightarrow F_3$	'Organisms within a species have both similar and different DNA' supports 'organisms get one allele per parent, and traits can be predicted'
$G2_1 \rightarrow H_2$	'Species look and function differently' supports 'environment can affect traits or functions'
$G2_2 \rightarrow G1_3$	'Organisms within a species look and function differently' supports 'organisms have different DNA even within a species'
$G2_2 \rightarrow H_2$	'Organisms within a species look and function differently' supports 'environment can affect traits or functions'
$G2_4 \rightarrow H_5$	'DNA changes can be beneficial, neutral, or harmful, and can change protein structure/function' supports 'environment can change type and amount of proteins that influence cell function'
$H_2 \rightarrow G2_3$	'Environment can affect traits or functions' supports 'changes to an organism can be beneficial or harmful'
$J_4 \rightarrow H_6$	'Gene expression can change at any point during one's life' supports 'environment can change genes which change proteins, or change gene expression of proteins'

concepts in a legible manner. Condensed descriptions of hypothetical links between construct ideas are described in Table 2; all levels are in Table S1.

Arrows between the bubbles indicate relationships between concepts where we posit that understanding the concept at the tail of the arrow influences understanding the concept at the head of the arrow (Figure 4, Table 2). For example, when a student understands that genes are informational (construct B level 2; denoted B_2), we hypothesize that they are better positioned to understand that changes to genes change traits (construct C2 level 1; denoted $C2_1$); the arrow going from B_2 to $C2_1$ illustrates this hypothesized relationship. If students know that genes contain information, they can apply this to understanding that changing this information can lead to changes in traits. As another example (Figure 4, Table 2), the idea that the environment can change genes which change proteins or can change gene expression of proteins (construct H level 6; denoted H_6) is influenced by the understanding that changes to genes leads to changes in proteins functions, which lead to changes in traits (construct C2 level 6; denoted $C2_6$), and the understanding that gene expression can change at any point during one's life (construct J level 4; denoted J_4).

We used a variety of sources to inform our hypothesized connections between concepts. For example, $B_3 \rightarrow D_3$ and $B_3 \rightarrow C2_2$ were mainly informed by our knowledge of the domain and previous data within our own research (Todd, 2013; Todd & Kenyon, 2015). The Roseman genetics learning progression (Roseman et al., 2006) contains hypothesized connections between concepts; as such, we used these to inform our connections including, but not limited to $E_2 \rightarrow F_3$. We also used our most probable response map (Figure 3) to support our connections. For example, the logit measure where a student most likely has a level 5 understanding of construct E is at a lower logit measure than a level 5 of construct G2, indicating that it takes more ability to achieve level 5 of G2 than level 5 of E. We hypothesized these ideas are related using our previous research and thus connected these ideas from $E_5 \rightarrow G2_5$.

We hypothesize that these relationships between concepts exist, but want to be clear that students may not make all of these connections as they progress in their understanding. The numerous connections between concepts in Figure 4 indicate that the constructs are quite related to each other, providing evidence for the validity of defining genetics literacy in terms of being able to understand and integrate the three inter-related models (Stewart et al., 2005) as well as rationale for teaching the three models concurrently

(Duncan et al., 2009). Indeed, our finding that the LPA-MG provides a unidimensional scale corroborates this idea.

We also attempted to roughly order the levels of each construct in the hypothetical concept map in terms of the connections between the concepts. All the arrows point from left to right under the hypothesis that knowledge of certain concepts (i.e. $G1_1$, organisms have different traits or functions) influences knowledge of more complex concepts (i.e. F_1, organisms have different versions of traits; $G2_1$, species look and function differently; E_1, organisms can only get traits of their parents). We want to note that the overall left to right order of the concepts is rough in that we hypothesize knowledge of $G1_1$ helps students better understand or achieve learning performance F_1, but students may not have to understand $G1_1$ to understand F_1. Again, we wish to note that the distances between concepts are not indicative of the amount of 'conceptual leap' between concepts, but are due in part by keeping arrows pointing toward the right illustrating how concepts may influence one another, thus keeping the map legible.

There is also some uncertainty as to whether or not knowledge of D_1 and other concepts not hypothesized to be related to these occurs at the same time, before, or after knowledge of $G1_1$. Good fit of the progression, and individual items, with the Rasch model indicates that the data support the ordering of the hypothesized progression underlying each item. However, we also find that not all levels of the progression are likely to be expressed within a particular ability range. For example, we are most likely to see levels 0, 2, and 6 expressed for construct D. This does not mean that the other progression levels (1, 3, 4, and 5) are not expressed in the data, only that we are more likely to see 0, 2, and 6. And though our concept map contains many connections between concepts, that is not to say that students can and will make all of these connections between the various concepts. Our hypothetical concept map of an expert-level understanding of modern genetics models how the various constructs may be related to and influence each other, which is one of the goals of learning progression refinements (Shea & Duncan, 2013).

Findings from this study support our hypothetical concept map and the potential connections between the concepts. In the following sections, we use the expert-level concept map as a guide to explain the hypothesized concept patterns of a student at the median logit measure (average-level understanding) and the most probable understandings of a student at the first quartile logit measure (lower-level understanding).

Average-level understanding

Using our concept map, we will more qualitatively describe the most probable understandings of a student at the median logit measure (Figure 3, line in box and middle line) as an example of what the average college student in our sample likely understands about modern genetics. Since we had three assessment items for each construct, we used the most frequent probable level for each construct at the median student logit measure to construct a concept map that illustrates the most probable responses for each construct (Figure 5). The levels above the most probable response of the median student were removed (i.e. E_{4-5}) and constructs were completely removed if the most probable response was a 0 (constructs G1 and G2). Hypothesized connections between concepts remained if both concepts were present and were removed if one or both of the concepts were absent.

As may be expected based on research on expert-novice differences (Hmelo-Silver & Pfeffer, 2004), the most striking contrast between the expert-level concept map (Figure

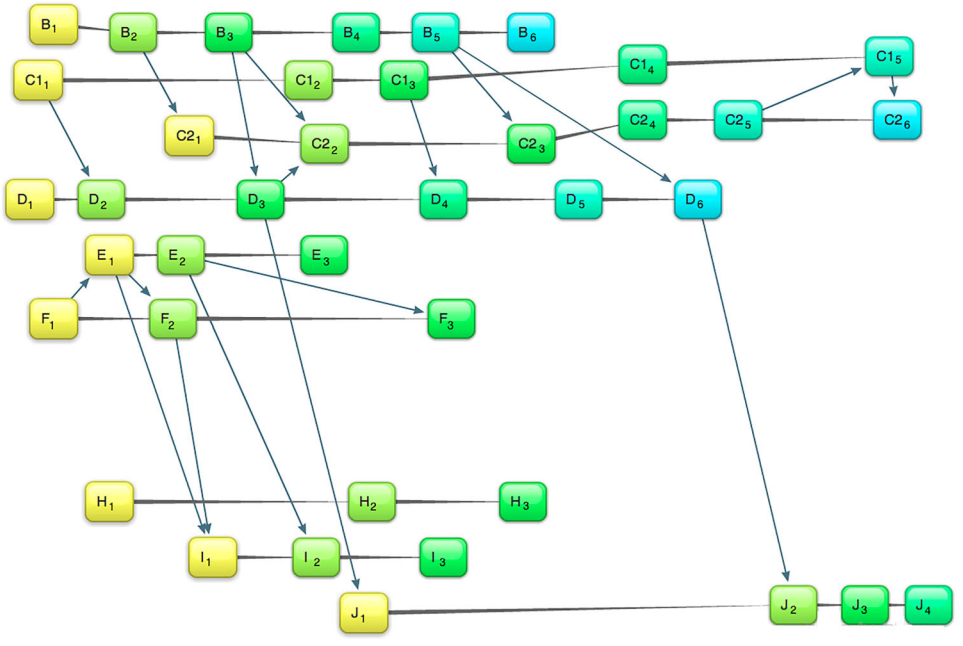

Figure 5. Hypothetical concept map of student at the median logit measure. Using the expert-level concept map as a guide (Figure 4), the levels above the most probable response for a student at the median logit measure (Figure 3) for each construct were removed. Connections between concepts were retained if both concepts were present but removed if one or both of the concepts were not present.

4) and the median student concept map (Figure 5) is the decrease in the amount of hypothesized connections between the concepts. The decrease in connections essentially yields two different clusters of inter-related ideas. The top cluster contains ideas in the *molecular* model of genetics that deal with proteins and their functions. The average student likely has a good understanding of proteins, how proteins connect genes to traits, and how cells express different DNA to produce the proteins needed for their functions (constructs B, C1, C2, and D). The average student also likely has connections between these concepts and the idea that gene expression can change during an organisms' life (construct I). These concepts are likely to be very inter-related for the average student. Unlike our average student, Shi et al. (2010) found that their population of introductory molecular and cell biology students had difficulties with the concept of gene expression. However, their items for this concept probe a much more detailed and mechanistic understanding of gene expression including the role of promoters, while our expert-level understanding for the gene expression construct (construct D) simply consists of the understanding that somatic cells have the same DNA to express different proteins (Table S1, D_6). Given that our population of students performed well on this construct, it may be reasonable for future revisions of the learning progression to include additional higher levels to this construct, including the mechanistic details of gene expression.

The middle cluster contains ideas in the *genetic* and *meiotic* models of genetics that discuss patterns of inheritance. The average student likely understands how to do

Punnett squares (F_3), but sorts alleles independently of chromosomes without consideration of the details of meiosis (E_3). The average student also likely has connections between these concepts and the concept that mutations can be passed to offspring (construct I). It is very interesting that we hypothesize that a median student makes no connections between these two clusters of the more *molecular* ideas (constructs B, C1, C2, D, and J) and the more *genetic* and *meiotic* ideas (constructs E, F, and I). The average student likely has an understanding that the environment can affect our cells (H_3), but we hypothesize that the average student tends to make no connections between this concept and the other concepts. The average student also likely has no understanding of how DNA varies between and within species (construct G1) or, similar to Smith et al. (2008) findings, how DNA changes lead to increased variation and evolution of a species (construct G2).

The clustering seen in Figure 5 indicates the average college student in our sample is likely unable to make connections between the *molecular* model of genetics with the *genetic* and *meiotic* models, meaning they cannot use their knowledge of proteins and their functions to explain patterns of inheritance or understand how DNA conserved between species codes for important gene products. While these students may understand protein functions and patterns of inheritance separately, they do not see the connections. Similarly, the average student likely understands that the environment can affect our cells (H_3), but likely does not understand how the environment can change gene expression (H_6) or how genetic changes can drive evolution of a species (connection with constructs G2 and C2). They see modern genetics as less unified than an expert-level student may (Figure 4).

Lower-level understanding

Similar to the average-level understanding, the most probable understandings of a student at the first quartile (Figure 3, left side of box) represent fragmented clusters of ideas and few connections between concepts. The concept map of a student at this lower-level understanding (logit measure at the first quartile) is shown in Figure 6. This concept map essentially shows four different clusters of ideas: (1) construct C1; (2) constructs E, F, and I; (3) construct H; and (4) construct J.

A student at the first quartile logit measure likely has a good understanding about what proteins do ($C1_5$), but does not understand that genes code for proteins (concept B is missing), so is unable to describe that how proteins connect genes and traits (concept C2 is missing), similar to findings by Marbach-Ad (2001). A low-level student likely understands that the environment can affect cells/organs/tissues (H_3), but does not understand that the environment can change entities inside the cell such as genes, proteins, or protein expression (higher levels of construct H) consistent with the lack of knowledge about what genes do (constructs B and C2) and gene expression (construct D). Interestingly, low-level students likely do understand that gene expression can change at any point during one's life (J_4) but our concept map shows no connections between this concept and others, consistent with the student having no knowledge of how cells express different genes (construct D) or that the environment can alter gene expression (higher levels of H). This may indicate that low-level students believe that gene expression may change, but do not understand how it can be altered or the outcomes of such alteration.

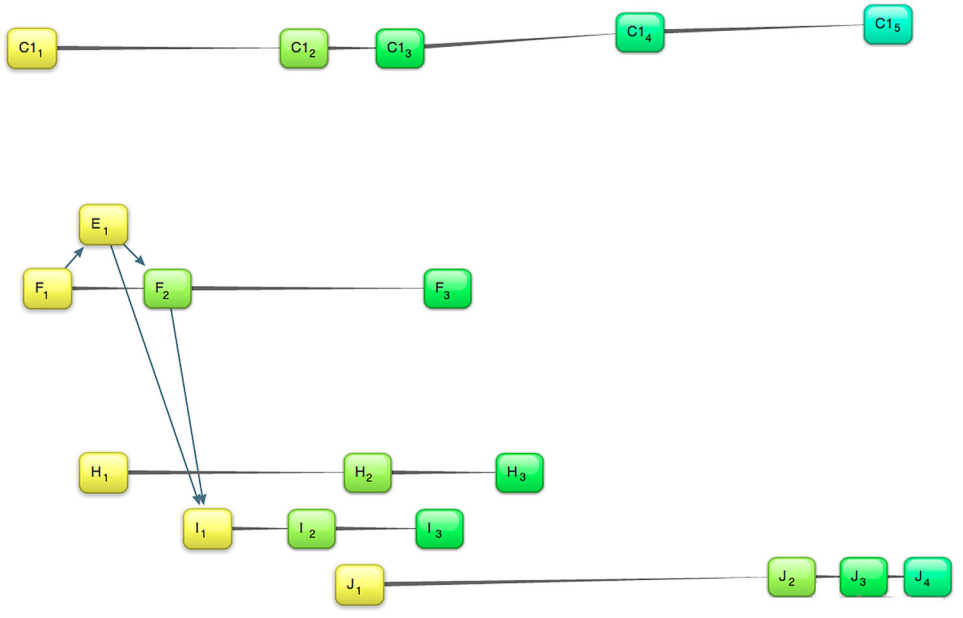

Figure 6. Hypothetical concept map of student at the first quartile logit measure. Using the expert-level concept map as a guide (Figure 4), the levels above the most probable response for a student at the first quartile logit measure (Figure 3) for each construct were removed. Connections between concepts were retained if both concepts were present but removed if one or both of the concepts were not present.

The potential connections between concepts of a low-level student are those between constructs E, F, and I. Low-level students are predicted to have a good understanding of how to construct Punnett squares (F_3) and that traits are inherited and mixed from parents in offspring (E_1, F_2) – the *genetic* model – and may use this knowledge to understand that a change in traits can be passed to offspring (I_1) and that mutations in gametes can be passed to offspring (I_3). Aside from these potential connections, low-level students likely have very disconnected knowledge. Similar to the average-level students, there are no hypothetical connections between ideas in the *molecular* model and ideas in the *genetic* or *meiotic* models, indicating that these students cannot integrate these models. Our data suggest that a low-level student likely has poor understanding of the *molecular* model, no understanding of the *meiotic* model, but a working knowledge of the *genetic* model (F_3). Our concept map indicates low-level students conceptualize modern genetics as even more disjointed clusters of unrelated concepts in comparison to average-level and expert-level students.

Though we used a learning progression targeted to students in grades 5–10 to assess modern genetics understanding in college students, we found that the students had levels of understanding that were captured well by the progression. Our student population represented a range in understanding from novice to expert, indicating that there is a great deal of variation in genetics knowledge in introductory biology courses. Using the most probable response map (Figure 3), we found that the average student in our sample (represented by the median logit measure) likely had good knowledge of the

molecular and *genetic* models but a poor understanding of the *meiotic* model as well as a poor understanding of how these models relate to each other. A low-level student in our sample (first quartile logit measure) likely had poor knowledge of the *molecular* model, no knowledge of the *meiotic* model, working knowledge of the *genetic* model, and no understanding of how these models relate to each other.

As genetics literacy consists of being able to understand and integrate these models (Stewart et al., 2005) and *Vision and Change* (AAAS, 2011) calls for college students to understand how entities at the molecular level drive evolution and macroscopic structures and functions, our data suggest that introductory biology should focus genetics instructional time on the process of meiosis and how the *molecular*, *genetic*, and *meiotic* models fit together. Instructors could build upon students' apparent knowledge of the *molecular* model of genetics (proteins and their functions) to make connections between that model and the *genetic* and *meiotic* models such as how protein interactions can account for dominant/recessive relationships.

Conclusion

We present Version 2 of our LPA-MG and demonstrate that it is a valid, reliable, and unidimensional instrument that can assess understandings of introductory biology students. Our data suggest that the assessment was written at an appropriate level for our population of college students and revealed that the average student in our sample likely had a good understanding of the *molecular* and *genetic* models of molecular genetics, but a poor understanding of the *meiotic* model and how the three models relate. We also proposed a hypothetical concept map of an expert-level understanding of modern genetics and used the concept map to illustrate the likely understandings of a student at the median logit measure and a student at the first quartile logit measure. Though based on data from three prior studies, the links between constructs in the concept map remain hypothetical until empirical testing is done. We encourage other researchers to empirically examine our hypothetical relationships between concepts and to further explore the validity of the LPA-MG instrument in novel contexts.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCiD

Amber Todd http://orcid.org/0000-0002-5091-8800

Notes on contributors

Amber Todd is a postdoctoral research fellow in the Department of Biological Sciences at Wright State University in Dayton, Ohio.

William Romine is an Assistant Professor in the Department of Biological Sciences at Wright State University in Dayton, Ohio.

References

AAAS. (2011). *Vision and change in undergraduate biology education*: *A call to action*. Washington, DC: Author.

Alonzo, A. C. (2012). Eliciting student responses relative to a learning progression. In A. C. Alonzo & A. W. Gotwals (Eds.), *Learning progressions in science: Current challenges and future directions* (pp. 241–254). Rotterdam: Sense.

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, *43*(4), 561–573.

Boone, W. J., Townsend, J. S., & Staver, J. (2011). Using Rasch theory to guide the practice of survey development and survey data analysis in science education and to inform science reform efforts: An exemplar utilizing STEBI self-efficacy data. *Science Education*, *95*(2), 258–280.

Bowling, B. V., Acra, E. E., Wang, L., Myers, M. F., Dean, G. E., Markle, G. C., … Huether, C. A. (2008). Development and evaluation of a genetics literacy assessment instrument for undergraduates. *Genetics*, *178*, 15–22.

Briggs, D. C., Alonzo, A. C., Schwab, C., & Wilson, M. (2006). Diagnostic assessment with ordered multiple-choice items. *Educational Assessment*, *11*(1), 33–63.

Corcoran, T. B., Mosher, F. A., & Rogat, A. D. (2009). *Learning progressions in science*: *An evidence-based approach to reform*. New York, NY: Columbia University, Teachers College, Center on Continuous Instructional Improvement.

Couch, B. A., Wood, W. B., & Knight, J. K. (2015). The Molecular Biology Capstone Assessment: A concept assessment for upper-division molecular biology students. *CBE-Life Sciences Education*, *14*(1), ar10.

Dikmenli, M., Cardak, O., & Kiray, S. A. (2011). Science student teacher's ideas about the 'gene' concept. *Procedia Social Behavioral Sciences*, *15*, 2609–2613.

Donovan, J., & Venville, G. (2014). Blood and bones: The influence of the mass media on Australian primary school children's understandings of genes and DNA. *Science & Education*, *23*, 325–360.

Dougherty, M. J. (2009). Closing the gap: Inverting the genetics curriculum to ensure an informed public. *The American Journal of Human Genetics*, *85*, 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., & Hmelo-Silver, C. E. (2009). Learning progressions: Aligning curriculum, instruction, and assessment. *Journal of Research in Science Teaching*, *46*(6), 606–609.

Duncan, R. G., & Reiser, B. J. (2007). Reasoning across ontologically distinct levels: Students' understandings of molecular genetics. *Journal of Research in Science Teaching*, *44*(7), 938–959.

Duncan, R. G., Rogat, A., & 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.

Elmesky, R. (2013). Building capacity in understanding foundational biology concepts: A K-12 learning progression in genetics informed by research on children's thinking and learning. *Research in Science Education*, *43*(3), 1155–1175.

Elrod, S. L. (2007). *Genetics concept inventory*. Retrieved July 14, 2015, from http://bioliteracy.colorado.edu/Readings/papersSubmittedPDF/Elrod.pdf.

Fisher, K. M. (1992). Improving high school genetics instruction. In M. U. Smith & P. E. Simmons (Eds.), *Teaching genetics: Recommendations and research proceedings of a national conference* (pp. 24–28). Cambridge, MA: National Science Foundation.

Gericke, N. M., & Smith, M. U. (2014). Twenty-first-century genetics and genomics: Contributions of HPS-informed research and pedagogy. In Michael R. Matthews (ed.), *International Handbook of Research in History, Philosophy and Science Teaching* (pp. 423–467). Netherlands: Springer.

Gericke, N. M., & Wahlberg, S. (2013). Clusters of concepts in molecular genetics: A study of Swedish upper secondary science students' understanding. *Journal of Biological Education*, *47*(2), 73–83.

Gollust, S. E., Wilfond, B. S., & Hull, S. C. (2003). Direct-to-consumer sales of genetic services on the internet. *Genetics in Medicine*, *5*(4), 332–337.

Gotwals, A. W., & Songer, N. B. (2010). Reasoning up and down a food chain: Using an assessment framework to investigate students' middle knowledge. *Science Education*, *94*, 259–281.

Hallden, O. (1990). Questions asked in common sense contexts and in scientific contexts. In P. L. Lijnse, P. Licht, W. de Vos, & A. J. Waarlo (Eds.), *Relating macroscopic phenomena to microscopic particles* (pp. 119–130). Utrecht: CD-ß Press.

Hmelo-Silver, C. E., & Pfeffer, M. G. (2004). Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors, and functions. *Cognitive Science*, *28*(1), 127–138.

Howitt, S., Anderson, T., Costa, M., Hamilton, S., & Wright, T. (2008). A concept inventory for molecular life sciences: How will it help your teaching practice. *Australian Biochemist*, *39*(3), 14–17.

Hull, S. C., & Prasad, K. (2001). Reading between the lines: Direct-to-consumer advertising of genetic testing in the USA. *Reproductive Health Matters*, *9*(18), 44–48.

Klymkowsky, M. W., Underwood, S. M., & Garvin-Doxas, R. K. (2010). *Biological concepts instrument (BCI): A diagnostic tool for revealing student thinking*. arXiv:1012.4501v1.

Knippels, M. C. P. J. (2002). *Coping with the abstract and complex nature of genetics in biology education: The yo-yo teaching and learning strategy* (Doctoral dissertation). Utrecht University, Utrecht, the Netherlands. Utrecht: CD-ß Press.

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.

Lewis, J., Leach, J., & Wood-Robinson, C. (2000). All in the genes? – Young people's understanding of the nature of genes. *Journal of Biological Education*, *34*(2), 74–79.

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

Linacre, J. M., & Tennant, A. (2009). More about critical eigenvalue sizes in standardized-residual principal components analysis (PCA). *Rasch Measurement Transactions*, *23*(3), 1228.

Marbach-Ad, G. (2001). Attempting to break the code in student comprehension of genetic concepts. *Journal of Biological Education*, *35*(4), 183–189.

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

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, *47*(2), 149–174.

McElhinny, T. L., Dougherty, M. J., Bowling, B. V., & Libarkin, J. C. (2014). The status of genetics curriculum in higher education in the United States: Goals and assessment. *Science & Education*, *23*(2), 445–464.

NGSS Lead States. (2013). *Next Generation Science Standards: For the states, by the states*. Achieve, Inc. on behalf of the twenty-six states and partners that collaborated on the NGSS. Retrieved from http://www.nextgenscience.org/next-generation-science-standards.

Plummer, J. D., & Krajcik, J. (2010). Building a learning progression for celestial motion: Elementary levels from an earth-based perspective. *Journal of Research in Science Teaching*, *47*(7), 768–787.

Rogat, A., Anderson, C. A., Foster, J., Goldberg, F., Hicks, J., Kanter, D., … Wiser, M. (2011). *Developing learning progressions in support of new science standards: A RAPID workshop series*. Consortium for Policy Research in Education. Retrieved from http://eric.ed.gov/?id=ED536834.

Romine, W. L., Schaffer, D. L., & Barrow, L. (2015). Development and Application of a Novel Rasch-based Methodology for Evaluating Multi-Tiered Assessment Instruments: Validation and utilization of an undergraduate diagnostic test of the water cycle. *International Journal of Science Education*, *37*(16), 2740–2768.

Roseman, J., Caldwell, A., Gogos, A., & Kurth, L. A. (2006). *Mapping a coherent learning progression for the molecular basis of heredity*. Paper presented at the Annual Meeting of the National Association of Research in Science Teaching, San Francisco, CA.

Sadler, T. D., & Zeidler, D. I. (2005). The significance of content knowledge for informal reasoning regarding socioscientific issues: Applying genetics knowledge to genetic engineering issues. *Science Education*, *89*, 71–93.

Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Archer, A., Fortus, D., … Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, *46*(6), 632–654.

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

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.

Shi, J., Wood, W. B., Martin, J. M., Guild, N. A., Vicens, Q., & Knight, J. K. (2010). A diagnostic assessment for introductory molecular and cell biology. *CBE-Life Sciences Education*, *9*, 453–461.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological assessment*, *8*(4), 350.

Smith, L. A., & Williams, J. M. (2007). 'It's the X and Y thing': Cross-sectional and longitudinal changes in children's understanding of genes. *Research in Science Education*, *37*, 407–422.

Smith, M. K., Wood, W. B., & Knight, J. K. (2008). The genetics concept assessment: A new concept inventory for gauging student understanding of genetics. *CBE-Life Sciences Education*, *7*, 422–430.

Stevens, S. Y., Delgado, C., & Krajcik, J. S. (2010). Developing a hypothetical multi-dimensional learning progression for the nature of matter. *Journal of Research in Science Teaching*, *47*(6), 687–715.

Stewart, J., Cartier, J. L., & Passmore, C. M. (2005). Developing understanding through model-based inquiry. In M. S. Donovan & J. D. Branford (Eds.), *How students learn* (pp. 515–565). Washington, DC: National Research Council.

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

Todd, A. N. (2013). *The Molecular Genetics Learning Progressions: Revisions and Refinements Based on Empirical Testing in Three 10th Grade Classrooms* (Doctoral dissertation, Wright State University).

Todd, A., & Kenyon, L. (2015). Empirical refinements of a molecular genetics learning progression: The molecular constructs. *Journal of Research in Science Teaching*. doi:10.1002/tea.21262.

Todd, A., Romine, W. L., & Cook Whitt, K. (in press). Development and validation of the Learning Progression-based Assessment of Modern Genetics (LPA-MG) in a high school context. *Science Education*. doi:10.1002/sce.21252

Tsui, C., & Treagust, D. F. (2010). Evaluating secondary students' scientific reasoning in genetics using a two-tier diagnostic instrument. *International Journal of Science Education*, *32*, 1073–1098.

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.

Weems, G. H., Onwuegbuzie, A. J., Schreiber, J. B., & Eggers, S. J. (2003). Characteristics of respondents who respond differently to positively and negatively worded items on rating scales. *Assessment & Evaluation in Higher Education*, *28*(6), 587–606.

Wilson, M. (2009). Measuring progressions: Assessment structures underlying a learning progression. *Journal of Research in Science Teaching*, *46*, 716–730.

Wright, T., & Hamilton, S. (2008). Assessing student understanding in the molecular life sciences using a concept inventory. *ATN Assessment*, *8*, 216–224.

Wynne, C. F., Stewart, J., & Passmore, C. (2001). High school students' use of meiosis when solving genetics problems. *International Journal of Science Education*, *23*(5), 501–515.

Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching*, *39*(1), 35–62.