Genetics Concepts Inventory (GenCI) Development

Susan Elrod, Ph.D.
October 15, 2007
Biological Sciences Department, California Polytechnic State University,
San Luis Obispo 93407 (selrod@calpoly.edu) 805/756-2875

Abstract

A genetics concept inventory (GenCI) is under development as a tool to measure students understanding in different learning environments. The motivation for developing the GenCI is as one of several assessment strategies for evaluating an innovative literature-based case study (LBCS) approach for teaching introductory college-level genetics at California Polytechnic State University, San Luis Obispo (Cal Poly). The initial development of the GenCI and the LBCS approach are described in a paper presented at the Conceptual Assessment in Biology workshop (1). The LBCS approach also incorporates a significant scientific information literacy component (2). This paper outlines the main concepts (i.e., big ideas) in genetics, the further development and use of the GenCI assessment tool and the challenges associated with its development and use.

Big Ideas in Genetics

The big ideas in genetics that were used to develop some of the GenCI questions are summarized in Table 1. There are four main categories: 1) nature of the genetic material, 2) gene expression and regulation, 3) transmission, and 4) variation and evolution. Current GenCI questions are focused primarily in areas 1 and 3, and additional questions in areas 2 & 4 are in development.

Table 1. Major Concepts in Genetics[1]

Major Concept Area	Information Content (Structure)	Information Flow (Process)
1.0 Nature of Genetic Material	Genes are the basic unit of biological inheritance Biological inheritance information is located in the chromosomes, which	Genes contain the information for products, such as proteins (polypeptides) or RNA molecules, that contribute to the determination of traits within cells and

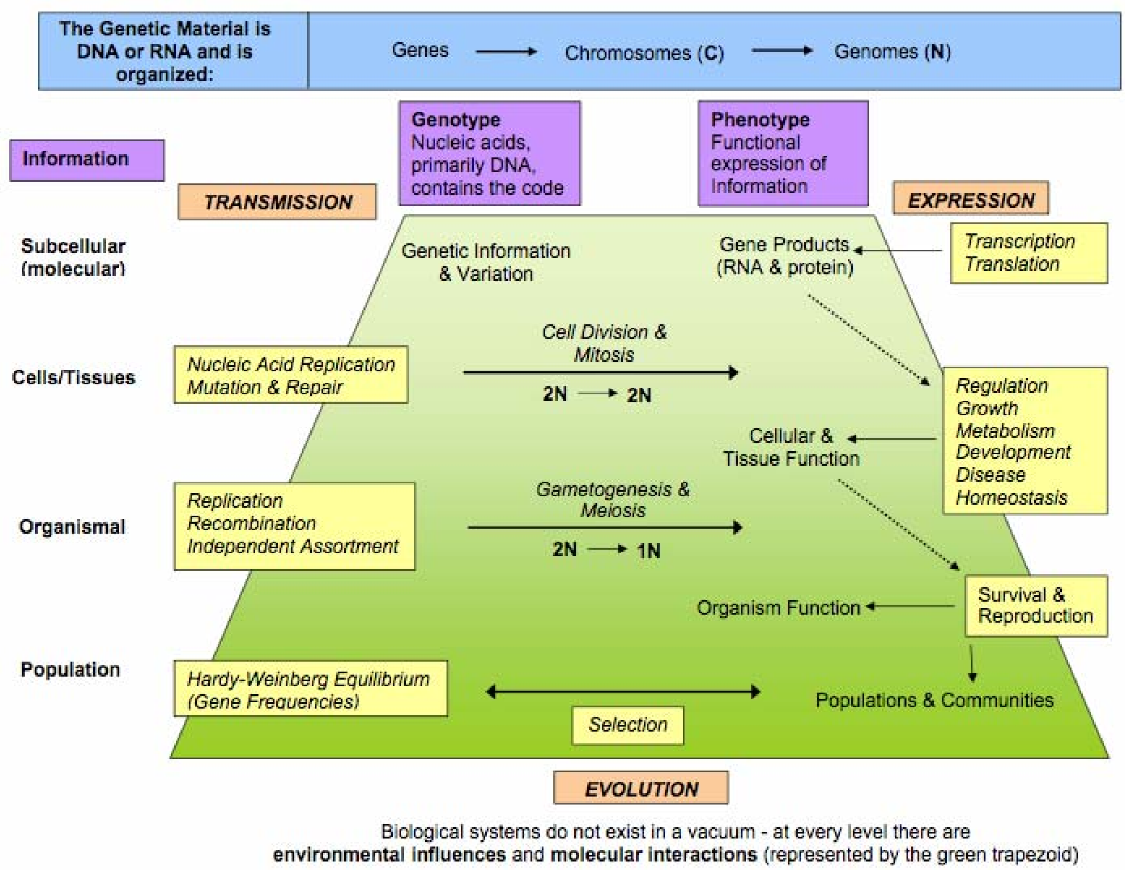
	are made of DNA and proteins and located in nucleus of eukaryotic cells; all cells have chromosomes (all living organisms are made up of cells) DNA is composed of deoxyribonucleotides (ACGT); RNA of ribonucleotides (ACGU); nucleotides are composed of a 5-carbon sugar, base, phosphate group; proteins are comprised of amino acid subunits Chromosomes in diploid organisms are grouped in pairs, which called homologous chromosomes; homologous chromosomes contain gene pairs, located in the same position on each homologue (locus); mitochondria and chloroplasts contain their own DNA There are two types of chromosomes, autosomal and sex (e.g., X and Y); all animal cells contain sex chromosomes; eggs of mammals contain a single X and sperm can contain either an X or a Y	organisms; enzymes are an example of one class of proteins; enzymes catalyze chemical reactions within cells; chemical reactions result in the production of molecules used for various cellular processes (e.g., amino acids used to make proteins); abnormal gene expression can lead to disease (e.g., cancer); one gene, one gene product (not enzyme, protein, polypeptide) Most traits are determined by the action of many different genes, as well as the interaction of gene products with their environment; genes and the environment interact to determine phenotype
2.0 Gene Expression and Regulation	The information within DNA molecules contains the code used by cells to produce RNA and/ or protein molecules The genetic code refers to a triplet sequence of DNA that codes for a specific amino acid during protein synthesis; one codon specifies one amino acid in a polypeptide chain The physical composition of an organism is called the phenotype; genetic composition called the genotype	The central dogma describes how genetic information functions within the cells of an organism; transcription results in the production of RNA; translation results in the production of proteins Inside cells, different genes are expressed over the course of development, in different cell types and in response to environmental stimuli specifying the cellular type, shape and function
3.0 Transmission	Somatic cells of an organism carry the same number of chromosomes and the same inheritance information Gametes carry half the chromosomes and, consequently half the inheritance information; the combination of alleles (or inheritance information) in each gamete is unique and different from the parents	Meiosis is the process by which genetic information is passed from generation to generation in sexually reproducing organisms; however, mitosis (asexual reproduction) is used by some organisms as a means of reproduction Each homologous chromosome is inherited from one of the diploid organism's parents'; homologous chromosomes separate from one another during meiosis (Principle of Segregation); different chromosomes (and thus genes and alleles) will separate independently of one another during meiosis (Principle of Independent Assortment), however, genes (alleles) located close to one another on a chromosome are linked and more are likely to be inherited together

		than those farther apart or on another chromosome; recombination disrupts linkage relationships A zygote is formed by the process of fertilization of two gametes in both plants and animals Plants also use mitosis and meiosis for production of daughter cells and gametes, respectively; bacteria reproduce by a process known as fission which is different from mitosis as well as meiosis
4.0 Variation and Evolution	Mutations are heritable changes in the sequence of DNA; can involve single or a few DNA base pairs or large regions of chromosomes; spontaneous or environmental agents (chemicals, radiation) Genes may take on different forms, which are called alleles; alleles are composed of slightly different sequences of DNA and encode gene products of different sequence as determined by the changes in DNA; gene products from different alleles may have drastically or moderately different function For a given gene, diploid organisms can be homozygous (same alleles) or heterozygous (different alleles) Diploid organisms may contain only two different alleles for any given gene at a specified locus Genes and environment together are responsible for the phenotypic variation between species	Processes that cause individuals of a species to be genetically unique: recombination, independent assortment, and mutation Allelic relationships can be complete dominance and recessive, incomplete dominance, semi-dominant, co-dominant; dominance means that trait is observed in the phenotype of a diploid organism regardless of the nature of the homologous allele; recessive means that trait is not observed in a diploid organism if the homologous allele is dominant Cells contain mechanisms for repairing damage to DNA Natural selection acts upon genetic variation within populations resulting in changes in gene frequencies, and thus evolution

[1] Major conceptual areas were identified and the information content and flow were elaborated. Concept descriptions were modified from 3- 11 based on a framework developed by the Michigan State University group (12).

Hott *et al.*, (2) have described six genetics conceptual areas in the context of an introductory biology course for non-science majors: 1) the nature of the genetic material, 2) transmission, 3) gene expression, 4) gene regulation, 5) evolution and 6) genetics and society. A visual representation of genetics concepts was developed using a variety sources to enable better understanding of the relationships between major concepts (Figure 1).

Figure 1. Genetic Concepts



This figure was developed in discussion with Joyce Parker, John Merrill, Brett Merritt and Merle Heideman (Michigan State University) and Mike Klymkowsky and Kathy Garvin-Doxas (University of Colorado, Boulder).

The Genetics Concept Inventory (GenCI): Results and Use

The GenCI was given in its original form as a pre- and post-test to upper division genetics students during Spring 2006 (Adv I). A revised version was given as a post-test in a lower division cell and molecular biology course during Fall 2006 (Intro) and as a pre-test in an upper division genetics course during Fall 2007 (Adv II). It will be given again as a post-test in the same Fall 2007 genetics course. Overall results are shown in Table 2. The Adv II students had the entire first week of class to take the test, which may explain why their average score was slightly higher. A recommendation for future pre-tests is to have students take it before the class period.

Table 2. GenCI statistics

Test	Average	Standard Deviation	Range (High/Low)
Pre-test Adv I (n=49)	50% (38/76 points)	11	64/16
Post-test Adv I (n=32)	65% (49/76 points)	9	68/25
Intro Post-test (n=107)	63% (25/40 points)	5.5	36/10
Pre-test Adv II (n=50)	59% (23/39 points)	5.7	35/14

Adv I = Version 1.0 given in Genetics course during Spring 2006 contained 38 questions each worth 2 points; Intro and Adv II = Version 2.1 given in Intro cell and molecular course in Fall 2006 contained 40 questions each worth 1 point.and in Genetics course during Fall 2007 contained 39 questions each worth 1 point.

Analysis of Selected GenCI Questions

A few GenCI questions are highlighted in this analysis, results of which are shown in ascending order of version 1.0 in Table 3. The question numbers changed between version 1.0 and 2.1, and some questions changed within version 2.1 as indicated in the data tables and discussion.

Table 3. GenCI questions of interest

Q # (A)	***Q # (B)***	***Question Content***	***Intro post %***	***Adv II Pre %***	***Adv I Pre %***	***Adv I Post %***
16	8	Nature of allele	37	46	31	81
26	29	Relationship of genetic information in egg and ovary cell after oogenesis	41	44	49	50
29	12*	Cell types that contain genetic information for eye color	73	34	20	52
31	19	Nature of genetic recombination	22	18	42	75
34	34	Products of translation are proteins	31	52	37	75
35	33**	Products of transcription	77	18	59	88

*A= Adv I GenCI (version 1.0); B = Intro & Adv II (version 2.1)****; * = question revised from 1.0 to 2.1; ** = question revised from Intro to Adv II***

Question 16/8 asks students to identify the correct response to the question, “what is an allele?” Less than half the Intro, Adv II pre and Adv I pre students choose “an alternate form of a gene.” Other answers chosen are “part of a gene,” “one of two genes in a diploid organism,” or “a region of chromosome where a gene is located” (Table 3). It may be easy to see why they would choose "one of two genes in a diploid organism" but interviews will help us better understand this result. A corollary question that asks students what the maximum number of alleles of a single gene that can exist within a diploid organism with 6 chromosomes reveals that students may not fully understand the meaning of the word allele: a minority 20% of Adv II students chose the correct answer (2), while the majority of students chose 12 (58%).

There are several questions that deal with the results of chromosome and genetic information consequences of mitosis and meiosis. Students show a high degree of understanding regarding the chromosome and genetic information consequences of mitosis; that is, on the pre-test 80% correctly state that the chromosome number is the same before and after mitosis and 86% correctly state that the genetic information between daughter cells is the same after mitosis. And, students are very clear on the fact that egg and sperm cells from the same organism have the same number of chromosomes (data not shown). However, students are essentially split on their responses regarding the genetic information content of ovary and egg cells after meiosis (Question 26/29, Table 3). Perhaps students are confused with the intent of the question, i.e., an egg cell may be interpreted as the same as an ovary cell, or perhaps they truly are confused about the relationship between the genetic information in these two different cell types. Interestingly, greater than 65-70% of students accurately answered that an egg cell has half the chromosomes as an ovary cell after meiosis. Interviews with students will shed more light on the reasons for these apparently incongruous answers. Also, more robust statistical analysis (beyond that available in the Blackboard learning management system) will allow correlations between pre- and post-test answers. For example, it will be informative to know whether the students who still think that egg cells contain the same genetic information as ovary cells are the same students who cannot correctly identify the number of chromosomes in an egg cell.

A breakdown of specific responses to questions 29/12 is shown in Table 4. This question relates to whether or not students understand that all cell types contain the same genetic information, a misconception identified from the 9th grade/high school research. It appears that a significant number of college juniors maintain this idea even by the end of the term. While 55% of Adv I students start the quarter thinking that only gametes contain the genetic information for eye color, a large number of students (32%) still think this at the end of the term. In the Adv II pre-test results, 30% of students think gametes are the only cells than contain eye color genes; however, 16% think it is eye and gamete cells and 34% give the correct answer (all cell types listed). Interestingly, responses to a corollary question that asks whether genes are found in all cell types in an organism show that 78-80% of Adv I and Adv II students answer correctly. Student interviews as well as statistical correlations will be required to further analyze this discrepancy. In this case, it would be informative to find out whether students making up the 32% who still think that gametes are the only cell type to contain eye color genes are the same students who are confused regarding which cell types contain genes.

Table 4. Responses to Question 29 and 12

Possible Answer	Adv I Pre %	Adv I Post %	Adv II Pre %
All	20	52	34
Eye & gamete	12	0	16
Gamete only	55	32	30

With respect to question 31/19 regarding recombination (Table 3), students usually choose “the process of crossing over during meiosis” over the correct response of “sorting of alleles into new combinations.” This is perhaps understandable as these two concepts are often discussed at the same time and the word recombination and crossing over may also be used interchangeably.

The next set of questions deal with the products and location of translation and transcription (Question 34 and 35/33 in Table 3, respectively). Intro results indicate that 31% of students correctly identify proteins as the products of translation, and 37% or 52% of students (Adv I and Adv II, respectively) do the same (Table 5).

Table 5. Responses to Question 34 (products of translation)

Possible Answer	**Intro Post %**	***Adv II Pre %***	**Adv Pre %**	**Adv Post %**
Proteins	31	52	37	75
Messenger RNAs or ribosomal RNAs	17	16	33	9
Amino acids	13	22	18	0
DNA	16	10	8	3
Amino acids and proteins	22	nd	nd	nd

Fisher (1985) observed that at least half of students in a similar study responded with "amino acids" to a similarly-worded question. That isn't the case here: less than or close to 20% give this response. And, there also appears to be confusion between mRNAs and proteins suggesting that students may be simply mixing up the terms translation and transcription. Fisher's question was worded slightly differently and used "activating enzymes" instead of "proteins" which may also be a factor. Alternatively, perhaps instruction has changed over the past decades to clear up the misconception originally described. It may also be that students really don't know what happens in translation. Student interviews on this topic, including the origin of amino acids will be instructive. With respect to transcription (Question 35/33 in Table 3), Intro and Adv I students have a generally good grasp that mRNA molecules are the proper product. In the Adv II scenario, rRNA was added as a possible choice in a multiple answer question and, in this case, few students (only 18%) identified both mRNA and rRNA as products of transcription.

In the Adv II version of the exam, students were asked to choose where within cells mRNAs, proteins and amino acids were synthesized in three separate questions. Around 60% of students know the correct compartment for mRNA and protein synthesis but the class is split on the location of amino acids synthesis (Table 6). Given that students generally take genetics before biochemistry, their understanding of metabolic pathways may be limited, including those involved in amino acid production. Finally, perhaps adding a question that asks students to identify which cellular process gives rise to amino acids would also be useful.

Table 6. Adv II responses to where mRNAs, proteins and amino acids are made.

Possible Answer	mRNAs	Proteins	Amino Acids
Ribosomes	32	62	46
Free in the cytoplasm	8	30	40
Nucleus	60	6	14

One of the questions asked students to complete the phrase, "One gene encodes…" in a short answer format in Adv I pre and post test, then those short answers were used to construct multiple choice options for the Intro and Adv II test takers. The answers students gave with the corresponding percentages are summarized in Table 7. Common Intro and initial Adv 1 or II answers are that genes encode traits or a single protein. The responses to this question at the Adv I post-test reveal that student thinking becomes more complex over the course of the term: after class, they were less likely to say that a gene encoded a trait and more likely to say that a gene encoded many proteins or a transcriptional unit that could produce many different products. However, it is disturbing that the percentage of students who still believe that one gene encodes one protein or polypeptide increased over the course of the quarter. Perhaps they are still holding on to the famous phrase "one gene, one enzyme." Interestingly, three responses to this short answer Adv I post-test stated that a gene encodes many amino acids. This ties back to Question 34 regarding understanding of translation. Even though there were no students giving amino acids as the products of translation in question 34 at the end of the quarter (Table 5), a few revealed a possible confusion or lack of full understanding in this short answer question.

Table 7. Responses to question, "One gene encodes…"

Possible Answer	Intro Post %	*Adv II Pre %*	Adv I Pre %	Adv I Post %
A trait or characteristic	35	32	43	10
Many traits	14	14	10	6
One enzyme	-	24	0	10
One protein or polypeptide	27	-	19	30
Many proteins	-	6	0	23
A transcriptional unit…	20	24	0	20
Other (short answer only)	-	-	28	1

The GenCI questions for all versions of the test will be provided at the CAB II meeting. Trials of the GenCI have been run at Michigan State University and data analysis of those results will proceed in the next 6 months. Additionally, further question development is underway. Approximately seven questions have been removed because students in all trials gave greater than 75% correct response or it was deemed that there were other more important conceptual areas to include. A plan for student interviews is being developed to better understand some of the results described above. After interviews, we plan to create another modification of the test with modified as well as new questions to pilot.

Challenges

The major challenges discussed at the CAB I meeting in Boulder, CO are still relevant. These challenges include creating clear and unambiguous test items, and writing test items that measure conceptual understanding not just structural or definitional understanding. Technological challenges include statistical analysis (validity and reliability testing; question pairing and correlations) and online test administration that is available to students at a variety of campuses that allows for access not only to the exam, but also to the raw data. It is also challenging to interpret the data, especially if multiple answer questions are used. Multiple answer questions are useful because they provide flexibility to allow students to choose a variety of answers according to what they are thinking. However, it is difficult to know whether or not they realized the question may have multiple answers. A worthy discussion surrounding the use of multiple answer questions would benefit this work. It is also difficult to know whether results from a particular test item indicate a serious misconception or simply a misunderstanding. Certainly interviews help in these cases, however that cannot be done for each test taker. A clearer definition of misconception in biology will be helpful in the development and use of these types of inventories. Additionally, the distinction between misconception and misunderstanding will be narrowed with validity analysis of test items.

Ways I could use help from the group

- Student interview protocol
- Design and statistical correlation analysis of paired or tiered questions; discussion of use of multiple answer questions
- Consensus definition of misconception
- Consensus definition of a concept in biology; description of major genetics concepts
- Additional trials of the GenCI instrument at other universities; possible incorporation of GenCI questions into the BCI.

References Cited

1. Elrod, S.L., 2007, *Genetics Concept Inventory*, Conceptual Assessment in Biology II workshop paper, Boulder, Colorado.
2. Elrod, S.L. and M.M. Somerville, 2007, *Literature-Based Scientific Learning: A Collaboration Model*, Journal of Academic Librarianship, in press.
3. Hott, A.M., C.A. Huether, J.D. McInerney, C. Christianson, R. Fowler, H. Bender, J. Jenkins, A. Wysocki, G. Markle, R. Karp, 2002, *Genetics Content in Introductory Biology Courses for Non-Science Majors: Theory and Practice*, BioScience 52 (11); 1024-1035.
4. Marbach-Ad, G., 2001, *Attempting to break the code in student comprehension of genetic concepts*, J of Bio Educ 35(4):183-189.
5. Marbach-Ad, G. and R. Stavy, 2000, *Students' Cellular and Molecular Explanations of Genetic Phenomena*, J Bio Educ, 34(4): 200-205.
6. Banet, E. and E. Ayuso, 2000, *Teaching Genetics at Secondary School: A Strategy for Teaching about the Location of Inheritance Information*, Sci Ed, 84: 313-351.
7. Wood-Robinson, C., Lewis, J. and Leach, J. 2000 *Young people's understanding of the nature of genetic information in the cells of an organism*, J Bio Educ 35(1): 29-36.
8. Lewis, J., Leach, J. and Wood-Robinson, C. 2000 *Chromosomes: The Missing Link - young people's understanding of mitosis, meiosis and fertilization*, JBio Educ 34(4): 189-199.
9. Lewis, J., Leach, J. and Wood-Robinson, C. 2000 *What's In a Cell? Young peoples understanding of the genetic relationship between cells, within an individual*, J Bio Educ 34(3): 129-132.
10. Lewis, J., Leach, J. and Wood-Robinson, C. 2000 *All In The Genes? Young people's understanding of the nature of genes*, J Bio Educ 34(2): 74-79.
11. Lewis, J. and Wood-Robinson, C. (2000) *Genes, chromosomes, cell division and inheritance - do students see a relationship?*, Int J Sci Educ 22(2): 177-195.
12. Parker, J. *et al*., 2007, *Teaching and Learning Biology at the Undergraduate Level*, Conceptual Assessment in Biology II workshop paper, Boulder, Colorado.
13. Fisher, K.M., 1985, *A Misconception in Biology: Amino Acids and Translation*, J Res Sci Teach 22(1): 53-62.