Chapter 16
The Contribution of History and Philosophy to the Problem of Hybrid Views About Genes in Genetics Teaching

Charbel N. El-Hani, Ana Maria R. de Almeida, Gilberto C. Bomfim, Leyla M. Joaquim, João Carlos M. Magalhães, Lia M.N. Meyer, Maiana A. Pitombo, and Vanessa C. dos Santos

16.1 Introduction

The gene concept has been one of the landmarks in the history of science in the twentieth century, which has been even characterized as "the century of the gene" (Gelbart 1998; Keller 2000). However, there are nowadays persistent doubts about the meaning and contributions of this concept, not only among philosophers of biology[1] but also among empirical scientists.[2] Moreover, by the mid-2000s concerns about the gene extended to the editorials of high-impact scientific journals (e.g., Pearson 2006).

[1] See, for example, Burian (1985), Falk (1986), Fogle (1990), Hull (1974), and Kitcher (1982).

[2] See, for example, Gerstein et al. (2007), Kampa et al. (2004), Venter et al. (2001), and Wang et al. (2000).

C.N. El-Hani (✉)
History, Philosophy and Biology Teaching Lab (LEFHBio),
Institute of Biology, Federal University of Bahia, Salvador, Brazil
e-mail: charbel@ufba.br; charbel.elhani@pesquisador.cnpq.br

Department of General Biology, Institute of Biology, Universidade Federal da Bahia,
Campus de Ondina. Rua Barão de Jeremoabo, s/n, 40170-115 Ondina, Salvador-BA, Brazil

A.M.R. de Almeida
History, Philosophy and Biology Teaching Lab (LEFHBio),
Institute of Biology, Federal University of Bahia, Salvador, Brazil

Graduate Studies Program in Plant Biology, University of California, Berkeley, USA

G.C. Bomfim • L.M. Joaquim • M.A. Pitombo • V.C. dos Santos
History, Philosophy and Biology Teaching Lab (LEFHBio),
Institute of Biology, Federal University of Bahia, Salvador, Brazil

J.C.M. Magalhães
Department of Genetics, Federal University of Paraná, Curitiba, Brazil

M.R. Matthews (ed.), *International Handbook of Research in History, Philosophy and Science Teaching*, DOI 10.1007/978-94-007-7654-8_16,

There are negative and positive reactions to the problem of the gene or, as El-Hani (2007) describes, attempts to eliminate this concept from biology or to keep it although radically reconceptualized. Keller (2000), for instance, suggested that maybe the time was ripe to forge new words and leave the gene concept aside (see also Portin 1993; Gelbart 1998). More optimistic views are found, for example, in Hall (2001), who argued that, despite published obituaries, the gene was not dead, but alive and well, and seeking a haven from which to steer a course to its "natural" home, the cell as a fundamental morphogenetic unit, or in Knight (2007), for whom "reports of the death of the gene are greatly exaggerated."

The crisis of the gene concept is mostly related to its interpretation as *a stretch of DNA that encodes a functional product, a single polypeptide chain or RNA molecule*, that is, the so-called classical molecular gene concept (Neumann-Held 1999; see also Griffiths and Neumann-Held 1999; Stotz et al. 2004). Under the influence of this concept, simple and straightforward one-to-one relationships (function = gene = polypeptide = continuous piece of DNA = cistron) were regarded as acceptable in understanding the functioning of the genetic system from the 1940s to the 1970s (Scherrer and Jost 2007a, b). These relationships were captured in a manner that was heuristically powerful in genetics and molecular biology, which benefited from treating the gene as an uninterrupted unit in the genome, with a clear beginning and a clear ending and with a single function ascribed to its product (and, thus, indirectly to the gene). The explanatory and heuristic power of this concept follows from how it brought together structural and functional definitions of the gene, alongside with an easily understandable mechanics. With the introduction of an informational vocabulary in molecular biology and genetics (Kay 2000), genes were also regarded as informational units, leading to what has been called the informational conception of the gene (Stotz et al. 2004), a popular notion in textbooks, the media, and public opinion.

This picture changed since the 1970s, as the view of the gene as a structural and functional unit was increasingly challenged by anomalies resulting from research mostly conducted in eukaryotes, in which we find nothing like the tight physical complex linking transcription and translation observed in bacteria. We can classify these anomalies in three kinds, all related to counterevidence for a unitary relationship between genes, gene products, and gene function: (i) *one-to-many* correspondences between DNA segments and RNAs/polypeptides (as, for instance,

L.M.N. Meyer
History, Philosophy and Biology Teaching Lab (LEFHBio),
Institute of Biology, Federal University of Bahia, Salvador, Brazil

Department of Biosciences, Federal University of Sergipe, Itabaiana, Brazil

in alternative splicing,[3] Black 2003; Graveley 2001), (ii) *many-to-one* correspondences between DNA segments and RNAs/polypeptides (as in genomic rearrangements, such as those involved in the generation of diversity in lymphocyte antigen receptors in the immune system[4]; see Cooper and Alder 2006; Murre 2007), and (iii) *lack of correspondence* between DNA segments and RNAs/polypeptides (as we see, e.g., in mRNA editing[5]; see Hanson 1996; Lev-Maor et al. 2007).

Another key issue related to the gene concerns conceptual variation and ambiguities throughout its history (e.g., Carlson 1966). As Rheinberger (2000) argues, genes can be regarded as "epistemic objects" in genetics and molecular biology, entities introduced and conceived as targets of research, whose understanding is framed by the set of experimental practices used by particular scientific communities. Thus, conceptual variation can be explained as a consequence of different experimental practices used by diverse communities of scientists who deal with genes as epistemic objects (Stotz et al. 2004). For instance, population geneticists often work with an instrumental view of genes as determinants of phenotypic differences, since this is often enough to deal with the relationship between changing gene frequencies in populations over time and changes in the phenotypes of the individuals making up those populations. They tend to emphasize, thus, genes as markers of phenotypic effects, taking a more distal view on gene function. Molecular biologists, in turn, focus their attention on genes in DNA and their molecular products and interactions, emphasizing the structural nature of genes and their role in the cellular system they are part of. They take a more proximal view of genes and tend to be reluctant to identify a gene by only considering its contributions to relatively distant levels of gene expression (Stotz et al. 2004).

The phenomenon at stake here is gene function, and consequently, we will refer to multiple models of gene function, in the structure of which a central element is the gene concept.[6] The experimental practices used by diverse scientific communities

[3] In alternative splicing, a pre-mRNA molecule is processed – in particular, spliced – in a diversity of manners, so that different combinations of exons emerge in the mature mRNA. In this manner, several distinct mRNAs and, thus, polypeptides can be obtained from the same DNA sequence. In *Drosophila melanogaster*, for instance, DSCAM alternative splicing can lead to ca. 38,016 protein products (Celotto and Graveley 2001).

[4] The generation of the diverse antigen receptors found in lymphocytes and, consequently, of antibody specificity depends on a combinatorial set of genomic rearrangements between different DNA segments called variable segments, constant segments, and diversity and joining segments.

[5] mRNA editing is an alteration of mRNA nucleotides during processing, resulting in lack of correspondence between nucleotide sequences in mature mRNA and nucleotide sequences in DNA.

[6] "Model" is a polysemous term, with diverse meanings that capture distinct relationships between elements of knowledge (e.g., Black 1962; Grandy 2003; Halloun 2004, 2007; Hesse 1963). We treat models here as constructs created by the scientific community in order to represent relevant aspects of experience, i.e., phenomena and processes/mechanisms that can explain and/or predict them. In these terms, models capture the relationship between a symbolic system (a representation) and phenomena, processes, and mechanisms ontologically treated as being part of the world or nature. Models are built through processes of generalization, abstraction, and idealization that crucially involves selecting a number of entities, variables, relationships associated with a specific class of phenomena and processes/mechanisms to be included in the model, while others are

lead to variation in models of gene function and gene concepts. The expression "conceptual variation" describes, then, the range of different meanings ascribed to a concept, not necessarily all of them outdated, since they may still be used in different contexts.

Conceptual variation has been heuristically useful in the history of genetics.[7] Different gene concepts and different gene function models have been and still are useful in different areas of biology, with different theoretical commitments and research practices. Nevertheless, while recognizing that conceptual variation is a desirable feature in our understanding of genes, several authors stress that we should clearly distinguish between different concepts and models, with diverse domains of application.[8] After all, conceptual variation may also lead to misconceptions and misunderstandings. Falk (1986, p. 173), for instance, considers that the pluralism found in the current picture about genes "… brought us […] dangerously near to misconceptions and misunderstandings." Fogle (1990, p. 350) argues that "despite proposed methodological advantages for the juxtaposition of 'gene' concepts it is also true […] that confusion and ontological consequences follow when the classical intention for 'gene' conjoins a molecular 'gene' with fluid meaning." Keller (2005) argues that many problems arise from ambiguities in the usage of the term "gene," calling attention to difficulties with gene counting, since the value obtained may vary by two, three, or more orders of magnitude depending on how genes are defined.

Diversity in meaning and heterogeneity in reference potential can lead to semantic incommensurability, although this is not necessarily so. In the history of genetics, ideas associated with different ways of understanding genes and their roles in living systems have been sometimes merged in the construction of new concepts and models. However, one needs to consider that conceptual change often leads to scientific concepts with heterogeneous reference potentials and, thus, to models with diverse meanings, and as a result, there can be semantic incommensurability between concepts and models. When semantically incommensurable models and concepts, or even some of their features, are mixed up, logical inconsistencies and conceptual incoherence can appear.

In science, conceptual variation and the combination of ideas related to different models are usually (but not always) less problematic, since researchers usually develop a sophisticated understanding of the knowledge base of their research field (even though much can remain tacit) and also learn epistemic practices that stabilize

selected out. These entities, variables, and relationships are captured by scientific concepts, and thus, a model can be seen as a system of related concepts. Concepts gain meaning by being used in model construction, as contributors to model structure (Halloun 2004). If we understand scientific theories as families of models – according to a semantic approach (e.g., Develaki 2007; Suppe 1977; van Fraassen 1980) – concepts will form a network of relationships as a consequence of their participation in a series of models, and ultimately, the meaning of a concept will be constructed out of its relationship with other concepts in a network of models.

[7] See, for example, Burian (1985), Falk (1986), Griffiths and Neumann-Held (1999), Kitcher (1982), and Stotz et al. (2004).

[8] For instance, El-Hani (2007), Falk (1986), and Griffiths and Neumann-Held (1999).

to a significant extent the use of concepts and models. They are embedded in a community committed to a specific set of epistemic practices that make it more likely that they employ particular meanings ascribed to gene concepts and gene function models, which properly operates in a given domain of investigation. They also tend to recognize the prospects and limits of different concepts and models. This does not mean that concepts and models are per se stabilized when they emerge in the scientific community. On the contrary, they usually appear in a more rudimentary way, and if they are adopted by the scientific community, they can be elaborated and eventually stabilized by the practice of using them to guide research. When concepts and models are fused, the scientific community may be able to work out possible incoherence. However, as the diversity of concepts and models expands – as we see in the case of genes in the post-genomic era – difficulties are more likely to arise, particularly in the absence of a clear and explicit demarcation between those diverse meanings. This means that we should not remain content with the tacit usage of distinct meanings in different research settings, but rather worry about the clear demarcation of their domains of application (El-Hani 2007).

Certainly, teachers and students are embedded in a number of communities just as scientists are part of the scientific community. Every human being participates in a number of communities, which shape their understanding of the world. They can be described, if we follow Wenger (1998), as "communities of practice" (CoPs), cohesive groups of individuals mutually engaged in a joint enterprise, who exhibit distinct sets of knowledge, abilities, and experiences, and are actively involved in collaborative processes, sharing information, ideas, interests, resources, perspectives, activities, and, above all, practices, such that they build a shared repertoire of knowledge, attitudes, values, etc. (see also Lave and Wenger 1991). In the scientific community, we can find CoPs which generate a shared repertoire of knowledge, epistemic practices, and values that can stabilize the understanding of theories, models, and concepts to varying degrees. This means that scientists build a collective empiricism (Daston and Galison 2010) that often allows them to deal with a variety of models and concepts in a more consistent way. Or, to put it differently, persons tend to form "thought collectives," communities that mutually exchange ideas and develop a given "thought style" (Fleck 1979/1935). What is at stake here, then, is that scientists, teachers, and students pertain to different communities of practice, if we follow Wenger's formulation or, thought collectives, if we follow Fleck's and, thus, will tend to assume different perspectives on the diversity of scientific models and concepts. And the fact that scientists can be embedded in communities that generate that very diversity is of the utmost importance here.

When we turn to science education, we have additional reasons to worry about conceptual variation about genes and their function and the hybridization of different gene concepts and gene function models, as argued by Gericke and Hagberg (2007, 2010a, b), Gericke et al. (in press), and Santos et al. (2012). After all, even though teachers and students are themselves embedded in CoPs or thought collectives, they are not embedded in those scientific communities that generate knowledge about genes and their function. Moreover, in educational settings conceptual variation tends to be greater than in the scientific community, since both scientific and

everyday meanings are represented and interact with each other within classrooms (Mortimer and Scott 2003), and disciplinary boundaries which may stabilize meaning making are not always present. In sum, when compared to the scientific community, there is much more potential to indiscriminate mixture of semantically incommensurable scientific concepts and models in the science classroom and, thus, a much bigger potential that logical inconsistencies and conceptual incoherence emerge. This is particularly true when science is taught without due attention to its history and philosophy.

It is important, therefore, to investigate whether and how conceptual variation related to the gene concept and gene function models is present in school science and also what potential problems it may bring to genetics teaching and learning. In this chapter, we will survey the results of a research program conducted in our lab in the last 7 years, focusing on how ideas about genes and gene function are treated in school knowledge, as represented in textbooks and students' views. Moreover, following our usual approach to research on science education, we move from descriptive to intervention studies, i.e., from diagnosing views on genes to investigating a teaching strategy implemented in a classroom setting with the goal of changing higher education students' views and, in particular, improving their understanding of scientific models and conceptual variation around genes and their function. Here, we will first consider results from investigations on how higher education and high school textbooks deal with genes, gene function, and their conceptual variation. Second, we will report unpublished results concerning how higher education biology students deal with genes and gene function. Third, we will present findings of an unpublished intervention study in which we investigate design principles for teaching sequences about genes and their function, considering conceptual variation in genetics and molecular biology, the crisis of the gene concept, and current proposals for revising its meaning. As a background for these empirical researches, we will turn to their theoretical underpinnings, resulting from both the literature on philosophy of biology/theoretical biology and the educational literature.

16.2 Genes and Gene Function Through the History of Genetics

The term "gene" was created in 1909, by Johannsen, following his distinction between genotype and phenotype, which told apart two ideas embedded in the term "unit character," then largely used, (1) a visible character of an organism which behaves as an indivisible unit of Mendelian inheritance and, by implication, (2) the idea of that entity in the germ cell that produces the visible character (Falk 1986). Johannsen proposed, then, the existence of basic units composing the genotype and phenotype, respectively, "genes" and "phenes." While the latter term never gained currency in biology, the former became central in newborn genetics and marked its development throughout the twentieth century.

Genes were seen instrumentally in the beginnings of genetics. Johannsen conceived "gene" as a very handy term with no clearly established material counterpart (Johannsen 1909). Although accepting that heredity was based on physicochemical processes, he warned against the conception of the gene as a material, morphologically characterized structure. For Johannsen, "the gene is […] to be used as a kind of accounting or calculating unit" (Johannsen 1909; See Falk 1986; Wanscher 1975). At that period, the gene (that "something" which was the potential for a trait) could only be inferred from its "representative," the trait. That is, the gene was defined top-down, based on the phenotype.

This way of understanding genes is part of the Mendelian model of gene function, as reconstructed by Gericke and Hagberg (2007). According to this model, the gene is the unit of transmission (or inheritance) and function, treated as an abstract entity interpreted instrumentally as a phenotype in miniature. The function of the gene is of minor importance in the Mendelian model, focused on explaining genetic transmission. Moreover, due to the instrumental nature of the gene and its definition from the phenotype, this model conceives the gene as a necessary and sufficient condition for the manifestation of a trait, with no consideration of environmental or any other factor besides those instrumental entities. Thus, it assumed a unitary relationship between genes and traits, and the idea of genes as units became central in Mendelian genetics, thereafter substantially influencing twentieth century biology.

With the establishment of the chromosome theory of heredity by T. H. Morgan and his group, a new understanding of genes emerged (Carlson 1966). This understanding amounts to Gericke and Hagberg's (2007) classical model of gene function. The gene acts, in this model, as the unit of genetic transmission, inheritance, function, mutation, and recombination (Mayr 1982). Two additional important ideas are that genes exist in different variants (alleles) and consist or act as enzymes that produce traits. Since the molecular structure of genes was unknown, this latter idea was vague, and genes and their function were still inferred from traits. This model treated genes, however, as more active in the determination of traits than the Mendelian model did. Due to the development of linkage maps by Alfred Sturtevant, from Morgan's group, genes came to be interpreted in terms of the beads-on-a-string concept. Those quantified particles in the chromosomes were increasingly seen in realist rather than instrumentalist way, despite Morgan's hesitation (Falk 1986). Another notorious member of Morgan's group, Herman J. Muller, was one of the first supporters of the idea that genes were material units, "ultramicroscopic particles" in the chromosomes, arguing against the description of the gene as "a purely idealistic concept, divorced from real things" (quoted by Falk 1986). This view paved the way for subsequent steps in a research program aiming at elucidating the material bases of inheritance.

With a minor modification resulting from biochemical studies on the nature of genes, what Gericke and Hagberg (2007) call the biochemical-classical model of gene function emerged. The gene was treated, then, as being responsible for the production of a specific enzyme, which produced a trait. Also, as increased knowledge on biochemical reactions became available, the focus shifted from transmission to gene action and function. The biochemical-classical model explained gene function

by reducing it to the relationship between a specific enzyme produced by the gene and the determination of a phenotypic trait. The model did not explain, however, the biochemical processes involved, and consequently, it still used the conceptual tools of classical genetics. The biochemical-classical gene was still an entity with unknown molecular structure.

The biochemical-classical model is the origin of the famous "one gene-one enzyme" hypothesis, which suffered several reformulations with increasing knowledge: when it was shown that the gene product was not always an enzyme, there was a shift to the "one gene-one protein" hypothesis, and when it was shown that proteins could be composed by several polypeptides, the "one gene-one polypeptide" hypothesis emerged. Finally, when it was established that RNAs could also be final gene products, the "one gene-one polypeptide or RNA" hypothesis prevailed. Notice, however, that an important shared content in all these hypotheses is that genes are treated as units.

At first, the gene was conceived as a unit of transmission, recombination, function, and mutation, but this did not hold. Benzer (1957) showed that units of function (his "cistrons") are typically much larger than units of recombination ("recons") and mutation ("mutons"). The terms "muton" and "recon" were deleted from the vocabulary of genetics, but "cistron" survived to these days and is often used in the place of "gene," indicating that the idea that prevailed was that of the gene as "unit of function."

The molecular-informational model (Santos et al. 2012)[9] was the culmination of a series of investigations about the material nature of the gene, which ultimately led to the proposal of the double helix model of DNA by Watson and Crick (1953). This model explained in one shot the nature of the linear sequence of genes, the mechanism of gene replication and RNA synthesis from DNA sequences, and the separation of mutation, recombination, and function at the molecular level. It was responsible for the wide acceptance of a realist view about genes, since there was now a clear material counterpart for the gene concept. The stage was set for a molecular definition of genes, in which genes were not defined anymore in a top-down manner, based on phenotypic traits, but in a bottom-up approach, focused on nucleotide sequences in DNA. This was accomplished through a concept named by Neumann-Held (1999) the classical molecular concept of the gene. According to it, a gene is a DNA segment encoding one functional product, which can be either a RNA molecule or a polypeptide. This concept superimposed a molecular understanding onto the idea of a hereditary unit supported by Mendelian genetics (Fogle 1990) and played an important role in the transition from classical genetics to a new era in which genetics and molecular biology became inseparable.

In the classical molecular concept, the gene is a continuous and discrete DNA segment, with no interruption or overlap with other units, showing a clear-cut beginning and end, and a constant location. Genes can be treated, then, as units of structure and, provided that they codify a single RNA molecule or polypeptide with a single function, also as units of function. And, with the introduction of information talk in

[9] This corresponds to Gericke and Hagberg's (2007) neoclassical model of gene function.

biology (Kay 2000) and in connection with the so-called central dogma of molecular biology, the gene became also a unit of information, simultaneously a chemical and a program for running life.[10] However, this idea is hardly trivial: despite the widespread usage of informational terms in molecular biology and genetics (say, "genetic information," "genetic code," "genetic message," "signaling,"), they can be still regarded as metaphors in search of a theory (El-Hani et al. 2006, 2009; Griffiths 2001). We do not have yet a sufficient and consistent theory of biological information, despite the utility of Shannon and Weaver's (1949) mathematical theory of communication for several purposes in biological research (Adami 2004). The non-semantic understanding of information in this theory seems insufficient for a theory of biological information. Many authors argue that biology needs a theory of information including syntactic, semantic, and pragmatic dimensions (e.g., El-Hani et al. 2006, 2009; Hoffmeyer and Emmeche 1991; Jablonka 2002). Notwithstanding, genes are frequently treated as informational units, leading to the informational conception of the gene (Stotz et al. 2004), which is often superimposed onto the classical molecular concept even though it does not have a clear meaning.

As discussed in the introduction, several findings of genetic, molecular, and genomic research challenged in the last three decades the molecular-informational model, posing problems for the understanding of a gene as a unit of structure, function, and/or information. Even though the crisis of this model was more widely recognized in the last two decades of the twentieth century, many-to-many relationships were known to classical genetics already. Benzer, for instance, regarded the gene as a "dirty word" (Holmes 2006).

Gericke and Hagberg (2007) introduce a "modern model" to encompass these challenges, in which the gene is treated as a combination of DNA segments that acts in a process that defines the function. This stretch of DNA contains regulating sequences and a transcription unit, made of coding sequences, but also introns and flanking sequences. It is expressed to produce one or several functional products, either RNAs or polypeptides. Smith and Adkison (2010) complemented this account by considering two further elements: (1) the findings of the Human Genome Project, such as the relatively limited number of genes in human and other genomes, when compared to previous estimates, and the similarity in gene numbers between humans and other animals, and (2) the definition of gene proposed by the ENCyclopedia Of DNA Elements (ENCODE) project.[11] We need to be careful,

[10] This shows the connection between the informational conception of the gene and genetic determinism (Oyama 2000/1985), a common element of the "gene talk" (Keller 2000) that pervades the media and the public opinion. With the central dogma, DNA became a sort of reservoir from where all "information" in a cell flows and to which it must be ultimately reduced. Through their connection with the doctrine of genetic determinism, the conceptual problems related to genes and genetic information have important consequences for public understanding of science and several socioscientific issues related to genetics and molecular biology (say, genetic testing, cloning, genetically modified organisms).

[11] The ENCODE project is an international consortium of scientists trying to identify the functional elements in the human genome sequence, with significant impact on our understanding about genes and genomes. The ENCODE database can be reached at http://www.genome.gov/10005107#4.

however, in referring to a "modern model," since this may mask the fact that there is no prevailing model nowadays. The gene concept is now in flux, changing meanings as researchers produce novel interpretations of the structure and dynamics of the genomic system.

Several proposals for reformulating the gene concept appeared in the last 20 years. We will just mention some of them here, with no intention of being exhaustive or providing any detailed discussion.[12] Some authors argued against the idea of genes as units and proposed, instead, views about genes as combinations of nucleic acid sequences that correspond to a given product (Fogle 1990, 2000; Pardini and Guimarães 1992) and might be located in processed RNA molecules (Scherrer and Jost 2007a, b). These proposals accommodate anomalies such as overlapping and nested genes by denying the idea of genes as units in DNA.

Other authors put forward a process-oriented view of genes.[13] In Neumann-Held's "process molecular gene concept," for instance, genes are not treated as "bare DNA" but as the whole molecular process "… that leads to the temporally and spatially regulated expression of a particular polypeptide product" (Griffiths and Neumann-Held 1999, p. 659). Since different epigenetic conditions that affect gene expression are in this way built into the gene, this proposal can accommodate anomalies such as alternative splicing or mRNA editing.

Moss (2001, 2003) distinguished between two meanings ascribed to genes and, consequently, demarcated two concepts, gene-P and gene-D, which have been usually conflated throughout the twentieth century. Gene-P amounts to the gene as determinant of phenotypes or phenotypic differences. It is an instrumental concept, not accompanied by any hypothesis of correspondence to reality, and this is what allows one to accept the simplifying assumption of a preformationist determinism (as if the trait was already contained in the gene, albeit in potency). Gene-P is useful to perform a number of relevant tasks in genetics, such as pedigree analysis or genetic improvement by controlled crossing methods. Gene-D amounts to the gene as a developmental resource in causal parity (Griffiths and Knight 1998) with other such resources (say, epigenetic ones). It is conceived as a real entity defined by some molecular sequence in DNA which acts as a transcription unit and provides molecular templates for the synthesis of gene products, being in itself indeterminate with respect to the phenotype (Moss 2003, p. 46). Gene-D is in accordance, thus, with the classical molecular concept. Moss argues that genes can be productively conceived in these two different ways, *but nothing good results from their conflation* (Moss 2001, p. 85). This conflation is one of the main sources of genetic determinism, with important consequences to socioscientific issues, since it leads to the idea of

The participants of the ENCODE can be found at http://www.genome.gov/26525220. See also The ENCODE Project Consortium (2004).

[12] For detailed discussion, see Meyer et al. (2013). When we consider these views about genes and their function, it is worth pondering about the school level to which they can be adequately transposed. This issue is also discussed by Meyer et al. (2013).

[13] See, for example, El-Hani et al. (2006, 2009), Griffiths and Neumann-Held (1999), Keller (2005), and Neumann-Held (1999, 2001).

genes as major or even single causal determinants of phenotypic traits, even highly complex traits, such as sexual orientation, intelligence, or aggression.

Among the contributions of the ENCODE project, we find a new definition of gene: "… *a union of genomic sequences encoding a coherent set of potentially overlapping functional products*" (Gerstein et al. 2007, p. 677, emphasis in the original). In this definition, different functional products of the same class (proteins or RNAs) that overlap in their usage of the same primary DNA sequences are combined in the same gene, and thus, several anomalies are accommodated by challenging the unitary relationship between genes, gene products, and gene function embedded in the classical molecular concept.

Some works strive for solving the gene problem by building new languages that cut up the genetic system into novel categories, organizing our understanding into different sets of concepts (Keller and Harel 2007; Scherrer and Jost 2007a, b). On the one hand, this may solve, or dissolve, problems and limits posed by our current language about genes. On the other, there is an expected difficulty of translation between the new languages and the one already established in the fields of genetics and molecular biology, which may hamper researchers' understanding of those new ways of speaking and, thus, their acceptance. To maintain sufficient bridges between new and older ways of speaking seems crucial, then, for the success of these proposals.

When we consider these new views about genes and their function, it is worth pondering about the school level to which they can be adequately transposed. This is not the space, however, to enter this discussion (see Meyer et al. 2013).

16.3 Methods

16.3.1 Textbook Studies[14]

16.3.1.1 Sample

We analyzed higher education and high school textbooks. A sample of higher education Cell and Molecular biology textbooks was selected through a survey of 80 course syllabi of 67 universities located in the 5 continents, randomly chosen in Google® searches performed in 2004. We analyzed three of the most used textbooks, respectively, Lodish et al. (2003, n = 33 syllabi, the most used), Alberts et al. (2002, n = 28, the second most used), and Karp (2004, n = 5, the fifth most used). In many countries these textbooks are used in their original language, although it is possible to find translations. Thus, we analyzed them in the original language.

Eighteen biology textbooks (see Appendix 1) submitted by publishing companies to the Brazilian National Program for High School Textbooks (PNLEM) (El-Hani et al. 2007, 2011) were analyzed. This sample shows external validity regarding

[14] For more details, see Santos et al. (2012) and Pitombo et al. (2008).

Brazilian textbooks. PNLEM is a huge governmental initiative, providing textbooks to students enrolled in public high schools throughout the country. These textbooks are aimed at general high school biology courses attended by all students, covering all areas of biology. Besides being distributed to public schools by PNLEM, most of these textbooks are also used by private schools.

16.3.1.2 Textbook Content Analysis

Each textbook was analyzed as a whole using categorical content analysis (Bardin 2000). The procedure involved, first, the decomposition of the texts into units of analysis (recording units), from which categories were built through regroupings of text elements sharing characteristics identified by semantic criteria, i.e., by the presence of the same meaning in different text elements, not by the occurrence of specific keywords or sentences. First, an exploratory reading was performed to plan the decomposition of the texts, data treatment, and categorization. Besides the units of recording, we also considered units of context, larger segments of text embedding the units of recording, which provided a background for interpreting them. Recording units were the basic units for categorization and frequency calculation, varying in size from a single statement to a whole paragraph.

Since different areas of biology use particular epistemic practices, which lead to the creation of distinct ways of thinking and speaking about genes, most units of context were related to biological subdisciplines. In high school textbooks, they were characterization of life and/or living beings (i.e., the introductory chapters in the textbooks), cell and molecular biology, genetics, evolution, and glossary.[15] In higher education textbooks, the following units of context were employed: classical genetics, developmental genetics, evolutionary/population genetics, genetics of microorganisms, genetics of eukaryotes, medical genetics, molecular biology/molecular genetics, cell biology, biochemistry, cell signaling, genetic engineering, genomics, introduction, history of science, and glossary.[16]

Higher education textbooks were analyzed by using categories informed by the historical, philosophical, and scientific literature about genes. In high school textbooks, we employed three analyzing procedures: (1) analysis of gene concepts and (2) analysis of function ascription to genes, both based on the abovementioned literature, and (3) analysis of historical models of gene function, as described by Gericke and Hagberg (2007). In the latter analysis, we used the research instrument built by these authors, with some changes, to investigate how the variants associated with each of the seven epistemological features of the historical models were found in the recording units (Table 16.1).

Depending on the combination of epistemological feature variants used in an explanation of gene function, the explanation present in the recording unit can be classified into the historical models (Table 16.2). However, in school science, models are often reconstructed in a nonhistorical way, due to neglect of their historical

[15] Only 4 textbooks had a glossary. All other units of contexts were present in all textbooks.

[16] A glossary was present in all the textbooks.

Table 16.1 Description of the epistemological feature variants used in the high school textbooks analyses

Epistemological features			Epistemological feature variant
1	The structural and functional relation to the gene	**1a**	The gene is an abstract entity and, thus, has no structure
		1b	The gene is a particle on the chromosome
		1c	The gene is a DNA segment
		1d	The gene consists of one or several DNA segments with various purposes
		1e	**The gene is a carrier, bearer, and/or unit of information**
2	The relationship between organization level and definition of gene function	**2Ia**	The model has entities at the phenotypic level and abstract concepts[a]
		2Ib	The model has entities at the phenotypic and cell levels[a]
		2Ibx	The model has entities at the phenotypic, cell, and molecular levels[a]
		2Ic	The model has entities at the molecular level
		2Icx	The model has entities at the cell and molecular levels
		2Icy	**The model has entities at the phenotypic and molecular levels[a]**
		2IIa	The correspondence between gene and its function is one-to-one
		2IIb	The correspondence between gene and its function is many-to-many
3	The 'real' approach to define the function of the gene	**3a**	The function of the gene is defined "top-down"
		3b	The function of the gene is defined "bottom-up"
		3c	The function of the gene is defined by an underlying process related to the capacity of expressing a particular gene product[b]
4	The relationship between genotype and phenotype	**4a**	There is no separation between genotype and phenotype
		4b	There is a separation, without explanation, between genotype and phenotype
		4c	There is a separation between genotype and phenotype with enzyme as intermediate causal explanation[b]
		4d	There is a separation between genotype and phenotype, explained by biochemical processes
5	The idealistic *versus* naturalistic relationships in the models	**5Ia**	There are idealistic relations in the model, with no reference to natural processes[b]
		5Ib	There are naturalistic relations in the model, with a detailed description of the biochemical process of gene expression[b]
		5IIa	The relations in the model are causal and mechanistic (chemical interactions of genes determine traits independently of context)[b]
		5IIb	The relations in the model are process oriented and holistic (the function of the gene depends on the context in which it is embedded)[b]
6	The reduction explanatory problem	**6a**	There is explanatory reduction from the phenotypic level to abstract concepts[a]
		6b	There is explanatory reduction from the phenotypic to the cell level[a]
		6bx	There is explanatory reduction from the phenotypic level to the molecular level[a]
		6c	There is no explanatory reduction
7	The relationship between genetic and environmental factors [in development and the construction of the phenotype]	**7a**	Environmental entities are not considered
		7ax	Environmental entities + genetic entities result in a trait/product/function[a]
		7b	Environmental entities are implied by the developmental system
		7c	Environmental entities are shown as part of a process

Variants in gray were introduced by Santos et al. (2012) in the original research instrument constructed by Gericke and Hagberg (2007)

[a]Changes in terminology introduced by Santos et al. (2012) in the epistemological feature variants

[b]Variants modified by Santos et al. (2012) in order to make some aspects more explicit

[c]The relationship is understood in additive terms, each factor being related to the product, but with no significant mutual influence between them

and epistemological backgrounds during didactic transposition (Justi and Gilbert 1999). Thus, hybrid models are often found, i.e., explanatory models consisting of aspects belonging to different historical models, which may be incoherent if incommensurable aspects are mixed up. We calculated the degree of model hybridization in textbook explanations of gene function, by ascertaining the frequency of

Table 16.2 Models of gene function and their epistemological feature variants (Gericke and Hagberg 2007, modified by Santos et al. 2012)

	Epistemological feature variants								
Models of gene function	**1**	**2I**	**2II**	**3**	**4**	**5I**	**5II**	**6**	**7**
Mendelian model	1a	2Ia	2IIa	3a	4a	5Ia	5IIa	6a	7a
Classical model	1b	2Ib	2IIb	3a	4b	5Ia	5IIa	6b	7a
Biochemical-classical model	1b	2Ib	2IIa and 2IIb	3a and 3b	4c	5Ia	5IIa	6b	7a
Neoclassical (or molecular-informational) model	1c and 1e	2Ic	2IIa	3b	4d	5Ib	5IIa	6c	7b
Modern model	1d	2Ic	2IIa	3c	4d	5Ib	5IIb	6c	7c

false-historical (i.e., belonging to the wrong historical model) and nonhistorical (i.e., not present in any of the historical models) feature variants.

We analyzed the presence of historical models of gene function in the textbooks in two different ways. In a previous study (Santos et al. 2012), we identified feature variants related to these models in each set of chapters related to the domain of a biological subdiscipline and, then, checked the model to which most of the epistemological feature variants were linked. We identified, thus, the prevailing model at that set of chapters, while the other feature variants, either false historical or nonhistorical, allowed us to calculate the degree of model hybridization at that same portion of the textbook. In a subsequent work (Gericke et al. in press), we described which models of gene function prevailed in each chapter and, then, calculated the degree of hybridization based on false-historical and nonhistorical feature variants. In this work, we will consider only the latter analysis.

The analyses of higher education textbooks were performed by the same researcher in order to increase their reliability, while two other researchers examined all the analyses, comparing part of the results with the original textbooks. In the study about high school textbooks, internal reliability was increased by carrying out independent analyses of the recording units by two researchers (cf. LeCompte and Goetz 1982). Inter-rater agreement between these analyses was high, reaching 89.9 %. The two raters and a senior researcher discussed the diverging categorizations, looking for shared agreement, such that the findings amount to consensus reached by those three researchers. In four instances where no consensus was reached, the recording units were excluded from the analysis.

16.3.2 Study on Higher Education Students' Views About Genes and Their Functions

16.3.2.1 Sample

We investigated the views of 112 biology undergraduate students of two Brazilian universities (Federal University of Paraná, UFPR, hereafter U1 – 60,

Federal University of Bahia, UFBA, hereafter U2 – 52 students) on genes and their functions. The sample from each university was subdivided according to whether or not they had already attended Genetics courses. All students that had already attended Genetics courses had also previously attended Cell and Molecular biology courses.[17]

16.3.2.2 Data Gathering Tool

We employed a questionnaire constructed and validated by ourselves, comprising three sections: (A) students' personal data, including information on his/her experiences on teaching and research training; (B) open and closed questions on genes, challenges to the classical molecular gene concept, and biological information; and (C) closed questions on the gene concept. Sections (B) and (C) contained 11 questions. Due to space constraints, we will consider only the results for three of them. The first is deliberately open ended and divergent, aiming at eliciting a diversity of answers: "In your view, what is a gene?" The other two are closed-ended questions, which were partly derived from Stotz et al. (2004). Both presented the same options for the students to mark, but one was a forced choice, while the other was a free-choice question. Here are the statements that the students could choose with the understanding of genes closer to each shown within brackets (information not available for the students): (a) A gene is a heritable unit transmitted from parents to offspring [Mendelian concept]. (b) A gene is a sequence of DNA which codes for a functional product, which can be a polypeptide or an RNA [Classical molecular concept]. (c) A gene is a structure which transmits information or instructions for development and organic function from one generation to another [Informational conception]. (d) A gene is a determinant of phenotypes or phenotypic differences [Gene-P]. (e) A gene is a developmental resource, side to side with other equally important resources (epigenetic, environmental) [Gene-D]. (f) A gene is a process that includes DNA sequences and other components, which participate in the expression of a particular polypeptide or RNA product [Process molecular concept]. (g) A gene is any segment of DNA, beginning and ending at arbitrary points on the chromosome, which competes with other allelomorphic segments for the region of chromosome concerned [Evolutionary gene concept, sensu Dawkins]. (h) A gene is a sequence of DNA with a characteristic structure [Classical molecular concept]. (i) A gene is a sequence of DNA with a characteristic function [Classical molecular concept]. (j) A gene is a sequence of DNA containing a characteristic information [Informational conception].

The study was approved by the Research Ethics Committee of the Institute of Collective Health/UFBA and by the National Committee of Research Ethics (recording number 12112), and the participants gave informed consent to participate.

[17] In both universities, the biology curriculum includes two courses on Genetics and one course on Cell and Molecular biology.

16.3.2.3 Data Analysis

For analyzing the students' responses to the open-ended question, we used the same technique described in the study about textbooks, categorical content analysis, following the same procedures. In the closed questions, we tabulated the frequencies of the alternatives marked in the forced- and free-choice items.

In order to increase internal reliability, two researchers performed independent analyses of the students' answers to the open-ended questions. Inter-rater agreement between these analyses was not very high, reaching 60 %. It was very important, then, to discuss the differences in categorization between those two researchers. This was done by a group of four researchers, including two senior researchers not involved in the previous analyses. We included in the final analysis only those answers for which shared agreement was possible.

The hybrid answers to the open-ended question were recategorized by three researchers who strived for reaching a consensus concerning the prevailing meaning. Once each answer was classified into a single category, they were analyzed statistically through a chi-square test in order to ascertain whether there were significant differences between the views of students who had attended or not the Genetics and Cell and Molecular biology courses. Thus, we could test the influence of the courses on students' ideas about genes in both universities, including also data from the closed questions. The null hypothesis (H_0) was that the two variables would be independent, i.e., the fact that the students had attended the courses would not affect their views about genes and their functions. H_0 would be rejected when the calculated chi-square was equal to or greater than 9.48, and the alternative hypothesis (H_1) would be accepted, showing influence of the courses on students' views. The significance level (α) was 0.05 and for all questions the degree of freedom was equal to 4.

16.3.3 Investigating a Teaching Sequence on the Problem of the Gene

16.3.3.1 Construction of the Teaching Sequence

The study was conducted in two classes of Medicine freshmen students, who attended in the second semester of 2009 the Cell and Molecular biology course under the responsibility of a teacher-researcher involved in the study, at Federal University of Bahia, located in Northeast Brazil. One class (11 students, 18–24 years) followed an approach employed by the teacher for many years, with no explicit discussion about gene function models and gene concepts (hereafter, class A). In another class (13 students, 15–23 years), the new teaching sequence was implemented, including an explicit discussion on those models and concepts, in a modest but explicit approach to the nature of science (NOS) (Matthews 1998; Abd-El-Khalick and Lederman 2000) (class B). Most students came from households with high and middle income.

Table 16.3 Framework proposed by Mortimer and Scott (2003) for the analysis of interactions and meaning making in science classrooms

Analytical aspects	
i. Teaching focus	1. Teaching purposes 2. Content
ii. Approach	3. Communicative approach
iii. Actions	4. Patterns of interaction
	5. Teacher's interventions

The Cell and Molecular biology course is traditionally divided into two modules, molecular and cellular. Usually, the course includes theoretical and practical lessons and students' seminars. Theoretical lessons comprise a short quiz; an activity oriented by a study guide; teacher's exposition, in which he makes the students feel free to pose questions and raise doubts; small group work, in which selected texts are discussed; and whole class discussion. Practical lessons aim at allowing students to observe cell phenomena and offering them an initiation to lab practices. In the seminars, students are divided into small groups to present selected scientific papers.

The teaching sequence was built collaboratively with the teacher, who has B.Sc. in Biological Sciences and M.Sc. and Ph.D. in Pathology. At the time of the study, he had 17 years of experience teaching this same course.

To construct the teaching sequence, we considered three a priori analytical dimensions (Artigue 1988; Méheut 2005): (1) epistemological, related to the contents to be learned, the problems they can solve and their historical genesis; (2) psycho-cognitive, considering the students' cognitive characteristics; and (3) didactic, linked to the constraints posed by the functioning of the teaching institution (programs, timetables, etc.). The first dimension followed from the historical and philosophical background used in the research program. The second benefited from the collaboration with the teacher, who has a wealth of knowledge on students' previous conceptions, difficulties, etc. Finally, we deliberately constructed the teaching sequence to be compatible with the typical constraints involved in undergraduate Cell and Molecular biology courses, which typically have extensive syllabi in Brazilian universities, with much content to be covered usually in 45–60 h. We planned the teaching sequence to fit into the time made available by the teacher, 5 h distributed in 2 days of classes. Within these time constraints, he assured us, it would be more feasible that the proposal could be used in most similar courses.

We used a discourse analysis perspective to plan the activities, designing communicative approaches and interaction patterns to be used by the teacher. The framework for classroom discourse analysis developed by Mortimer and Scott (2003) was adapted for this goal. It is based on five interrelated aspects that focus on the teacher's role, grouped in three dimensions: *teaching focus*, *communicative approach*, and *actions*. The *communicative approach* is the central element, since it is through it that we understand how the teaching focuses are worked, i.e., the *teaching purposes* and *contents*, by means of which actions, the *pedagogical interventions*, which result in certain *patterns of interaction* (Table 16.3).

The investigation was framed in the context of educational design research (Baumgartner et al. 2003; Plomp 2009; van den Akker et al. 2006), which aims at both developing educational interventions and advancing our knowledge about their characteristics and the processes of designing and developing them. The main research question in educational design research is to establish what are the characteristics or design principles of an intervention *X* for obtaining the outcome $Y(Y_1, Y_2, \ldots, Y_n)$ in context *Z* (Plomp 2009). Design principles are initially derived by us from the relevant literature and practitioner knowledge, and as the investigation of a series of prototypes of the teaching sequences proceeds, we not only test the initial design principles but also derive additional principles from the empirical results.

At this point, we tested just the first prototype of the teaching sequence in a single classroom. The following design principles were used: (1) The classroom discursive interactions were planned to flow from a dialogical approach, in which students' ideas played a prominent role in meaning making, to a more authoritative approach, in which the diversity of ideas raised was subjected to evaluation and selection in order to construct in the classroom the perspective of school science; (2) in classroom discursive interactions, the teacher stressed key ideas when they appeared, in order to construct the school science perspective around them; (3) texts produced by ourselves, aiming at the didactic transposition of debates on genes and their functions, were provided to small groups of students, alongside with guiding questions; (4) the teaching sequence used a historically and philosophically informed approach, putting emphasis on the role of models in science, their relation with reality, and the importance of their demarcation; (5) several historical models of gene function and gene concepts were explicitly addressed and differentiated; (6) the crisis of the classical molecular concept was explicitly discussed, as well as reactions to it; (7) in order to discuss this crisis, the teacher used molecular phenomena already addressed in his classes previously, even though at that point no conceptual consequences related to genes were derived.

16.3.3.2 The Teaching Sequence

The teaching sequence adopted an explicit approach to the NOS in the context of teaching about genes and their function, seeking to promote learning *with* models and *about* models.

The first class begins with the teacher asking the students what is a gene, an open-ended and divergent question intended to raise as many students' conceptions as possible. The teacher avoids evaluative comments or gestures, in order to maintain the dialogical interaction with the pupils. As the students offer their answers, the teacher copies them in the blackboard to be used later. This activity is followed by an exposition about models and their role in science. Even though the teacher speaks most of the time, he prompts the students to participate by posing questions. The students are divided into small groups and receive the first text prepared by our team, "historical models of the gene concept" (text contents are similar to those

found in Sect. 16.2 above), followed by a number of guiding questions for cooperative discussion. The answers are used by the teacher to promote whole class discussion, which creates the opportunity to prompt sharing of the discussions in the small groups, to check the students' understanding and to stress key ideas for the construction of the intended perspective on genes. He goes back then to the students' initial answers, available in the blackboard, discussing the relationship between their ideas and the historical models about gene function. Now he evaluates their answers, showing when they are closer to one or another model and pointing out which models are still accepted and in what features. He also stresses which answers are distant from any scientific model and brings to the fore the hybrid models, if they are present in the students' answers. The expectation is that, at the end of the class, the diversity of students' ideas raised and the diversity of scientific models about genes have been systematized.

In the second class, the teacher begins by briefly reviewing the previous session and posing questions for the students in order to evaluate their understanding. Then, he makes an exposition on the crisis of the classical molecular concept, using challenging phenomena that were already discussed in the previous classes, such as alternative splicing and gene overlapping. The students are divided again into small groups, receiving the second text we prepared, "proposals for the gene concept" (text contents are similar to those in Sect. 16.2), with guiding questions. Again, this is followed by whole class discussion. The class ends with a discussion on the current status of our understanding about genes, in which the teacher highlights the idea that the classical molecular concept is in crisis, but none of the proposals discussed in the second text are widely accepted by the scientific community. The intended perspective on genes is arguably clear for the students: the gene concept is now changing under our very noses, with all directions of change still being debated. The teacher also takes a last opportunity to stress the existence of a diversity of gene concepts and models of gene function, claiming that several models show greater explanatory and heuristic powers than a single, overarching definition of gene, provided that we properly demarcate their domains of application.

16.3.3.3 Teaching Sequence Validation

We performed a posteriori internal and external validation of the teaching sequence (Artigue 1988; Méheut 2005). In the internal validation, we compared the effects of the teaching sequence in relation to its goals, by comparing the students' learning outcomes with the planned learning goals. To perform this comparison, we investigated how the students mobilized ideas about genes and their function in a discursive context structured by a subset of the items from the questionnaire used to investigate students' views (see above), with some modifications validated in a pilot test. Here we will discuss the same three questions mentioned above. In the closed questions, the alternative (g), related to the evolutionary gene concept, was excluded in this study. The questionnaire was used in three moments: in the second lesson of the whole course, when we could probe students' views with no influence of the course

(pretest); at the end of the molecular module, which coincided with the last day of the teaching sequence (posttest); and two months after the intervention (retention test). The classes have also been video recorded to provide raw material for the analysis of classroom discursive interactions, but these data have not been treated yet.

In the internal validation, we are evaluating if the teaching sequence does reach the planned learning goals. If we use the framework presented by Nieveen et al. (2006), this is a development study, aiming at solving educational problems by focusing on the proposal and testing of broadly applicable design principles. The goal is to understand how and why a given intervention functions in the particular context in which it was developed. It is this knowledge that is summarized in design principles (Reeves 2006; van den Akker et al. 2006), or intervention or design theories (Barab and Squire 2004), which are expected to generalize beyond the context of the study. Although we cannot expand further on the topic here, we should mention that this knowledge is conceived by us as generalizing in two (not mutually exclusive) ways: (1) through situated generalization (Simons et al. 2003), i.e., the transformation of data gathered in a context into evidence transferable to other contexts, so as to indicate a course of action or be incorporated in judgments preceding action, due to teachers' perception of a connection between the investigated context and the context of their pedagogical work, and (2) as a generalization resulting from maximizing the variation of qualitatively different investigated cases (Larsson 2009). As we investigated only the first prototype of the teaching sequence, the second kind of generalization is not yet at reach. However, the first kind of generalization is already feasible, since other college and university teachers may perceive the same problems discussed here in their classrooms and, eventually, see in the teaching sequence a putative approach to their pedagogical practice.

We also performed a *preliminary* external validation of the sequence by comparing the effects of the teaching sequence with the approach employed for many years in the course. The same questionnaire was applied for class A in the same moments mentioned above. Using Nieveen and colleagues' (2006) framework, this is an effectiveness study, which can provide evidence for the impact of the intervention by comparing its effectiveness in relation to another teaching approach. As Brown (1992) argues, our goal in such a study should be to accommodate variables rather than controlling them, since research needs to occur within the natural constraints of real classrooms. One manner of accommodating confounding variables is to use sufficient numbers of replicas of each treatment such that we can distinguish between the effects of the intervention and confounding variables randomly assorted to the replicas, such as students' motivation, the quality of their previous knowledge, and teacher-students relationships. But when we do research in real educational contexts, we often do not count with enough number of classes for replicating treatments. This was the case in our study, since there was only one teacher interested in engaging in it, and he had only two courses under his responsibility. This means that we cannot sufficiently distinguish between the effects of the teaching sequence and confounding variables, although we had the same teacher and similar sets of students in the two classes. Nevertheless, the results revealed interesting patterns, although preliminary and to be taken with a grain of salt.

16.3.3.4 Data Analysis

The answers to the questions included in the tool were treated through categorical analysis (open-ended item) and tabulation (closed item) as described above. Internal reliability was increased in the open-ended question by independent analyses by two researchers, with high inter-rater agreement (89.1 %). Differences in categorization were discussed with two other researchers (one of them also the teacher of the course), and the final analysis included only those answers in which shared agreement was reached.

16.4 Results and Discussion

16.4.1 Textbook Studies

16.4.1.1 Views About Genes in Higher Education Cell and Molecular Biology Textbooks

Figure 16.1 shows the distribution of gene concepts in the three higher education Cell and Molecular biology textbooks we analyzed (Pitombo et al. 2008).

In Karp's (2004) textbook, there were 73 recording units explicitly addressing gene concepts, considerably more than in the other two books (35, Alberts et al. 2002, 23, Lodish et al. 2003). This follows from the fact that the former book focuses on concepts and experiments, as shown by its subtitle, giving more attention to history. Symptomatically, the Mendelian conception, according to which the gene is a unit of inheritance, showed the highest prevalence (31.5 %), and most of these occurrences were in sections discussing the history of genetics. The Mendelian conception is mostly treated in this textbook as a view on genes that is historically relevant, but is not often used to account for current perspectives on genes, which are frequently represented by the second most frequent view (24.6 %), the informational conception, in which the gene is seen as a unit or carrier of information. Since information is a metaphorical notion that still needs theoretical clarification in genetics (El-Hani et al. 2009; Griffiths 2001), it is problematic to appeal mainly to this idea to explain what genes are. The third more frequent concept in Karp was gene-P (20.5 %), which was mostly used in sections about the history of genetics and medical genetics, where it usefully abstracts away from the complexities of the genotype-phenotype relationship, focusing on the predictive relationship between gene loci and pathological conditions. Finally, the classical molecular concept appeared in 13.7 % of the recording units, distributed in a wide variety of contexts, including molecular biology, evolutionary genetics, genetic engineering, and genomics, besides historical narratives about genetics. We can say, therefore, that in this textbook, when genes are described in molecular terms and from an updated perspective, the molecular-informational model of gene function prevails.

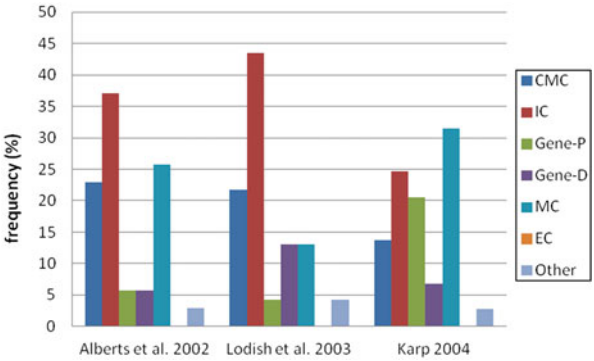

Fig. 16.1 Distribution of gene concepts in three higher education Cell and Molecular biology textbooks. *CMC* classical molecular concept, *IC* informational conception, *MC* Mendelian conception, *EC* evolutionary concept

Alberts et al. (2002) and Lodish et al. (2003) are much less diversified in their treatment of genes, even though they still show conceptual variation. In these textbooks, the informational conception was remarkably predominant (37.1 %, Alberts et al.; 43.5 %, Lodish et al.), being frequently associated with the classical molecular concept (22.9 %, Alberts et al.; 21.7 %, Lodish et al.). Their basic message about the nature of genes amounts, thus, to a combination of the metaphor of information and the idea of the gene as unit of structure and/or function in DNA, which is characteristic of the molecular-informational model.

In all the textbooks, the classical molecular gene concept was predominantly used when they were addressed contents related to Molecular biology and Molecular Genetics. This concept was also used by the three textbooks in their glossaries, in order to define genes. The informational conception, in turn, was found in more diversified contexts in the textbooks, when compared with the classical molecular gene concept, indicating how widespread this conception is, despite its lack of solid theoretical background.

However, the prevalence of the molecular-informational model sounds strange in the three textbooks, when we consider that they discuss the anomalies challenging it in the last decades. The conceptual lessons following from these empirical findings are not taken into account, yet another indication of a largely atheoretical and ahistorical treatment of the contents. Despite the presence of conceptual variation, these textbooks do not provide clues for teachers and students about the distinct origins, domains of application, and meanings of concepts related to different models along the history of genetics and molecular biology. Thus, hybridization of incommensurable aspects of different models and semantic confusion are likely to happen. This is a good case in point regarding the harmful consequences of teaching science without teaching about science. The students do not have much chance of learning with models and about models, since these textbooks address the contents as if they referred to reality themselves, as discovered by science, not to models about reality, historically constructed by the scientific community. The relationship between model and reality becomes unclear when most of the explanations just consider what *is* in the world, not how we interpret what *is* in the world based

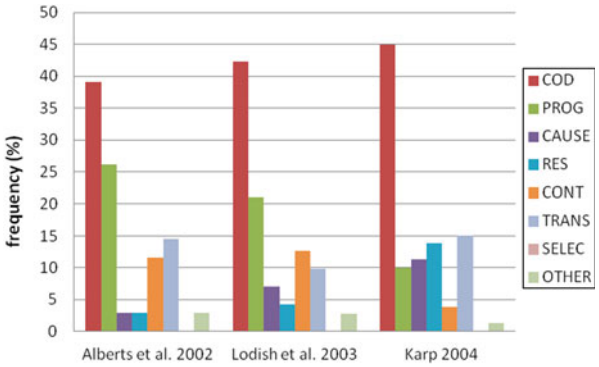

Fig. 16.2 Distribution of functions attributed to genes in three higher education Cell and Molecular biology textbooks. *COD* codifying the primary structure of polypeptides or RNAs (classical molecular concept), *PROG* program or instruct cellular functioning and/or development (informational conception), *CAUSE* cause or determine phenotype or difference between phenotypes (gene-P), *RES* act as a resource for development (gene-D), *CONT* control cell metabolism (informational conception), *TRANS* transmit hereditary traits (Mendelian conception), *SELEC* act as unit of selection (evolutionary concept)

on theoretically laden evidence and inferences (which are often conflated in the textbooks with observations).

As an example, the following definition hybridizes features related to the Mendelian and the informational conception:

> Gene - Physical and functional unit of heredity, which carries information from one generation to the next (Lodish et al. 2003, Glossary, G-9).

This sentence, in turn, hybridizes gene-P and the informational conception:

> These instructions are stored within every living cell as its genes, the information-containing elements that determine the characteristics of a species as a whole and of the individuals within it (Alberts et al. 2002, p. 191).

The harmful consequences of combining these different features of historical models become apparent, as the idea of "genetic information" is taken to imply a reduction of the development of all characteristics of the species and the individuals to the DNA nucleotide sequences. We can explicitly see the connection between the genetic determinism that often marks gene talk in the social arena and the way genes are treated in these textbooks.

The interpretation that the molecular-informational model prevails in these textbooks is reinforced when we examine the functions attributed to genes (Fig. 16.2). In all of them, the function most frequently ascribed is codifying the primary structure of polypeptides or RNAs, aligned with the classical molecular concept (39.1 %, Alberts et al.; 42.3 %, Lodish et al.; 45 %, Karp). In the former two textbooks, the second most frequent function, to program or instruct cellular function and/or development, is also related to that model, namely, to the informational conception (26.1 %, Alberts et al.; 21.1 %; Lodish et al.). In Karp, to transmit hereditary

traits is the second most common function (15 %), consistently with the high prevalence of the Mendelian conception.

Generally speaking, we observe a proliferation of meanings attached to genes as we progress from context to context in these textbooks, with no model unification or demarcation. This happens both in gene concepts and function ascription to genes.

16.4.1.2 Views About Genes in High School Biology Textbooks

Figure 16.3 shows the distribution of gene concepts in 18 Brazilian high school biology textbooks, including those approved and not approved by the Brazilian National Program for High School Textbooks (PNLEM) (Santos et al. 2012).

In these textbooks, three gene concepts were significantly more prevalent: the classical molecular concept, the informational conception, and the gene-P. In 12 of the 18 textbooks, gene-P was the most frequent, answering for more than 40 % of the recording units in 4 textbooks. The classical molecular concept and the informational conception were more prevalent in 3 textbooks each.

The fact that gene-P is so often used in these textbooks follows from the extensive content of the genetics chapters, where we find several examples of pedigree analyses and estimates of the inheritance probability of phenotypic traits. Here is an example of a recording unit showing gene-P:

> The gene for brown eyes located in the chromosome is an allele of the gene for green eyes, located in the homologous chromosome (T2, vol. 3, p. 15).[18]

Gene-P is often employed in the textbooks just as it was used in classical genetics, when genes were inferred from phenotypes. However, these statements are framed in an "updated" language, and thus, teachers and students cannot figure out that the textbook is using a way of understanding genes that was frequently employed when there was no established knowledge on the nature of the genetic material. Moreover, a key requirement for a valid usage of genes-P is not found in these textbooks, namely, a clear understanding of the distinction between this instrumental concept and a realist interpretation of the genetic material. In the absence of this distinction, gene-P is simply conflated with the classical molecular gene concept, which provides then a molecular background to understand genes as determinants of phenotypes, as expressed by gene-P. The kind of conflation that Moss (2001, 2003) identifies as a source of genetic determinism, between a preformationist instrumental concept (gene-P) and a molecular realist concept (gene-D), is favored by the way these textbooks deal with genes.

It is this sort of hybridization between features related to different models that can lead to semantic confusions, hampering students' understanding and favoring ideas with important socioscientific implications, such as genetic determinism. If a student learns that genes determine phenotypes in the absence of a historically and

[18] All translations of textbook passages from Portuguese were made by the authors of the present paper. Commentaries by the authors are shown in brackets.

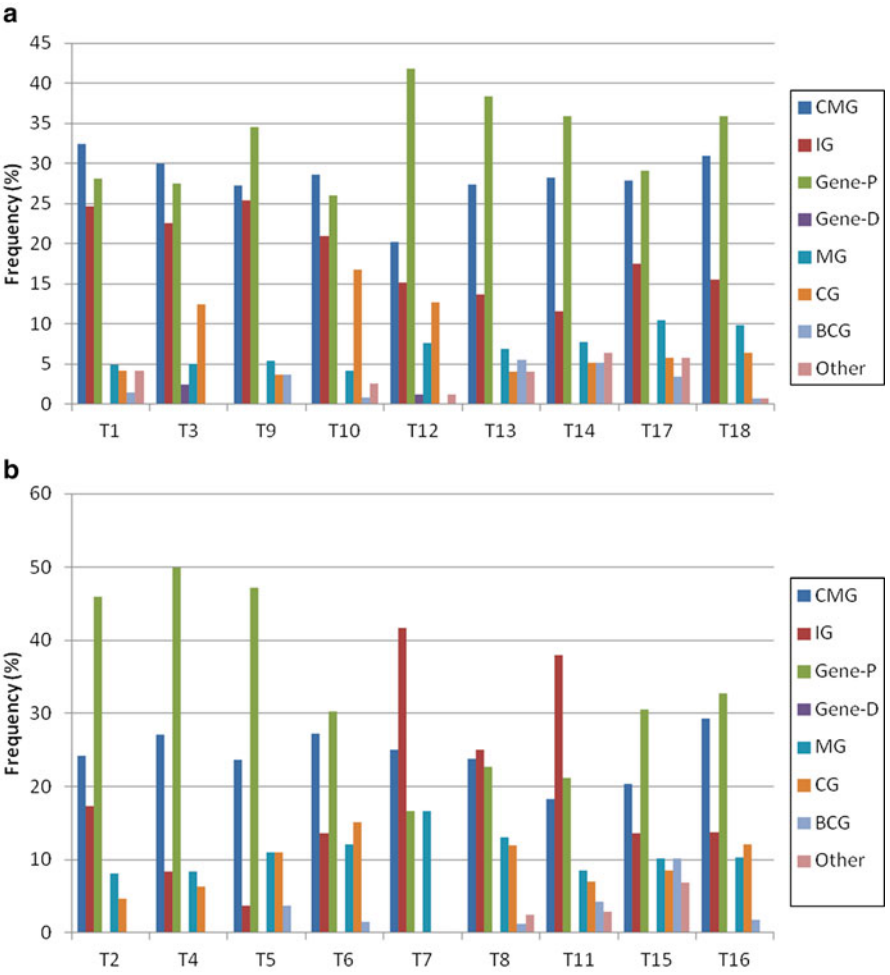

Fig. 16.3 Distribution of gene concepts in Brazilian high school biology textbooks. *CMG* classical molecular gene, *IG* informational gene, *MG* Mendelian gene, *CG* classical gene, *BCG* biochemical-classical gene. (**a**) Textbooks approved; (**b**) textbooks not approved by PNLEM. Textbooks are indicated by the codes listed in Appendix 1

epistemologically informed discussion of the role of this instrumental concept in classical genetics and then moves on to study about genes depicted in a realist manner as structural and functional units in DNA, the conflation between these two concepts and the resulting semantic confusions seem almost inevitable.

Symptomatically, in all textbooks in which gene-P prevails, the second most frequent concept was the classical molecular gene. Moreover, in 39.1 % of the recording units where we found the classical molecular gene, gene-P was also present. The classical molecular concept only entails colinearity between a gene

and the primary structure of a protein or RNA but does not fix the relationship between genes and phenotypes at a higher level. This relationship enters the textbook explanation through the hybridization with gene-P, predictably leading to genetic determinism. The passage below illustrates the hybridization between the classical molecular gene and gene-P, with clear determinist undertones:

> Currently we know that the gene […] is a sequence of nucleotides in DNA. Each gene is responsible for the synthesis of a protein and, consequently, for one or more characteristics of the individual, since proteins can have structural and regulatory functions in metabolism. Genes are located in chromosomes and are didactically represented by letters, numbers, and symbols. For instance, the gene for normal skin color is symbolized by *A* and the gene for albinism, by *a* (T6, p. 283).

This amalgam of a preformationist view of the gene as determinant of phenotypes and a molecular view of the gene as information carrier located in DNA is the major picture of the gene in these textbooks. The classical molecular concept, in particular, was found in the most diverse contents in the textbooks, in all three high school years, with relatively high frequency (Santos et al. 2012).

In Fig. 16.4, we can see the functions attributed to genes in the high school biology textbooks we analyzed. In almost all textbooks (17), genes are most often regarded as codifiers of the primary structure of polypeptides or RNAs (in accordance with the classical molecular concept) and determinants of phenotypes (in line with gene-P).

All the historical models identified by Gericke and Hagberg (2007) were found in the textbooks (Fig. 16.5), showing how they are marked by conceptual variation. The molecular-informational model was dominant, in keeping with the prevalence of the classical molecular concept and the informational conception in the textbooks. However, the difference of prevalence between the four most frequent models is in fact quite small, highlighting how the predominant feature of these textbooks is, in fact, conceptual variation, with no clear demarcation between the different models and their domains of application. Gericke and colleagues (in press) compared the distribution of these historical models in a large and significant sample of Swedish and Brazilian textbooks, as well as in 7 textbooks used in English-speaking countries. Despite some differences, the distribution of the different models within the textbooks of the different countries was very similar. They interpret this finding as showing that the conceptual variation in genetics is captured in a similar textbook discourse that is culturally independent, that is, didactic transposition (Chevallard 1989) leads to similar end products in those different countries, maybe as a consequence of the influence of the higher education textbooks used by textbook authors to learn about genetics and cell and molecular biology.

Half of the high school textbooks analyzed (9) discussed split genes. To our understanding, six of them treated split genes and splicing in a satisfactory manner. However, only three considered alternative splicing, and among the latter, only two discussed the conceptual implications of this phenomenon to the way genes are conceived.[19] This indicates that, in spite of the overwhelming predominance of an outdated

[19] It is worth noting, however, that none of the higher education cell and molecular biology textbooks offered such a discussion.

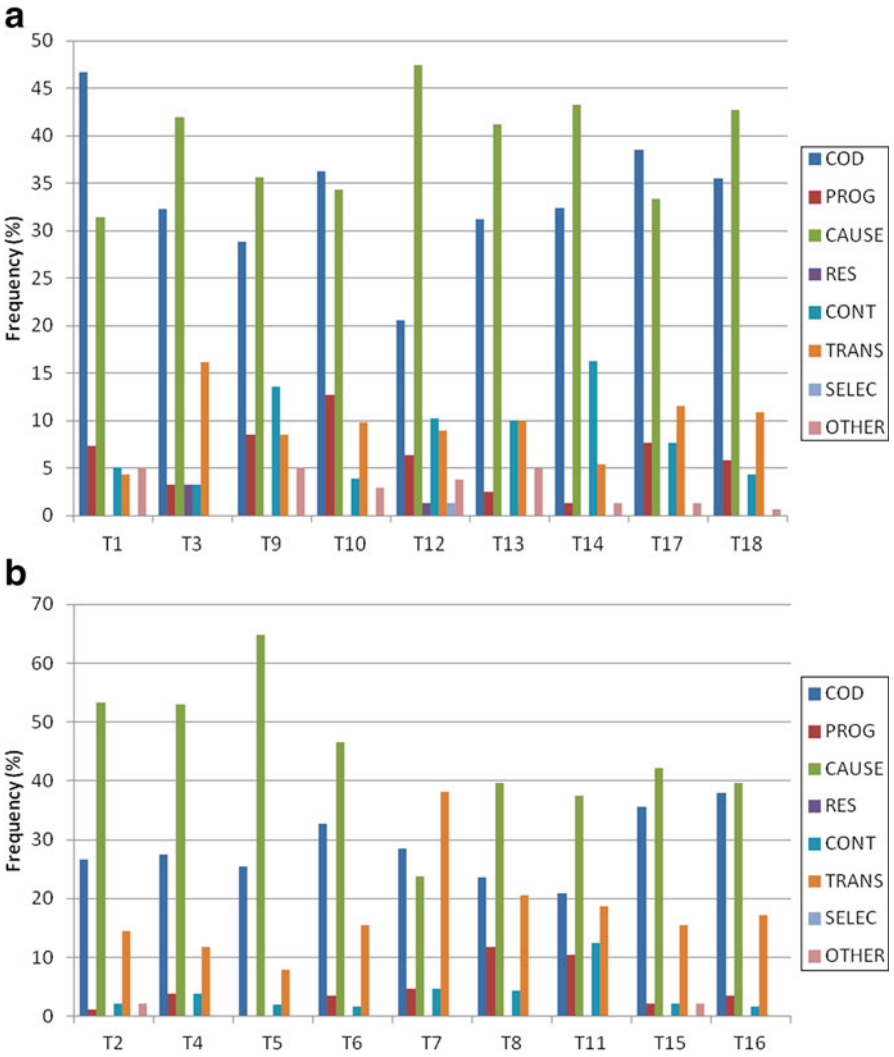

Fig. 16.4 Distribution of functions attributed to genes in Brazilian high school biology textbooks. *COD* codifying the primary structure of polypeptides or RNAs (classical molecular concept), *PROG* program or instruct cellular functioning and/or development (informational conception), *CAUSE* cause or determine phenotype or difference between phenotypes (gene-P), *RES* act as a resource for development (gene-D), *CONT* control cell metabolism (informational conception), *TRANS* transmit hereditary traits (Mendelian conception), *SELEC* act as unit of selection (evolutionary concept). (**a**) Textbooks approved; (**b**) textbooks not approved by PNLEM. Textbooks are indicated by the codes listed in Appendix 1

approach to the gene concept, at least in some textbooks, there seems to be an ongoing transition to a more updated treatment. However, in the majority of the high school textbooks, the case is similar to that of higher education textbooks: when the challenges to the classical molecular concept are discussed, relatively

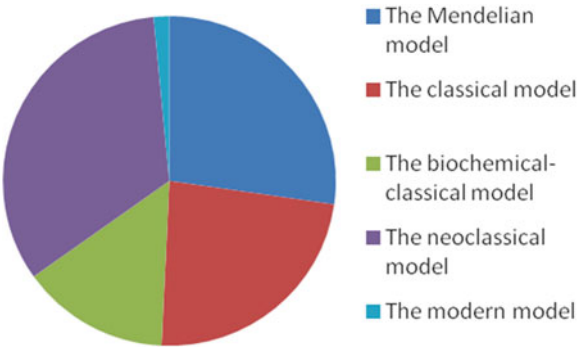

Fig. 16.5 Distribution of the historical models identified by Gericke and Hagberg (2007) in Brazilian high school biology textbooks (in percentage)

obvious conceptual consequences are not considered. This can be seen as a consequence of the way the textbooks typically approach scientific knowledge, as a list of isolated facts, building a fragmented rhetoric of conclusions (Schwab 1964).

When using the vast majority of these textbooks, students and teachers cannot get even a glimpse of the state of affairs in current discussions about genes. Some may think that it is too much to demand that school science considers these recent developments at high school. However, for most students this may be the last opportunity to learn about genes and their function and, thus, to build a critical stance towards gene talk in socioscientific issues, from the safety of genetically modified organisms to the use of genetic testing in society.

We also did a systematic analysis of model hybridization in the high school biology textbooks, finding a widespread use of hybrid models for describing gene function (Table 16.4), often combining features of models focusing on the molecular and cellular level with features of models dealing with the phenotypic level, derived from classical genetics. As Santos and colleagues (2012) show, the molecular-informational model seems to be taken as a basis by the textbooks, with features from a variety of models being hybridized with it. Thus, conceptual variation, although present in the textbooks, is not explicitly dealt with, being difficult for teachers and students to realize that different aspects of gene function are mixed up and, in particular, to take notice of the ambiguities, logical inconsistencies, and semantic confusions that may follow.

16.4.2 Higher Education Students' Views About Genes and Their Functions

The Biological Sciences students who participated in the study about their views about genes and their functions were divided into two groups, depending on whether they attended (YG) or not (NG) Genetics courses. In one of the universities investigated, located at the South part of Brazil (UFPR, U1), the distribution was 32 students in group YG and 28 in NG. In another university included in the study,

Table 16.4 Hybridization frequency of textbook models

	Mendelian model	Classical model	Biochemical-classical model	Molecular-informational model	Modern model
Level of hybridization (%)[a]	7.7	18.4	9.5	41.8	

[a]The level of hybridization equals the frequency of exchanged epistemological feature variants, calculated as the number of incorrect historical feature variants (nonhistorical and false historical) divided by the total number of feature variants in the textbook models

located in the Northeast region of Brazil (UFBA, U2), we had 19 students in YG and 33 in NG.

The chi-square test performed to statistically analyze the influence of the Genetics course on students' ideas about genes in both universities resulted in the values 9.83 and 10.07 in U1 and U2, respectively. Thus, in both universities, a significant relationship was found between the students' attendance to the Genetics courses and the views about genes expressed in their answers.

Figure 16.6 shows the distribution of the answers in the categories obtained in the analysis of the open-ended and divergent question "In your view, what is a gene?" for the two universities and the two groups.

Regarding the classical molecular concept and the informational conception, the results show similar effects of the Genetics courses on Biological Sciences students' views in the two universities. They led to a significant increase in the percentage of answers committed to the classical molecular concept and a decrease in the students' commitment to the informational conception, with the difference that only a slight decrease took place at U1.

On the one hand, if we consider that basically all the challenges faced by the classical molecular concept are addressed by those courses, we can suspect that no connection is made between examining empirical findings in genetics and cell and molecular biology and reflecting on their conceptual implications. This may be a consequence of the lack of an epistemological and historical dimension in the teaching practice in those courses. On the other hand, the impact they had on the students' appeal to the informational conception is a positive consequence of the courses, which can be attributed to the fact that the students are stimulated to delve into more details regarding the structure and function of the genetic material. This can be associated to both the increase in their allegiance to the classical molecular concept and the decrease in their use of the informational conception.

As an example of a student's answer committed to the classical molecular concept, we can quote[20]:

> It is a fragment of DNA responsible for codifying a polypeptide chain or RNA (U1, student 20, YG).

[20]The answers were freely translated from Portuguese to English by the authors of the paper.

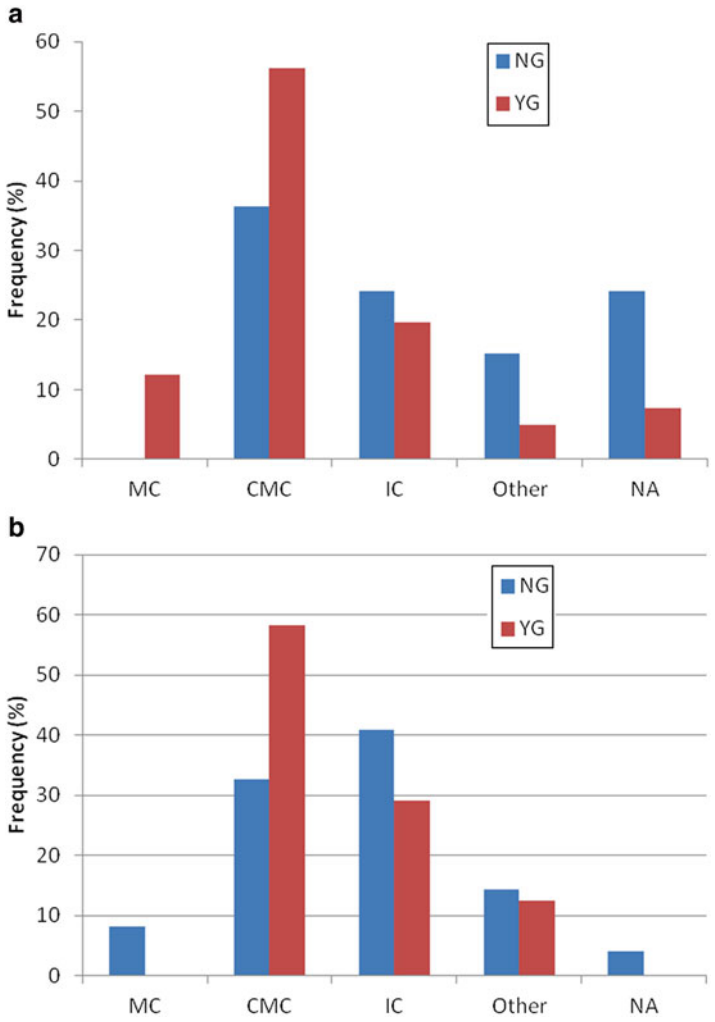

Fig. 16.6 Distribution of answers given by students of two Brazilian universities to the question "In your view, what is a gene?" *MC* Mendelian conception, *CMC* classical molecular concept, *IC* informational conception. (**a**) U1 (UFPR); (**b**) U2 (UFBA). The number of answers is larger than the number of students because there were answers which combined more than one view about genes and, thus, were classified in more than one category

Here is an example, in turn, of an answer exhibiting the informational conception:

Hereditary informational unit (U1, student 3, YG).

Different views about genes were often hybridized by the students in their answers (21.7 % of the answers in U1, 38.5 %, in U2). This suggests that the students may be reproducing the hybrid views about genes found in textbooks (see

Sect. 16.4.1). As there was no trend of decrease of such hybridization after the Genetics courses, classroom teaching and learning seems to be unable to overcome this difficulty posed by the treatment of genes and their functions in the textbooks.

In the closed questions, we used the classification of the alternatives into gene concepts shown in the methods section and, additionally, gathered less represented answers, related to gene-P, gene-D, and the evolutionary gene concept, into a single category, other gene concepts. When considering the forced-choice question, we can see the same pattern observed in the open-ended question regarding the prevalence of the classical molecular concept (particularly, item b, Fig. 16.7. In items h and i, also related to this concept, there were no important changes) and the decrease of the informational conception (items c and j, Fig. 16.7) after the students attended the courses.

In both universities, the students' commitment to the Mendelian conception, as shown by the closed questions, decreased (item a, Fig. 16.7). This may be a consequence of the impact of the molecular treatment of genes during the courses.

Now, compare Fig. 16.7 with Fig. 16.8, which shows the results for the very same closed question, but in a free-choice format. The pattern that is readily apparent is that the students marked a large variety of views about genes when they are allowed to do so. To our understanding, this is a striking evidence that conceptual variation regarding genes, as represented in higher education and high school textbooks, can be translated into students' allegiance to several different accounts about genes and their functions. In itself, the results from these two questions do not allow us to conclude that students are facing difficulties with this conceptual variation, for instance, not knowing what views about genes are more adequate to deal with what sorts of problems, or being entangled in semantic confusions and ambiguities following from combining incommensurable perspectives embraced by different models and concepts. But consider that teaching about genes in those courses uses the textbooks we analyzed, where a historically and epistemologically informed approach to models about genes and their function is typically lacking. It is at least plausible, then, to interpret the fact that the students marked so many different views about genes in the free-choice question as meaning that they are prone to conflate incommensurable aspects of models and, also, to misapply these models, using them outside their domain of validity.

16.4.3 From Diagnosis to Intervention: A Teaching Sequence on the Problem of the Gene

Our previous study on higher education students' views about genes and their functions suggested several shortcomings in teaching about genes at Genetics courses in two Brazilian universities. Part of the limitations of these courses could be attributed to the lack of an epistemological and historical dimension in the treatment of the contents, in particular, to an insufficient attention to teaching both *with* models and *about* models (Gericke and Hagberg 2007).

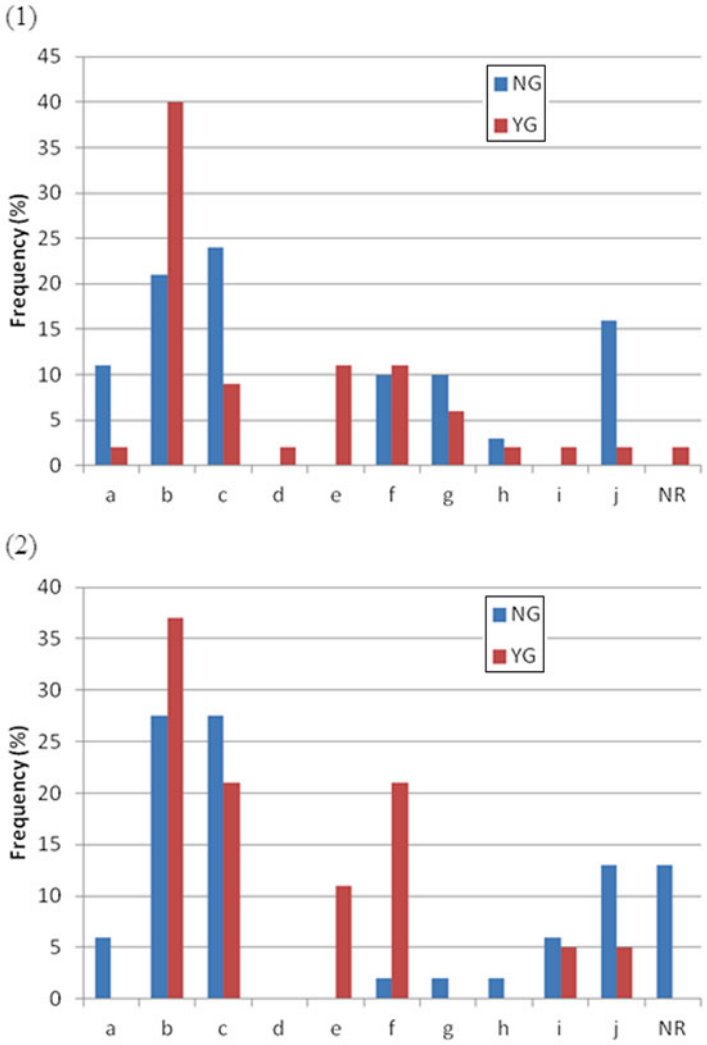

Fig. 16.7 Distribution of answers given by students of two Brazilian universities to a forced-choice closed question presenting several alternatives concerning the nature of genes: (**a**) Mendelian; (**b**), (**h**), and (**i**) classical molecular; (**c**) and (**j**) informational; (**d**) gene-P; (**e**) gene-D; (**f**) process molecular gene; (**g**) evolutionary gene concept. *NR* no response. (1) U1 (UFPR); (2) U2 (UFBA)

Therefore, it seemed natural to us to move from diagnosis to intervention, through the development and investigation of a teaching sequence built collaboratively with a higher education Cell and Molecular biology teacher at the Federal University of Bahia, located in Northeast Brazil. As presented in the Methods section, this teaching sequence explicitly addressed NOS contents, in particular, the historical construction and nature of gene function models and gene concepts. Our intention

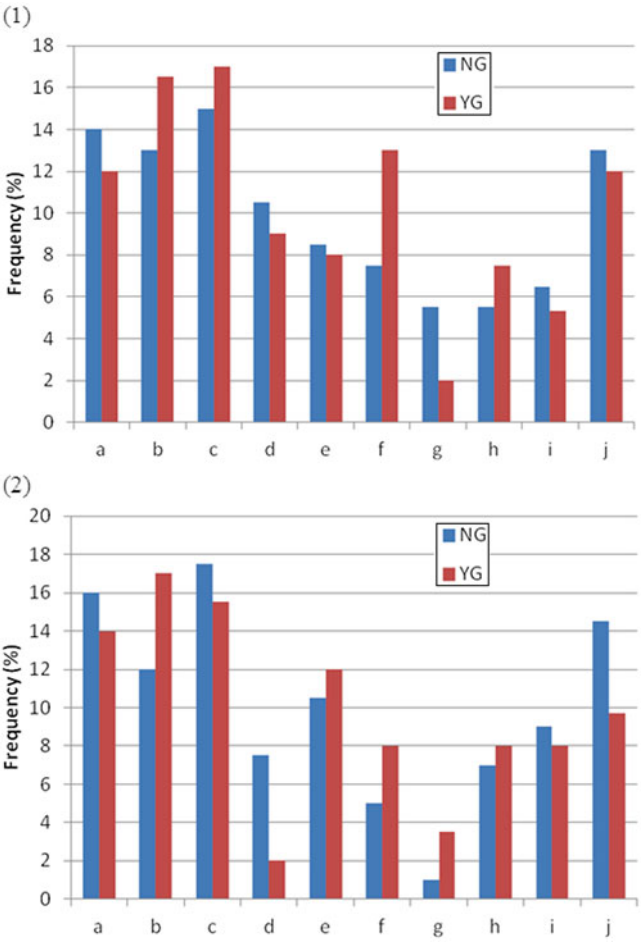

Fig. 16.8 Distribution of answers given by students of two Brazilian universities to a free-choice closed question presenting several alternatives concerning the nature of genes: (**a**) Mendelian; (**b**), (**h**), and (**i**) classical molecular; (**c**) and (**j**) informational; (**d**) gene-P; (**e**) gene-D; (**f**) process molecular gene; (**g**) evolutionary gene concept. (1) U1 (UFPR); (2) U2 (UFBA)

was not to deal with complex historical, philosophical, or sociological issues, but just to teach with models and about models when dealing with genes, as a way of providing conditions for the students to understand that genes have been and are still conceived in different ways in distinct subfields of biology, as a consequence of different epistemic practices that characterize the works of diverse scientific communities.

Figure 16.9 shows the distribution of the answers in the categories obtained in the analysis of the open-ended question "In your view, what is a gene?" in the three

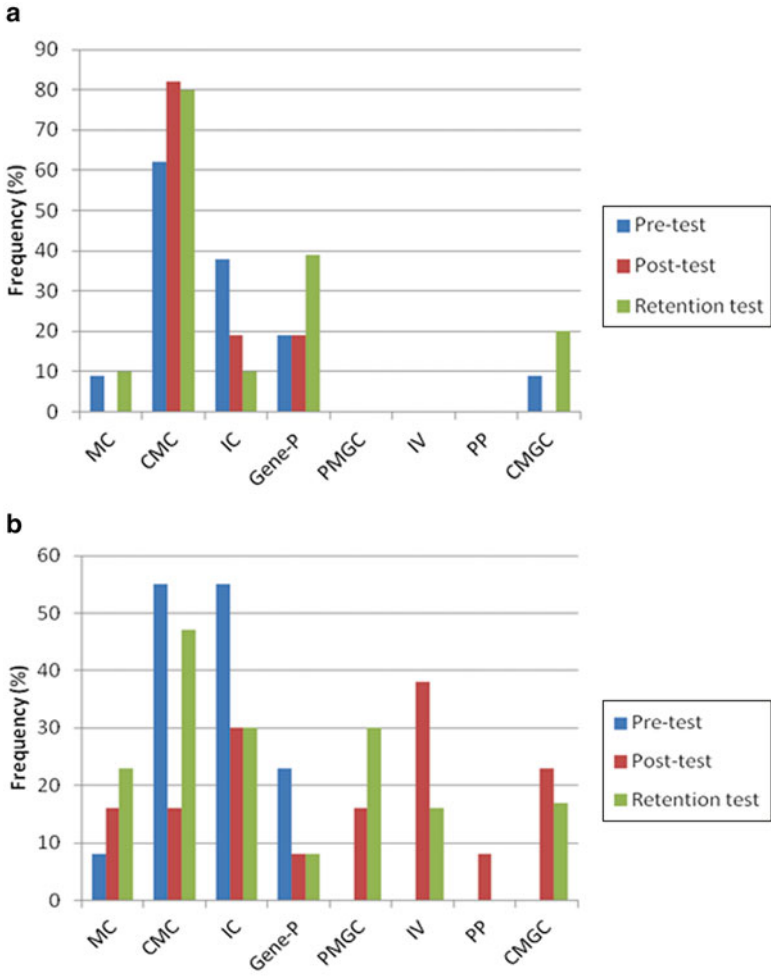

Fig. 16.9 Distribution of answers given to the question "In your view, what is a gene?" by the students of the classes investigated. *MC* Mendelian conception, *CMC* classical molecular concept, *IC* informational conception, *PMGC* process molecular gene concept, *IV*, instrumental view about genes, *PP* perception of the problem, *CMGC* contemporary molecular gene concept (The "contemporary molecular gene concept" amounts to a conservative response to the problem of the gene, which regards the gene as a linear DNA sequence but abandons the idea that it has a single developmental role, defining it, for instance, as "a DNA sequence corresponding to a single 'norm of reaction' of genes products across various cellular conditions" (Griffiths and Neumann-Held 1999, p. 658)). (**a**) Class A (usual approach to the course, with no explicit discussion on gene function models and gene concepts); (**b**) class B (where the teaching sequence was implemented). The number of answers is larger than the number of students because there were answers which combined more than one view about genes and, thus, were classified in more than one category

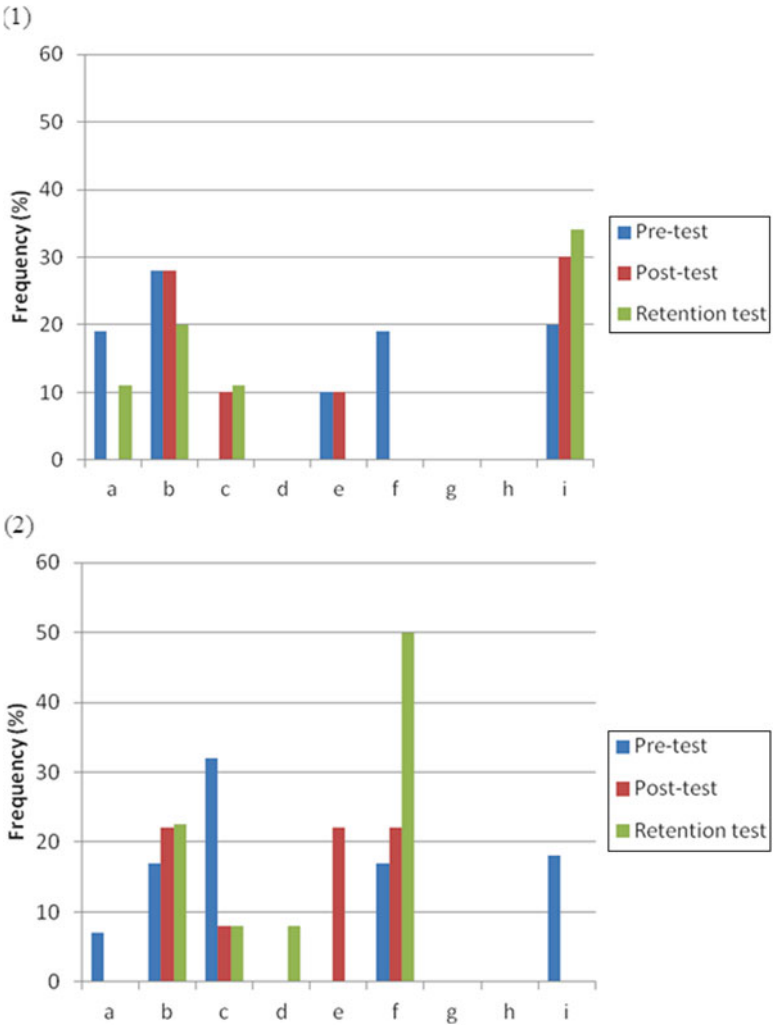

Fig. 16.10 Distribution of answers given by the students of the classes investigated to a forced-choice closed question presenting several alternatives concerning the nature of genes: (**a**) Mendelian; (**b**), (**g**), and (**h**) classical molecular; (**c**) and (**i**) informational; (**d**) gene-P; (**e**) gene-D; (**f**) process molecular gene. (1) Class A (usual approach to the course, with no explicit discussion on gene function models and gene concepts); (2) class B (where the teaching sequence was implemented)

moments in which the data were gathered. It is interesting to look at these results alongside with those for the closed forced-choice question, which allowed us to survey students' ideas about genes using a different kind of tool. We can see the distribution of answers in the pretest, posttest, and retention test in Fig. 16.10.

Considering, first, the internal validation of the teaching sequence, we can see some positive learning outcomes, compared to the intended learning goals: first, the informational conception was successfully challenged by the teaching sequence, falling in the posttest and maintaining the lower frequency in the retention test, when compared with the pretest, both in the open and in the closed forced-choice question. Here is an example of a students' answer committed to the informational conception and, also, showing a close relationship between this conception and genetic determinism:

> Gene is the unit of data storage of the species. The union of the genes (which are in DNA) forms the genome, where we find all the information for the development of the being (Student 2, Class B, pre-test).

Second, the students showed an enriched repertoire of views about genes after the intervention. For instance, the process molecular gene concept increased in frequency in the posttest, reaching an even higher frequency in the retention test, both in the open and the closed forced-choice question. An instrumental view about genes was considered by a significant proportion of the students in the answers to the open question in the posttest, and despite the frequency dropped in the retention test, it still reached 16 % of the answers. An example of the instrumental view and the process molecular gene concept can be found in the following students' answer:

> The gene concept is relative and depends on the way the gene will be studied. It can be understood as a physical structure that originates RNAs and proteins or as the fruit of a process or the very process, for instance (Student 1, Class B, post-test).

There were also limits, however, regarding the planned learning goals. The most important concerns the fact that, even though the commitment to the classical molecular concept significantly decreased among the students in the posttest, this was just a transitory effect. Almost the same frequency of students' answers to the open question related to this concept was found in the pretest and retention test. If we consider alternative (b) in the closed forced-choice question, we see a similar pattern, with a slight increase in the posttest that is maintained in the retention test. The following answer is a straightforward example of a student's rendering of the classical molecular gene concept:

> Gene is a nucleotide sequence that determines the synthesis of a protein (Student 5, Class B, post-test).

The return of the classical molecular concept in the retention test is not surprising. It just reveals that 5 h of lessons are not enough to challenge a view so deep rooted in the students' views, as a consequence of its reinforcement during years of schooling (as indicated by our results for high school biology textbooks). This is one example of students' prior conceptions that are resistant to change even when specifically targeted in teaching interventions. Interestingly enough, this is a prior conception that is itself a product of previous schooling. In order to reach a successful change in students' commitment to the classical molecular concept, it would be necessary to defy it repeatedly in the intervention, in several different contexts, going far beyond what was possible in the short time range of the intervention.

There was considerable overlapping of ideas related to different gene concepts in the students' answers in all the moments in which the data gathering tool was applied. In class A, 36.4 % of the answers in the pre- and posttest showed category overlapping, with this frequency increasing to 40 % in the retention test. In class B, there were 38.5 % of answers with category overlapping in the pre- and posttest, with an increase to 53.8 % in the retention test. Thus, neither the usual course nor the teaching sequence seemed to be successful in demarcating different gene concepts. This interpretation is reinforced by the analysis of the data for the free-choice closed question, shown in Fig. 16.11. Just as we saw in the study on students' views about genes, when they were free to choose several views about genes, they marked a lot of alternatives. As remarked above, conceptual variation regarding genes as represented in textbooks seems to be translated into students' allegiance to several different accounts about genes and their functions. Even though these results cannot by itself lead to the conclusion that students are wrapped up by semantic confusions and ambiguities by appealing to such a variety of views about genes, if we combine them with our findings in the textbook studies, we can have reasons to worry about this potential hybridization of different ideas regarding genes and their functions.

If we now turn to the external validation of the teaching sequence, some interesting patterns can be discerned, although we need to see them with a grain of salt, given the constraint that the experimental design included only two classes. The classical molecular concept increased in frequency in the students' answers after the intervention, not only in the posttest but also in the retention test. This finding is in agreement with our previous finding that Genetics and Cell and Molecular biology courses in the same university lead to an increase in this much challenged view about genes, despite the fact that the anomalies faced by it are addressed in those very courses. Moreover, the usual approach followed in the course did not produce even the transitory decrease in students' commitment to this concept found in the teaching sequence explicitly addressing gene function models and gene concepts.

As in the case of the intervention, the informational conception dropped in frequency in the answers to the open question when the usual approach was employed in the course, corroborating the findings of the prior investigation of students' views in the same university. But in this case the closed forced-choice question showed an opposite tendency.

Finally, a significantly smaller diversity of views about genes was observed in class A when compared with class B, in the answers to both the open and the closed forced-choice question. This is not surprising since those views were explicitly discussed in the latter but not in the former class.

Some design principles underlying the construction of the teaching sequence were not tested in this study, such as the proposed pattern of classroom discursive interactions, which require for its testing a treatment of the video-recorded material that we did not perform yet. If we consider the didactic material elaborated to the course, the historically and philosophically informed approach, the treatment of models of gene function and gene concepts, and the discussion of the crisis of the classical molecular concept using molecular phenomena already addressed

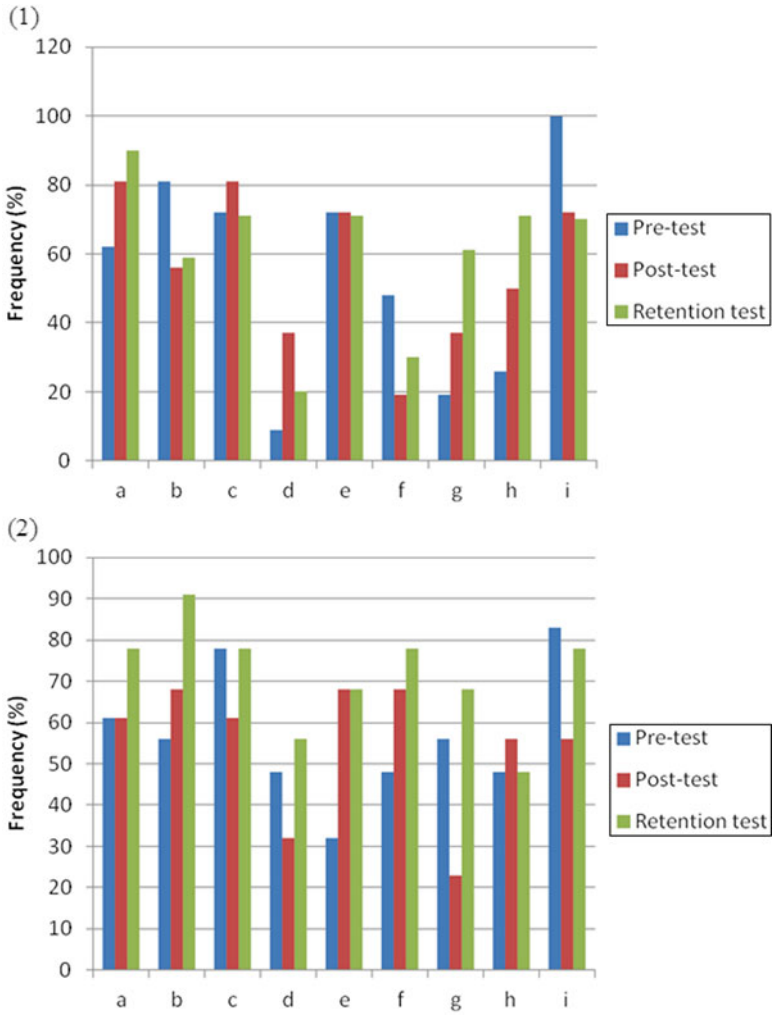

Fig. 16.11 Distribution of answers given by the students of the classes investigated to a free-choice closed question presenting several alternatives concerning the nature of genes: (**a**) Mendelian; (**b**), (**g**), and (**h**) classical molecular; (**c**) and (**i**) informational; (**d**) gene-P; (**e**) gene-D; (**f**) process molecular gene. (1) Class A (usual approach to the course, with no explicit discussion on gene function models and gene concepts); (2) class B (where the teaching sequence was implemented)

in the course, the results showed both contributions and limitations. The failures of the intervention are particularly interesting at this step of our research, since they indicated the need to introduce changes in the teaching sequence: for instance, a stronger challenge to the classical molecular concept and a more efficient discussion of the nature of models in connection with the historical construction of our understanding about genes, in order to decrease the hybridization of ideas

related to different models and concepts by the students. Nevertheless, the detected advances show that it is promising to continue the investigation with a revised prototype of the teaching sequence.

16.5 Conclusion

We have been engaged in the last 7 years in a research program on the treatment of conceptual variation regarding genes and their function in school science. Following the approach to research on science education used in our lab, we took as a starting point a number of descriptive studies aiming at diagnosing views about genes found in textbooks and students and moved to intervention studies, investigating a teaching strategy for improving higher education students' understanding of scientific models and conceptual variation around genes and their function. This teaching strategy is aligned with a contextual approach to science education, using a historically and philosophically informed approach to teach not only with but also about gene function models.

Our investigations on textbooks showed the prevalence of the molecular-informational model and a significant degree of hybridization between features from different models, even when they are incommensurable. This was found in both higher education Cell and Molecular biology textbooks and high school biology textbooks. Moreover, even when the empirical findings challenging the molecular-informational model of gene function are discussed by the textbooks, conceptual lessons are not often derived from them. In high school biology textbooks, another worrisome finding was that gene-P was often used and, more than that, was often conflated with the molecular-informational model. To treat genes as determining phenotypic traits is a conceptual tool for abstracting away the complexity of the genotype-phenotype relationship in tasks like pedigree analysis, often found in high school textbooks. However, genes are most often regarded by these textbooks as codifiers of the primary structure of polypeptides or RNAs (in accordance with the classical molecular concept) and determinants of phenotypes (in line with gene-P), showing how these textbooks consistently hybridize these two gene concepts. The conflation with a molecular account of the gene transposes the deterministic assumption to DNA sequences that only determines the phenotype at its lowest level, namely, the primary structure of proteins (sometimes, also their three-dimensional structure) and the structure of RNAs. It is lost from sight, thus the complexity of development, which mediates between genotype and phenotype and involves epigenetic and environmental factors as resources in causal parity with genes (Arthur 2011; Griffiths and Knight 1998).

This provides an example of a conflation of gene concepts leading to serious consequences in genetics teaching. As gene-P, an instrumental concept depicting genes as determining phenotypes, is conflated with a realist understanding of genes as molecular units in the genome, genetic deterministic views are very likely to develop: the molecular units become determiners of phenotypes and not

entities contributing to development in complex causal pathways involving other developmental resources. Preformationism lingers, then, in this manner of speaking about genes, as if traits themselves were somehow coded in the genome, and not constructed by complex developmental processes. As statements about genes-P are framed in an "updated" language, which connects it with molecular views about genes, and a historical and philosophical treatment of models is largely absent, students and teachers have no chance of understanding the instrumental nature of that concept and the explanatory context in which its usefulness is observed. The conflation between features of different gene function models not only leads to consequential problems in students' understanding of genes and their role in living beings – such as the commitment to a hyperbolic, overextended view of what DNA and genes do in cell systems – but also has implications to popular discourses about genes (or, in Keller's [2000] words, "gene talk") found in the media and even in textbooks themselves.[21]

As learning about genes becomes deeply contaminated by genetic deterministic views, students are less likely to develop a critical appraisal of socioscientific issues (Sadler 2011) related to genetics or to become capable of socially responsible decision making (Santos and Mortimer 2001) in situations involving knowledge about genes and their functions in living systems. After all, as Nelkin and Lindee (1995, p. 197) discuss,

> the findings of scientific genetics – about human behavior, disease, personality and intelligence – have become a popular resource precisely because they conform to and complement existing cultural beliefs about identity, family, gender and race […] the desire for prediction, the need for social boundaries, and the hope for control of the human future […] Whether or not such claims are sustained in fact may be irrelevant; their public appeal and popular appropriation reflect their social, not their scientific power.

Genetics is connected with socioscientific issues of central importance, such as cloning, stem cell research, genetically modified organisms, genetic engineering, use of genetic tests in society, human genetic improvement (eugenics), and reprogenetics. Sadler and Zeidler (2005) found that students' reasoning patterns in genetic engineering socioscientific issues are influenced by their knowledge of genetics, showing the importance that they properly learn about genes for their future life, not only as students but also as citizens that need to be informed by a consistent scientific understanding of the subject in order to actively and fully participate in democratic decision making.

The way these high school and higher education textbooks deal with conceptual variation can be regarded, thus, as a key problem in genetics teaching. For instance, all the historical models identified by Gericke and Hagberg (2007) were found in the high school textbooks and hybridization of features from different models was very frequent, showing how much conceptual variation was embedded in the treatment of genes, despite the prevalence of the molecular-informational model.

[21] See, for example, Condit et al. (1998, 2001), Carver et al. (2008), Keller (2000), and Nelkin and Lindee (1995).

As observed in Swedish high school textbooks and also in textbooks from four English-speaking countries, such conceptual variation is present in the explanations about genes with no clear demarcation between multiple historical models and their domains of application (Gericke et al. in press). Features related to different models are integrated in a single, linear narrative about genes, in such a manner that no conceptual variation seems to exist.

In a study of students' views about genes in two Brazilian universities (Federal University of Paraná and Federal University of Bahia), we compared biology students who had attended Genetics courses and those who did not and found that these courses increased their commitment to the classical molecular concept while decreasing their appeal to the informational conception. Again, no connection seemed to be properly made between the treatment of molecular phenomena that put into question the classical molecular gene and their conceptual implications. Students had difficulties in dealing with conceptual variation about genes, often hybridizing features from different models, even when they were incommensurable. Moreover, the degree of such hybridization was largely unaffected by Genetics courses, probably as an effect of the textbooks used, which included those analyzed here.

The convergence between our results concerning textbooks at two educational levels and higher education students' views is indicative of the reinforcement of the students' commitment to the molecular-informational model by the textbooks, as well as of the tendency to conflate features from different historical models. As we did not analyze pedagogical practice in the Genetics course of either of the universities, we cannot show data about how that practice was influenced by the textbooks used. However, our own acquaintance with these courses allows us to say that pedagogical work is significantly framed by the textbooks, making it likely the reinforcement hypothesis proposed above. Needless to say, it will be necessary to investigate classroom work in these courses to advance a more reliable conclusion to this effect.

A significant part of the problem with the treatment of conceptual variation about genes in higher education and high school textbooks results from the lack of a historically and philosophically approach to science education. In the absence of a clear discussion of models and either their role in science or their relation with reality, teachers and students are encouraged to address genes in a naïve realist manner and, also, to conflate features of different concepts as models as if they could be simply added as descriptive hallmarks of a reality being simply presented (rather than represented) in scientific theories and models. When using these textbooks, teachers and students do not have much chance of understanding the distinct origins, domains of application, and meanings of gene concepts and gene function models. Meanings ascribed to gene are simply accumulated as genes are discussed from different perspectives chapter after chapter, with the textbooks offering on the whole a thorough mixture of ideas originating from different models, often incommensurable with one another. The gene function models offer a particularly striking example of how the use of multiple models in science teaching can generate learning problems if not taught explicitly (Chinn and Samarapungavan 2008).

It seems necessary, thus, to change the treatment of genes in both textbooks and courses towards a more contextual approach, in which students must learn not only with gene function models but also about such models. If we do so, we can also address important NOS contents in connection with the history of the gene concept. After all, the transition from the understanding of genes in classical genetics to the molecular gene with the advent of molecular biology, as well as the crisis of the gene concept and the various approaches proposed to overcome it, compose a very interesting case of conceptual change and, also, provide a window into how theoretical entities are investigated and represented in science. This does not mean that one has to deal with complex historical, philosophical, or sociological issues when writing about genes in textbooks or teaching about genes in the classroom. We take the more modest position of proposing that one needs to write and teach about gene function models in a more explicit manner, paying attention to some basic aspects, such as the nature of models and their complex relation with reality, or the variation between gene function models and gene concepts in different subfields of biology.

To argue against the indiscriminate conflation of features related to different historical models of gene function does not imply that one should defend some single and all-encompassing gene concept or model of gene function. No such single model or concept could ever capture the diversity of meanings and epistemic roles associated with genes since the beginnings of the twentieth century. The idea is rather of a coexistence of a diversity of gene concepts and gene function models in school science, but with well-delimited domains of application (Burian 2004; El-Hani 2007). It is very important to provide students with a structured, organized view about the variety of meanings ascribed to genes and their functions, in order to avoid semantic confusions and indiscriminate mixtures of meanings related to different scientific contexts. To deal with conceptual variation, it is not enough to just say that "it may not be important to know what the precise meaning of 'gene' is" (Knight 2007, p. 300). To entertain the importance of a clear treatment of different gene concepts and gene function models, we need just to rephrase this statement by considering a plurality of ways of understanding genes: even though it is not really important to provide a single precise meaning of "gene," we need, still, to provide a clear and precise understanding of the several different meanings of "gene," since they cannot be all put to each and every use. Conceptual variation is not in itself the problem, but the absence of a proper historical and philosophical treatment of models about genes and their functions, which favors the extensive hybridization of ideas related to different models.

The lack of a historical and philosophical treatment of genes is also partly the explanation for the intriguing finding that neither textbooks nor students derive conceptual lessons from the challenges to the molecular-informational model that gave rise to the so-called crisis of the gene concept. Certainly, the textbooks could derive such lessons if they were more conceptually and theoretically oriented, even if they did not give much attention to history or philosophy of science. But this orientation is also typically lacking in these textbooks.

If a contextual approach to teaching about genes, with due attention to teaching with and about models, was in place, students and teachers would have a greater chance of building an understanding of genes and their roles in living systems that

could be richer and more aligned with what we currently know about the complex dynamics and architecture of the genome or the dependence of gene function on the cellular and supracellular context. This complexity is usually abstracted away in school science in favor of deterministic views, emphasizing one-to-one relationships between genes, functional proteins, and phenotypes, despite the overwhelming evidence that these relationships do not hold in most of the cases.[22] Textbook discourse should come closer to the knowledge structure of the academic disciplines of genetics and molecular biology in this case (Gericke et al. in press). It is not that high school textbooks should be necessarily updated with the last words in scientific knowledge. Since at high school students have to learn the basics of scientific disciplines, it may be more important to teach about developments of the past, which established the grounds of a way of thinking in a scientific domain, than to pursue an updated curriculum for its own sake. We need to introduce recent developments of science in school when they make an important difference for the way the students think about a domain of phenomena. This is, in our view, precisely the case with the developments of genetics and molecular biology in the last two decades. More attention should be given in genetics teaching to the current situation of the classical molecular concept, instead of just presenting it as if it was as accepted and coherent as it was in the past. At least, the fact that there are serious debates about what is a gene in the scientific community deserves attention in genetics teaching, even at the high school level. Our data do not show, however, the gene concept being treated as a controversial subject matter in either high school or higher education.

We need to investigate ways of introducing into school science the current understanding of the anomalies challenging the classical molecular concept and at least some of the alternatives to this way of understanding genes (Meyer et al. 2013). In the case of high school biological education, we think it is possible to create conditions for the students to understand that, even though the classical molecular concept has been quite important in the history of biology, it has ended up showing consequential limitations. Moreover, the concepts of gene-P and gene-D, the necessity of demarcating between them, and a critique of genetic determinism would be important additions to the high school genetics curriculum. If school science took into consideration the complex mapping between genotype, development, and phenotype (Arthur 2011), this might make a difference to students' thinking, creating conditions for the development of more informed and critical attitudes towards the deterministic talk about genes that pervades several spheres of society.

It was evident to us, then, that we needed to build and investigate an educational intervention based on a number of educated guesses about how to deal with conceptual variation about genes, which could be used as design principles for teaching interventions and, then, empirically tested in the classroom. One of the key design principles is to give a central role to a historical and philosophical approach to gene. We built such a teaching sequence in collaboration with a higher education Cell and

[22] See, for example, El-Hani (2007), El-Hani et al. (2009), Fogle (1990), Keller (2000), Moss (2003), and Scherrer and Jost (2007a, b).

Molecular biology teacher at a Brazilian University (Federal University of Bahia) and investigated it in accordance with design-based research. The teaching sequence was oriented towards a contextual approach, explicitly addressing the historical and philosophical dimensions of science, with a particular focus on the historical construction and nature of models of gene function and gene concepts. The internal validation of the teaching sequence showed some positive learning outcomes, but also some limits in attaining the planned learning outcomes. In particular, we managed to obtain just a transitory decrease of the classical molecular concept, an outcome that was not really surprising given the fact that – as our results in the diagnostic studies showed – this view has been reinforced throughout the lives of the students at school. Moreover, we did not reach success regarding the demarcation between gene concepts and gene function models, with the same high levels of hybridization observed in the diagnostic studies being also found in the intervention studies. Even though the external validation of the teaching sequence was constrained by the number of classes available for the study, the comparison between the usual way of teaching about genes in the course and the new intervention gave some hints of positive changes: the usual approach did not lead even to a transitory decrease of the classical molecular concept, and the students' views on genes have been enriched by the teaching sequence. The first result seems robust, since it is in strict accordance with the findings of our study on students' views about genes in the same university. The second finding amounts to the major difference brought about by the teaching sequence. Nevertheless, this outcome should be accompanied by a proper understanding of models and their demarcation, in order to lead to genuine gains for the students. But this was not observed in this first prototype of the teaching sequence.

These findings gave us clear clues about changes in the intervention for its second prototyping: the classical molecular concept needs to be challenged in a stronger way, and the discussion about models, their historical construction, and the necessity of their demarcation should be reformulated in order to reach a higher level of efficacy. Needless to say, the greatest challenge will be to accommodate these changes in the limited time available for the intervention, as a consequence of the overstuffed curricula of Genetics and Molecular biology courses at the university level.

Acknowledgments We are thankful to the Brazilian National Council for Scientific and Technological Development (CNPq) and the Research Support Foundation of the State of Bahia (FAPESB) for support during the development of the research reported in this paper.

Appendix 1: List of Analyzed Higher Education Textbooks

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. & Walter, P. (2002). *Molecular biology of the cell* (4th Ed). New York, NY: Garland.

Karp, G. (2004). *Cell and molecular biology: Concepts and experiments* (4th Ed). New York, NY: John Wiley and Sons.

Lodish, H., Kaiser, C. A., Berk, A., Krieger, M., Matsudaira, P. & Scott, M. P. (2003). *Molecular cell biology* (5th Ed). New York, NY: W. H Freeman.

Appendix 2: List of Analyzed High School Textbooks

T1 – Amabis, J. M. & Martho, G. R. (2005). *Biologia*. São Paulo: Moderna.
T2 – Borba, A. A. & Cançado, O. F. L. (2005). *Biologia*. Curitiba: Positivo.
T3 – Borba, A. A., Crozetta, M. A. S. & Lago, S. R. (2005). *Biologia*. São Paulo: IBEP.
T4 – Boschilia, C. (2005). *Biologia sem segredos*. São Paulo: RIDEEL.
T5 – Carvalho, W. (2005). *Biologia em foco*. São Paulo: FTD.
T6 – Cheida, L. E. (2005). *Biologia integrada*. São Paulo: FTD.
T7 – Coimbra, M. A. C., Rubio, P. C., Corazzini, R., Rodrigues, R. N. C. & Waldhelm, M. C. V. (2005). *Biologia – Projeto escola e cidadania para todos*. São Paulo: Editora do Brasil.
T8 – Faucz, F. R. & Quintilham, C. T. (2005). *Biologia: Caminho da vida*. Curitiba: Base.
T9 – Favaretto, J. A. & Mercadante, C. (2005). *Biologia*. São Paulo: Moderna.
T10 – Frota-Pessoa, O. (2005). *Biologia*. São Paulo: Scipione.
T11 – Gainotti, A. & Modelli, A. (2005). *Biologia*. São Paulo: Scipione.
T12 – Laurence, J. (2005). *Biologia*. São Paulo: Nova Geração.
T13 – Linhares, S. & Gewandsznajder, F. (2005). *Biologia*. São Paulo: Ática.
T14 – Lopes, S. & Rosso, S. (2005). *Biologia*. São Paulo: Saraiva.
T15 – Machado, S. W. S. (2005). *Biologia*. São Paulo: Scipione.
T16 – Morandini, C. & Bellinello, L. C. (2005). *Biologia*. São Paulo: Atual.
T17 – Paulino, W. R. (2005). *Biologia*. São Paulo: Ática.
T18 – Silva-Júnior, C. & Sasson, S. (2005). *Biologia*. São Paulo: Saraiva.

References

Abd-El-Khalick, F. & Lederman, N. G. (2000). Improving science teachers' conceptions of nature of science: A critical review of the literature. *International Journal of Science Education*, 22, 665–701.

Adami, C. (2004). Information theory in molecular biology. *Physics of Life Reviews*, 1, 3–22.

Arthur, W. (2011). *Evolution: A developmental approach*. Chichester: Wiley-Blackwell.

Artigue, M. (1988). Ingéniérie didactique. *Recherches en didactique des mathemátiques*, 9, 281–308.

Barab, S. & Squire, K. (2004). Design-based research: putting a stake in the ground. *Journal of the Learning Sciences*, 13, 1–14.

Bardin, L. (2000). *Análise de conteúdo* (Content analysis). Lisboa: Edições 70.

Baumgartner, E., Bell, P., Bophy, S. et al. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, 32, 5–8.

Benzer, S. (1957). The elementary units of heredity. In W. McElroy and B. Glass (Eds.), *The chemical basis of heredity* (pp. 70–93). Baltimore, MD: John Hopkins Press.

Black, D. L. (2003). Mechanisms of alternative pre-messenger RNA splicing. *Annual Review of Biochemistry*, 72, 291–336.

Black, M. (1962). *Models and metaphors: Studies in language and philosophy*. Ithaca, NY: Cornell University Press.

Brown, A. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences*, 2, 141–178.

Burian, R. M. (1985). On conceptual change in biology: The case of the gene. In D. J. Depew & B. H. Weber (Eds.), *Evolution at a crossroads: The new biology and the new philosophy of science* (pp. 21–24). Cambridge, MA: The MIT Press.

Burian, R. M. (2004). Molecular epigenesis, molecular pleiotropy, and molecular gene definitions. *History and Philosophy of Life Sciences*, 26, 59–80.
Carlson, A. E. (1966). *The gene. A critical history*. Philadelphia, PA: W. B. Saunders.
Carver, R., Waldahl, R. & Breivik, J. (2008). Frame that gene – A tool for analyzing and classifying the communication of genetics to the public. *EMBO reports*, 9, 943–947.
Celotto, A. & Graveley, B. (2001). Alternative splicing of the *Drosophila* DSCAM pre-mRNA is both temporally and spatially regulated. *Genetics*, 159, 599–608.
Chevallard, Y. (1989). On didactic transposition theory: Some introductory notes. Paper presented at the International symposium on selected domains of research and development in mathematics education, *Proceedings* (pp. 51–62). Bratislava, Slovakia. Retrieved October 29, 2011 from: http://yves.chevallard.free.fr/spip/spip/article.php3?id_article=122
Chinn, A. C. & Samarapungavan, A. (2008). Learning to use scientific models: Multiple dimensions of conceptual change. In R.A. Duschl & R.E. Grandy (Eds.), *Teaching scientific inquiry* (pp. 191–225). Rotterdam: Sense Publishers.
Condit, C.M., Ofulue, N. & Sheedy, K.M. (1998). Determinism and mass-media portrayals of genetics. *American Journal of Human Genetics*, 62, 979–984.
Condit, C. M., Ferguson, A., Kassel, R., Tadhani, C., Gooding, H. C. & Parrot, R. (2001). An explanatory study of the impact of news headlines on genetic determinism. *Science Communication*, 22, 379–395.
Cooper, M. D. & Alder, M. N. (2006). The evolution of adaptive immune systems. *Cell*, 124, 815–822.
Daston, L. & Galison, P. (2010). *Objectivity*. Brooklyn, NY: Zone Books.
Develaki, M. (2007). The model-based view of scientific theories and the structuring of school science programs. *Science & Education*, 16 (7–8), 725–749.
El-Hani, C. N. (2007). Between the cross and the sword: The crisis of the gene concept. *Genetics and Molecular Biology*, 30, 297–307.
El-Hani, C. N., Queiroz, J. & Emmeche, C. (2006). A semiotic analysis of the genetic information system. *Semiotica*, 160, 1–68.
El-Hani, C. N., Queiroz, J. & Emmeche, C. (2009). *Genes, Information, and Semiosis*. Tartu: Tartu University Press, Tartu Semiotics Library.
El-Hani, C. N., Roque, N. & Rocha, P. B. (2007). Brazilian high school biology textbooks: Results from a national program. In: *Proceedings of the IOSTE International Meeting on Critical Analysis of School Science Textbook* (pp. 505–516). Hammamet, Tunisia: University of Tunis.
El-Hani, C. N., Roque, N. & Rocha, P. L. B. (2011). Livros didáticos de Biologia do ensino médio: Resultados do PNLEM/2007. *Educação em Revista*, 27, 211–240.
Falk, R. (1986). What is a gene? *Studies in the History and Philosophy of Science*, 17, 133–173.
Fleck, L. (1979/1935). *Genesis and development of a scientific fact*. Chicago, IL: The University of Chicago Press.
Fogle, T. (1990). Are genes units of inheritance? *Biology and Philosophy*, 5, 349–371.
Fogle, T. (2000). The dissolution of protein coding genes. In P. Beurton, R. Falk & H-J. Rheinberger (Eds.), *The concept of the gene in development and evolution* (pp. 3–25). Cambridge: Cambridge University Press.
Gelbart, W. (1998). Databases in genomic research. *Science*, 282, 659–661.
Gericke, N. M. & Hagberg, M. (2007). Definition of historical models of gene function and their relation to students' understandings of genetics. *Science & Education*, 16, 849–881.
Gericke, N. M. & Hagberg, M. (2010a). Conceptual incoherence as a result of the use of multiple historical models in school textbooks. *Research in Science Education*, 40, 605–623.
Gericke, N. M. & Hagberg, M. (2010b). Conceptual variation in the depiction of gene function in upper secondary school textbooks. *Science & Education*, 19, 963–994.
Gericke, N. M., Hagberg, M., Santos, V. C., Joaquim, L. M. & El-Hani, C. N. (in press). Conceptual variation or Incoherence? Textbook discourse on genes in six countries. *Science & Education*.

Gerstein, M. B., Bruce, C., Rozowsky, J. S., Zheng, D., Du, J., Korbel, J. O., Emanuelsson, O., Zhang, Z. D., Weissman, S., & Snyder, M. (2007). What is a gene, post-ENCODE? History and updated definition. *Genome Research*, 17, 669–681.

Grandy, R. E. (2003). What are models and why do we need them? *Science & Education*, 12(8), 773–777.

Graveley, B. R. (2001). Alternative splicing: Increasing diversity in the proteomic world. *Trends in Genetics*, 17, 100–107.

Griffiths, P. E. (2001). Genetic information: A metaphor in search of a theory. *Philosophy of Science*, 68, 394–403.

Griffiths, P.E. & Knight, R.D. (1998). What is the developmental challenge? *Philosophy of Science*, 65, 2, 253–258.

Griffiths, P. E. & Neumann-Held, E. (1999). The many faces of the gene. *BioScience*, 49, 656–662.

Hall, B. K. (2001). The gene is not dead, merely orphaned and seeking a home. *Evolution and Development*, 3, 225–228.

Halloun, I. A. (2004). *Modeling theory in science education*. Dordrecht: Kluwer Academic Publishers.

Halloun, I. A. (2007). Mediated modeling in science education. *Science & Education*, 16, 653–697.

Hanson, M. R. (1996). Protein products of incompletely edited transcripts are detected in plant mitochondria. *The Plant Cell*, 8(1), 1–3.

Hesse, M. B. (1963). *Models and analogies in science*. London: Seed and Ward.

Hoffmeyer, J. & Emmeche, C. (1991). Code-duality and the semiotics of nature. In M. Anderson & F. Merrell (Eds.), *On semiotic modeling* (pp. 117–166). Berlin: Mouton de Gruyter.

Holmes, F. L. (2006). *Reconceiving the gene: Seymour Benzer's adventures in phage genetics*. New Haven, CT: Yale University Press.

Hull, D. L. (1974). *Philosophy of biological science*. Englewood Cliffs, NJ: Prentice-Hall.

Jablonka, E. (2002). Information: Its interpretation, its inheritance, and its sharing. *Philosophy of Science*, 69, 578–605.

Johannsen, W. (1909). *Elemente der exakten erblichkeitslehre*. Jena: Gustav Fischer. Retrieved August 23, 2012 from: http://caliban.mpiz-koeln.mpg.de/johannsen/elemente/johannsen_elemente_der_exakten_erblichkeitslehre_2.pdf.

Justi, R. S. & Gilbert, J. K., (1999). A cause of ahistorical science teaching: Use of hybrid models. *Science Education*, 83, 163–177.

Kampa, D., Cheng, J., Kapranov, P., Yamanaka, M., Brubaker, S., Cawley, S., Drenkow, J., Piccolboni, A., Bekiranov, S., Helt, G., Tammana, H. & Gingeras, T. R. (2004). Novel RNAs identified from an in-depth analysis of the transcriptome of human chromosomes 21 and 22. *Genome Research*, 14, 331–342.

Kay, L. E. (2000). *Who wrote the book of life? A history of the genetic code*. Stanford, CA: Stanford University Press.

Keller, E. F. (2000). *The century of the gene*. Cambridge, MA: Harvard University Press.

Keller, E. F. (2005). The century beyond the gene. *Journal of Biosciences*, 30, 3–10.

Keller, E. F. & Harel, D. (2007). Beyond the gene. *PLoS One*, 2, e1231.

Kitcher, P. (1982). Genes. *British Journal for the Philosophy of Science*, 33, 337–359.

Knight, R. (2007). Reports of the death of the gene are greatly exaggerated. *Biology and Philosophy*, 22, 293–306.

Larsson, S. (2009). A pluralist view of generalization in qualitative research. *International Journal of Research & Method in Education*, 32, 25–38.

Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York, NY: Cambridge University Press.

LeCompte, M. & Goetz, J. (1982). Problems of reliability and validity in ethnographic research. *Review of Educational Research*, 52(1), 31–60.

Lev-Maor, G., Sorek, R., Levanon, E. Y., Paz, N., Eisenberg, E. & Ast, G. (2007). RNA-editing-mediated exon evolution. *Genome Biology*, 8, R29.

Matthews, M. R. (1998). In defense of modest goals when teaching about the nature of science. *Journal of Research in Science Teaching*, 35, 161–174.

Mayr, E. (1982). *The growth of biological thought: Diversity, evolution and inheritance*. Cambridge, MA: Harvard University Press.

Méheut, M. (2005). Teaching-learning sequences tools for learning and/or research. In K. Boersma, M. Goedhart, O. De Jong & H. Eijkelhof (Eds.), *Research and the quality of science education* (pp. 195–207. Dordrecht: Springer.

Meyer, L. M. N., Bomfim, G. C. & El-Hani, C. N. (2013). How to understand the gene in the 21st century. *Science & Education*, 22, 345–374.

Mortimer, E. F. & Scott, P. H. (2003). *Meaning making in secondary science classrooms*. Maidenhead: Open University Press.

Moss, L. (2001). Deconstructing the gene and reconstructing molecular developmental systems. In S. Oyama, P. E. Griffiths & R. D. Gray (Eds.), *Cycles of contingency: Developmental systems and evolution* (pp. 85–97). Cambridge, MA: MIT Press.

Moss, L. (2003). *What genes can't do*. Cambridge, MA: The MIT Press.

Murre, C. (2007). Epigenetics of Antigen-receptor gene assembly. *Current Opinion in Genetics & Development*, 17, 415–421.

Nelkin, D. & Lindee, S. M. (1995). *The DNA mystique: The gene as a cultural icon*. New York, NY: Freeman.

Neumann-Held, E. (1999). The Gene is dead – Long live the gene: Conceptualizing genes the constructionist way. In P. Koslowski (Ed.). *Sociobiology and bioeconomics: The theory of evolution in biological and economic thinking* (pp. 105–137). Berlin: Springer.

Neumann-Held, E. (2001). Let's talk about genes: The process molecular gene concept and its context. In S. Oyama, P. E. Griffiths & R. D. Gray (Eds.), *Cycles of contingency: Developmental systems and evolution* (pp. 69–84). Cambridge, MA: MIT Press.

Nieveen, N., McKenney, S. & Van den Akker, J. (2006). Educational design research: The value of variety. In J. Van den Akker; K. Gravemeijer; S. McKenney & N. Nieveen (Eds). *Educational design research* (pp. 151–158). London: Routledge.

Oyama, S. (2000/1985). *The ontogeny of information: Developmental systems and evolution* (2nd ed). Cambridge: Cambridge University Press.

Pardini, M. I. M. C. & Guimarães, R. C. (1992). A systemic concept of the gene. *Genetics and Molecular Biology*, 15, 713–721.

Pearson, H. (2006). What is a gene? *Nature*, 441, 399–401.

Pitombo, M. A., Almeida, A. M. R., & El-Hani, C. N. (2008). Gene concepts in higher education cell and molecular biology textbooks. *Science Education International*, 19(2), 219–234.

Plomp, T. (2009). Educational design research: An introduction. In: T. Plomp & N. Nieveen (Eds.). *An introduction to educational design research* (pp. 9–35). Enschede: SLO – Netherlands Institute for Curriculum Development.

Portin, P. (1993). The concept of the gene: Short history and present status. *Quarterly Review of Biology*, 56, 173–223.

Reeves, T. C. (2006). Design research from a technology perspective. In J. Van den Akker, K. Gravemeijer, S. McKenney & N. Nieveen (Eds.). *Educational design research* (pp. 52–66). London, Routledge.

Rheinberger, H.-J. (2000). Gene concepts: Fragments from the perspective of molecular biology. In: P. Beurton, R. Falk & H.-J. Rheinberger (Eds.). *The concept of the gene in development and evolution* (pp. 219–239). Cambridge: Cambridge University Press.

Sadler, T. D. (Ed.). (2011). *Socioscientific issues in the classroom: Teaching, learning and research*. Dordrecht: Springer.

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

Santos, V. C., Joaquim, L. M. & El-Hani, C. N. (2012). Hybrid deterministic views about genes in biology textbooks: A key problem in genetics teaching. *Science & Education*, 21, 543–578.

Santos, W. L. P. & Mortimer, E. F. (2001). Tomada de decisão para ação social responsável no ensino de ciências (decision making for responsible social action in science teaching). *Ciência e Educação*, 7, 95–111.

Scherrer, K. & Jost, J. (2007a). The gene and the genon concept: A functional and information-theoretic analysis. *Molecular System Biology*, 3, 1–11.

Scherrer, K. & Jost, J. (2007b). The gene and the genon concept: Coding versus regulation. A conceptual and information-theoretic analysis of genetic storage and expression in the light of modern molecular biology. *Theory in Biosciences*, 126, 65–113.

Schwab, J. (1964). Structure of the disciplines: Meaning & significances. In G. W. Ford & L. Pugno (eds.), *The structure of knowledge & the curriculum* (pp. 6–30). Chicago, IL: Rand, McNally & Co.

Shannon, C. E. & Weaver, W. (1949). *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.

Simons, H.; Kushner, S.; Jones, K. & James, D. (2003). From evidence-based practice to practice-based evidence: the idea of situated generalization. *Research Papers in Education*, 18, 347–364.

Smith, M. U. & Adkison, L. R. (2010). Updating the model definition of the gene in the modern genomic era with implications for instruction. *Science & Education*, 19, 1–20.

Stotz, K., Griffiths, P. E. & Knight, R. (2004). How biologists conceptualize genes: An empirical study. *Studies in the History and Philosophy of Biological & Biomedical Sciences*, 35, 647–673.

Suppe, F. (1977). *The structure of scientific theories*. Urbana, IL: University of Illinois Press.

The ENCODE Project Consortium (2004). The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science*, 306, 636–640.

van den Akker, J., Gravemeijer, K., McKenney, S. & Nieveen, N. (2006). *Educational design research*. London, Routledge.

van Fraassen, B. (1980). *The scientific image*. Oxford: Clarendon Press.

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G. et. al. (2001). The sequence of the human genome. *Science*, 291, 1305–1351.

Wang, W., Zhang, J., Alvarez, C., Llopart, A. & Long, M. (2000). The origin of the jingwei gene and the complex modular structure of its parental gene, yellow emperor, in *Drosophila melanogaster*. *Molecular Biology and Evolution*, 17, 1294–1301.

Wanscher, J. H. (1975). The history of Wilhelm Johannsen's genetical terms and concepts from the period 1903 to 1926. *Centaurus*, 19(2), 125–147.

Watson, J. D. & Crick, F. H. C. (1953). A structure for deoxyribose nucleic acid. *Nature*, 171, 737–738.

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. New York, NY: Cambridge University Press.

Charbel N. El-Hani is Professor of History, Philosophy, and Biology Teaching at the Institute of Biology, Federal University of Bahia (UFBA), Brazil, and receives a Productivity Research Grant of the National Council for Scientific and Technological Development (CNPq). He was trained as a biologist and subsequently did his Ph.D. in Education at University of São Paulo. He is affiliated with the Graduate Studies Programs in History, Philosophy, Science Teaching (Federal University of Bahia and State University of de Feira de Santana), Ecology and Biomonitoring (Federal University of Bahia), and Genetics (Federal University of Bahia). He coordinates the History, Philosophy, and Biology Teaching Lab at UFBA. His research interests

are in science education research, philosophy of biology, biosemiotics, ecology, and animal behavior. He is a member of editorial boards of Brazilian and international journals in science education and philosophy of biology. Among his recent publications one finds: Santos, V. C.; Joaquim, L. M. & El-Hani, C. N. (2012). Hybrid deterministic views about genes in biology textbooks: A key problem in genetics teaching. *Science & Education* 21: 543–578. DOI: 10.1007/s11191-011-9348-1; Arteaga, J. M. S. & El-Hani, C. N. (2012). Othering Processes and STS Curricula: From 19th Century Scientific Discourse on Interracial Competition and Racial Extinction to Othering in Biomedical Technosciences. *Science & Education* 21: 607–629. DOI: 10.1007/s11191-011-9384-x; Mortimer, E. F., Scott, P. & El-Hani, C. N. (2012). The heterogeneity of discourse in science classrooms: the conceptual profile approach. In: Fraser, B., Tobin, K. & McRobbie, C. (Eds.). *Second International Handbook of Science Education* (vol. 1, pp. 231–246). Dordrecht: Springer; El-Hani, C. N. & Sepulveda, C. (2010). The relationship between science and religion in the education of protestant biology preservice teachers in a Brazilian university. Cultural Studies of *Science Education* 5(1): 103–125. DOI: 10.1007/s11422-009-9212-7.

Ana Maria R. de Almeida is a Ph.D. candidate at the University of California at Berkeley in the Department of Plant and Microbial Biology where she studies the evolution of morphological diversity of tropical monocot flowers, especially regarding developmental processes and their relationship to evolution (Evo-Devo). She is also a member of the Freeling Lab, at the same Department, where she works on genome evolution. Ana obtained her Master's Degree in History, Philosophy, and Science Teaching at the Federal University of Bahia (UFBA), Brazil. During her Masters she studied Philosophy of Ecology, more specifically the Biodiversity-Ecosystem Functioning Program. Her research interests are in the relationship between evolution, ecology and developmental biology, genome evolution, and science teaching. Among her most recent publications, there are: D'Hont, A. et al. (2012). The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature*, 488: 213–219 DOI: 10.1038/nature11241; Specht, C.D.; Yockteng, R.; Almeida, A.M.; Kirchoff, B.; Kress, J. (2012). Homoplasy, Pollination, and Emerging Complexity During the Evolution of Floral Development in the Tropical Gingers (Zingiberales). *Bot. Rev*. (2012) 78:440–462, DOI: 10.1007/s12229-012-9111-6; Almeida, A.M.; Brown, A.; Specht, C.D. Tracking the Development of the Petaloid Fertile Stamen in *Canna indica*: Insights into the origin of androecial petaloidy in the Zingiberales (submitted to AoB PLANTS on November 2012).

Gilberto C. Bomfim is Professor of Cell and Molecular Biology and Genetics Teaching at the Institute of Biology, Federal University of Bahia, Brazil. He is a biologist and did his Ph.D. in Pathology at Federal University of Bahia. As a researcher, he participates in the History, Philosophy, and Biology Teaching Lab at UFBA. Among his recent publications, one finds: Meyer, L. M. N.; Bomfim, G. C. & El-Hani, C. N. (2013). How to understand the gene in the 21st century. *Science & Education*, 22, 345–374. DOI: 10.1007/s11191-011-9390-z.

Leyla M. Joaquim is a Ph.D. candidate in the Programme of History, Philosophy, and Science Teaching at the Federal University of Bahia (UFBa). She is part of the History, Philosophy, and Biological Teaching Lab at UFBA and the "Lacic group" (Laboratory Science as Culture) headed by Olival Freire Jr. After obtaining her Bachelor and Teaching Degree in Biological Sciences at Federal University of Paraná (UFPR), she started her graduate studies in the field of Genetics Teaching. She joined projects investigating how ideas about genes and gene function are treated in school knowledge, as represented in textbooks and students' views. Currently, she is dedicated to her Ph.D. dissertation in History of Science, which is a study about the interest of physicists in biological problems and uses oral history as one of the methodological tools. She had research stays at Max Planck Institute for the History of Science (MPIWG – Berlin, Germany) at Department I and II in 2010 and 2012, respectively. Accordingly, her research interests lie mainly in science education research, history of biology, and history of physics. Among her recent publications, one finds: Santos, V. C.; Joaquim, L. M. & El-Hani, C. N. (2012). Hybrid deterministic views about genes in biology textbooks: A key problem in genetics teaching. *Science & Education* 21: 543–578. DOI: 10.1007/s11191-011-9348-1.

João Carlos M. Magalhães is Professor of Genetics, Population Genetics, and Methodology of Science at the Federal University of Parana, Brazil. He graduated with Biological Sciences, Master in Genetics, and Ph.D. in Genetics in the Federal University of Paraná. He is affiliated to the Graduate Studies Program in Genetics at the same university. He has research experience in Genetics and Philosophy of Science, working mainly in the following areas: Genetics of Human Populations and Philosophy of Biology. His doctoral thesis discusses the logical foundations of Genetics and Evolution. He also has experience in application and development of software simulation of evolutionary processes to improve learning outcomes of Biology students. His recent publications include: Tureck, L. V., Santos, L. C., Wowk, P. F., Mattos, S. B., Silva, J. S., Magalhaes, J. C. M., Roxo, V. M. S., Bicalho, M. G. (In press). HLA-G Regulatory polymorphisms in Afro-Brazilians HLA-G. Tissue Antigens, 2012; Magalhaes, J. C. M. & Gondro, C. A contemporary view of Population Genetics in Evolution. In: Krause, D., Videira, A. A. P. (Org.). *Brazilian Studies in Philosophy and History of Science: an account of recent works*. 1ed. New York: Springer, 2011, v. 2990, p. 281–290; Alle, B. R.; Furtado-Alle, L., Gondro, C., Magalhaes, J. C. M. Kuri: A Simulator of Ecological Genetics for Tree Populations. *Journal of Artificial Evolution and Applications*, v. 2009, p. ID 783647, 2009; Gondro, C. & Magalhaes, J. C. M. A simple genetic algorithm for studies in mendelian populations. In: H. Abbass, T. Bossamaier; J. Wiles. (Org.). *Recent Advances in Artificial Life*. 1ed. London: World Scientific Publishing, 2005, p. 85–98.

Lia M. N. Meyer is Professor of Science Teaching Methodology at the Department of Biosciences – Campus Itabaiana, Federal University of Sergipe, Brazil. She is a member of the History, Philosophy, and Biology Teaching Lab at UFBA. Her research interests are in science education research, teacher education, and philosophy of

biology. Among his recent publications, one finds: Meyer, L. M. N.; Bomfim, G. C. & El-Hani, C. N. (2013). How to understand the gene in the 21st century. *Science & Education*, 22, 345–374. DOI: 10.1007/s11191-011-9390-z.

Maiana A. Pitombo is a graduate student at Federal University of Rio Grande do Norte, Brazil. She has a Bachelor's Degree in Biology from the Federal University of Bahia. Her research interests are in Genetics and Immunology. Her publications include: Pitombo, M. A.; Almeida, A. M. R. de; El-Hani, C. N. Gene concepts in higher education cell and molecular biology textbooks. *Science Education International*, v. 19, p. 219–234, 2008.

Vanessa C. dos Santos is a Visiting Professor of History, Philosophy, and Biology Teaching at the State University of Paraiba, Brazil. She was trained as a biologist (Federal University of Bahia) and has a Master's Degree in the Graduate Studies Program in History, Philosophy, and Science Teaching (Federal University of Bahia and State University of de Feira de Santana). Her research interests are in history and philosophy of biology, science education research, biosemiotics, and situated cognition. Among her recent publications, one finds: Santos, V. C.; Joaquim, L. M. & El-Hani, C. N. (2012). Hybrid deterministic views about genes in biology textbooks: A key problem in genetics teaching. *Science & Education* 21: 543–578. DOI: 10.1007/s11191-011-9348-1.