Res Sci Educ (2014) 44:81–108
DOI 10.1007/s11165-013-9375-9

Teaching Genetics in Secondary Classrooms: a Linguistic Analysis of Teachers' Talk About Proteins

Karin Thörne · Niklas Gericke

Published online: 31 July 2013

Abstract This study investigates Swedish biology teachers' inclusion of proteins when teaching genetics in grade nine (students 15–16 years old). For some years, there has been a call to give attention to proteins when teaching genetics as a means of linking the concepts 'gene' and 'trait'. Students are known to have problems with this relation because the concepts belong to different organizational levels. However, we know little about how the topic is taught and therefore this case study focuses on how teachers talk about proteins while teaching genetics and if they use proteins as a link between the micro and macro level. Four teachers were recorded during entire genetics teaching sequences, 45 lessons in total. The teachers' verbal communication was then analyzed using thematic pattern analysis, which is based in systemic functional linguistics. The linguistic analysis of teachers' talk in action revealed great variations in both the extent to which they used proteins in explanations of genetics and the ways they included proteins in linking genes and traits. Two of the teachers used protein as a link between gene and trait, while two did not. Three of the four teachers included instruction about protein synthesis. The common message from all teachers was that proteins are built, but none of the teachers talked about genes as exclusively encoding proteins. Our results suggest that students' common lack of understanding of proteins as an intermediate link between gene and trait could be explained by limitations in the way the subject is taught.

Keywords Classroom study · Genetics education · Proteins · Secondary school · Teachers language

Introduction

The gene–trait relation is central in genetics, but known to be educationally challenging (e.g., Duncan and Reiser 2007; Duncan and Tseng 2011; Gericke and Hagberg 2007; Gericke et al. 2013; Lewis and Kattmann 2004; Marbach-Ad and Stavy 2000; Venville and Treagust 1998). Several studies have shown that students generally lack understanding about the mechanistic steps between genes and traits (Lewis and Kattmann 2004; Marbach-Ad and Stavy 2000) and have alternative understandings about the relation between genes and

K. Thörne (✉) · N. Gericke
Department of Environmental and Life Sciences, Karlstad University,
Universitetsgatan 2, 651 88 Karlstad, Sweden
e-mail: karin.thorne@kau.se

traits, and more specifically about how genes confer physical traits (Lewis and Kattmann 2004; Venville and Treagust 1998). Furthermore, the mechanisms whereby genes confer traits are generally not clearly presented in textbooks (Flodin 2009; Gericke and Hagberg 2010a, b). To address these issues there have been calls to highlight the importance of proteins in the linkage between gene and trait (Duncan and Reiser 2007; Lewis and Kattmann 2004). However, researchers have paid little attention to how genetics is taught in school classrooms and the topics genetics teachers cover.

To understand why students experience difficulties with the gene–trait relation, it is important to explore teaching situations and the language used in situ, in addition to the material presented in textbooks. Thus, the aim of the study presented here is to contribute to knowledge about the roles that teachers ascribe to proteins while teaching genetics in secondary school. Using thematic pattern analysis, an analytical approach developed within the framework of systemic functional linguistics (SFL), we have investigated how four Swedish teachers talk about proteins in their teaching of genetics in grade 9 (15–16 years). We first describe problems associated with our focal science content, then describe our theoretical underpinnings, which are based on the work of Lemke (1990) and Halliday (e.g. Halliday 1993; Halliday and Martin 1993).

Genetics Education

Several areas of difficulties have been identified in genetics education. One of the most central of these problematic areas is the complex nature of genetics where concepts and processes belong to different organizational levels, also known as the micro–macro problem (Knippels, 2002). According to Marbach-Ad and Stavy (2000), genetic phenomena can be described at three levels: macro (e.g., visible traits), micro (cellular phenomena), and sub-micro (biochemical structures, e.g., genes). Hence, 'gene' and 'trait' are concepts associated with different levels of organization. Researchers have argued for the importance of fostering students' ability to move between different biological organizational levels (Schönborn and Bögeholtz 2009), which includes teaching how genetic information is translated into visible traits via protein synthesis (Duncan et al. 2009; Venville and Treagust 2002). As argued by Duncan and Tseng (2011), students have to understand that genes do not directly affect visible traits in order to understand the underlying mechanistic explanations at lower levels.

In science, *mechanisms* are central to understanding and describing phenomena (Machamer et al. 2000). A mechanism is a complex entity or, more generally, a process that produces a regular phenomenon, for example a trait of an organism. Within biology, mechanistic explanations have been important in understanding of molecular genetics such as, DNA replication, protein synthesis, and gene expression (Darden, 2008). Therefore, reasoning about mechanisms should be included in genetics education because descriptions of mechanisms are part of an answer to the questions about why and how a biological object or process have occurred. Accordingly, proteins can be used in mechanistic explanations to describe how and why genes are related to traits.

Bridging the Gap Between Organizational Levels

Venville and Donovan (2005) interviewed nine genetics experts about how the gene concept should be taught in order to educate a population appropriately for the biotechnology age. The central conclusion was that teachers should explicitly and clearly teach students that genes consist of DNA, which provides a chemical code for polypeptides, and in turn polypeptides build up proteins. Further conclusions were that the importance of proteins should be highlighted, and proteins' multiple functions in organisms' bodies should be taught. The experts also emphasized the complexity of gene function, in contrast to the common notion that one gene determines one character, although one thought it would be reasonable to teach this notion initially, although it is a simplification (Venville and Donovan 2005). Moreover, a modern, complex understanding of this relation, as described in contemporary molecular genetics, is probably too difficult for students in school genetics classes to comprehend (Gericke and Hagberg 2007).

Duncan and Reiser (2007) argue that students need to be informed that genes contain instructions for the structure of an amino acid sequence. Futher, teachers' instruction should emphasize the central role of proteins and mechanistic explanations about how proteins are responsible for phenomena at higher organizational levels. The important role of proteins is also stressed in the learning progression about modern genetics, from grade 5 to grade 10, presented by Duncan et al. (2009). They also offer teaching recommendations based on what they claim are the *Big Ideas* in genetics, arguing that it is better to focus on some core ideas rather than a lot of details. Two of these Big Ideas are: (1) genetic information specifies polypeptides and (2) the central role of proteins and that genetic effects are mediated through interactions between molecules. Thus, according to Duncan et al. (2009), it is important to teach students about the mechanistic and causal premise of a phenotypic trait because a lack of this understanding will increase the risk of students acquiring the deterministic view that one gene directly determines a whole trait. Duncan (2007) argues for the importance of focusing more on mechanistic explanations of gene–trait relations, rather than separate details of (for example) transcription and translation, since this focus helps students to understand the general connections between gene–protein–trait concepts (Duncan and Reiser 2007).

Allchin (2000) argues that teachers should explicitly pay attention to the pathway from gene via protein to physical traits in their teaching and link different aspects of genetics together with good examples, for example, alkaptonuria or albinism, or examples that illustrate underlying molecular causes of traits in typical Mendelian models, such as wrinkled peas and white-eyed fruit flies, as described by Guilfoile (1997). A key conclusion we draw from this review is that genetic education research literature highlights the importance of including information about proteins when teaching genetics, particularly as a means of connecting the gene and trait concepts when teaching gene function.

Student (Mis)understandings About Proteins

Most students are not aware of proteins' central role in effectuating the genetic code and often view genes as containing information about both structure and function at multiple biological organizational levels. That is, students do not see genes as exclusively coding for amino acid sequences,[1] but in addition believe that genes have multiple effects affecting different biological organizational levels (Duncan and Reiser 2007). This could mean that genes code not only for proteins but also-directly-for traits such as eye color and nose shape. Neither is the link between

[1] Of course, DNA does not exclusively code for amino acid sequences, but also for tRNA, rRNA, and regulatory sequences. However, for the level of education (introductory course in genetics at lower secondary school), we consider it reasonable to be excluded. So, for convenience, hereafter we regard DNA as encoding exclusively amino acids (in contrast to traits).

genes and proteins clear for the students. Even if they have developed an understanding of genes as instructions for proteins they still think that genes are instructions for other things at the same time. They do not realize the central role of proteins as entities mediating the effects of genes observed at higher organizational levels (Duncan and Reiser 2007). Gericke and Wahlberg (2013) recently found that upper secondary students seldom connect the protein concept to the gene concept, but rather use the DNA concept as an intermediate link between the two in the context of protein synthesis. This indicates that an even simpler mechanistic explanation is not easily accessible for students. Students' difficulties in linking genes with traits, grasping mechanistic explanations that link genes and traits, and their tendency to view genes as both Mendelian genes that determine traits and (separately) as molecules that code for proteins have also been addressed by van Mil et al. (2013). These authors present a framework for teaching students about links between molecular and cell, organ, and organism levels, but mainly focus on links between molecular and cellular phenomena.

Venville and Treagust (1998) showed that students mostly regard a gene as a particle and few develop an understanding of the connections between the genetic code, proteins, and phenotype. Students do not usually connect genes with a product (Lewis et al. 2000). Few students think of genes as codes for protein production, although they may have detailed knowledge about the structure of DNA and steps in protein synthesis. Marbach-Ad and Stavy (2000) showed that Israeli students in grade 9 had difficulties connecting phenomena at the macro and micro levels. Students in grade 12 were able to make connections to larger-scale extents, but without being able to explain the intermediate processes. In a later study by Marbach-Ad (2001) students were questioned about relations between concepts. Students in grade 9 mainly thought of genes as being inseparable from traits, while students in grade 12 viewed genes as determining traits. Only the students who had taken the most advanced courses, pre-service teachers, thought of genes as information for building proteins. Very few included in the study made connections between genes or DNA and proteins, and they failed to make the connection between proteins and traits. Similarly, Lewis and Kattmann (2004) found that 14–19-year-old German and English students viewed genes, as small trait-bearing particles whose main function was transfer between generations. Venville et al. (2005) also found that students comprehend genes and traits as being the same and argue in line with Lewis and Kattmann (2004) that this view could be a barrier to learning about the intermediate mechanisms involving proteins as the functional units. This is because if students view genes and traits as being the same, they will not need an explanation of how one affects the other. The results by Lewis and Kattmann (2004) and Venville et al. (2005) are in accordance with Venville and Treagust (1998) and Marbach-Ad (2001), indicating that this way of understanding gene function is a widespread phenomenon. Even if the students recognized a connection between genes and traits in which the genes affect traits, they did not hold a conception of the intermediate stages.

Instructional Issues and Suggestions

The suggestion from research is that teachers should help students to understand the relationship between different concepts and different levels of organization. Some problems are addressed in teaching and textbooks. For example, Venville and Treagust (2002) observed the teaching in eight high school classes (years 8, 10, and 12) and interviewed the students. They found that students had difficulty connecting different parts of genetics to obtain a broader understanding due to disconnected instruction of different aspects of genetics. This resulted in students having knowledge about protein synthesis but still being unable to connect genes with proteins. Students also showed an inability to give mechanistic explanations for the relationship between the gene and trait concepts.

Verhoeff et al. (2009) studied cell biology sections in two Dutch textbooks used in pre-university biology courses and showed that the content dealing with cellular and sub-cellular levels was detailed, but few connections were made to organism level phenomena. Further, in textbooks used in many countries (including European, North American, and South American countries and Australia), 37–51 % of the explanatory models of gene function are based on Mendelian or classical genetics (dos Santos et al. 2012; Gericke and Hagberg 2010a, b; Gericke et al. 2012). Thus, a major proportion of their content presents gene function in contexts that do not use proteins as a means to link the gene and trait concepts.

Some studies have focused on the design of genetics courses and the effects of including molecular processes in them. In a study of students' genetics conceptions by Tsui and Treagust (2007), classes in grades 10 and 12 were taught genetics in a course that included use of a computer program dealing with phenomena at different organizational levels, including molecular mechanisms. The results showed that students developed understanding of genetics, but very few had a conception of genes as information for making proteins even after instruction. Duncan and Tseng (2011) designed a genetics unit with four learning goals: (1) that the genetic information only specifies the structure of proteins; (2) proteins are the central biological elements that mediate genetic effects; (3) the sequence of amino acids in a protein affects the protein's overall structure and properties, which in turn both confer and constrain its function; and (4) the ability to reason across levels to generate explanations of genetic phenomena. Most of the students in grade 9 who participated in the study developed an understanding of genes as information for proteins. Their understanding of proteins as intermediates between genes and traits also increased. The students became more aware of proteins' diverse functions in the body, but not to the expected extent since they still seemed to prefer references to previously known protein structures, such as hair, nails, or enzymes. Based on their positive results of the students' learning, Duncan and Tseng (2011) suggest that mechanistic explanations of genetic phenomena could be used in middle school since they do not need to involve complex molecular details. In addition, many children aged 10–12 years already know about DNA from mass media, which often provides dubious ideas about genetics, thus raising the need for schools to start teaching genetics earlier (Donovan and Venville 2012). Eklund et al. (2007) designed a teaching unit with an explicit focus on the connections between genes and proteins. Their study was not focused on language, but it revealed linguistic challenges in the teaching of the topic and they therefore address the need for focusing on language.

In summary, previous science education research has contributed to our knowledge about the nature of students' difficulties with the gene–trait relation (e.g., Lewis and Kattmann 2004; Marbach-Ad and Stavy 2000) as well as how textbooks deal with this topic (e.g. Flodin 2009; Gericke and Hagberg 2010a, b; Gericke et al. 2012; Verhoeff et al. 2009). A few empirical studies have also included investigations of how genetics is taught (Duncan and Tseng 2011; Venville and Treagust 1998; Venville and Treagust 2002). However, we have found no published studies that have explicitly focused on how genetics is communicated during teaching situations within the classroom. This study investigates teachers' oral communication of genetics content in action, providing a genuine contribution to our knowledge of how genetics is really taught in the classroom.

Theoretical Framework

We consider the communication between the teacher and students to be at the core of teaching and thus think it is important to pay attention to the language used in teaching situations in order to understand student difficulties. The crucial role that language plays in science education has been stressed by many researchers (Fang 2005; Lemke 1990; Mortimer and Scott 2003; Ogborn et al. 1996; Wellington and Osborne 2001). However, given that teachers have an important role in helping students to master the language of science, with its specific terms and style of talking (Mortimer and Scott 2003), and that teachers are a major information source in the classroom, their use of language in teaching has been surprisingly little researched. It has been investigated in a few studies (e.g., Brown and Spang 2008; Dagher and Cossman 1992; Oyoo 2011; Wilson 1999), but few have specifically focused on the language used to communicate the science content.

In one of these studies, Nygård-Larsson (2008) analyzed teachers' language when teaching systematic biology, using the analytical framework presented by Lemke (1990). A central aspect of this type of analysis is to examine how different meanings of words are realized by the way they are connected to each other, i.e., their semantic relations. This is grounded in the theoretical framework of SFL developed by Halliday (1993). SFL is concerned with language as a source of meaning and considers the functions of grammar in creating and expressing meaning (Halliday and Matthiessen 2004). The analytical unit in SFL is the clause, which includes a process (something that happens) and at least one participant (someone or something involved in the process, e.g., an actor). The framework presented by Lemke (1990) is based on spoken language, and although the core of the framework is the clause, he expands it to examine how larger patterns of connections between words or concepts are built up into what he calls *thematic patterns*. A thematic pattern imparts the meaning of a specific scientific concept, in our study the concept of the protein.

Our objective was to map and characterize all cases where four recorded teachers talked about proteins in their teaching of genetics. The starting point of our analysis was therefore the word protein (or other words representing proteins, e.g., enzyme) to systematically explore the different ways proteins were included in the teaching. From this we could discern the typical thematic patterns in the four teachers' verbal communication (the analysis is described in further detail in the "Method" section). We believe this type of analysis could provide important information about possible causes for the difficulties students experience while learning genetics.

Aim

The aim of this study was to discern how teachers *talk* about proteins in their teaching of genetics and if they included the concept of proteins in linking different organizational levels, as recommended in the research literature (e.g., Duncan and Reiser 2007; Duncan and Tseng 2011; Venville and Treagust 2002). By exploring what is taught to the students through teachers' spoken language, we want to contribute to understanding of what lies behind students' challenges in linking the gene and trait concepts and their difficulties in understanding the role of proteins in this linkage.

Research Questions

While teaching genetics:

- to what extent do teachers refer to proteins?
- what role do teachers ascribe to proteins in their verbal communication?
- how do they use proteins as links between micro and macro level phenomena?

Methods

Data Collection

Four teachers, with teaching experience ranging from 6 to 12 years, teaching ninth grade students were observed over their entire genetics teaching sequences extending from 3 to 5 weeks per teacher. The criteria for the participating teachers were that they should:

- Be teaching genetics at the time of the data collection
- Have a teacher qualification in biology
- Have a minimum of 5 years teaching experience
- Work at schools that were located at a distance possible for the researcher to reach several times a week.

The teachers who agreed to participate worked at two schools. We denote the teachers *T1*, *T2*, *T3*, and *T4* respectively. T1 and T2 were teaching at one school and T3 and T4 at another. Both pairs planned their teaching together. All four teachers were educated at the same university. Swedish teacher education is regulated centrally; however, students can to some degree choose between different courses. All teachers in this study had taken university courses in cell biology that included genetics. T2 and T3 had also taken a university course in biochemistry including genetics. In addition, T1 and T3 had studied advanced courses which included genetics. Thus, the teachers in this study all have an adequate education for teaching biology at secondary level.

The teachers carried a microphone during the recorded lessons so everything they said during their whole teaching sequence was recorded, including dialogues with students. Additional data included were collected by observation and audio/video recordings of 45 lessons (41 h) in total. Lessons where the students had a substitute teacher were not recorded. Nor were lessons where they took tests. All audio recordings were transcribed. In Tables 1, 3, 5, and 7 presented below, the lessons are numbered chronologically as they occurred in the teaching sequence, i.e., the first lesson is denoted *1*, the second lesson *2*, etc.

The study was guided by the Swedish Research Council's ethical guidelines. Both the teachers and students were informed about the study and signed a consent form as an agreement to participate in it. They were also informed (orally and in the written agreement) that they could withdraw their participation at any point.

Data Analysis

Teaching Sequence Analysis

As a first step of the analysis we identified all the occasions where the teachers talked about proteins, i.e., all the text in the monologues and dialogues (in which the teacher participated) that related to the word protein was collected. Some of the sequences were several pages long, for example when a teacher was talking about protein synthesis. Other sequences were just a few sentences, for example when a teacher briefly mentioned that there are proteins in cell walls. From the collection of text excerpts, we assessed the extent of each teacher's talk about proteins, examined how it was sequenced over the lessons, and obtained an overview of the content. Collectively, this provides an idea of the prominence of the roles ascribed to proteins in the teaching. The results are summarized in Tables 1, 3, 5, and 7, where we describe the types of activity in which proteins were mentioned (e.g., *whole class instruction* or *students' individual* *work*) and briefly describe the science content of the teachers' spoken communications regarding proteins in each case (e.g., *protein synthesis* or *teacher mentioning protein threads in the context of mitosis*).

SFL Analysis

In a second step, we conducted an SFL analysis (Halliday and Matthiessen 2004) of the selected excerpts including talk about proteins. The analysis was centered around the word protein. However, although our starting point was the word *protein*, cognate words like *protein synthesis* or *protein threads* were also included. We also included words like *enzyme*, *hemoglobin* etc., i.e., examples of proteins. The *semantic relations* between these words and the surrounding words were then analyzed. Thus, our unit of analysis is protein (or enzyme, etc.) and surrounding words semantically related to protein. This was often within clauses, but could also be within a sub-ordinate clause, e.g., a circumstance. For example, in the sentence *Proteins build up muscles in the eye*, the word proteins is semantically related to muscles as an actor (that which is responsible for the action) in a process (the building of muscles), and then comes a circumstance, *in the eye*, which semantically relates the protein to a location. Thus, in a single sentence protein could be semantically related to several other words. This analysis enabled systematic exploration of ways the teachers talked about proteins' actions and locations, their explanations of how proteins are acted upon, how they defined proteins, and how explicitly they dealt with the taxonomy of proteins. We sorted the semantic relations as follows:

1. Process of doing and happening (see Halliday and Matthiessen 2004, pp 179–182 for further details). This category included all clauses where proteins were described as acting entities or affected entities. An example of such a clause is *The enzyme cuts the protein*. The process here is cuts (a process is typically described by a verb), while *the enzyme* is the *actor*, that which is responsible for the occurrence of the process. *Protein* in this case is the *goal*, the entity that is acted upon. As another example, in *Proteins give rise to skin color*, *protein* is the actor in the process *gives rise*, while *skin color* is the goal. Hence, analysis of clauses in this category clarified the functions of the mentioned proteins or processes affecting them.
2. Process of creation (see Halliday and Matthiessen 2004, pp 182–186), a subcategory of process of doing and happening. Since the teachers talked about protein synthesis and proteins as being built, we wanted to separate this talk from cases where they talked of proteins' involvement in processes to apprehend if this was the main content in their spoken communication, or if they also connected proteins to other processes of doing and happening. Therefore, we sorted talk about proteins being created into this separate category rather than including it in the process of doing-and-happening category.
3. Relational process (see Halliday and Matthiessen 2004, pp 210–214). This category refers to clauses that define and establish relations between two participants, which are commonly used in scientific texts. The use of the word process here could be somewhat misleading, but the process included in this type is typically *be* or *have*. That is, this type of process does not unfold through time and no input of energy is involved. Examples are *Muscles are proteins*, where one thing is said to be a type of, equal to, or the same as the other, or *Food consists of proteins.*
4. Locations (see Lemke 1990, p. 223) describe where something is located, e.g., *in the ribosome* (note that ribosome is not always a location, semantically, for instance in *The*

ribosome constructs proteins, *The ribosome* is an actor). Such clauses may also indicate something about the micro–macro perspective, depending on the type of structure the protein is connected to.

5. Hypernym/hyponym (see Lemke 1990, p. 221) is a category of taxonomic relations. We included in this category all examples of proteins and protein classes, if they were cited as such. For instance, in the sentence *An enzyme is a sort of protein*, enzyme is a hyponym of protein, that is, *enzyme* is explicitly related to proteins as a sub-category. However, in the example *DNA codes for enzymes and proteins*, enzymes and proteins are not mentioned hierarchically. Thus, some cases where words like *enzyme*, *insulin*, or *hemoglobin* were mentioned in the teachers' communications were not included in this category because they were not taxonomically related to proteins as hyponyms.

The results of the SFL analysis of each teacher's communications are summarized in separate tables (Tables 2, 4, 6, and 8). This analysis clarified, in detail, the typical ways in which each of the four teachers talked about proteins.

Thematic Pattern Analysis Addressing Research Question 1

We then identified the main messages put forward by the teachers, i.e., *the thematic patterns* that were communicated. This was done by taking both the results from the overall sequences of the content and the results of the SFL analysis into account. That is, based on how the content occurred in the teaching sequence (repeatedly or rarely) and the teachers' typical ways of talking about proteins as revealed by the SFL analysis, we could discern their dominant communication about proteins. Taken together, the findings demonstrate the roles ascribed to proteins in their teaching of genetics.

Thematic Pattern Analysis Addressing Research Question 2

To explore how the teachers used proteins as links between micro and macro level phenomena, we used data from the SFL analysis described above. By collecting and examining all the utterances including proteins from the SFL analysis, we were able to analyze the ways in which proteins were used (or not used) as links between micro and macro levels. We regarded visible traits as macro level phenomena in our analysis and other gene products as belonging to the micro level. For example, we identified a teacher talking about hemoglobin as a protein as an example of connecting proteins to micro level phenomena. However, we identified a teacher talking about proteins being responsible for eye color as use of protein to connect micro and macro levels. If a teacher said that a protein was responsible for a trait's occurrence, we saw this as an example of linking micro and macro levels, albeit in a very general rather than concrete manner.

In addition to examining the data from the SFL analysis, we also read through all the material obtained from the 45 recorded lessons to identify sequences with a biochemical focus where we expected to find talk related to micro–macro issues, but the word protein (and cognate words of proteins, or examples of proteins) was excluded. This additional analysis enabled us to identify occasions where a teacher talked about how different combinations of bases in DNA give rise to different traits, without including proteins (examples including proteins, cognates of proteins, and examples of proteins were all identified in the original SFL analysis).

Validity and Reliability

Observing, recording and transcribing a whole teaching sequence is time consuming, but since we wanted to explore how teachers talk about specific content, we think it was important to capture their use of language in action in the classroom while teaching. The validity is higher than it would have been if we had used an indirect methodology, for example, if we had interviewed teachers about their teaching. Our intention in this study was to study authentic teaching situations. That is, we did not want to affect the teaching as it is ordinarily undertaken and no intervention was made. However, the presence of a researcher in the classroom observing and recording the teaching inevitably influences the teaching, a phenomenon known as the Hawthorne effect (Robson 2002). It is difficult, if not impossible, to quantify the degree to which this effect compromises the authenticity of teaching, but to minimize the influence of the research design the teachers were not informed about the exact research focus, i.e., that we were going to explore how they talked about proteins. The teachers were just told that we were interested in studying genetics teaching in the classroom and that our study might reveal both positive and negative aspects of the teaching. Prolonged involvement is usually used in qualitative research to increase its validity (Robson 2002). We observed many lessons, which should have minimized the observer effect since we used a situated approach over a long-time period.

Member-checking was also used to enhance the validity of the results (Robson 2002). We used this methodology to confirm that the thematic patterns we identified in the teachers' spoken communications were reasonable in the teachers' opinions, by discussing the results of our analysis with one of them. The teacher thought the results were reasonable and accurately described how he/she taught the genetics content. Further, an external researcher with experience of linguistics analysis within the field of SFL was contacted and asked to conduct a parallel, independent SFL analysis of our dataset to secure intercoder reliability. The results were then discussed with this researcher until consensus was reached regarding the identified semantic relations.

Results

Of the 45 recorded lessons, 25 included talk about proteins. The teachers talked about proteins in various ways, with varying degrees of detail and varying emphasis on proteins as functional entities. In the following text and tables, all the results related to teachers 1–4 are described in consecutive order. First, we summarize how the teacher in question included protein in their teaching, in what lessons references to proteins occurred, the context, and an overview of the protein content. This is summarized in Tables 1, 3, 5, and 7. The results of the SFL analysis are then presented, which reveals the science content the teachers expressed, together with excerpts of specific interest. The main thematic patterns (the emphasized messages) communicated by each of the teachers are recapitulated in one or two sentences. Finally, we summarize how each teacher used proteins to link micro and macro level phenomena.

Teacher 1

Teacher 1 talked relatively little about proteins— mainly during instruction about protein synthesis, which occurred in two lessons. Proteins were also briefly mentioned a few of other times. The occasions when proteins were mentioned during this teacher's sequence of lessons are described and the contexts are summarized in Table 1.

Table 1 Occasions where talk of proteins occurred during T1's teaching sequence, consisting of 13 recorded lessons

Lesson no.	General description of activities and content including proteins
1	• Whole class instruction. Reasoning about whether the following teaching will include protein synthesis or not. Teacher mentions protein as a constituent of food
4	• Whole class repetition of the former lesson (which, amongst other things, was about the DNA molecule, including bases and their combination). Teacher mentions that DNA can have information about proteins
	• Whole class instruction about mitosis. Teacher talk about protein threads in cell division
6	• Whole class instruction about protein synthesis
	• Students' individual work. Students' questions about protein synthesis
7	• Whole class instruction. Repetition of protein synthesis

The lessons are numbered according to the order of the recording. The numbers (left column) indicate the lessons, for instance 4 indicates the fourth lesson, while the bullets indicate different occasions within lessons, in cases where talk about proteins was interrupted by other activities or other topics were covered

The SFL analysis shows that teacher 1 typically talked about proteins as things that are built. Proteins are included in 16 creative processes describing proteins coming into existence, like *proteins are constructed*, *build*, or *become* (see Table 2, column 2). Sometimes there is another actor in these clauses, *we* or *amino acids*, which are responsible for the action's occurrence, as in *We have amino acids that tie together and become a protein*.

The teacher also mentioned other processes that proteins are involved in, besides being created, saying for example that proteins are *transported*. However, the only action any protein was said to do (in the form of *protein threads*) was to *come out* in cell division, which occurred four times (see Table 2, column 1 under proteins in doing-and-happening processes).

Besides one occasion, where the teacher said *enzymes cut and break things*, proteins are not represented as actors in processes affecting another entity, a goal. Thus, the teacher did not point out other functional roles of proteins in the body. There were occasions that could have been starting points for discussion about proteins' other functions in the body, as in the example shown below, where a student and the teacher talked about the construction of a protein. The student then asked where the constructed protein goes. However, instead of explaining the fate of the protein and its possible tasks, the teacher's response is rather vague, saying that the protein *then goes where it should be*. After this response, the protein is said to be *transported*, and again the protein is involved in a process without affecting any goal:

> Student: And then it disappears, where does it end up, the protein goes like out…?
> Teacher: Exactly. Then it depends where the pr, well, if it is there it should be, then it goes where it should be.
> Student: Yes.
> Teacher: Is transported away.

In relational processes (see Table 2, column 3), we see that the teacher talks about the protein as *being needed*, as in *we need proteins* and *proteins are needed in different processes*, indicating that proteins have important functions, but as shown above, this is not taken any further except for the case of protein threads.

Several of the mentioned relational processes focus on the creation of proteins, as in *Protein synthesis is the construction of proteins.* Thus, even in the relational processes, in which the relations between participants are expressed statically, the main message appears to be that proteins are things that are built.

Table 2 Categories of semantic relations to proteins (or examples of proteins) used by teacher 1

Proteins in processes of doing and happening (actors and goals other than proteins in italics, process in bold)	Proteins in creative processes of coming into existence (actors other than proteins in italics, process in bold)	Proteins in relational processes (The related participants in italics, the relational verb in bold)
Proteins **are transported** (2)	Proteins are **constructed/built/become** (11)	*Proteins* **are** *part of food*
Proteins **go** where they should be	*You/we* **build/construct** the protein (2)	*Proteins* **are needed** *in processes*
Protein threads **come out/grow out/get stuck** (4)	*Amino* acids **are linked together/joined** and build/become protein (3)	*We* **need** *proteins*
Enzymes **cut and break** *things*		*They* **are** *protein threads*
We washed proteins		*DNA* **contains** *information* about *our inheritance or proteins*
The sequence **tells** what protein it's going to be		*Protein synthesis* **is** *the construction of protein*
		The process **is** *protein synthesis*
		Protein construction **is named** *protein synthesis*

Hyponymy	Locations of proteins
None	From the poles of the cell (2)
	In the ribosome (2)
	In our body (2)
	Where it should be
	Out into the cell into the vacuoles in the nucleus

Numbers in brackets indicate the frequency of each type of expression

The analysis showed that this teacher did not give any example of a protein, i.e., hyponymy. The teacher talked about enzymes, but drew no taxonomic connection between *enzyme* and *protein*, i.e., did not talk explicitly about enzymes as *examples* of proteins. *Protein* and *enzyme* were merely talked about as different entities, with different functions, treating enzymes as actors in the process of constructing proteins, and proteins merely as something constructed in protein synthesis. According to the connection between gene and trait, the gene–protein–trait path stops with the protein, i.e., we found no link between protein and trait in any of the processes mentioned in teacher 1's talk.

The additional reading of the transcripts (as described in the "Method" section) revealed that teacher 1 had a biochemical focus in one of the lessons where we could have expected to find some talk about proteins, but proteins were not mentioned. In this sequence, the teacher talked about the combination of bases in DNA as causing traits, but without mentioning proteins. The teacher explains in some detail how the bases are in certain combinations (A–T and C–G) in DNA and that the combination of the bases is what *tells*, or *determines* what we will look like:

> Teacher: (…) How these [bases] are combined then, we can see here, so it depends on how they are put together, so they constitute a gene, and it is the genes that..that is what we inherit you could say, it is what tells how we become and what we will look like.
> (…)
> Teacher: How these [bases] then are combined on this stretch, that is what determines the trait.
> (…)
> Teacher: It's a specific sequence of the DNA string that constitutes a gene and depending on how these bases are combined we get certain traits.

Later in the same lesson in an individual conversation with a student, this is further established:

> Teacher: But depending on how these are combined..
> Student: It's a gene, I mean a part of the DNA is a gene…
> Teacher: Yes exactly.
> Student: Okay.
> Teacher: And then depending on what (not audible) combines you get a certain trait.
> Student: Hmm.
> Teacher: So the gene for eye color maybe is..say it's located here.
> Student: Aha!
> Teacher: So if it is combined in one way it will be blue eyes..
> Student: Yes.
> Teacher: And if it is combined in another way it will be brown.
> Student: Clever.

The teacher repeatedly goes directly from the combination of the bases to the visible trait, without including proteins. However, the student seems pleased and apparently, from his/her reaction, thinks it is a reasonable explanation.

Teacher 1 also talks as if DNA contains information about protein *or* other things, saying:

> Teacher: (…) DNA that contains information about our inheritance, or it could be about a specific protein that should be constructed, actually about anything.

Regarding locations of proteins, teacher 1 said they are either connected to the cell (proteins are constructed on the ribosome) or more generally *in our body*, as in *proteins are needed in* *our body* (see Table 2, column 2 about locations). Proteins did not appear to link micro and macro level phenomena in any case, except for their unspecific location *in our body.*

To summarize, the main thematic pattern emphasized by this teacher was that *proteins are built*. Besides talking about proteins as being built, the proteins were not ascribed any other significance according to the semantic analysis. Teacher 1 did not use proteins as links between micro and macro levels, but jumped directly from combinations of bases in DNA or genes to protein and talked as if DNA contains information for proteins *or* other things. The path from micro to macro level phenomena did not involve proteins.

Teacher 2

Teacher 2 also mainly talked about proteins in connection to protein synthesis, but also talked about proteins more generally in other contexts, starting in the first lesson. Talk about DNA and genes, protein synthesis, and proteins' connections with traits occurred several times. The occasions when proteins were mentioned during this teacher's sequence of lessons are described and the contexts are summarized in Table 3.

The SFL analysis indicates that teacher 2, like teacher 1, typically talked about proteins as things that *are created*. For example, proteins were mentioned in 20 creative processes of coming into existence (see Table 4, column 2). The actors responsible for the creative process were *amino acids* or the general *you*, but teacher 2 also talked about *the skin* as constructing proteins, which could be seen as a way of connecting micro and macro levels, a structure at the macro level being responsible for the creation of a protein.

Teacher 2 also talked about proteins in processes of doing and happening (see Table 4, column 1) and repeatedly represented protein as an actor in clauses with a goal, e.g., eye color. That is, this teacher's way of talking ascribed active functions in the body to proteins. In ten of the processes of doing and happening mentioned, there is also a link between protein and a trait

Table 3 Occasions where talk of proteins occurred during teacher 2's teaching sequence, consisting of seven lessons in total

Lesson no.	General description of activities and content including proteins
1	• Whole class instruction about the cell and its constituents. The teacher says we need proteins to build our bodies and that they are built by ribosomes
2	• Whole class instruction about mitosis. The teacher mentions *protein threads*
3	• Whole class instruction about the DNA molecule, genes as information about proteins and proteins' connections to traits
	• Whole class instruction. The teacher mentions that protein threads drag chromosomes in cell division
6	• Whole class instruction about protein synthesis, including talk about gene–protein–trait connections
	• Students' individual work about protein synthesis, students' questions
	• Whole class instruction, summarizing protein synthesis
	• Students' individual work, where the teacher mentions that protein threads drag proteins in cell division
	• Individual student question about proteins in DNA
7	• Individual student question about where there is information about proteins in the textbook

The numbers (left column) indicate the lessons, for instance 6 indicates the sixth lesson, while the bullets indicate different occasions within lessons, in cases where talk about proteins was interrupted by other activities or other topics were covered

Table 4 Categories of semantic relations to proteins (or examples of proteins) used by teacher 2

Proteins in processes of doing and happening (actors and goals other than proteins in italics, process in bold)	Proteins in creative processes of coming into existence (actors other than proteins in italics, process in bold)	Relational (the related participants in italics, the relational verb in bold)
The protein **gives**, **carries** out a trait (4)	Proteins are **constructed** (5)	*Ribosomes* **are** *protein factories*
Protein **colors** *the eye* brown/blue (2)	*Ribosomes* **construct** proteins (4)	*A gene/part of DNA* **is** a *recipe for protein* (2)
Proteins **give** *us* brown color	*Ribosomes* **construct** proteins so we can grow and build our body	*Traits* **consist** of *proteins*
Proteins **settle** in the shallow cells in the skin	*Amino acids* **are joined together**—we get a protein (2). *Amino acids* **connect** and we get a protein (2)	*The brown color, that* **is** *a brown protein*
Proteins **build up** *muscles* in the eye	*We/you* **construct** protein (5)	*Proteins* **are** built of *amino acids* (2)
Proteins **protect** *us* from skin damage	*The skin* ***constructs*** protein	*A gene, a bit of DNA* **has** the *recipe to construct a protein*
Protein threads **pull** *them* [chromosomes] (2)		*Genes* **are** *recipe codes for the construction of proteins*
Specific places on the DNA molecule **tell** *what proteins* should be		*There are places that* **are responsible** *for the recipe to construct protein*

Hyponymy (hypernym in bold)	Location
Protein—melanin (2)	In the eye (3)
	In the body, in the cells
	In the shallow cells in the skin
	In the skin cells
	In skin
	To the ribosome in the cell
	In the ribosomes that are in all cells in the body

Numbers in brackets indicate the frequency of each type of expression. The square brackets around chromosome indicate that the word is mentioned implicitly, referring to another clause at the macro level. Moreover, the relation between protein and trait is manifested in different relational clauses, e.g., when the trait is said to *consist* of proteins.

The next example shows how the teacher emphasized the function of proteins as contributing to traits. The relation between protein and trait is manifest in clauses of doing-and-happening and relational clauses, where the trait is said to *consist* of proteins:

> Teacher: All our traits you could say consist basically of proteins. Our eye color for example, if you have brown eyes, that is proteins which color the eye brown. And if you have blue eyes, a protein colors the eye blue.

Or, as in the following example:

> Teacher: Yes but what the genes really do, they give us certain traits, for example eye color and everything, hair color (…) What they really do is that they are the recipe for proteins that are constructed so that the body, it is proteins that give us, that carry out so to say the traits and give us the brown color, that is a brown pigment, so it is proteins that build up the muscles in the eye that makes this specific color right there. And the issue I talked about once when you become brown and you sunbathe, melanin is the name of that protein, that is a gene, that is, a part of DNA that has the recipe for us to construct those in the skin cells when you are exposed to sun. So genes are recipes for the construction of certain proteins that give us certain traits.

As we can see in the example above, this teacher cited one example of protein, i.e., melanin, which is explicitly related to protein as a hyponym.[2]

The eye is a location mentioned several times (see Table 4, column 2). Another location that protein is connected to in this lesson is *the shallow cells in the skin*. This also connects proteins to the macro level by placing them in a macro level structure. A connection between genes (or DNA) and proteins is made seven times, e.g., in relational processes like *A gene is a recipe for a protein* (see Table 4, column 3).

To summarize, the general thematic pattern in this teacher's utterances could be rephrased as *proteins are constructed from the information contained in the DNA/gene and proteins are responsible for traits like eye color or pigments*. This teacher gives an overall picture of how genes are related to traits with proteins as intermediates, and although few examples are used (eye and skin color), he/she repeatedly and explicitly makes the gene–protein–trait path clear. The proteins are described as active, functioning entities in cells and thus used as links between micro and macro level phenomena.

Teacher 3

Teacher 3 talked about proteins in several lessons, as shown in Table 5. However, the talk was brief and not detailed. Furthermore, although the teacher talked about protein being constructed by ribosomes, no instruction about protein synthesis was included. Instead, the teacher suggested that students search for information about protein synthesis on their own at home. He/she also repeated on several occasions that DNA codes for proteins, but also has another function.

The results of the SFL analysis showed that teacher 3 also mostly talked about proteins as being built. For example, proteins (or examples of proteins) were mentioned in 16 creative processes of proteins coming into existence (see Table 6, column 2). Proteins were also defined as *substances that are built* in a relational clause (see Table 6, column 3). This might give the

[2] Melanin is actually not a protein, but a polymer constructed from the amino acid trypsin, catalyzed by the enzyme tyrosinase (protein). However, melanin is an intermediate that links the gene with the trait.

Table 5 Occasions where talk of proteins occurred during teacher 3's teaching sequence, consisting of 12 lessons in total

Lesson no.	General description of activities and content including proteins
3	• Students' individual work. The teacher talks with an individual student about lysosomes transporting protein
5[a]	• Whole group instruction. The teacher says that this is one of DNA's functions (the other is to code for cell division). Insulin is given as an example of a protein
6[a]	• Whole class instruction about ribosomes constructing proteins and enzymes. Insulin is mentioned. DNA is said to code for proteins, and the teacher says that this is one function, while the other is to copy itself
	• Whole class instruction. The teacher mentions again that ribosomes construct proteins, adding that proteins are building blocks for the material in the body, and repeating that DNA's other function is to code for cell division
7	• Students' individual work. A student asks the teacher if ribosomes have something to do with building proteins, the teacher answers that they do, that the ribosomes build proteins which are put together to form enzymes and other things
8	• A student mentions protein in connection to talk about food
	• A student mentions something about protein-rich
9[a]	• Whole class instruction. The teacher says that students could read about how proteins are made, because that is *also* a consequence of the DNA code
	• Students' individual work. The teacher talks with an individual student about muscles and proteins
	• Students' individual work. The teacher talks to individual students about the genetic code as having two functions: as a recipe for proteins and replicating itself through new individuals
10[a]	• Whole class instruction. The teacher says that DNA codes for proteins, its function is not just spreading genes, talks about insulin in gene technology
	• Students' individual work: The teacher mentions crops as being *protein-rich* while talking about GMO with an individual student
12	• Whole class instruction about what the students are supposed to work with. The teacher tells the students that they can read about how proteins are constructed, because DNA is *also* a code for the proteins that should be built

The lessons are indicated by numbers and the bullets indicate different occasions where the content included proteins

[a] This class was sometimes divided into two groups that were given almost identical lessons. Lesson 5 and 6 as well as 9 and 10 were such similar lessons

impression that teacher 3 talked a lot about proteins, but since it was in short utterances without further explanation (e.g., there was no instruction about protein synthesis), we do not see that this teacher ascribed an important role to proteins.

Proteins were mentioned in a few processes of doing and happening, e.g., as being transported. Protein is not referred to in the context of acting, although in one example protein is indirectly ascribed a function in a clause where *insulin* (previously mentioned by the teacher as an example of a protein) is said to regulate blood sugar levels. The following excerpt illustrates the thematic pattern of proteins being things that are built, and the teacher gives an example of a protein, insulin:

> Teacher: (…) what are proteins?
> Student: Vitamins…
> Teacher: Yes but they are substances that are built in different places in the body. It could be for example insulin.

Table 6 Categories of semantic relations to proteins (or examples of proteins) used by teacher 3

Proteins in processes of doing and happening (actors and goals other than proteins in italics, process in bold)	Proteins in creative processes of coming into existence (actors other than proteins in italics, process in bold)	Relational (the related participants in italics, the relational verb in bold)
[Insulin] **regulates** *blood sugar levels*	Proteins are **built, constructed/put together** (7)	*Proteins* **are** *building blocks* (2)
Lysosomes **transport** protein *(2)*	*Ribosomes* **construct** proteins (4)	*Proteins* **are** *building blocks for all material*
DNA **codes** for proteins *(2)*	*Ribosomes* **construct** proteins and enzymes	*Protein* **are** *substances that are built*
	Ribosomes **construct** for example *insulin*	*Muscles* **are** *proteins*
	DNA **controls** what protein should be constructed	*Genes* **are** *codes for different proteins*
	Chromosomes **code** for what proteins **should be built**	*DNA* **is** *a code for proteins*
	Pigs **can produce** insulin	

Hyponymy (hypernym in bold)	Location
Protein—insulin	In the body (3)
Hormone—insulin	On the ribosomes

Numbers in brackets indicate the frequency of each type of expression. The square brackets around insulin indicate that the word is mentioned implicitly, referring to another clause

The teacher thus expresses taxonomic relations between proteins and hyponyms, such as insulin, and between proteins and a class of substances (vitamins) that are not proteins. However, in some cases, the taxonomic relations are vague or inconsequent and it is not clear whether another mentioned substance is a type of protein. For example, in the following sequence, the teacher expresses a taxonomic relation but here *insulin* is a hyponym of *hormone*, without any reference to hormones being proteins.

> Teacher: What decides that it is specifically this protein, or hormone I should say, insulin is a hormone.

In the next example, proteins and enzymes are placed at the same taxonomic level by the wording, while insulin is a hyponym, but of enzyme, and it is not clearly stated that enzymes are a class of protein.

> Teacher: The ribosomes! They construct proteins and enzymes, for example insulin.

Genes, DNA, and chromosomes are related to proteins a couple of times by the teacher; the DNA is portrayed as an actor in the process of controlling which protein should be built. When reading the transcripts to get an overview of the presented content, we also noted that this teacher repeatedly communicated that DNA has two functions, to code for proteins and to code in cell division. The teacher probably intended to refer to replication, but interestingly used the word code:

> Teacher: So, this is one of the things that DNA does. That is to code for proteins and that is really, that is a really big assignment. But then the DNA has another assignment as well, to code for something else.
> (…)
> Teacher: What do cells do when they become more?
> Student: They divide.
> Teacher: They divide yes. So the other thing that DNA has, it is to code in cell division.

In one example, an explicit link is made between protein and a macro level structure, in a relational process that defines muscles as being proteins (see Table 6, column 3), but no connections are made between proteins and traits. The location coupled to the proteins is *ribosomes* or generally *in the body* without relating proteins to a macro level (see Table 6, column 2 about location).

In summary, the main thematic pattern in teacher 3's utterances is that *DNA has two functions: to code for proteins that are built, and to copy itself and spread the genes.* That is, the teacher emphasizes that the construction of proteins is one of two functions of DNA. The gene–protein–trait path stops with protein, except for the general utterance that proteins are building blocks for all the material in the body.

Teacher 4

This teacher talked considerably more about proteins than the others. Proteins were mentioned in several lessons and in various contexts; for example, proteins were included in talk about the cell, protein synthesis, and gene technology. The instruction included scientific content at a relatively detailed level (Table 7).

The SFL analysis showed that proteins occurred in numerous clauses, of various categories, used by teacher 4 (see Table 8). He/she talked about proteins as being constructed; they

Table 7 Occasions where talk of proteins occurred during teacher 4's teaching sequence, consisting of 13 lessons in total

Lesson no.	General description of activities and content including proteins
2	• Whole class instruction about the cell and its constituents, the teacher says that ribosomes construct proteins, which gives the cell structure and builds our bodies
3[a]	• Whole class instruction. The teacher mentions that genes correspond to athletic ability etc. and protein
	• Whole class instruction. The teacher talks about proteins as being constructed of amino acids and that we have to eat proteins, which are degraded into amino acids that are put together to proteins specific to us with the help of the DNA code
4[a]	• Whole class instruction. The teacher mentions that genes correspond to hair color etc. and proteins
	• Whole class instruction. The teacher talks about proteins being constructed of amino acids in the ribosome with the help of the DNA code. Proteins are building blocks for us humans
6	• Lab. activity: DNA extraction. The teacher mentions that the students should get rid of the proteins, mentions proteins as being in the cell walls in lab
9	• Whole class instruction. The teacher mentions milk protein in a dialogue with students about lactose intolerance
	• Whole class instruction. The teacher talks about how protein affects tomato skin in a GMO context
10	• Whole class instruction. The teacher mentions that a cell must have proteins to be able to build new cells
11	• Whole class instruction. The teacher talks about diseases that occur due to lack of specific proteins (SCID and hemophilia)
	• Whole class instruction. The teacher talks about growth hormone, insulin in connection to deficiencies and medicines
12	• Whole class instruction. The teacher talks about proteins in connection to doping and red blood cells
	• Whole class instruction. The teacher mentions proteins in connection to talk about transplantation

The lessons are indicated by numbers in the left column and the bullets indicate different occasions where the content included proteins

[a] Some of the lessons were given in half class, which resulted in two very similar lessons. Lessons 3 and 4 were such lessons

occur in 12 creative processes (see Table 8 column 2). There were several different actors responsible for the construction of the proteins mentioned, e.g., *amino acids*, *genes*, *cells*, *the thyroid gland*, and *we*. Thus, this teacher cited several structures, both micro and macro, that are involved in the construction of proteins.

This teacher talked not only about the construction of proteins but also their degradation; about how we eat proteins, the proteins are broken down into amino acids in our body and the amino acids are subsequently put together as suitable proteins:

> Teacher: So these amino acids are building blocks for proteins and we must eat proteins and proteins then they are broken down in the body, or are degraded to amino acids and then our cells can, with the help of this DNA-code, put together these amino acids in the right order so we get our protein right. So we get the right protein.

The teacher also mentioned proteins as being participants in 19 different processes of doing and happening, where proteins were often actors, highlighting different functions of proteins

Table 8 Categories of semantic relations to proteins (or examples of proteins) used by teacher 4

Proteins in processes of doing and happening (actors and goals other than proteins in italics, process in bold)	Proteins in creative processes of coming into existence (actors other than proteins in italics, process in bold)	Relational (the related participants in italics, the relational verb in bold)
Proteins **give** *the cells* structure	*Amino acids* **build up** our proteins (2)	*Amino acids* **are** *building blocks in proteins* (3)
Proteins **build** muscles	*Amino acids* **are joined together** [into a protein]	*Proteins* **are** *building blocks* in humans
Proteins **build up** hair	*We* **construct** proteins	*Proteins* **are** building blocks
Hemoglobin **controls, helps** in oxygen transport	*Cells* **construct** the protein	*The whole of us* **consists of** *proteins*
Proteins **build up** insulin, hemoglobin or hormones	*Gene that cannot* **construct** protein	*Hair* **consists of** *proteins*
The testosterone **makes** *you* build germ cells	*Genes* **construct** *the protein*	*Skin cells* **consist of** *proteins*
Proteins that **break down** *the shell*	*The thyroid gland* **produces** *growth hormone*	*We* **need** *proteins*
The protein **participates** in the healing process	*You* **produce** *growth hormone*	*It* [*the cell*] must **have** *new proteins* to build new cells
Protein **build** *red blood cells*	Hemoglobin or insulin **are built**	*The body* **needs** *hemoglobin*
Proteins **break down, are degraded**	*You* **produce** testosterone molecules	*The cell* **needs** *amino acids*
Protein **breaks down**	*We* **can build** *hemoglobin or insulin*	*Gene* **corresponds to** *protein*
The proteins **are packed**		*Gene* **corresponds to** *eye color, hair color, protein*
We **eat** protein (3)		*Don't* **have** *much of that protein you don't grow in the way you should*
The body **cuts** proteins apart		*Ribosomes,* [*are*] *the protein factory of the body*
Part of the DNA molecule **codes** for a protein		*A gene* **is** *a recipe for a specific protein (2)*
The code **tells** what the protein should look like		*A sequence* **corresponds to, is** *a recipe for, a protein*
Gene that cannot **break down** milk sugar		*Part of the DNA molecule* **corresponds** *to a protein*
		Genes **correspond to** *eye color, athletic ability, muscle mass and proteins*

Hyponymy (hypernym in bold)	Location
Protein—hemoglobin or insulin	In the body (7)
Protein—female sex hormones **Protein**—male sex hormone (2)	In blood
Protein—hemoglobin (2)	Out from the nucleus, to the protein factory
Medicine—insulin, growth hormone	To the cell
Protein—growth hormone	In the cell plasma

Numbers in brackets indicate the frequency of each type of expression. The square brackets around words indicate that the words are mentioned implicitly, referring to another clause (see Table 8, column 1). For example, proteins (or examples of proteins) *build red blood cells*, *help in oxygen transport*, or *build insulin*. These are micro level functions, but macro level structural functions of proteins were also mentioned in cases such as *gives the cells structure*, *builds muscles*, or *build up hair*.

Proteins were also connected to macro level structures, like hair, in clauses expressing several relational processes:

> Teacher: Well it's the proteins that give us..the cells' structure…Our hair consists of proteins..so we need proteins for our hair to grow. Our skin cells consist of proteins, well the whole of us consists of proteins, is built of proteins (…)

In the processes of doing and happening mentioned, we see that genes are also portrayed as the actors in utterances about genes coding for proteins. However, genes are also represented as actors when the teacher says that genes cannot break down milk sugar (see Table 8, column 1), implying that the genes act directly (or do not) in a micro level structure, without the involvement of a protein.

Teacher 4 also expressed taxonomic relations to examples of proteins in several ways (see Table 8, column 2), strengthening our view that this teacher 4 ascribed the most prominent roles to proteins. The overall thematic pattern in this teacher's utterances is not easy to discern, since he/she communicates so many things about proteins. However, a message repeatedly communicated is that *amino acids are put together with the help of DNA, to become proteins that have different functions in the body.*

The teacher expressed the gene–protein–trait path and used proteins as links between micro and macro phenomena in some cases. However, most examples remained at the micro level. Furthermore, the teacher ascribes two functions to genes: coding for proteins *and* macro level traits, as in the following excerpt:

> Teacher: So on these chromosomes there are different genes that correspond to eye color.. athletic ability…muscle mass, eh..and certain proteins that we need in the body.

Thus, this teacher made connections between genes and proteins but did not apparently explicitly talk about genes as exclusively coding for proteins.

Another example found was in the context of gene technology, when the teacher jumped from gene function to macro level phenomena (human speech and walking upright):

> Teacher: Another such a change is actually speech, that we can talk as we do, put words to things. That is something that we have not always been able to do. It is a change in our genes that made us able to communicate with each other, as well being able to walk upright, on two legs. We came from four-legged creatures.

The above example does not mention proteins and was identified in the additional reading of the transcripts.

Summary

To summarize the results from the teaching of all four teachers, we can say that teacher 1 talked relatively little about proteins and when doing so did not indicate that proteins have many functions; the main thematic pattern discerned is simply that *proteins are built*. Teacher 1 talked about proteins as being needed, without giving any further explanation regarding their functions, and without using proteins to link micro and macro level phenomena (jumping instead from genes to macroscopic traits, or from genes to proteins (see Fig. 1)). Teacher 2 ascribed more prominent roles to proteins, the main thematic pattern in the teaching being that *proteins*

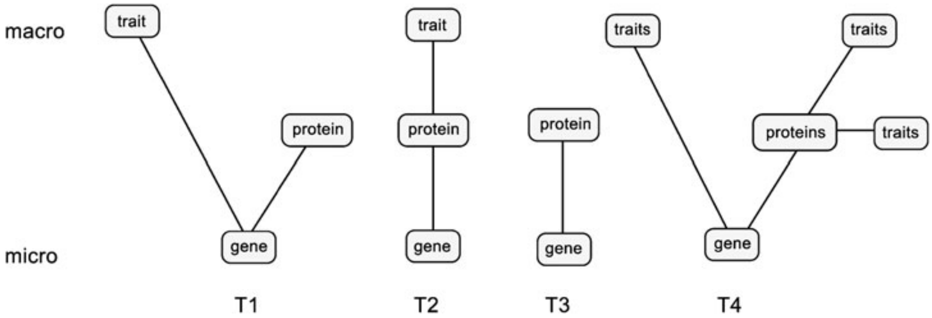

Fig. 1 Illustration of how the teachers used proteins as links between micro and macro level phenomena. Teachers 1 to 4 (T1 to T4) from *left* to *right*. Teacher 1 either jumped from gene (combination of bases in DNA) directly to macroscopic traits or stopped at protein. Teacher 2 explicitly used proteins as links between the levels. Teacher 3 stopped at protein. Teacher 4 provided numerous examples of protein functions and used proteins as links between levels, but mostly used micro level examples of protein functions. Teacher 4 also jumped directly from gene to trait, without involving proteins, as an alternative path

are constructed from the information contained in the DNA/gene and proteins are responsible for traits like eye color or pigmentation. Proteins act as functional entities that affect other entities, such as eye color and skin color and thus use protein as a link between micro and macro level phenomena (see Fig. 1). Teacher 3 did not emphasize proteins, but talked about proteins as being built. The thematic pattern discerned is that *DNA has two functions; to code for proteins, which are built, and to copy itself and spread the genes.* Teacher 3 jumped from gene to protein without making any link between proteins and macro level phenomena (see Fig. 1). Teacher 4 talked most about proteins and in various contexts. One thematic pattern repeatedly communicated was that *amino acids are put together with the help of DNA, to become proteins that have different functions in the body.* Thus, this teacher used proteins as links between micro and macro level phenomena, but typically described proteins as functioning at the micro level. He/she also expressed the idea that genes code for traits *and* proteins without mentioning the intermediary involvement of proteins in the former (see Fig. 1).

Discussion

The messages expressed by the four teachers when talking about proteins differed substantially, as summarized in the last part of the "Results" section. If four teachers talk so diversely about proteins, regarding both their importance and functions, it seems likely that more generally teachers' talk about this subject varies greatly. We identified the main message emphasized by each teacher as a thematic pattern and found that this pattern differed considerably among the teachers. This implies that students are presented widely differing content regarding the role of proteins in genetics lessons and hence have widely differing opportunities to learn about the subject. However, there are also some similarities amongst the four teachers' communications.

As shown by the SFL analysis, all four teachers highlighted the importance of the creation of proteins through numerous creative process clauses. This seems to be the common base for the four teachers. For example, teacher 1 talked relatively little about proteins and when doing so did not ascribe the proteins many functions other than being constructed.

Teachers 1, 2, and 4 included instruction about protein synthesis in their teaching. Notably, however, teacher 3 excluded this topic from the teaching and recommended that students look for information on the internet. This is a very demanding task for a student. Even if the students had been able to understand some detailed aspects on their own, they probably would not have been able to assemble the information to form a coherent overall understanding of the process, since a focus on details of (for example) transcription and translation does not help students to understand the general connections between genes, proteins, and traits (Duncan and Reiser 2007). Nevertheless, despite excluding instruction about protein synthesis, teacher 3 emphasized the construction of proteins since creative process clauses were the most abundant category of semantic relations in this teacher's utterances concerning proteins, as shown in the SFL analysis (Table 6).

A further striking feature of all four teachers' communications is what they did *not* say. None of the teachers explicitly talked about genes as exclusively coding for proteins (or sequences of amino acids). In the previously cited study by Duncan and Reiser (2007), students who had been explicitly taught about the central dogma, without understanding exactly what genes contain information about, thought of genes as coding for different products at different organizational levels at the same time. That is, the students thought of DNA as coding for proteins *and* traits. Similarly, three of the four teachers we recorded talked about genes as providing information for the construction of not only proteins but also other things, and two implied that genes are associated with (or act directly at) macro as well as micro levels, as in utterances like *Genes correspond to eye color, muscle mass and protein.* Teacher 3, especially, emphasized that protein construction is one of two DNA functions. This is not the same as saying that DNA codes for components, structures, or features of different organizational levels (or traits), but it certainly blurs the key message that DNA codes for amino acid sequences (*not* traits).

The teachers differed in their talk about the function of proteins. Teacher 1 talked about proteins as being needed, but without giving any further explanations about the functions of proteins. Similarly, Duncan and Reiser (2007) found that students know that proteins are important, but cannot say for what reason. Teacher 3 did not talk much about proteins as functional units. This lack of emphasis on proteins' roles in genetic teaching cannot be expected to help students acquire understanding of the importance of proteins and linking genes and traits.

The other two teachers included proteins in processes of doing and happening to a much greater extent, that is, they talked more about proteins as functional units. Teacher 2 talked about proteins as giving rise to different traits at the macro level, while teacher 4 mainly talked about proteins at the micro level, in the form of hormones, insulin, etc. In one respect, teacher 4 gave a holistic picture by talking about the whole chain from proteins being eating and degraded into amino acids, which are re-assembled into proteins (e.g., hemoglobin) with DNA as a recipe. On the other hand, this teacher went into numerous details and consequently presented rather fragmented content. As shown by the SFL analysis of teacher 4's utterances, numerous actors were mentioned, in both processes of creation (e.g., amino acids, genes, cells, and our whole bodies) and in processes of doing and happening (proteins, the code, the body), which were not explicitly connected to each other. Moreover, this teacher implied that genes act directly, as in the example of breaking down milk sugar (see Table 8, column 1).

The teachers' examples of proteins also varied in explicitness. For example, our SFL analysis identified examples of classes of proteins being mentioned, but not semantically related to protein as hyponyms (notably enzymes); thus, the wording did not indicate that enzymes are a class of protein. We believe it is important to pay attention to these kinds of issues because as teachers we might not always think carefully enough about the ways we are talking. Simply mentioning hormones, enzymes, and insulin does not help students understand the relations between them.

The teachers also varied in their use of proteins as links between micro and macro level phenomena (see Fig. 1), giving students' different opportunities to understand the links. Lewis and Kattmann (2004) argue that being able to distinguish between genes and traits as different things is highly important in order to be able to learn about the biological and causal mechanisms linking them. They found that despite teaching about genetics, students still have difficulties with this and that some of the students' erroneous views come from everyday life experience and ways of talking. However, based on our results, we would like to raise another possibility; that students may think in certain ways not *despite* the teaching, but *because* of the teaching. To put it simply, maybe students learn what they are taught. This is supported by previous findings that in classes where teachers emphasize the particle model of the gene, the students do not develop an understanding of the gene as an instructional unit for proteins (Venville and Treagust 1998). Our results indicate that what teachers communicate in the classroom is not very different from previously identified naive students' conceptions (Lewis and Kattmann 2004; Marbach-Ad, 2001). We cannot say anything about how the four teachers' communications we recorded affected the students' understanding since that was beyond the scope of our study. However, we can conclude that students' learning about contemporary genetics, including the connections between the gene–protein–trait concepts, as called for by many researchers (e.g., Duncan and Reiser 2007; Duncan and Tseng 2011; Venville and Treagust 2002), is unlikely to be helped by some elements of teachers' verbal communication of genetics within the classroom (this study) or descriptions and explanatory models of gene function in textbooks (dos Santos et al. 2012; Gericke and Hagberg 2010a, b; Gericke et al. 2012).

Duncan and Reiser (2007) mention that the teacher in their study focused on DNA's structure, rather than its function, which might at least partly explain why students referred to genes' structure rather than function, i.e., *genes as particles* rather than *productive sequences of information that code for proteins*. Similarly, Venville and Treagust (1998) found that students typically have a particle view of the gene, even after teaching about genetics, prompting them to conclude (p. 1052):

> Perhaps it is time for teachers to give that image a vibrant and productive boost by stepping out of the 1950s and the fixation with the double helix, and into a world where genetic process is important.

Even if we would not say it is a *fixation*, we can see how this exemplifies the type of teaching that focuses on details, e.g., about the DNA molecule, without providing an overall picture incorporating proteins' central role and the links they provide between different organizational levels. Similar results are reported in a study by Gericke and Wahlberg (2013) who found that upper secondary students emphasized the structural entities of protein synthesis rather than the functional, and also the students did not generally link the protein concept to gene concept, but preferred the DNA concept. The particle gene concept seems to be an obstacle for learning functional genetics that includes proteins.

According to Duncan (2007), it is not necessary to have extensive knowledge of details to be able to reason about genetic phenomena, but an understanding of the central role of proteins and proteins as links between genes and traits is crucial and a key resource in genetic reasoning. Similarly, van Mil et al. (2013) conclude that proteins' functions can be highlighted without going into exhaustive molecular details and Allchin (2000) states that "A few well-chosen examples can illustrate the link between genes, proteins and physiological traits." (p. 638). Teacher 2 talked about proteins as functional entities affecting other phenomena, such as eye color and skin color, and used proteins as links between micro and macro levels (see Fig. 1). Although the examples were limited and basic, this seems to be a possible way of presenting proteins as links between organizational levels at a suitable level for students in compulsory school. Therefore, we see this teacher's way of including proteins in the link between micro and macro levels as an example of how this could be done in a simplified, appropriate manner at this age level. However, the validity of this hypothesis remains to be explored.

Although we see connections between students' difficulties described in earlier research with the ways these four teachers communicated the subject, we want to emphasize that this is not a critique of the teachers. The teachers all talked differently about proteins, incorporating them to varying degrees in their teaching about genetics. The teachers have similar educational backgrounds and worked in pairs at their schools. The teachers at the same schools used the same textbooks and planned the teaching together. Nevertheless, they talked about and included proteins in their teaching in different ways. This may have been partly because the Swedish curriculum at the time of the study (Swedish National Agency for Education 2006) was very general, open to interpretation, and gave little guidance about the specific content that should be taught in genetics in general and none about proteins. The current Swedish curriculum, which was recently implemented (Swedish National Agency for Education 2011), does not go into specific details about content either and does not say anything about proteins in the context of genetics. Thus, the call for attention to proteins (e.g., Duncan and Tseng 2011) in science education research has not reached Swedish curricula. Proteins are also sparsely treated in the textbooks used by the teachers. Since the teachers do not have any support, either in the curricula or the teaching tradition, they cannot be expected to talk about proteins in the way researchers recommend. However, we still think it is important to focus on how teachers treat this subject; indeed, the absence of support for the teachers in this respect may increase its importance.

Using linguistic analysis, we were able to show how teachers talk about proteins when teaching genetics. We see the language used by the teacher as the very core of teaching. According to Hand et al. (2003), we need to pay careful attention to the scientific language used in education, and many other researchers agree that the language used in teaching is important (Lemke 1990; Mortimer and Scott 2003; Ogborn et al. 1996; Wellington and Osborne 2001). We therefore suggest that more studies should focus on the verbal communication in classrooms during the teaching of other science disciplines, in addition to genetics. The methodology applied in this study could be used as an example of a fruitful way to analyze how teachers teach specific content in the classrooms. There have been very few such studies using an SFL approach to analyze how science is taught in action, and there is a need for more studies in this neglected area since we still know very little about how teachers talk about different science topics in actual teaching situations.

Teacher education has an important task to prepare students for their coming professions, and we advocate a greater focus on language because of the central role it plays in teaching. As a teacher, it is crucial to be aware of how different expressions create different meanings, so it is important to pay attention to how we semantically relate words to each other. If, for example, we talk about proteins and enzymes at a taxonomically equal level, neither will have hyponymous status as an example of the other. As a further example, if we say that genes are codes for traits *and* proteins, this will imply that these are two, independent alternatives. We argue that addressing these issues will highlight important linguistic features that have profound consequences for teaching and learning, not just in genetics, but also science in general.

Acknowledgments We would like to thank The Hasselblad Foundation for financing this research. We would also like to thank the teachers and students who participated in the study.

References

Allchin, D. (2000). Mending Mendelism. *The American Biology Teacher, 62*(9), 632–640.

Brown, B. A., & Spang, E. (2008). Double talk: synthesizing everyday and science language in the classroom. *Science Education, 92*(4), 708–732.

Dagher, Z., & Cossman, G. (1992). Verbal explanations given by science teachers: their nature and implications. *Journal of Research in Science Teaching, 29*(4), 361–374.

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

Donovan, J., & Venville, G. J. (2012). Blood and bones: the influence of the mass media on Australian primary school children's understandings of genes and DNA. *Science Education*. doi:10.1007/s11191-012-9491-3.
dos 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*(4), 543–578.
Duncan, R. G. (2007). The role of domain specific knowledge in generative reasoning about complicated multileveled phenomena. *Cognition and Instruction, 25*(4), 271–336.
Duncan, R. G., & Reiser, B. J. (2007). Reasoning across ontologically distinct levels: studentst' understandings of molecular genetics. *Journal of Research in Science Teaching, 44*(7), 938–959.
Duncan, R. G., Rogat, A. D., & Yarden, A. (2009). A learning progression for deepening students' understanding of modern genetics across the 5th–10th grades. *Journal of Research in Science Teaching, 46*(6), 655–674.
Duncan, R. G., & Tseng, K. A. (2011). Designing project-based instruction to foster generative and mechanistic understandings in genetics. *Science Education, 95*(1), 21–56.
Eklund, J., Rogat, A.D., Alozie, N., & Krajcik, J.S. (2007). Promoting student scientific literacy of molecular genetics and genomics. *Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, April 2007, New Orleans*
Fang, Z. (2005). Scientific literacy: a systemic functional linguistics perspective. *Science Education, 89*(2), 335–347.
Flodin, V. (2009). The necessity of making visible concepts with multiple meanings in science education: the use of gene concept in a biology textbook. *Science Education, 18*(1), 73–94.
Gericke, N. M., & Hagberg, M. (2007). Definition of historical models of gene function and their relation to students' understanding of genetics. *Science & Education, 16*(7–8), 849–881.
Gericke, N. M., & Hagberg, M. (2010a). Conceptual incoherence as a result of the use of multiple historical models in school textbooks. *Research in Science Education, 40*(4), 605–623.
Gericke, N. M., & Hagberg, M. (2010b). Conceptual variation in the depiction of gene function in upper secondary school textbooks. *Science & Education, 19*(10), 963–994.
Gericke, N. M., Hagberg, M., & Jorde, D. (2013). Upper secondary students' understanding of the use of multiple models in biology textbooks—The importance of conceptual variation and incommensurability. *Research in Science Education, 43*(2), 755–780.
Gericke, N. M., Hagberg, M., Santos, V. C., Joaquim, L. M., & El-Hani, C.N. (2012). Conceptual variation or incoherence? Textbook discourse on genes in six countries. *Science & Education*, doi: 10.1007/s11191-012-9499-8
Gericke, N. M. & Wahlberg, S. (2013). Clusters of concepts in molecular genetics: a study of Swedish upper secondary science students' understanding. *Journal of Biological Education, 47*(2), 73–83.
Guilfoile, P. (1997). Wrinkled peas and white-eyed fruit flies: the molecular basis of two classical genetic traits. *The American Biology Teacher, 59*(2), 92–95.
Halliday, M. A. K. (1993). Towards a language-based theory of learning. *Linguistics and Education, 5*(2), 93–116.
Halliday, M. A. K., & Martin, J. R. (1993). *Writing science: literacy and discursive power*. Bristol: Falmer.
Halliday, M. A. K., & Matthiessen, C. M. I. M. (2004). *An introduction to functional grammar*. London: Arnold.
Hand, B., Alvermann, D., Gee, J., Guzzetti, B., Norris, S., Phillips, L., et al. (2003). Message from the "Island Group": what is literacy in science literacy? *Journal of Research in Science Teaching, 40*(7), 607–615.
Knippels, M.C.P.J. (2002). *Coping with the abstract and complex nature of genetics in biology education—the yo-yo learning and teaching strategy.* Dissertation, CD-β Press, Utrecht.
Lemke, J. L. (1990). *Talking science: language, learning, and values*. Norwood: Ablex.
Lewis, J., Leach, J., & Wood-Robinson, C. (2000). All in the genes? Young people's understanding of the nature of genes. *Journal of Biological Education, 34*(2), 74–79.
Lewis, J., & Kattmann, U. (2004). Traits, genes, particles and information: re-visiting students' understandings of genetics. Research report. *International Journal of Science Education, 26*(2), 195–206.
Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science, 67*, 1–25.
Marbach-Ad, G. (2001). Attempting to break the code in student comprehension of genetic concepts. *Journal of Biological Education, 35*(4), 183–189.
Marbach-Ad, G., & Stavy, R. (2000). Students' cellular and molecular explanations of genetic phenomena. *Journal of Biological Education, 34*(4), 200–205.
Mortimer, E. F., & Scott, P. (2003). *Meaning making in secondary science classrooms*. Buckingham: Open University Press.
Nygård-Larsson, P. (2008). Hur systematisk är systematiken? Några semiotiska resursers betydelse för skapandet av explicita tematiska mönster i lärartext och lärobok. In P. Juvonen (Ed.), *Language and learning. Papers from the ASLA Symposium in Stockholm, 7–8 November* (pp. 71–83). Uppsala: Svenska föreningen för tillämpad språkvetenskap.

Ogborn, J., Kress, G., Martins, I., & Mc Gillicuddy, K. (1996). *Explaining science in the classroom*. Buckingham: Open University Press.

Oyoo, S.O. (2011). Language in science classrooms: an analysis of physics teachers' use of and beliefs about language. *Research in Science Education*. doi:10.1007/s11165-011-9228-3

Robson, C. (2002). *Real world research: a resource for social scientists and practitioner-researchers*. Oxford: Blackwell.

Schönborn, K. J., & Bögeholtz, S. (2009). Knowledge transfer in biology and translation across external representations: experts' views and challenges for learning. *International Journal of Science and Mathematics Education, 7*(5), 931–955.

Swedish National Agency for Education. (2006). *Curriculum for the compulsory school system, the pre-school class and the leisure-time centre* (94th ed.). Stockholm: Swedish National Agency for Education.

Swedish National Agency for Education. (2011). *Curriculum for the compulsory school system, the pre-school class and the leisure-time centre 2011*. Stockholm: Swedish National Agency for Education.

Tsui, C.-Y., & Treagust, D. F. (2007). Understanding genetics: analysis of secondary students' conceptual status. *Journal of Research in Science Teaching, 44*(2), 205–235.

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

Venville, G. J., & Donovan, J. (2005). Searching for clarity to teach the complexity of the gene concept. *Teaching Science, 51*(3), 20–24.

Venville, G. J., Gribble, S. J., & Donovan, J. (2005). An exploration of young children's understandings of genetics concepts from ontological and epistemological perspectives. *Science Education, 89*(4), 614–633.

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

Venville, G. J., & Treagust, D. F. (2002). Teaching about the gene in the genetic information age. *Australian Science Teachers Journal, 48*(2), 20–24.

Verhoeff, R. P., Boerwinkel, D. J., & Waarlo, A. J. (2009). Genomics in school. *EMBO Reports, 10*(2), 120–124.

Wellington, J., & Osborne, J. (2001). *Language and literacy in science education*. Buckingham: Open University Press.

Wilson, J. M. (1999). Using words about thinking: content analyses of chemistry teachers' classroom talk. *International Journal of Science Education, 21*(10), 1067–1084.