

# Making authentic science accessible—the benefits and challenges of integrating bioinformatics into a high-school science curriculum

Yossy Machluf, Hadas Gelbart, Shifra Ben-Dor and Anat Yarden

Corresponding author. Yossy Machluf, The Department of Science Teaching, Weizmann Institute of Science, 234 Herzl St. Rehovot, P.O. box 26, 7610001, Israel. Tel.: +972-8-9342273; Fax: +972-8-9342681; E-mail: Yossy.machluf@gmail.com

## Abstract

Despite the central place held by bioinformatics in modern life sciences and related areas, it has only recently been integrated to a limited extent into high-school teaching and learning programs. Here we describe the assessment of a learning environment entitled 'Bioinformatics in the Service of Biotechnology'. Students' learning outcomes and attitudes toward the bioinformatics learning environment were measured by analyzing their answers to questions embedded within the activities, questionnaires, interviews and observations. Students' difficulties and knowledge acquisition were characterized based on four categories: the required domain-specific knowledge (declarative, procedural, strategic or situational), the scientific field that each question stems from (biology, bioinformatics or their combination), the associated cognitive-process dimension (remember, understand, apply, analyze, evaluate, create) and the type of question (open-ended or multiple choice). Analysis of students' cognitive outcomes revealed learning gains in bioinformatics and related scientific fields, as well as appropriation of the bioinformatics approach as part of the students' scientific 'toolbox'. For students, questions stemming from the 'old world' biology field and requiring declarative or strategic knowledge were harder to deal with. This stands in contrast to their teachers' prediction. Analysis of students' affective outcomes revealed positive attitudes toward bioinformatics and the learning environment, as well as their perception of the teacher's role. Insights from this analysis yielded implications and recommendations for curriculum design, classroom enactment, teacher education and research. For example, we recommend teaching bioinformatics in an integrative and comprehensive manner, through an inquiry process, and linking it to the wider science curriculum.

**Key words:** bioinformatics education; secondary school; domain-specific knowledge; Bloom's taxonomy; assessment

## Introduction

Bioinformatics, an emerging interdisciplinary field, applies principles of computer sciences and information technologies to make the vast, diverse and complex life sciences data more

understandable and useful, and to help realize its full potential [1]. Bioinformatics education can be broadly defined as the teaching and learning of the use of computer databases to gather, store, organize and index biomedical data, and of

**Yossy Machluf** is a Senior Intern in the Department of Science Teaching, Weizmann Institute of Science. He is a researcher, science educator, curriculum developer and teachers' trainer in the fields of bioinformatics and biotechnology.

**Hadas Gelbart** is an Associate Director of International research at the National Authority for Measurement and Evaluation in Education (RAMA).

**Shifra Ben-Dor** is a Senior Staff Scientist in the Bioinformatics Unit, Weizmann Institute of Science. In addition to research, she has extensive experience teaching at the graduate school level.

**Anat Yarden** is an Associate Professor, Head of the Department and head of the Life Sciences Group in the Department of Science Teaching, Weizmann Institute of Science.

**Submitted:** 28 September 2015; **Received (in revised form):** 19 November 2015. **Accepted:** 11 December 2015

© The Author 2016. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)

specialized tools to view, interpret, analyze and integrate data to solve biological problems [2–4]. Of the various branches of bioinformatics, namely ‘research, development, or application of computational tools and approaches for expanding the use of biological, medical, behavioral or health data’ [1], each of which implies different activities and requires a different set of skills and knowledge, the later one (application of tools) is predominantly, not to say exclusively, represented in school science curricula. Here, it is referred to as the ‘bioinformatics approach’. Bioinformatics, which naturally suits computer-based scientific-inquiry activities, is emblematic of the top recommendations for undergraduate life sciences education [5–9]. It provides students with the opportunity to engage in large, scientifically authentic data sets and tools that reflect the nature of modern research in the life sciences. It also allows students to solve complex and real-world problems through integration of scientific knowledge, theories and technologies, while enabling them to develop computational and technological competencies as well as quantitative and modeling skills.

The field of bioinformatics is integral to modern scientific research, and it should become as influential for science education [10], at both the tertiary and secondary educational levels. This should be of interest and concern to both the scientific community, which is eager for bioinformatics-literate individuals equipped with 21st-century skills and knowledge that will pursue careers in the life sciences, and to the educational community, which focuses on preparing the next generation of informed citizens. However, there is a growing disparity between the rapidly evolving world of research and the entrenched culture of science education, which is a barrier to realizing the full potential of publicly accessible electronic data [11]. This is reflected in the tremendous accumulation of bioinformatics tools and databases to empower scientific research, compared with a relatively minor increase in the number of educational resources in bioinformatics [12, 13]. While most bioinformatics educational resources and courses were designed for the undergraduate level (for example [14]), bioinformatics educational reform is sporadic at the secondary level [15, 16]. In addition, better incorporation of bioinformatics into states’ science standards is required [17].

The learning landscape and needs, of both bioinformatics continuing education programs [18] and high-school resources, is changing. The challenge of bringing the complex and contemporary science of bioinformatics to the high-school classroom has only begun to be addressed in the last decade. It was first identified by individual educators who developed bioinformatics teaching materials, which were integrated into their lessons [19, 20]. Today, diverse bioinformatics-based online resources, such as educational tools [21], educational games [22], outreach programs [23–25], collaborations with researchers [26, 27], lessons and teaching modules ([28–32, 33]; e.g.) are being offered to high-school students. Dedicated bioinformatics training courses are available for secondary-school science teachers [34]. Diverse activities, including workshops, courses, meetings and even a task force were established within the Global Organisation for Bioinformatics Learning, Education and Training, GOBLET [24, 35], to collect and share materials and examples of best practices for introducing bioinformatics to high-school teachers and to the public at large. Theoretical and practical aspects of bioinformatics integration into high-school biology classes have been addressed in a joint effort through educational activities in international societies and conferences [36, 37], as well as in scientific journals [31, 38–41]. Recently, rules were set for the

development [42] and teaching [43–45] of bioinformatics training courses, which can also be applied at the high-school level. Nevertheless, only a fraction of high-school students are exposed to these materials, in part owing to the limited integration of bioinformatics into science curricula [17], and the major challenges involved [16, 46–48].

Research in bioinformatics education is often labeled as being in its infancy. Consistent with this idea, recent studies identified the need to clearly define the bioinformatics curriculum and the content that should be taught together with suitable pedagogical approaches, to appropriately orchestrate the desired learning outcomes, to identify the evidence of the learning, to pinpoint the pedagogical approaches of delivery, and to assess the impact on students [16]. Assessment of cognitive, affective and psychomotor learning targets was previously reported (reviewed in [15]). However, recent survey studies have raised concerns on both the assessment and actual impact of most bioinformatics programs: not only a minority of bioinformatics education research papers reported on assessment, but also the focus of assessment and the quality of evidence provided have revealed weaknesses, as assessment in many of these papers focused on curricular change and students’ perceptions and attitudes, rather than on their learning gains. Moreover, instruments that mostly do not provide any reliability or validity evidence were used in many studies [15, 16]. Nevertheless, bioinformatics education in high school revealed a positive effect on students’ knowledge and confidence in the bioinformatics domain [33], expansion of students’ understanding of certain genetic concepts [32, 49] and increased motivation and interest in science, technology, engineering and mathematics careers [48]. On the other hand, students’ difficulties were less investigated. These complementary aspects of evidence-based cognitive (learning gains and comprehension difficulties) and affective (attitudes and perceptions) outcomes are crucial to both assessing the impact of bioinformatics modules and curricula, and to drawing recommendations for effective modes of instruction and for integration of bioinformatics into science curricula.

To make steps toward this end, we aim in this study to assess the cognitive and affective outcomes of bioinformatics education in high school. We previously reported on the introduction of a bioinformatics learning environment, entitled ‘Bioinformatics in the Service of Biotechnology’ into the Israeli national biotechnology curriculum [47]. We also proposed a framework for characterizing bioinformatics teaching units, based on three criteria [47]: the type of domain-specific knowledge required to answer each question (declarative, procedural, situational, strategic; following [50]), the scientific field that each question stems from (biology, bioinformatics or their combination) and the associated cognitive-process dimension (remember, understand, apply, analyze, evaluate, create; following [51]). The feasibility of this framework was confirmed by two independent analyses: (i) analysis of an assessment tool for bioinformatics learning, which revealed that it represents characteristics of authentic modern scientific research that is connected to the ‘traditional’ curriculum [52]; and (ii) analysis of the questions embedded in two selected scientifically authentic bioinformatics activities (hereafter, shortly referred to as ‘activities’) from the bioinformatics learning environment, which demonstrated an agreement between the design of the activities and their pedagogical goals [47]. In this study, the same framework is used to assess not the learning materials (the bioinformatics learning environment or the assessment tool) but the impact of using these materials on students’ cognitive outcomes, in terms of both students’ successes and

challenges. In addition, students' attitudes toward learning bioinformatics and the learning environment were probed. The specific research questions for this study were:

#### *Cognitive outcomes*

- Are the bioinformatics activities suitable to high-school students' cognitive abilities?
- What are the difficulties encountered by students? Which types of questions were difficult to deal with?
- If and to what extent do the bioinformatics activities promote acquisition of knowledge, as well as appropriation of the bioinformatics approach?

#### *Affective outcomes*

- What are students' attitudes toward the bioinformatics learning environment and the bioinformatics lessons?
- How do students perceive the process of learning bioinformatics?

To enable better understanding of the context of this study, key aspects related to the longitudinal processes of the bioinformatics learning environment development, its integration into the science curriculum in Israel and the implementation process will first be described.

## Context

### Bioinformatics in upper secondary schools—rationale and objectives

In Israel, approximately 1600 10th-grade students choose to join a 3-year track specializing in biotechnology each year (of approximately 100 000 10th-grade students overall across the country). The curriculum for biotechnology majors includes (i) obligatory topics such as introduction to biotechnology, genetic engineering, molecular biology and biochemistry; (ii) advanced laboratories; and (iii) elective topics such as industrial fermentation processes, immunodiagnosics and immunotherapy, tissue culture, environmental biotechnology, bio-nanotechnology and an inquiry project [53]. The laboratories, designed to reflect modern research, are composed of practical and computerized components. The later component comprises about 12% of the total grade in the final matriculation examination. Traditionally, it involved a 'Computerized research problems' unit, allocated 60 h during the last year of high school (12th grade). Bioinformatics was introduced into the curriculum as an alternative to this unit.

The Israeli curriculum for biotechnology majors shares many topics with the Australian one [54] and the online teaching resources for biotechnology teachers in Australia (<http://archive.industry.gov.au/Biotechnologyonline.gov.au/topitems/resources.html>) and New Zealand (<http://biotechlearn.org.nz/>). Noteworthy, the level of the 12th-grade senior high-school biotechnology curriculum is in a way similar to the advanced placement biology in the United States. Therefore, the following description, analysis and implications are not restricted to the Israeli setting, but are relevant to bioinformatics curriculum designers, science educators and teachers at both high-school and undergraduate levels worldwide.

A steering committee (composed of experts from the fields of bioinformatics, biology, science education, computer sciences, the national supervisor of biotechnology education at the Israeli Ministry of Education and leading teachers) outlined the desired concepts, core ideas and scientific practices to be

acquired through bioinformatics learning. It also set practical objectives for bioinformatics education in accordance with the biotechnology curriculum, target population, time frame and available resources. The main objective was to properly contextualize the bioinformatics content, skills and thinking approaches as part of a broader scientific vision and research toolbox to promote students' scientific mastery of biotechnology. It was essential to develop a deep sense of the nature of scientific investigation and address key actions in bioinformatics-integrated research, such as data retrieval, analysis, visualization and modeling through authentic hands-on and mind-on activities.

### The bioinformatics learning environment

The bioinformatics learning environment (<http://stwww.weizmann.ac.il/g-bio/bioinfo/>) was described in detail elsewhere [47]. Its core component consists of activities that are interrelated and linked to other units of the learning environment and to the pre-existing biotechnology curriculum (Supplementary Appendix S1). The original materials are in Hebrew. To provide the bioinformatics community (including teachers and students) with an opportunity to experience the learning environment first hand, an English version of a typical activity, including links to animated tutorials (with English subtitles) that present the relevant bioinformatics tools, is provided in Supplementary Appendix S2.

The bioinformatics learning environment makes use of the most widely used bioinformatics tools and databases [55]. The spread of the selected bioinformatics tools (namely Entrez, Blast-N, Blast-P, ClustalW, ORF Finder, Primer3Plus, Prosite and Jmol) throughout the core units of the learning environment is portrayed in Table 1. The considerations for the selection of the bioinformatics tools were scientific, pedagogical and technical [47]. These fundamental tools are widely used by scientists and enable acquisition of central bioinformatics practices [2] such as organization, retrieval and analysis of records in diverse databases, alignment of sequences, designing of primers for polymerase chain reaction (PCR), prediction of open reading frames in a given nucleotide sequence and the deduced amino acid sequences, searching for protein domains, functional sites and families, and visualizing the three-dimensional structure of molecules. They are tightly linked to knowledge and skills in the biotechnology curriculum, such as multiple modeling representations (of a gene or a protein), the transition between DNA-RNA-protein levels, and are directly connected to techniques (PCR). These tools were well suited in both basic and applied research (Supplementary Appendix S1). They are basic, well-supported tools with decent interfaces, which teachers and students can manage in a reasonable manner. All bioinformatics tools are web-based (with the exception of Jmol) free tools. In light of our previous experience [32, 56], the learning environment only makes use of authentic web-based databases, rather than databases locally installed in the educational site. The challenges arising from the dynamic nature of the rapidly changing databases, and dealing with discrepancies between the findings obtained in the bioinformatics tools in real time compared with these displayed in the learning environments will be described in detail in 'The implementation process' section.

In the learning environment, students are actively engaged in activities in biotechnology, addressing real-world problems and using diverse bioinformatics tools and databases while acquiring and applying modern scientific practices (skills,

Table 1. Spread of the bioinformatics tools throughout the main units of the learning environment

Bioinformatics learning environment units and their components	Bioinformatics tools					Databases	Predominant type of knowledge
	Entrez Blast	Nuc. Blast	Protein Blast	ClustalW Primer3Plus	ORF-finder		
Activities 'in-depth' unit	+						
Single-step activities			+				P, D, Si
Extended lifespan						+	
Structure-function relationship in antifreeze proteins							
Multi-step activities	+			+			P, D, Si, St
Mitochondrial DNA and the mystery of the royal family							
Identifying alleles of hemoglobin conferring resistance to malaria	+			+			
Screening for novel genes involved in antibiotic biosynthesis		+	+				
Exploring the genotype-phenotype relations in cystic fibrosis					+		
Searching for a competitive inhibitor of the anthrax toxin	+++	+++		+++	From + to +++		St, Si, D, P
'Integrated' activities						+++	
Characterizing the sequence and 3D structure of green fluorescent protein				+++	+++	+++	
Introduction unit							
Biology							D
Biotechnology							D, Si
Bioinformatics							D
Evolution							D
Virtual 'Bioinformatics toolbox'	+++	+++		+++	+++	+++	P, D, Si, St
Educational games							
Data integration							D, St, P
Sequence alignment	Concepts	Concepts	Concepts	Concepts	Concepts	+	D
Dilemmas	From concepts to applications of a specific tool in a given scenario (depending on whether the game is played before experiencing activities in the learning environment or following it)						D
Overview—terms and concepts	+++	+++	+++	+++	+++	+++	D, P, Si

'+' indicates learning the basis of using the bioinformatics tool, mainly focusing on procedures and concepts (knowing 'what' and 'how').  
 '++++' indicates a more advanced learning focusing on broader aspects of integrating bioinformatics tools into research and their contribution (knowing 'where' and 'when' and 'why').  
 The types of domain-specific knowledge are displayed according to the level of predominance, and indicated by their initiatives: Procedural, Declarative, Situational and Strategic knowledge.

knowledge and modes of thinking). Such practices resemble authentic scientific research as they are carried out in the scientific community, and offer students the opportunity to develop a contextualized and deep understanding of scientific knowledge and of how this knowledge is acquired, evaluated and developed. They can also invoke the reasoning used by scientists and the epistemology underlying authentic scientific research [57]. Through the inquiry process, students are required to coordinate between different types of knowledge from different scientific disciplines, to recall prior content knowledge, to apply technical skills in using bioinformatics tools, to reason scientifically, to make decisions following a strategic plan and to evaluate and justify the scientific process and its steps [47].

### The implementation process

The development process was iterative and intertwined with the implementation process. Initially, two prototype activities were designed and developed in a close collaboration with members of the steering committee. Then a two-stage study was conducted during the spring of 2010 and spring 2011, enrolling biotechnology students and teachers from one and four different schools, respectively. After feedback was obtained, a set of recommendations and guidelines were drawn and incorporated into the design and development process of the bioinformatics learning environment. During the first years of implementation, a continual bidirectional process of directing and supporting the teachers, as well as updating and receiving feedback of their experiences in practice, was conducted through a forum, phone and e-mail contacts, as well as visits to schools and class observations. This process led to further important modifications in the learning environment and reshaping of the training workshops.

Diverse teachers' professional development (TPD) programs were established, aiming at recruiting teachers, training them and promoting leadership among the community of teachers. These included (i) six nationwide continuing-training workshops (56 h) to provide teachers with the relevant theoretical knowledge and technical skills in bioinformatics; (ii) a workshop for developing teaching materials (28 h); (iii) a 1-year-long TPD program (280 h) for outstanding senior teachers [52], which ran for 4 years. It was aimed at building the teachers' personal vision as reform enhancers as well as promoting their bioinformatics mastery, while they were involved in the processes of

design and development, testing and distribution. Four of them participated in the second international bioinformatics high-school teachers' training course [34]; (iv) to increase teachers' interest in bioinformatics education, two symposia, on proteomics and on personalized medicine and nutrition, were organized. Within 4 years, almost all biotechnology teachers in Israel took part, at least once, in a nationwide continuing-training workshop, many of them also participated in the symposia, and elite teachers joined TPD programs for leading teachers.

The number of biotechnology teachers who chose to teach bioinformatics is steadily increasing (Figure 1A), and consequently so are the schools that select this topic and the number of students who are assessed in the matriculation examination in bioinformatics (Figure 1B). The numbers more than doubled each year, reflecting teachers' own initiative, internal motivation and choice to adopt and teach bioinformatics. More teachers are choosing to teach bioinformatics in as early as the 11th grade, solely or in parallel to teaching it in the 12th grade. This is owing to both the teachers' increased confidence in their ability and the availability of teaching materials that are suitable for 11th graders (such as single-step 'in-depth' activities and educational games). Note that some of these materials were initially developed by teachers in the TPD programs, then adopted, edited and enriched by us, and are now an integral part of the bioinformatics learning environment.

Owing to the dynamic nature of the rapidly changing bioinformatics tools and databases, during the first 3 years of implementation, the developers often modified the learning environment to keep it relevant, updated and in congruence with the outputs obtained while using the bioinformatics tools in real time. This was crucial to the teachers' confidence, and it required a tremendous effort. During this period, most teachers realized that many students noticed the discrepancies but were less disturbed by them, managed to properly answer the questions and continued the bioinformatics activity. Today, teachers are only notified of major modifications in the interfaces of tools or the obtained findings and are instructed on how to deal with them. The teachers use the changes in bioinformatics tools and databases as an educational opportunity rather than a pitfall—to enable students to realize the nature of the tools and databases, the accumulation of information, the improvement in tools interface (operational friendly interface, more options) and performance (accuracy, speed).

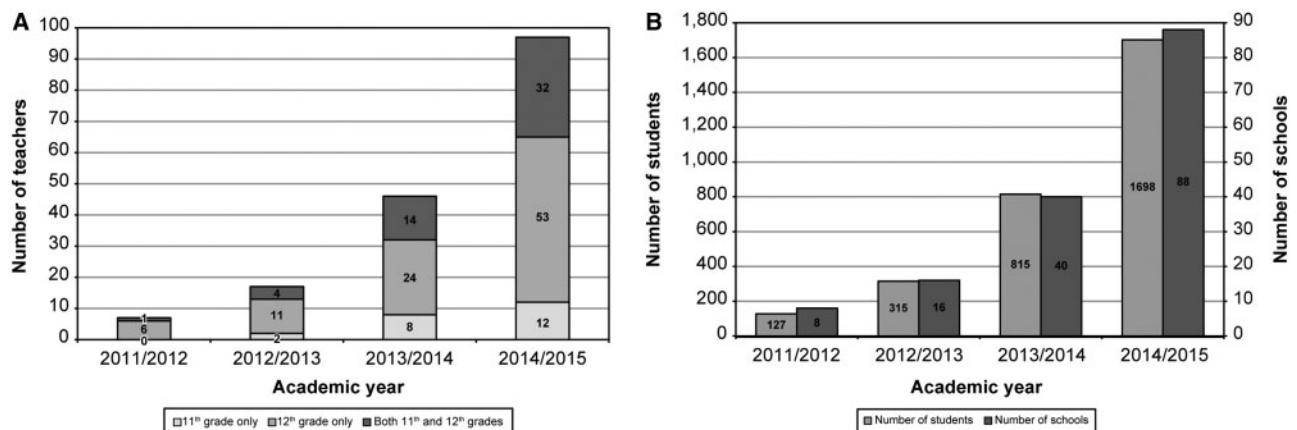


Figure 1. The implementation process from the teachers' and students' perspectives. (A) The number of teachers who taught bioinformatics in 11th grade only (light gray), 12th grade only (intermediate gray) or both grades (dark gray), during four academic years. (B) The number of students who were assessed in the matriculation examination (left, light gray) and the schools that selected this elective topic (right, dark gray), during four academic years.

The community of bioinformatics teachers is growing in terms of both numbers and activities. Altogether, these diverse activities and their products served not only to train the teachers, but also to empower them, to expand the repertoire of resources, to link bioinformatics to the biotechnology curriculum, to ensure that the materials meet their needs (pedagogically, cognitively, scientifically, technologically) and adjusted to their students (in terms of interests, language, design, rationale and pedagogy), to set realistic and practical expectations from both the students and teachers that match the scope of the curriculum and to develop teachers' sense of ownership of the bioinformatics learning environment. It also allowed us to step out gradually, while leading teachers took our place and supported other teachers. In parallel, an important shift in the status of bioinformatics in the science curriculum occurred: bioinformatics was introduced as an elective topic into the curriculum of 11th and 12th graders during the 2011–12 academic year [47], and became a required topic from the 2014–15 academic year. It is important to note that providing evidence on the impact of the bioinformatics learning materials on students' cognitive and affective outcomes (see research questions above), played a key role in this process, as well as on drawing recommendations for teaching bioinformatics in high-school settings.

## Methods

### Population

The study population consisted of 11th-grade high-school biotechnology majors ( $n = 44$ , 16–17 years of age, 21 males and 23 females) from four different schools who were introduced to one of two 'in-depth' activities of the bioinformatics learning environment (<http://stwww.weizmann.ac.il/g-bio/bioinfo/>). These students had no prior engagement with either the field of bioinformatics or the topics of the activities. The students were instructed by their teachers (all teachers had participated in a 1-year TPD program) in a 5-h-long bioinformatics session. The questions of one activity and their analysis following the framework appear in [Supplementary Appendix S3](#).

### Data source

Students' answers to questions embedded in the activities were automatically documented and archived in a predesigned database of the bioinformatics learning environment. The answers were evaluated and scored independently by two science educators, and discussed until 100% agreement was reached.

Students were also asked to respond to questionnaires (described in [Supplementary Appendix S4](#)), before and after exposure to the bioinformatics learning environment (pre- and post-activity questionnaire, respectively). The questionnaires were composed of three parts: (i) students were asked to define, schematically draw and explain biological terms in genetics, such as gene, open-reading frame, coding sequence, exon and promoter. These terms are central to the 11th grade curriculum, but not the focus of the activities (Part A in [Supplementary Appendix S4](#)); (ii) students encountered a research problem that differed from those appearing in their textbooks or in the bioinformatics learning environment. Students were asked to design a study using any of the methodologies and techniques with which they are familiar, justify it and make predictions (Part B in [Supplementary Appendix S4](#)). It should be noted that use of the bioinformatics approach and tools was not obligatory for solving the research problem. These two parts, which were common to both the pre- and post-activity questionnaires, were

used to assess students' prior knowledge, their newly acquired knowledge and their ability to apply this knowledge in a novel context; (iii) students were asked to respond to 16 Likert-type items (1–5 scale), aiming at elucidating their attitudes toward bioinformatics and the bioinformatics learning environment. This part was exclusive to the post-activity questionnaire (Part C in [Supplementary Appendix S4](#)).

The later affective component was also assessed by 10 focus-group interviews of two to three randomly selected pairs from each class and subsequent two whole-class interviews. These were used to enable students to freely express their views and perceptions of their bioinformatics learning experience (examples of questions from the semi-structured interviews are provided in [Supplementary Appendix S5](#)).

The whole-day experience was video- and audio-taped and teaching artifacts (such as lesson plans and presentations) were collected and analyzed.

Teachers' attitudes and views toward the issues that were raised by the students were probed by diverse communication methods, such as post-activity interviews and questionnaires (with each of the four teachers who participated in this study), informal discussions during TPD programs (a few dozen teachers) and e-mails. Analysis of these sources is only provided in light of students' attitudes and views, and is not the main scope of this manuscript.

### Data analysis

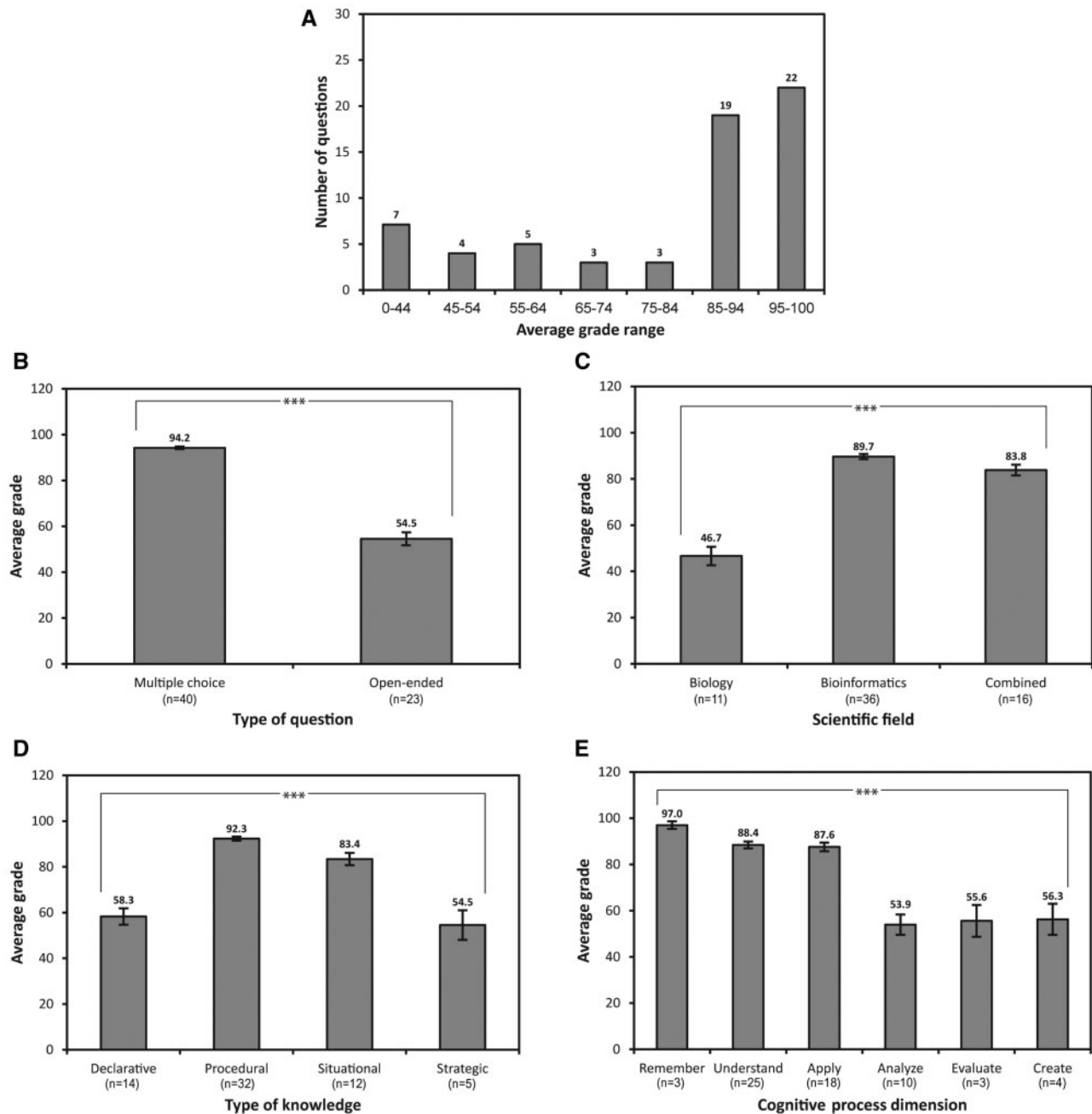
Learning cognitive outcomes, as evidenced by the students' answers, were assessed with regard to question characteristics, following the abovementioned criteria, using a nonparametric (Wilcoxon test) and parametric (analysis of variance, ANOVA) statistical tests. The paired *t*-test was used to compare students' answers between pre- and post-activity questionnaires. Learning affective outcomes were measured by mean, standard error and distribution of answers to Likert-type items. Video and audio recordings of the enactment and interviews were transcribed, and qualitatively analyzed using a 'bottom-up' approach.

## Results and discussion

### Cognitive outcomes

To discern whether the bioinformatics activities are suitable to the high-school biotechnology students' cognitive abilities, students' answers to the questions embedded in the two selected activities were analyzed. The analysis revealed that the educational demands of the bioinformatics learning environment are in line with students' abilities, as the average grade for 65% of the questions (41 of 63) was within the range of 85–100 ([Figure 2A](#)). Nevertheless, 25% of the questions (16 of 63) were challenging, with an average grade below 65 ([Figure 2A](#)).

We subsequently characterized students' performance through the activities, using the proposed framework for classification of questions [47]. It also enabled us to discover the difficulties students encountered, namely the types of questions that were difficult to deal with. This analysis revealed higher average achievements in multiple-choice questions ([Figure 2B](#)), in questions that stem from the bioinformatics field, solely or in combination with the biology field ([Figure 2C](#)), in questions that require the use of procedural or situational knowledge ([Figure 2D](#)) and in questions that are associated with lower



**Figure 2.** Assessment of students' performance during the bioinformatics activities. An analysis of the answers obtained from all students in four classes to questions embedded in two bioinformatics activities within the bioinformatics learning environment. (A) General assessment of learning performance in each question—distribution of average grade for each question. The average grade of students' answers to questions was classified according to (B) the type of question, (C) the scientific field, (D) the required type of knowledge or (E) the associated cognitive-process dimension. In brackets, n represents the number of questions in each category. Mean  $\pm$  standard error are presented. Statistics: A Kruskal-Wallis one-way ANOVA was applied to compare the achievements of students in the different categories. Significance implies that the achievement for at least one of the categories is different from the other. \*\*\* $p < 0.001$ .

cognitive-process dimensions such as 'Remember', 'Understand' or 'Apply' (Figure 2E). On the other hand, open-ended questions (Figure 2B), questions that stem from the biological field (Figure 2C), questions that require the use of declarative or strategic knowledge (Figure 2D) and questions that are associated with higher cognitive-process dimensions such as 'Analyze', 'Evaluate' or 'Create' (Figure 2E) were all significantly more challenging, as evidenced by lower average achievements. While some of these findings were expected, others were surprising and required deeper and broader discussion.

One possible explanation for these findings relies on the nature of the questions. We previously reported on clusters of question characteristics (categories of the classification framework). For example, providing answers to most multiple-choice questions necessitated the use of procedural knowledge, which stems mainly from the bioinformatics field and is associated with the cognitive process 'Understand'. Furthermore, providing answers to most open-ended questions necessitated the use of declarative knowledge, which mainly stems from the biological field and is associated with the cognitive process 'Analyze' [47].

While students performed well in all question categories of the former cluster, they performed poorly in all question categories of the latter cluster (Figure 2B–E). All questions necessitating strategic knowledge were previously reported as open-ended and mainly associated with the cognitive process ‘Create’—these were the most challenging questions. Combinatorial analysis among two (or more) categories was only partially possible owing to a limited number of questions in some rubrics. Nevertheless, this analysis revealed a higher average grade for multiple-choice questions compared with the open-ended questions in all categories of scientific field, domain-specific knowledge or the associated cognitive-process dimension. The differences in students’ performance based on scientific field, namely lower performance on questions that stems from the field of biology compared with those that stem from either bioinformatics or a combination of both fields, were evident and significant for open-ended questions, but not for multiple-choice questions. Similarly, the differences in students’ performance in the question category of domain-specific knowledge (namely lower performance on questions requiring the use of declarative or strategic knowledge, compared with those requiring the use of either procedural or situational knowledge), as well as with regard to questions category of the associated cognitive-process dimension (namely lower performance on questions associated with high-order thinking compared with those associated with low-order thinking), were only evident and significant for open-ended questions, but not for multiple-choice questions. These findings may be a result of insufficient range of grades for the multiple-choice questions, which were closed to the score limitation at the top of a scale, indicating a possible ceiling effect within this type of question. Analysis of students’ performance stratified by both domain-specific knowledge and scientific field further supported the notions that (i) questions that stem from the field of biology are harder to deal with, compared with those that stem from either bioinformatics or a combination of both fields, regardless of the required domain-specific knowledge; (ii) questions that require the use of declarative or strategic knowledge are harder to deal with, compared with those requiring the use of either procedural or situational knowledge, almost completely regardless of the scientific field they stem from.

Bioinformatics is a complex field, necessitating broad procedural skills to use tools and databases and coupled with factual knowledge and strategic thinking. The topic of bioinformatics and the coordination of the related knowledge and skills are not abundant or typical in school tasks. Both teachers, during the TPD programs, and students, as evident in the interviews, used the terms ‘new world’ and ‘old world’ to distinguish between two kinds of demands in the activities. In their perspective, ‘new world’ refers to tasks and questions that stem from the bioinformatics field (solely or in combination with the field of biology) and require mainly the use of procedural or situational knowledge. In contrast, ‘old world’ refers to tasks and questions that stem from the biology field and require mainly the use of declarative knowledge. In this perspective, the coordination of different types of knowledge, which may stem from both the biology and bioinformatics fields, can be viewed as an integration of both ‘worlds’. We decided to adopt this authentic perspective also for the purpose of discussing the findings. Our findings revealed that students dealt properly with questions addressing the knowledge and skills of the ‘new world’, whereas those of the ‘old world’ or combined ‘worlds’ were harder to deal with. This finding was surprising. It suggests students’ difficulties in integrating declarative knowledge,

mostly in biology, with procedural knowledge, mostly in bioinformatics, and strategic knowledge of the research plan with both knowledge types, along with the application of technical skills of using bioinformatics tools and higher-order thinking, namely evaluation and interpretation of findings to provide meaning. This integration can also be framed in light of the roles scientific models play in the learning process. Based on biological knowledge and experiences with scientific models, students make their own mental models of scientific phenomena to display their understanding. These mental models can serve as explanatory tools or be used to formulate theories and make predictions [58]. The use of bioinformatics tools may guide students toward being capable of using their mental models to explain and predict natural phenomena. A correct explanation or prediction, which is leaning on both the use of bioinformatics tools and databases and in light of the biological model, represents such integration.

Yet, this may also reflect characteristics of the teaching materials and/or lesson design. Remarkably, during the teacher professional development programs, many teachers, who are experts in biotechnology or biology education but novices in bioinformatics education, expressed their concerns about teaching bioinformatics and specifically its procedures. Furthermore, almost all teachers predicted that students would struggle with questions stemming from the ‘new world’, mainly questions that require the use of procedural knowledge or stemming from the field of bioinformatics. This prediction is in contrast with the findings on student performance (and perception, see below). Consequently, many of the teachers devoted most of their efforts to teaching technical and procedural aspects of using bioinformatics tools, which was revealed through analysis of their teaching materials and time allocation during the observed lessons. During the recent TPD programs, the discrepancy between students’ performance and teachers’ predication of students’ difficulties is at the heart of a confrontation session, in which teachers are asked to provide alternative plans for bioinformatics lessons (and materials), derived and justified from this and other studies, to support student needs. Analysis of teachers’ artifacts and interviews with teachers (at the end of the TPD programs and during the following years) uncovered that most teachers perceive bioinformatics education as a three-step challenge: first, and most importantly, building the procedural (mainly) bioinformatics foundations of the ‘new world’, then connecting it to the ‘old world’ to draw a deeper biological/biotechnological meaning and finally integrating the bioinformatics tools and databases as part of strategic plan of biological research (data not shown). The order of these steps reflects importance and priority. The first, second and last steps may be viewed as a short-, intermediate- and long-term educational goals of the curriculum, respectively. Typically, teachers who teach bioinformatics for the first time tend to focus mainly on the short-term goals and to a lesser extent on the intermediate-term goals, while almost neglecting the long-term goals. The more teachers are experienced and confident in their knowledge and skills in the bioinformatics era, and self-perceive themselves as advanced bioinformatics trainees, the connections and integration between the two ‘worlds’ (steps 2 and 3) are done earlier in the curriculum, in wider and deeper contexts, depending of course on the level of students. Acting in a community of teachers and developers, where institutional and cultural barriers were removed, enabled the sharing of experiences and insights, and therefore promoted the change in focus on learning goals from short-term goals to intermediate- and long-term ones.

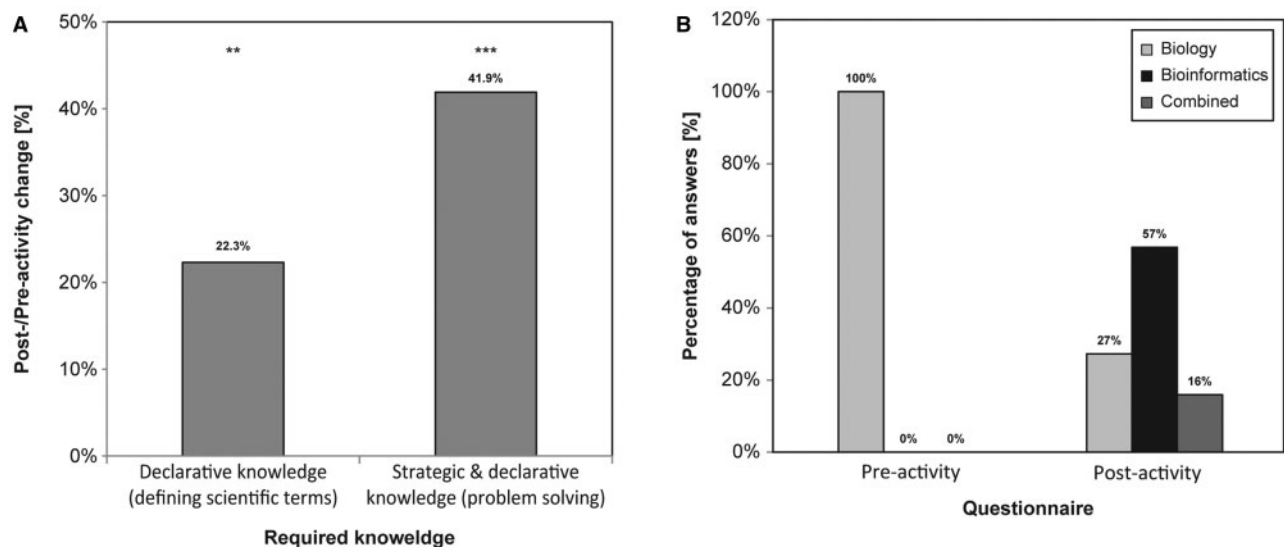


These findings are in line with previous ones where we reported that engagement in a simulation of authentic scientific research [32] requires continuous integration of scientific facts and bioinformatics procedures while reasoning scientifically and making decisions [59]. The ability to coordinate and use different knowledge dimensions, and mainly strategic knowledge, are at the heart of performing authentic scientific research. Such coordination is not typically found in regular school tasks [57, 60], and requires support and guidance from the teacher [59, 61]. Based on our current and previous findings regarding student difficulties and the observed discrepancy between the students' performance and the teachers' predictions, we recommend teaching bioinformatics in an integrative and comprehensive manner, and linking it to the wider science curriculum.

A case study by Wefer and Anderson [62] highlighted the pronounced individual differences in the way students contextualize and process domain-specific knowledge and skills associated with a bioinformatics unit. In their view, bioinformatics mastery requires integration of concrete factual information with procedural skills, analytical skills and higher-order knowledge in a well-rationalized and coherent way. Accordingly, recommendations were made for teachers 'to design lessons that help students comprehensively integrate these cognitive dimensions in a holistic fashion' [62]. Another aspect of integration was also demonstrated recently when a few high-school students who experienced a computational biology course persisted in thinking of separate and irrelevant 'computer' and 'biology' units, rather than seeing the connections between them [30]. Similar findings have been observed at the undergraduate level, where students were found to differ in their ability to understand and articulate the biological significance of the bioinformatics results, as well as to understand the differences between listing a fact, carrying out a procedure or running a computer program, and drawing inferences and insights from them [63]. Buttigieg [64] claimed that to give the audience some perspective on 'why and how' bioinformatics research is carried out, one should develop the factual knowledge of biology alongside the procedural knowledge of computer science and

mathematics in a cohesive manner. Likewise, providing students with the technical understanding of the databases (the 'how') and the conceptual knowledge about the overall annotation process (the 'why') were identified as critical ingredients for the success of genome annotation in the bioinformatics classroom [65]. This body of evidence demonstrates the importance of integration of different types of knowledge and across different disciplines for drawing inferences and scientific insights, as well as for knowledge organization in a well-rationalized and coherent manner among both school [30, 62] and academia [63, 64] students.

In addition to the direct evidence from students' answers to questions embedded in the bioinformatics activities from the learning environment, we wished to assess if and to what extent the bioinformatics activities promote acquisition of knowledge, as well as appropriation of the bioinformatics approach. These can be considered as aspects of content mastery, in both general (scientific) and specific (bioinformatics) domains. To this end, pre- and post-activity questionnaires were distributed and analyzed (Supplementary Appendix S4). First, students were asked to define, draw and explain biological terms in genetics. To answer this task, one may use mainly declarative knowledge. A 22% increase in the average grade was observed (Figure 3A, left column), indicating that students had acquired biological content knowledge in the course of the bioinformatics lessons. Then, students were asked to solve a problem by designing a study using any of the methodologies and techniques with which they are familiar (biological technique and/or bioinformatics tools), justify it and make predictions. To answer this task, one may coordinate different types of knowledge, mainly strategic and declarative knowledge, and to a lesser extent procedural and situational knowledge. A 40% increase was observed in the correctness of students' answers in solving a research problem in a novel inquiry setting after the bioinformatics lesson (Figure 3A, right column). Besides the correctness of the study design proposed by the students, we also analyzed the scientific field on which the design was based. Following the 5-h-long bioinformatics experience, most of the students'



**Figure 3.** Assessment of students' performance through the questionnaires. An analysis of students' answers in pre- and post-activity questionnaires. (A) Knowledge acquisition—change of mean score between post- and pre-activity questionnaires aimed at assessing declarative knowledge (defining scientific terms, left column) or both strategic and declarative knowledge (problem-solving, right column). (B) Appropriation of the bioinformatics approach—distribution of the scientific approaches applied by students in the proposed study design while solving problems. Statistics: A paired t-test was applied to compare the achievements of students in post- versus pre-activity questionnaires. \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

study designs were based on a bioinformatics approach, solely (57%) or in combination (16%) with a biological approach (Figure 3B). Considering that the inquiry settings were novel, and could be answered by previous biological knowledge, these findings suggest that students appropriate the bioinformatics approach as part of their ‘scientific research toolbox’, namely perceive bioinformatics knowledge as useful and use bioinformatics principles, techniques and tools in various research contexts. It can signify adoption of authentic scientific research practices. These are two complementary facets of knowledge acquisition.

We [32, 49] and others [66, 67] have found that learning bioinformatics through activities or research simulations can complement and expand the understanding of biological content, such as of the genetics domain, as well as to expose students to new information, novel technologies and different pedagogy. It can also promote recognition of the research practices, namely help students reflect on the complexity of the cognitive processes and epistemology that guide authentic genetic research. Interestingly, at both the high-school and university levels, two types of learners—research-oriented and task-oriented—were identified on the basis of the differences in the ways they seized opportunities to recognize the research practices, which in turn influenced their learning outcomes [49, 63]. In this study, a few students, mostly—but not exclusively—the task-oriented ones, claimed that the excess of text, explained print-screens and animated tutorials led them to read only the questions, while skipping the explanations, especially when their answers were correct. One pair of research-oriented students looked for the differences between various amino acids, to pinpoint a possible biochemical explanation for the effect of a mutation within the beta-globin protein underlying resistance to malaria—an explanation that they delivered to the whole class at the end of the activity. It is possible that instruction by the teacher and delivery modes as well as the culture of computer-based lessons in the classroom may affect the proportion of these types of learners, and subsequently their learning gains.

### Affective outcomes

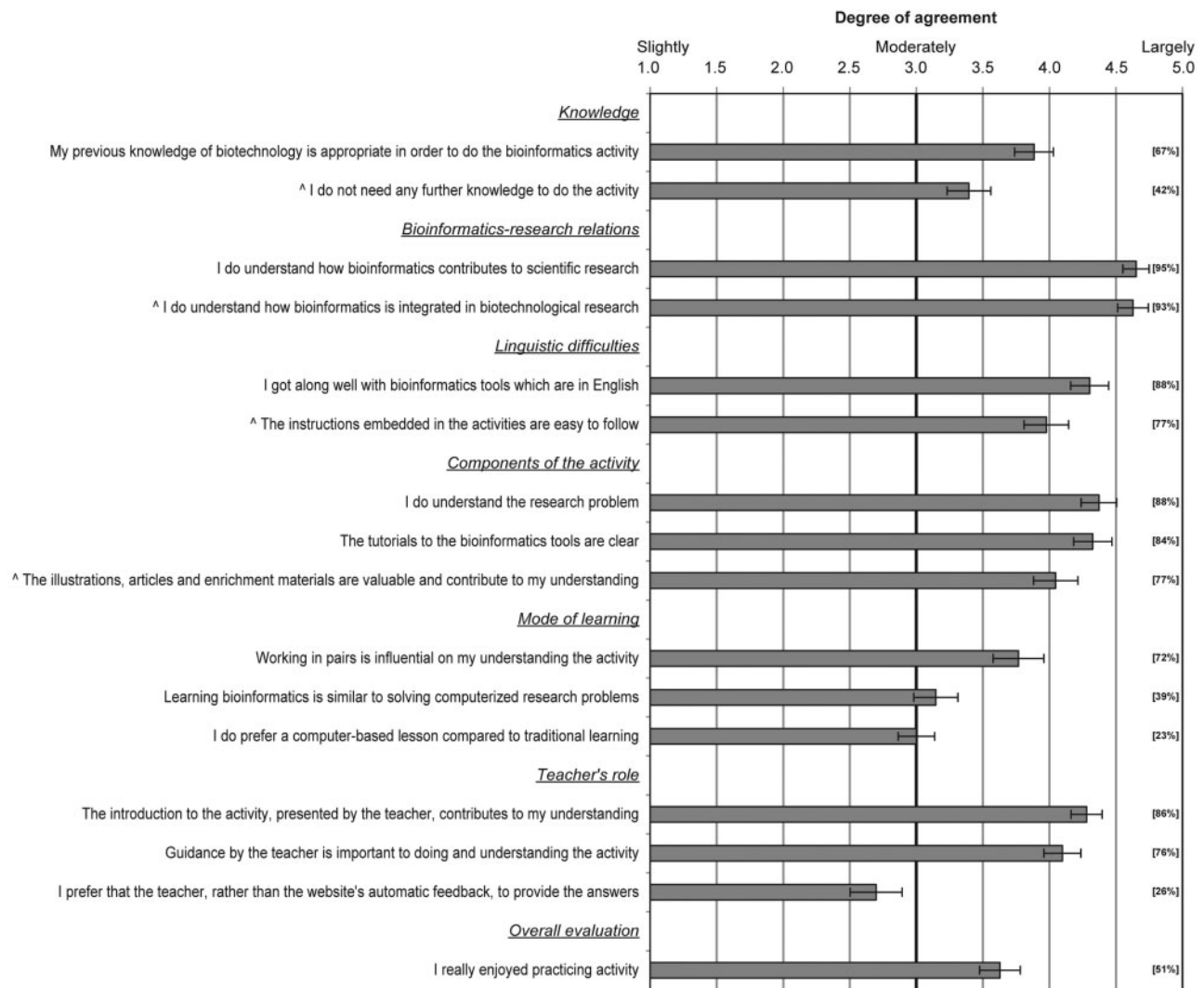
To capture students’ attitudes toward the bioinformatics learning environment and bioinformatics lessons, we used a post-activity questionnaire (Supplementary Appendix S4) and interviews (Supplementary Appendix S5). Students were positive toward the bioinformatics learning environment (Figure 4). Half of them agreed that it was an enjoyable experience (average 3.6), almost all (95%) claimed that they understood the inquiry process and the relationship between bioinformatics and scientific research (average 4.6) and most (~80%) found the diverse units of the bioinformatics learning environment valuable (average > 4).

Most importantly, while students appreciated the diverse scaffolding means, the automated and immediate feedback to their answers to multiple-choice questions and working in pairs, they were divided in their preferences for computer-based lessons compared with a traditional learning mode. They also highlighted the significance of the teacher’s guidance throughout the activity, from the introduction to the closing of the lesson. Students felt more comfortable when teachers overviewed the research strategy and stages involving bioinformatics tools within this plan, as well as the outcomes of the inquiry. Teachers have been previously shown to play a key role in supporting the learning of bioinformatics [59], as well as the enactment of adapted primary literature [61], and in promoting

students’ use of strategic knowledge, and the coordination between strategic and declarative knowledge, in a manner similar to scientists in the course of performing authentic research. Regarding the questions requiring various types of knowledge, one student claimed that ‘some questions, mainly at the beginning of the activity, were quite silly, sorry for the expression... it was too easy and technical!... But toward the end we faced questions that made us think, it was interesting!’—many other students shared this view. There was a wide agreement among students regarding their wish to reduce the number of words, in both the narrative part of the activity and the explanations, and this was even more pronounced among non-native Hebrew speakers, such as new immigrants and Arabs, as was reported by teachers during training courses.

With regard to students’ perception of the process of learning bioinformatics—it is worth noting that the students liked the subjects of the activities, which were attractive and relevant to their fields of interest, and the use of authentic tools ‘like real scientists’. Yet, some students wished to have a more open inquiry, in the sense of investigating a new research question, with an unknown final answer, and to discover ‘new things for the whole scientific community’ themselves—representing a wish to be knowledge producers rather than consumers. This was a rather individual disposition, as some preferred the structured activity (see also [30]). Teachers, on the other hand, rejected this option owing to lack of experience, knowledge, time and confidence, and felt more comfortable when both the starting and end points of the inquiry were known (personal communication with numerous teachers during TPD programs). These two problem-based approaches, termed ‘free form’ and ‘workbook’, respectively, were previously discussed in the context of incorporating bioinformatics into existing undergraduate courses [68]. Also at the university level, a flexible and open-ended course design can be stressful and time-consuming for the instructors, and may not work well for all students [69].

In the interviews, students described their day-long bioinformatics experience by the adjectives ‘interesting’, ‘fun’, ‘fascinating’ and ‘enriching’, regardless of the type of bioinformatics activity or class settings. These refer to both the subject of the bioinformatics activities, the place of students and teachers in the learning process and the delivery mode, which were ‘different than what we usually experience at school’. Students from different classes pointed out that in contrast to other school computational inquiry activities, which ‘are based on and repeat what we learned in class’, the bioinformatics activities promote the acquisition ‘of additional knowledge beyond what we already learned in class, and can explain and clarify other stuff we leaned... you need to apply previous knowledge, and acquire new knowledge by yourself’. Students concluded by recommending, ‘I wish I could experience more such [bioinformatics] activities... [you should] bring it [the bioinformatics learning environment] to more schools!’. One student perceived the experience in a broader perspective of career development in science: ‘It [the bioinformatics experience] did something to me! if I wish to understand the world... and to figure out if I really want it [STEM career] for myself in the future or not... that the whole point of these biotechnology programs—to attract students [to this field] that will later in the future become scientists. It did something to me! A sense and feeling of.. yes! It is interesting for me!’. Altogether, these lines of evidence, although not based on systematic and thorough analysis, may suggest that our learning environment raised students’ motivation to learn bioinformatics, and can promote the interest in developing science careers. In light of the central place of high



**Figure 4.** Assessment of students' attitudes. Summary of students' responses to post-activity questionnaire Likert items (scale: 1–5) aimed at assessing their attitudes toward learning bioinformatics through the bioinformatics learning environment. Bars and error bars represent mean  $\pm$  standard error, while numbers in brackets represent percentage of students whose answer was 4 or 5. ^ indicates items that were formulated in a negative voice in the questionnaire.

schools in influencing students' impressions of science and impacting their future career decisions [70] and the potential of bioinformatics modules and science teachers to promote interest in and awareness of STEM careers [48, 71, 72], it appears that bioinformatics education in secondary school holds great promise toward answering the need for young people who will pursue STEM careers in the academy or industry.

## Conclusions and future directions

In summary, we describe the implementation and assessment of a bioinformatics learning environment aimed at introducing bioinformatics into a high-school biotechnology majors' curriculum.

Within 4 years of implementation, the number of students, teachers and schools who chose to learn and teach bioinformatics was doubled every year, and bioinformatics became a required topic in the science curriculum of biotechnology majors. Teachers are agents of change [73] in the educational system, and central to all stages of integrating bioinformatics into the high-school science curriculum. The diverse TPD

programs were aimed at recruiting teachers and establishing both a community of qualified teachers and a leadership of entrepreneurs. Proper training enabled teachers to successfully adapt sustainable and scalable innovations in the bioinformatics curriculum. *Sustainability* refers to curriculum innovations that will be used in the classroom after the developers have left, whereas *scalability* refers to successful enactment by teachers who have had no contact with the developers [74].

Our assessment revealed the suitability of the bioinformatics activities to students' abilities, students' difficulties as well as the learning gains in terms of knowledge acquisition (in both bioinformatics and biology/biotechnology) and appropriation of the bioinformatics approach as part of the students' scientific 'toolbox'. In addition to the cognitive outcomes, it revealed students' positive attitudes toward the bioinformatics learning environment, and their perception of the learning process. The framework for characterizing the questions embedded in the bioinformatics activities uncovered the key difficulties encountered by students. Notably, these difficulties were different from what the teachers had predicted, thereby serving to draw specific recommendations for instruction and scaffolding, for both

teachers and teacher educators. For example, teachers usually emphasize the procedural aspects of using bioinformatics tools, whereas students find these aspects easy, technical and even boring. However, the considerations involved in selecting a specific bioinformatics tool, based on the research goal, strategy and available data, are less clear to students. Therefore, suggestions for, and modeling of, teacher intervention to clarify these points are offered during the TPD programs, as part of the teachers' role in ensuring that students understand the scope and objectives of the activity, while providing appropriate guidance. It would be of high interest to assess students' performance and attitudes (at the whole-class and individual-student levels) in different classes, and check how these are related to different teaching strategies and styles, as well as teachers' beliefs about bioinformatics education. Nowadays, the learning environment is shifting to a more flexible, free and open-source platform termed Moodle (modular object-oriented dynamic learning environment [75]). With its comprehensive, customizable and source-learning management features, each teacher can adapt the course to his/her educational approach and classroom characteristics, and create a personalized and dynamic online course.

Our bioinformatics inquiry-based activities address authentic biological/biotechnological questions, which are relevant to the school science curriculum, using authentic research plans and tools. The learning environment is aimed at providing students with a toolbox of technical skills and thinking abilities in bioinformatics (like [33]), rather than the computation and its use in biology (unlike [30]). Although the algorithms are not investigated, their principles of action are discussed and even modeled (such as calculating scores in BLAST, using paper and pencil). In the future, multidisciplinary and collaborative projects might be developed for students who are learning biology/bioinformatics and computer sciences, similar to the 'in-concert teaching approach' [76], or programs for gifted students who excel in both sciences. Mixing students from different backgrounds in a group to solve a complex task can create stimulating learning opportunities [77]. Meanwhile, we hope teachers will adopt an open-ended inquiry approach, structured or nonstructured [68], which will focus on students' own questions and interests. Such individualization may increase students' involvement and independence, which in turn should enable a high-level inquiry project, namely open inquiry [78], increase students' ownership of both the learning process and its products and empower students.

Based on our and others cumulative experience, we propose key rules that have emerged as being central to integrating bioinformatics into high-school science curricula in general and for effective learning in particular (Supplementary Appendix S6). These rules are at the interface of several related processes, such as policy making, defining standards and objectives, designing and developing learning materials, actual teaching and pedagogical delivery modes, assessment and training teachers. The rules related to 'Actual teaching and pedagogical delivery modes' and 'Assessment' were derived from the findings of this study. Integration is a key theme, but harbors diverse facets, which are independent yet related to each other. At the curricular level, as appears in the steering committee declarations and teachers' views [52], the bioinformatics curriculum should be linked to the preexisting science curricula. This allows a more interdisciplinary approach for learning, where fundamental concepts, ideas and practices from different disciplines, such as bioinformatics, biology and statistics, can be connected. Moreover, different types of domain-specific knowledge (procedural, declarative, situational and strategic knowledge)

should be taught in a comprehensive and coherent manner, and ideally should be linked to computational (if possible) and technological competencies, as well as quantitative and modeling skills (in line with the standards). Such integrative learning is a major challenge for students and necessitates the guidance of the teachers, as indicated by students' achievements and perception. The teachers should be aware of the individual needs and rhythms of students, therefore address multiple learning styles, continuously monitor their performance and actions in real time, act as facilitators who guide inquiry that enables students to construct scientific knowledge by themselves and propose leading questions to assess students' understanding and probe their difficulties. In addition, teachers should recognize the junctions where instruction is beneficial, such as introducing the research plan, evaluating the contribution and limitations of each bioinformatics tool, etc. as suggested by the students, and at the same time to allow independent inquiry of the students. This is reminiscent of the recommendation for teachers to explain and review the purpose of inquiry lab activities in pre-lab and post-lab sessions, respectively [79]. The teachers praised the importance of prior awareness to the explicit and clear requirements and curricular demands, and specify the assessment tool as a model that shapes and directs the design of lessons as well as expectations, purposes and behaviors of students. Assessment of the impact of the materials and delivery modes on students' knowledge, skill acquisition and attitudes, in both the national and local school settings, may allow an improvement in the materials or the way they are used, as well as promote broader integration of bioinformatics into schools, in different science subjects (using Jmol for chemistry education, in either science or interdisciplinary projects, etc.). We attempt to provide a broader and comparative view of the importance and relevance of these rules not only to high-school level students, but also to undergraduate and graduate students, as well as a clinical audience. While some rules are as important for teaching bioinformatics in both secondary schools and universities, it seems that the former share much more similarity to teaching bioinformatics to a clinical audience. A key pre-requisite for successful implementation and integration of bioinformatics into the high-school science curriculum is collaboration and communication between policy makers, stakeholders, inspectors, developers, scientists, science educators, teachers' trainers, teachers and students—along all stages of design, development, enactment, assessment and refinement. Such collaboration is also crucial in bridging the gap between teachers' perceived goals for teaching bioinformatics and students perceptions of this unit (and one may add developers' goals and policy makers' goals, as part of the gap between the intended and implemented curriculum). Such mismatches in expectations often occur in school laboratory activities [79, 80]. A literature survey suggested that science teachers across many countries hold expectations (perception of learning goals) and consequently exhibit behaviors (practical work and assessment) that do not use effectively and make the most of the unique environment of the school laboratory [79]. Bioinformatics is part of the modern laboratory, in academy and classroom settings, and bioinformatics teachers are part of the science teaching community. However, in the context of bioinformatics education in high schools in Israel, we could identify indications that teachers' perceptions, expectations and behavior might promote meaningful learning consistent with contemporary standards. This was evident, only in some aspects, already in our day-long pilot study, and more prominently among teachers with experience in teaching

bioinformatics (data not shown). Supporting evidence came from the students who only experienced a day-long bioinformatics activity but already appropriate the bioinformatics approach and gained knowledge and skills. Whereas previous studies suggested that many students perceived manipulating equipment and measuring as goals of laboratory activities [79], here biotechnology students perceived manipulating ideas as central goal of bioinformatics activities—‘What is required here [in the bioinformatics activity] is thinking, a lot of thinking . . . and in a broader sense’. We envision that these shared lessons, insights and recommendations (Supplementary Appendix S6) might be valuable for fostering the incorporation of other related emerging scientific fields in high-school (and even higher) education, as a first step toward an integrated curricula in science, mathematics and technology.

Among the limitations of our study, one may stress the small sample sizes of students and teachers, as well as their experience in the activities and embedded questions. Nevertheless, it lays the groundwork for current research on a nationwide level. Preliminary results, from a cumulative 4 years of experience including students from over 100 classes, strengthen the validity of the results presented here.

#### Key Points

- Endeavors to incorporate bioinformatics into high-school science classrooms should be accompanied by educational research at the interface of bioinformatics curriculum design, teaching processes, as well as students’ cognitive and affective outcomes.
- A framework for characterizing the type of knowledge, the scientific field that each question stems from and the associated cognitive-process dimension required to answer questions embedded in bioinformatics learning environments is used to characterized students’ difficulties and knowledge acquisition.
- Analysis of students’ cognitive outcomes revealed learning gains in bioinformatics and related scientific fields, as well as appropriation of the bioinformatics approach as part of the students’ scientific ‘toolbox’.
- For students, questions stemming from the biology field and requiring declarative or strategic knowledge were harder to deal with. This stands in contrast to teachers’ prediction!
- Analysis of students’ affective outcomes revealed positive attitudes toward bioinformatics and the learning environment, as well as their perception of the teacher’s role.

#### Supplementary data

Supplementary data are available online at <http://bib.oxfordjournals.org/>.

#### Funding

The development of the bioinformatics learning environment and training courses were funded by the Ministry of Education and The Israeli Science Teaching Center Grant. Teachers’ professional development program, within the framework of the Rothschild-Weizmann Program for Excellence in Science Education, was generously supported by Caesaria Foundation.

#### Acknowledgments

We wish to thank Orna Dahan, Amir Mitchell and Carmit Avidan-Shpalter for their collaborative efforts in designing and developing the bioinformatics learning environment, and Eilat Avraham, the national supervisor for biotechnology teaching at the Israeli Ministry of Education, who had the vision and enterprise to lead the way in bringing bioinformatics to high-school classrooms. We would like to thank Yetty Varon for her statistical advice, Ziv Ariely for graphic design, Avi Tal for desktop publishing, and Avi Hofstein for a helpful discussion.

#### References

1. National Institutes of Health [NIH]. NIH working definition of bioinformatics and computational biology. <http://www.bisti.nih.gov> (17 July 2000 date last accessed).
2. National Center for Biotechnology Information [NCBI]. A science primer: bioinformatics. <http://www.ncbi.nlm.nih.gov/About/primer/bioinformatics.html> (15 June 2011 date last accessed).
3. Counsell D. A review of bioinformatics education in the UK. *Brief Bioinform* 2003;4:7–21.
4. Koch I, Fuellen G. A review of bioinformatics education in Germany. *Brief Bioinform* 2008;9:232–42.
5. National Research Council [NRC]. *BIO2010: Transforming Undergraduate Education for Future Research Biologists*. Washington, DC: The National Academies Press, 2003.
6. National Research Council [NRC]. *A New Biology for the 21st Century: Ensuring the United States Leads the Coming Biology Revolution*. Washington, DC: The National Academies Press, 2009.
7. American Association for the Advancement of Science [AAAS]. Vision and change in undergraduate biology education. A call to action. <http://visionandchange.org/files/2011/03/Revised-Vision-and-Change-Final-Report.pdf>, 2011.
8. Woodin T, Carter VC, Fletcher L. Vision and change in biology undergraduate education, A call to action—initial responses. *CBE Life Sci Educ* 2010;9:71–3.
9. Labov JB, Reid AH, Yamamoto KR. Integrated biology and undergraduate science education: a new biology education for the twenty-first century?. *CBE Life Sci Educ* 2010;9:10–16.
10. Pevzner P, Shamir R. Computing has changed biology—biology education must catch up. *Science* 2009;325:541–2.
11. Donovan S. Big data: teaching must evolve to keep up with advances. *Nature* 2008;455:461.
12. Brazas MD, Yamada JT, Ouellette BF. Evolution in bioinformatic resources: 2009 update on the Bioinformatics Links Directory. *Nucleic Acids Res* 2009;37:W3–5.
13. Brazas MD, Yamada JT, Ouellette BF. Providing web servers and training in Bioinformatics: 2010 update on the Bioinformatics Links Directory. *Nucleic Acids Res* 2010;38:W3–6.
14. Chapman BS, Christmann JL, Thatcher EF. Bioinformatics for undergraduates: steps toward a quantitative bioscience curriculum. *Biochem Mol Biol Educ* 2006;34:180–6.
15. Campbell CE, Nehm RH. A critical analysis of assessment quality in genomics and bioinformatics education research. *CBE Life Sci Educ* 2013;12:530–41.
16. Magana AJ, Taleyarkhan M, Alvarado DR, et al. A survey of scholarly literature describing the field of bioinformatics education and bioinformatics educational research. *CBE Life Sci Educ* 2014;13:607–23.

17. Wefer SH, Sheppard K. Bioinformatics in high school biology curricula: a study of state science standards. *CBE Life Sci Educ* 2008;7:155–62.
18. Brazas MD, Ouellette BFF. Navigating the changing learning landscape: perspective from bioinformatics.ca. *Brief Bioinform* 2013;14:556–2.
19. Gelbart H, Yarden A. Learning genetics through an authentic research simulation using a web-based learning environment in bioinformatics. In: *The Vth Conference of European Researchers in Didactics of Biology (ERIDOB)*. Patras University Press, Patras, Greece, 2004, 193–204.
20. Wefer SH. Name that gene. *Am Biol Teach* 2003;65:610–13.
21. Toby I, Pope A. Bioinformatics tools available for K-12 students to engage in research. *Aviat Space Environ Med* 2014;85:484–5.
22. Schneider MV, Jimenez RC. Teaching the fundamentals of biological data integration using classroom games. *PLoS Comput Biol* 2012;8:e1002789.
23. McQueen J, Wright JJ, Fox JA. Design and implementation of a genomics field trip program aimed at secondary school students. *PLoS Comput Biol* 2012;8:e1002636.
24. Atwood TK, Bongcam-Rudloff E, Brazas ME, et al. GOBLET: the global organisation for bioinformatics learning, education and training. *PLoS Comput Biol* 2015;11:e1004143.
25. Dutta-Moscato J, Gopalakrishnan V, Lotze MT, et al. Creating a pipeline of talent for informatics: STEM initiative for high school students in computer science, biology, and biomedical informatics. *J Pathol Inform* 2014;5:12.
26. Pavesi G, Siccardi A, Viale G, et al. Hedgehogs, humans and high-school science. *EMBO Rep* 2008;9:208–11.
27. Munn M, O'Neil Skinner P, Conn L, et al. The involvement of genome researchers in high school science education. *Genome Res* 1999;9:597–607.
28. Campbell AM, Zanta CA, Heyer LJ, et al. DNA microarray wet lab simulation brings genomics into the high school curriculum. *CBE Life Sci Educ* 2006;5:332–9.
29. Jungck JR, Donovan SS, Weisstein AE, et al. Bioinformatics education dissemination with an evolutionary problem solving perspective. *Brief Bioinform* 2010;11:570–81.
30. Gallagher SR, Coon W, Donley K, et al. A first attempt to bring computational biology into advanced high school biology classrooms. *PLoS Comput Biol* 2011;7:e1002244.
31. Lewitter F, Bourne PE. Teaching bioinformatics at the secondary school level. *PLoS Comput Biol* 2011;7:e1002242.
32. Gelbart H, Yarden A. Learning genetics through an authentic research simulation in bioinformatics. *J Biol Educ* 2006;40:107–12.
33. Marques I, Almeida P, Alves R, et al. Bioinformatics projects supporting life-sciences learning in high schools. *PLoS Comput Biol* 2014;10:e1003404.
34. Wood L, Gebhardt P. Bioinformatics goes to school—new avenues for teaching contemporary biology. *PLoS Comput Biol* 2013;9:e1003089.
35. Corpas M, Jimenez RC, Bongcam-Rudloff E, et al. The GOBLET training portal: a global repository of bioinformatics training materials, courses and trainers. *Bioinformatics* 2015;31:140–2.
36. International Society for Computational Biology [ISCB]. 18th Annual international conference on Intelligent Systems for Molecular Biology [ISMB]: Tutorial on Teaching bioinformatics in high school biology courses. <http://www.iscb.org/archive/conferences/iscb/ismb2010-program/ismb2010-tutorials.html#AM1>. <http://www.iscb.org/archive/conferences/iscb/ismb2010-program/ismb2010-tutorials.html#AM1> (3 May 2015 date last accessed).
37. Global Organisation for Bioinformatics Learning ETG. GOBLET AGM 2014 & STAO Conference. <http://mygoblet.org/about-us/goblet-events/goblet-agm-2014-stao-conference>. <http://mygoblet.org/about-us/goblet-events/goblet-agm-2014-stao-conference> (3 May 2015 date last accessed).
38. Lewitter F. Welcome to PLoS Computational Biology “Education”. *PLoS Comput Biol* 2006;2:e1.
39. Lewitter F. Moving education forward. *PLoS Comput Biol* 2007;3:e19.
40. Neil Sarkar I. Editorial: bioinformatics education in the 21st century. *Brief Bioinform* 2010;11:535–6.
41. Schneider MV, Jungck JR. Editorial: international, interdisciplinary, multi-level bioinformatics training and education. *Brief Bioinform* 2013;14:527.
42. Via A, De Las Rivas J, Attwood TK, et al. Ten simple rules for developing a short bioinformatics training course. *PLoS Comput Biol* 2011;7:e1002245.
43. Form D, Lewitter F. Ten simple rules for teaching bioinformatics at the high school level. *PLoS Comput Biol* 2011;7:e1002243.
44. Costa MJ, Galembeck E, Marson GA, et al. A quick guide for computer-assisted instruction in computational biology and bioinformatics. *PLoS Comput Biol* 2008;4:e1000035.
45. Brazas MD, Lewitter F, Schneider MV, et al. A quick guide to genomics and bioinformatics training for clinical and public audiences. *PLoS Comput Biol* 2014;10:e1003510.
46. Cummings MP, Temple GG. Broader incorporation of bioinformatics in education: opportunities and challenges. *Brief Bioinform* 2010;11:537–43.
47. Machluf Y, Yarden A. Integrating bioinformatics into senior high school: Design principles and implications. *Brief Bioinform* 2013;14:648–60.
48. Kovarik DN, Patterson DG, Cohen C, et al. Bioinformatics education in high school: Implications for promoting science, technology, engineering, and mathematics careers. *CBE Life Sci Educ* 2013;12:441–59.
49. Gelbart H, Brill G, Yarden A. The impact of a web-based research simulation in bioinformatics on students' understanding of genetics. *Res Sci Educ* 2009;39:725–51.
50. de Jong T, Ferguson-Hessler MGM. Types and qualities of knowledge. *Educ Psychol* 1996;31:105–13.
51. Krathwohl DR. A revision of Bloom's taxonomy: an overview. *Theory Pract* 2002;41:212–18.
52. Machluf Y, Gelbart H, Yarden A. High-school teachers' appropriation of an innovative curriculum in bioinformatics. In: *The 9th Conference of European Researchers in Didactics of Biology (ERIDOB)*. Berlin, Germany, 2012, 213–25.
53. Israeli Ministry of Education. *Syllabus of biotechnological studies (10th-12th Grade)*. Jerusalem, Israel: State of Israel Ministry of Education Curriculum Center (In Hebrew), 2005.
54. Kidman G. Attitudes and interests towards biotechnology: the mismatch between students and teachers. *Eurasia J Math Sci Technol Educ* 2009;5:135–43.
55. Furge LL, Stevens-Truss R, Moore DB, et al. Vertical and horizontal integration of bioinformatics education. *Biochem Mol Biol Educ* 2009;37:26–36.
56. Gelbart H, Yarden A. *Bioinformatics—Deciphering the Secrets of the Genome*. Rehovot, Israel: The Amos de-Shalit Center for Science Teaching, 2001. <http://stwww.weizmann.ac.il/bioinformatics>
57. Chinn CA, Malhotra BA. Epistemologically authentic inquiry in schools: a theoretical framework for evaluating inquiry tasks. *Sci Educ* 2002;86:175–218.

58. Treagust DF, Chittleborough G, Mamiala TL. Students' understanding of the role of scientific models in learning science. *Int J Sci Educ* 2002;**24**:357–68.
59. Gelbart H, Yarden A. Supporting learning of high-school genetics using authentic research practices: the teacher's role. *J Biol Educ* 2011;**45**:129–35.
60. Yarden A. Guest editorial—reading scientific texts: adapting primary literature for promoting scientific literacy. *Res Sci Educ* 2009;**39**:307–11.
61. Falk H, Yarden A. “Here the scientists explain what I said.” Coordination practices elicited during the enactment of the Results and Discussion sections of adapted primary literature. *Res Sci Educ* 2009;**39**:349–83.
62. Wefer SH, Anderson RO. Identification of students' content mastery and cognitive and affective percepts of a bioinformatics miniunit: a case study with recommendations for teacher education. *J Sci Teach Educ* 2008;**19**:355–73.
63. Honts JE. Evolving strategies for the incorporation of bioinformatics within the undergraduate cell biology curriculum. *Cell Biol Educ* 2003;**2**:233–47.
64. Buttigieg PL. Perspectives on presentation and pedagogy in aid of bioinformatics education. *Brief Bioinform* 2010;**11**:587–97.
65. Shapiro C, Ayon C, Moberg-Parker J, et al. Strategies for using peer-assisted learning effectively in an undergraduate bioinformatics course. *Biochem Mol Biol Educ* 2013;**41**:24–33.
66. Holtzclaw JD, Eisen A, Whitney EM, et al. Incorporating a new bioinformatics component into genetics at a historically black college: outcomes and lessons. *CBE Life Sci Educ* 2006;**5**:52–64.
67. Grisham W, Schottler NA, Valli-Marill J, et al. Teaching bioinformatics and neuroinformatics by using free web-based tools. *CBE Life Sci Educ* 2010;**9**:98–107.
68. Boyle JA. Bioinformatics in undergraduate education: practical examples. *Biochem Mol Biol Educ* 2004;**32**:236–8.
69. Temple L, Cresawn SG, Monroe JD. Genomics and bioinformatics in undergraduate curricula: contexts for hybrid laboratory/lecture courses for entering and advanced science students. *Biochem Mol Biol Educ* 2010;**38**:23–8.
70. Subotnik RFT, Robert H, Rochelle R, John A. Specialized Public High Schools of Science, Mathematics, and Technology and the STEM Pipeline: what do we know now and what will we know in 5 years?. *Roeper Rev* 2010;**32**:7–16.
71. Cohen C, Patterson DG, Kovarik DN, et al. Fostering STEM career awareness: emerging opportunities for teachers. *Wash State Kappan* 2013;**7**:12–17.
72. Cohen C, Patterson DG, Kovarik DN, et al. The emerging role of science teachers in fostering STEM career awareness. *Wash State Kappan* 2013;**6**.
73. Fullan MG. Why teachers must become change agents. *Educational Leadersh* 1993;**50**:12–17.
74. Fishman BJ, Krajcik J. What does it mean to create sustainable science curriculum innovations? A commentary. *Sci Educ* 2003;**87**:564–73.
75. Dougiamas M, Taylor P. Moodle: using learning communities to create an open source course management system. In: D Lassner, C McNaught (eds). *World Conference on Educational Multimedia, Hypermedia and Telecommunications 2003*. Honolulu, Hawaii: AACE, 2003, 171–8.
76. Goodman AL, Dekhtyar A. Teaching bioinformatics in concert. *PLoS Comput Biol* 2014;**10**:e1003896.
77. Abeln S, Molenaar D, Feenstra KA, et al. Bioinformatics and Systems Biology: bridging the gap between heterogeneous student backgrounds. *Brief Bioinform* 2013;**14**:589–98.
78. Schwab J, Brandwein P. *The Teaching of Science: The Teaching of Science as Inquiry*. Cambridge, MA: Harvard University Press, 1962.
79. Hofstein A, Lunetta VN. The laboratory in science education: foundations for the twenty-first century. *Sci Educ* 2004;**88**:28–54.
80. Lunetta VN, Hofstein A, Clough MP. Learning and teaching in the school science laboratory: an analysis of research, theory, and practice. In: SK Abell, NG Lederman (eds). *Handbook of Research on Science Education*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc., 2007, 393–442.