Integrating bioinformatics into senior high school: design principles and implications

Yossy Machluf and Anat Yarden

Submitted: 5th February 2013; Received (in revised form): 4th April 2013

Abstract
Bioinformatics is an integral part of modern life sciences. It has revolutionized and redefined how research is carried out and has had an enormous impact on biotechnology, medicine, agriculture and related areas. Yet, it is only rarely integrated into high school teaching and learning programs, playing almost no role in preparing the next generation of information-oriented citizens. Here, we describe the design principles of bioinformatics learning environments, including our own, that are aimed at introducing bioinformatics into senior high school curricula through engaging learners in scientifically authentic inquiry activities. We discuss the bioinformatics-related benefits and challenges that high school teachers and students face in the course of the implementation process, in light of previous studies and our own experience. Based on these lessons, we present a new approach for characterizing the questions embedded in bioinformatics teaching and learning units, based on three criteria: the type of domain-specific knowledge required to answer each question (declarative knowledge, procedural knowledge, strategic knowledge, situational knowledge), the scientific approach from which each question stems (biological, bioinformatics, a combination of the two) and the associated cognitive process dimension (remember, understand, apply, analyze, evaluate, create). We demonstrate the feasibility of this approach using a learning environment, which we developed for the high school level, and suggest some of its implications. This review sheds light on unique and critical characteristics related to broader integration of bioinformatics in secondary education, which are also relevant to the undergraduate level, and especially on curriculum design, development of suitable learning environments and teaching and learning processes.

Keywords: bioinformatics education; high school; authenticity; learning environment; domain-specific knowledge; revised Bloom's taxonomy

INTRODUCTION
Dramatic progress in biological understanding, coupled with major advances in experimental techniques, novel approaches and computational analyses, are transforming the sciences of biology, biotechnology and medicine. Yet, these exciting new fields and areas of science are rarely integrated into science classrooms or textbooks. This pattern leaves high school science education lagging behind cutting-edge scientific discoveries, which hold great potential for supporting students’ understanding and eliciting their interest and motivation to learn science.

Biology in the 21st century is expanding from a purely laboratory-based science to an information-aided one [1]. Massive growth in information, due to experimental and technological advances, has led to ‘an absolute requirement for computerized databases to store, organize, and index the data and for specialized tools to view and analyze the data’ [2]. 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 [3]. The ultimate goal of bioinformatics 'is to enable the discovery of new biological insights as well as to create a global perspective from which unifying principles in biology can be discerned' [2]. Bioinformatics is now an integral part of modern biology. It has revolutionized and redefined how research is carried out [2, 4–6] and has had an enormous impact on biotechnology, medicine, agriculture, industry and related areas.

In light of the ‘information revolution’, the lagging, or limited bioinformatics educational units from high school science curricula, is of growing concern to both the scientific community, which is eager for bioinformatics-literate individuals who will pursue careers in the life sciences, and the educational community, which focuses on preparing the next generation of informed citizens.

In this article, we review recent initiatives to integrate the rapidly growing interdisciplinary field of bioinformatics into secondary science classrooms. We discuss insights and lessons from the accumulating experiences in the teaching and learning of bioinformatics in high schools. We present a novel approach, which we suggest as a uniform and comprehensive platform, to guide the design and characterization of bioinformatics learning materials and demonstrate its feasibility using our own bioinformatics learning environment. Implications and recommendations for bioinformatics curriculum design, teacher training and means of improving students’ learning processes are suggested.

**BIOINFORMATICS EDUCATION—AN OVERVIEW**

Despite the tremendous growth in the number of bioinformatics tools and databases to empower scientific research, only minor increase have been seen in the number of educational resources [7, 8]. Donovan [9] claimed that ‘given the growing disparity between the rapidly evolving world of research and an entrenched culture of science education, the future of science depends on our commitment to preparing future scientists to work with “big data”’.

The origin of bioinformatics education lies in self-teaching and apprenticeship-like models, where pioneers in the field taught themselves and each other, relying on personal experience and key articles (e.g. [10]). Today, training programs are being established for bioinformatics services and faculties [11–16]. The need to prepare 21st-century scientists has led to a paradigm shift in biology and bioinformatics education [17–23], striving to mirror today’s research trends and keeping science curricula current. Initially, efforts were invested in developing structured certificate and degree programs to teach bioinformatics at the graduate [24–27] and undergraduate [14, 28–32] levels. However, the challenge of bringing the complex and contemporary science of bioinformatics to the high school classroom is only now being addressed (see further on).

A key question is ‘what are the standards of bioinformatics education at each educational level (from high school to secondary and tertiary levels)?’ The standards, in turn, should be integrated into policy, curriculum, instruction and assessment to support meaningful learning. Standards can be defined in terms of scientific practices, unifying cross-cutting concepts, and discipline-related core ideas [33]. Key themes to foster students’ realization of the real-life contribution of bioinformatics, to promote their understanding and to increase their interest are ‘integration’ and ‘context’ [18, 23, 29]. The term ‘integration’ means that fundamental concepts and ideas (or knowledge) as well as competencies (or practices) of each discipline (biology, computer sciences, mathematics, etc.) should be connected and integrated, rather than presented as separate discipline-specific units. The term ‘context’ means that the concepts, ideas and practices should be taught in relevant scientific contexts, using a problem-based approach [34–36], rather than as a collection of instructions in a recipe book. Of note, although the graduate-level programs are mainly designed to teach the fundamentals of bioinformatics, focusing on sophisticated computation, mathematics and informatics [27, 29], at the high school level, curricula usually use ‘simple’ bioinformatics as a teaching tool and provide students with a toolbox of technical skills and thinking abilities in bioinformatics [37]. Standards and objectives of bioinformatics education should be consistent with the educational framework (university versus high school, formal versus non-formal), target population (age, abilities, background, etc.), time frame (annual topic, short course, etc.) and resources (technical, educational, materials). Nevertheless, in all cases, it should strive 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 minds-on activities.

TEACHING AND LEARNING BIOINFORMATICS IN HIGH SCHOOL

The need to incorporate bioinformatics into secondary school science classes was first identified by individual teachers and educators who integrated bioinformatics into their lessons. It is only in recent years that joint international efforts have been devoted to teaching bioinformatics in secondary schools, in the following forms:

• educational activities in international societies and conferences. For example, during the 2010 meeting of the International Society for Computational Biology Education Committee, a major topic of discussion was how to get bioinformatics integrated into high school biology classes, organize a tutorial aimed at secondary school biology and chemistry teachers interested in learning about bioinformatics and how to include it in their curricula and built a collection of useful resources for secondary school educators and students looking for information on Bioinformatics teaching tools (see http://www.iscb.org/high-school-secondary-school-resources, and reported in [38]).

• an Education section in PLoS computational biology [39] and special issues in other academic journals such as Briefings in Bioinformatics [19], provide practical and theoretical information of educational programs and materials, also for secondary school educators, teachers and inquisitive minds [40].

• and funds for developing bioinformatics learning materials.

Today, a plethora of bioinformatics-based online resources, learning modules and outreach programs (usually at universities, e.g. [41]) are being offered to high school students (reviewed in [38]), and bioinformatics-based lessons and curricular elements have been developed [14, 38, 42–53]. Yet, only a fraction of high school students are exposed to these materials, partially due to its limited integration in scientific curricula. Recently, rules were set for the development [13] and teaching [37, 54] of bioinformatics courses, which can be implied also at the high school level. As construction of knowledge, acquisition of skills and motivation are primarily achieved when students are engaged and pursue a goal [55], it was recommended to develop a series of inquiry-based research activities that are build on each other and are linked to preexisting science curricula. Each activity should be simple, in terms of its objectives, and based on a familiar and relevant biological context. The problem-solving process should address multiple learning styles, enable the modeling of the hidden processes and principles of actions of the bioinformatics tools using paper and pencil. Moreover, it should allow different rates of progression and provide opportunities for independent thinking and individualization to increase students’ involvement. In addition, it should allow students to discover new information and concepts themselves, which in turn should empower students and serve as a basis for a product of the learning process. Students’ performance should be assessed and evaluated. The teacher should ensure that students understand the scope and objectives of the activity while providing appropriate guidance. Teacher should also ensure ahead of time the availability, compatibility, preparedness and relevance of the bioinformatics activities and learning materials, the bioinformatics resources and the computational equipment. Given the availability of diverse web-based publicly accessible research databases and tools, as well as educational resources and recommendations, it is now practical, appropriate and desirable to broadly incorporate bioinformatics into the high school classroom.

Opportunities and challenges

Incorporation of bioinformatics into high school curricula offers great opportunities and major challenges. Among the main benefits is the applicability of bioinformatics to life sciences curricula and the inherent features of bioinformatics tools and databases to promote student learning [38, 56]. Bioinformatics naturally fit computer-based scientific inquiry activities and allow active engagement with big data sets and current technologies in real-world problem-solving contexts while using 21st century skills. It may also enable interactive instruction accompanied by authentic visualizations and modeling, which are in line with active learning and ‘less traditional’ learning styles. Moreover, students today are digital ‘natives’, that is to say they have grown up with computers, internet and technology, and they therefore possess natural proficiency and technological competence [56]. These ‘natives’ have a different approach to
learning, one which is concerned with speed of access, instant gratification and multitasking abilities. Obviously, independent of a possible increase in students’ motivation and interest in life sciences research, students can develop a research toolbox, built up of knowledge skills and thinking approaches that can serve them as informed citizens and in future science studies.

Based on the categorization proposed by Cummings and Temple [56], we map the major challenges for broader incorporation of bioinformatics in high school education into two interrelated categories: (i) infrastructure and curriculum and (ii) the human factor.

**Infrastructure and curriculum**

Teaching bioinformatics inherently requires computational infrastructure, including, for example, computers, internet access, appropriate bandwidth and bioinformatics tools, the availability of which is not trivial. In addition, most high schools do not provide support policies or the qualified and knowledgeable personnel required for installation and maintenance of the hardware and software. In many places, bioinformatics is not integral in the national science curriculum, which implies a lack of standards, of common and recommended learning materials and of assessment tools. Teachers often need to develop their own learning materials or to choose among limited options of learning materials available online, which are mainly designed for the undergraduate level. These are usually static and not modular and cannot be adapted to individual teacher objectives. In addition, they usually provide neither scaffolding for learning and instruction nor feedback and reports on students’ performance. Moreover, the materials are often outdated and contain broken links, the latter due to the dynamic nature of the rapidly changing bioinformatics tools and databases. The authentic bioinformatics resources are also complex and rich in professional scientific terminology, usually have a non-user-friendly interface and are not customizable. These requirements, features and weaknesses pose difficulties for managers, teachers and students.

**The human factor**

Teachers often lack any prior experience in bioinformatics research or its instruction and have only little (if any) background in associated disciplines, such as computation and statistics. Long and continuing professional development programs and short training courses in bioinformatics, which combine knowledge, skills and pedagogy in technology-rich environments, are infrequent and rare. For some teachers who are not native English speakers, language is also a barrier. Therefore, many teachers feel less comfortable with bioinformatics. Teachers who struggle by themselves to cope with biological context, the complex and dynamic bioinformatics tools and databases and interpretations of the findings—also need to support their students, with diverse learning styles and difficulties and with minimal-to-no-teaching assistance. Students are impacted by the same difficulties as teachers, related to learning to use bioinformatics resources in diverse research contexts.

**Design principles of an authentic bioinformatics learning environment**

To successfully implement and broadly incorporate bioinformatics in high school science, we must recognize, tackle and overcome these challenges. In our opinion, an essential preliminary step is to determine a policy that bioinformatics should be integrated as an elective or obligatory topic in the national science standards and curriculum. This would necessitate the definition of practical educational goals, which are consistent with the target population, the timeframe, the available resources and the national standards. In addition, minimal technical and computational requirements should be characterized and provided to schools. Obviously, the rules and recommendations for development of learning materials and teaching bioinformatics, as described before, should be implemented. Support, to both students and teachers, is another key issue. The learning materials should encompass diverse scaffolding means and real-time feedback to support students’ knowledge construction, at different grades following the spiral learning model. Short- and long-term training programs as well as a continuous support for teachers should be offered. Moreover, an assessment tool, which is consistent with the learning processes, should be designed. Above all, considering the challenges and opportunities, one should ensure collaboration and communication between policy makers, inspectors, developers, scientists, science educators and teachers—along all stages of design, development, enactment, evaluation and refinement.

To illustrate means of approaching some of the aforementioned challenges, we describe the rationale, guidelines and design principles underlying the development of a web-based learning environment [52] that is aimed at introducing bioinformatics into a
high school biotechnology majors curriculum (60 h, 12th grade) in Israel (for additional information, see Machluf and Yarden [57]). The curriculum for biotechnology majors includes obligatory topics such as genetic engineering and biochemistry, and the elective topics of immunodiagnostics and immunotherapy, tissue culture, environmental biotechnology, bio-nanotechnology and advanced laboratories, as well as bioinformatics [58]. It has been recently (2011–12 academic year) introduced into the national matriculation examination in Biotechnology in Israel as an innovative online problem-solving type assessment comprising ~12% of the total grade. Noteworthy, the level of the 12th-grade senior high school biology curriculum is in a way similar to the advanced placement biology in the United States. Therefore, we believe that the following description, analysis and implication are relevant to bioinformatics curriculum designers, science educators and teachers at both high school and undergraduate levels.

In the learning environment (Figure 1), both pedagogy and technology were recruited for first-hand active learning processes and educational purposes aimed at engaging students in scientifically authentic inquiry-based activities in biotechnology, approaching real-world problems, using diverse bioinformatics tools and databases, while acquiring and applying modern scientific practices (skills, knowledge and mode of thinking). The learning environment was designed with the canonical perspective on authentic science education in mind (following Buxton [59]), namely, practices that resemble authentic scientific research, as they are carried out by the scientific community. Authentic practices can offer students opportunities to develop a contextualized and deep understanding of scientific knowledge and of how this knowledge is acquired, evaluated and developed [60–62], as well as to invoke the reasoning that scientists use and the epistemology underlying authentic inquiry [63].

The activities in the learning environment were developed based on primary research articles, which were tailored and adapted to high school cognitive level and knowledge, and were selected according to (i) the relevance of the scientific context to students’

---

Figure I: The learning environment’s home page. The various units are presented, as well as the scientifically authentic inquiry activities and the bioinformatics tools.
interests; (ii) a clear biotechnological application; (iii) use of a variety of bioinformatics tools and databases that are suitable for the high school students’ cognitive level; (iv) leading-edge high-impact subjects that are broadly covered in the popular scientific literature and in the public media, indicating scientific importance and accessibility; (v) clear connections to principles and techniques in the biotechnology syllabus; and (vi) representation of diverse organisms and different molecules (DNA, RNA and proteins). The topics of the activities focus on authentic investigations that are aimed at improving humans’ life quality and expectancy: (i) identifying alleles of hemoglobin conferring resistance to malaria; (ii) screening for novel genes involved in antibiotic biosynthesis; (iii) exploring the genotype–phenotype relations in cystic fibrosis; (iv) searching for a competitive inhibitor of the anthrax toxin; (v) characterizing the sequence and 3D structure of green fluorescent protein. More details on the activities, the research objectives and rationale, questions addressed and choice of tools can be found in Supplementary Materials (Supplementary Appendix 1).

In three activities (termed ‘in-depth activities’, i–iii above), the students become familiar with bioinformatics tools and databases for the first time, and therefore emphasis is placed on prompting understanding via hand-in-hand guidance through each step of the inquiry process, mainly focusing on procedures and concepts. The remaining two activities (termed ‘integrated activities’, iv and v above) are based on this previous experience, hence highlighting the research strategy, considerations of selecting bioinformatics tools and their contribution to basic and applied scientific research. Students use most tools once in an ‘in-depth activity’ and once in an ‘integrated activity’ (Table 1). Though activities are modular, the ‘integrated activities’ are connected to and build on the experience acquired in the ‘in-depth activities’.

Each multistep activity begins with a detailed background for the scientific investigation, including images and external links and animations, presenting the rationale, goal, design and 2–4 key tasks of the investigation. Throughout the inquiry process, which is embedded within a large number of questions and assignments, the students experience different scientific practices; they 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.

The selected bioinformatics tools (Entrez, Blast-N, Blast-P, ClustalW, ORF Finder, Primer3Plus, Prosite and Jmol) are basic yet fundamental; namely, they are well-supported tools with decent interfaces, which are not difficult for teachers and students to use; they are widely used by scientists and enable acquisition of central bioinformatics principles and approaches. Moreover, the tools are freely available on the web, and the databases are frequently updated, both reflecting key practices of

<table>
<thead>
<tr>
<th>Bioinformatics tools</th>
<th>Activities</th>
<th>‘In-depth activities’</th>
<th>‘Integrated activities’</th>
<th>‘In-depth activities’</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Identifying alleles of hemoglobin conferring resistance to malaria</td>
<td>Screening for novel genes involved in antibiotic biosynthesis</td>
<td>Exploring the genotype–phenotype relations in cystic fibrosis</td>
<td>Characterizing the sequence and 3D structure of green fluorescent protein</td>
</tr>
<tr>
<td>Entrez</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Blast-N</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Blast-P</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>ClustalW</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Primer3+</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>ORF Finder</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Prosite</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Jmol</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
the scientific community, such as sharing resources, and the dynamic nature of scientific knowledge. The tools are introduced to the learners in a virtual ‘Bioinformatics Toolbox’ that includes detailed interactive tutorials and written texts. Although the algorithms are not demonstrated, their principles of operation are discussed.

An introductory unit lays the groundwork for learning, comprising subject matter knowledge in biology, biotechnology and bioinformatics. It covers basic concepts and core ideas in biology such as the central dogma, from a DNA sequence to protein function, and regulation of gene expression at all levels. The bioinformatics field is presented from historical and research perspectives. Its contribution is illustrated by the paradigm shift in evolutionary research and modes of integration in modern basic and applied biotechnological research. More details on the introductory unit can be found in the Supplementary Materials (Supplementary Appendix 2).

Scaffolding is provided by automatic reflections on students’ answers to multiple-choice questions (offering corrections, explanations and empowerment), textual descriptions and graphical annotations to illuminate bioinformatics tools’ interfaces and result pages and reference to a scientific dictionary. To make the scientific thinking visible (and epistemologically transparent), a summary of each investigation, including the use and contribution of each bioinformatics tool to each task, appears at the beginning and end of each activity.

An assessment unit includes short exercises and matriculation exam-like questions to provide students with a deeper understanding of the bioinformatics approach and to improve their skills in bioinformatics practices. Documentation of curricular requirements of bioinformatics mastery is essential to direct instruction and learning.

Altogether, the bioinformatics learning environment meets the challenges [56] and fulfills the guidelines [13, 37] for bioinformatics education, to enable high school students to acquire a scientific toolbox containing bioinformatics-specific knowledge and skills, as well as general research practices. These are important components of scientific literacy for both informed citizens and future life science researchers.

Teachers’ perspective—training and support

To support the implementation of bioinformatics in the biotechnology syllabus, teachers were recruited as agents of change. To this end, outstanding senior teachers, with either a MSc or a PhD degree in biology, were involved in the design of the learning environment, in the development of learning materials, teachers’ guides, instructional means and assessment tools and in other teachers’ preparation and support. This was achieved by a 1-year long teachers’ professional development program in bioinformatics within the framework of Rothschild-Weizmann Program for Excellence in Science Education, which is now running for the 3rd year. The program is aimed at establishing a community of teachers who desire to be innovative, and collaborate in enhancing the implementation of bioinformatics into high schools [64]. Moreover, training courses (56 h) and learning and teaching materials design workshops (28 h) for teachers are offered on a nationwide scale, together with monitoring, supervision and constant guidance at the personal level. The close collaboration between the teachers and developers and researchers contributed to linking bioinformatics to preexisting biotechnology curricula, to setting realistic and practical expectations from students and teachers that match the syllabus’s scope, to teachers’ sense of ownership of the learning environment and to developing their identity as reform-minded science teachers. These actions have led to the establishment of a community of teachers, developers and researchers who collaborate in adapting the new learning environment and promoting its implementation and distribution.

A teachers’ guide is inherent to the learning environment. Suggestions for instruction, enactment and evaluation are provided for each activity, as well as answers to the questions, a closed forum for teachers, presentations and instructional materials. A teachers’ interface for analysis of students’ performance, as answers to questions that are embedded in the learning environment and stored online in a database, is also integrated into the learning environment. Examples and more details on the teachers’ guide can be found in the Supplementary Materials (Supplementary Appendix 3).

Research and lessons

Endeavors to incorporate bioinformatics into high school science classrooms are only seldom accompanied by educational research at the interface of bioinformatics curriculum design, teaching and learning processes. Bioinformatics is a complex field in nature,
necessitating wide procedural skills to use tools and databases coupled with factual knowledge and strategic thinking.

A seminal case study by Wefer and Anderson [65] highlighted the individual and marked differences in the way students contextualize and process domain-specific knowledge and skills associated with a bioinformatics unit. Their analysis was based on four components: specific knowledge and facts (concrete factual knowledge), higher-order knowledge (principles, theories and generalizations related to bioinformatics that draw holistically from declarative and procedural knowledge), procedural skills (skills needed to use bioinformatics programs) and analytical thinking skills (skills needed to apply factual and procedural knowledge to interpret and apply bioinformatics information in various situations). In this view, bioinformatics mastery requires integration of factual information with procedural knowledge and analytical skills in a well-rationalized and coherent way. Accordingly, recommendations were made for teachers to design lessons in a way that ‘helps students to comprehensively integrate these cognitive dimensions in a holistic fashion’ [65].

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 [49]. Similar findings were 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 inference and insight from them [66]. Buttigieg [67] claimed that to grant the audience 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.

We previously developed a research simulation in genetics and bioinformatics for high school biology majors [51] and used it to characterize the teaching approaches and to examine how learning through the environment influences students’ acquisition of genetics knowledge and their comprehension of scientific practices [68–70]. We found that engagement in an authentic scientific research simulation requires continuous application of facts and procedures while reasoning scientifically and making decisions. Thus, the ability to coordinate and use different knowledge dimensions, mainly conditional knowledge (following [71]), were shown to be at the heart of performing authentic scientific research. Teachers were shown to play a key role in supporting the learning of bioinformatics and in promoting students’ use of conditional knowledge in a manner similar to scientists in the course of performing authentic research [70]. This study was based on the definitions proposed by Alexander and Judy [71] of declarative knowledge (knowing ‘what’—factual knowledge), procedural knowledge (knowing ‘how’—compilation of declarative knowledge into functional units that incorporate domain-specific strategies) and conditional knowledge (knowing ‘when and where’ to access certain facts or use particular procedures). We [68, 69] and others [32, 72] found that learning bioinformatics can complement and enhance understanding of biological content, such as genetics. 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 [66, 69].

In addition to the benefits and challenges involved in the incorporation of a modern interdisciplinary field into the high school classroom, this literature survey brings to light the disparity in theoretical perspectives and terminologies related to knowledge and skills. Clearly, different terms are used to refer to similar types of knowledge, and similar terms are used to refer to different types of knowledge. For example, what Wefer and Anderson [65] referred to as ‘analytical thinking skills’, others refer to as ‘conditional knowledge’ [71] or ‘strategic knowledge’ [73]. Yet strategic knowledge was also considered a special form of procedural knowledge [71]. We need a uniform comprehensive platform to guide the design of bioinformatics curricula, to enable the characterization of learning materials (see further on), to analyze students’ learning styles, outcomes and needs, to portray teaching strategies and means, to draw recommendations on curriculum design, teaching and learning processes and to promote bioinformatics mastery among secondary school students.
TYPES OF DOMAIN-KNOWLEDGE AND COGNITIVE PROCESSES INVOLVED IN BIOINFORMATICS COMPETENCY

The bioinformatics field is rich in a wide range of domain-specific knowledge and skills, which originate from different disciplines, as well as in general scientific practices and cognitive processes. These parameters are usually required within a school bioinformatics curriculum, as students study using a bioinformatics activity, and are evaluated by questions embedded within the bioinformatics resources. We suggest a set of criteria to gain a comprehensive view of the characteristics of the questions embedded in bioinformatics learning materials. Such a characterization enables performing a quantitative analysis of students’ learning outcomes and can serve as a basis to describe the teaching approaches involved. A unified platform, which is based on knowledge, cognition and scientific approach, can serve for the analysis of curriculum design and content, as well as learning and teaching processes, outcomes and personal attitudes (see later in the text). The findings and insights may lead to the construction of operational recommendations for policy makers, developers of learning and teaching materials, teachers–trainers and teachers, which in turn may enable to improve and expand the educational resources, their proper use, and broader integration into high schools. This general platform can be supplemented with specific features of each bioinformatics learning module/course, such as the software tools and databases, the topics of bioinformatics activities, module design and format and so forth. The proposed criteria are as follows.

Types of knowledge

Frequent attempts have been made to construct a theoretical framework for knowledge classification, and various types of knowledge have been termed. This criterion is based on a view of scientific knowledge as consisting of four distinguishable but related elements [74]:

- Declarative knowledge—knowledge of ‘what’ facts, concepts, laws, principles and theories are [74]. This knowledge is also known as conceptual, propositional [75], substantive or content knowledge.
- Procedural knowledge—knowing ‘how’ to perform valid actions in certain processes or routines. It also includes an understanding of techniques, methods and appropriate usage and experimental design [76], as well as perceptual motor skills [75]. This knowledge was also termed ‘concepts of evidence’ [77] or ‘thinking behind the doing’ [78].
- Situational knowledge—knowledge of situations that enable the solver to sift relevant features out of the problem statement and to supplement information in the statement [74].
- Strategic knowledge—a general plan of action for the problem-solving process [74] or knowing the conditions ‘when and where’ knowledge would be applicable [71, 79]. These are domain-specific heuristics that aid in the regulation, execution and evaluation of a task [71, 73].

This platform fits the characteristics of bioinformatics research. For instance, when a research question or objective is defined, the researcher uses heuristics and designs a strategy to follow them to get closer to a potential solution. Accordingly, the researcher selects appropriate experimental techniques and bioinformatics tools, considering both the current knowledge base of facts, concepts and theories and the mode of action, demands and application of each bioinformatics tool (and database). Then the researcher uses bioinformatics tools while making decisions on both the diverse options of using each tool and the interpretation and making sense of the results. This is a multistep, iterative and reflective process.

In the bioinformatics learning environment, declarative knowledge is mainly presented in the introductory unit (general concepts and models in biology, bioinformatics and biotechnology), and in the dictionary, procedural knowledge is mainly displayed in the tutorials (specific to each bioinformatics tool), whereas the activities and exercises require the construction, application and integration of these two types of knowledge along with situational knowledge and strategic knowledge.

Scientific approach

The scientific approach or discipline from which the knowledge stems. We distinguish between questions that require the use of either biological topic/context-associated knowledge or bioinformatics-related knowledge and skills or a combination of both. When higher resolution is required, for instance when computational programming and statistical analysis are integral components of the activity, one
might add additional categories, e.g. computation/programming and statistics, respectively.

**Cognitive domain**
Classifications of cognitive demand schemes are largely based on categorizations of knowledge types and associated mental processes that are used to describe educational objectives or assessment tasks. We suggest following the well-known revised Bloom’s Taxonomy [80], which identifies six hierarchical categories of cognitive process dimensions: Remember, Understand, Apply, Analyze, Evaluate and Create.

Analogous to this taxonomy for learning, and despite some criticism, Jungck suggested a taxonomy for ‘quantitative reasoning in biology’, namely, an ‘implementable sequence of steps that is intended to develop students’ metacognitive growth in mathematical thinking through explicit activities’ [81].

**CHARACTERISTICS OF THE BIOINFORMATICS LEARNING ENVIRONMENT**
In the ‘Bioinformatics in the Service of Biotechnology’ learning environment, declarative knowledge in biology and bioinformatics is mainly provided in the ‘Introductory’ and ‘Dictionary’ units. The ‘Tool-Box’ deals mainly with procedural aspects of operating bioinformatics tools and to a lesser extent with declarative concepts of their operation, as well as scientific strategies or conditions for using them. In the ‘Activities’ and ‘Questions’ units, bioinformatics tools are used in biological contexts in problem-solving approaches, where strategic knowledge and situational knowledge are required.

To examine the feasibility of the proposed set of criteria, the questions embedded in two activities (1 and 2 above, n = 63 questions) from the learning environment were characterized. We also noted the type of question, distinguishing between open-ended and multiple-choice. Importantly, although many questions required a combination of two (or more) types of knowledge, questions were classified based on the predominant type of knowledge (examples of questions’ classification are provided in the Supplementary Materials—Supplementary Appendix 4).

The distribution of the questions in those two activities was in accordance with the goal of introducing students to the bioinformatics approach and the procedures for use of bioinformatics tools. Providing answers to approximately half of the questions in these ‘in-depth activities’ requires the use of procedural knowledge, whereas questions that require the use of either declarative or situational knowledge are almost equally represented, and those that require the use of strategic knowledge are the least common (Table 2). Similarly, most questions deal with the bioinformatics approach, solely or in combination with a biological approach. Logically, most questions are associated with the cognitive processes ‘Understand’ and ‘Apply’ for the newly acquired knowledge. Higher-order thinking questions are also represented, yet to a lesser extent. Multiple-choice questions are overrepresented with respect to open-ended questions, probably because they allow automated feedback to students’ answers. Interestingly, providing answers to most multiple-choice questions necessitates the use of procedural knowledge, which mainly stems from the bioinformatics approach and is associated with the cognitive process of ‘Analyze’. On the other hand, providing answers to most open-ended questions necessitates the use of declarative knowledge, which mainly stems from the biological approach and is associated with the cognitive process ‘Understand’ (data not shown).

<table>
<thead>
<tr>
<th>Classification</th>
<th>Categories</th>
<th>Number of questionsa</th>
<th>Significanceb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of question</td>
<td>Multiple choice</td>
<td>40 (63.5%)</td>
<td>$P &lt; 0.05$</td>
</tr>
<tr>
<td></td>
<td>Open ended</td>
<td>23 (36.5%)</td>
<td></td>
</tr>
<tr>
<td>Scientific approach</td>
<td>Biological approach</td>
<td>11 (17.5%)</td>
<td>$P &lt; 0.01$</td>
</tr>
<tr>
<td></td>
<td>Bioinformatics approach</td>
<td>36 (57%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Combined approach</td>
<td>16 (25.4%)</td>
<td></td>
</tr>
<tr>
<td>Domain-specific knowledge</td>
<td>Declarative knowledge</td>
<td>14 (22.2%)</td>
<td>$P &lt; 0.001$</td>
</tr>
<tr>
<td></td>
<td>Procedural knowledge</td>
<td>32 (50.8%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Situational knowledge</td>
<td>12 (19%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Strategic knowledge</td>
<td>5 (7.9%)</td>
<td></td>
</tr>
<tr>
<td>Revised Bloom’s taxonomy of cognitive processes</td>
<td>Remember</td>
<td>3 (4.8%)</td>
<td>$P &lt; 0.001$</td>
</tr>
<tr>
<td></td>
<td>Understand</td>
<td>23 (37.7%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Apply</td>
<td>18 (28.6%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Analyze</td>
<td>10 (15.9%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Evaluate</td>
<td>3 (4.8%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Create</td>
<td>4 (6.3%)</td>
<td></td>
</tr>
</tbody>
</table>

*aThe actual number (left) and percentage (right, in parentheses) of questions in each classification category. bChi-square test was applied. Significance implies that there is a difference between the expected and observed frequency of at least one of the question categories.*
IMPLICATIONS FOR CURRICULUM DESIGN, TEACHERS’ TRAINING AND LEARNING PROCESSES

Such characterization of learning materials may also be used to analyze students’ performance (by the answers they provide to these questions) and attitudes toward bioinformatics (by questionnaires and interviews), e.g. which practices and types of knowledge they perceive as harder to deal with or requiring more support and so forth. Likewise, teachings strategies and teachers’ beliefs on the incorporation of bioinformatics into the high school classroom can also be analyzed with reference to these criteria.

Continuous coordination between different types of knowledge and sources of information is among the characteristics that differentiate real authentic scientific inquiry from tasks commonly used in schools [63]. The learning environment ‘Bioinformatics in the Service of Biotechnology’, its activities and the questions embedded in it comprehensively integrate and present unique features of modern authentic bioinformatics-combined scientific research, which is rich in diverse procedural skills coupled with the declarative knowledge and strategic thinking required to understand and master bioinformatics approaches and applications.

The analysis proposed here for the bioinformatics learning environment may shed light on the means by which high school students perceive and deal with scientifically authentic inquiry activities, as well as on the means used for instruction of such materials. Attempts to make science learning better resemble authentic scientific practices have led to educational reforms, at least since Dewey [82]. Nevertheless, only a few educational researchers have adopted authentic scientific practices, as they are practiced by the scientific community, as a basis for examining the teaching and learning of science (e.g. [83, 84]). We envision that this framework for analyzing curriculum design and content, as well as learning and teaching processes, might be valuable to foster bioinformatics education in high schools (and universities). For instance, it might uncover the kinds of questions students (and teachers) struggle with, and the derived recommendations for adequate instruction may be integrated into the learning environment as well as the teachers’ training courses. It can also serve as a monitoring tool for quality-control processes for developers and policy makers to ensure that the bioinformatics resource characteristics are in line with their educational and scientific goals.

SUPPLEMENTARY DATA

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

Key Points

- The evolution of bioinformatics education should keep up with advances in bioinformatics research.
- Given the availability of publicly accessible research databases and tools, as well as educational resources, it is now practical, appropriate and desirable to broadly incorporate bioinformatics into the high school classroom.
- Bioinformatics resources for high school students should be designed and developed to make the most of educational opportunities and to overcome the challenges involved.
- Endeavors to incorporate bioinformatics into high school science classrooms should be accompanied by educational research at the interface of bioinformatics curriculum design, teaching and learning processes.
- Characterizing the types of knowledge, the scientific approach and the cognitive domain required to answer questions embedded in bioinformatics learning environments is proposed as a platform for analysis of students’ learning outcomes.

Acknowledgements

We wish to thank Hadas Gelbart and Shifra Ben-Dor for fruitful discussions; Orna Dahan, Amir Mitchell, Carmit Avidan-Shpalter, and Hadas Gelbart for collaborative efforts to design and develop the bioinformatics curriculum, and Eilat Avraham, the national supervisor for biotechnology teaching at the Ministry of Education, who had the vision and enterprise to lead the way in bringing bioinformatics to high-school classrooms.

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.

References


