

Section A4 Secondary education	
UDC 378.011.3-051:57:004.6	
Date of initial submission	2026-08-25
Date of acceptance for publication	2026-09-30
Date of publication/release	2026-09-30

Bioinformatics Training of Pre-Service Biology Teachers: From Classroom Learning to Research and Project-Based Activities

Iryna Upatova

Doctor of Pedagogical Sciences, Professor,
Municipal Establishment «Kharkiv Humanitarian-Pedagogical Academy»
of the Kharkiv Regional Council, Kharkiv, Ukraine
e-mail: handiy63@hgpakh.ukr.education
<https://orcid.org/0000-0002-0060-1186>

Vitalii B. Moskalov

Doctor of Philosophy in Biology,
ME “Kharkiv Humanitarian and Pedagogical Academy”
of the Kharkiv Regional Council, Kharkiv, Ukraine
e-mail: vitmoskalov93@gmail.com
<https://orcid.org/0000-0002-9831-3173>

Olena Shyrokykh

Candidate of Biological Sciences, Associate Professor,
Municipal Establishment «Kharkiv Humanitarian-Pedagogical Academy»
of the Kharkiv Regional Council, Kharkiv, Ukraine
e-mail: elena.dekhtiarova@hgpakh.ukr.education
<https://orcid.org/0000-0001-9617-3333>

Iryna Kadenko

Candidate of Pedagogical Sciences, Associate Professor,
Municipal Establishment «Kharkiv Humanitarian-Pedagogical Academy»
of the Kharkiv Regional Council, Kharkiv, Ukraine
e-mail: elena.dekhtiarova@hgpakh.ukr.education
<https://orcid.org/0000-0002-6525-3485>

Andrii Mukhin

Master's Student in the Specialty A4.05 Biology and Human Health,
ME “Kharkiv Humanitarian and Pedagogical Academy”
of the Kharkiv Regional Council, Kharkiv, Ukraine
e-mail: andrey380959316511@gmail.com
<https://orcid.org/0009-0006-1694-6492>

Abstract. The purpose of the study is to develop and theoretically substantiate a conceptual model of bioinformatics training for pre-service biology teachers that integrates classroom learning, computational analysis of biological data, independent research, research-project design, and the transfer of acquired competencies to the supervision of student research and project-based activities. The study has a theoretical and conceptual character and is based on competence-based, inquiry-oriented, research-based, and systems approaches. The research methods included a structured literature search and purposive selection of relevant

publications, conceptual, comparative, and structural-functional analysis, classification and systematization of educational practices, and conceptual modelling. The source base comprised scientific publications and international policy documents addressing bioinformatics competencies, computational thinking in life science education, inquiry- and research-based bioinformatics learning, teacher preparation, secondary-school bioinformatics, sustainable education, and open science.

The study substantiates a professional continuum of bioinformatics training comprising five interconnected functional roles: classroom learner, guided computational investigator, independent researcher, research-project designer, and research mentor. Their progression is determined primarily by increasing research autonomy and responsibility for biological problem formulation, data selection, analytical decision-making, interpretation, and research design rather than by increasing computational complexity. The model integrates four interrelated components of professional preparation: biological, computational and data, research, and pedagogical. Experimental and computational approaches are treated as complementary components of biological inquiry, while programming languages, biological databases, and analytical environments are selected according to the biological question, data characteristics, educational objectives, and learner preparation. The transition from independent researcher to project designer and research mentor represents a qualitative change in professional development: personal research autonomy becomes a resource for supporting the research autonomy of another learner.

The scientific novelty of the study lies in conceptualizing bioinformatics training as a continuous professional trajectory connecting classroom learning, guided computational inquiry, independent biological research, research-project design, and student research mentoring. The proposed model identifies research autonomy and responsibility for decision-making, rather than computational complexity, as its principal developmental axis and specifies the transition from acquiring personal research autonomy to creating conditions for the research autonomy of school students.

The practical significance of the findings lies in the possibility of applying the proposed framework to the design of bioinformatics training for pre-service biology teachers and to the organization of school-level research and project-based activities. It can guide the selection of computational tools and biological data according to research questions and learner preparation, support the progressive transfer of research responsibility from instructor to learner, and facilitate the design and mentoring of feasible student projects. The framework may be particularly applicable to research competitions and work with gifted or highly motivated students, including educational settings where open biological data and computational resources can complement limited experimental infrastructure.

Keywords: digitalization, digital transformation, pre-service biology teachers, critical thinking, computational thinking, data analysis, research-based learning, research autonomy, project-based learning, student research, research mentoring, gifted students.

Біоінформатична підготовка майбутніх учителів біології: від аудиторного навчання до дослідницької та проєктної діяльності

Москальов Віталій Борисович

Доктор філософії з біології,

КЗ «Харківська гуманітарно-педагогічна академія» Харківської обласної ради, Харків, Україна

e-mail: vitmoskalov93@gmail.com

<https://orcid.org/0000-0002-9831-3173>

Широких Олена Олександрівна

кандидат біологічних наук, доцент,

КЗ «Харківська гуманітарно-педагогічна академія» Харківської обласної ради, Харків, Україна

e-mail: elena.dekhtiarova@hgpakh.ukr.education

<https://orcid.org/0000-0001-9617-3333>

Упатова Ірина Петрівна

доктор педагогічних наук, професор,

КЗ «Харківська гуманітарно-педагогічна академія» Харківської обласної ради, Харків, Україна

e-mail: handiy63@hgpakh.ukr.education

<https://orcid.org/0000-0002-0060-1186>

Каденко Ірина Валеріївна

кандидат педагогічних наук, доцент,

КЗ «Харківська гуманітарно-педагогічна академія» Харківської обласної ради, Харків, Україна

e-mail: kadenko19@ukr.net

<https://orcid.org/0000-0002-6525-3485>

Мухін Андрій Миколайович

Магістрант за спеціальністю А4.05 Біологія та здоров'я людини,

КЗ «Харківська гуманітарно-педагогічна академія» Харківської обласної ради, Харків, Україна

e-mail: andrey380959316511@gmail.com

<https://orcid.org/0009-0006-1694-6492>

Анотація. Метою дослідження є розроблення та теоретичне обґрунтування концептуальної моделі біоінформатичної підготовки майбутніх учителів біології, яка інтегрує аудиторне навчання, обчислювальний аналіз біологічних даних, самостійну дослідницьку діяльність, проєктування дослідницьких робіт і перенесення набутих компетентностей у керівництво учнівськими дослідженнями та проєктною діяльністю. Дослідження має теоретико-концептуальний характер і ґрунтується на компетентнісному, дослідницько-орієнтованому, науково-дослідницькому та системному підходах. Методи дослідження охоплювали структурований пошук і цілеспрямований добір релевантних публікацій, концептуальний, порівняльний і структурно-функціональний аналіз, класифікацію та систематизацію освітніх практик і концептуальне моделювання. Джерельну базу становили наукові публікації та міжнародні програмні документи, присвячені біоінформатичним компетентностям, обчислювальному мисленню в природничій освіті, дослідницько-орієнтованому навчанню біоінформатики, підготовці вчителів, біоінформатиці в середній освіті, освіті для сталого розвитку та відкритій науці.

У дослідженні обґрунтовано професійний континуум біоінформатичної підготовки, що охоплює п'ять взаємопов'язаних функціональних ролей: здобувач освіти в аудиторному навчанні, дослідник під керівництвом із використанням обчислювальних методів, самостійний дослідник, розробник дослідницького проєкту та наставник дослідницької

діяльності. Перехід між цими ролями визначається насамперед зростанням дослідницької автономії та відповідальності за формулювання біологічної проблеми, добір даних, прийняття аналітичних рішень, інтерпретацію результатів і проектування дослідження, а не підвищенням обчислювальної складності. Модель інтегрує чотири взаємопов'язані компоненти професійної підготовки: біологічний, обчислювально-інформаційний, дослідницький і педагогічний. Експериментальні та обчислювальні підходи розглядаються як взаємодоповнювальні складові біологічного дослідження, а мови програмування, біологічні бази даних та аналітичні середовища добираються відповідно до біологічного питання, характеристик даних, освітніх цілей і рівня підготовки здобувачів освіти. Перехід від самостійного дослідника до розробника дослідницького проекту та наставника дослідницької діяльності становить якісну зміну професійного розвитку: особиста дослідницька автономія стає ресурсом для підтримки дослідницької автономії іншого здобувача освіти.

Наукова новизна дослідження полягає в концептуалізації біоінформатичної підготовки як безперервної професійної траєкторії, що поєднує аудиторне навчання, кероване дослідження із застосуванням обчислювальних методів, самостійне біологічне дослідження, проектування дослідницьких робіт і наставництво в учнівській дослідницькій діяльності. У запропонованій моделі головною віссю професійного розвитку визначено дослідницьку автономію та відповідальність за прийняття рішень, а не обчислювальну складність; конкретизовано перехід від набуття власної дослідницької автономії до створення умов для дослідницької автономії учнів.

Практичне значення одержаних результатів полягає в можливості застосування запропонованої моделі для проектування біоінформатичної підготовки майбутніх учителів біології та організації дослідницької і проектної діяльності учнів. Модель може слугувати орієнтиром для добору обчислювальних інструментів і біологічних даних відповідно до дослідницьких питань та рівня підготовки здобувачів освіти, забезпечувати поступове передавання відповідальності за дослідження від викладача до здобувача освіти, а також сприяти проектуванню та наставницькому супроводу здійснення учнівських дослідницьких проектів. Запропонований підхід може бути особливо доцільним для підготовки учнівських робіт до дослідницьких конкурсів і роботи з обдарованими або високомотивованими учнями, зокрема в освітніх умовах, де відкриті біологічні дані та обчислювальні ресурси можуть доповнювати обмежену експериментальну інфраструктуру.

Ключові слова: цифровізація, цифрова трансформація, майбутні вчителі біології, критичне мислення, обчислювальне мислення, аналіз даних, дослідницько-орієнтоване навчання, дослідницька автономія, проектне навчання, учнівські дослідження, дослідницьке наставництво, обдаровані учні.

Introduction

Relevance of the Problem. The increasing integration of computational approaches into the life sciences has changed the ways in which biological questions are formulated, investigated, and interpreted. Modern biological research increasingly combines experimental observations with the acquisition, processing, visualization, and computational analysis of biological data. Consequently, the ability to work with biological data is becoming an important component of life science education rather than a competence restricted to professional bioinformaticians. This transition has stimulated efforts to define core bioinformatics competencies for different categories of users and incorporate them into undergraduate life science curricula [1–3].

From an educational perspective, however, computational competence should not be reduced to learning a programming language or mechanically applying ready-made software. Computational thinking requires students to formalize biological problems, identify relevant variables and data, select analytical procedures, evaluate outputs, and relate computational results to the original biological question. Rubinstein and Chor [4] emphasized computational

thinking as a distinct component of life science education, with programming serving as a means of developing algorithmic and logical reasoning rather than an end in itself. Bioinformatics provides a suitable environment for integrating these processes within a research workflow in which students formulate biological questions, obtain or generate data, analyse and visualize them, and interpret the resulting patterns biologically. Open databases and computational environments further allow such activities to be organized around the logic of biological data analysis rather than mastery of particular software products.

This approach has specific implications for the preparation of pre-service biology teachers. Their bioinformatics training has a dual purpose: future teachers need sufficient disciplinary and computational competence to understand how contemporary biological evidence is generated and analysed, but they must also be able to transform this knowledge into learning and research activities appropriate for school students.

This educational trajectory also intersects with broader objectives of sustainable and open education. Sustainable Development Goal 4 emphasizes inclusive and equitable quality education, relevant skills, and education for sustainable development, including its integration into teacher education [5]. Bioinformatics can contribute to these objectives by developing competencies in problem formulation, critical work with data, evidence evaluation, and contemporary research methods. At the same time, the UNESCO Recommendation on Open Science emphasizes access to scientific knowledge, methods, infrastructures, and opportunities for scientific education and capacity development [6]. Open biological databases and open-source computational environments can reduce some resource-related barriers to authentic data-based research, although their educational use still depends on hardware, connectivity, teacher competence, and methodological support.

Analysis of Recent Research and Publications. Previous studies have demonstrated both the perceived need to incorporate bioinformatics into biology teacher education and the feasibility of dedicated learning environments for pre-service teachers [7, 8]. Such preparation can include work with biological databases, sequence analysis, phylogenetic reconstruction, and computational tools while simultaneously addressing their subsequent application in secondary-school biology education [8].

Exposure to bioinformatics tools, however, represents only one stage of professional preparation. A more consequential transition occurs when students move from following predefined analytical procedures to using computational methods to answer biological questions. Inquiry-based approaches facilitate this transition by organizing learning around questions, evidence, analysis, and interpretation. Bioinformatics has been incorporated into inquiry-oriented laboratory education in which computational analysis complements experimental work and real biological datasets are used to address research questions [9]. Experimental and computational approaches can therefore be treated as complementary: experimental observations may generate data and hypotheses for computational analysis, while computational results may reveal patterns, generate hypotheses, or guide subsequent experiments.

This progression also connects classroom learning with increasingly independent research. Structured tasks can initially introduce the logic of biological data analysis under instructor guidance and subsequently provide greater autonomy in selecting data, analytical procedures, visualization, and interpretation. Ultimately, computational approaches can become part of an independent biological investigation rather than remain isolated exercises. Inquiry-based modules and course-based research experiences demonstrate that students can progress towards authentic research practices involving biological data analysis, interpretation, and communication of results [10–12].

For pre-service teachers, this progression has an additional professional dimension: experience acquired as learners and researchers can subsequently be transformed into the capacity to design feasible research tasks and supervise their implementation by school students. Bioinformatics has already been introduced at the secondary-school level [13],

including sequence analysis and phylogenetic reconstruction in programmes for gifted high-school students [14]. Such approaches are relevant to research competitions and project-based activities because access to authentic biological data can extend the range of feasible research questions beyond those determined solely by locally available experimental infrastructure.

Unresolved Aspects of the Problem. Existing research therefore establishes the principal components required for bioinformatics-oriented teacher preparation. Bioinformatics competencies have been systematized for life science education [1–3], computational thinking has been conceptualized within life science education [4], and inquiry-oriented and course-based models demonstrate the integration of bioinformatics with authentic research [9–12]. Bioinformatics instruction has also been developed specifically for pre-service biology teachers [7, 8], while secondary-school applications demonstrate its feasibility for younger learners [13, 14].

These strands, however, have predominantly been considered as separate educational contexts or transitions. In particular, the progression from learning bioinformatics methods, through their application in independent biological research, to designing and mentoring research projects for school students remains insufficiently conceptualized as a continuous component of pre-service biology teacher education. A conceptual framework is therefore needed that connects the acquisition of bioinformatics competence with increasing research autonomy and its subsequent transformation into the capacity to organize and mentor student research.

Purpose of the Study. Accordingly, this study aims to develop and theoretically substantiate a conceptual model of bioinformatics training for pre-service biology teachers that integrates classroom learning, computational analysis of biological data, independent research, research-project design, and the transfer of acquired competencies to the supervision of student research and project-based activities.

Scientific Novelty. The scientific novelty of the study lies not in introducing bioinformatics into biology teacher education or advocating particular programming languages, software environments, or databases, but in conceptualizing its development as *a professional continuum linking classroom learning, guided computational inquiry, independent research, research-project design, and student research mentoring*. Within this continuum, the future teacher progresses from learner to computational investigator, independent researcher, project designer, and research mentor. The framework places biological problem formulation and interpretation at the center of bioinformatics training, treats experimental and computational approaches as complementary components of biological inquiry, and conceptualizes the transition from personal research autonomy to the capacity to support the research autonomy of another learner.

Practical Significance. The practical significance of the proposed framework lies in its applicability to the design of bioinformatics training within pre-service biology teacher education and to the organization of research- and project-based student activities. It provides a basis for selecting computational activities according to biological questions and learner preparation rather than the complexity of particular software tools; for progressively transferring responsibility for data selection, analysis, interpretation, and research design from instructor to learner; and for preparing future teachers to design and mentor feasible school-level research projects. The framework may be particularly applicable to research competitions and work with gifted or highly motivated students, including settings in which access to laboratory infrastructure is limited but open biological data and computational resources are available.

Methodology

Research Methods. This study employed a theoretical-conceptual research design aimed at developing a structured model of bioinformatics training for pre-service biology teachers. The methodological approach combined a structured literature search and purposive selection of relevant publications with conceptual, comparative, and structural-functional analysis,

classification and systematization of educational practices, and conceptual modelling. Framework development followed the general logic of conceptual analysis and synthesis, in which concepts identified across heterogeneous bodies of literature are compared, grouped, and integrated according to their relationships and functions within the phenomenon under investigation [15].

Conceptual synthesis focused on recurrent educational functions, relationships, and transitions identified across the selected studies. The resulting elements were organized according to changes in learner autonomy and professional responsibility. This analysis provided the basis for constructing a developmental continuum linking classroom bioinformatics learning, guided computational inquiry, independent research, research-project design, and pedagogical transfer to the supervision of student research. The resulting framework was subsequently examined for internal coherence by mapping each stage to its dominant educational purpose, characteristic learning activities, expected competencies, and the changing professional role of the pre-service biology teacher.

Data Sources. The source base comprised peer-reviewed scientific publications and relevant international policy documents. The literature search addressed six interconnected areas: (1) bioinformatics education and bioinformatics competencies; (2) bioinformatics in the preparation of pre-service biology and science teachers; (3) computational thinking, programming, and biological data analysis in life science education; (4) inquiry-, research-, and project-based approaches to biology and bioinformatics education; (5) bioinformatics in secondary education, including research activities involving gifted students; and (6) open science and sustainable-development principles relevant to access to contemporary scientific and educational resources. Methodological publications concerning conceptual framework development and literature synthesis were considered separately.

Relevant publications were identified using combinations of the terms: *bioinformatics education, bioinformatics competencies, pre-service biology teachers, pre-service science teachers, computational thinking, biology education, biological data analysis, inquiry-based learning, research-based learning, course-based undergraduate research experience, project-based learning, secondary school bioinformatics, gifted students*, and related terms. Searches prioritized publications indexed in major scholarly databases and available through publishers' platforms or established bibliographic repositories. Recent studies were prioritized to represent current approaches, while earlier publications were retained when they represented foundational competency frameworks, conceptual approaches, or established educational models. Official United Nations and UNESCO documents provided the sustainable-development and open-science context.

Publications were included when they contributed directly to at least one component of the proposed educational trajectory: defining bioinformatics competencies; integrating computational methods with biological learning; supporting inquiry or independent research; preparing future biology teachers; transferring bioinformatics approaches to secondary education; or supporting student research and project activities. Studies focused primarily on advanced omics training, specialist bioinformatics degree programmes, or computational method development without a relevant educational component were excluded from the conceptual synthesis. This delimitation reflected the focus on bioinformatics as a component of biology teacher preparation rather than specialist training for professional bioinformaticians.

Analytical Tools. The selected literature was analysed using an analytical matrix comprising nine dimensions: *target learner group; educational context; type of biological data or problem; degree of learner autonomy; function of computational tools; relationship between computational and experimental activities; type of research activity; expected competencies or learning outcomes; and opportunities for transferring acquired competencies to teaching or student research supervision*. Particular attention was paid to transitions between educational roles and activities rather than to individual software environments.

Comparative analysis was used to identify common and divergent characteristics of the educational approaches represented in the literature; classification and systematization were used to organize recurrent functions and forms of bioinformatics training; and structural-functional analysis was used to examine relationships among biological, computational/data, research, and pedagogical components. Conceptual modelling was then applied to integrate these elements into a coherent professional trajectory. The principal analytical criterion for differentiating successive roles was the redistribution of autonomy and responsibility for biological problem formulation, data selection, analytical decision-making, interpretation, research design, and subsequent mentoring of student research.

Study Limitations. The study is theoretical and conceptual rather than experimental, and the proposed model represents a theoretically substantiated framework derived from the analysis and synthesis of published research rather than an empirically validated progression of competence development. Its stages should therefore be interpreted as functional components of bioinformatics teacher preparation rather than as rigid or universally sequential levels.

The literature selection was designed to support conceptual framework development and was not intended as an exhaustive systematic review or bibliometric analysis. The resulting synthesis is consequently dependent on the scope of the selected literature and the analytical categories used to interpret it. In addition, the framework deliberately excludes advanced specialist and omics-oriented bioinformatics training and focuses on approaches potentially applicable to biology teacher education and subsequent school-level research activities. Empirical validation of the proposed continuum, including operationalization of assessment criteria and evaluation in pre-service teacher education, remains a task for subsequent research.

Results

Conceptual Structure of Bioinformatics Training. The analysis of the selected literature indicates that bioinformatics training for pre-service biology teachers cannot be adequately represented as the acquisition of a discrete set of software skills. Existing competency frameworks distinguish between different profiles and levels of bioinformatics users, while educational studies demonstrate that computational methods acquire educational value when they are embedded in biological problem solving, inquiry, and research activities. On this basis, bioinformatics training in the present model is conceptualized as the interaction of four functional components: *biological, computational and data, research, and pedagogical* ones (Fig. 1).

The *biological component* provides the disciplinary basis of the model. It includes the ability to formulate biologically meaningful questions, identify appropriate biological objects and variables, understand the origin and limitations of biological data, and interpret computational results in relation to biological processes. This component determines the purpose of computational activity: analytical methods are selected in response to a biological question rather than introduced as isolated technical procedures. Consequently, programming proficiency alone is not treated as an indicator of bioinformatics competence within the proposed framework.

The *computational and data component* comprises the ability to obtain, structure, process, analyse, visualize, and document biological data using appropriate computational environments. Its scope depends on the biological problem and may range from basic manipulation and visualization of experimental datasets to retrieval and analysis of nucleotide or protein sequences. The model therefore does not prescribe mastery of a single programming language or software package. Python, Biopython, and R represent complementary tools that can support different stages of the analytical workflow: general data processing and automation, operations with biological sequences and database records, and statistical analysis and visualization, respectively. Their educational function is defined by the biological task to be solved.

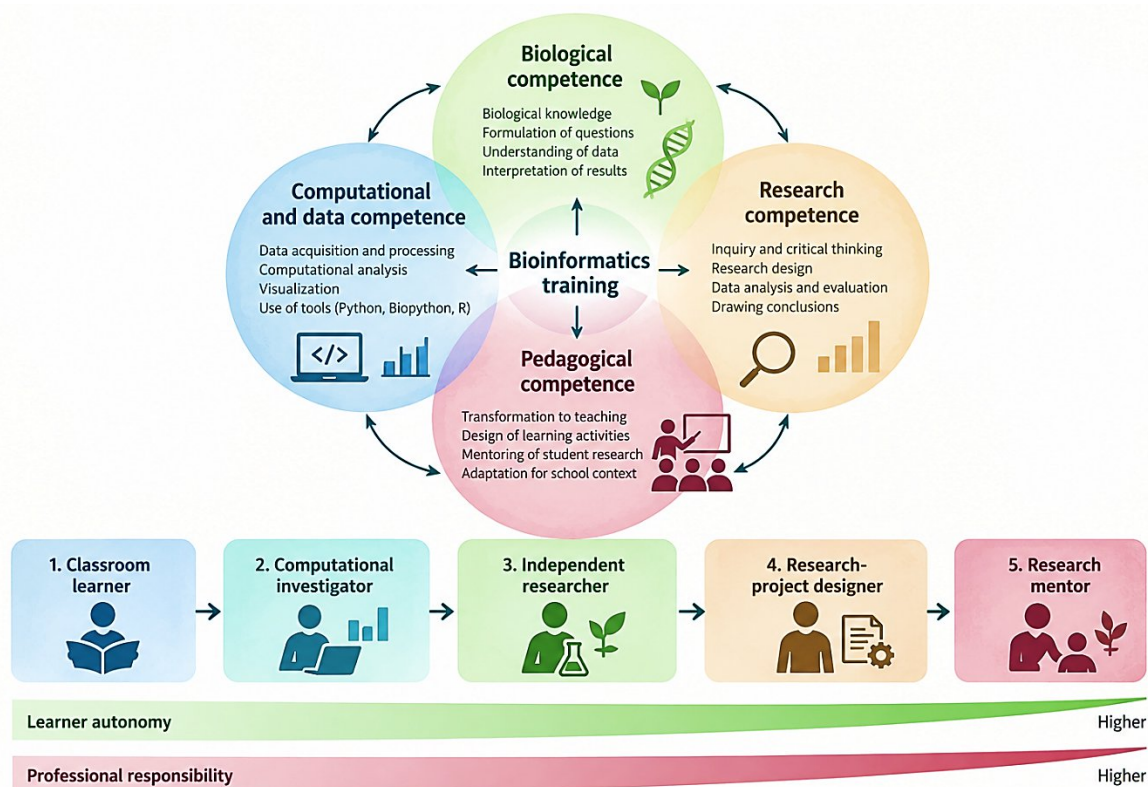


Figure 1. Conceptual structure and developmental continuum of bioinformatics training for pre-service biology teachers

The *research component* transforms computational skills into instruments of scientific inquiry. At this level, students are expected not only to reproduce a demonstrated analytical procedure but also to formulate a research question, select or obtain suitable data, justify an analytical strategy, evaluate the reliability and limitations of the resulting evidence, and formulate biologically meaningful conclusions. Computational analysis may be based on openly available biological data, data obtained during laboratory or field activities, or a combination of both. This creates a methodological connection between experimental and computational approaches rather than treating them as separate forms of biological investigation.

The *pedagogical component* distinguishes the proposed framework from bioinformatics training designed for life science students in general. For a pre-service biology teacher, successful completion of a bioinformatics analysis is not the final professional outcome. The acquired disciplinary, computational, and research experience must subsequently be transformed into the ability to select age-appropriate biological problems, adjust the complexity of data and analytical procedures, design feasible research tasks, provide methodological support, and supervise student interpretation of results. Bioinformatics competence therefore acquires a second function: it becomes a resource for organizing research-oriented biology education.

These four components are not considered consecutive levels. They interact throughout training, while their relative contribution changes as learner autonomy and professional responsibility increase. A classroom exercise may predominantly develop biological and computational understanding; guided inquiry adds elements of research decision-making; independent research requires integration of biological, computational, and research competencies; and the design and supervision of a school-level research project additionally requires pedagogical transformation of this experience.

Accordingly, the proposed model describes bioinformatics training as a *developmental continuum of professional roles* rather than as a sequence of increasingly complex software skills. Five functional stages can be distinguished: (1) *classroom learner*, (2) *guided*

computational investigator, (3) independent researcher, (4) research-project designer, and (5) research mentor. Movement along this continuum is characterized by increasing autonomy in biological problem formulation and analytical decision-making and, at the later stages, by a transition from solving a research problem personally to designing conditions under which another learner can investigate it.

This distinction is important because the highest level of computational complexity does not necessarily correspond to the highest level of professional preparation for a future teacher. A technically sophisticated analysis performed according to a predefined protocol may require less research autonomy than a comparatively simple analysis independently selected and justified for answering a biological question. Similarly, adapting a technically modest bioinformatics task for a school research project may require substantial biological, methodological, and pedagogical judgement. The progression proposed here is therefore defined primarily by *autonomy, integration of competencies, and professional responsibility*, rather than by the complexity of programming or computational methods.

Integration of Computational and Experimental Approaches in Biological Inquiry. Within the proposed framework, computational and experimental approaches are considered complementary components of biological inquiry rather than separate educational trajectories. Their integration is organized around a biological problem and the iterative relationship between hypothesis formulation, data acquisition, computational analysis, and biological interpretation. This principle allows bioinformatics training to remain biologically oriented while introducing computational methods as functional instruments of research (Fig. 2).

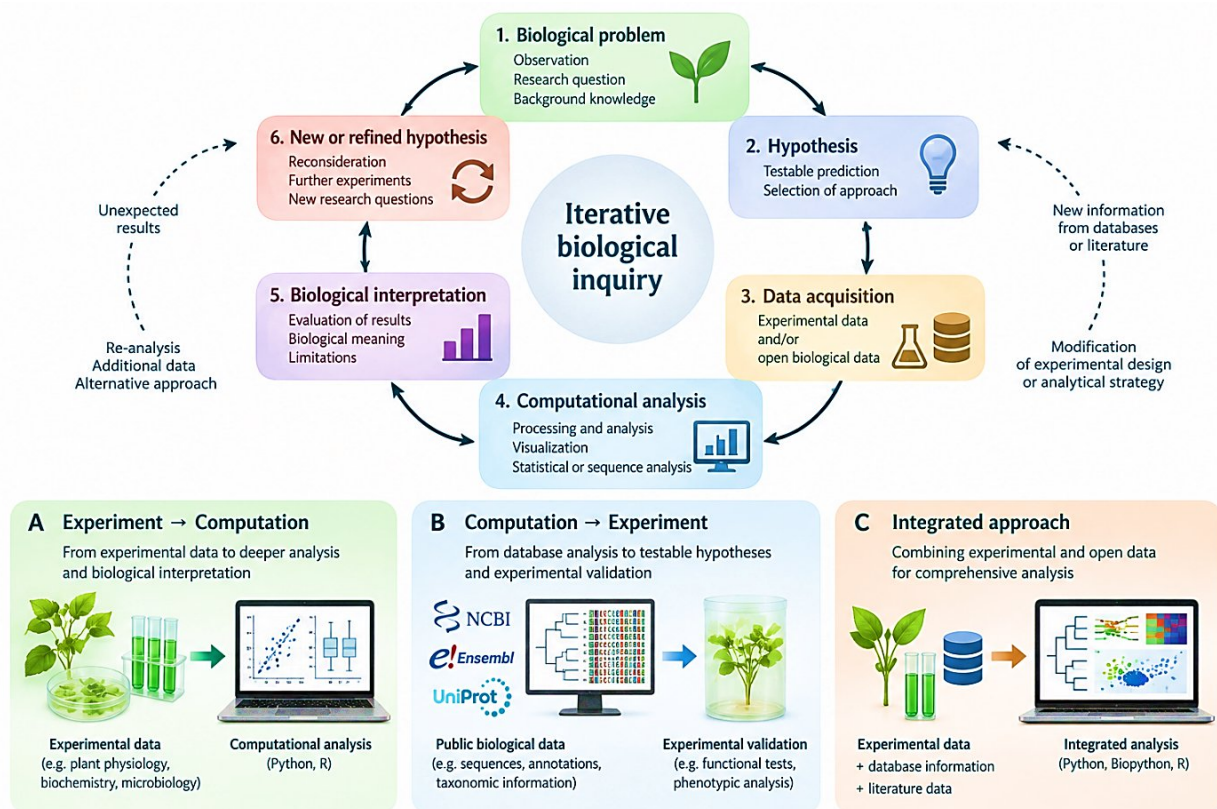


Figure 2. Integration of experimental and computational approaches within an iterative biological research workflow

The starting point of the workflow is a *biological question* rather than a computational technique. Students first identify the biological phenomenon or relationship to be investigated and formulate a hypothesis that can be examined using appropriate evidence. The required data may originate from two principal sources. The first is *experimentally generated data*,

including measurements obtained in laboratory, field, or educational experiments. The second is *open biological data*, such as nucleotide or protein sequences, annotated database records, taxonomic information, or other publicly accessible datasets. These sources can be used independently or combined within a single research project.

When students work with experimentally generated data, computational methods extend the analytical stage of the experiment. Raw observations can be organized, checked, transformed, statistically analysed, and visualized before biological interpretation. In this context, the computational component does not replace the experiment but increases the range and reproducibility of analytical procedures that can be applied to its results. R and Python can support data organization, exploratory analysis, statistical testing, visualization, and reproducible processing workflows. The choice between environments should be determined by the structure of the data and the analytical objective rather than by the requirement to master a particular programming language (Table 1).

Table 1. Examples of computational activities integrating biological questions, data sources, and research-oriented learning

Biological problem / activity	Data source	Computational approach	Appropriate tools	Expected output	Possible research extension
Analysis of an experimental biological response	Measurements obtained in laboratory, field, or educational experiments	Data cleaning and structuring; descriptive and inferential statistics; visualization; comparison of experimental groups	R, Python	Statistical estimates, plots, identification of response patterns	Formulation of additional hypotheses; extension of experimental design; analysis of additional factors or time points
Quantitative characterization of nucleotide or protein sequences	FASTA sequences obtained from public databases or selected records	Sequence parsing; calculation of length, nucleotide/ amino-acid composition, GC content and other basic sequence characteristics	Python, Biopython	Reproducible quantitative description and comparison of sequences	Comparison among genes, proteins, species, or functional groups; formulation of hypotheses concerning sequence variation
Comparative sequence analysis	Homologous nucleotide or protein sequences from public databases	Sequence retrieval and processing; pairwise or multiple comparison; identification of conserved and variable regions	Biopython, Python; external alignment tools where appropriate	Alignment-derived characteristics; identification and visualization of sequence similarities and differences	Comparative analysis of related taxa or sequence variants; selection of regions for further functional or experimental investigation
Phylogenetic reconstruction based on sequence data	Homologous sequences representing selected taxa	Sequence preparation and alignment; calculation or import of evolutionary relationships;	Biopython, Python, R; specialized phylogenetic tools where required	Phylogenetic tree and biologically reasoned interpretation of relationships	Testing alternative taxon sets, genes, or markers; comparison of molecular

		tree construction and visualization			relationships with taxonomic or phenotypic information
Database-driven biological investigation	GenBank/NCBI, UniProt, or other curated biological resources	Retrieval and filtering of records; extraction of annotations and selected features; comparison across records	Biopython, Python	Structured dataset derived from biological database records	Generation of a testable hypothesis; selection of biological objects for subsequent computational or experimental analysis
Integrated experimental-computational project	Student-generated experimental data combined with relevant open biological or reference data	Statistical analysis and visualization of experimental results; retrieval and analysis of complementary database information; integrated biological interpretation	R, Python, Biopython as required by the research question	Combined evidence from experimental and computational analyses	Independent research project; refinement of the hypothesis; design of follow-up experiment or adaptation into a school-level research project

A second pathway begins with biological information available in public databases. Here, computational analysis may itself constitute a substantial part of the investigation. Students can retrieve biological records, extract selected features, compare nucleotide or amino acid sequences, calculate sequence characteristics, identify similarities and differences among biological objects, or use sequence data as a basis for phylogenetic inference. Biopython is particularly suitable for introducing this type of workflow because it connects Python-based computation with common biological data formats, sequence operations, and access to bioinformatics resources. Importantly, the educational outcome is not the execution of individual functions but the ability to connect each computational operation with a biologically interpretable question.

These two pathways can interact in both directions. *Experiment-to-computation* workflows begin with observations or measurements and use computational analysis to identify patterns, test hypotheses, and quantify relationships. Conversely, *computation-to-experiment* workflows may begin with database-derived information or preliminary computational analysis, from which students formulate predictions that can subsequently be examined experimentally. A third, hybrid form combines experimentally generated and publicly available data, for example by interpreting an observed biological response in relation to sequence, taxonomic, or reference information obtained from external resources. Thus, computational analysis can function as an analytical continuation of an experiment, a source of hypotheses for further experimentation, or an independent mode of investigation when the research question can be addressed using existing biological data.

This reciprocal relationship is represented in the proposed framework as an *iterative research cycle: biological problem* → *hypothesis* → *data acquisition* → *computational analysis* → *biological interpretation* → *hypothesis refinement* → *further experimental or computational investigation*.

The cycle is deliberately non-linear. An unexpected analytical result may require verification of data quality, modification of the analytical procedure, reconsideration of the original hypothesis, acquisition of additional data, or formulation of a new biological question. Such iteration is

educationally important because it shifts attention from obtaining a predetermined “correct” output towards evaluating evidence and making justified research decisions.

The functional roles of Python, Biopython, and R can therefore be differentiated without constructing separate instructional trajectories for each tool. *Python* provides a general environment for data processing, automation, reproducible calculations, and implementation of relatively simple analytical procedures. *Biopython* extends this environment towards biological sequences, standard bioinformatics formats, and database-oriented operations. *R* is particularly useful for statistical analysis, data exploration, and graphical representation of experimental and other structured biological datasets. Their functions partly overlap, and the proposed framework does not require all three environments to be used within every learning task or research project. Instead, students should progressively learn to select the least unnecessarily complex toolset that adequately addresses the biological question.

This principle is especially important in pre-service teacher education. Excessive computational complexity can shift cognitive attention from biological reasoning towards syntax and software troubleshooting, whereas excessive simplification may reduce bioinformatics to mechanical execution of predefined commands. The proposed model therefore places *biological interpretability, methodological transparency, and reproducibility* above computational sophistication. A comparatively simple script or statistical analysis that students can explain, reproduce, critically evaluate, and relate to a biological hypothesis is educationally more valuable than a technically advanced workflow whose logic remains opaque to them.

Accordingly, the integration of experimental and computational approaches serves two functions within bioinformatics training. First, it expands the methodological repertoire available for investigating biological problems. Second, it demonstrates that biological evidence can be generated and evaluated through interconnected experimental and computational procedures. This integrated research logic provides the methodological basis for the subsequent transition from guided classroom activities to increasingly independent research and project-based work.

Developmental Continuum from Classroom Learning to Independent Research. The transition from classroom bioinformatics learning to independent research in the proposed model is based on a progressive increase in *learner autonomy and responsibility for research decisions*. Computational complexity alone does not define this progression. Instead, the principal distinction between its stages concerns who formulates the biological question, selects the data and analytical strategy, evaluates intermediate results, and determines how the investigation should proceed. Three successive roles can therefore be distinguished within the research-oriented part of the continuum: *classroom learner, guided computational investigator, and independent researcher* (Table 2).

Table 2. Progression of research autonomy

Dimension	Classroom learner	Guided computational investigator	Independent researcher
Biological question	Instructor-defined	Partly defined / selected	Formulated or substantially refined by student
Data	Provided	Choice among alternatives / partly selected	Independently selected or generated
Analytical strategy	Predetermined	Choice and modification	Independently justified
Interpretation	Guided	Increasingly independent	Independent and critical
Response to unexpected results	Instructor-supported	Selects among alternatives	Revises analysis/design

Instructor role	Instruction	Scaffolding	Research supervision
Student responsibility	Execution and understanding	Analytical decision-making	Complete research design

At the *classroom learner* stage, the principal objective is to establish a connection between biological reasoning and computational operations. The biological problem, dataset, analytical procedure, and expected type of output are largely defined by the instructor. Students learn how biological information is represented as data, how appropriate computational procedures transform these data, and how analytical outputs should be interpreted biologically. Tasks at this stage may include processing a small experimental dataset, retrieving and examining a biological sequence, calculating basic sequence characteristics, or producing and interpreting a statistical visualization. The emphasis is placed on transparency of the workflow and understanding why each analytical step is performed rather than on programming complexity.

At the *guided computational investigator* stage, instructional support is retained but some research decisions are transferred to the student. The instructor may define the general biological problem while students select among alternative datasets, variables, analytical procedures, or visualization strategies. Alternatively, a common dataset may be provided while students formulate different questions that can be addressed using it. Importantly, the analytical pathway is no longer completely predetermined. Students may encounter missing or heterogeneous data, alternative analytical options, unexpected patterns, or results that do not support the initial hypothesis. They are therefore required to justify methodological choices and distinguish between a computational output and the biological evidence that can legitimately be inferred from it.

The transition between these first two stages requires *progressive reduction of instructional scaffolding*. Support can initially include prepared datasets, partially completed code, templates for data processing, explicit analytical sequences, or predefined interpretation questions. These elements are gradually reduced as students become capable of performing the corresponding decisions independently. The purpose of scaffolding is therefore not to simplify the biological problem indefinitely but to make a complex research workflow accessible while preserving its internal logic. A successful transition occurs when students no longer depend on a fixed sequence of commands and can explain, modify, and justify the analytical workflow.

At the *independent researcher* stage, responsibility shifts from completing an analytical task to designing and conducting a coherent investigation. Students formulate or substantially refine the research question, determine what evidence is required, select appropriate experimental or open data, choose computational methods, document the workflow, and interpret results in relation to the original hypothesis. They must also recognize methodological limitations and determine whether additional analysis, new data, or modification of the research design is required. Depending on the problem, an independent project may be computational, experimental–computational, or based on the integration of student-generated and publicly available biological data.

Research autonomy does not imply the absence of instructor support. Rather, the function of support changes from prescribing analytical steps to providing *research supervision*. The instructor may discuss feasibility, methodological validity, data quality, interpretation, or ethical and reproducibility issues without determining every analytical decision. This distinction is important in pre-service teacher education because students simultaneously experience two perspectives: conducting research under supervision and observing how effective research supervision can be organized. Their own progression towards autonomy thus provides experience that can subsequently be transferred to mentoring school-level projects.

The model consequently treats independence as a multidimensional characteristic. A student may demonstrate autonomy in data processing while still requiring substantial assistance in formulating a biologically meaningful question; conversely, a strong biological hypothesis may be investigated using relatively simple computational procedures. Progression should therefore be evaluated across several dimensions: *problem formulation, data selection,*

analytical decision-making, biological interpretation, evaluation of limitations, reproducibility, and responsibility for the overall research design. These dimensions provide a more appropriate basis for assessing development than the technical sophistication of the software or code employed.

The first three stages of the continuum thus transform bioinformatics from a classroom subject into a research instrument. The classroom learner acquires the logic of biological data analysis; the guided computational investigator begins to make and justify analytical decisions; and the independent researcher integrates biological, computational, and methodological reasoning within a complete investigation. This transition establishes the prerequisite for the final professional transformation addressed in the model: moving from conducting research personally to *designing research opportunities for others and mentoring their implementation.*

Pedagogical Transfer: From Researcher to Project Designer and Mentor. The final stage of the proposed continuum involves a qualitative change in the function of bioinformatics competence. Whereas the preceding stages focus primarily on the ability of pre-service biology teachers to conduct their own computational or integrated biological investigations, pedagogical transfer requires them to use this experience to *design research opportunities for other learners.* The central professional transition can therefore be expressed as a shift from *"How can I investigate this biological problem?"* to *"How can I enable a student to investigate a meaningful biological problem?"*

The first role within this transition is that of a *research-project designer.* At this stage, pre-service teachers select a biologically meaningful problem and transform it into a feasible student research project. This transformation requires more than simplification of a university-level analytical workflow. The biological question, data requirements, computational procedures, expected degree of autonomy, duration of the project, and criteria for interpretation must be aligned with the learner's prior knowledge and available educational resources. The resulting project should preserve the essential logic of scientific inquiry while reducing technical complexity that does not contribute directly to answering the biological question.

Bioinformatics provides considerable flexibility for such project design. A research project may be based entirely on open biological data, combine publicly available information with observations or measurements generated by students, or extend a conventional experimental project through computational analysis. Examples include comparative characterization of nucleotide or protein sequences, analysis of sequence variation among taxa, reconstruction and interpretation of phylogenetic relationships, quantitative analysis of experimental biological responses, and integration of experimental observations with information retrieved from biological databases. The appropriate level of computational complexity can therefore be adjusted without necessarily reducing the conceptual significance of the biological problem.

A critical design principle is *methodological feasibility.* A technically sophisticated workflow is not inherently more appropriate for student research than a simpler one. The selected methods should allow the learner to understand the relationship between the biological question, the data, the analytical procedure, and the resulting evidence. For this reason, a project based on relatively simple sequence comparisons or statistical analysis may provide greater research value than an advanced computational pipeline that the student can execute but cannot explain or critically evaluate. The pre-service teacher must therefore distinguish between complexity that supports inquiry and complexity that merely increases the technical burden of the project.

The subsequent role of *research mentor* shifts responsibility from designing the project to supporting another learner's increasingly independent investigation. Effective mentoring does not imply performing difficult analytical steps on behalf of the student or predetermining the desired result. Instead, the mentor helps the learner refine the research question, evaluate the suitability and quality of data, consider alternative analytical decisions, identify methodological limitations, interpret unexpected findings, and determine whether the available evidence

supports the conclusions. The same principle of progressive scaffolding used in the pre-service teacher's own training can thus be transferred to the supervision of student research.

This distinction is particularly important when working with *gifted or highly motivated students* and preparing research projects for competitions. Such students may be capable of working with more complex biological questions or computational procedures, but increasing technical difficulty should not substitute for research independence. The mentor's task is to create conditions in which the student retains intellectual ownership of the project: understanding the origin of the data, justifying methodological decisions, reproducing the analysis, interpreting the results, and defending the conclusions. Bioinformatics is particularly suitable for differentiated research activity because the depth of analysis can be adjusted while the learner continues to work with authentic biological data and scientifically meaningful questions.

The use of open biological databases and open-source computational environments also allows project design to be adapted to different institutional conditions. Where laboratory resources are available, computational analysis can extend experimental work and provide additional levels of interpretation. Where access to laboratory infrastructure is limited, carefully designed data-driven projects can still provide experience in hypothesis formulation, evidence evaluation, reproducible analysis, and scientific communication. Computational projects should not therefore be regarded merely as substitutes for laboratory work, but as an additional form of biological investigation with its own methodological possibilities and limitations.

Pedagogical transfer also requires explicit attention to *reproducibility and research integrity*. A student project should retain information about data provenance, selection criteria, analytical procedures, software or computational environments, and transformations applied to the original data. Code or analytical scripts, where used, should be sufficiently documented to allow the procedure to be repeated. At the same time, conclusions should remain proportional to the evidence generated by the project. These requirements are particularly important in educational research activities because they shift assessment from the visual attractiveness or apparent sophistication of a final product towards transparency of the research process and defensibility of the biological interpretation.

The transition from independent researcher to project designer and research mentor completes the professional continuum proposed in this study. Bioinformatics competence is thereby transformed from an individual research capability into a *pedagogical resource*. The pre-service biology teacher first experiences the progression from structured learning to independent inquiry and subsequently applies the same principles when creating and supervising research opportunities for school students. The complete trajectory can therefore be represented as *learner → computational investigator → independent researcher → research-project designer → research mentor*, with increasing autonomy during the first part of the continuum followed by increasing responsibility for enabling the autonomy of another learner.

Generalization of the Results. The proposed model integrates bioinformatics training of pre-service biology teachers into a continuous professional trajectory rather than treating it as a separate set of computational skills. It combines biological, computational and data, research, and pedagogical components and represents their development through five functional roles: *classroom learner → guided computational investigator → independent researcher → research-project designer → research mentor*. Progression through the first three roles is characterized primarily by increasing autonomy in formulating biological questions, selecting and analysing data, choosing appropriate methods, interpreting results, and making research decisions. Computational complexity is therefore treated as a variable characteristic of particular tasks rather than as the principal criterion of professional development.

The subsequent transition from independent researcher to project designer and research mentor changes the direction of this development: personal research competence becomes a resource for creating conditions for the research autonomy of another learner. Within this trajectory, experimental and computational approaches can function reciprocally, while the

choice of Python, Biopython, R, databases, or other analytical resources is determined by the biological question and educational purpose. Thus, the model connects *bioinformatics learning, independent biological research, and the design and mentoring of student research projects* within a single framework for pre-service biology teacher education.

Discussion

Interpretation of the Results. The proposed framework indicates that bioinformatics competence in pre-service biology teacher education should be understood more broadly than technical proficiency in particular programming languages or computational environments. Its central organizing element is the biological problem: computational tools are selected according to the question being addressed, the characteristics of the available data, and the required analytical procedure. Biological, computational/data, research, and pedagogical components therefore function as interrelated dimensions of professional competence rather than as separate areas of training.

This interpretation also changes the criterion by which progression in bioinformatics training is assessed. Computational complexity alone does not adequately represent professional development. A learner who independently formulates a relatively simple biological question, selects appropriate data and methods, evaluates limitations, and interprets the resulting evidence may demonstrate greater research competence than one who executes a technically sophisticated but predefined analytical pipeline. The five roles identified in the proposed continuum should therefore be interpreted primarily through changes in research autonomy, integration of biological and computational reasoning, and responsibility for decision-making.

The progression from classroom learner to guided computational investigator and independent researcher reflects a gradual transfer of responsibility from instructor to learner. This includes responsibility for problem formulation, data selection, analytical decisions, interpretation, and, ultimately, research design. Scaffolding remains important during this transition but changes in function: rather than permanently simplifying the analytical process, it progressively enables the learner to make and justify research decisions independently.

Experimental and computational approaches occupy complementary positions within this process. Experimental observations may generate data and hypotheses for computational analysis, whereas computational analysis may reveal patterns, refine hypotheses, or guide subsequent experimentation. At the same time, their integration does not require every research activity to contain both components. Open biological data may provide a sufficient basis for some investigations, whereas other projects may combine experimentally generated data with computational or statistical analysis. Thus, methodological configuration should be determined by the biological question and educational objective rather than by an expectation of maximum computational complexity.

For pre-service biology teachers, independent research is not the endpoint of this trajectory. The transition from independent researcher to research-project designer and research mentor changes the direction of professional development: personal research autonomy becomes a resource for supporting the research autonomy of another learner. Project design requires the transformation of a biological problem into a feasible educational investigation without losing its scientific logic, whereas mentoring requires decisions about when support is necessary and when analytical and interpretative responsibility should remain with the student.

Comparison with Previous Studies. This interpretation is consistent with existing bioinformatics competency frameworks while extending them towards the specific professional functions of future biology teachers. Bioinformatics competencies have increasingly been differentiated according to user roles and educational needs. Mulder et al. [2] distinguished different categories of bioinformatics users, Wilson Sayres et al. [1] defined core competencies for undergraduate life science education, and the updated ISCB framework differentiated knowledge, skills, and attitudes across educational and professional contexts [3].

These approaches support the view that pre-service biology teachers do not require specialist preparation equivalent to professional bioinformaticians but need a competency profile appropriate to biological inquiry and subsequent educational practice.

The priority given in the proposed framework to biological problems rather than individual software tools is also consistent with previous educational approaches. Computational thinking in life science education should not be equated with programming itself; programming can instead function as a means of formalizing and investigating biological problems [4]. Bioinformatics instruction similarly requires integration of biological and computational reasoning rather than technical operations in isolation [16]. Project-based bioinformatics training further emphasizes authentic problems, learner initiative and independence, accompanied by a transition of the instructor towards an advisory role [17].

Previous inquiry- and research-oriented studies also support the transition from structured learning towards increasing research autonomy. Brown [10] demonstrated a student-led inquiry model in which undergraduates selected research targets and applied scientific research processes within a structured framework. Course-based research experiences have incorporated bioinformatics analysis and visualization into authentic investigations [11], while classroom preparation has been connected with subsequent independent bioinformatics research [12]. These studies support individual elements of the proposed progression, whereas the present model makes the transfer of responsibility for research decisions an explicit organizing dimension.

The complementary relationship between computational and experimental investigation is supported by Goller et al. [9], who incorporated bioinformatics tools and real datasets alongside wet-laboratory experiments within inquiry-based molecular biology modules. Their approach demonstrates that computational analysis can participate directly in biological investigation rather than function as an isolated digital addition to laboratory education. The present framework extends this principle by explicitly allowing reciprocal transitions between experimental and computational investigation as well as projects based predominantly on either type of evidence.

Evidence from teacher and secondary education further supports the feasibility of the pedagogical-transfer component. Bioinformatics has been introduced at the secondary-school level using authentic biological questions and contemporary scientific resources [13]. Machluf et al. [18] demonstrated that its integration into a high-school science curriculum can provide access to authentic scientific practices while also revealing pedagogical and implementation challenges relevant to teacher education. Bioinformatics preparation of pre-service biology teachers has likewise been connected explicitly with subsequent school biology teaching [8], while sequence alignment and phylogenetic reconstruction have been implemented with gifted high-school students [14]. Existing studies therefore support the principal components of the proposed framework, but predominantly address them as separate educational contexts. The present model connects these components within a continuous trajectory from bioinformatics learning through independent research to project design and research mentoring.

Scientific Novelty. The scientific novelty of the proposed framework can be specified at three interrelated levels. First, it integrates processes that have commonly been considered separately—bioinformatics learning, inquiry-based computational activity, independent biological research, and pedagogical application—within a single professional continuum: *classroom learner* → *guided computational investigator* → *independent researcher* → *research-project designer* → *research mentor*. Bioinformatics competence is thereby conceptualized not as mastery of a fixed set of tools but as a developing capacity to select and apply computational approaches within biological inquiry and subsequently transfer this experience to educational practice.

Second, the model identifies *research autonomy and responsibility for decision-making*, rather than computational complexity, as the principal developmental axis. This allows technically different activities to occupy comparable positions in the framework when they

require similar levels of independence in problem formulation, data selection, methodological choice, interpretation, and evaluation of limitations.

Third, the model identifies a qualitative change after the stage of independent research. During the initial stages, development involves increasing autonomy of the pre-service teacher; during pedagogical transfer, professional competence is increasingly expressed through the capacity to create conditions for the autonomy of another learner. This transition from *acquiring research autonomy to enabling research autonomy* provides the conceptual link between personal bioinformatics competence and the professional functions of research-project design and mentoring.

Practical Significance. The proposed framework can be used to organize bioinformatics training in pre-service biology teacher education without making particular programming languages or software environments the primary curricular objective. Python, Biopython, R, biological databases, and other computational resources can be selected according to biological questions, data types, learner preparation, and the degree of research autonomy expected at a particular stage. The same principle permits flexible combinations of computational, experimental, and integrated research activities.

The framework also provides a basis for designing and mentoring school-level research projects. This is particularly relevant to gifted or highly motivated students and competition-based activities, for which greater analytical depth may be appropriate but should not substitute for genuine student participation in research. Project quality can instead be considered in terms of the defensibility of the biological question, transparency of data provenance, methodological justification, reproducibility, critical interpretation, and correspondence between evidence and conclusions. The practical objective is therefore not maximum computational sophistication but increasing student responsibility for meaningful research decisions.

Open biological databases and open-source computational environments can broaden opportunities for such activities. This is consistent with the UNESCO Recommendation on Open Science, which emphasizes access to scientific knowledge, methods, infrastructures, and opportunities for scientific education [6]. Data-driven projects may reduce dependence on specialized laboratory infrastructure and expand the range of feasible investigations. However, open access does not eliminate barriers associated with hardware, connectivity, digital literacy, teacher expertise, and methodological support.

These practical implications should be considered in light of the limitations of the present study. The proposed framework is a *conceptual model rather than an empirically validated developmental sequence*, and its five roles should be regarded as functional transitions rather than rigid stages. It deliberately prioritizes accessible biological data analysis over advanced specialist and omics-oriented workflows, increasing its relevance to teacher education while limiting its applicability as a general model of professional bioinformatics training. Implementation will also depend on curriculum structure, prior computational preparation, institutional resources, and educational context.

Empirical validation is therefore the principal next step. Assessment criteria should operationalize biological problem formulation, data selection, analytical decision-making, reproducibility, interpretation, research design, and mentoring competence. Future studies should determine whether research-oriented bioinformatics training improves not only future teachers' computational competence but also their capacity to design feasible student investigations and support learner autonomy. In this sense, the practical potential of the proposed continuum lies in connecting *learning bioinformatics, conducting biological research, and enabling student research* within a coherent model of pre-service biology teacher education.

Conclusions

This study developed a conceptual model of bioinformatics training for pre-service biology teachers that connects classroom learning with research and subsequent pedagogical practice. The model positions bioinformatics not as a discrete set of digital or programming skills, but as a methodological component of contemporary biological inquiry in which computational tools are selected according to the biological question, data characteristics, and analytical purpose. Accordingly, biological reasoning and interpretation remain central, while Python, Biopython, R, biological databases, and other computational resources perform complementary functional roles.

The proposed framework integrates four interrelated components of professional preparation: biological, computational and data, research, and pedagogical competence. Their development is represented through a continuum of five functional roles: classroom learner → guided computational investigator → independent researcher → research-project designer → research mentor. Progression through the first three roles is primarily characterized by increasing autonomy in formulating biological questions, selecting data and analytical approaches, interpreting results, and making research decisions. The subsequent transition to project designer and mentor changes the direction of professional responsibility: competence is expressed increasingly through the ability to design feasible research opportunities and support the research autonomy of another learner.

The model also integrates experimental and computational approaches within a common iterative logic of biological inquiry. Computational analysis may extend experimental work, open biological data may provide an independent basis for investigation or generate hypotheses for subsequent experiments, and both sources of evidence may be combined within integrated research projects. Such flexibility makes it possible to adjust the methodological complexity of learning and research activities while preserving authentic biological questions and research logic.

For teacher education, this approach provides a basis for connecting university-level bioinformatics training with research- and project-based activities in school biology. It may be particularly useful for supervising research projects and competition-oriented activities involving gifted or highly motivated students, because analytical depth and computational complexity can be adapted to learner preparation and available resources. The principal criterion should not be technical sophistication, but the learner's meaningful participation in formulating the problem, obtaining or selecting data, making analytical decisions, ensuring reproducibility, and interpreting biological evidence.

The proposed continuum is theoretical and should not be interpreted as an empirically established sequence of competence development. Its further application requires operational criteria for assessing research autonomy and mentoring competence, followed by empirical evaluation in pre-service biology teacher education. Such validation would make it possible to determine how effectively the transition from learning bioinformatics, through conducting research, to enabling student research contributes to the professional preparation of future biology teachers.

Список використаних джерел

1. Wilson Sayres M. A., Hauser C., Sierk M., et al. Bioinformatics core competencies for undergraduate life sciences education. *PLoS ONE*. 2018. Vol. 13, no. 6. Article e0196878. <https://doi.org/10.1371/journal.pone.0196878>
2. Mulder N., Schwartz R., Brazas M. D., et al. The development and application of bioinformatics core competencies to improve bioinformatics training and education. *PLoS Computational Biology*. 2018. Vol. 14, no. 2. Article e1005772. <https://doi.org/10.1371/journal.pcbi.1005772>
3. Brooksbank C., Brazas M. D., Mulder N., et al. The ISCB competency framework v. 3: a revised and extended standard for bioinformatics education and training. *Bioinformatics Advances*. 2024. Vol. 4, no. 1. Article vbae166. <https://doi.org/10.1093/bioadv/vbae166>

4. Rubinstein A., Chor B. Computational Thinking in Life Science Education. *PLoS Computational Biology*. 2014. Vol. 10, no. 11. Article e1003897. <https://doi.org/10.1371/journal.pcbi.1003897>
5. United Nations. *Transforming our world: the 2030 Agenda for Sustainable Development*. New York: United Nations, 2015. URL: <https://sdgs.un.org/2030agenda>
6. UNESCO. *UNESCO Recommendation on Open Science*. Paris: UNESCO, 2021. URL: <https://unesdoc.unesco.org/ark:/48223/pf0000379949>
7. Dong H., Joo H., Lee K.-J. Inservice and Preservice Biology Teachers and Professors' Perception of Introducing Bioinformatics in Biology Education. *Biology Education*. 2013. Vol. 41, no. 4. P. 601–617. <https://doi.org/10.15717/bioedu.2013.41.4.601>
8. Sari I. J., El Islami R. A. Z. STEM literacy-oriented bioinformatics education through the development of e-BIMO for pre-service biology teachers in the 21st century. *Discover Education*. 2025. Vol. 4. Article 366. <https://doi.org/10.1007/s44217-025-00836-5>
9. Goller C. C., Srougi M. C., Chen S. H., Schenkman L. R., Kelly R. M. Integrating Bioinformatics Tools into Inquiry-Based Molecular Biology Laboratory Education Modules. *Frontiers in Education*. 2021. Vol. 6. Article 711403. <https://doi.org/10.3389/feduc.2021.711403>
10. Brown J. A. L. Evaluating the effectiveness of a practical inquiry-based learning bioinformatics module on undergraduate student engagement and applied skills. *Biochemistry and Molecular Biology Education*. 2016. Vol. 44, no. 3. P. 304–313. <https://doi.org/10.1002/bmb.20954>
11. Gao L., Guo M. A course-based undergraduate research experience for bioinformatics education in undergraduate students. *Biochemistry and Molecular Biology Education*. 2023. Vol. 51, no. 2. P. 189–199. <https://doi.org/10.1002/bmb.21710>
12. Bartlett B., Stitt-Bergh M., Kantar M., Bingham J.-P. A data science practicum to introduce undergraduate students to bioinformatics for research. *Biochemistry and Molecular Biology Education*. 2023. Vol. 51, no. 5. P. 520–528. <https://doi.org/10.1002/bmb.21762>
13. Lewitter F., Bourne P. E. Teaching Bioinformatics at the Secondary School Level. *PLoS Computational Biology*. 2011. Vol. 7, no. 10. Article e1002242. <https://doi.org/10.1371/journal.pcbi.1002242>
14. Park Y., Cha H. Design and Implementation of an Instructional Bioinformatics Program for Gifted High School Students: Emphasizing Phylogenetic Reconstruction through Sequence Alignment. *Biology Education*. 2025. Vol. 53, no. 4. P. 584–598. <https://doi.org/10.15717/bioedu.2025.53.4.584>
15. Jabareen Y. Building a Conceptual Framework: Philosophy, Definitions, and Procedure. *International Journal of Qualitative Methods*. 2009. Vol. 8, no. 4. P. 49–62. <https://doi.org/10.1177/160940690900800406>
16. Tapprich W. E., Reichart L., Simon D. M., Duncan G., McClung W., Grandgenett N., Pauley M. A. An instructional definition and assessment rubric for bioinformatics instruction. *Biochemistry and Molecular Biology Education*. 2021. Vol. 49, no. 1. P. 38–45. <https://doi.org/10.1002/bmb.21361>
17. Emery L. R., Morgan S. L. The application of project-based learning in bioinformatics training. *PLoS Computational Biology*. 2017. Vol. 13, no. 8. Article e1005620. <https://doi.org/10.1371/journal.pcbi.1005620>
18. Machluf Y., Gelbart H., Ben-Dor S., Yarden A. Making authentic science accessible—the benefits and challenges of integrating bioinformatics into a high-school science curriculum. *Briefings in Bioinformatics*. 2017. Vol. 18, no. 1. P. 145–159. <https://doi.org/10.1093/bib/bbv113>

References

1. Wilson Sayres, M. A., Hauser, C., Sierk, M., Robic, S., Rosenwald, A. G., Smith, T. M., et al. (2018). Bioinformatics core competencies for undergraduate life sciences education. *PLoS ONE*, 13(6), e0196878. <https://doi.org/10.1371/journal.pone.0196878>
2. Mulder, N., Schwartz, R., Brazas, M. D., Brooksbank, C., Gaeta, B., Morgan, S. L., et al. (2018). The development and application of bioinformatics core competencies to improve bioinformatics training and education. *PLoS Computational Biology*, 14(2), e1005772. <https://doi.org/10.1371/journal.pcbi.1005772>
3. Brooksbank, C., Brazas, M. D., Mulder, N., et al. (2024). The ISCB competency framework v. 3: A revised and extended standard for bioinformatics education and training. *Bioinformatics Advances*, 4(1), vbae166. <https://doi.org/10.1093/bioadv/vbae166>
4. Rubinstein, A., & Chor, B. (2014). Computational thinking in life science education. *PLoS Computational Biology*, 10(11), e1003897. <https://doi.org/10.1371/journal.pcbi.1003897>
5. United Nations. (2015). *Transforming our world: The 2030 Agenda for Sustainable Development*. United Nations. URL: <https://sdgs.un.org/2030agenda>
6. UNESCO. (2021). *UNESCO Recommendation on Open Science*. UNESCO. URL: <https://unesdoc.unesco.org/ark:/48223/pf0000379949>
7. Dong, H., Joo, H., & Lee, K.-J. (2013). Inservice and preservice biology teachers and professors' perception of introducing bioinformatics in biology education. *Biology Education*, 41(4), 601–617. <https://doi.org/10.15717/bioedu.2013.41.4.601>
8. Sari, I. J., & El Islami, R. A. Z. (2025). STEM literacy-oriented bioinformatics education through the development of e-BIMO for pre-service biology teachers in the 21st century. *Discover Education*, 4, 366. <https://doi.org/10.1007/s44217-025-00836-5>
9. Goller, C. C., Srougi, M. C., Chen, S. H., Schenkman, L. R., & Kelly, R. M. (2021). Integrating bioinformatics tools into inquiry-based molecular biology laboratory education modules. *Frontiers in Education*, 6, 711403. <https://doi.org/10.3389/educ.2021.711403>
10. Brown, J. A. L. (2016). Evaluating the effectiveness of a practical inquiry-based learning bioinformatics module on undergraduate student engagement and applied skills. *Biochemistry and Molecular Biology Education*, 44(3), 304–313. <https://doi.org/10.1002/bmb.20954>
11. Gao, L., & Guo, M. (2023). A course-based undergraduate research experience for bioinformatics education in undergraduate students. *Biochemistry and Molecular Biology Education*, 51(2), 189–199. <https://doi.org/10.1002/bmb.21710>
12. Bartlett, B., Stitt-Bergh, M., Kantar, M., & Bingham, J.-P. (2023). A data science practicum to introduce undergraduate students to bioinformatics for research. *Biochemistry and Molecular Biology Education*, 51(5), 520–528. <https://doi.org/10.1002/bmb.21762>
13. Lewitter, F., & Bourne, P. E. (2011). Teaching bioinformatics at the secondary school level. *PLoS Computational Biology*, 7(10), e1002242. <https://doi.org/10.1371/journal.pcbi.1002242>
14. Park, Y., & Cha, H. (2025). Design and implementation of an instructional bioinformatics program for gifted high school students: Emphasizing phylogenetic reconstruction through sequence alignment. *Biology Education*, 53(4), 584–598. <https://doi.org/10.15717/bioedu.2025.53.4.584>
15. Jabareen, Y. (2009). Building a conceptual framework: Philosophy, definitions, and procedure. *International Journal of Qualitative Methods*, 8(4), 49–62. <https://doi.org/10.1177/160940690900800406>
16. Tapprich, W. E., Reichart, L., Simon, D. M., Duncan, G., McClung, W., Grandgenett, N., & Pauley, M. A. (2021). An instructional definition and assessment rubric for bioinformatics instruction. *Biochemistry and Molecular Biology Education*, 49(1), 38–45. <https://doi.org/10.1002/bmb.21361>

17. Emery, L. R., & Morgan, S. L. (2017). The application of project-based learning in bioinformatics training. *PLoS Computational Biology*, 13(8), e1005620. <https://doi.org/10.1371/journal.pcbi.1005620>
18. Machluf, Y., Gelbart, H., Ben-Dor, S., & Yarden, A. (2017). Making authentic science accessible—the benefits and challenges of integrating bioinformatics into a high-school science curriculum. *Briefings in Bioinformatics*, 18(1), 145–159. <https://doi.org/10.1093/bib/bbv113>