International Journal of
Molecular Sciences

MDPI

Article

BioSuiteT: A Unified Tool for Biological Sequences Analysis

Victor Terron-Macias [1], Jezreel Mejía[2*], Mirna Muñoz[2*], Miguel Terron-Hernández[3*], Miguel Canseco-Pérez[4*], Roberto Hernán[4], and Yadira Quiñonez[5]

[1] Monterrey Institute of Technology and Higher Education, Academic Unit of Computer Science; A00843371@tec.mx
[2] Research Centre in Mathematics, Software Engineering Unit ; jmejia@cimat.mx (J.M.); mirna.munoz@cimat.mx (M.M.)
[3] Technological University of Tlaxcala, Industrial Maintenance Engineering Department; miguel.terron@uttlaxcala.edu.mx
[4] Polytechnic University of Chiapas, Agroindustrial Engineering Department; mcanseco@ia.upchiapas.edu.mx (M.C.-P.), rhernan@ia.upchiapas.edu.mx (R.H.)
[5] Autonomous University of Sinaloa, Informatics Department; yadiraqui@uas.edu.mx
* Correspondence: jmejia@cimat.mx (J.M.); mirna.munoz@cimat.mx (M.M.); miguel.terron@uttlaxcala.edu.mx (M.T.-H.); mcanseco@ia.upchiapas.edu.mx (M.C.-P.)

Simple Summary: Bioinformatics or people related to biology field working with biological information often struggle to analyze their data because they need to use multiple complex computer programs that require specialized technical knowledge. We created a user-friendly application called BioSuiteT that combines many essential tools for analyzing biological information into a unified platform. Our goal was to help users to spend more time on their research and less time dealing with technical difficulties. We tested our platform with 73 users and found that they spent significantly less time setting up and switching between different tools compared to existing software. BioSuiteT allows to study DNA sequences, protein structures, compare genetic information between species, and perform other important analyses through a simple web browser, without needing advanced computer skills but also by providing the local installer if required. By removing technical barriers, BioSuiteT helps scientists focus on solving important biological questions that can benefit human health and our understanding of life.

Abstract: Bioinformatics tools' increasing complexity and fragmentation present significant challenges for researchers, particularly those without extensive programming expertise. This research presents BioSuiteT v1.0, a unified web-based platform that integrates multiple bioinformatics tools into a single, user-friendly environment. BioSuiteT incorporates twelve core functionalities, including DNA and protein sequence analysis, transcription processes, sequence alignment, BLAST integration, PDB visualization and analysis, phylogenetic tree construction, MOTIFS analysis, and regular expression searching. The platform was developed using the Django framework and MongoDB, following ISO/IEC 29110 standard. Performance testing demonstrated the platform's capability to handle concurrent users while maintaining stable performance with a reduced memory usage. Validation testing involving eight researchers, five bioinformatics experts, and sixty practitioners showed a 95% reduction in technical support requirements compared to traditional installable tools and a 90% reduction in tool switching time. BioSuiteT effectively addresses key challenges identified in bioinformatics software usage, including installation barriers, programming expertise requirements, and tool fragmentation, while maintaining the necessary functionality for biological sequence analysis.

Keywords: proteomics; genomics; omics analysis tool; bioinformatics tools.

Academic Editor: Firstname Lastname

Received: date
Revised: date
Accepted: date
Published: date

Citation: To be added by editorial staff during production.

1. Background

Biological sequence analysis represents one of the most important activities in the bioinformatics field [1], significantly enhanced by advances in the computational field due the voluminous and incremental datasets and analysis methods [2]; it spans a wide range of disciplines, each focusing on the comprehensive study of a specific type of biological sequence; this includes: genomics, transcriptomics, proteomics [3, 4], which are one of the most prominent omics fields, each contributing to insights into the complex interactions of biological systems [5, 6, 7]. Integrating multiple types of biological sequence data is often referred to as multi-omics which is used to analyze different biological processes [8].

Moreover, researchers face significant challenges in data analysis due to the large amount of biological data and the integration of multiple tools that can efficiently complete the analysis [9]. Some bioinformatics tools focus on specific data types and lack interoperability between them; to analyze these data, it is required to integrate these tools properly [10].

Nevertheless, recent advances in bioinformatics have led to the development of integrated platforms that aim to overcome these limitations by providing a unified environment for data analysis [11]. Notable examples include ***ExPASy*** [12], a global web server for the biology field supplied by the Swiss Bioinformatics Institute, which focuses on offering access to a wide range of databases and analytical tools [13]. Another notable example is ***NCBI Tools*** [14], which provides various tools for analyzing, downloading, and visualizing biological data [15].

Additionally, there is a diverse range of bioinformatics tools such as ***BioPython***, a Python library that comprises a set of modules for performing different types of biological tasks [16], ***Cogent3***, a Python library that analyzes biological sequence data [17], ***Cactus***, an open-source software program used for performing genome alignments [18], ***PyCogent***, a software library for genomic biology that provides a framework for analyzing biological sequences, plot generation, among others tasks [19], ***PyEvolve***, an open-source Python framework for implementing genetic algorithms to model and simulate evolutionary processes within biological sequences [20], and ***FungiRegEx***, a web application with a GUI used for regular expression detection in fungal genomes [21], among others tools that mostly are focused into performing a specific task.

However, many existing bioinformatics software require computational expertise due to its complexity, limiting its accessibility to a broader scientific audience that could not be directly involved in the computational field. This includes platforms that integrate multiple tools into a unified environment but remain challenging when used [9].

Moreover, many surveys and studies confirm that the challenges associated with using bioinformatics software are related mostly to the following issues:

1. *Poor design and documentation*: The lack of documentation and a proper design make it difficult to install and use any tool effectively [22, 23, 24].
2. *Lack of maintenance and support*: Once a tool is released, it may not be maintained or updated, leading to unresolved bugs and issues [22, 23, 24].
3. *Compatibility and standardization:* This can lead to compatibility issues [24], especially when integrating multiple tools into a workflow.
4. *Overabundance and redundancy*: The number of available tools can be overwhelming, with many tools performing similar functions [22, 23].
5. *Learning curves*: Some tools have varying interfaces and functionalities, and this can create a steep learning curve for users, particularly those new to bioinformatics [22], especially if code skills are required.
6. *Lack of community feedback and reviews*: This is to help users make informed decisions about which tools to use [22].

7. *Interpretability of results*: Questioning whether the software is accurately processing data [22, 24].
8. *Resource allocation*: Some tools often take precedence over improving or supporting existing ones, leading to a fragmented ecosystem where resources are spread thin [22, 23].
9. *Platform-specific tools:* Tools are designed to work optimally with specific data types or experimental setups, which can limit their general applicability and require users to adapt their workflows accordingly [22].
10. *User-friendliness and accessibility:* Many bioinformatics tools are not user-friendly, hindering their effective use [24], and most software tools lack a Graphic User Interface (GUI), making them difficult to use.

According to the survey by J. Cazier [22], around 52% of responses explicitly or implicitly mentioned that there are too many tools and emphasized the challenges this creates, especially for uses without programming knowledge (this perception is shown in **Figure 1**).

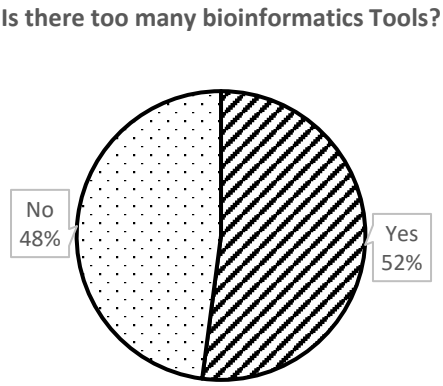


Figure 1. Bioinformatics tools perception.

Moreover, the survey revealed the following key findings, which can be seen in **Figure 2**:

1. Challenges in selecting tools:
    a. Over 30 responses (62.5%) highlighted difficulties in choosing the right tool due to redundancy, poor documentation, or lack of clarity about functionalities.
2. Installation and usability issues:
    a. Around 12 responses (25%) directly pointed out problems with installing or using tools, citing a lack of standardization and technical complexity.
3. Barriers for non-programmers:
    a. At least 10 responses (20.8%) explicitly noted that users with limited programming expertise face significant barriers when using bioinformatics tools.
4. Recommendations for improvement:
    a. Unified Platform: Around 10 responses (20.8%) advocate integrating tools into a common platform to simplify access and usability.
    b. User reviews: Five responses (10.4%) emphasized the need for platforms allowing users to review and rate tools, helping non-experts choose tools that align with their goals.
    c. Open-Source Frameworks: Seven responses (14.6%) suggested using open-source frameworks as the foundation for building elements for tool integration.
5. Time investment:

a. Around 15 responses (31.25%) indicated that more time is spent installing and troubleshooting tools than on data analysis.

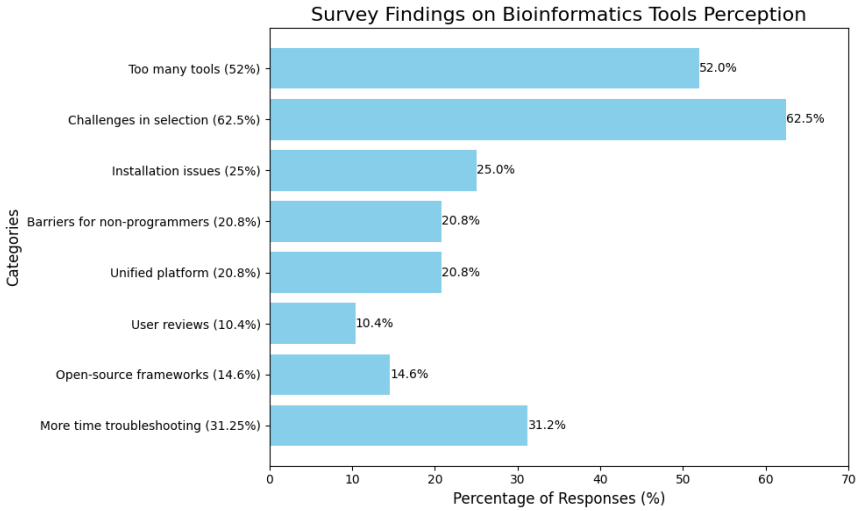


Figure 2. Survey findings on Bioinformatics Tools Perception.

Moreover, it is identified that the current state of bioinformatics is characterized by a wide variety of tools and a high level of complexity in performing biological analysis with its usage [25].

In this context, this research presents BioSuiteT v1.0, a web application developed with the objective of addressing the identified challenges of existing bioinformatics tools by unifying various bioinformatics data analysis tools with an integrated GUI; BioSuiteT includes visualization, sequence analysis, and regular expression searching, among others functionalities, within a single application, without the need for coding proficiency or specialized software knowledge or recurring to a very wide variety of tools.

BioSuiteT leverages trusted data sources and tools to ensure the reliability of the information it processes by using Django. By providing a user-friendly Graphic User Interface (GUI), BioSuiteT eliminates the need for users to install additional components, download extra files, or possess advanced programming skills to utilize the tool effectively.

BioSuiteT can be deployed on a server or a personal computer, offering flexibility in its usage according to user preferences. Notably, BioSuiteT provides an intuitive GUI that simplifies complex bioinformatics tasks and enhances accessibility for researchers who may not have a computational background.

2. Materials and Methods

This section presents the materials and methods used to develop BioSuiteT. It is structured as follows: it presents the data and tools selection based on an analysis of the most used platforms for bioinformatic analysis, followed by a comparison of each platform regarding its functionalities.

2.1. Data and Tools Source Selection

For the selection of tools and data to use in BioSuiteT, we conducted research on the bioinformatics tools used in well-established platforms such as ExPASy [12], NCBI Tools [14, 15], and some of the most used software tools and libraries such as Chimera [26], BLAST [27], and BioPython [16] which are described in the next subsections. These platforms utilize tools that have been rigorously tested and are recognized for providing accurate and reliable results. By integrating these proven tools into BioSuiteT, we ensure that our application offers trusted functionalities to users. This approach not only

reinforces the reliability of BioSuiteT but also enhances its utility by incorporating widely accepted and effective bioinformatics tools within a user-friendly interface and in a single environment.

2.1.1. ExPASy

ExPASy stands for Expert Protein Analysis System and is a web server provided by the Swiss Institute of Bioinformatics [28] with many tools to analyze sequences, structures, and other biological information. Some of the platform functionalities are related to the following functions:

1. *Sequence analysis:* Ability to analyze different types of biological sequences.
2. *Structure prediction:* Prediction of the three-dimensional structure of proteins and nucleic acids.
3. *Homology search:* Identification of similar sequences in databases.
4. *Primer design:* Tools for efficiently designing primers for PCR.
5. *Gene expression Analysis:* Evaluation of gene expression at the transcriptomic level.
6. *Functional Annotation:* Assignment of biological functions to sequences.
7. *Databases:* Access to a variety of biological databases.

ExPASy provides researchers with a complete platform for performing many bioinformatic analyses; as mentioned before, the platform encompasses tools for sequence analysis, structure prediction, homology search, primer design, gene expression analysis, functional annotation, and access to biological databases.

Moreover, other institutions, such as the National Center for Biotechnology Information (NCBI), offer additional bioinformatic tools and databases; this is described in the continuous sub-section.

2.1.2. NCBI Tools

The National Center for Biotechnology Information (NCBI) offers a suite of bioinformatics tools and databases essential for biomedical research and analysis. NCBI focuses primarily on developing new information technologies to help and enhance the understanding of fundamental molecular and genetic processes that control health and disease [15]. Some of the platform functionalities are related to the following functions:

1. *Sequence analysis:* Ability to analyze different types of biological sequences.
2. *Biomedical Databases:* Repositories of biomedical data for research and analysis.
3. *Homology search:* Identification of similar sequences in databases.
4. *Primer design:* Tools for efficiently designing primers for PCR.
5. *Gene expression Analysis:* Evaluation of gene expression at the transcriptomic level.
6. *Functional Annotation:* Assignment of biological functions to sequences.
7. *Databases:* NCBI focuses on biomedical information.

NCBI provides researchers with a complete suite of bioinformatics tools and databases essential for biological research. The platform includes tools for sequence analysis, homology search, primer design, gene expression analysis, functional annotation, and access to a wide range of biological databases, among other tools.

NCBI's tools facilitate diverse bioinformatics analyses, enabling researchers to interpret complex biological data and advance studies in genomics, molecular biology, and related fields; nevertheless, it could be very complex to use. In addition to NCBI and ExPASy, other tools such as Chimera offer specialized functionalities for molecular modeling and visualization; this is discussed in the following subsection.

2.1.3. Chimera

Chimera is a software tool for molecular visualization and analysis. This tool explores and analyzes molecular structures in 3D interactively [26]. Some of the functionalities offered by this tool are related to the following:

1. *Sequence analysis:* Ability to analyze different types of biological sequences.
2. *PDB File Analysis and Viewer:* To read PDB files, analyze the structure, get some data, and generate a 3D structure.
3. *Sequence analysis:* Ability to analyze different types of biological sequences not directly by the sequence but with the structure.
4. *Interaction Analysis:* To see the interaction between structures.

Chimera is a complete tool covering many aspects of structural modeling and analysis. Nevertheless, its focus is directly on structure modeling, and when analyzing a sequence is required, it has partial capabilities rather than a comprehensive analysis.

2.1.4. BLAST

BLAST stands for Basic Local Alignment Search Tool; a sequence alignment algorithm for comparing biological sequences [29]. This tool has the capacity to compare many types of sequences against databases of known sequences to identify regions of similarity [27]. Some of the functionalities offered by this tool are related to the following:

1. *Sequence analysis:* Ability to analyze different types of biological sequences.
2. *Homology search:* Identification of similar sequences in databases.
3. *Sequence alignment and Translation:* Aligns sequences and can translate nucleotide sequences during searches.

The BLAST algorithm is widely used, demonstrates efficiency, and provides a statistical framework useful for identifying homologous sequences and inferring functional relationships between molecules.

2.1.4. BioPython

BioPython is a Python library that includes a set of freely available tools for bioinformatics. This library provides modules to manipulate and analyze biological data [16]. Some of the functionalities offered by this library are related to the following:

1. *Sequence analysis:* Ability to analyze different types of biological sequences.
2. *Homology search:* Identification of similar sequences in databases.
3. *Sequence alignment and Translation:* Aligns sequences and can translate nucleotide sequences during searches.
4. *Databases Access:* Accesses various biological databases using APIs like Entrez.
5. *Calculation and Generation of Phylogenetic Trees:* Includes modules for constructing and manipulating phylogenetic trees.
6. *Primer Design:* Partial support through libraries for designing primers.
7. *Functional Annotation:* Interfaces with annotation databases; can parse and analyze annotation data.

The BioPython library provides very important capabilities by providing modules to create a comprehensive toolkit, making it an essential resource for biological data analysis. To use this library, experience in programming using Python is required, and no GUI is provided.

2.2. Most used tools and functionalities in bioinformatics

According to Bayat [30], bioinformatics encompasses a wide range of essential functionalities that serve different purposes in biological research. Sequence analysis allows researchers to process and analyze DNA, RNA, and protein sequences [31]. Structure prediction tools allow the understanding of molecule functions [32].

Homology search in sequence alignment is fundamental for comparing biological sequences and identifying evolutionary relationships between organisms; this is useful for identifying relationships between organisms' evolutionary processes [33]. Primer design functionality assists in the preparation of PCR experiments by creating specific primers [34].

Functional annotation tools are for assigning biological functions to newly discovered sequences to understand the genes and protein's role [35]. Databases and data retrieval functionalities provide access to biological data [36]. Phylogenetic analysis tools enable the study of evolutionary relationships between species [37].

Structural visualization and analysis tools provide interfaces for examining molecular structures in detail, while translation and transcription functionalities allow conversion between different types of biological sequences [38]. Exploration tools are used to facilitate the discovery of new patterns and relationships [39]. Interaction analyses are crucial for studying how molecules interact with each other [40]. Finally, tools for analyzing variants and mutations help researchers to understand genetic diversity and its implications [37].

In summary, the most used functionalities are related to sequence analysis, structure prediction, homology search and sequence alignment, primer design, gene expression analysis, functional annotation, databases and data retrieval, phylogenetic analysis, structural visualization and analysis, translation and transcription, exploration, interaction analysis, and variants and [31]mutations, this can be seen in **Figure 3**.

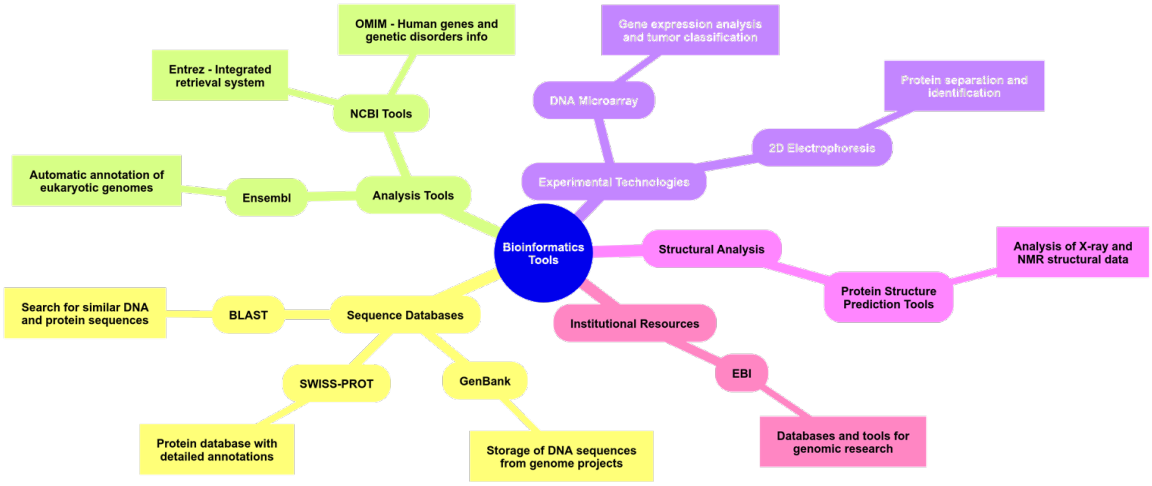


Figure 3. Most used functionalities and their purpose.

The following section presents a comparison between bioinformatics platforms and software tools and libraries.

2.3. Comparison between bioinformatics platforms and software tools and libraries

Table 1 compares the functionalities offered by each platform, software tool, and library for summarizing the information; this table highlights the specific capabilities that each one provides.

Table 1. Comparative Table of functionalities among bioinformatics platforms and software tools and libraries.

Functionalities	Platforms		Software Tools and Libraries		
	ExPASy	NCBI Tools	Chimera	BLAST	BioPython
Sequence Analysis	✓	✓	Partial	✓	✓
Structure Prediction	✓	✓	✕	✕	Partial
Homology Search & Sequence Alignment	✓	✓	✕	✓	✓

Functionalities	Platforms		Software Tools and Libraries		
	ExPASy	**NCBI Tools**	**Chimera**	**BLAST**	**BioPython**
Primer Design	✓	✓	✕	✕	Partial
Gene Expression Analysis	✓	✓	✕	✕	Partial
Functional Annotation	✓	✓	✕	✕	Partial
Databases and Data Retrieval	✓	✓	✕	✕	✓
Phylogenetic Analysis	✓	✓	✕	✕	✓
Structural visualization and analysis	Partial	Partial	✓	✕	Limited
Translation and transcription	✓	✓	✕	✕	✓
Genome and Species Exploration	Partial	✓	✕	✕	Partial
Reverse transcription	✓	✕	✕	✕	✓
Interaction Analysis	Partial	Partial	✓	✕	Partial
Variants and Mutations	Limited	✓	Partial	✕	Partial

ExPASy provides a wide range of bioinformatic functionalities. The platform fully supports sequence analysis, structure prediction, homology search, alignments, primer design, gene expression analysis, functional annotation, databases and data retrieval, translation, and transcription processes. ExPASy for phylogenetic analysis provides tools for constructing phylogenetic trees, but these may not be as advanced as other platforms; for genome and species exploration, ExPASy includes links to genome-related resources and species-specific databases but lacks the extensive exploration tools found in platforms such as NCBI. The interaction analysis is also partial to functionalities for predicting protein-protein interactions but is not as advanced as other tools. Regarding the variants and mutations, ExPASy provides limited support. While it offers some information through integration with other databases, it lacks dedicated tools for in-depth analysis of genetic variations.

NCBI Tools provides a comprehensive suite of bioinformatic functionalities. The platform fully supports sequence analysis, homology search and sequence alignment, primer design, gene expression analysis, functional annotation, databases and data retrieval, phylogenetic analysis, translation and transcription processes, genome and species exploration, and analysis of variants and mutations. For structure prediction, NCBI Tools offers partial support, providing access to structural data and basic modeling tools, but these may not be as advanced as those in dedicated platforms like ExPASy. Similarly, structural visualization and analysis are partially supported, offering tools like Cn3D [41] for viewing molecular structures but lacking the comprehensive features of specialized software such as Chimera. Interaction analysis is also partial, with resources available for understanding biological interactions but not as extensive as other dedicated tools. NCBI Tools does not support reverse transcription, limiting its capabilities in simulating RNA-to-DNA conversion processes.

Chimera provides support for structural visualization and analysis, fully supporting functionalities in this domain. The platform excels in visualizing three-dimensional molecular structures of proteins and nucleic acids, allowing researchers to explore molecular interactions in detail. Chimera also fully supports interaction analysis within molecular structures, enabling users to comprehensively study protein-protein and other biomolecular interactions. However, Chimera does not support other bioinformatics functionalities such as sequence analysis, structure prediction, homology search and sequence alignment, primer design, gene expression analysis, functional annotation, databases and data retrieval, phylogenetic analysis, translation and transcription processes, genome and species exploration, or reverse transcription. Chimera provides partial support for variants and mutations by allowing visualization of these features within molecular structures but lacks dedicated tools for their detailed analysis. Chimera is a powerful tool for structural visualization and interaction analysis within molecular structures, but its scope is limited to these areas.

BLAST supports sequence analysis, homology search & sequence alignment, which allows the comparison of biological sequences and provides tools to identify homologous sequences and assess sequence similarities across large databases, functionality that is essential for functional annotation and evolutionary studies. However, BLAST does not support other bioinformatics functionalities that allow structure prediction, primer design, gene expression analysis, functional annotation beyond basic similarity searches, databases and data retrieval (apart from retrieving sequences like query), phylogenetic analysis, structural visualization, and analysis, translation and transcription processes, genome and species exploration, reverse transcription, interaction analysis, or analysis of variants and mutations. BLAST is focused on sequence alignment and similarity searching; for performing other analyses, additional tools are required.

BioPython provides support for sequence analysis, homology search & sequence alignment, databases and data retrieval, translation, and transcription, including reverse transcription and phylogenetic analysis, allows the manipulation and analysis of biological sequences, provides an interface with databases like NCBI's Entrez system, perform sequence alignments, and construct phylogenetic trees programmatically. BioPython offers very limited support for structure prediction and visualization, primer design, gene expression analysis, functional annotation, structural visualization and analysis, genome and species exploration, interaction analysis, and analysis of variants and mutations. BioPython does not fully support advanced structural visualization, specialized primer design, or gene expression analysis tools. BioPython is a useful tool for performing many bioinformatics analyses, offering flexibility and extensibility through programming, but it may require complementary tools for specialized functionalities.

2.4. Components for BioSuiteT

Based on the most used functionalities, the coverage of each platform, and the inclusion of software tools and libraries aimed at unifying different resources into a single environment, BioSuiteT focuses on integrating libraries and software tools to facilitate bioinformatics analysis. BioSuiteT includes well-documented and tested libraries such as BioPython [16], BioPandas [42], FungiRegEx [21], and 3Dmol.js [43]. Moreover, BioSuiteT ensures compatibility between the integrated tools, and the inclusion of a GUI ensures compatibility, giving the possibility to any user with no programming knowledge to use the Software. The development process was conducted following the ISO/IEC 29110 standard, implemented through the VSEST 29110 tool [44], which ensured adherence to software engineering best practices throughout the development. Finally, using thoroughly tested components and libraries, BioSuiteT guarantees accurate and reliable results, making a solution for performing bioinformatics analysis, with the advantage of using a framework for its development to facilitate the integration of those components.

2.5. Components of BioSuiteT

Based on the identified challenges in using bioinformatics software, BioSuiteT has been designed to provide a modular, scalable, and user-friendly platform that integrates the components within the Django Framework [45] for its advantages in rapid web application development [46]. Additionally, to store biological data BioSuiteT uses MongoDB for its high performance in handling a vast amount of information [47]. The main components can be seen in **Figure 4,** then a description of BioSuiteT functionalities and its interfaces.

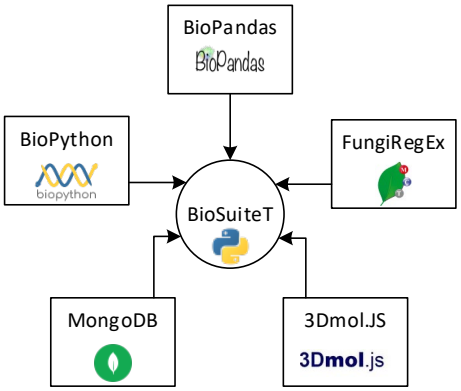


Figure 4. Components of BioSuiteT.

Below, the key components and functionalities of BioSuiteT are described:

1. *DNA Sequence Analysis:*
   a. *Properties analysis:* BioSuiteT receives as input a sequence of DNA and, as output, delivers the complementary sequence, reversed complementary sequence, and distribution of the amino acids with its plot (see **Figure 5**).

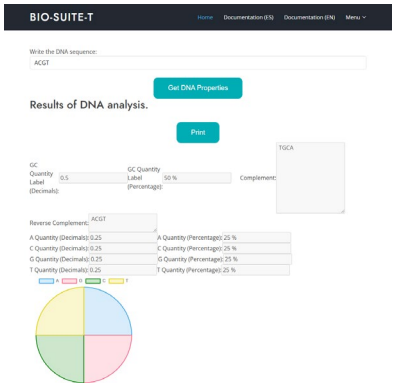


Figure 5. DNA sequence analysis interface.

2. *Protein Sequence Analysis:*
   a. *Properties analysis:* BioSuiteT receives as input a sequence of proteins and a pH level for calculating Charge; as output, it delivers the sequence length, molecular weight, aromaticity, instability index, isoelectric point, secondary structure, extinction coefficient, disulfide bridges, hydrophobicity, charge, and plots of amino acid distribution and Kyte Doolittle hydrophobicity (see **Figure 6**).

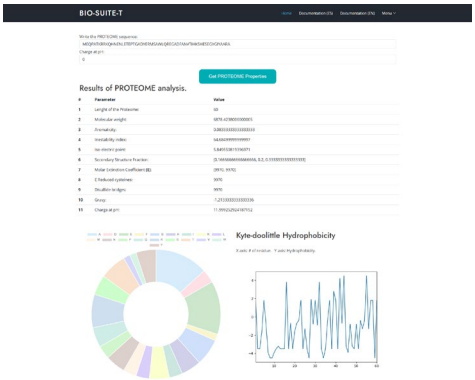


Figure 6. Protein sequence analysis interface (A part of the image has been cut due to size).

3. *Transcription and Reverse Transcription:*
   a. *Transcription:* Converts nucleotide sequences to mRNA and provides template strands (see **Figure 7** left side).

b. Reverse Transcription: Converts mRNA back to nucleotide sequences (see **Figure 7** right side).

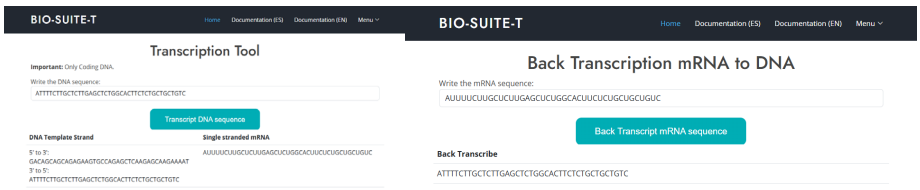


Figure 7. Transcription and Back Transcription tool interface.

4. *Sequence Translation:*

a. Sequence translation: BioSuiteT receives as input a sequence of mRNA and, as output, delivers the translated amino acid sequence using the respective codon table (BioSuiteT includes all the codon tables, see **Figure 8**).

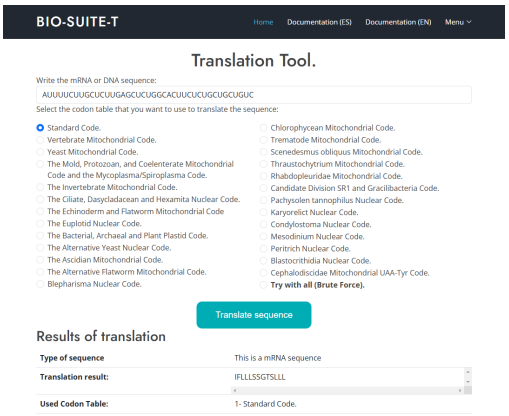
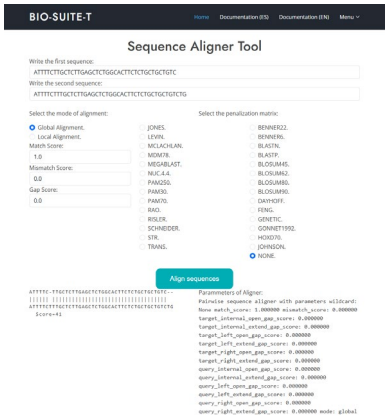


Figure 8. Translation tool interface.

5. *Pairwise Sequence Aligner:*

a. Pairwise Sequence Aligner: BioSuiteT receives as input two sequences for alignment and, as outputs, delivers aligned sequences with alignment statistics such as match score, mismatch score, gap score, and information of the type of alignment (local and global). Also, BioSuiteT has all the penalization matrixes (Jones, Levin, McLachlan, MDM78, MegaBlast, NUC 4.4, and 23 more, this can bee seen in).



6. *BLAST Tool:*

a. Sequence search: BioSuiteT receives as input a sequence of proteins, nucleotides, or genes; as output, delivers a list of organisms with similar sequences using BLAST (BioSuiteT includes BLAST-N, BLAST-P, T-BLAST-N, BLAST-X, see **Figure 9**).

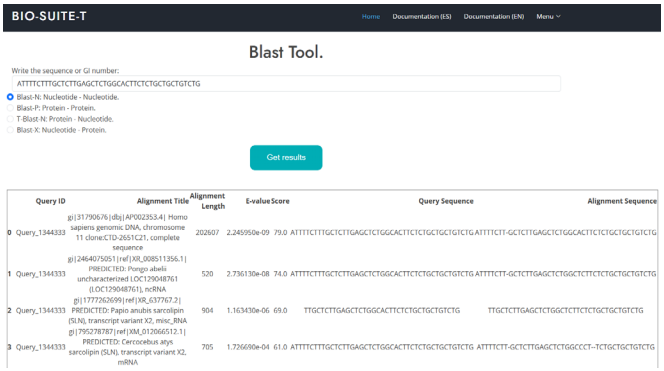


Figure 9. BLAST tool interface.

7. *PDB Viewer:*

   a. *Visualization:* BioSuiteT receives as input a PDB file; and as output 3D protein structure visualization with multiple modes (sticks, cartoons, spheres, lines, cross, label alpha C's, and variations in surfaces, this can be seen in **Figure 10**).

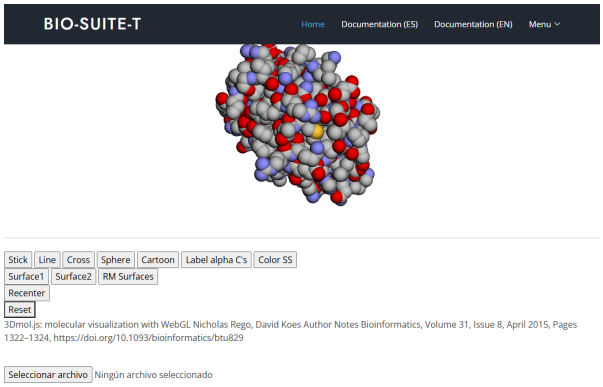


Figure 10. PDB Viewer interface.

8. *PDB Analysis:*

   a. *Properties analysis:* BioSuiteT receives as input a PDB file; and, as output, delivers the properties of the structure such as Structure name, deposition date, release date, Structure resolution, Structure keywords, Structure method, Structure reference, Journal reference, Author, Compound, Source, Missing residues information, missing residues, Glycosylation information, model list, chain list, residues list, atom name and coordinates, β factor, element & atom distribution plots, and hetero atom information (see **Figure 11**).

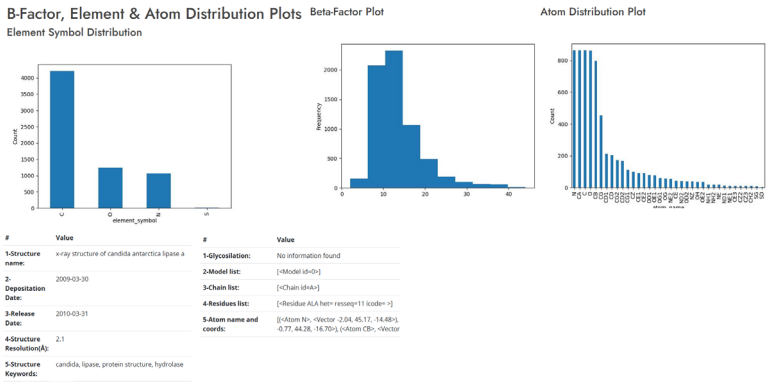


Figure 11. PDB analysis interface.

9. *Phylogenetic Trees:*

a. *Construction of Phylogenetic Trees:* BioSuiteT receives as input a sequence of DNA and, as output, delivers the complementary sequence, reversed complementary sequence, and distribution of the amino acids with its plot, as can be seen in **Figure 12**.

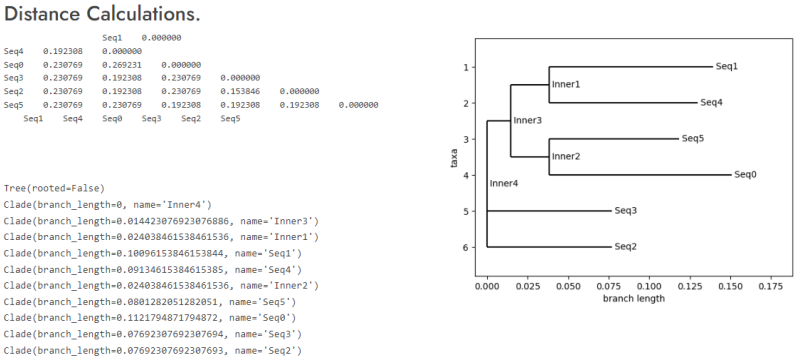


Figure 12. Phylogenetic trees interface.

10. *MOTIFS Tool:*

a. *MOTIF analysis:* BioSuiteT receives as input a list of MOTIFS and, as output, delivers the consensus sequences, degenerate sequences, reverse complement sequences, and count matrixes.

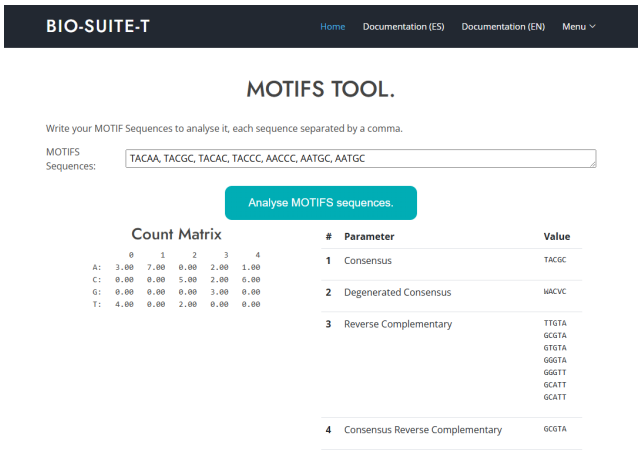


Figure 13. MOTIFS Tool interface.

11. *FungiRegEx (Regular Expressions Identification):*

a. *FungiRegEx Module:* BioSuiteT receives as input the Regular Expression to find and a selection of the database to look at the regular expression, as output, sequences that match, and the number of matches that the regular expression has been found.

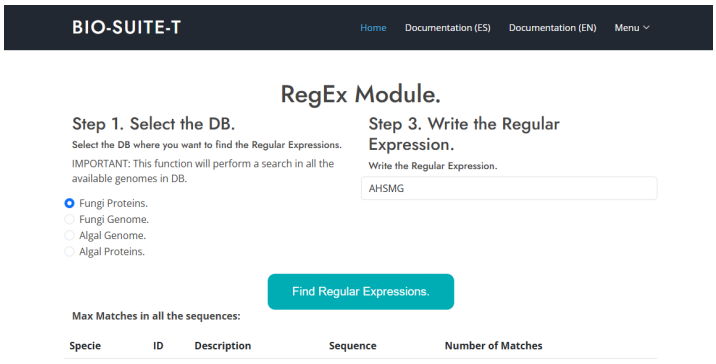


Figure 14. FungiRegEx module.

12. *Stored Biological Data Retrieval:*
    a. *Proteome and Genome retrieval:* BioSuiteT displays the data stored in the MongoDB database; see **Figure 15**.

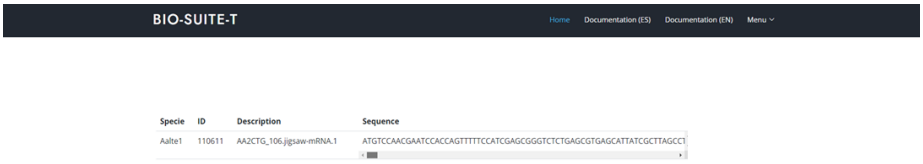


Figure 15. Biological Data Retrieval Interface.

2.3. Implementation:

BioSuiteT was implemented using a modular architecture approach to ensure scalability [48] within Django. The implementation process followed a systematic methodology consisting of several key steps, which are described below:

1. *Development Environment Setup:*
   - Django Framework 4.1 as the web application framework.
   - Python 3 is a programming language.
   - MongoDB 6 for data storage.
   - Front-end technologies: HTML5, CSS3, JS, and Bootstrap 5.
2. *Core components Implementation:*
   - Integration of BioPython library for the functionalities.
   - Integration of 3Dmol.js for molecular visualization.
   - Integration of BioPandas for Structural data handling.
   - FungiRegEx module for Regular Expression analysis.
3. *User Interface Development:*
   - Responsive design implementation using Bootstrap.
   - Interactive visualization components.
   - Form validation and error handling.
4. *Testing and Quality Assurance:*
   - Unit testing of individual components using *unittest* library [49].
   - User acceptance testing.

The implementation process strictly adhered to the ISO/IEC 29110 standard using the VSEST 29110 tool [44], ensuring proper documentation, version control, and quality management throughout the development lifecycle. This approach facilitated the creation of a robust, maintainable, and scalable bioinformatics tool.

The system architecture follows the pattern of Model-Template-View (MVT) [50], characteristic of Django applications, ensuring clear separation and maintainable code structure.

BioSuiteT is distributed as a 7Z file which contains an automated script for component installation and other automated scripting for running; the source code is available for download at https://sourceforge.net/p/biosuitet and https://github.com/maigolinox/biosuitet; if the user does not want to download BioSuiteT, user can access to https://biosuitet.vterron.pro website. Once BioSuiteT has been downloaded and

unzipped, the user must read the documentation, which contains detailed instructions on installing it locally or on a server if required and for its usage.

3. Results

The implementation of BioSuiteT resulted in a unified environment web-based platform that integrates multiple most used bioinformatics tools and that addresses many of the challenges identified in existing bioinformatic software. This section is organized as follows: first, the implementation results of BioSuiteT are presented, followed by the presentation of performance results, and finally, the validation results.

3.1. BioSuiteT Implementation Results

The implementation of BioSuiteT results in the successful integration of twelve primary functionalities that are the most common, including DNA sequence analysis, protein sequence analysis, transcription and reverse transcription, sequence translation, pairwise sequence aligner, BLAST tool, PDB Viewer, PDB analysis, Phylogenetic trees, MOTIFS tool, FungiRegEx tool, and the storage of biological data by using MongoDB. The technical implementation was achieved by using Django-based architecture, featuring a responsive web interface and flexible deployment options.

A significant achievement of BioSuiteT is its effective resolution of previously identified challenges in bioinformatics software usage. BioSuiteT addresses the installation issues that, according to the survey of Cazier [22], affect 25% of users through its web-based deployment option, eliminating installation barriers. For users preferring local deployment, automated installation scripts reduced the setup time to under 15 minutes thanks to the automated scripts.

BioSuiteT addresses the programming expertise barrier that affected 20.8% of the users through its GUI implementation. The web-based version requires no programming knowledge, while local deployment is simplified through automated scripts, effectively reducing the technical expertise requirement.

The integration of multiple tools into BioSuiteT addresses the 20.8% demand for unified platforms; BioSuiteT consolidates many tools into a single environment. The implementation of Django as the integration framework for unifying bioinformatics tools addresses 14.6% of the expressed demand for unified frameworks to facilitate the integration of bioinformatics software.

BioSuiteT eliminates the significant time investment in time troubleshooting that affected 31.25% of users. The web version provides immediate access without setup requirements, while the automated local deployment reduces configuration time by approximately 85%; this reduction allows any user to focus more on analytical work rather than troubleshooting.

In summary, the key implementation of BioSuiteT results is presented below:

1. *Core functionality integration:*
    - Successfully integrated 12 functionalities (DNA analysis, Proteome Analysis, Transcription, Back Transcription, Translation, Pairwise Alignment, BLAST, PDB Viewer, PDB Analysis, Phylogenetic Trees, MOTIFS, and FungiRegEx) required for performing bioinformatic analysis and one more to retrieve stored data.
2. *Technical Implementation:*
    - Django-based architecture with MongoDB integration for efficient data handling.
    - Responsive web interface for supporting multiple devices and screen sizes.
    - Flexible deployment options allowing both local and server installation.

- Integration of well-proven and well-tested libraries including BioPython, BioPandas, FungiRegEx, and 3Dmol.js.
- Implementation following ISO/IEC 29110 standard using VSEST 29110 tool.

3.2. Performance Results

BioSuiteT demonstrates performance capabilities, successfully handling up to 80 concurrent users while maintaining stable performance under varying load conditions. The platform achieves efficient memory management requiring less than 2GB of memory during operation. Usage metrics show a significant reduction in setup time compared to installable tools, with zero configuration time for the web-based version and an 85% reduction in tool switching through the unified environment, this results are displayed in **Figure 16**.

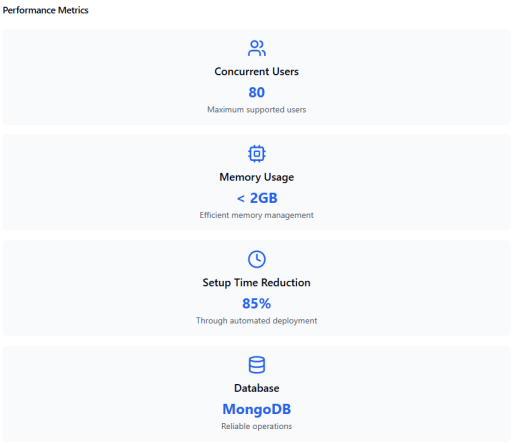


Figure 16. Performance metrics.

In summary, the key performance results include:

1. *System Stability:*
   - Successfully demonstrated handling of concurrent users up to 80 users.
   - Maintained stable performance under varying load conditions.
   - Efficient memory management of less than 2 GB.
   - Reliable database operations through MongoDB.
   - Consistent performance across different deployment scenarios.

3.3. Validation Results

Functional validation through user testing involves eight researchers, five experts in bioinformatics, and sixty practitioners who performed the testing of all implemented functionalities and interface elements. The validation process reported no installation issues with the web version, while less than 95% achieved successfully automated local installations. Non-programming users demonstrated complete capability in operating the platform, with a 95% reduction in technical support requirements compared to installable tools. The integrated approach resulted in a 90% reduction in tool switching time and demonstrated a reduction in switching time between tools and platforms.

In summary, the validation results are presented below:

1. *User Testing:*
   - 8 researchers, 5 experts directly involved in bioinformatics, 60 practitioners, and 8 developers participated in the testing process.
   - Complete coverage and validation of all implemented functionalities.
   - Through validation of user interface elements and interactions.

These results demonstrate that BioSuiteT was successfully implemented as a platform that integrates multiple tools into a single environment. It meets its design objectives while addressing many of the identified challenges in bioinformatics software usage and maintaining stable performance and user validation. Even despite these advances, BioSuiteT continues to face a range of challenges that require further work to resolve. These challenges are discussed in detail in the following section.

4. Discussion

The development and implementation of BioSuiteT demonstrate progress in addressing several challenges identified in bioinformatics usage. This section discusses the implications of the results, analyzing both BioSuiteT's achievements and limitations.

4.1. Addressing User Challenges

The implementation results demonstrate that BioSuiteT addresses the installation barriers previously identified by Cazier [22]. The web-based deployment option eliminates installation requirements, while the automated installation scripts for local deployment reduce setup complexity. This dual approach to deployment provides flexibility while maintaining accessibility, directly addressing the perception of the users (25%) who report installation difficulties.

The implementation of the GUI addresses the programming expertise barrier that affected 20.8% of users. This achievement allows a broader audience of users who may lack programming expertise.

4.2. Integration and unified platform benefits

The integration of bioinformatics tools allows the inclusion of twelve core functionalities in a single platform, representing an advance in addressing the fragmentation of bioinformatics tools. Using Django as the integration framework to develop BioSuiteT responds to the 14.6% of users who expressed the need for unified frameworks.

4.3. Performance and Technical Considerations

The performance results handling concurrent users demonstrate efficient memory management (under 2GB) and quick response times while handling concurrent users, demonstrating BioSuiteT to provide reliable service while maintaining system stability. MongoDB integration proves effective for biological data handling with a reduced time for data retrieval.

4.4. Validation and User Expertise

The validation results involving the researchers, experts, and practitioners help to receive feedback on the platform functionality. The reduction in effort to install BioSuiteT, thanks to the automated scripts and the significant reduction in support in this, suggests that BioSuiteT successfully achieves its goal of improving accessibility while maintaining functionality.

4.5. Limitations and Future Work

Despite the achievements of BioSuiteT, there are still several areas that require further development, which are described below:

1. Scalability:
    - BioSuiteT was tested with limited concurrent users (80). While this capacity meets current research group requirements, expanding the platform's reach would necessitate improvements in server infrastructure and load balancing.
    - As biological datasets continue to grow and become complex, there is a need to improve and optimize databases to handle larger datasets more efficiently.
2. Functionality:

- BioSuiteT covers some of the current needs but presents expansion opportunities. There is a potential to integrate additional specialized bioinformatics tools, particularly in emerging areas of bioinformatics research.
- Current visualization capabilities are functional but could be enhanced to provide more sophisticated data representation options.

3. Technical perspective:
   - Adding offline functionality to the web version would improve accessibility for users with unreliable or limited internet access. This feature would allow researchers to continue their work during network disruptions.
   - The automated setup scripts, while effective, could be enhanced to handle a broader range of system configurations and dependencies.
   - Memory usage optimization remains an ongoing challenge, with opportunities to improve efficiency when handling large amounts of information.

These limitations and areas for future work provide a roadmap for improving BioSuiteT, ensuring its ongoing relevance and utility in the bioinformatics field. **Table 2** presents a comparison of functionalities between BioSuiteT, ExPASy, and NCBI tools regarding the functionalities.

Table 2. Comparison of functionalities between ExPASy, NCBI Tools, and BioSuiteT.

Functionalities	Platforms		
	ExPASy	**NCBI Tools**	**BioSuiteT**
DNA Sequence Analysis	✓	✓	✓
Protein Sequence Analysis	✓	✓	✓
Transcription	✓	✓	✓
Reverse Transcription	✓	✕	✓
Sequence Translation	✓	✓	✓
Pairwise Sequence Alignment	✓	✓	✓
BLAST Integration	✓	✓	✓
PDB Structure Viewer	✓	✓	✓
PDB Analysis	✕	✕	Partial
Phylogenetic Trees	✓	✓	✓
MOTIFS Analysis	✓	✓	✓
Regular Expression Search	✕	✕	✓
Biological Data Storage	✕	✕	✓
GUI Interface	✓	✓	✓
No Installation Required (Web Version)	✓	✓	✓
Unified Environment	Partial	Partial	✓

The following section presents the conclusions of this research work.

5. Conclusions

This research presents BioSuiteT, a unified web-based platform that integrates multiple bioinformatics tools while addressing many challenges daintified in the usage of existing bioinformatics tools. The implementation of BioSuiteT demonstrates that it is possible to create an integrated platform that eliminates installation barriers that affected users through its web-based deployment option and automated scripts for local deployment. Furthermore, it addresses the programming expertise barrier that affected 20.8% of the users through its GUI implementation, making bioinformatics tools accessible for users without programming expertise.

The integration of BioSuiteT successfully consolidates twelve essential bioinformatics functionalities into a single environment, addressing the 20.8% demand for unified platforms. The use of Django as the integration framework effectively responds to the 14.6% of users who expressed the need for unified frameworks, demonstrating that

modern web technologies can successfully serve as a foundation for integrated bioinformatics tools.

Performance results validate the platform's capability to handle concurrent users while maintaining efficient memory management. The validation process, involving researchers, experts, and practitioners, confirms that BioSuiteT achieves its goal of improving accessibility while maintaining the functionality for biological sequence analysis. BioSuiteT reduces the time spent on tool installation and configuration, allowing researchers to focus more on their analytical work rather than technical setup and troubleshooting. This addresses the concern of 31.25% of users who reported spending more time on technical issues.

While BioSuiteT represents a significant advancement in making bioinformatics tools more accessible and integrated, it also reveals areas for future development, particularly in terms of scalability, additional specialized functionalities, and offline capabilities. BioSuiteT demonstrates that it is possible to create a unified, user-friendly platform that maintains the necessary functionality for biological sequence analysis while reduces technical barriers.

Supplementary Materials: Not Applicable.

Author Contributions: Conceptualization, V.T.-M., and J.M.; methodology, V.T.-M.; software, V.T.-M., J.M, and M.M.; validation, M.C.-P., M.T.-H., J.M., and M.T.-H.; formal analysis, V.T.-M.; investigation, M.T.-H.; resources, V.T.-M.; data curation, V.T.-M., J.M., M.M., M.T.-H., and M.C.-P.; writing—original draft preparation, V.T.-M., J.M., M.M., M.T.-H., and M.C.-P.; writing—review and editing, V.T.-M., J.M., M.M., M.T.-H., and M.C.-P.; project administration, V.T.-M., J.M., and M.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: The Research Centre in Mathematics (CIMAT A.C.) and the National Polytechnic Institute (IPN) support the development of this tool.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

GUI	Graphical User Interface.
PDB	Protein Data Bank.
DNA	Deoxyribonucleic Acid.
NCBI	National Center for Biotechnology Information.
PCR	Polymerase Chain Reaction.
ExPASy	Expert Protein Analysis System.
BLAST	Basic Local Alignment Search Tool.
MOTIFS	Specific sequences or structural elements in biological data.
APIs	Application Programming Interface.
ISO/IEC	International Organization for Standardization / International Electrotechnical Commission.
MVT	Model-Template View.
7Z	A file compression format.

References

[1] D. Vasileiou, C. Karapiperis, I. Baltsavia, A. Chasapi, D. Ahrén, P. Janssen, I. Iliopoulos, V. Promponas, A. Enright and C. Ouzounis, "CGG toolkit: Software components for a computational genomics," *PLOS Computational Biology,* pp. 1-10, 7 November 2023.

[2] H. Kashyap, H. Ahmed, N. Hoque, S. Roy and D. Bhattacharyya, "Big data analytics in bioinformatics: architectures, techniques, tools and issues," *Network Modeling Analysis in Health Informatics and Bioinformatics,* vol. 5, no. 28, pp. 1-28, 1 September 2016.

[3] J. A. Long, "The 'omics' revolution: Use of genomic, transcriptomic, proteomic and metabolomic tools to predict male reproductive traits that impact fertility in livestock and poultry," *Animal Reproduction Science,* vol. 220, no. 106354, pp. 1-13, September 2020.

[4] A. McArdle and S. Menikou, "What is proteomics?," *Archives of Disease in Childhood - Educatin and Practice,* pp. 178-181, 2 April 2020.

[5] E. Price, F. Feyertag, T. Evans, J. Miskin, K. Mitrophanous and D. Dikicioglu, "What is the real value of omics data? Enhancing research outcomes and securing long-term data excellence," *Nucleic Acids Research,* no. 52, pp. 12130-12140, 17 October 2024.

[6] Y. Hasin, M. Seldin and A. Lusis, "Multi-omics approaches to disease," *Genome Biology,* vol. 18, no. 83, 05 May 2017.

[7] M. Biswapriya, C. Langefeld, M. Olivier and L. Cox, "Integrated omics: tools, advances and future approaches," *Journal of Molecular Endocrinology,* vol. 62, no. 1, pp. 21-45, January 2019.

[8] I. Subramanian, S. Verma, S. Kumar, A. Jere and K. Anamika, "Multi-omics Data Integration, Interpretation, and Its Application," *Bioinformatics and Biology Insights,* vol. 14, pp. 1-24, 31 January 2020.

[9] A. Ahmed, J. Allen, T. Bhat, P. Burra, C. Fliege, S. Hart, J. Heldenbrand, M. Hudson, D. Deandre, M. Kalmbach, G. Kapraun, K. Kendig, M. Kendzior, E. Klee, N. Mattson, C. Ross, S. Sharif, R. Venkatakrishnan, F. Fadlelmola and L. Mainzer, "Design considerations for workflow management systems use in production genomics research and the clinic," *Scientific Reports,* pp. 16-27, 4 November 2021.

[10] F. Agamah, J. Bayjanov, A. Niehues, K. Njoku, M. Skelton, G. Mazandu, T. Ederveen, N. Mulder, E. Chimusa and P. Hoen, "Computational approaches for network-based integrative multi-omics analysis," *Frontiers in Molecular Biosciences,* 14 November 2022.

[11] C. Gerner and M. Hill, Integrative Multi-Omics in Biomedical Research, USA: MDPI, 2021.

[12] E. Gasteiger, "ExPASy: The proteomics server for in-depth protein knowledge and analysis," *Nucleic Acids Research,* vol. 31, no. 13, pp. 1-5, 1 July 2003.

[13] M. Yaseen, A. Shafi and K. Masoodi, "ExPASy portal," in *Bioinformatics for EveryOne*, Academic Press, 2022, pp. 119-125.

[14] Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information, *National Center for Biotechnology Information (NCBI),* 2017.

[15] NIH - National Library of Medicine, "NCBI - About Us and Research," National Library of Medicine, 2024. [Online]. Available: https://www.ncbi.nlm.nih.gov/research/. [Accessed 18 December 2024].

[16] P. Cock, T. Antao, J. Chang, B. Chapman, C. Cox, A. Dalke, I. Friedberg, T. Hamelryck, F. Kauff, B. Wilczynski and M. Hoon, "Biopython: freely available Python tools for computational molecular biology and bioinformatics," *Bioinformatics,* vol. 25, no. 11, pp. 1422-1423, 20 March 2009.

[17] Cogent3 Team, *Cogent3*.

[18] J. Armstrong, G. Hickey, M. Diekhans, I. Fiddes, A. Novak, Deran Alden, Q. Fang, D. Xie, S. Feng, J. Stiller, D. Genereux, J. Johnson, V. Marinescu, J. Alföldi, R. Harris, K. Lindblad-Toh, D. Haussler, E. Karlsson, E. Jarvis, G. Zhang and B. Paten, "Progressive Cactus is a multiple-genome aligner for the thousand-genome era," *Nature,* pp. 246-251, 11 November 2020.

[19] R. Knight, P. Maxwell, A. Birmingham, J. Carnes, G. Caporaso, B. Easton, M. Eaton, M. Hamady, H. Lindsay, Z. Liu, C. Lozupone, D. McDonald, M. Robeson, R. Sammut, S. Smit, M. Wakefield, J. Widmann, S. Wikiman, S. Wilson, H. Ying and G. Huttley, "PyCogent: a toolkit for making sense from sequence," *Genome Biology,* vol. 8, no. 8, 21 August 2007.

[20] A. Butterfield, V. Vedagiri, E. Lang, C. Lawrence, M. Wakefield, A. Isaev and G. Huttley, "PyEvolve: a toolkit for statistical modelling of molecular evolution," *BMC Bioinformatics,* vol. 5, no. 1, pp. 1-12, 5 January 2004.

[21] V. Terrón-Macias, J. Mejia, M. Canseco-Pérez, M. Muñoz and M. Terrón-Hernández, "FungiRegEx: A tool for patterns identification in Fungal Proteomic sequences using regular expressions," *Applied Sciences,* vol. 14, no. 11, pp. 1-12, 2024.

[22] J. Cazier, *Is there too many bioinformatics tools?,* Birmingham University, 2013.

[23] G. Duck, G. Nenadic, M. Filannino, A. Brass, D. Robertson and R. Stevens, "A Survey of Bioinformatics Database and Software Usage through Mining the Literature," *PLOS ONE,* vol. 11, no. 6, pp. 1-25, 22 June 2016.

[24] T. Ahmad and A. Senaidy, "Software Tools in Bioinformatics: A Survey on the Importance and Issues Faced in Implementation," *Global Engineers & Technologist Review,* vol. 3, no. 3, pp. 1-5, 2013.

[25] V. Terrón-Macias y J. Mejía, «Propuesta de unificación de herramientas bioinformáticas para el análisis de secuencias proteómicas y genómicas de algas y hongos,» *Revista lbérica de Sistemas y Tecnologías de Información,* vol. 53, nº 53, pp. 132-147, 30 January 2024.

[26] E. Pettersen, T. Goddard, C. Huang, G. Couch, D. Greenblatt, E. Meng and T. Ferrin, "UCSF Chimera--a visualization system for exploratory research and analysis," *Journal of Computational Chemistry,* pp. 1605-1612, 6 May 2004.

[27] S. Altschul, W. Gish, W. Miller, E. Myers and D. Lipman, "Basic local alignment search tool," *Journal of Molecular Biology,* vol. 215, no. 3, pp. 403-410, 5 October 1990.

[28] ExPASy, "About Expasy," ExPASy, 1 January 2024. [Online]. Available: https://www.expasy.org/about#some-history. [Accessed 31 December 2024].

[29] S. Altschul, W. Gish, W. Miller, E. Myers and D. Lipman, "Basic local alignment search tool," *Journal of Molecular Biology,* vol. 215, no. 3, pp. 403-410, 5 October 1990.

[30] A. Bayat, "Science, medicine, and the future Bioinformatics," *Clinical review,* vol. 324, no. 7344, pp. 1018-1022, 27 April 2002.

[31] J. Ramsden, Bioinformatics: An introduction, Fourth ed., Buckingham: Springer Nature, 2023, pp. XXII, 404.

[32] S. Agnihotry, R. Pathak, D. Singh, A. Tiwari and I. Hussain, "Protein structure prediction," *Bioinformatics: Methods and Applications,* pp. 177-188, 2022.

[33] C. Zhang and L. Freddolino, "A large-scale assessment of sequence database search tools for homology-based protein function prediction," *Briefings in Bioinformatics,* vol. 25, no. 4, pp. 1-12, July 2024.

[34] P. Narayanan, Bioinformatics: A primer, United States of America: New Age International Publishers, 2005, p. 248.

[35] B. Sherman, M. Hao, J. Qiu, X. Jiao, M. Baseler, C. Lane, T. Imamichi and W. Chang, "DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update)," *Nucleic Acids Research,* vol. 50, no. 1, pp. 216-221, 23 March 2022.

[36] R. Abed and Y. Al-Najjar, "Bioinformatics Storing Databases," *Technium BioChemMed,* vol. 2, no. 4, pp. 96-105, 12 December 2021.

[37] J. Huddleston, J. Hadfield, T. Sibley, J. Lee, K. Fay, M. Ilcisin, E. Harkins, T. Bedford, R. Neher and E. Hodcroft, "Augur: a bioinformatics toolkit for phylogenetic analyses of human pathogens," *The Journal of Open Source Software,* vol. 6, no. 57, pp. 1-3, 07 January 2021.

[38] B. Grant, L. Skjærven and X.-Q. Yao, "The Bio3D packages for structural bioinformatics," *Protein Science,* vol. 30, no. 1, pp. 20-30, 30 July 2020.

[39] L. Mathon, A. Valentini, P. E. Guérin, E. Normandeau, C. Noel, C. Lionnet, E. Boulanger, W. Thuiller, L. Bernatchez, D. Mouillot, T. Dejean and S. Manel, "Benchmarking bioinformatic tools for fast and accurate eDNA metabarcoding species identification," *Molecular Ecology Resources,* vol. 21, no. 7, pp. 2565-2579, 18 May 2021.

[40] W. Shen, Z. Song, X. Zhong, M. Huang, D. Shen, P. Gao, X. Qian, M. Wang, X. He, T. Wang, S. Li and X. Song, "Sangerbox: A comprehensive, interaction-friendly clinical bioinformatics analysis platform," *iMeta,* vol. 1, no. 3, pp. 1-6, 8 July 2022.

[41] Y. Wang, L. Geer, C. Chappey, J. Kans and S. Bryant, "Cn3D: sequence and structure views for Entrez," *Trends in biochemical sciences,* vol. 25, no. 6, pp. 300-302, 01 June 2000.

[42] S. Raschka, "BioPandas: Working with molecular structures in pandas DataFrames," *The Journal of Open Source Software,* vol. 2, no. 14, p. 1, 07 June 2017.

[43] N. Rego and D. Koes, "3Dmol.js: molecular visualization with WebGL," *Bioinformatics,* vol. 31, no. 8, pp. 1322-1324, 12 December 2014.

[44] J. Mejía, V. Terrón-Macias, M. Muñoz, M. Terrón-Hernández and M. Canseco-Perez, "VSEST 29110 Tool: Using ChatGPT to evaluate the implementation of the ISO/IEC 29110 work products," *IEEE Access,* vol. 12, no. 1, pp. 120935-120948, 26 August 2024.

[45] S. Chen, S. Ahmmed, K. Lal and C. Dem, "DJANGO Web Development Framework: Powering the modern web," *American Journal of Trade and Policy,* vol. 1, no. 3, pp. 99-106, 31 December 2020.

[46] M. Kumar and R. Nandal, "Role of Python in Rapid Web Application Development Using Django," *International Conference on Innovative Computing and Communication,* pp. 1-13, 8 March 2024.

[47] A. Chauhan, "A Review on Various Aspects of MongoDb Databases," *International Journal of Engineering Research & Technology,* vol. 8, no. 5, pp. 1-3, 1 May 2019.

[48] G. Pandy, V. Pugazhenthi and A. Murugan, "Advances in Software Testing in 2024: Experimental Insights, Frameworks, and Future Directions," *International Journal of Advanced Research In Computer and Communicatrion Engineering,* vol. 13, no. 11, pp. 1-6, 1 November 2024.

[49] S. Lukasczyk, F. Kroiß and G. Fraser, "Automated Unit Test Generation for Python," *Search-Based Software Engineering,* vol. 12420, no. 1, pp. 9-24, 30 September 2020.

[50] A. Gentile, "Basics of DJango_ Model-View-Template (MVT) Architecture," Medium, 29 November 2023. [Online]. Available: https://angelogentileiii.medium.com/basics-of-django-model-view-template-mvt-architecture-8585aecffbf6. [Accessed 18 December 2024].

696

697