Development of a bioinformatic platform for the efficient management and biotransformation of agro-industrial waste through microalgae cultivation

Vítor Silva[1], Filipe Maciel[2], and Pedro Geada[2]

[1] School of Engineering, University of Minho, Portugal
[2] Centre of Biological Engineering, University of Minho, Portugal

Abstract The increasing volume of agro-industrial byproducts represents a significant environmental challenge, with one-third of global food production being wasted annually. This study presents an innovative bioinformatic platform that optimizes the biotransformation of agro-industrial waste through targeted microalgae cultivation. The platform integrates databases of sugar compositions with enzymatic profiles of microalgae species, enabling optimal microalgae-waste substrate pairings. Using BioPython, NCBI, KEGG and UniProt databases, the methodology focuses on hydrolytic enzymes involved in the hydrolysis of key sugars (lactose, sucrose, fructose, maltose, starch and cellulose) across five algal phyla. The web-based Streamlit platform identified 59 algae species with confirmed enzymatic capabilities, retrieved 22 scientific articles, and established 11 algae-enzyme matches. This approach aligns with UN Sustainable Development Goals 2, 9, and 12, providing researchers and industry with a tool for sustainable waste management and high-value biocompound production.

1 Introduction

Food waste, which is defined as any food that is not eaten or waste produced during food preparation in homes or businesses, is becoming a bigger issue in many parts of the world. This problem has important ramifications from an ethical and economic standpoint in addition to an environmental one. There is a huge gap between the effective distribution of food and its production, as evidenced by the fact that so much food is wasted worldwide—enough to feed everyone who is malnourished. Food waste and world hunger are related [1].

The management of food waste is, therefore, a fundamental process to minimise the environmental impact of the waste generated by society. This process involves several stages, such as planning, collection, recycling, processing, and disposal of food waste, to reduce the amount of discarded food and promote its efficient reuse. The implementation of appropriate management practices not only reduces waste but also contributes to sustainability, helping to lessen the

pressure on natural resources and the energy required for the production of new food [1].

Given the problems with food waste, creative solutions that not only reduce waste but also enhance the value of the processes' byproducts are desperately needed. In this regard, growing microalgae in agro-industrial waste has become a novel and sustainable way to increase the value of waste from different industries. The cultivation of microalgae is a particularly promising method for adding value to waste, in comparison to techniques like anaerobic digestion, aerobic digestion, and thermal degradation are currently employed. This method provides a circular economy-based solution by producing a range of sustainable goods, including animal feed, pigments, lipids, food supplements, biofuels, and biofertilizers [2].

2 State of the art

2.1 Agro-industrial waste

One of the biggest problems facing the environment and the economy today is food waste. An estimated one-third of the world's food production is wasted, placing a heavy burden on the environment and leading to serious resource inefficiencies [3]. From the stage of cultivation to the point of final consumption, this waste happens at every stage of the production chain, impacting not only the availability of food but also the sustainability of food systems [4].

Agro-industrial waste includes a wider range of byproducts than just food-related waste. These wastes fall into three primary categories: hazardous wastes, non-recyclable/non-compostable wastes, and recyclable/compostable wastes [5]. For instance, wastes from pits and slaughterhouses are classified as secondary wastes, while recyclable wastes from animal manure and pruning are classified as primary wastes [5]. Hazardous wastes need to be managed strictly in accordance with regulations, and non-recyclable wastes from construction and agricultural mechanisation pose major management challenges [5]. Additionally, many agro-industrial wastes are disposed of by incineration or landfill, leading to environmental and socioeconomic problems [6], and their accumulation exceeds 2 billion tonnes globally, making it imperative to research their utilisation to maximise benefits [7].

The dairy industry is one of the most prominent examples. Whey and wastewater production have significantly increased as a result of the dairy industry's increased production due to rising demand for dairy products [8]. Because they reduce the amount of dissolved oxygen in water bodies, these wastes, which are distinguished by their high organic content, are harmful to the environment. Depending on the product type and operational procedures, dairy wastewater has high concentrations of organic components like lactose, fat, and whey protein [8].

In addition, the dairy industry contributes significantly to pollution through effluent emissions and waste disposal [9]. This problem is made worse by projections that the global dairy market will reach US$1.243 trillion by 2028, which is

predicted to result in an even higher rise in waste production [9]. Although it is one of the biggest polluters, especially when it comes to water consumption, the dairy industry plays a vital role in ensuring the world's food security by turning milk into necessary products [10]. Despite its widespread pollution, the industry has the potential to make a sustainable contribution to global food security [10].

A growing emphasis on the valorisation of waste in the framework of the circular economy is a result of the apparent need for an efficient method of managing food waste. Current strategies encourage resource reduction, recycling, reuse, and recovery in order to maximise material and energy recovery rather than merely eliminating waste [11]. This focus reduces the ecological footprint of food production by enabling the conversion of food waste into useful products like electricity, biofuels, and biofertilizers [11]. But in order for these tactics to work, it is crucial to comprehend the extent of the issue and its root causes, which differ depending on the national context. Only with a comprehensive view of food waste, it will be possible to develop effective policies and solutions tailored to different realities [11].

2.2   Circular economy

The circular economy has been promoted as an effective response to the growing scarcity of natural resources and as a driver for a more sustainable economic system. This approach aims to use resources more efficiently, closing loops and reusing them repeatedly, in order to eliminate the need for virgin raw materials [12]. Furthermore, the circular economy has been widely recognized as a solution to reduce pressure on the environment while simultaneously driving economic growth [13]. However, perspectives on the circular economy are often fueled by imprecise definitions, which hinder a comprehensive understanding of its true application [12].

Although the environmental and economic dimensions of the circular economy are widely discussed, it is crucial to integrate its social dimension for the approach to become truly sustainable. Sustainability is generally recognised as a three-dimensional approach, composed of the economic, environmental, and social dimensions, with the latter often being marginalised in academic discourse [14]. Effectively, a holistic understanding of the circular economy, which takes into account social impacts, is essential to ensure that this approach is not only environmentally efficient but also fair and inclusive for communities [14].

Optimising the social, environmental, technical, and economic values of materials and products in society is essential to restoring and regenerating the environment and realising the full potential of the circular economy. In terms of natural resources and population well-being, this integrative approach can make a substantial contribution to global sustainability [13]. It is crucial to stress that sustainability is not as well incorporated into studies on the topic or into the actual application of the circular economy as it would be ideal. There is a need for more integration of sustainable development principles into discussions and actions related to the circular economy, as evidenced by the fact that only 38% of publications on the topic are in line with these principles [13].

2.3   Microalgae

Microalgae have emerged as a promising solution for various industries, especially when taking into account the circular economy and sustainable development. One of their main advantages is their ability to grow under a range of conditions and produce beneficial compounds like vitamins, proteins, lipids, and pigments. Additionally, because of their exceptional adaptability, microalgae only need light and basic nutrients to fix large amounts of CO_2 through photoautotrophy and metabolise other carbon sources heterotrophically at the same time. They are a very promising platform for applications related to the circular economy because of these two features[15]. An effective method with significant nutritional potential has been found for the cultivation of microalgae using agro-industrial waste, such as that from the processing industries of starches, fruits, vegetables, meat, dairy, olive oil, wine, and beer [1]. This utilization of waste contributes to the valorisation of agro-industrial byproducts, reducing waste and providing a sustainable source of biomass.

Microalgae not only play an important role in the generation of biofuels and high-value biocompounds, but also have significant applications in environmental biotechnology, particularly in the treatment of agro-industrial wastewater. This wastewater, often characterised by high concentrations of organic matter, nitrogen, and phosphorus, can cause serious environmental problems if not adequately treated. The cultivation of microalgae can overcome the limitations of other treatment methods, as it promotes the removal of nutrients and the in situ production of oxygen, making it an effective solution for pollution caused by agro-industrial effluents [16].

In terms of sustainable development, microalgae align with the United Nations Sustainable Development Goals (SDGs), contributing to SDG 2 (Zero Hunger), SDG 9 (Industry, Innovation and Infrastructure), and SDG 12 (Responsible Consumption and Production). They provide sustainable sources of food, renewable bioenergy, and high-value biological compounds, assisting in waste remediation and the creation of products such as fertilizers, proteins, and pigments [15]. Finally, the most common monosaccharides in the polysaccharides of microalgae, such as glucose, galactose, and fructose, are versatile, biodegradable, and biocompatible compounds, with properties that make them suitable for a variety of applications, including the production of chemicals, food, and animal feed [17]. These characteristics highlight the importance of microalgae in the circular economy, which aims to reduce waste and efficiently reuse resources.
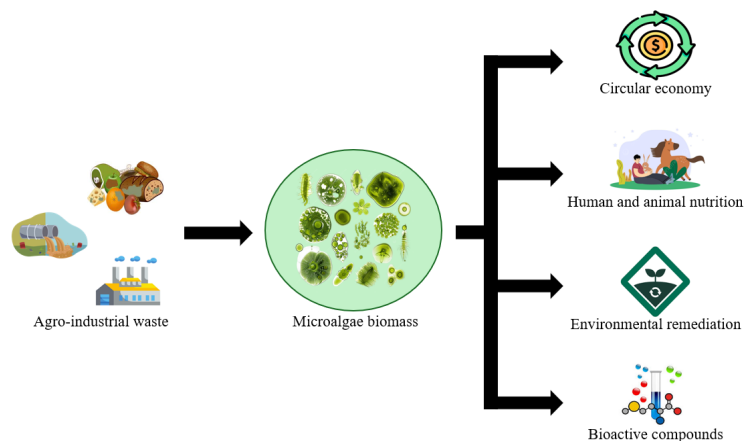
Figure 1. Schematic representation of the biotechnological valorisation of agro-industrial waste through microalgal biomass production, and its potential applications in circular economy strategies, nutritional supplementation for humans and animals, environmental remediation processes, and the synthesis of bioactive compounds.

3   Objectives

Following the United Nations Sustainable Development Goals 2, 9, and 12, this work plan proposes an innovative approach to reduce the impacts of agro-industrial activities and food retail. The focus is on developing bioprocesses that mitigate and enable the reuse of organic carbon-rich waste, using microalgae. For these processes to be implemented quickly and effectively, it is essential to employ bioinformatics tools that allow for the integration of the most advantageous characteristics of agri-food waste with the identification of microalgae strains with the greatest potential for valorisation of each type of waste.

This study aims to:

1. Create a platform that connects a specific sugar derived from agro-industrial waste to a microalgal strain that carries the enzyme necessary to hydrolyse it into simple sugars (e.g. glucose);
2. Incorporate information about the enzymes produced by different species of microalgae on the same platform;
3. Use the platform to identify which byproducts can serve as a substrate for the cultivation of microalgae, pointing out the species with the greatest potential for growth in each case.

4   Methodology

A bioinformatics platform is being developed that initially aims to identify microalgae with enzymes that are capable of degrading sugars such as lactose and

sucrose. These are the main targets, but fructose, maltose and other oligosac-charides such as starch and cellulose are also valid. Using specific queries in the Proteins section, NCBI is useful to locate and extract enzymes that are present in microalgae. To get reliable results, these search terms aid in reducing the size of the search space. In the next stage, the set of enzymes is organized and indexed, associating each microalga with the enzymes it encodes. This mapping is carried out automatically with the help of tools such as BioPython, which allows the reading and organization of the data to establish a correspondence between microalgae and the enzymes responsible for the degradation of sugars. Biopython also enables the retrieval of PubMed articles relating to microalgae and enzymes, as well as those linking microalgae to each enzyme and vice versa, and affords programmatic access to each organism's taxonomic information in a structured manner. Retrieving enzyme data from UniProt and KEGG Enzyme, such as EC (Enzyme Commission) numbers, biological functions and catalysed reactions further enhances the platform by adding comprehensive catalytic and mechanistic insights for reliable analysis. Finally, a user interface is created using Python's Streamlit library, allowing users to visualise and filter the results in real time. The application's backend performs automatic searches in the databases, returning organised results and facilitating the identification of microalgae with the potential for waste biotransformation.
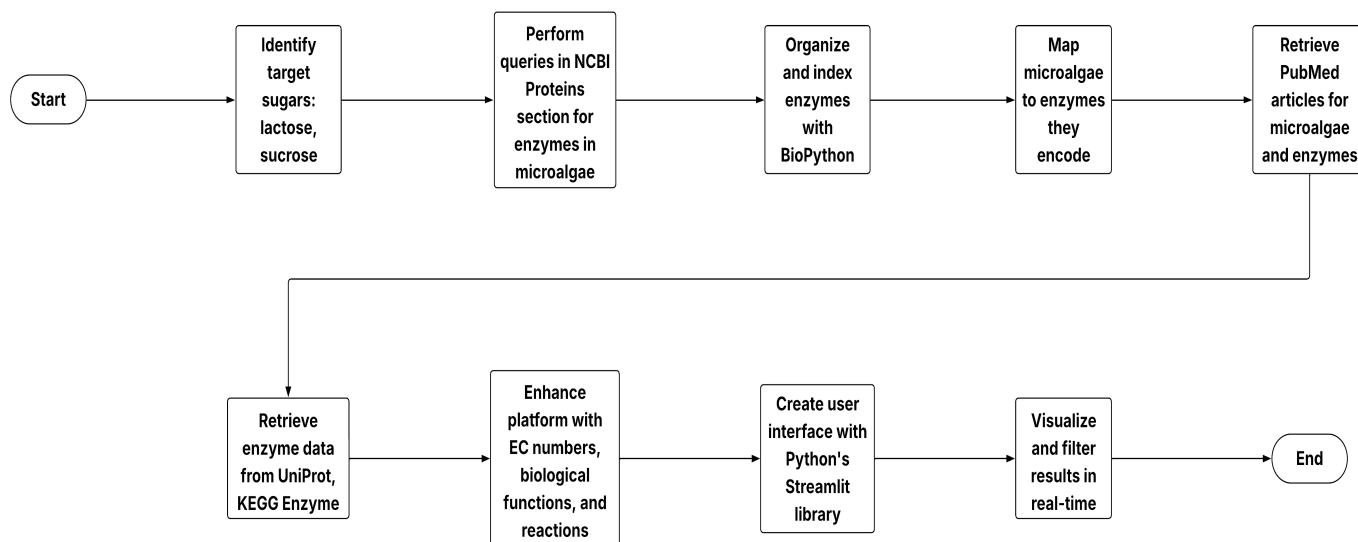

Figure 2. Flowchart illustrating the stages of acquisition, analysis, and presentation of data from microalgae and enzymes, from the identification of target sugars to the final user interaction with the platform.

5   Results

5.1   Enzymes

For this study, the following enzymes were used, each identified by its corresponding EC (Enzyme Commission) number. EC number is numerical classification devised by the International Union of Biochemistry and Molecular Biology to categorise enzymes according to the reactions they catalyse, thereby providing an unambiguous system for enzyme identification.

- **Lactose**
    - β-galactosidase (EC 3.2.1.23)
    - lactase (EC 3.2.1.23)
- **Sucrose**
    - β-fructofuranosidase (EC 3.2.1.26)
    - invertase (EC 3.2.1.26)
    - sucrase (EC 3.2.1.26)
- **Fructose**
    - fructokinase (EC 2.7.1.4)
- **Maltose**
    - maltase (EC 3.2.1.20)
- **Starch**
    - β-amylase (EC 3.2.1.2)
    - α-amylase (EC 3.2.1.1)
- **Cellulose**
    - cellulase (EC 3.2.1.4)

5.2   Algae

The main algal phyla relevant to this study are listed below, along with a brief description of each:

- **Chlorophyta** – green algae
- **Rhodophyta** – red algae
- **Glaucophyta** – freshwater algae
- **Bacillariophyta** – diatoms
- **Haptophyta** – golden algae

5.3   Platform

This platform compiles and arranges information on enzymes made by different algae species as well as sugars obtained from agro-industrial waste. It seeks to determine which sugar fractions are suitable as substrates for various algae, emphasising those with the greatest capacity for biomass production and enzymatic activity.

The core features of the platform include:

- **Interactive data exploration:** Users can filter data by algae species and target sugars, with real-time updates on the number of enzymes confirmed, species diversity, and filtering statistics.
- **Scientific Literature Integration:** Automated search on PubMed (via NCBI API) for articles that mention combinations of algae species and specific enzymes.
- **Taxonomic insights:** Taxonomy information of algae species is retrieved from the NCBI Taxonomy database, aiding biological classification and lineage understanding.
- **External database retrieval:**
  - *KEGG:* Provides enzymatic information such as reactions, EC numbers, definitions, and pathway associations.
  - *UniProt:* Offers biochemical properties including catalytic activities, cofactor requirements, subunit structures, and metabolic pathways.
- **User-friendly interface:** Built with Streamlit, the application provides an accessible and intuitive interface for researchers, enabling data visualization, literature browsing, and enzyme-pathway analysis.

Figure 3 shows that the "Enzymes in Algae" platform's "Data" tab is divided between a main dataframe view and a left-hand sidebar. Two dropdowns allow users to filter by algal species or target sugars, and the "Apply" and "Clear" buttons allow users to update the view. The "About this platform" panel in the sidebar describes how agro-industrial waste data is combined with algal enzyme data. Three metrics—the number of algae species, the percentage of confirmed enzymes, and the total number of records—appear above a scrollable table on the right. An enzyme entry is listed in each row, along with the species, enzyme name, EC number, target sugar, a brief description, and a colour-coded icon that indicates the status of confirmation. This arrangement facilitates quick filtering, browsing, and interpretation of the dataset.
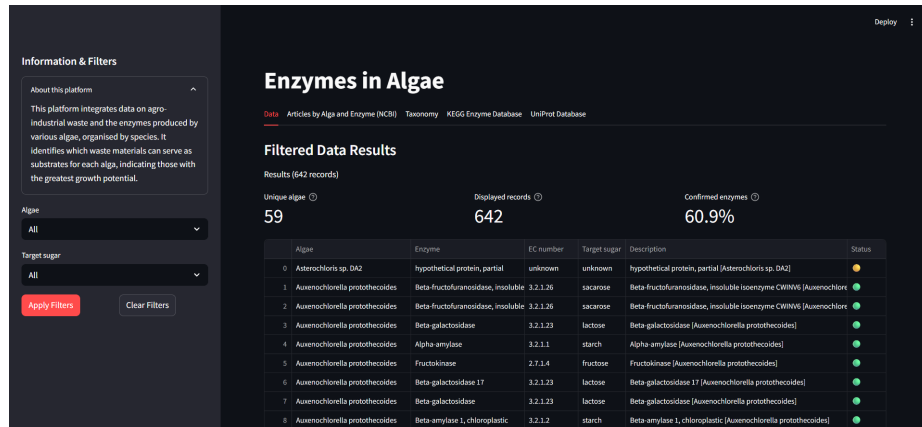

Figure 3. Main interface of the "Enzymes in Algae" platform, displayed under the "Data" tab. In this case, the filters are set so that all possible results are displayed.

Figure 4 displays the "Articles by Alga and Enzyme (NCBI)" section of the platform. This interface provides a summary of scientific articles from the NCBI database related to specific algae–enzyme combinations. At the top, key metrics are presented, including the total number of algae species (59), the number of articles retrieved (22), and the number of successful algae–enzyme matches (11). Below, the interface lists individual algae species along with the number of associated combined articles. These entries can be expanded to reveal further details about each article. The articles are organised by enzyme and it is possible to view the title, authors, publication venue, a link to PubMed and, finally, the abstract.
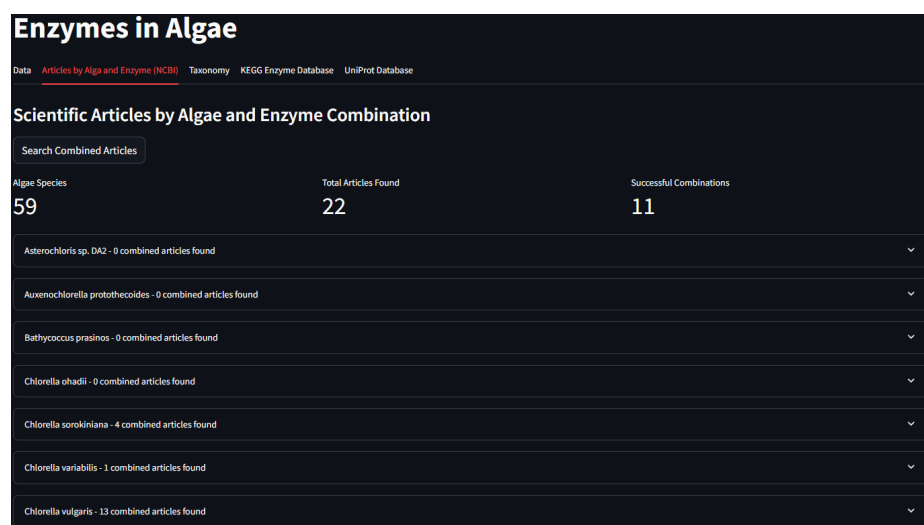

Figure 4. "Articles by Alga and Enzyme (NCBI)" tab.

Figure 5 presents the "Taxonomy" section of the platform. This interface enables users to retrieve taxonomic information for each algal species listed. By clicking the "Load Taxonomy Data" button, users can view detailed classification data for individual algae. Each species appears in a collapsible menu format, allowing for an organised and compact display of the taxonomic hierarchy. The taxa are organised from top to bottom, from the most general to the most specific.
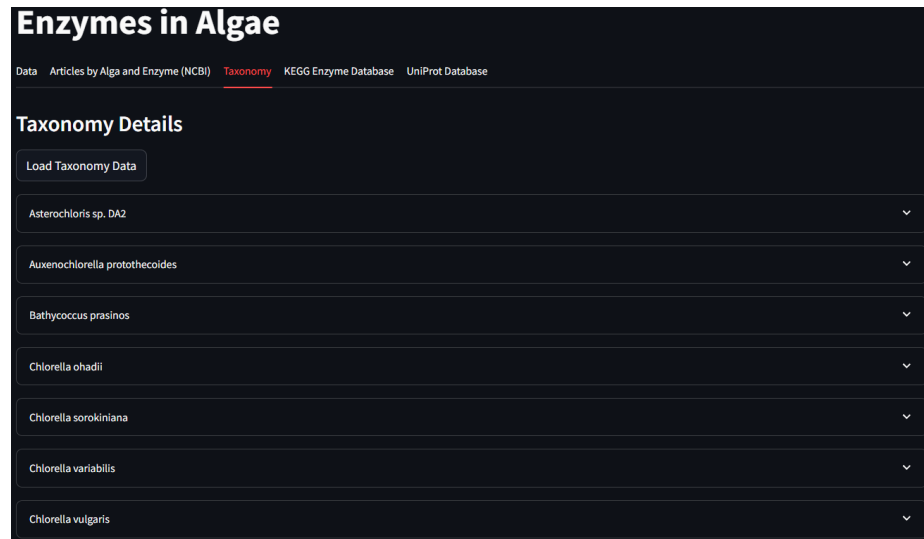
Figure 5. "Taxonomy" tab.

Figure 6 illustrates the "KEGG Enzyme Database" section of the platform. When the user clicks "Load KEGG Data," the system retrieves enzyme information linked to algal species from the KEGG database. In this instance, data for 10 enzymes were successfully retrieved, corresponding to 7 unique EC numbers. Each enzyme is shown with its name and EC classification, and additional details can be expanded. Also, a direct link accompanying each enzyme directs the user to its entry in the KEGG Enzyme database. This section helps explore enzymatic functions and classifications as defined by the KEGG system, offering insights into the metabolic capabilities of algae.
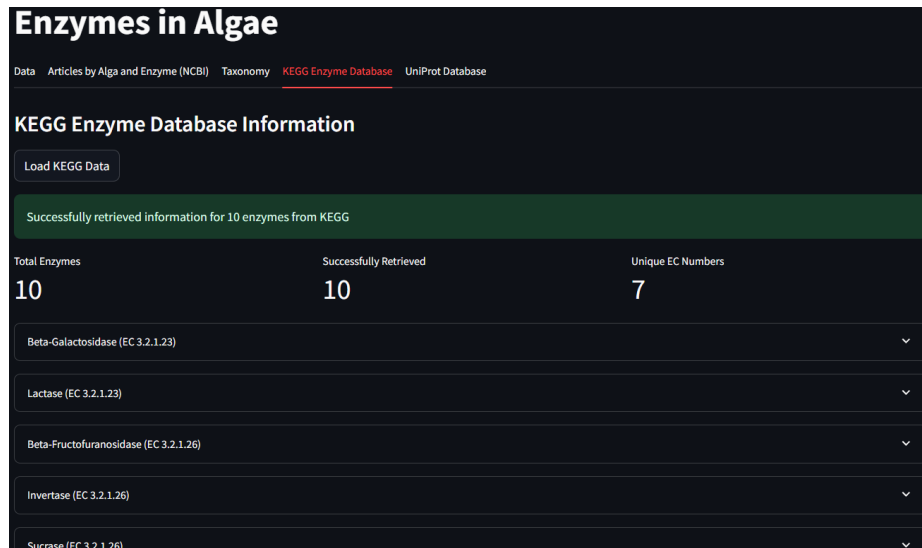
Figure 6. "KEGG Enzyme Database" tab.

Figure 7 illustrates the "UniProt Database" section of the platform. After clicking "Load UniProt Data," the system fetches protein-related information for the selected enzymes from the UniProt database. Here, data for all 10 enzymes were retrieved, resulting in 300 UniProt entries analysed—an average of 30 per enzyme. Each enzyme is listed with its name and EC number, alongside the number of entries analysed. Users can click the link for each enzyme to search for it in the UniProt database. This section provides access to protein-level data, including sequences and annotations, offering a deeper understanding of the enzymatic landscape in algal species.
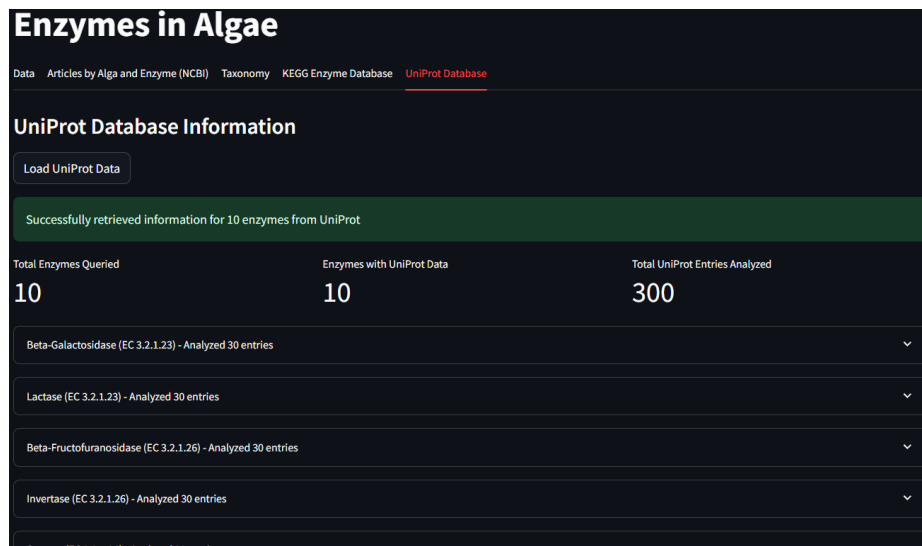
Figure 7. "UniProt Database" tab.

6   Conclusion

This research successfully developed a bioinformatic platform that addresses agro-industrial waste management through strategic microalgae cultivation. The platform integrates multiple biological databases to match microalgae species with specific waste substrates based on enzymatic capabilities, creating a practical tool for biotechnological waste valorization.

The systematic approach demonstrated the effective integration of BioPython, NCBI, KEGG and UniProt databases to create a user-friendly Streamlit interface. The platform identified 59 algae species with confirmed enzymatic capabilities for processing various sugar substrates, establishing a solid foundation for targeted waste biotransformation processes.

This work contributes significantly to circular economy strategies and sustainable development goals by transforming waste materials into valuable resources. The platform democratizes access to complex biological data, enabling researchers and industry professionals to make informed decisions about microalgae cultivation for specific waste streams.

The modular design creates opportunities for future enhancements, including expanded databases, predictive modeling capabilities, and pilot-scale integration. This research establishes a robust framework for systematic agro-industrial waste valorization, providing both immediate practical benefits and a foundation for continued development in environmental biotechnology.

References

1. Melikoglu, M., Lin, C., Webb, C.: Analysing global food waste problem: pinpointing the facts and estimating the energy content. Open Engineering **3**, 157–164 (2013)
2. Kumar, Y., Kaur, S., Kheto, A., Munshi, M., Sarkar, A., Pandey, H.O., Tarafdar, A., Sindhu, R., Sirohi, R.: Cultivation of microalgae on food waste: Recent advances and way forward. Bioresource Technology **363**, 127834 (2022)
3. Corrado, S., et al.: Food waste accounting methodologies: Challenges, opportunities, and further advancements. Global Food Security **20**, 93–100 (2019)
4. Sarker, A., et al.: A comprehensive review of food waste valorization for the sustainable management of global food waste. Sustainable Food Technology **2**, 48–69 (2024)
5. Chikezie, C., Okey, S.: Agro-Industrial Waste Management: The Circular and Bioeconomic Perspective. (2023)
6. Prado-Acebo, I., Cubero-Cardoso, J., Lu-Chau, T.A., Eibes, G.: Integral multivalorization of agro-industrial wastes: A review. Waste Management **183**, 42–52 (2024)
7. Singh, R., et al.: Utilisation of agro-industrial waste for sustainable green production: a review. Environmental Sustainability **4**, 619–636 (2021)
8. Ahmad, T., et al.: Treatment and utilization of dairy industrial waste: A review. Trends in Food Science & Technology **88**, 361–372 (2019)
9. Udourioh, G.A., Solomon, M.M., Okolie, J.A.: A Review of the Valorization of Dairy Industry Wastes through Thermochemical, Biological, and Integrated Processes for Value-Added Products. Food Science of Animal Resources **45**, 375–408 (2025)
10. Feil, A.A., et al.: Sustainability in the dairy industry: a systematic literature review. Environmental Science and Pollution Research **27**, 33527–33542 (2020)
11. Tamasiga, P., et al.: Food Waste and Circular Economy: Challenges and Opportunities. Sustainability **14**, 9896 (2022)
12. Figge, F., Thorpe, A.S., Gutberlet, M.: Definitions of the circular economy: Circularity matters. Ecological Economics **208**, 107823 (2023)
13. Velenturf, A.P.M., Purnell, P.: Principles for a sustainable circular economy. Sustainable Production and Consumption **27**, 1437–1457 (2021)
14. Mies, A., Gold, S.: Mapping the social dimension of the circular economy. Journal of Cleaner Production **321**, 128960 (2021)
15. Olabi, A.G., et al.: Role of microalgae in achieving sustainable development goals and circular economy. Science of The Total Environment **854**, 158689 (2023)
16. Posadas, E., et al.: Microalgae-based agro-industrial wastewater treatment: a preliminary screening of biodegradability. Journal of Applied Phycology **26**, 2335–2345 (2014)
17. Franco-Morgado, M., et al.: Microalgae and cyanobacteria polysaccharides: Important link for nutrient recycling and revalorization of agro-industrial wastewater. Applied Food Research **3**, 100296 (2023)