Journal Pre-proof

Global WWTP Microbiome-based Integrative Information Platform: From experience to intelligence

Fuzhong Xiong, Zhiguo Su, Yushi Tang, Tianjiao Dai, Donghui Wen

PII: S2666-4984(23)00135-7

DOI: https://doi.org/10.1016/j.ese.2023.100370

Reference: ESE 100370

To appear in: *Environmental Science and Ecotechnology*

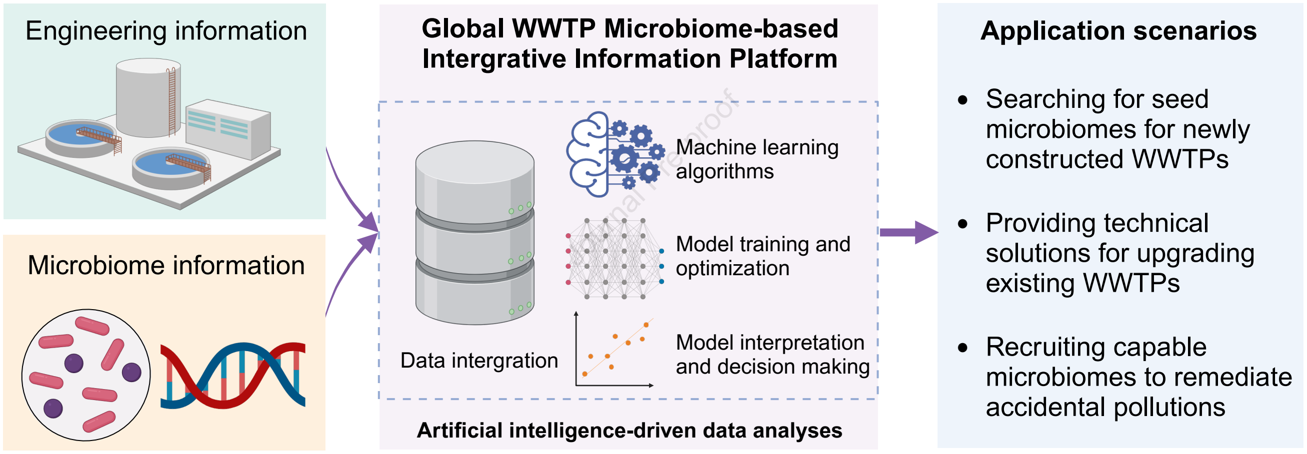
Received Date: 16 August 2023

Revised Date: 11 December 2023

Accepted Date: 12 December 2023

Please cite this article as: F. Xiong, Z. Su, Y. Tang, T. Dai, D. Wen, Global WWTP Microbiome-based Integrative Information Platform: From experience to intelligence, *Environmental Science and Ecotechnology* (2024), doi: https://doi.org/10.1016/j.ese.2023.100370.

An artificial intelligence-driven strategy for microbiome engineering in wastewater treatment



Global WWTP Microbiome-based Integrative Information Platform: From experience to intelligence

Fuzhong Xiong[a], Zhiguo Su[b], Yushi Tang[c], Tianjiao Dai[d], Donghui Wen[a,*]

[a] College of Environmental Sciences and Engineering, Peking University, Beijing, 100871, China

[b] School of Environment, Tsinghua University, Beijing, 100084, China

[c] Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ, 08544, USA

[d] School of Water Resources and Environment, China University of Geosciences (Beijing), Beijing, 100083, China

[*] Corresponding author.

E-mail address: dhwen@pku.edu.cn (D. Wen).

Abstract

Domestic and industrial wastewater treatment plants (WWTPs) are facing formidable challenges in effectively eliminating emerging pollutants and conventional nutrients. In microbiome engineering, two approaches have been developed: a top-down method focusing on domesticating seed microbiomes into engineered ones, and a bottom-up strategy that synthesizes engineered microbiomes from microbial isolates. However, these approaches face substantial hurdles that limit their real-world applicability in wastewater treatment engineering. Addressing this gap, we propose the creation of a Global WWTP Microbiome-based Integrative Information Platform, inspired by the untapped microbiome and engineering data from WWTPs and advancements in artificial intelligence (AI). This open platform integrates microbiome and engineering information globally and utilizes AI-driven tools for identifying seed microbiomes for new plants, providing technical upgrades for existing facilities, and deploying microbiomes for accidental pollution remediation. Beyond its practical applications, this platform has significant scientific and social value, supporting multidisciplinary research, documenting microbial evolution, advancing Wastewater-Based Epidemiology, and enhancing global resource sharing. Overall, the platform is expected to enhance WWTPs’ performance in pollution control, safeguarding a harmonious and healthy future for human society and the natural environment.

Keywords: Wastewater treatment plants; Microbiome; Environmental engineering; Information platform; Artificial intelligence

1. Introduction

Domestic and industrial wastewater treatment plants (WWTPs) are sophisticated engineering systems designed to eliminate pollutants, ensure human health, and maintain ecological sustainability. Within WWTPs, biological treatment methods are widely employed as central technological units [1, 2]. The utilization of microbiome engineering in wastewater treatment can be traced back to the late 19th century [3, 4]. Since the mid-20th century, activated sludge and anaerobic digestion have emerged as worldwide benchmarks for treating domestic and industrial wastewater [5, 6]. Currently, these microbiome-based WWTPs play a vital role in connecting the societal and natural cycles of elements on our planet.

However, the escalating complexity of pollutants in modern wastewater poses significant challenges for WWTPs. Biological treatment processes often inadequately eliminate emerging pollutants [7-10], and the transformation and removal of conventional nutrients still lack stability [11, 12]. This issue renders WWTPs into pollution point sources with considerable risk worldwide. In recent decades, there has been a remarkable expansion in our understanding of the microbial world, driven by significant advancements in fundamental disciplines, such as microbiology, biochemistry, molecular biology, and microbial ecology. This burgeoning scientific knowledge has promoted the development of two strategies for microbiome engineering in the context of wastewater treatment: the bottom-up and top-down approaches (Fig. 1) [13, 14].

The bottom-up approach starts with specific microbial isolates with known physiological characteristics and metabolic pathways. Then, based on the effective performances of individual microbes and the well-organized interspecific interactions, a consortium with desired functions (i.e., degrading target pollutants in wastewater) could be artificially assembled through co-cultivating processes with rational designing and regulating strategies [15, 16]. Conversely, the top-down approach initiates with a seed microbiome containing culturable and unculturable microorganisms, which usually originate from activated sludge in engineering systems or contaminated sediment and soil in natural environments. Then, techniques like enrichment, artificial selection, and directed evolution are employed to domesticate this seed microbiome [17-19]. Finally, it will be shaped into an engineered microbiome with desired functions, such as efficient degradation of target pollutants and elevated tolerance to unfavorable conditions. No matter which approach is adopted, the engineered microbiomes for wastewater treatment could be assembled and improved through an iterative “design-build-test-learn” (DBTL) cycle [13, 14]. However, both approaches are currently constrained to theoretical frameworks, as they have encountered practical bottlenecks in engineering applications.

At the “bottom”, research on pollutant-degrading strains, functional genes, and metabolic pathways has become increasingly comprehensive. In controlled laboratory environments, carefully constructed microbial consortia can indeed manifest desired functions in simulated wastewater treatment experiments [20, 21]. However, these consortia often fall short when dealing with the real-world complexity of domestic and industrial wastewater [22, 23]. This inefficacy stems from several key factors. (1) Individual microbes adapt, survive, and reproduce within specific habitats. When inoculating these microbes from a laboratory environment into an engineering environment, their metabolic functions will be unpredictable during the adaptation process [24-26]. (2) The simple co-cultivation of different isolates does not guarantee combining their functions due to metabolic division of labor [27, 28]. Hence, establishing a microbial consortium that can stably perform its intended functions requires a clear understanding of each isolate's metabolic pathways and their respective roles in the division of labor. In addition, these isolates must develop mutualistic interactions in specific environments, which is often fraught with uncontrollability [13, 26]. (3) The constructed microbial consortia lack a variety of unculturable and rare species that possess unknown yet indispensable functions [29, 30], making it difficult to form stable engineered microbiomes for long-term use. Consequently, the bottom-up approach usually proves inefficient and unfeasible for real-world wastewater treatment engineering.

At the "top", a multitude of research concerning microbial community composition, functional traits, and ecological interactions consistently guides the testing and upgrading of wastewater treatment engineering. The seed or original engineered microbiomes could be shaped into more capable ones via different strategies [31-33]. However, these microbiomes have to face the challenge of fluctuating environmental factors in wastewater, and real-world engineering also needs to carry out multi-objective start-ups and adjustments. Therefore, microbiome engineering often falls into a time-consuming and low-efficiency cycle of trial and error [34, 35]. This is attributable to several factors. (1) The engineered microbiomes contain numerous obscure species (i.e., microbial dark matter) and biomolecules with as-yet-undetermined functions and interactions [29, 36, 37]. Hence, this often necessitates a deep reliance on empirical knowledge for effective modification. (2) Microbial communities in natural states exhibit both functional redundancy and multifunctionality [38]. The temporal variations of community composition and its ecological interactions in response to wastewater quality will introduce uncertainties in the overall functionality [39-41]. Accordingly, the simplified community structures and interactions might cause the engineered microbiomes to lose their redundancy and "multithreading" capabilities in addressing the complicated and variable stress of actual wastewater [14]. (3) Additionally, specific wastewater treatment systems (e.g., refractory industrial wastewater treatment reactors, anaerobic granular sludge reactors, and nitrifying reactors) face challenges in finding suitable seed microbiomes. The processes of domesticating and enriching often take a few months to years to yield a capable engineered microbiome [42-44]. Overall, the above pitfalls have notably limited the efficacy of the top-down approach. As a result, research within this field is dedicated to fulfilling these knowledge gaps and exploring more feasible solutions.

The development of WWTPs, with microbiome engineering as the core, has evolved over a century and formed a foundational theoretical and technological system. However, a vast array of data and information from engineering practices has remained unexploited within WWTPs, largely due to the absence of an effective data integration and analysis platform. Although some public databases collect microbial genomic information from various environments, including WWTPs, their primary focus is preserving and sharing gene sequence data without directly linking to real-world microbiome engineering. The current advancements in integrated science, big data computing, and artificial intelligence (AI) inspire us to propose a novel approach: firstly, collect and consolidate comprehensive datasets of engineering parameters and microbiome information from operating WWTPs worldwide; secondly, harness the capabilities of big data engines to guide the search of seed microbiomes suitable for newly established WWTPs and provide technical strategies for shaping them into engineered microbiomes; thirdly, employ AI-driven modeling and multi-objective optimization analyses to diagnose issues within existing WWTPs and propose optimized regulation schemes that can upgrade the current low-efficiency engineered microbiomes to high-efficiency ones.

This novel approach has notable advantages that can overcome the limitations of the bottom-up and top-down approaches. We do not need to start from scratch in strain screening and engage in intricate co-cultivation designs or conduct time-consuming artificial domestication to assemble the desired engineered microbiomes, which can save substantial time and economic costs. The seed microbiomes, directly acquired *in* *situ* from global WWTPs, possess stable ecological structures and encompass microbial dark matter vital for community functioning. This is crucial for the stable operation of newly established or upgraded microbial engineering systems in the long term. Additionally, advanced computational tools based on AI modeling and big data mining greatly elevate the efficiency and intelligence of microbiome engineering. Thus, we can swiftly obtain practical suggestions for the operation or upgrading of WWTPs.

As an initial step to achieve this approach, there is a pressing need to establish an integrative information platform that integrates engineering and microbiome data from global WWTPs and provides deep analytical capabilities based on AI-driven models and big data computations.

2. Establishment of Global WWTP Microbiome-based Integrative Information Platform

We propose establishing an open platform for sharing and service by extensively collecting information from domestic and industrial wastewater treatment systems worldwide. This platform will serve as a microbial ecology database for numerous large-scale microbiome engineering projects globally, particularly those involving WWTPs, recording the survival and effectiveness of “professional decomposers” on Earth. Additionally, this platform, featuring environmental engineering context, will provide economically and technically feasible solutions for establishing and optimizing WWTPs.

The information covered by this platform includes two parts: (1) basic information and engineering parameters of WWTPs and (2) microbiome information in the WWTPs. The first part (Fig. 2, green panels) can be further divided into: (1) WWTP backgrounds, including geographical location, source of wastewater, industrial type, treatment capacity, discharge destination, etc.; (2) WWTP technologies, including treatment processes, types of biological reactors (e.g., anaerobic upflow anaerobic sludge blanket [UASB] or expanded granular sludge bed [EGSB], aerobic conventional activated sludge [CAS], sequencing batch reactor [SBR]), and operational parameters, such as flux, hydraulic retention time (HRT), sludge retention time (SRT), mixed liquor suspended solids (MLSS), sludge reflux ratio, carrier type, etc.; (3) WWTP performances, including removal efficiency for key pollutants, such as chemical oxygen demand (COD), biochemical oxygen demand (BOD), total organic carbon (TOC), ammonium, total nitrogen, total phosphorus, refractory organics, antibiotic resistance genes, and effluent toxicity reduction; (4) Environmental factors, i.e., physicochemical characteristics of the microbial environment, including water quality indicators (temperature, pH, dissolved oxygen, salinity, pollutant concentrations, etc.) and sludge characteristics (particle size, settleability, hydrophobicity, zeta potential, etc.).

On the other hand, the platform will encompass microbiome information derived from various technological means, such as high-throughput amplicon sequencing, metagenomic sequencing, single-cell sequencing, and the corresponding annotations for species and functions. Additionally, using multi-omics approaches that combine metagenomics, transcriptomics, proteomics, and metabolomics, the platform can provide profound insights into microbial metabolic pathways and crucial substance transformation mechanisms. More importantly, after obtaining the above information, the platform will employ big data-driven computational methods like multivariate statistical analysis, machine learning, and interpretive modeling to conduct in-depth analyses (Fig. 2, yellow panel). These will establish a deep coupling framework between the engineering information and microbiome information, ultimately guiding the start-up and operational maintenance of WWTPs based on research results obtained from the microbiome information.

We designate this archival library and technical service platform as the Global WWTP Microbiome-based Integrative Information Platform (hereinafter referred to as "the Platform"). The Platform stands apart from existing microbial strain repositories and online microbial genome databases. Its distinctiveness originates from the following aspects. Firstly, due to the inability to simulate habitats, the microbiomes remain *in situ* in the WWTP engineering systems rather than preserved in physical cryogenic storage. Secondly, the Platform incorporates not only microbial genomic sequences, but also a wealth of background information, and actual parameters associated with the engineering. More importantly, complex statistical analyses and up-to-date computational processes, including machine learning and model training, are employed in the Platform to dig into the multidimensional information and develop feasible solutions for engineering systems.

During the establishment and operation of the Platform, it is crucial to foster deep collaboration and data sharing with existing public databases, such as the National Center of Biotechnology Information (NCBI) in the USA, the European Bioinformatics Institute of European Molecular Biology Laboratory (EMBL-EBI) in Europe, the DNA Data Bank of Japan (DDBJ) in Japan, the National Genomics Data Center (CNCB-NGDC) in China, the Global Survey of Activated Sludge Microbiome by Global Water Microbiome Consortium [2], and the Microbial Database for Activated Sludge (MiDAS)

[45]. Some of these databases collect microbial genomic information from various global environments and hosts, while some are specifically established for wastewater treatment systems. However, their primary focus is preserving and sharing gene sequence data rather than directly associating with engineering practice. Hence, they can offer substantial data support to the Platform. Meanwhile, the Platform requires the ongoing contributions of operational engineers and researchers across various disciplines for tasks such as sample collection, daily monitoring of system status and pollutant-removing efficiency, genome extraction, high-throughput sequencing, data analysis, and data collation and uploading (Fig. 2, blue panels).

Ultimately, the Platform will apply cloud technology to store, update, and share all information and corresponding results from big data computational analyses globally. With continuous data updates, the Platform will routinely recalculate and retrain its models to provide the most accurate and state-of-the-art results. Furthermore, based on the differing user requirements, the Platform will modularize the data upload, storage, and computation processes and develop a suite of professional, efficient, and user-friendly toolkits to maximize the Platform’s usability.

3. Promising applications of the Platform

3.1. Search for seed microbiomes suitable for newly constructed WWTPs

Based on the established Platform, the engineers can input basic information about newly built WWTPs, excluding the microbiome information. The Platform will then quickly identify several similar WWTPs, considering factors such as economy, safety, and convenience to pinpoint the optimal choice. The microbial community from the activated sludge or digested sludge of this selected WWTP will serve as the seed microbiome for the new plant. Concurrently, the Platform will conduct a comprehensive statistical analysis of microbial community structures, environmental conditions, and engineering parameters from various relevant WWTPs. This analysis will be used to propose optimized start-up conditions and operational parameters, ensuring the efficient performance of the seed microbiome. Ultimately, the Platform will aid in proposing a multi-objective optimization and precise regulatory scheme for wastewater treatment engineering (Fig. 3). This method of assembling "well-structured" microbial communities overcomes the limitations of traditional empirical approaches and transcends the unknown details of synthetic biology, such as complex metabolic pathways, species interactions, and division of labor among microbial isolates [46].

3.2. Provide targeted solutions for upgrading existing WWTPs

As global environmental regulations tighten, wastewater discharge standards have become increasingly stringent [47]. Many existing WWTPs struggle to comply with these enhanced requirements. In response, the Platform emerges as an essential tool, providing pivotal support for WWTPs to upgrade their operations. It employs big data analysis and machine learning models to evaluate the operational efficiency of existing WWTPs, pinpointing specific impediments that hinder compliance with required standards. Subsequently, the Platform offers a wide range of targeted upgrade schemes and conducts multi-scenario analyses, considering efficiency, economy, and safety factors. Finally, the Platform will intelligently generate precise control schemes for efficient operation of existing WWTPs and propose reasonable and feasible technical optimization directions (Fig. 3). For instance, the Platform can guide WWTPs to effectively address issues such as substandard pollution indicators, sludge bulking, biological foam, and the loss of active microorganisms [48]. This can be accomplished through adjusting operational parameters, modulating environmental factors, and reforming treatment units.

3.3. Recruit microbiomes to respond to sudden and accidental environmental pollution

Once the Platform fully grasps the profound connection between microbiome profiles and pollutant degradation, the engineered microbiome in WWTPs can act as the “regular forces” in addressing the bulk of pollutants derived from anthropogenic emissions. These “regular forces” often include “specialized units” capable of degrading hazardous xenobiotics, a major concern in unforeseen environmental incidents and emergent pollution. Sudden environmental events, such as offshore oil spills, hazardous chemical leaks, and unintentional wastewater discharges, can momentarily inundate the natural environment with pollutants [49]. Currently, physical and chemical methods are predominantly employed to address these incidents swiftly, yet they incur high costs and risk secondary pollution. In such cases, the Platform offers a prompt, eco-friendlier alternative by efficiently mobilizing appropriate microbiomes and devising effective remediation strategies (Fig. 3). Initially, prioritizing safety and adaptability, the Platform selects potential microbiomes from global WWTPs, taking into account crucial information about the sudden environmental event, including the primary pollutants, emission volume, location of the accident, and local hydrological conditions. The Platform then applies up-to-date big data analysis techniques and machine learning models to intelligently determine potential costs, feasible techniques, and pollutant degradation efficiency according to the input conditions. Eventually, the Platform offers an optimized solution as an effective microbial remediation strategy for sudden environmental incidents.

Taken together, with the assistance of the Platform, the well-matched seed microbiomes can accelerate their adaptation in newly established engineering systems and finally form capable engineered microbiomes that meet our expectations. Meanwhile, the inefficient engineered microbiomes in the existing WWTPs can evolve into more efficient ones through multi-objective optimizations. In this way, big data-driven intelligent models will achieve a leap in developing wastewater treatment microbiome engineering. Moreover, the Platform will greatly contribute to the rapid biological remediation of accidental environmental pollution.

4. Scientific and social significance of the Platform

4.1. Provide integrated data resources for scientific research in environmental biotechnology, microbial ecology, energy, and climate change

Although wastewater treatment biotechnology, an artificial augmentation of the microbial “decomposers” functioning on Earth, has been broadly utilized for over a century, the biological processes and mechanisms involved remain a “black box” [29]. With the advancement of various technologies and the pursuit of sustainable development goals, next-generation WWTPs need to be intelligent, energy-producing, and resource-recycling [1]. The Platform provides valuable data sources for scientists to delve into the operational rules of WWTPs, thus supporting automated control and intelligent operation and maintenance. The Platform facilitates the discovery of novel microbial metabolic pathways and the development of advanced biotechnologies. It also seeks potential technological pathways to reduce greenhouse gas emissions and recover resources to alleviate the anthropogenic impact on the natural ecosystem [50]. Based on the Platform, theoretical research on artificial systems can promote the development of microbial ecology and can even lead to the emergence of a new discipline — intelligent ecology. This discipline will employ AI-driven models and algorithms to elucidate complex ecological processes, predict ecological outcomes, and address ecological challenges.

4.2. Record the co-evolution processes between industrial civilization and microorganisms at the global scale

Due to the constant progression of industrial civilization, industrial production is undergoing rapid changes, characterized by the continuous emergence of new manufacturing processes and products [51]. As important participants in decomposing industrial synthetic substances, wastewater treatment microorganisms are evolving in tandem with the corresponding industrial processes [22, 25]. Professor Stephen

Palumbi, the author of “The Evolution Explosion: How Humans Cause Rapid Evolutionary Change,” once posited that hospitals provide an ideal setting to observe the evolution of bacterial resistance to antibiotics [52]. Similarly, WWTPs are also ideal places to track the evolution of microbial tolerance and degradation capabilities to emerging chemicals in a long historical period. Hence, the Platform, maintained and improved over the years, will chronicle the evolutionary trajectory of microorganisms in the Anthropocene era.

4.3. Monitor and provide early warning of public health risks in wastewater collection areas

It is worth noting that the massive microbial metagenomic sequences within the Platform, processed through AI-based computational analysis, can assist us in predicting potential health risks from the potential microbial mutations or the emergence of novel pathogenic organisms in our environment. The best example in this context is the surveillance of the SARS-CoV-2 virus in wastewater [53]. As stated in a report from the US Centers for Disease Control and Prevention, sewage monitoring is a powerful tool that provides early warning signals for virus transmission and helps track the complete development trends and patterns of each epidemic wave [54]. In addition, other respiratory viruses like influenza and respiratory syncytial virus can also be detected in time-series wastewater samples [55]. Since the Platform continuously collects microbiome information from WWTPs worldwide, it can use powerful bioinformatic tools and existing pathogen databases to identify the putative pathogenic species and related metabolic pathways or processes. Therefore, the Platform can provide comprehensive and timely data support for Wastewater-Based Epidemiology (WBE) or Wastewater-Based Surveillance (WBS) by routinely detecting significant changes in high-risk infectious disease markers [55, 56].

4.4. Provide a global communication and trading platform for the optimization and integration of techniques and resources

Human society has formed a mature and complex production chain comprising raw material collecting, upstream primary processing, midstream refinement, and downstream manufacturing [57]. With the progression of economic globalization, factories worldwide are united into a giant "global producer network", while people from different countries or areas are integrated into a "global consumer network". However, the sustainable and stable functioning of these two networks relies on the "global decomposer network" underpinned by modern biotechnologies, especially within wastewater treatment engineering [58]. In this regard, the Platform will serve as a seed bank and a trading platform for engineered microbiomes. Based on the principles of information sharing and fair trading, the Platform encourages technical guidance and resource sharing from developed countries to developing countries in fields such as industrial policy, production and consumption, ecological and environmental protection, and Environmental, Social, and Governance (ESG) for sustainable development [59, 60]. Doing so can minimize the costs of trial and error in developing countries and the potential risks of pollution to the earth, thereby enhancing the collective well-being of all humankind.

5. Concluding remarks

The performances of modern WWTPs significantly determine the extent of human interference with nature. With the evolution of WWTP automation, driven by advancements in online monitoring technology, there is a surge in the production of detailed engineering operational data. Meanwhile, the rapid development of next-generation sequencing and multi-omics technologies has deepened our understanding of microbial community structures and functions within WWTPs. This intersection offers an invaluable opportunity to harness and maximize the potential of this wealth of engineering and microbiome data through advanced AI-driven data processing techniques. In this context, the ultimate aim of the Global WWTP Microbiome-based Integrative Information Platform (which might be abbreviated as GWMII Platform) is to overcome the inherent limitations of traditional empirical approaches, establishing a robust framework for intelligently guaranteeing optimized wastewater treatment performances. Thus, the GWMII Platform paves the way for fostering a balanced and sustainable synergy between human society and the natural environment.

CRediT authorship contribution statement

Fuzhong Xiong: Conceptualization, Investigation, Writing - Original Draft, Visualization. **Zhiguo Su:** Conceptualization, Writing - Original Draft, Visualization. **Yushi Tang:** Writing - Original Draft. **Tianjiao Dai:** Writing - Original Draft, Validation. **Donghui Wen:** Conceptualization, Supervision, Project Administration, Writing - Review & Editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was financially supported by the National Natural Science Foundation of China (grant number 51938001, 52170185, 42007291, and 52070111) and the China Postdoctoral Science Foundation (grant number 2022M721815). We extend our appreciation to BioRender.com for the invaluable support in creating Fig. 2, Fig. 3, and the Graphical Abstract presented in this article.

References

1. Jafarinejad S. A framework for the design of the future energy-efficient, cost-effective, reliable, resilient, and sustainable full-scale wastewater treatment plants. Current Opinion in Environmental Science & Health. 2020;13:91-100. doi: 10.1016/j.coesh.2020.01.001.
2. Wu L, Ning D, Zhang B, Li Y, Zhang P, Shan X, et al. Global diversity and biogeography of bacterial communities in wastewater treatment plants. Nature Microbiology. 2019;4(7):1183-95. doi: 10.1038/s41564-019-0426-5.
3. Alleman JE, Prakasam TBS. Reflections on Seven Decades of Activated Sludge History. Journal (Water Pollution Control Federation). 1983;55(5):436-43.
4. Ardern E, Lockett WT. Experiments on the oxidation of sewage without the aid of filters. Journal of the Society of Chemical Industry. 1914;33(10):523-39.
5. Jenkins D, Wanner J. Activated sludge-100 years and counting: IWA publishing; 2014.
6. Lettinga G. Anaerobic digestion and wastewater treatment systems. Antonie van Leeuwenhoek. 1995;67(1):3-28. doi: 10.1007/BF00872193.
7. Byrns G. The fate of xenobiotic organic compounds in wastewater treatment plants. Water Research. 2001;35(10):2523-33. doi: 10.1016/S0043-1354(00)00529-7.
8. Kasprzyk-Hordern B, Dinsdale RM, Guwy AJ. The removal of pharmaceuticals, personal care products, endocrine disruptors and illicit drugs during wastewater treatment and its impact on the quality of receiving waters. Water Research. 2009;43(2):363-80. doi: 10.1016/j.watres.2008.10.047.
9. Michael I, Rizzo L, McArdell CS, Manaia CM, Merlin C, Schwartz T, et al. Urban wastewater treatment plants as hotspots for the release of antibiotics in the environment: A review. Water Research. 2013;47(3):957-95. doi: 10.1016/j.watres.2012.11.027.
10. Montecinos S, Gil M, Tognana S, Salgueiro W, Amalvy J. Distribution of microplastics present in a stream that receives discharge from wastewater treatment plants. Environ Pollut. 2022;314:120299. doi: 10.1016/j.envpol.2022.120299.
11. Li C, Liu S, Ma T, Zheng M, Ni J. Simultaneous nitrification, denitrification and phosphorus removal in a sequencing batch reactor (SBR) under low temperature. Chemosphere. 2019;229:132-41. doi: 10.1016/j.chemosphere.2019.04.185.
12. Campo R, Sguanci S, Caffaz S, Mazzoli L, Ramazzotti M, Lubello C, et al. Efficient carbon, nitrogen and phosphorus removal from low C/N real domestic wastewater with aerobic granular sludge. Bioresource Technology. 2020;305:122961. doi: 10.1016/j.biortech.2020.122961.
13. Lawson CE, Harcombe WR, Hatzenpichler R, Lindemann SR, Löffler FE, O'Malley MA, et al. Common principles and best practices for engineering microbiomes. Nature Reviews Microbiology. 2019;17(12):725-41. doi: 10.1038/s41579-019-0255-9.
14. Hu H, Wang M, Huang Y, Xu Z, Xu P, Nie Y, et al. Guided by the principles of microbiome engineering: Accomplishments and perspectives for environmental use. mLife. 2022;1(4):382-98. doi: 10.1002/mlf2.12043.

15. Jia X, Liu C, Song H, Ding M, Du J, Ma Q, et al. Design, analysis and application of synthetic microbial consortia. Synthetic and Systems Biotechnology. 2016;1(2):109-17. doi: 10.1016/j.synbio.2016.02.001.
16. De Roy K, Marzorati M, Van den Abbeele P, Van de Wiele T, Boon N. Synthetic microbial ecosystems: an exciting tool to understand and apply microbial communities. Environmental Microbiology. 2014;16(6):1472-81. doi: 10.1111/1462-2920.12343.
17. Huang Y, Wen L, Zhang L, Xu J, Wang W, Hu H, et al. Community-integrated multi-omics facilitates the isolation of an organohalide dehalogenation microorganism. The Innovation. 2023;4(1):100355. doi: 10.1016/j.xinn.2022.100355.
18. Sánchez Á, Vila JCC, Chang C-Y, Diaz-Colunga J, Estrela S, Rebolleda-Gomez M. Directed Evolution of Microbial Communities. Annual Review of Biophysics. 2021;50(1):323-41. doi: 10.1146/annurev-biophys-101220-072829.
19. Mueller Ulrich G, Juenger Thomas E, Kardish Melissa R, Carlson Alexis L, Burns Kathleen M, Edwards Joseph A, et al. Artificial Selection on Microbiomes To Breed Microbiomes That Confer Salt Tolerance to Plants. mSystems. 2021;6(6):e01125-21. doi: 10.1128/mSystems.01125-21.
20. Feng N-X, Yu J, Xiang L, Yu L-Y, Zhao H-M, Mo C-H, et al. Co-metabolic degradation of the antibiotic ciprofloxacin by the enriched bacterial consortium XG and its bacterial community composition. Science of The Total Environment. 2019;665:41-51. doi: 10.1016/j.scitotenv.2019.01.322.
21. Zhang L, Qiu X, Huang L, Xu J, Wang W, Li Z, et al. Microbial degradation of multiple PAHs by a microbial consortium and its application on contaminated wastewater. Journal of Hazardous Materials. 2021;419:126524. doi: 10.1016/j.jhazmat.2021.126524.
22. Raper E, Stephenson T, Anderson DR, Fisher R, Soares A. Industrial wastewater treatment through bioaugmentation. Process Saf Environ. 2018;118:178-87. doi: 10.1016/j.psep.2018.06.035. PubMed PMID: WOS:000453939000017.
23. Kapoor RT, Danish M, Singh RS, Rafatullah M, H.P.S AK. Exploiting microbial biomass in treating azo dyes contaminated wastewater: Mechanism of degradation and factors affecting microbial efficiency. J Water Process Eng. 2021;43:102255. doi: 10.1016/j.jwpe.2021.102255.
24. Eydallin G, Ryall B, Maharjan R, Ferenci T. The nature of laboratory domestication changes in freshly isolated Escherichia coli strains. Environmental Microbiology. 2014;16(3):813-28. doi: 10.1111/1462-2920.12208.
25. Steensels J, Gallone B, Voordeckers K, Verstrepen KJ. Domestication of Industrial Microbes. Current Biology. 2019;29(10):381-93. doi: 10.1016/j.cub.2019.04.025.
26. Palková Z. Multicellular microorganisms: laboratory versus nature. EMBO reports. 2004;5(5):470-6. doi: 10.1038/sj.embor.7400145.
27. Tsoi R, Wu F, Zhang C, Bewick S, Karig D, You L. Metabolic division of labor in microbial systems. P Natl Acad Sci USA. 2018;115(10):2526-31. doi: 10.1073/pnas.1716888115.

28. Thommes M, Wang T, Zhao Q, Paschalidis Ioannis C, Segrè D. Designing Metabolic Division of Labor in Microbial Communities. mSystems. 2019;4(2):e00263-18. doi: 10.1128/msystems.00263-18.
29. Zhang Y, Wang Y, Tang M, Zhou J, Zhang T. The microbial dark matter and "wanted list" in worldwide wastewater treatment plants. Microbiome. 2023;11(1):59. doi: 10.1186/s40168-023-01503-3.
30. Zamkovaya T, Foster JS, de Crécy-Lagard V, Conesa A. A network approach to elucidate and prioritize microbial dark matter in microbial communities. The ISME Journal. 2021;15(1):228-44. doi: 10.1038/s41396-020-00777-x.
31. Pechaud Y, Peyre Lavigne M, Bessiere Y, Ochoa JC, Queinnec I, Paul E. Influence of shear stress, organic loading rate and hydraulic retention time on the biofilm structure and on the competition between different biological aggregate morphotypes. Journal of Environmental Chemical Engineering. 2022;10(3):107597. doi: 10.1016/j.jece.2022.107597.
32. Jiang Y, Shi X, Ng HY. Aerobic granular sludge systems for treating hypersaline pharmaceutical wastewater: Start-up, long-term performances and metabolic function. Journal of Hazardous Materials. 2021;412:125229. doi: 10.1016/j.jhazmat.2021.125229.
33. Liu Y, Tay J-H. The essential role of hydrodynamic shear force in the formation of biofilm and granular sludge. Water Research. 2002;36(7):1653-65. doi: 10.1016/S0043-1354(01)00379-7.
34. Tejaswini ESS, Panjwani S, Gara UBB, Ambati SR. Multi-objective optimization based controller design for improved wastewater treatment plant operation. Environmental Technology & Innovation. 2021;23:101591. doi: 10.1016/j.eti.2021.101591.
35. Lawson Christopher E. Retooling Microbiome Engineering for a Sustainable Future. mSystems. 2021;6(4):e00925-21. doi: 10.1128/msystems.00925-21.
36. Solden L, Lloyd K, Wrighton K. The bright side of microbial dark matter: lessons learned from the uncultivated majority. Current Opinion in Microbiology. 2016;31:217-26. doi: 10.1016/j.mib.2016.04.020.
37. Jiao J-Y, Liu L, Hua Z-S, Fang B-Z, Zhou E-M, Salam N, et al. Microbial dark matter coming to light: challenges and opportunities. National Science Review. 2021;8(3):nwaa280. doi: 10.1093/nsr/nwaa280.
38. Louca S, Polz MF, Mazel F, Albright MBN, Huber JA, O'Connor MI, et al. Function and functional redundancy in microbial systems. Nature Ecology & Evolution. 2018;2(6):936-43. doi: 10.1038/s41559-018-0519-1.
39. Al Ali AA, Naddeo V, Hasan SW, Yousef AF. Correlation between bacterial community structure and performance efficiency of a full-scale wastewater treatment plant. J Water Process Eng. 2020;37:101472. doi: 10.1016/j.jwpe.2020.101472.
40. Xu R-z, Fang S, Zhang L, Huang W, Shao Q, Fang F, et al. Distribution patterns of functional microbial community in anaerobic digesters under different operational circumstances: A review. Bioresource Technology. 2021;341:125823. doi: 10.1016/j.biortech.2021.125823.

41. Zhang B, Ning D, Yang Y, Van Nostrand JD, Zhou J, Wen X. Biodegradability of wastewater determines microbial assembly mechanisms in full-scale wastewater treatment plants. Water Research. 2020;169:115276. doi: 10.1016/j.watres.2019.115276.
42. Escudié R, Cresson R, Delgenès J-P, Bernet N. Control of start-up and operation of anaerobic biofilm reactors: An overview of 15 years of research. Water Research. 2011;45(1):1-10. doi: 10.1016/j.watres.2010.07.081.
43. Bagchi S, Biswas R, Nandy T. Start-up and stabilization of an Anammox process from a non-acclimatized sludge in CSTR. Journal of Industrial Microbiology and Biotechnology. 2010;37(9):943-52. doi: 10.1007/s10295-010-0743-4.
44. Ji J, Peng Y, Wang B, Mai W, Li X, Zhang Q, et al. Effects of salinity build-up on the performance and microbial community of partial-denitrification granular sludge with high nitrite accumulation. Chemosphere. 2018;209:53-60. doi: 10.1016/j.chemosphere.2018.05.193.
45. Dueholm MKD, Nierychlo M, Andersen KS, Rudkjøbing V, Knutsson S, Arriaga S, et al. MiDAS 4: A global catalogue of full-length 16S rRNA gene sequences and taxonomy for studies of bacterial communities in wastewater treatment plants. Nat Commun. 2022;13(1):1908. doi: 10.1038/s41467-022-29438-7.
46. McCarty NS, Ledesma-Amaro R. Synthetic Biology Tools to Engineer Microbial Communities for Biotechnology. Trends in Biotechnology. 2019;37(2):181-97. doi: 10.1016/j.tibtech.2018.11.002.
47. Smith K, Guo S, Zhu Q, Dong X, Liu S. An evaluation of the environmental benefit and energy footprint of China's stricter wastewater standards: Can benefit be increased? J Clean Prod. 2019;219:723-33. doi: 10.1016/j.jclepro.2019.01.204.
48. Faisal M, Muttaqi KM, Sutanto D, Al-Shetwi AQ, Ker PJ, Hannan MA. Control technologies of wastewater treatment plants: The state-of-the-art, current challenges, and future directions. Renewable and Sustainable Energy Reviews. 2023;181:113324. doi: 10.1016/j.rser.2023.113324.
49. Cao G, Yang L, Liu L, Ma Z, Wang J, Bi J. Environmental incidents in China: Lessons from 2006 to 2015. Science of The Total Environment. 2018;633:1165-72. doi: 10.1016/j.scitotenv.2018.03.271.
50. Du W-J, Lu J-Y, Hu Y-R, Xiao J, Yang C, Wu J, et al. Spatiotemporal pattern of greenhouse gas emissions in China’s wastewater sector and pathways towards carbon neutrality. Nature Water. 2023;1(2):166-75. doi: 10.1038/s44221-022-00021-0.
51. Kosek K, Luczkiewicz A, Fudala-Książek S, Jankowska K, Szopińska M, Svahn O, et al. Implementation of advanced micropollutants removal technologies in wastewater treatment plants (WWTPs) - Examples and challenges based on selected EU countries. Environmental Science & Policy. 2020;112:213-26. doi: 10.1016/j.envsci.2020.06.011.
52. Palumbi SR. Evolution explosion: how humans cause rapid evolutionary change: WW Norton & Company; 2002.
53. Fernandez-Cassi X, Scheidegger A, Bänziger C, Cariti F, Tuñas Corzon A, Ganesanandamoorthy P, et al. Wastewater monitoring outperforms case numbers as a tool

to track COVID-19 incidence dynamics when test positivity rates are high. Water Research. 2021;200:117252. doi: 10.1016/j.watres.2021.117252.

54. Larsen DA, Wigginton KR. Tracking COVID-19 with wastewater. Nature Biotechnology. 2020;38(10):1151-3. doi: 10.1038/s41587-020-0690-1.

55. Toribio-Avedillo D, Gómez-Gómez C, Sala-Comorera L, Rodríguez-Rubio L, Carcereny A, García-Pedemonte D, et al. Monitoring influenza and respiratory syncytial virus in wastewater. Beyond COVID-19. Science of The Total Environment. 2023;892:164495. doi: 10.1016/j.scitotenv.2023.164495.

56. Sims N, Kasprzyk-Hordern B. Future perspectives of wastewater-based epidemiology: Monitoring infectious disease spread and resistance to the community level. Environment International. 2020;139:105689. doi: 10.1016/j.envint.2020.105689.

57. Lanza G, Ferdows K, Kara S, Mourtzis D, Schuh G, Váncza J, et al. Global production networks: Design and operation. CIRP Annals. 2019;68(2):823-41. doi: 10.1016/j.cirp.2019.05.008.

58. Liwarska-Bizukojc E, Bizukojc M, Marcinkowski A, Doniec A. The conceptual model of an eco-industrial park based upon ecological relationships. J Clean Prod. 2009;17(8):732-41. doi: 10.1016/j.jclepro.2008.11.004.

59. Rajesh R. Exploring the sustainability performances of firms using environmental, social, and governance scores. J Clean Prod. 2020;247:119600. doi: 10.1016/j.jclepro.2019.119600.

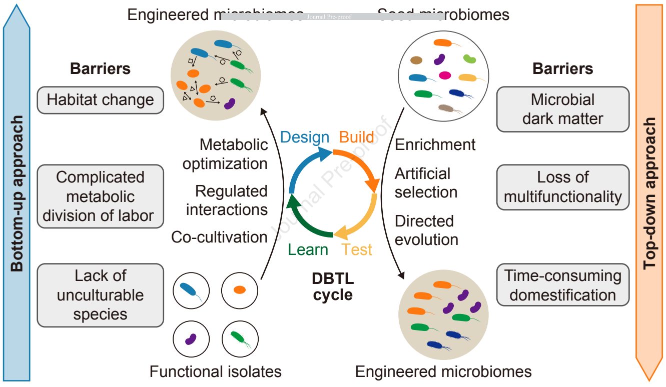
60. Ascensão F, Fahrig L, Clevenger AP, Corlett RT, Jaeger JAG, Laurance WF, et al. Environmental challenges for the Belt and Road Initiative. Nature Sustainability. 2018;1(5):206-9. doi: 10.1038/s41893-018-0059-3.

Figure captions

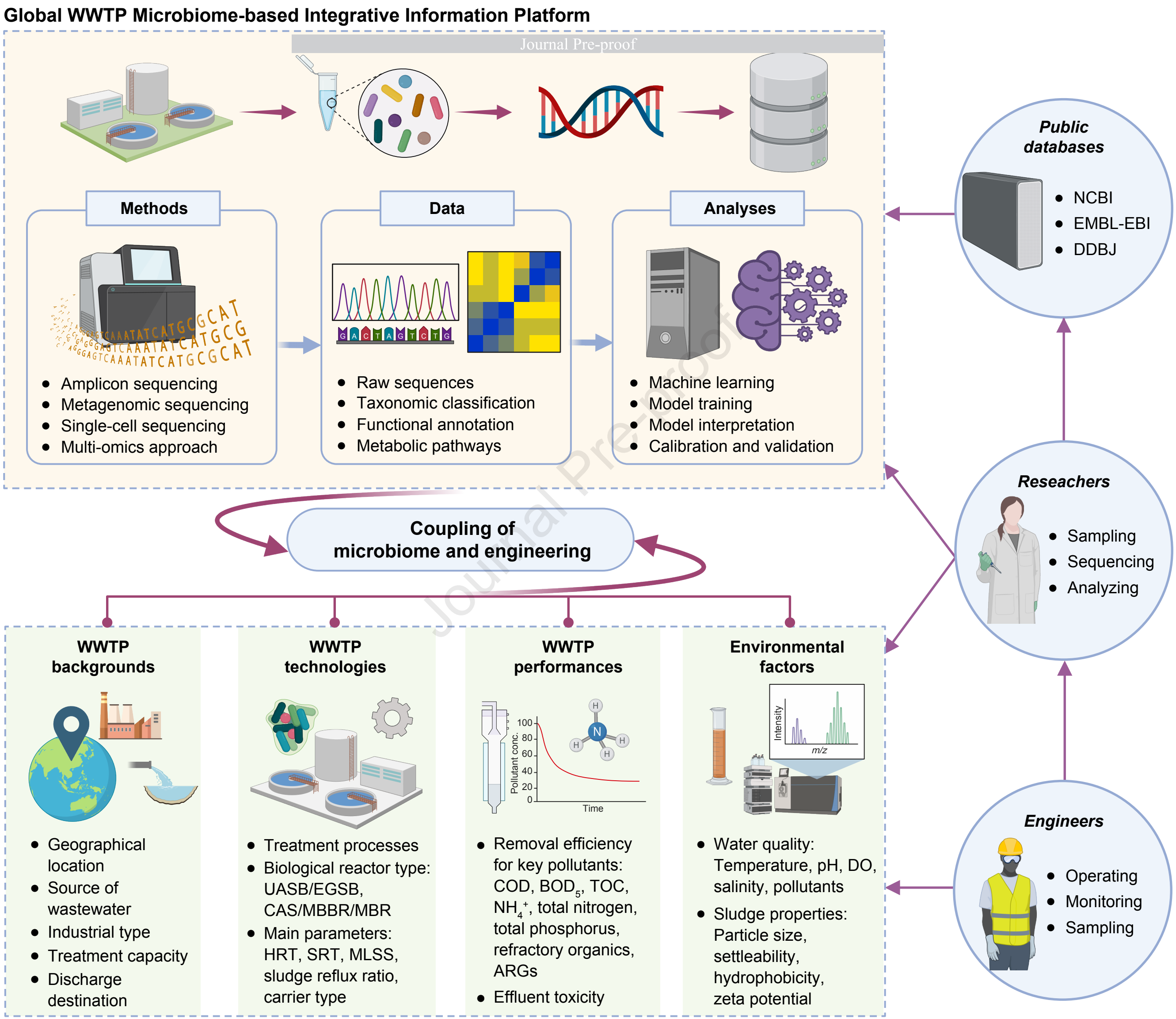
Fig. 1. The bottom-up and top-down approaches applied in microbiome engineering for wastewater treatment and their potential barriers. DBTL cycle: “design-build-test-learn” cycle.

Fig. 2. The conceptual diagram of Global WWTP Microbiome-based Integrative Information Platform. The yellow panel illustrates the structure of microbiome information and in-depth analysis processes within the Platform; the green panels describe the main contents of engineering information; and the blue panels denote the possible data sources and cooperative contributions from relevant parties.

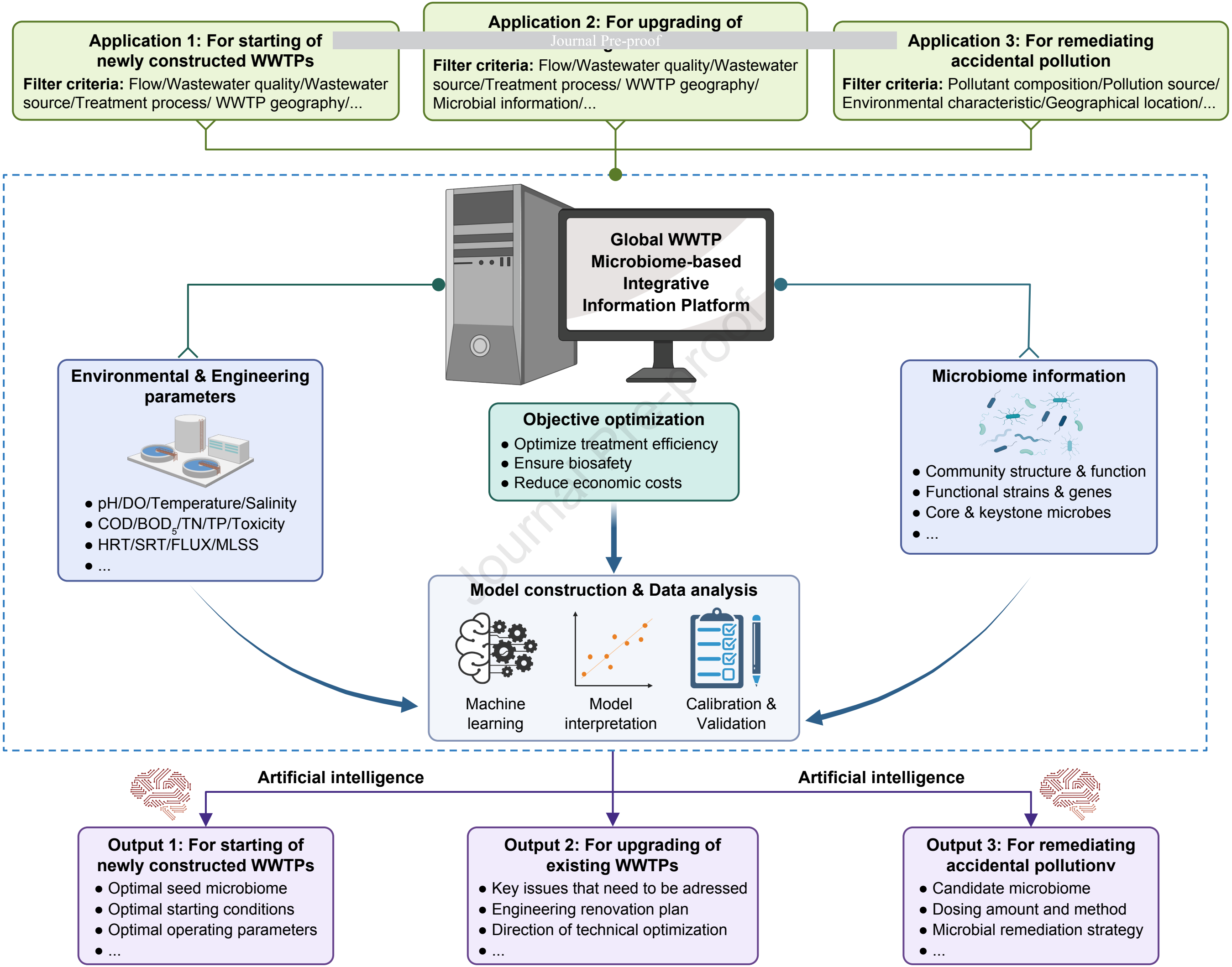
Fig. 3. The application scenarios of Global WWTP Microbiome-based Integrative Information Platform. This panel illustrates the input and output contents and the working pattern of the Platform in the three application scenarios: starting up new WWTPs, upgrading existing inefficient WWTPs, and remediating accidental environmental pollution.



Bottom-up approach
Barriers
Habitat change
Complicated metabolic division of labor
Lack of unculturable species
Metabolic optimization
Regulated interactions
Co-cultivation
Functional isolates
Design
Build
Test
Learn
DBTL cycle
Enrichment
Artificial selection
Directed evolution
Engineered microbiomes
Barriers
Microbial dark matter
Loss of multifunctionality
Time-consuming domestification
Top-down approach



Global WWTP Microbiome-based Integrative Information Platform
Methods
Amplicon sequencing
Metagenomic sequencing
Single-cell sequencing
Multi-omics approach
Data
Raw sequences
Taxonomic classification
Functional annotation
Metabolic pathways
Analyses
Machine learning
Model training
Model interpretation
Calibration and validation
Public databases
NCBI
EMBL-EBI
DDBJ
Reseachers
Sampling
Sequencing
Analyzing
Engineers
Operating
Monitoring
Sampling
Coupling of microbiome and engineering
WWTP backgrounds
Geographical location
Source of wastewater
Industrial type
Treatment capacity
Discharge destination
WWTP technologies
Treatment processes
Biological reactor type: UASB/EGSB, CAS/MBBR/MBR
Main parameters: HRT, SRT, MLSS, sludge reflux ratio, carrier type
WWTP performances
Pollutant conc.
Time
Removal efficiency for key pollutants: COD, BOD_5, TOC, NH_4^+, total nitrogen, total phosphorus, refractory organics, ARGs
Effluent toxicity
Environmental factors
Intensity
m/z
Water quality: Temperature, pH, DO, salinity, pollutants
Sludge properties: Particle size, settleability, hydrophobicity, zeta potential


Application 1: For starting of newly constructed WWTPs
Filter criteria: Flow/Wastewater quality/Wastewater source/Treatment process/ WWTP geography/...
Application 2: For upgrading of
Filter criteria: Flow/Wastewater quality/Wastewater source/Treatment process/ WWTP geography/ Microbial information/...
Application 3: For remediating accidental pollution
Filter criteria: Pollutant composition/Pollution source/ Environmental characteristic/Geographical location/...
Global WWTP Microbiome-based Integrative Information Platform
Environmental & Engineering parameters
pH/DO/Temperature/Salinity
COD/BOD5/TN/TP/Toxicity
HRT/SRT/FLUX/MLSS
...
Objective optimization
Optimize treatment efficiency
Ensure biosafety
Reduce economic costs
Microbiome information
Community structure & function
Functional strains & genes
Core & keystone microbes
...
Model construction & Data analysis
Machine learning
Model interpretation
Calibration & Validation
Artificial intelligence
Artificial intelligence
Output 1: For starting of newly constructed WWTPs
Optimal seed microbiome
Optimal starting conditions
Optimal operating parameters
...
Output 2: For upgrading of existing WWTPs
Key issues that need to be adressed
Engineering renovation plan
Direction of technical optimization
...
Output 3: For remediating accidental pollutionv
Candidate microbiome
Dosing amount and method
Microbial remediation strategy
...

Highlights

- A "Global WWTP Microbiome-based Integrative Information Platform" is proposed.
- The Platform employs AI-driven modeling and analyzing for WWTP-relevant parties.
- The Platform aims to enhance the biodegradation efficiency of pollutants on Earth.
- The Platform will be of significance for our human society and natural environment.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: