



UNIVERSITÀ DI PARMA

UNIVERSITA' DEGLI STUDI DI PARMA

DOTTORATO DI RICERCA IN
SCIENZE DEGLI ALIMENTI

CICLO XXXVI

Leveraging Lactic Acid Bacteria for New Sustainable Processes

Coordinatore:

Chiar.ma Prof.ssa Chiara Dall'Asta

Tutore:

Chiar.mo Prof. Erasmo Neviani

Co-tutori:

Chiar.ma Prof.ssa Camilla Lazzi

Chiar.ma Prof.ssa Alessia Levante

Dottoranda: Laura Troiani

Anni Accademici 2020/2021 – 2022/2023

*“Le savant n’est pas l’homme qui fournit les vraies réponses,
c’est celui qui pose les vraies questions.”*

C. Lévi-Strauss

Al mio papà

TABLE OF CONTENTS

Abstract	III
Sommario	V
Introduction	1
Rationale and Aims	13
Chapter 1 - Zooming into the metabolic biodiversity of a microbial core community of food origin.....	15
Chapter 2 - Understanding of lactic acid bacteria tolerance to extreme environments and antimicrobial compounds	57
Chapter 3 - Machine Learning approach to explore the microorganisms' biodiversity: bridging experimental datasets and predictive models.....	87
Chapter 4 - Optically pure D-lactic acid production: lactic acid bacteria as key players in the bioplastic industry	111
Chapter 5 - Investigation of the capabilities of <i>Enterococcus faecalis</i> strains to use glycerol as alternative carbon source.....	137
Conclusions	163
About the author.....	169
Scientific Activities	171

Abstract

Microorganisms are present in every ecosystem, whose balance and functionality might strongly depend on microbial diversity. Biodiversity is changing worldwide because of anthropogenic and natural drivers with several consequences on our planet's well-being and human existence itself. Recognising the importance of biodiversity, Culture Collections and microbial Biological Resource Centres have been created to preserve microbial diversity and facilitate research and industrial purposes. The University of Parma's Culture Collection is mainly composed of lactic acid bacteria (LAB) isolated from various food matrices. LAB are a wide group of ubiquitous Gram-positive bacteria with remarkable versatility and relevance, especially in food and health, thanks to their metabolic diversity and adaptability to different environmental conditions. Changes in phenotype depend on the complex interactions between environmental factors and genotype, and the understanding of LAB's phenotypic features is crucial for exploiting their potential in diverse industries. Current technologies in genome sequencing, bioinformatics, and high-throughput screening techniques applied to large collections of microbial strains give us new opportunities to move towards a sustainable economic system. Producing value-added compounds of industrial interest by biological methods is one of the strategies of the circular economy. The implementation of a circular economy model enables the recovery of waste and by-products globally generated every day, thus providing both economic and environmental benefits.

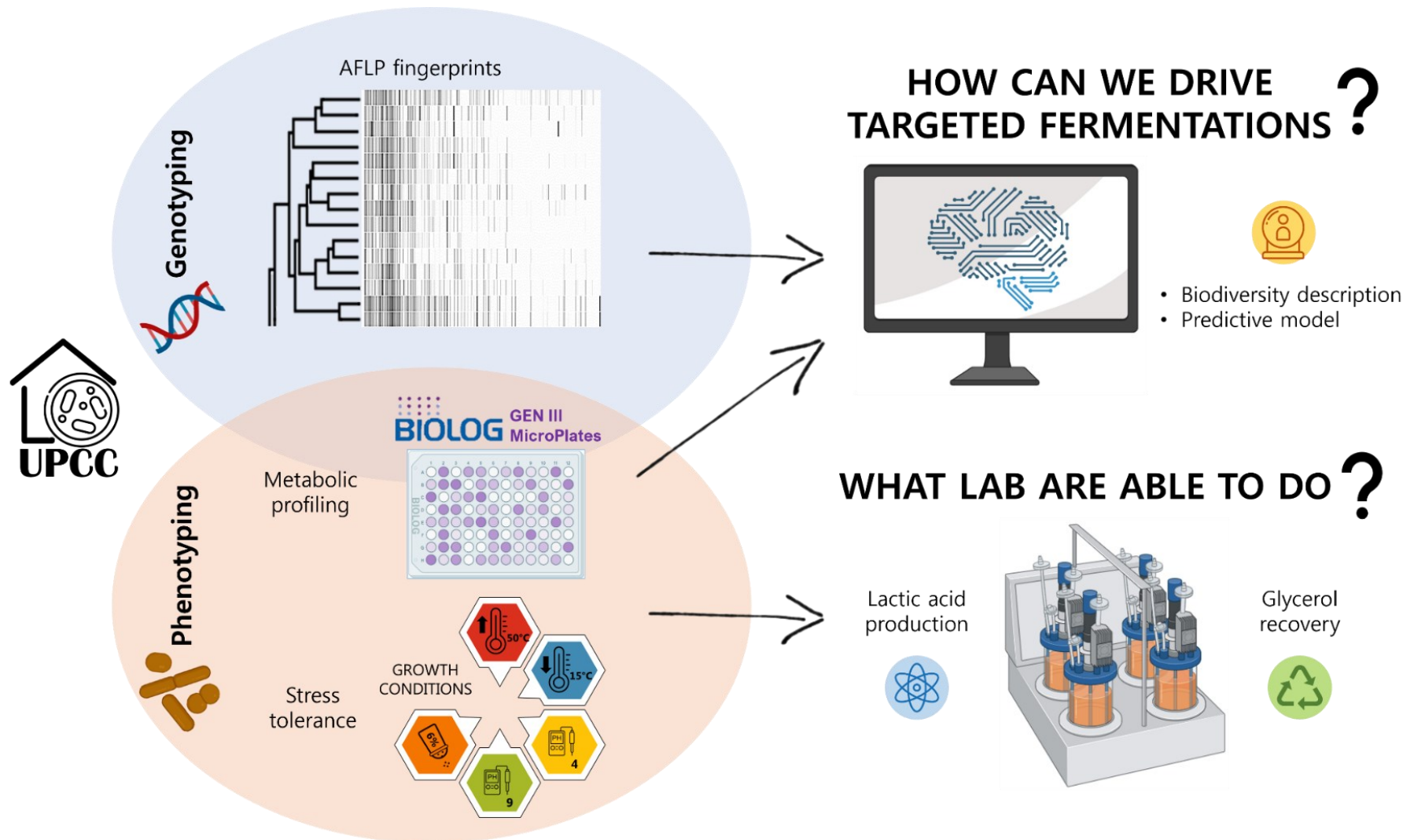
Within this scenario, the present Ph.D. thesis focused on the phenotypic characterisation of a representative selection of LAB from the University of Parma's Culture Collection. This approach allowed us to objectively describe the microbial diversity and obtain predictive machine learning models to design targeted processes. Furthermore, the phenotyping enabled the exploration of LAB's technological potential focusing on lactic acid production for bioplastics and biodiesel by-products' valorisation.

Sommario

I microrganismi sono presenti in ogni ecosistema, influenzandone fortemente l'equilibrio e le funzionalità. La biodiversità sta subendo modificazioni a livello globale a causa di fattori antropici e naturali con diverse conseguenze sul benessere del nostro pianeta e sull'esistenza umana stessa. Considerata l'importanza della biodiversità, anche nel mondo microscopico, sono state create Collezioni Microbiche e Centri di Risorse Biologiche per preservare la diversità microbica e facilitare il raggiungimento di scopi sia scientifici che industriali. La Collezione Microbica dell'Università di Parma (UPCC-University Parma Culture Collection) è costituita principalmente da batteri lattici (LAB) isolati da varie matrici alimentari. I LAB rappresentano un ampio gruppo di batteri Gram-positivi ubiquitari notevolmente versatili e rilevanti, soprattutto in ambito alimentare e sanitario, grazie alla loro diversità metabolica e adattabilità alle diverse condizioni ambientali. I cambiamenti fenotipici dipendono dalle complesse interazioni tra fattori ambientali e genotipo, e la comprensione delle caratteristiche fenotipiche dei LAB è cruciale per sfruttarne il potenziale in diversi settori. Le attuali tecnologie di sequenziamento del genoma, bioinformatica e tecniche di screening ad alto rendimento applicate a grandi collezioni microbiche ci offrono nuove opportunità per muoverci verso un sistema economico sostenibile. Produrre composti ad alto valore aggiunto di interesse industriale grazie all'impiego di microrganismi è una delle strategie dell'economia circolare. L'implementazione di un modello di economia circolare consente il recupero, mediante approcci biotecnologici, di scarti e sottoprodotti generati a livello globale ogni giorno, fornendo così benefici sia economici che ambientali.

All'interno di questo scenario, la presente tesi di Dottorato si è concentrata sulla caratterizzazione fenotipica dei batteri lattici dalla Collezione Microbica dell'Università di Parma. Questo approccio ha permesso di descrivere oggettivamente la diversità microbica e ottenere modelli predittivi al fine di ideare processi mirati. Inoltre, la fenotipizzazione ha permesso di esplorare il potenziale tecnologico dei LAB concentrandosi sulla produzione di acido lattico come precursore di bioplastiche e la valorizzazione dei sottoprodotti del biodiesel.

GRAPHICAL ABSTRACT



Introduction

Microorganisms are fundamental to life on Earth. They are present in every ecosystem affecting the balance with and among all other living beings (De Vero et al., 2019). The viability of such ecosystems heavily depends on the biodiversity they contain. But what does biodiversity mean? Biodiversity is the variety of all living entities, where variety can refer to multiple aspects (e.g. genetic, taxonomic, phylogenetic, functional...) (Díaz & Malhi, 2022), and can be described at different levels (e.g. genes, species, ecosystems...) (Ruggiero et al., 2015) and by several metrics and indices (e.g. richness, distribution of abundance...) (Daly et al., 2018). Prokaryotes (Bacteria and Archaea kingdoms) are the eldest and most ubiquitous form of life. The description of their diversity is quite controversial due to difficulties in their isolation (only 1% isolated in culture) and issues about species boundaries because of lateral gene transfer (Louca et al., 2019). Both genomics and metabolism define the ecological differentiation among bacterial strains, modelling adaptive strategies and consequently the microbial niches. Thanks to the increasing number of sequenced genomes, the concept of bacterial species is being reshaped by the substantial functional diversity revealed even among closely related strains (Koch et al., 2020). Considering three of the different ways to measure biodiversity, the Bacteria kingdom is collocated at first place for phylogenetic diversity, and second place for both amount of biomass, after the Plant kingdom, and species diversity, after the Animalia kingdom (Díaz & Malhi, 2022), despite the actual global biological diversity in terms of number of species remains uncertain and underestimated (Wiens, 2021).

Biological diversity is changing worldwide because of anthropogenic and natural drivers, such as land and sea use changes, natural resources exploitation and pollution, followed by climate change and invasive alien species (Gonzalez et al., 2023; Jaureguiberry et al., 2022). In the modern era, anthropogenic activities pose a serious threat not only to biodiversity (Wu et al., 2022), but also to human existence itself. Indeed, the maximization of short-term advantages turned out unsustainable since our actions negatively affect human life-sustaining foundations in the long term (Silvestro et al., 2022). The demographic increase by 2050 coupled with the loss of biosphere integrity, altered biogeochemical cycles, land-system change, climate change, and plastic and chemical pollution (Peixoto et al., 2022) will impose significant pressure on the Earth's final native ecosystems and the species living within them (Silvestro et al., 2022). For these reasons, in 1992, the United Nations established the Convention on Biological Diversity recognizing the importance of conserving biodiversity, including the maintenance of microbial

genetic resources of the planet (Vitorino & Bessa, 2018). Therefore, frameworks for systemic policies and action targets, such as the UN Convention on Biological Diversity's Global Biodiversity Framework (GBF) and the Sustainable Development (SD) Goals 2030 of the United Nations, are necessary to foster a shared methodology for assessing and understanding biodiversity patterns and the fundamental aspects of biological conservation. Achieving the GBF and SD objectives will require the identification of enhancements in biodiversity and ecosystem trends, alongside evaluations of evolving driver impacts (Gonzalez et al., 2023; Jaureguiberry et al., 2022). Facing environmental changes, biodiversity plays a pivotal role in sustaining and defending the resilience, stability and function of ecosystems which affect our lives (Hong et al., 2022).

In this context, Culture Collections (CCs) and microbial Biological Resource Centres (mBRCs) have been created with the primary mission of conserving microbial diversity over time and studying microorganisms' characteristics for research and industrial purposes, thus harnessing microbial biodiversity also to increase sustainability. Accordingly, these institutions ensure the responsible distribution of microbial resources and the associated data (De Vero et al., 2019; Overmann, 2015).

In Europe, more than 50 of these microbial domains (CCs, mBRCs, research institutes) from ten European countries, including Italy, are reunited under the Microbial Resource Research Infrastructure – European Research Infrastructure Consortium (MIRRI-ERIC), whose aims are the preservation, systematic investigation, provision and valorisation of microbial resources and biodiversity. It has been recognised as a “Landmark” on the European Strategy Forum on Research Infrastructures (ESFRI) Roadmap 2021 and gained the legal status of ERIC in June 2022 by the European Commission (Commission Implementing Decision (EU) 2022/1204). MIRRI examines the landscape and investigates the prospects within these domains, intending to proactively identify gaps and potential benefits to effectively meet the needs of its user communities, assisting them in maximizing the value and impact of their projects, technologies, and products. These are the pillars of MIRRI's Strategic Research & Innovation Agenda 2021-2030, a document that is dynamically reviewed in response to the global challenges, the research and innovation landscape and user needs (<https://www.mirri.org>).

The University of Parma owns a wide microbial collection (over 6000 isolates), denominated “University of Parma's Culture Collection” (UPCC), mainly composed of lactic acid bacteria (LAB) isolated from different food matrices.

LAB are a large group of Gram-positive, non-sporing, cocci or rod-shaped, anaerobic but aerotolerant bacteria, Generally Recognised As Safe (GRAS) by the Food and Drug

Administration or with the Qualified Presumption of Safety (QPS)-status by the European Food Safety Authority (O'Connell et al., 2022; Sun et al., 2015), that produce lactic acid as the primary final product of the carbohydrate metabolism (Morante-Carriel et al., 2023; Sauer et al., 2017). LAB can be discriminated based on the morphology, cocci or rod bacteria, or their metabolism, homofermentative or heterofermentative bacteria. Homofermentative LAB use the Embden-Meyerhof Pathway (EMP) to mostly produce lactic acid from sugars, while heterofermentative bacteria produce lactic acid, ethanol/acetic acid, and carbon dioxide via the Phosphoketolase Pathway (PP) (Buron-Moles et al., 2019). From a taxonomical point of view, LAB belong to several genera of the order Lactobacillales, considering the recent taxonomical update of the genus *Lactobacillus* by Zheng et al. (2020), who described 23 novel genera. However, this microbial group also includes a wide range of other genera (*Lactococcus*, *Enterococcus*, *Pediococcus*, *Leuconostoc*, *Streptococcus*, *Oenococcus*, *Tetragenococcus*, *Vagococcus*, *Weissella*, and *Carnobacterium*) thus making taxonomy itself not comprehensive when environmental niches, applications and functions of LAB are considered, as they switch their metabolism to cope with several conditions (George et al., 2018). LAB are ubiquitous, commonly associated with the food and health sector, and employed in many different applications (Duar et al., 2017). They are present wherever carbohydrates are available, including in plants and animals, even facing various adverse conditions, for instance during food processing and in the gastrointestinal tract (Mbye et al., 2020). These stress significantly and differently affect the biological functions of microorganisms, compelling them to employ several strategies for survival (Yang et al., 2021), develop species-specific adaptation mechanisms or stress responses that protect them from challenging environmental conditions (Matsumoto et al., 2022) and, sometimes, favour the secondary metabolism thus paving the way for alternative processes (Perry et al., 2022; Wang et al., 2020; Yeshi et al., 2022).

Because of this flexibility, phenotypic characterisation usually represents the initial step to identify lactic acid bacteria before delving into the molecular characterisation of their essential traits in bacterial taxonomy (Jeyagowri et al., 2023).

The term 'phenotype' does not only refer to the series of observable traits of an organism. Phenotyping aims to investigate strains' niche-specific metabolic features associated with cell physiology and growth, for instance, stress tolerance and substrate consumption, that suggest potential adaptations to specific environmental conditions. Changes in phenotype depend on the intricate interactions between environmental factors and genotype. Moreover, the exhaustive understanding of phenotype lags behind the genome characterisation (Houle et al., 2010). Koch et al. (2020) found a correlation between phenotypic variability, competition in

co-culture, and geographic distribution, thus implying an interplay between intraspecific diversity, microbial interactions, and biogeography. Hence, the ecological distribution of microorganisms is strongly shaped by the variability in their genomic and metabolic traits which affect their adaptive strategies. Environmental pressure determines the organism's evolutionary path due to prolonged exposure to specific conditions. The resulting evolutionary pressure drives the niche adaptation of many microorganisms by leaving a clear metabolic fingerprint in their genomes, as happened for many LAB (Acin-Albiac et al., 2020). LAB can metabolise several carbon sources and tolerate various environmental conditions, allowing them to colonize many habitats. The unique fermentation potential of LAB derives from a few metabolic pathways that are conveniently activated thanks to this metabolic diversity (Buron-Moles et al., 2019; Okoye et al., 2022). The microbial ability to run the metabolic paths depends on the enzymatic portfolio that is species-specific or strains-specific (Filannino et al., 2018). Current technologies in genome sequencing, bioinformatics, and high-throughput screening techniques applied to large microbial collections provide us with new opportunities in the discovery of products (Steele et al., 2019), such as precursors of bioplastics, or in the recovery of industrial waste to move towards a sustainable economic system (Bosco et al., 2021).

The implementation of a circular economy model enables the recovery of waste and by-products globally generated every day. This helps to reduce the excessive use of natural resources and mitigate the environmental impact (Venkata Mohan et al., 2020), by providing significant economic benefits (Socas-Rodríguez et al., 2021) and promoting sustainable development (Picot-Allain et al., 2020). Producing value-added compounds of industrial interest by biological methods is one of the strategies of the circular economy, especially in the agri-food sector where food waste and by-products can be used as alternative fermentation matrices (Hadj Saadoun et al., 2022; Santagata et al., 2021; Usmani et al., 2021). Understanding the metabolic capabilities of LAB and the environmental influence on their growth and physiology helps us to describe their distribution in nature, control them as we already do in health and food, but also to promote their growth in unusual matrices. Non-optimal fermentation matrices can be represented by industrial waste and by-products that most of the time need pretreatments to be suitable for microorganisms or to recover as much as possible from the biomass (Balasundaram et al., 2023). Based on the matrix and the substances to recover, pretreatments include physical, mechanical, chemical, and biological methods or combinations of them (Ojo & De Smidt, 2023; Rajesh Banu et al., 2020). These procedures alter the matrix composition and frequently yield degradation compounds that can have inhibitory effects on bacteria. Additionally, the purification of treated matrices can be expensive

thus increasing the final product's cost (Gubelt et al., 2020). Accordingly, the selection of bacteria capable of metabolising what the matrix offers as it is or able to tolerate specific environmental factors allows to avoid the substrates' treatment or microorganisms' adaptation, processes that require time and labour, especially in an industrial context where the maximization of every factor is the main mission.

Within this scenario, this Ph.D. thesis dealt with two main research questions. How can we drive targeted fermentation? What are lactic acid bacteria able to do? To answer these questions, a characterisation of a selected core of LAB from the University of Parma's Culture Collection was carried out. Results were elaborated to obtain an objective description of biodiversity and used for both the implementation of the knowledge about our microbial resources and the construction of a microbial platform designed for prediction. Moreover, the technological properties of LAB strains were evaluated, pointing out applications beyond the food field such as the bioplastic industry through the production of lactic acid or their employment for the recovery of biodiesel by-products like glycerol.

The present Ph.D. thesis was hence organized into three main activities: the first one examined the metabolic patterns of LAB strains and their behaviour when subjected to abiotic stress, the second one focused on the use of a Machine Learning model for predictive studies, and the last one dealt with the selection of promising strains for biochemicals production and waste valorisation.

The biodiversity within the microbial core selected during this Ph.D. project was investigated by using phenotypic approaches focused on the different metabolic patterns of LAB and their behaviour under non-physiological conditions. Methods such as the BIOLOG® System enable the simultaneous screening of numerous phenotypes (Acin-Albiac et al., 2020), thus reducing the experiment time and standardizing the phenotyping to all microorganisms involved. Furthermore, findings were used for training machine learning models to predict microbial functionalities exploiting the sole genomic fingerprints generated by the Amplified Fragment Length Polymorphisms (AFLP) technique. The employment of LAB has been challenging due to their metabolic diversity and confusing identity (Sun et al., 2015; Zheng et al., 2020), thus highlighting the complexity of studying a such valuable group of microorganisms. Phenotypic diversity results from intricate interactions between genotype and environment, creating a "geno-phenotype" network that remains enigmatic without comprehensive phenotypic data essential for studying these interactions. The use of a computational approach serves to reduce the number of experiments, thus making easier the selection process, restricting experiments to the validation phase, and designing targeted workflow (Sadybekov & Katritch, 2023). Machine

learning (ML) is a broad field of advanced computational and statistical methods whose models can help to describe complex data relationships and facilitate the interpretation of large sets of data in all sectors, including microbiology (Mahesh, 2019; Özel Duygan & Van Der Meer, 2022). It plays an important role in data management, exploration, and analysis, delivering valuable insights that guide decision-making and facilitating the achievement of targeted goals (Esposito et al., 2019; Steele et al., 2019). A combination of genotypic and phenotypic data provides an insight into the metabolic capabilities of industrially relevant strains, disclosing their potential application for the biotransformation of various substrates, including agro-industrial waste and by-products. Finally, the exploitation of LAB for industrial uses was evaluated through the screening of strains for the desired properties and the cultivation until the bioreactor scale.

This Ph.D. thesis seeks to highlight the effectiveness of phenotyping as an approach for the valorisation of microbial diversity and to point out the extraordinary functional potential of lactic acid bacteria for sustainable industrial processes, offering a promising path to address the pressing challenges of our time.

REFERENCES

- Acin-Albiac, M., Filannino, P., Gobetti, M., & Cagno, R. D. (2020). Microbial high throughput phenomics: The potential of an irreplaceable omics. *Computational and Structural Biotechnology Journal*, 18, 2290–2299. <https://doi.org/10.1016/j.csbj.2020.08.010>
- Balasundaram, G., Gahlot, P., Tyagi, V. K., & Kazmi, A. A. (2023). Advanced steam explosion as thermal hydrolysis process for high solids anaerobic digestion and enhanced methane yield: Proof of concept. *Sustainable Chemistry and Pharmacy*, 36, 101274. <https://doi.org/10.1016/j.scp.2023.101274>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Buron-Moles, G., Chailyan, A., Dolejs, I., Forster, J., & Mikš, M. H. (2019). Uncovering carbohydrate metabolism through a genotype-phenotype association study of 56 lactic acid bacteria genomes. *Applied Microbiology and Biotechnology*, 103(7), 3135–3152. <https://doi.org/10.1007/s00253-019-09701-6>
- Daly, A., Baetens, J., & De Baets, B. (2018). Ecological Diversity: Measuring the Unmeasurable. *Mathematics*, 6(7), 119. <https://doi.org/10.3390/math6070119>
- De Vero, L., Boniotti, M. B., Budroni, M., Buzzini, P., Cassanelli, S., Comunian, R., Gullo, M., Logrieco, A. F., Mannazzu, I., Musumeci, R., Perugini, I., Perrone, G., Pulvirenti, A., Romano, P., Turchetti, B., & Varese, G. C. (2019). Preservation, Characterization and Exploitation of Microbial Biodiversity: The Perspective of the Italian Network of Culture Collections. *Microorganisms*, 7(12), 685. <https://doi.org/10.3390/microorganisms7120685>
- Díaz, S., & Malhi, Y. (2022). Biodiversity: Concepts, Patterns, Trends, and Perspectives. *Annual Review of Environment and Resources*, 47(1), 31–63. <https://doi.org/10.1146/annurev-environ-120120-054300>
- Duar, R. M., Lin, X. B., Zheng, J., Martino, M. E., Grenier, T., Pérez-Muñoz, M. E., Leulier, F., Gänzle, M., & Walter, J. (2017). Lifestyles in transition: Evolution and natural history of the genus *Lactobacillus*. *FEMS Microbiology Reviews*, 41(Supp_1), S27–S48. <https://doi.org/10.1093/femsre/fux030>
- Esposito, S., Carputo, D., Cardi, T., & Tripodi, P. (2019). Applications and Trends of Machine Learning in Genomics and Phenomics for Next-Generation Breeding. *Plants*, 9(1), 34. <https://doi.org/10.3390/plants9010034>
- Filannino, P., Di Cagno, R., & Gobetti, M. (2018). Metabolic and functional paths of lactic acid bacteria in plant foods: Get out of the labyrinth. *Current Opinion in Biotechnology*, 49, 64–72. <https://doi.org/10.1016/j.copbio.2017.07.016>
- George, F., Daniel, C., Thomas, M., Singer, E., Guilbaud, A., Tessier, F. J., Revol-Junelles, A.-M., Borges, F., & Foligné, B. (2018). Occurrence and Dynamism of Lactic Acid Bacteria in Distinct Ecological Niches: A Multifaceted Functional Health Perspective. *Frontiers in Microbiology*, 9, 2899. <https://doi.org/10.3389/fmicb.2018.02899>
- Gonzalez, A., Chase, J. M., & O'Connor, M. I. (2023). A framework for the detection and attribution of biodiversity change. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 378(1881), 20220182. <https://doi.org/10.1098/rstb.2022.0182>

- Gubelt, A., Blaschke, L., Hahn, T., Rupp, S., Hirth, T., & Zibek, S. (2020). Comparison of Different Lactobacilli Regarding Substrate Utilization and Their Tolerance Towards Lignocellulose Degradation Products. *Current Microbiology*, 77(10), 3136–3146. <https://doi.org/10.1007/s00284-020-02131-y>
- Hong, P., Schmid, B., De Laender, F., Eisenhauer, N., Zhang, X., Chen, H., Craven, D., De Boeck, H. J., Hautier, Y., Petchey, O. L., Reich, P. B., Steudel, B., Striebel, M., Thakur, M. P., & Wang, S. (2022). Biodiversity promotes ecosystem functioning despite environmental change. *Ecology Letters*, 25(2), 555–569. <https://doi.org/10.1111/ele.13936>
- Houle, D., Govindaraju, D. R., & Omholt, S. (2010). Phenomics: The next challenge. *Nature Reviews Genetics*, 11(12), 855–866. <https://doi.org/10.1038/nrg2897>
- Jaureguiberry, P., Titeux, N., Wiemers, M., Bowler, D. E., Coscieme, L., Golden, A. S., Guerra, C. A., Jacob, U., Takahashi, Y., Settele, J., Díaz, S., Molnár, Z., & Purvis, A. (2022). The direct drivers of recent global anthropogenic biodiversity loss. *Science Advances*, 8(45), eabm9982. <https://doi.org/10.1126/sciadv.abm9982>
- Jeyagowri, N., Ranadheera, C. S., Manap, M. Y., Gamage, A., Merah, O., & Madhujith, T. (2023). Phenotypic Characterisation and Molecular Identification of Potentially Probiotic Lactobacillus sp. Isolated from Fermented Rice. *Fermentation*, 9(9), 807. <https://doi.org/10.3390/fermentation9090807>
- Koch, H., Germscheid, N., Freese, H. M., Noriega-Ortega, B., Lücking, D., Berger, M., Qiu, G., Marzinelli, E. M., Campbell, A. H., Steinberg, P. D., Overmann, J., Dittmar, T., Simon, M., & Wietz, M. (2020). Genomic, metabolic and phenotypic variability shapes ecological differentiation and intraspecies interactions of *Alteromonas macleodii*. *Scientific Reports*, 10(1), 809. <https://doi.org/10.1038/s41598-020-57526-5>
- Louca, S., Mazel, F., Doebeli, M., & Parfrey, L. W. (2019). A census-based estimate of Earth’s bacterial and archaeal diversity. *PLOS Biology*, 17(2), e3000106. <https://doi.org/10.1371/journal.pbio.3000106>
- Mahesh, B. (2019). *Machine Learning Algorithms -A Review*. <https://doi.org/10.21275/ART20203995>
- Matsumoto, S., Breniaux, M., Claisse, O., Gotti, C., Bourassa, S., Droit, A., Deleris-Bou, M., Krieger, S., Weidmann, S., Rudolf, J., & Lucas, P. (2022). The production of preconditioned freeze-dried *Oenococcus oeni* primes its metabolism to withstand environmental stresses encountered upon inoculation into wine. *International Journal of Food Microbiology*, 369, 109617. <https://doi.org/10.1016/j.ijfoodmicro.2022.109617>
- Mbye, M., Baig, M. A., AbuQamar, S. F., El-Tarabily, K. A., Obaid, R. S., Osaili, T. M., Al-Nabulsi, A. A., Turner, M. S., Shah, N. P., & Ayyash, M. M. (2020). Updates on understanding of probiotic lactic acid bacteria responses to environmental stresses and highlights on proteomic analyses. *Comprehensive Reviews in Food Science and Food Safety*, 19(3), 1110–1124. <https://doi.org/10.1111/1541-4337.12554>
- Morante-Carriel, L., Abasolo, F., Bastidas-Caldes, C., Paz, E. A., Huaquipán, R., Díaz, R., Valdes, M., Cancino, D., Sepúlveda, N., & Quiñones, J. (2023). Isolation and Characterization of Lactic Acid Bacteria from Cocoa Mucilage and Meat: Exploring Their Potential as Biopreservatives for Beef. *Microbiology Research*, 14(3), 1150–1167. <https://doi.org/10.3390/microbiolres14030077>

- O'Connell, L. M., Kelleher, P., van Rijswijk, I. M. H., de Waal, P., van Peij, N. N. M. E., Mahony, J., & van Sinderen, D. (2022). Natural Transformation in Gram-Positive Bacteria and Its Biotechnological Relevance to Lactic Acid Bacteria. *Annual Review of Food Science and Technology*, 13(1), 409–431. <https://doi.org/10.1146/annurev-food-052720-011445>
- Ojo, A. O., & De Smidt, O. (2023). Lactic Acid: A Comprehensive Review of Production to Purification. *Processes*, 11(3), 688. <https://doi.org/10.3390/pr11030688>
- Okoye, C. O., Dong, K., Wang, Y., Gao, L., Li, X., Wu, Y., & Jiang, J. (2022). Comparative genomics reveals the organic acid biosynthesis metabolic pathways among five lactic acid bacterial species isolated from fermented vegetables. *New Biotechnology*, 70, 73–83. <https://doi.org/10.1016/j.nbt.2022.05.001>
- Overmann, J. (2015). Significance and future role of microbial resource centers. *Systematic and Applied Microbiology*, 38(4), 258–265. <https://doi.org/10.1016/j.syapm.2015.02.008>
- Özel Duygan, B. D., & Van Der Meer, J. R. (2022). Recent advances in microbial community analysis from machine learning of multiparametric flow cytometry data. *Current Opinion in Biotechnology*, 75, 102688. <https://doi.org/10.1016/j.copbio.2022.102688>
- Peixoto, R. S., Voolstra, C. R., Sweet, M., Duarte, C. M., Carvalho, S., Villela, H., Lunshof, J. E., Gram, L., Woodhams, D. C., Walter, J., Roik, A., Hentschel, U., Thurber, R. V., Daisley, B., Ushijima, B., Daffonchio, D., Costa, R., Keller-Costa, T., Bowman, J. S., ... Berg, G. (2022). Harnessing the microbiome to prevent global biodiversity loss. *Nature Microbiology*, 7(11), Article 11. <https://doi.org/10.1038/s41564-022-01173-1>
- Perry, E. K., Meirelles, L. A., & Newman, D. K. (2022). From the soil to the clinic: The impact of microbial secondary metabolites on antibiotic tolerance and resistance. *Nature Reviews Microbiology*, 20(3), Article 3. <https://doi.org/10.1038/s41579-021-00620-w>
- Rajesh Banu, J., Merrylin, J., Mohamed Usman, T. M., Yukesh Kannah, R., Gunasekaran, M., Kim, S.-H., & Kumar, G. (2020). Impact of pretreatment on food waste for biohydrogen production: A review. *International Journal of Hydrogen Energy*, 45(36), 18211–18225. <https://doi.org/10.1016/j.ijhydene.2019.09.176>
- Ruggiero, M. A., Gordon, D. P., Orrell, T. M., Bailly, N., Bourgoin, T., Brusca, R. C., Cavalier-Smith, T., Guiry, M. D., & Kirk, P. M. (2015). A Higher Level Classification of All Living Organisms. *PLOS ONE*, 10(4), e0119248. <https://doi.org/10.1371/journal.pone.0119248>
- Sadybekov, A. V., & Katritch, V. (2023). Computational approaches streamlining drug discovery. *Nature*, 616(7958), Article 7958. <https://doi.org/10.1038/s41586-023-05905-z>
- Sauer, M., Russmayer, H., Grabherr, R., Peterbauer, C. K., & Marx, H. (2017). The Efficient Clade: Lactic Acid Bacteria for Industrial Chemical Production. *Trends in Biotechnology*, 35(8), 756–769. <https://doi.org/10.1016/j.tibtech.2017.05.002>
- Silvestro, D., Gorla, S., Sterner, T., & Antonelli, A. (2022). Improving biodiversity protection through artificial intelligence. *Nature Sustainability*, 5(5), Article 5. <https://doi.org/10.1038/s41893-022-00851-6>
- Steele, A. D., Teijaro, C. N., Yang, D., & Shen, B. (2019). Leveraging a large microbial strain collection for natural product discovery. *Journal of Biological Chemistry*, 294(45), 16567–16576. <https://doi.org/10.1074/jbc.REV119.006514>

- Sun, Z., Harris, H. M. B., McCann, A., Guo, C., Argimón, S., Zhang, W., Yang, X., Jeffery, I. B., Cooney, J. C., Kagawa, T. F., Liu, W., Song, Y., Salvetti, E., Wrobel, A., Rasinkangas, P., Parkhill, J., Rea, M. C., O'Sullivan, O., Ritari, J., ... O'Toole, P. W. (2015). Expanding the biotechnology potential of lactobacilli through comparative genomics of 213 strains and associated genera. *Nature Communications*, 6(1), 8322. <https://doi.org/10.1038/ncomms9322>
- Vitorino, L. C., & Bessa, L. A. (2018). Microbial Diversity: The Gap between the Estimated and the Known. *Diversity*, 10(2), Article 2. <https://doi.org/10.3390/d10020046>
- Wang, Y.-N., Meng, L.-H., & Wang, B.-G. (2020). Progress in Research on Bioactive Secondary Metabolites from Deep-Sea Derived Microorganisms. *Marine Drugs*, 18(12), Article 12. <https://doi.org/10.3390/md18120614>
- Wiens, J. J. (2021). Vast (but avoidable) underestimation of global biodiversity. *PLOS Biology*, 19(8), e3001192. <https://doi.org/10.1371/journal.pbio.3001192>
- Wu, L., Zhang, Y., Guo, X., Ning, D., Zhou, X., Feng, J., Yuan, M. M., Liu, S., Guo, J., Gao, Z., Ma, J., Kuang, J., Jian, S., Han, S., Yang, Z., Ouyang, Y., Fu, Y., Xiao, N., Liu, X., ... Zhou, J. (2022). Reduction of microbial diversity in grassland soil is driven by long-term climate warming. *Nature Microbiology*, 7(7), Article 7. <https://doi.org/10.1038/s41564-022-01147-3>
- Yang, H., He, M., & Wu, C. (2021). Cross protection of lactic acid bacteria during environmental stresses: Stress responses and underlying mechanisms. *LWT*, 144, 111203. <https://doi.org/10.1016/j.lwt.2021.111203>
- Yeshi, K., Crayn, D., Ritmejeriyè, E., & Wangchuk, P. (2022). Plant Secondary Metabolites Produced in Response to Abiotic Stresses Has Potential Application in Pharmaceutical Product Development. *Molecules*, 27(1), Article 1. <https://doi.org/10.3390/molecules27010313>
- Zheng, J., Wittouck, S., Salvetti, E., Franz, C. M. A. P., Harris, H. M. B., Mattarelli, P., O'Toole, P. W., Pot, B., Vandamme, P., Walter, J., Watanabe, K., Wuyts, S., Felis, G. E., Gänzle, M. G., & Lebeer, S. (2020). A taxonomic note on the genus *Lactobacillus*: Description of 23 novel genera, emended description of the genus *Lactobacillus* Beijerinck 1901, and union of *Lactobacillaceae* and *Leuconostocaceae*. *International Journal of Systematic and Evolutionary Microbiology*, 70(4), 2782–2858. <https://doi.org/10.1099/ijsem.0.004107>

WEB REFERENCES

<https://www.mirri.org>

Rationale and Aims

This Ph.D. project arose from three scientific and societal issues: the objectification of biodiversity and the prediction of microbial behaviour, the production of industrial biochemicals, and finally the recovery of industrial waste.

The first part of the research investigated the metabolic profiles and tolerance to abiotic stress of lactic acid bacteria strains. The second part of the project involved a Machine Learning model to build a microbial platform to predict the microorganisms' metabolic capabilities. The third one was aimed at selecting strains that could be industrially used for biochemical production and waste valorisation.

The final objectives of this thesis were the valorisation of the microbial collection, by revealing the metabolic diversity of microorganisms, and the exploitation of the functional potential of lactic acid bacteria for targeted fermentations.

- ❖ **Chapter 1** aimed to objectively describe the phenotypic biodiversity within a selected microbial core of lactic acid bacteria to explore the substrate utilisation capabilities and design fermentation processes according to their metabolic aptitude.
- ❖ **Chapter 2** aimed to evaluate the tolerance of lactic acid bacteria to adverse environmental conditions as well as their resistance to different antimicrobials, to finally understand the potential aptitude of microorganisms to develop targeted fermentations.
- ❖ **Chapter 3** aimed to exploit the fragment length distribution profile of bacterial genomes provided by the AFLP technology as a feature for computational predictive models.
- ❖ **Chapter 4** aimed to identify D-lactic acid-producing strains within a selected microbial core of lactic acid bacteria and boost their potentialities through cultivation in bioreactors.
- ❖ **Chapter 5** aimed to identify glycerol utilisers within a selected microbial core of lactic acid bacteria to produce lactic acid suitable for PLA production from a biodiesel by-product.

Chapter 1

Zooming into the metabolic biodiversity of a microbial core community of food origin

Laura Troiani^{1*}, Alessia Levante¹, Vincenzo Bonnici², Velia Pierdomenico², Alessandro Dal
Palù², Erasmo Neviani¹, Valentina Bernini¹, Camilla Lazzi¹

¹ Department of Food and Drug Sciences, Parma University, Parma, Italy

² Department of Mathematical, Physical and Computer Sciences, University of Parma, Parma,
Italy

Manuscript in preparation

ABSTRACT

Microorganisms constitute a vast reservoir of biodiversity in various ecological habitats and serve as important workhorses in the food, medical, and biotech industries. Therefore, the characterisation of microorganisms from culture collections is crucial for both basic and applied research. This study aimed to describe the genotypic and phenotypic biodiversity of a selected microbial core (SMC) consisting of 150 lactic acid bacteria from the University of Parma's Culture Collection (UPCC). To achieve this, ecological indices are proposed as an objective representation of strains' biodiversity, based on experimental data.

Amplified fragment length polymorphism (AFLP) methodology was employed for each strain to assess the genotypic variability. Additionally, the metabolic capabilities of the strains were tested over more than 70 different carbon sources using Biolog GENIII Microplates (BIOLOG Inc.©, USA). Genotypic profiling based on AFLP revealed consistent clustering according to species in agreement with genome-wide phylogenetic studies. However, the distribution of metabolic profiles, considering both species and substrate utilisation, exhibited greater variation, providing insights into the metabolic biodiversity of strains.

Phenotypic characterisation through these methods proves to be a powerful tool for studying microbial isolates, offering insights into their adaptation to specific niches. This knowledge can be leveraged to support advanced industrial applications, including the conversion of various types of waste and by-products, aligning with a circular economy-oriented perspective.

IMPORTANCE

In this work, aside from applying a genotyping approach, we highlighted the great potential of phenotyping to objectify the description of LAB biodiversity using simple, fast, easy output processing and cost-effective techniques compared to genomics. We remarked, through this study, that the metabolic biodiversity in the SMC is greater than that traditionally associated with the source of isolation and these findings can be extended to the entire microbial collection. Moreover, the most frequently consumed carbon sources are found in many agro-industrial wastes and by-products. We believe that the insights gained from this study have significant implications for the development of microbial platforms that, gathering information about microorganisms, enable the prediction and the manipulation of microbial behaviour to design targeted fermentation processes.

KEYWORDS

Phenotyping, Lactic acid bacteria, Genotyping, Microbial platform, Targeted fermentation, Waste recovery

ABBREVIATIONS

CCs = Culture Collections

mBRCs = microbial Biological Resource Centres

UPCC = University of Parma's Culture Collection

LAB = Lactic Acid Bacteria

GRAS = Generally Recognised As Safe

HomF = Homofermentative bacteria

HetF = Heterofermentative bacteria; F-HetF = facultatively; O-HetF = obligately

EMP = Embden-Meyerhof Pathway

PP = Phosphoketolase Pathway

SMC = Selected Microbial Core

AWCD = Average Well Colour Development

H' = Shannon diversity index

S = Richness

SR = Substrate Richness

E = Pielou's Evenness Index/substrate Evenness

D = Simpson Dominance index

D' = Simpson Dominance index

AFLP = Amplified Fragment Length Polymorphism

1. INTRODUCTION

Microorganisms are the most important representation of biodiversity in every ecological habitat. They affect the functionality of any ecosystem, keeping in healthy equilibrium the planet and all living beings. Furthermore, microorganisms have numerous applications and are involved in food, medical and biotech industries. Accordingly, to maintain natural ecosystems, for both research purposes and biotechnological exploitation, it is essential to characterise and preserve microbial biodiversity (De Vero *et al.*, 2019).

The aim of culture collections (CCs) and microbial biological resource centres (mBRCs) is to safeguard biodiversity by preserving microorganisms over time and studying all their features for research and industrial improvement, also guaranteeing a proper distribution of both microbial resources and associated data (Overmann, 2015). An essential part of this work is to understand the metabolic traits of different biotypes, which could help elucidate the interactions between them and with other biological subjects present in a specific ecosystem.

The University of Parma owns a wide microbial collection (over 6000 isolates), denominated University of Parma's Culture Collection (UPCC), mainly composed of lactic acid bacteria (LAB) isolated from different matrices, such as milk, whey, curd, cheeses, sourdough, fermented products, wastewater, meat, flour, fruits, and vegetables. LAB are a large group of Gram-positive, non-sporing, cocci or rod-shaped, preferably anaerobic but aerotolerant, Generally Recognised As Safe (GRAS) bacteria, which produce lactic acid as the major end-product of carbohydrate fermentation. LAB can be morphologically and metabolically subdivided into two groups: cocci or rods, and homofermentative or heterofermentative bacteria, respectively. Homofermentative LAB (HomF) produce mostly lactic acid from sugars via the Embden-Meyerhof Pathway (EMP), while heterofermentative bacteria (HetF; facultatively = F-HetF, or obligately = O-HetF) produce lactic acid, ethanol/acetic acid and carbon dioxide via the Phosphoketolase Pathway (PP) (Duar *et al.*, 2017; Halasz, 2009). LAB are widespread especially in the food and health sector, with many different applications (Buron-Moles *et al.*, 2019). From a taxonomical point of view, LAB represent members of the order Lactobacillales (George *et al.*, 2018), in which the genus *Lactobacillus* has been recently taxonomically updated by Zheng *et al.* (2020) who described 23 novel genera of LAB. The LAB group, however, also includes a wide range of other genera (*Lactococcus*, *Enterococcus*, *Pediococcus*, *Leuconostoc*, *Streptococcus*, *Oenococcus*, *Tetragenococcus*, *Vagococcus*, *Weissella*, and *Carnobacterium*), thus taxonomy itself is not comprehensive when we talk about

environmental niches, applications and roles of LAB, since they switch their metabolism to adapt to several conditions (George *et al.*, 2018).

Current technologies in genome sequencing, bioinformatics, and high-throughput screening techniques applied to large microbial strain collections give us new opportunities in natural product discovery (Steele *et al.*, 2019), such as aromatic compounds to enhance the organoleptic properties of food (Bancalari *et al.*, 2017; Lorn *et al.*, 2021), exopolysaccharides with several roles (Bancalari *et al.*, 2019), antimicrobial molecules to increase food shelf-life (Kuleaşan & Cakmakçi, 2002; Ricci, Bernini, *et al.*, 2019), bioactive compounds (Ricci, Cirlini, *et al.*, 2019), precursors of bioplastics to move towards a sustainable economy (Bosco *et al.*, 2021), and enzymes (Padmavathi *et al.*, 2018). Producing products of industrial interest by biological methods is one of the strategies of the circular economy, especially in the agri-food sector (Hadj Saadoun *et al.*, 2022; Santagata *et al.*, 2021; Usmani *et al.*, 2021). The implementation of a circular economy model would allow the recovery of most of the food waste and by-products produced every day all around the world and reduce the over-exploitation of natural resources and the environment (Venkata Mohan *et al.*, 2020), by providing significant economic benefits (Socas-Rodríguez *et al.*, 2021) and promoting sustainable development too (Picot-Allain *et al.*, 2020).

Environmental and societal changes in the last decades represent a factor of risk in terms of biodiversity loss, including microbial biodiversity, as reported in some ecosystems (L. Wu *et al.*, 2022). In this context, it becomes crucial the availability of CCs so that biodiversity can be preserved and leveraged to increase and improve the sustainability of industrial processes. Microbes can be used as a weapon against the depletion and deterioration of our planet. It is constantly indagated the quality and safety of waste to understand if they can be recycled and valorised for human and animal consumption (Socas-Rodríguez *et al.*, 2021). At the same time, it is likewise important to know which substrates microorganisms can metabolise to carry out targeted fermentation processes and maximize the end-product yield. Like most microorganisms, LAB prefer carbohydrates as the principal carbon and energy source for growing and producing biomass (Buron-Moles *et al.*, 2019). Within the wide range of carbon sources assimilated by strains, there could be potential alternatives for producing metabolites and biomass through fermentation (Mekonnen *et al.*, 2019).

Over the last two decades, high-throughput capabilities, lower costs, and advances in DNA sequencing and bioinformatics have generated an exponential increase in available data on microbial genome sequences (Steele *et al.*, 2019). Studies of comparative genomics and metagenomics have been extensively employed to indagate the potential functionality of

microbes, while (meta)transcriptomics especially focused on mechanisms for niche adaptation. Metabolic traits are related to the phenotype expression but, whilst data analysis methodologies and technologies for sequence-based omics quickly evolved, phenotype profiling stayed somehow hampered by some limitations (Houle *et al.*, 2010). Techniques like the Biolog System made possible to screen many phenotypes concurrently, introducing phenomics as the latest of the omics techniques (Acin-Albiac *et al.*, 2020).

LAB have evolved the ability to metabolise different carbon sources, which provided them important benefits of colonizing many ecosystems. The unique fermentation potential of LAB derives from a few metabolic pathways right activated by this metabolic variety (Buron-Moles *et al.*, 2019). The microbial ability to run certain metabolic paths depends on the enzymes set and it is species-specific or strains-specific (Di Cagno *et al.*, 2013). There are several commercial kits to know the substrate utilisation profiles of bacteria, which are less expensive and less difficult to use compared to culture-independent approaches (Mekonnen *et al.*, 2019). Among others, the Biolog system evaluates the capability of isolates to utilise and/or degrade certain molecules and allows to measure the degree of substrate assimilation by microbes based on the absorbance detection of a redox dye in the substrate. In this study, the Average Well Colour Development (AWCD) is used as an index of the total bioactivity of each strain. Additionally, other indices such as Shannon diversity index (H), Substrate Richness (SR), and substrate Evenness (E) were calculated (Acin-Albiac *et al.*, 2020; Dubey *et al.*, 2022; Feigl *et al.*, 2017; Ge *et al.*, 2018).

This work aimed to investigate the phenotypic biodiversity of a selected microbial core (SMC) of 150 LAB strains to explore the microorganisms' metabolic functionalities. To reach this goal, we proposed an easy and objective representation of genotypic and phenotypic biodiversity in microbial collections using ecological indices that derive from experimental data. The knowledge acquired through this approach about microbial metabolic aptitude can set the basis for a targeted selection of strains from a biobank useful to various industrial applications such as building microbial platforms for agro-industrial waste and by-product valorisation.

2. MATERIALS AND METHODS

2.1. Bacterial strain cultures preparation

A set of 150 LAB strains was selected from the University of Parma's Culture Collection (UPCC), by ensuring a large biodiversity of the sampled strains. As reported in Table 3, the screening involved species belonging to 12 genera, previously isolated across a 10-year temporal range (2012-2022) from several food matrices, especially of dairy origin.

Cultures were maintained as frozen stocks (-80°C) in their optimal growth medium added with 15% glycerol (w/v). Prior to use, they were propagated three times with 3% (v/v) inoculum in optimal growth media and incubated overnight at the optimal temperature in anaerobic conditions (AnaeroGen, ThermoScientific, Landsmeer, The Netherlands). All cultures were observed by microscopy to confirm their purity.

Table 3 Lactic acid bacteria selected from the UPCC for this study. HomF, Homofermentative metabolism; F-HetF, Facultatively Heterofermentative metabolism; O-HetF, Obligately Heterofermentative metabolism. Reference: Boone *et al.* (2001); Zheng *et al.* (2020); <https://atcc.org/>; <https://bacdiv.dsmz.de/>.

Genus	Species	No. strains	Medium ⁽¹⁾	Temp.	Metabolism	Isolation source
<i>Companilactobacillus</i>	<i>farciminis</i>	2	MRS	30°C	HomF	Sourdough
<i>Enterococcus</i>	<i>avium</i>	1	TSB	37°C	HomF	Sourdough
	<i>faecalis</i>	10				Cheese, cream, curd, milk
	<i>faecium</i>	6				Cheese, milk, whey starter
<i>Lacticaseibacillus</i>	<i>casei</i>	1	MRS	37°C	F-HetF	Cheese, milk
	<i>paracasei</i>	15				Cheese, milk
	<i>rhamnosus</i>	12				Cheese, curd, milk
	<i>zeae</i>	8				Cheese
<i>Lactiplantibacillus</i>	<i>plantarum</i>	6	MRS	37°C	F-HetF	Cheese
<i>Lactobacillus</i>	<i>amylolyticus</i>	1	MRS	37°C	HomF	Curd
	<i>delbrueckii</i>	11				Curd, milk, natural whey starter
	<i>gasseri</i>	1				Milk
	<i>helveticus</i>	9				Curd, milk, natural whey starter
<i>Lactococcus</i>	<i>lactis</i>	1	M17	30°C	HomF	Milk
<i>Latilactobacillus</i>	<i>curvatus</i>	10	MRS	30°C	F-HetF	Cheese

<i>Lentilactobacillus</i>	<i>diolivorans</i>	1	MRS	30°C	O-HetF	Milk
	<i>hilgardii</i>	2				Milk
	<i>otakiensis</i>	2				Milk
<i>Leuconostoc</i>	<i>citreum</i>	8	MRS	30°C	O-HetF	Milk, sourdough
	<i>lactis</i>	1				Milk
	<i>mesenteroides</i>	6				Cheese, milk
	<i>pseudomesenteroides</i>	1				Milk
<i>Levilactobacillus</i>	<i>brevis</i>	7	MRS	30°C	O-HetF	Sourdough
<i>Limosilactobacillus</i>	<i>fermentum</i>	8	MRS	37°C	O-HetF	Cheese, curd, whey starter
<i>Streptococcus</i>	<i>macedonicus</i>	2	M17	42°C	HomF	Milk
	<i>pasteurianus</i>	2	TSB	37°C		Milk
	<i>porcorum</i>	2				Milk
	<i>thermophilus</i>	9				Milk, cheese
	<i>uberis</i>	5				Milk

- 1) MRS and TSB media were purchased from VWR International bvba (Leuven, Belgium), while M17 medium from Merck (Darmstadt, Germany).

2.2. Genotypic characterisation

Automated DNA extraction was carried out by using the Kit “Pure Food Pathogen” of Maxwell® RSC Instrument (Promega, Milan, Italy) and following the supplied protocol. The 16S rRNA gene V1-V3 region was amplified by PCR composed of 10 µl of GoTaq® Master Mix 2X (Promega, Madison, WI, USA), 7 µl of sterile MilliQ water, 1 µl of 10 µM 46Fw Primer (5'-GCYTAACACATGCAAGTCGA-3'), 1 µl of 10 µM 536Rv Primer (5'-GTATTACCGCGGCTGCTGG-3'), and 1 µl of template DNA, with a GeneAmp® PCR System 2700 (Applied Biosystem, Foster City, CA, USA). The following thermal cycling conditions were used: denaturation at 95°C for 10 min, 30 cycles of 95°C for 30 sec, 59°C for 30 sec and 72°C for 30 sec for the annealing phase, and the final elongation step at 72°C for 10 min. Both the DNA and the PCR products were separated on 1% agarose gel electrophoresis in 1×TAE buffer and visualized by UV transillumination after staining with RedSafe DNA Gel Stain (Invitrogen Life Technologies, Carlsbad, California, USA). A GeneRuler 1Kb plus DNA ladder (Invitrogen Life Technologies, Carlsbad, California, USA) was used as molecular size standard. The PCR products were sequenced by MACROGEN Europe (Amsterdam, The Netherlands) and BLASTn (<https://blast.ncbi.nlm.nih.gov/>) was used for sequence identification to confirm the species.

To find possible clones within the Selected Microbial Core (SMC), the Amplified Fragment Length Polymorphism (AFLP) analysis was performed as suggested by Bertani *et al.* (2019). Restriction-Ligation (R/L) reactions, until a final volume of 50 µl, were composed of 10 µl of 250 ng/µl BSA (New England Biolabs, Ipswich, MA, USA), 5 µl of 1× T4 DNA ligase buffer with 1 µl of 1 mM ATP (New England Biolabs, Ipswich, MA, USA), 1 µl of 50 pmol/µl MseI adaptor, 1 µl of 5 pmol/µl EcoRI adaptor, 1 µl of 10 mM ATP (Invitrogen S.R.L., Milano, Italy), 0.5 µl of 10 U/µl MseI (New England Biolabs, Ipswich, MA, USA), 0.25 µl of 20 U/µl EcoRI (New England Biolabs, Ipswich, MA, USA), 0.1 µl of 200U/µl of T4 DNA ligase (New England Biolabs, Ipswich, MA, USA) and 15 µl of template DNA (500 ng). Sequences of the two adapters (BMR Genomics, Padova, Italy) are shown in Table 4. The reaction was incubated at 37°C for 4 hours. Two amplification reactions were executed after incubation (Tab. 4). The first pre-selective PCR included 10 µl GoTaq® Master Mix 2X (Promega, Madison, WI, USA), 0,6 µl of 10 µM EcoRI-0 Primer, 0.6 µl of 10 µM MseI-0 Primer and 2 µl of digested-ligated DNA, until a final volume of 20 µl. Amplified samples were diluted 1:10 in TE (10 mM TRIS HCl at pH=8, 0.1 mM EDTA at pH=8). The second selective amplification was composed of 10 µl GoTaq® Master Mix 2X (Promega, Madison, WI, USA), 0.6 µl of 10 µM EcoRI-A Primer, 0.6 µl of 10 µM MseI-C/MseI-A Primer and 2 µl of diluted DNA, until a final volume of 20 µl. All PCRs were performed in a GeneAmp® PCR System 2700 (Applied Biosystem, Foster City, CA, USA) by following the thermocycler program of Bertani *et al.* (2019). Both amplifications were sequentially confirmed by electrophoresis on a 1.2% agarose gel in 1×TAE buffer at 90 V/cm for 20 min and visualized as previously described for 16S rRNA gene PCR products.

Table 4 Primers used for AFLP analysis.

Primer name	Sequence (5'-3')
MseI-0	GATGAGTCCTGAGTAA
EcoRI-0	GACTGCGTACCAATTTC (labelled FAM 5')
MseI-A	GATGAGTCCTGAGTAAA
MseI-C	GATGAGTCCTGAGTAAC
EcoRI-A	GACTGCGTACCAATTCA (labelled FAM 5')

To choose the most suitable bases for the selective primers, an *In Silico* AFLP (<http://insilico.ehu.eus/AFLP/>) was performed by setting the studied organism, EcoRI and MseI as Restriction/Ligation enzymes, and 50-500 bp as desired sequence length. Different primer combinations were used, based on the different species analysed. The AFLP products were

analysed by capillary electrophoresis on an external service (BMR Genomics, Padova, Italy), and the resulting electropherograms were analysed through Bionumerics Software v 8.0 (Applied Maths NV, Belgium). Type strains of the species under investigation (*S. thermophilus* LMG 6896, *L. brevis* LMG 6906, *L. plantarum* LMG 6907, *L. curvatus* LMG 9198, *E. faecium* LMG 11423, *L. fermentum* LMG 11441, *E. faecalis* LMG 11638, *L. delbrueckii bulgaricus* LMG 13551, *L. delbrueckii lactis* LMG 18223, *L. helveticus* LMG 18225, *L. rhamnosus* LMG 18243, *L. casei* LMG 23516) were purchased from the American Type Culture Collection (Manassas, Virginia, USA), and DNA was extracted from cultures of two independent biological replicates allowing to assess the repeatability of the AFLP technique.

The AFLP protocol was applied to DNA extracted from type strains as well to establish the similarity percentage of profiles under which microorganisms can be considered clones (Lazzi *et al.*, 2009). AFLP profiles were compared through Bionumerics Software by setting GeneScan-500 [LIZ] as the size standard to normalize peaks. Bands from 50 to 500 bp were considered. AFLP profiles were clustered by Pearson Correlation. The similarity cut-off was set at 40%.

2.3. Phenotypic characterisation through BIOLOG® Technology

The Biolog GENIII Microplates (BIOLOG Inc.®, USA) were used to study the carbon sources utilisation, by simultaneously assaying the use of 71 carbon sources (columns 1-9) for each microorganism. All compounds tested are prefilled and dried into the microplate together with Tetrazolium Violet redox dye used as a colorimetric indicator of microbial respiration.

For the metabolic microarrays, some modifications were applied to the GENIII Microplates protocol. 1 ml of overnight cultures was collected for cell washing. It was centrifugated at 13400 rpm for 10 min and the supernatant was discarded. Cells were washed in Ringer's solution (VWR International bvba, Leuven, Belgium) twice to remove any effects of the growth medium.

For the inoculum preparation, washed cells were suspended in 12 ml of inoculation fluid IF-A (BIOLOG Inc.®, USA) to a final transmittance comprised between 90 and 98%. The average transmittance equal to 95% (= 0.02 Optical Density (OD) at 590 nm) corresponds to an initial inoculum of 10^5 cfu/ml.

The cell suspension was gently inoculated into the 96 wells with a multichannel pipette, by avoiding cross-contamination among substrates. The microplate was incubated at the optimal strain's temperature to allow the phenotypic fingerprint to form. The metabolic activity of strains was detected on the BIOLOG MicroReader Station (BIOLOG Inc.®, USA) by

absorbance reading at a wavelength of 590 nm, before incubation and after 72 h. Measurement was performed at a single end-point (72 h), and it was related to the absorbance of the reduced Tetrazolium Violet dye (purple coloured) that indirectly represents the increased metabolic activity in wells where cells used the substrate (Acin-Albiac *et al.*, 2020).

2.4. Data Analysis

2.4.1. Ecological description of the selected microbial core (SMC)

Considering the SMC as a microbial community, to estimate its overall biodiversity, common indices of microbial ecology were calculated as reported in Table 5.

Table 5 Measurements of biodiversity (Daly *et al.*, 2018).

Index	Formula	Description
Richness (S)		Total number of species in the microbial community
Log(S)		It sets the maximum value H' may achieve (Gophen, 2018; Shannon, 1948)
Shannon Diversity Index (H')	$H' = - \sum_{i=1}^n P_i * \ln P_i$	Entropy or disorder of the population $P_i = ni/N$ proportional abundance of species i where: ni = number of strains of species i N = total number of strains
Pielou's Evenness Index (E)	$E = H' / \log S$	Equitability of the population's species abundance distribution
Simpson Dominance Index (D)	$D = \sum_{i=1}^n P_i^2$	Probability that two strains randomly picked up from the community belong to the same species
Simpson Dominance Index (D')	$D' = 1/D$	It weights the common species more than the rare ones

2.4.2. Metabolic functional diversity of strains

Phenotypic raw data were elaborated with Rstudio (R version 4.3.1, Package “pheatmap” version 1.0.12) for statistical analysis and data visualization of assays, using the default algorithms for clustering. Having single end-point measurements, the Average Well Colour Development (AWCD) was calculated as reported in the literature (Acin-Albiac *et al.*, 2020; Dubey *et al.*, 2022). To reduce the noise levels, all absorbance values of carbon sources utilisation (columns 1-9) were referred against the negative control well (A1) and subsequently, all divided by the respective AWCD (Tab. 6). Negative values were set to 0. Normalized data were used for statistical analysis.

In parallel, the metabolic functional diversity of strains was estimated employing common indices of ecological diversity (Tab. 6), largely used for single end-point Biolog data analysis (Acin-Albiac *et al.*, 2020; Feigl *et al.*, 2017; Ge *et al.*, 2018). Data are reported as the average indices calculated within each species and were not calculated for species composed of less than 2 strains because they contribute less to the core community's diversity than others with many individuals (Daly *et al.*, 2018).

Significant differences observed between α -D-Glucose consumption with respect to other C sources were calculated by the Student's T-test ($\alpha = 0.05$).

Table 6 Data calculation to obtain values for statistical analysis and metabolic functional diversity indices (Acin-Albiac *et al.*, 2020; Dubey *et al.*, 2022; Feigl *et al.*, 2017; Ge *et al.*, 2018).

Estimation	Formula	C source utilisation
Average Well Colour Development (AWCD)	$AWCD = \sum_{i=1}^n (ODi - R)/n$	ODi = raw OD ₅₉₀ value of each well from 1 to 9 column R = OD ₅₉₀ value of negative control well n = number of substrates analysed (71)
Normalization	$Rsi = (ODi - R)/AWCD$	
Substrates Richness (SR)		Total number of substrates utilised by each strain (normalized value at 72h ≥ 1)
% SR	$\% SR = \left(\frac{SR}{n}\right) * 100$	Percentage of substrates assimilated
Shannon Diversity Index (H')	$H' = - \sum_{i=1}^n Pi * \ln Pi$	Diversity of substrate utilisation pattern where: $Pi = (ODi - R)/\sum(ODi - R)$ is the proportional colour development in the well i over the total colour development in all wells of a plate
Substrate Evenness (E)	$E = H' / \ln (SR)$	Equitability of C source assimilation efficiency over all utilised substrates

3. RESULTS

3.1. The SMC from the UPCC

Starting from over 6000 strains belonging to the UPCC, a reduced microbial core was selected to have a representative sample of the collection. To keep a high degree of biodiversity and to avoid potential clones, a total of 150 strains was taken considering different isolation matrices (mostly dairy products), research projects, and operators over the years. As a microbial community, the SMC's biodiversity can be described by several ecological indices. The total number of species considered in the study is 29 (S). The $\log(S)$ (1.46) is calculated to establish the maximum value the Shannon Diversity Index can reach ($H' = 1.32$). Furthermore, the Evenness Index of the SMC is $E = 0.90$. Finally, the dominance of some species over others is represented by $D = 0.05$ and $D' = 17.88$.

3.2. Genotypic characterisation

The 16S rRNA amplicon sequencing confirmed the identity of strains as recorded in the UPCC catalogue.

Comparing AFLP profiles of screened strains along with those of type strains, it was possible to dereplicate clones casually picked up from the UPCC. Hypothetical clones were found within the following species: *C. farciminis*, *E. faecalis*, *E. faecium*, *L. plantarum*, *L. delbrueckii*, *L. helveticus*, *L. curvatus*, *L. citreum*, *L. mesenteroides*, *L. brevis*, *L. fermentum*, *S. pasteurianus*, *S. porcorum*, *S. thermophilus*. The resulting AFLP profiles (Fig. 1) are labelled by species visualized through different colours of the dendrogram's branches. Profiles with more than 40% similarity were grouped into the same cluster and, in general, it is noted that strains belonging to the same species are close to each other. Clustering also reflects the division of strains according to the different type of metabolism that characterises them. Despite the overall evenness of clusters, some strains clustered with non-belonging species, especially when they come from different niches or when isolated temporally distant from each other. In the first case, an example is represented by *E. faecalis* UPCC 2206 and UPCC 529 isolated from a different matrix compared with other strains of the same species. The second instance, instead, is represented by *L. paracasei* UPCC 2333 isolated one year later than other strains isolated from the same matrix. Similarly, strain UPCC 256 that is the eldest *L. delbrueckii* of the SMC and clusters distantly from the other *L. delbrueckii* strains, and strains *S. macedonicus* UPCC 270, UPCC 2290, and *L. mesenteroides* UPCC 294 isolated in different moments compared with other strains of the belonging species. Clones of type strains used as reference for similarity

distance are highlighted in bold. They mostly fall within the same species' clusters, as shown in Figure 1. An exception is represented by the AFLP profiles of type strain clones of *L. curvatus* which are far from other profiles belonging to the same species. *L. casei* type strain clones are particularly close to *L. zae* profiles which confirms the high relatedness between these two species.

In Figure 1, concerning clusters that include more than 8-10 individuals (*L. paracasei*, *L. rhamnosus*, *L. fermentum*, *L. zae*, *L. curvatus*, *Leu. citreum*, *L. helveticus*, *E. faecalis*, *S. thermophilus*, *L. delbrueckii*), it is remarkable how some clusters show a higher level of biodiversity compared to others, especially the Lactobacillaceae's groups. This is particularly evident for *L. paracasei* and *L. rhamnosus* strains that were isolated in very different conditions. Here, the selection criteria of strains within the UPCC collection played a fundamental role as well as for those strains identified by more homogeneous AFLP profiles (e.g. *L. brevis*, *Leu. citreum*, *E. faecalis*, *S. thermophilus*), which are those isolated at the same time, from the same matrix and by the same operator.

To better represent this data, the pairwise AFLP profile percentage identity was plotted, grouped according to each species (Fig. 2). As can be seen from the graph, the isolates belonging to the genus *Lacticaseibacillus* show the greatest variability in terms of genotypic identity, followed by the genera *Lactobacillus* and *Enterococcus*. On the other hand, other genera show a greater genotypic identity among species members, such as the genera *Leuconostoc* or *Streptococcus* show more similar AFLP profiles, ranging between 50% and 100% similarity scores, or the species *L. plantarum*, *L. curvatus* and *L. brevis*, whose isolates AFLP profiles mostly show more than 75% identity. These similarities appear to be unrelated to the isolation source of the SMC strains.

Figure 1 Genotyping through AFLP profiles of the SMC isolates and type strains. The dendrogram represents the UPGMA clustering of AFLP profiles with Pearson correlation; the pseudogel of AFLP profiles of the isolates ranges between 50 and 500 bp. Dendrogram branches are coloured according to the strain species, as reported in the legend. The metabolism type of the isolates is reported alongside species description.

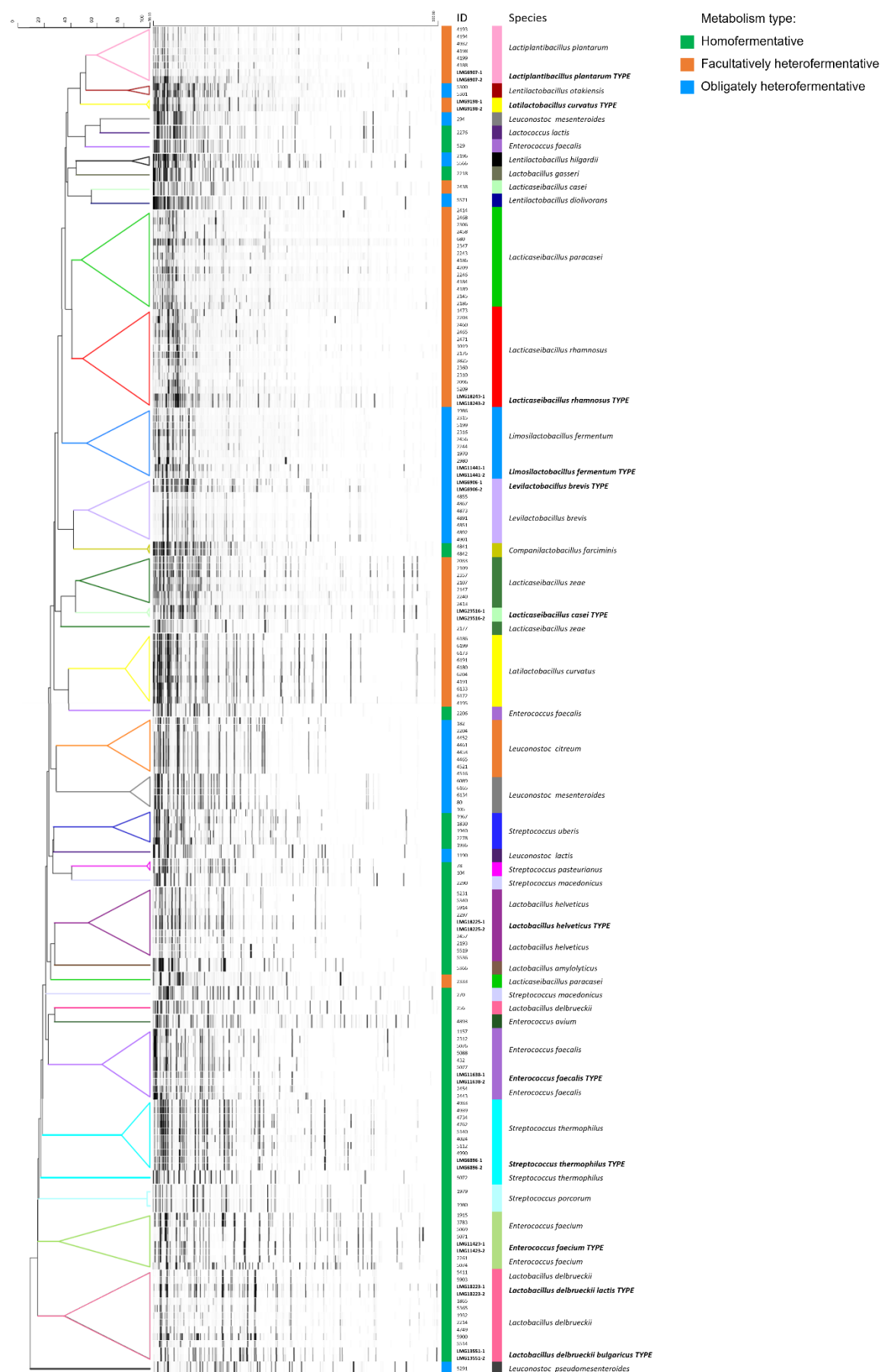
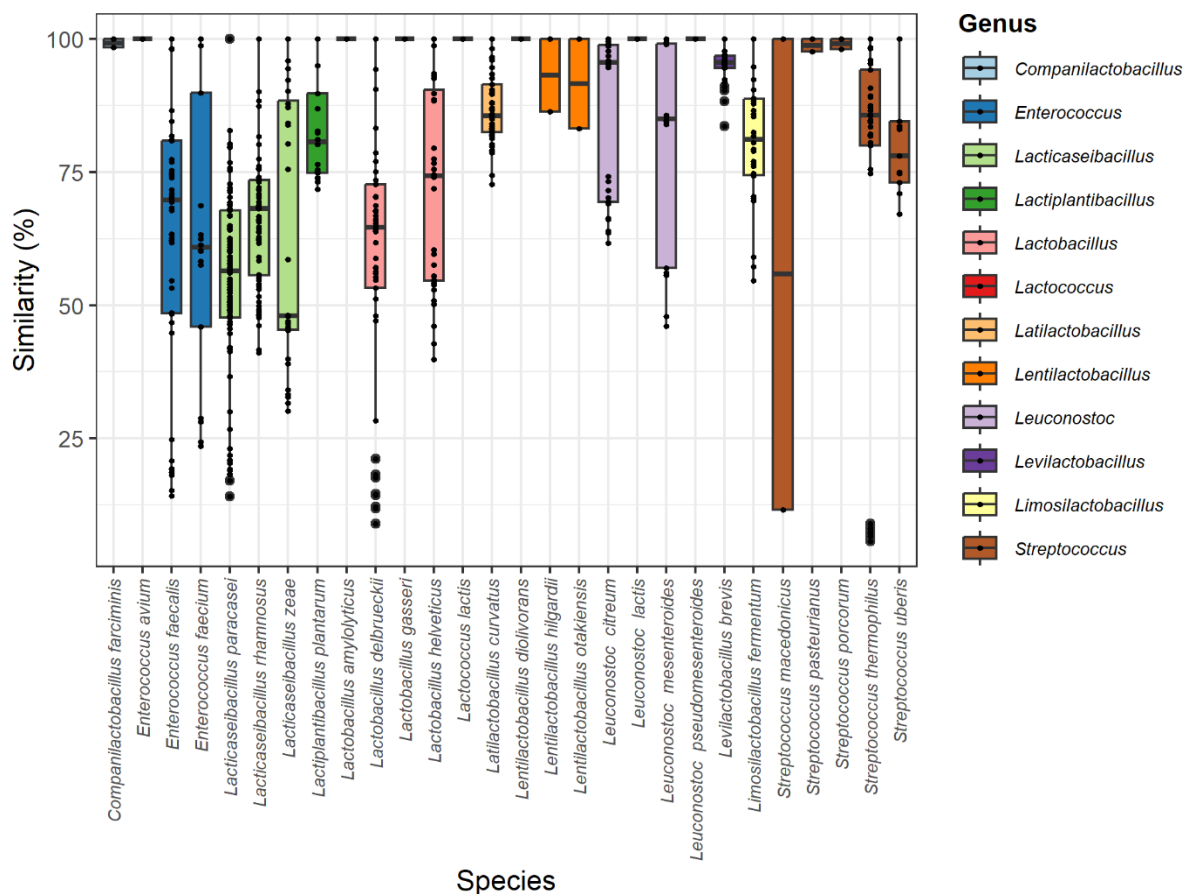


Figure 2 Range of pairwise intra-species percentage identity values by AFLP profiles. Boxplots are coloured according to the genera, as reported in the legend.

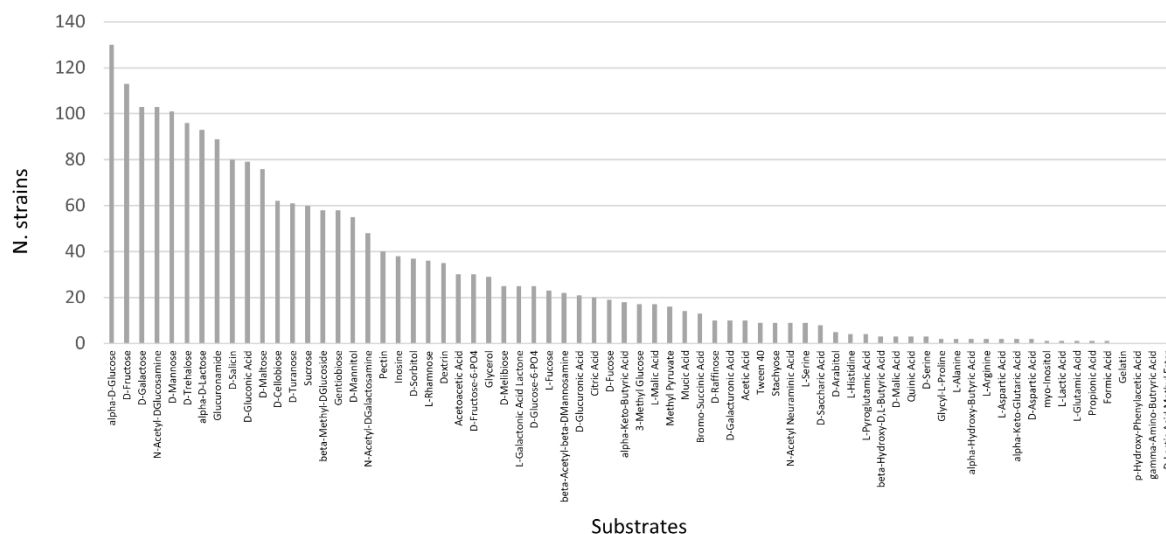


3.3. Phenotypic characterisation

3.3.1. Overall metabolic behaviour

The heatmap reported (Fig. 3) shows the carbon sources utilisation patterns of the SMC. The positive reaction in each well, indicated by the colour response, is related to the ability of bacteria to assimilate the compound in the well as the sole carbon source (Mekonnen *et al.*, 2019). There is a large variability in substrate assimilation levels among strains, also when they belong to the same species.

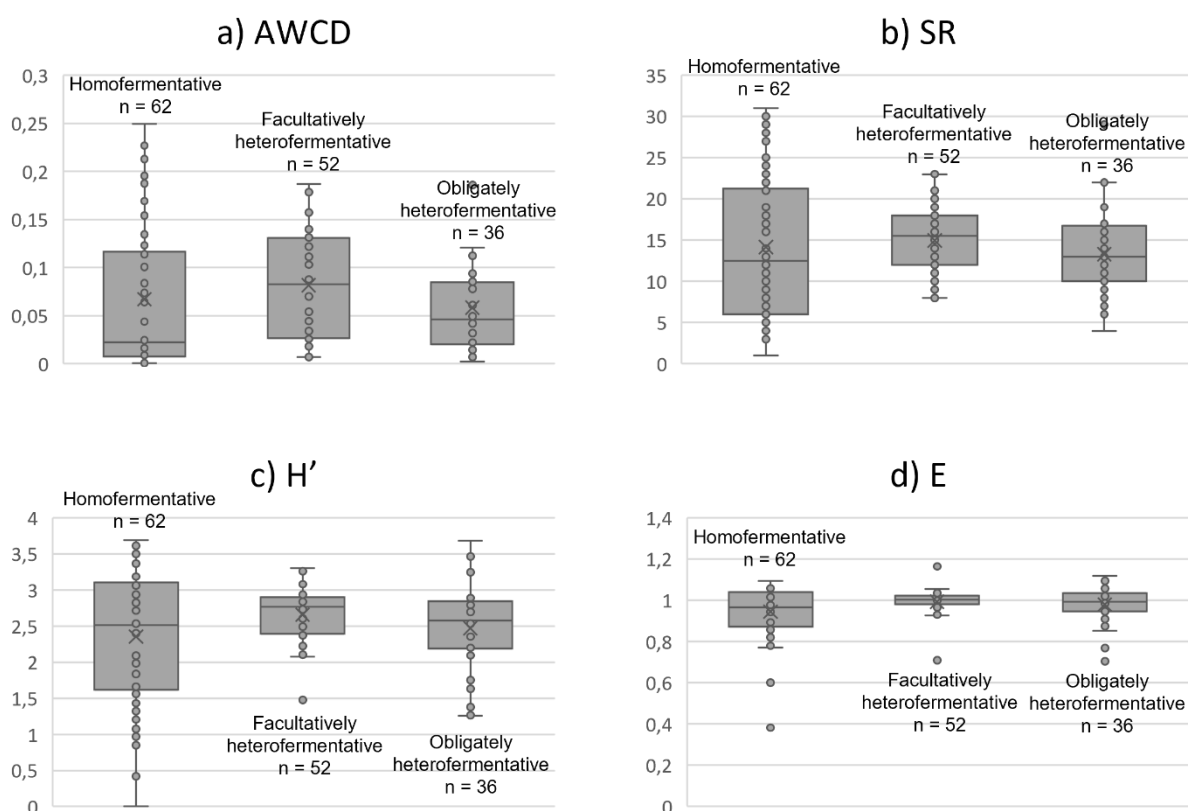
Figure 4 Frequency of substrates utilisation by the SMC. Each substrate was considered used when the normalized absorbance value was equal to or higher than 1.



Each strain's substrate utilisation profile can be described through the AWCD which measures the mean colour intensity, as the microbial metabolic capability is proportional to the degree of oxidation of the sugar in the GENIII Microplates (Dubey *et al.*, 2022). AWCD depends on the incubation time and should be estimated when the dye reduction has reached a plateau or when bacteria have theoretically assimilated the same amount of substrates (Acin-Albiac *et al.*, 2020). Within the SMC, HomF (62 strains) consists of a few strains more than Facultatively HetF (F-HetF) (52 strains), while Obligately HetF (O-HetF) are almost half of the previous groups (36 strains). As shown in Figure 5, in terms of intensity of substrates' utilisation, HomF strains are more phenotypically variable than F-HetF and O-HetF ones, as shown by the higher standard deviation (SD) in the measured AWCD (Fig. 5a). This result could be related to the higher phylogenetic distance between the HomF group species, compared to the HetF ones (Fig. 1). This heterogeneity inherent in the HomF group is reflected also in the substrate richness (SR) parameter that shows a great dispersion in the number of substrates that each strain can utilise, ranging between 3 and 31 used carbon sources (Fig. 5b), while strains belonging to F-HetF and O-HetF groups are less variable in terms of SR capability. Other information related to the substrate utilisation capability are given by the Shannon Index (H') and the Evenness (E) (Fig. 5). H' strongly depends on SR and depicts how each strain can differentially utilise the substrates supplied: the higher SR, the higher H' value and, accordingly, the higher H' , the higher the fluctuation of the intensity of substrates' consumption. As shown in Figure 5c, the variability of metabolic intensity into a microplate among strains is represented by the dispersion in the H' measure: it is elevated for HomF, followed by O-HetF, and rather reduced

for F-HetF. Finally, E describes the equitability of intensity among the C source metabolised: the greater E, the more similar the intensity of substrates consumed. In Fig. 5d we can see that F-HetF strains are the most homogeneous in their behaviour, meaning that the substrates utilised are consumed by all strains with similar intensity. Again, HomF strains are more variable in E to remark the different behaviour of phylogenetically divergent species. Summarizing, strains with a homofermentative metabolism are different from each other in terms of average intensity of consumption among all substrates (AWCD), number of substrates metabolised (SR), variability of intensity into the whole microplate (H'), and homogeneity of intensity among consumed substrates (E). On the opposite, heterofermentative bacteria are more similar to each other in their behaviour, with F-HetF performing better than O-HetF, thus remarking their relatedness.

Figure 5 Metabolic biodiversity of strains characterised by different metabolisms, represented by ecological diversity indices. a) AWCD represents the overall efficiency of consumption within the whole microplate of each strain; b) SR indicates the total number of consumed substrates by each strain; c) H' measures the variability in substrate utilisation efficiency within each microplate; d) E depicts the equitability of efficiency among the most frequently used substrate by each strain.



The same indices are reported for each species in Table 1, where it is possible to remark the variability into the three groups based on the metabolism type.

Within HomF, *E. faecalis* is the most efficient one since it strongly consumed ($AWCD = 0.126 \pm 0.038$) the major number of C sources ($SR = 28 \pm 3$), which implies a high entropy in the intensity ($H' = 3.584 \pm 0.094$) that is well distributed among used C sources ($E = 1.078 \pm 0.012$). On the other hand, *L. delbrueckii* and *L. helveticus* show the lowest AWCD (0.008 ± 0.006 and 0.006 ± 0.005 , respectively) associated with low SR (9 ± 3 and 6 ± 3) and slight fluctuation of the intensity of substrate consumption ($H' = 1.391 \pm 0.271$) due to the very low values of AWCD. Moreover, *S. thermophilus* is the least performing species ($AWCD = 0.017 \pm 0.007$ and $SR = 5 \pm 1$) since it only consumed a few substrates with different intensity ($E = 0.876 \pm 0.039$).

Considering the F-HetF group, *L. zae* metabolises the major number of C sources ($SR = 19 \pm 3$) with greater intensity than other species ($AWCD = 0.155 \pm 0.021$), very variable within the microplate ($H' = 2.981 \pm 0.192$) but constant among used substrates ($E = 1.014 \pm 0.018$). On the opposite in all the indices evaluated, there is *L. curvatus* that shows very different substrate assimilation capabilities.

Among O-HetF, *L. hilgardii* is the least interesting species as it has very scarce metabolic activity ($AWCD = 0.008 \pm 0.001$ and $SR = 6 \pm 0$). On the contrary, *Leuconostoc* species show the greatest AWCD (*Leu. citreum* = 0.086 ± 0.028 and *Leu. mesenteroides* = 0.098 ± 0.044), while *L. fermentum* shows the widest metabolic patterns ($SR = 19 \pm 6$) with a great intensity variance among all substrates ($H' = 3.000 \pm 0.444$) and uniformity among the consumed ones ($E = 1.028 \pm 0.064$).

Table 1 Metabolic functional diversity in different species, categorized by type of metabolism.

	AWCD	SR	% SR	H'	E
Homofermentative					
<i>Companilactobacillus farciminis</i>	0.017 ± 0.01	13 ± 1	18%	2.606 ± 0.153	1.032 ± 0.038
<i>Enterococcus faecalis</i>	0.126 ± 0.038	28 ± 3	39%	3.584 ± 0.094	1.078 ± 0.012
<i>Enterococcus faecium</i>	0.091 ± 0.020	18 ± 1	25%	2.871 ± 0.070	1.004 ± 0.027
<i>Lactobacillus delbrueckii</i>	0.008 ± 0.006	9 ± 3	13%	1.998 ± 0.326	0.922 ± 0.032
<i>Lactobacillus helveticus</i>	0.006 ± 0.005	6 ± 3	8%	1.237 ± 0.674	0.764 ± 0.160
<i>Streptococcus macedonicus</i>	0.054 ± 0.014	16 ± 2	22%	2.665 ± 0.183	0.973 ± 0.018
<i>Streptococcus pasteurianus</i>	0.220 ± 0.010	18 ± 0	25%	2.951 ± 0.005	1.021 ± 0.002
<i>Streptococcus porcorum</i>	0.135 ± 0.027	17 ± 1	24%	2.826 ± 0.154	0.998 ± 0.025
<i>Streptococcus thermophilus</i>	0.017 ± 0.007	5 ± 1	7%	1.391 ± 0.271	0.876 ± 0.039

<i>Streptococcus uberis</i>	0.192 ± 0.036	23 ± 1	32%	3.247 ± 0.069	1.036 ± 0.025
Facultatively heterofermentative					
<i>Lactocaseibacillus paracasei</i>	0.063 ± 0.039	13 ± 3	18%	2.515 ± 0.404	0.984 ± 0.089
<i>Lactocaseibacillus rhamnosus</i>	0.125 ± 0.016	17 ± 2	24%	2.856 ± 0.100	1.014 ± 0.013
<i>Lactocaseibacillus zeae</i>	0.155 ± 0.021	19 ± 3	27%	2.981 ± 0.192	1.014 ± 0.018
<i>Lactiplantibacillus plantarum</i>	0.041 ± 0.027	14 ± 4	20%	2.552 ± 0.329	0.981 ± 0.023
<i>Latilactobacillus curvatus</i>	0.018 ± 0.016	12 ± 4	17%	2.414 ± 0.460	0.967 ± 0.099
Obligately heterofermentative					
<i>Lentilactobacillus hilgardii</i>	0.008 ± 0.001	6 ± 0	8%	1.320 ± 0.800	0.737 ± 0.045
<i>Lentilactobacillus otakiensis</i>	0.035 ± 0.003	7 ± 1	9%	1.644 ± 0.021	0.881 ± 0.040
<i>Leuconostoc citreum</i>	0.086 ± 0.028	15 ± 2	21%	2.718 ± 0.139	1.011 ± 0.028
<i>Leuconostoc mesenteroides</i>	0.098 ± 0.044	13 ± 5	18%	2.605 ± 0.360	1.031 ± 0.047
<i>Levilactobacillus brevis</i>	0.020 ± 0.011	9 ± 3	12%	1.995 ± 0.407	0.951 ± 0.046
<i>Limosilactobacillus fermentum</i>	0.061 ± 0.046	19 ± 6	27%	3.000 ± 0.444	1.028 ± 0.064

3.3.2. Substrate utilisation patterns

Normalized absorbance values higher than 1 were used to determine the positive wells. Substrates metabolised by more than 30 strains were selected as favourite C sources. Among them, it is possible to distinguish monosaccharides (α -D-Glucose, D-Fructose, D-Galactose, D-Mannose, L-Rhamnose), disaccharides (D-Trehalose, α -D-Lactose, D-Maltose, D-Turanose, Gentiobiose, D-Cellobiose, and Sucrose), sugar derivatives (N-Acetyl-D-Glucosamine, D-Salicin, D-Gluconic acid, β -Methyl-D-Glucoside, D-Mannitol, N-Acetyl-D-Galactosamine, D-Sorbitol, Glucuronamide, Pectin, Dextrin, Acetoacetic Acid, D-Fructose-6-PO₄), and one purine nucleoside (Inosine). Glucose is the most frequently used C source, and Figure 6 shows how most of the other substrates can be actively consumed by LAB to levels comparable with the latter. However, some compounds are metabolised only by specific strains, and to a much higher degree than glucose.

Figure 6 Variability of the metabolic activity on the most utilised substrates displayed by variance among absorbance readings at 590 nm of the reduced Tetrazolium Violet dye into the microplate wells. Significant differences observed between α -D-Glucose consumption with respect to other C sources are represented by P values calculated through Student's T-test: ** P < 0.01, * P < 0.05.

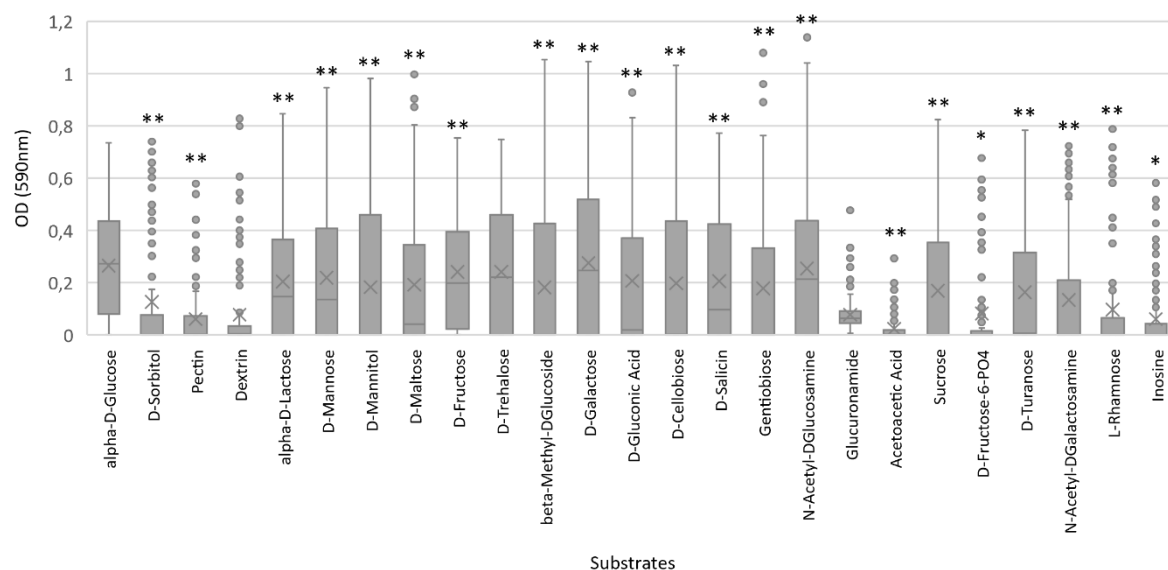


Figure 7 Heatmaps of the most frequently used C sources by the SMC. Strains were divided based on the microbial metabolism type and labelled by genera. The dendrogram based on the Euclidean distance measure represents the clustering of metabolic patterns and the colour intensity indicates the efficiency of strains to metabolise the substrates supplied.



Among HomF (Fig. 7a), it is possible to distinguish 5 main groups of utilisation patterns. The first two clusters are mainly composed of *Streptococcus* spp. and *E. faecium* strains able to

better use all the monosaccharides except L-Rhamnose, all the disaccharides except D-Turanose, and sugar derivatives such as D-Salicin, β -Methyl-D-Glucoside, N-Acetyl-D-Glucosamine, Pectin, Dextrin, as well as Glucuronamide for a few strains. Some bacteria within the second group have weak activity on D-Mannitol, D-Sorbitol, N-Acetyl-D-Galactosamine, and Inosine. The third one consists of three *L. delbrueckii* that metabolise all the monosaccharides except L-Rhamnose, disaccharides such as Sucrose, α -D-Lactose, D-Maltose and D-Trehalose, and some sugar derivatives (Pectin, Dextrin, Glucuronamide, Acetoacetic Acid, and D-Mannitol). This cluster remarks that although belonging to the same species, strains can have a different metabolic profile confirming the indices reported in Table 1. The fourth group consists of only *E. faecalis* that effectively metabolise Sucrose and derivatives like Pectin, Dextrin, Glucuronamide, and Acetoacetic Acid. They can also weakly utilise D-Fructose-6-PO₄, N-Acetyl-D-Galactosamine, Inosine, and contrarily to all other strains also L-Rhamnose and D-Turanose. The fourth cluster clearly displays what is estimated by the indices: *E. faecalis* uses the greatest number of substrates within the HomF strains ($SR = 28 \pm 3$), with moderated and variable efficiency ($AWCD = 0.126 \pm 0.038$ and $H' = 3.584 \pm 0.094$). The fifth group contains strains of *Lactobacillus* spp. and *S. thermophilus* that strongly assimilate α -D-Glucose as monosaccharide and Glucuronamide and Acetoacetic Acid as sugar derivatives. A few strains metabolise α -D-Lactose as disaccharides and polysaccharides like Pectin and Dextrin. Data concerning *S. thermophilus* in this heatmap confirm the low SR reported in Table 1.

F-HetF can be divided into four clusters. In the first group, it emerges a strong activity on Glucuronamide. The second one is dominated by *Lacticaseibacillus* spp. whose strains well assimilate most of all selected substrates, except Sucrose as disaccharides, and other compounds such as Acetoacetic acid, Pectin, Inosine, and D-Fructose-6-PO₄. The third cluster is characterised by the evident consumption of monosaccharides such as D-Fructose, D-Mannose, α -D-Glucose and D-Galactose, two disaccharides (D-Trehalose and α -D-Lactose), N-Acetyl-D-Glucosamine and Glucuronamide. The fourth cluster, mainly *L. paracasei* strains, beyond the same substrates utilised by the previous group, highlights a strong metabolic activity on D-Salicin and D-Turanose. For some strains, it is remarkable also the use of other sugar derivatives (N-Acetyl-D-Galactosamine, β -Methyl-D-Glucoside, D-Gluconic Acid, and D-Mannitol).

Within O-HetF there are three clusters. The first one is composed of *Leuconostoc* spp. strains and displays their ability to metabolise all the monosaccharides except L-Rhamnose, disaccharides like D-Turanose, D-Trehalose, α -D-Lactose, Sucrose and D-Maltose, and sugar derivatives such as D-Fructose-6-PO₄, Pectin, Dextrin, N-Acetyl-D-Glucosamine, D-Gluconic Acid, and Glucuronamide. The second group is the most heterogeneous containing all four

genera. Here it is evident the metabolic activity on Glucuronamide and D-Gluconic Acid. The latter is not used by *L. fermentum*, which shows instead a consumption of L-Rhamnose, Acetoacetic Acid, Inosine, and Dextrin. The last cluster shows metabolic profiles similar to the first one, but different in intensity. *Leuconostoc* again seems to consume the major number of substrates albeit weakly: this confirms the variable efficiency on metabolising substrates indicated by high values of H'.

3.4. Natural sources of the main metabolised C sources

The scientific literature states that many substrates characterised in this research for their microbial utilisation can be found in many foods and are often abundant in waste and by-products deriving from manufacturing processes (Tab. 2). Agrifood companies may be involved in a circular economy model where outputs, different from the main product of the process, can be recovered and reused as raw material in a new production process.

Table 2 Food industries involved in the circular economy model (Ranganathan *et al.*, 2020): A = Agricultural waste and lignocellulosic biomass, B = Seafood processing industry, C = Manufactures of coffee and tea, D = Dairy processing industry, E = Waste from brewery and distillery industries, F = Fruit and vegetables processing industries, G = Grain processing industry, H = Meat and poultry processing industries, I = insect feedstock, n.d.= not determined.

Substrate	Source	Food industries	References
N-Acetyl-D-Glucosamine	Component of chitin in yeast, fungi, seafood, algae, insects, component of cellulose and hemicellulose	A, B, I	(Cheba, 2011; Mieszkina <i>et al.</i> , 2021; Riemann & Azam, 2002)
D-Salicin	Bark of and leaves of willows, poplars, and various other plants	A	(Adamiak <i>et al.</i> , 2021)
alpha-D-Lactose	Milk, whey, hydrolysate, cheese residue, and other milk derivatives	D	(Akbas & Stark, 2016; Qi & Tester, 2020)
D-Maltose	Malted grains, broken grains, fine dust, husk, malt dust, grain separator waste, straw, bran, spent grains, coffee processing pulp and spent coffee grounds	C, E, G	(Qi & Tester, 2020)
D-Trehalose	Invertebrates, fungi, bacteria	n.d.	(Gallego-Lobillo <i>et al.</i> , 2021; Garcia

			& Gardner, 2021)
D-Cellobiose	Component of cellulose in vegetables, fruits, corn syrups, potato peel, mash, coffee bean husks, rice husk, vegetable and plant processing waste, peel, pomace, molasses, beet pulp, bagasse, lignocellulosic biomass	A, C, E, F	(Sasaki <i>et al.</i> , 2008; Ucar <i>et al.</i> , 2020)
Gentiobiose	Component of cellulose in vegetables, fruits, corn syrups, potato peel, mash, coffee bean husks, rice husk, vegetable and plant processing waste, peel, pomace, molasses, beet pulp, bagasse, lignocellulosic biomass	A, C, E, F	(Ucar <i>et al.</i> , 2020)
Sucrose	Plants, fruits and vegetables, fruit processing waste, pineapple peel, raw sugar beet juices, sugar beet processing by-products, molasses, coffee processing pulp	C, E, F	(Qi & Tester, 2020)
D-Turanose	Honey	n.d.	(Kim <i>et al.</i> , 2019)
alpha-D-Glucose	Starch, whey, components of cellulose and hemicellulose in fruit processing waste, pineapple peel, raw sugar beet juices, sugar beet processing by-products, molasses, corn processing wastes, coffee processing pulp	A, C, D, E, F, G	(Granström <i>et al.</i> , 2004)
D-Mannose	Component of mannan, hemicellulose, and cellulose in spent coffee grounds, potato peel, mash	A, C, E, F	(H. Wu <i>et al.</i> , 2019)
D-Fructose	Mature fruits, plant storage organs, honey, fruit and vegetables processing waste	F	(Georgelis <i>et al.</i> , 2018; Mieszkina <i>et al.</i> , 2021)
D-Galactose	Milk, whey, hydrolysate, cheese residue and other milk derivatives	D	(Granström <i>et al.</i> , 2004)
L-Rhamnose	Component of cellulose and hemicellulose in potato peel, mash, vegetable washing water, peel, pomace, molasses, and beet pulp. Component of Pectin in plants and of lipopolysaccharides in Gram-negative bacteria	A, F	(Mieszkina <i>et al.</i> , 2021; Reinhardt <i>et al.</i> , 2019)
N-Acetyl-D-Galactosamine	Milk and derivatives	D	(Feeney <i>et al.</i> , 2019)
beta-Methyl-D-Glucoside	Component of cellulose and hemicellulose	A, F, G	(Mieszkina <i>et al.</i> , 2021)
D-Sorbitol	Component of cellulose and hemicellulose in fresh and dried fruits, vegetables	F	(Dash <i>et al.</i> , 2019; Mieszkina <i>et al.</i> , 2021)

D-Mannitol	Component of cellulose and hemicellulose	A, F, G	(Mieszkin <i>et al.</i> , 2021)
Inosine	Meat products	H	(Wang <i>et al.</i> , 2020)
D-Gluconic Acid	Plants, fruits, rice, dairy products, vinegar, meat, wine, honey	B, D, F, G, H	(F. Kornecki <i>et al.</i> , 2020)
Pectin	Citrus (lemon, lime, orange) peels, apple pomace, sugar beet pulp, potato peel, mash	F	(Picot-Allain <i>et al.</i> , 2020)
Dextrin	Component of starch in whole potatoes, potato processing waste, sweet potato, potato peel, mash, broken grains, fine dust, husk, malt dust, grain separator waste, straw, bran	G	(Crini, 2020; Włodarczyk <i>et al.</i> , 2022)
Glucuronamide	Corn, lignocellulose	A, F, G	(Jonathan <i>et al.</i> , 2017)

4. DISCUSSIONS

This work aimed to set the basis for exploiting the invaluable microbial diversity that is preserved in microbial collections. With the outlook of researching industrial chemicals produced from food waste and by-products, the purpose of the study was phenotyping a selected microbial core of LAB strains to investigate their substrate utilisation capabilities and design fermentation processes according to their metabolic aptitude.

The ecological diversity indices used to describe the SMC estimate a good degree of biodiversity. Shannon Index (H') determines the amount of information in the population and it is related to the number of species (S) in the SMC: the higher S , the higher H' . Considering the logarithm of S ($\text{Log}(S) = 1.46$), a value of H' for the SMC corresponding to 1.32 indicates a good level of diversity proving that a proper selection approach to build the studied core starting from the UPCC was adopted. Therefore, we can assume that achieved knowledge about this reduced microbial community can be generalized to a wider one. The Evenness (E) heavily depends on S and gives the homogeneity of the community. A value of 0.90 means a comparable distribution of all species within the core, considering that $E = 1$ when they are all equally distributed. The Dominance value, which should range between 0 (maximum diversity) and 1 (absence of diversity), still confirms ($D = 0.05$) the diversity in the SMC, although $D' = 17.88$ underlines a strong dominance of some species that are more representative than others in terms of number of individuals (Daly *et al.*, 2018).

The UPCC, nowadays associated with the Joint Research Unit MIRRI-IT (<https://www.mirri-it.it/>), started to be implemented more than 20 years ago, therefore there are isolates identified a long time ago with several methods among which some could be outdated. By screening so many strains, this work started a worthwhile standardization of the procedure according to the most current technologies, both for genotyping and phenotyping. The AFLP clustering generally reflects the belonging to the species clustering. However, the fingerprints of individual strains show a certain level of intra-specific biodiversity and even more within the same genus. Strains included in clusters other than that of their species differ from each other in source and moment of isolation. This demonstrates that even a very limited community of the UPCC can have a high degree of genotypic variability. Despite AFLP can be considered a limited resolution technique if compared to genome-wide association studies, it is noteworthy that this genotyping technique captures the intraspecies variability inherent in the SMC. Indeed, the pairwise identity of AFLP profiles shown in Fig. 2 is in agreement with more complex indices, such as the pairwise amino acid identity of conserved genes (cAAI), that are widely

used in taxonomic studies. For instance, using cAAI, Zheng et al. (2020) have shown that the genera *Lacticaseibacillus* and *Lactobacillus* are among the most diverse groups, while the genera *Lactiplantibacillus*, *Latilactobacillus* and *Lentilactobacillus* have higher pairwise identity. The relevant note is that these findings are confirmed in a smaller subset of bacterial isolates, as is the case of UPCC collection, thus reaffirming the relevance of the approach used in this study.

Concerning phenotyping, the method used exploits the production of free electrons by bacteria during the utilisation of carbon sources, which reduces the dye whose colour change and intensity indirectly represent the substrate utilisation degree of microorganisms (Ge *et al.*, 2018). It follows that cells metabolically react by using the substance, but not necessarily by growing. Accordingly, BIOLOG® Microplates test metabolic activity rather than growth and are a more sensitive method to evaluate phenotypes (Mekonnen *et al.*, 2019).

Looking at Figure 3 it is possible to note that the distribution of metabolic profiles based on both species and metabolism type, is more scattered than the genotypic one. This is clear for the *Lacticaseibacillus* genus: Figure 1 shows that *L. paracasei* and *L. rhamnosus* clusters are close to each other while *L. zeae* is just below. Considering the genotypic fingerprints, instead of metabolic profiles, all strains group according to the species identification. The only exception is *L. paracasei* UPCC 2333, isolated in a different year compared to other *Lacticaseibacillus* strains. On the other side, the phenotyping (Fig. 3) remarks many differences, especially within the *L. paracasei* species where strains are scattered in 6 different positions along the dendrogram. The same comparison can be done for *Lactobacillus* and *Streptococcus* species. Similarly to the AFLP results, this biodiversity is even greater at the genus level. These differences have been already observed at the genome level for some isolates belonging to the UPCC collection (Levante *et al.*, 2021) and phenomena such as lifestyle adaptations or horizontal gene transfer (Duar *et al.*, 2017; Zheng *et al.*, 2020) can explain the genotypic variability observed in the SMC.

Taking into account the type of metabolism characterising the microorganisms, the composition of the SMC reflects the composition of the UPCC, largely resulting from isolation on fermented food where homofermentative bacteria are predominant. The AFLP profiling (Fig. 1) displays all HetF species in the upper part of the dendrogram and HomF in the lower one, with a general evenness. Nevertheless, metabolic profiles and diversity indices estimate an important phenotypic diversity for HomF strains, composed of genera belonging to three different families, as well as for HetF which are genetically closely related.

The similarity of metabolic profiles strongly depends on species or strain-specific abilities which is confirmed by the comparison between AFLP and GENIII clusters. This sustains the hypothesis of predicting substrate assimilation from the genetic determination of species (Mekonnen *et al.*, 2019). However, it is important to note that even closely related strains, sometimes clones, can have different C source assimilation patterns (Koch *et al.*, 2020). It means that unexpected correlations or metabolic functions can occur, likely due to evolutionary aspects linked to the isolation from complex ecosystems, e.g. processes of gene loss and gain. Species are characterised by a core of metabolic possibilities but the evolution of biotypes driven by the ecosystem promotes the occurrence of variants with wider or smaller potentialities in terms of C sources utilisation (Arnold *et al.*, 2022; O'Connell *et al.*, 2022). Our results seem to support this idea, therefore the biodiversity observed in this study would be due to selective pressures and not by chance. For instance, *Enterococcus* spp. has a greater number of variants due to the tendency of its species to acquire exogenous DNA (Tao *et al.*, 2023), which may be evolutionarily advantageous and allow microorganisms to survive and colonize a novel environment. The wideness of biodiversity and the number of variants in a species would be a measure of the species' ability to adapt and evolve.

In this work, we want to emphasize the presence of a significantly higher level of biodiversity compared to what is traditionally associated with the matrix of isolation. UPCC's isolates mostly derive from dairy matrices but it does not limit their ability to assimilate compounds of other origins, such as vegetable matrices (Ricci *et al.*, 2018; Sáez *et al.*, 2018). As described by Hayek *et al.* (2019) many agricultural waste and food by-products can support the growth of LAB similarly to the most common culture media, also in terms of biomass and lactic acid production, with the advantage of reducing costs and increasing the sustainability. Ecological microbial indices used to evaluate the metabolic diversity, help us in the selection process. The first indication goes at the AWCD which indicates the general potential metabolic activity of strains, and at the H' that estimates their physiological diversity. The selection process will then prioritize those strains able to metabolise as many targeted substrates, e.g those present in the matrix to ferment, as possible (high SR) also evaluating the efficiency of consumption (E): strains with high E if we desire to recover as much carbon as possible from the matrix, or alternatively, strains with low E if we want to favour one source over another.

Among the most frequently metabolised C sources, there are α -D-Glucose, D-Fructose and D-Galactose as monosaccharides, D-Trehalose, α -D-Lactose and D-Maltose as disaccharides (Fig. 4). These and other sugars are generally found in many wastes and by-products of different origin, which represent important carbon sources to produce key biopolymers for the bioplastic

industry (such as polylactic acid and polyhydroxyalkanoates) through bacterial fermentation. In line with the increasing market and the necessary increase of sustainability, bioplastics will be a dominant player in the future economy (Ranganathan *et al.*, 2020). Strains that use the major number of substrates mainly belong to *Enterococcus* spp., *Lactocaseibacillus* spp. and *Streptococcus* spp. and they could be scaled to a secondary screening to deeper investigate their abilities to ferment those sugars directly from the natural sources. Among sugar derivatives studied, many strains utilise N-Acetyl-D-Glucosamine which means that they can metabolise lysed cells (Sgarbi *et al.*, 2014). Moreover, there are compounds assimilated better than α -D-Glucose only by certain microorganisms. For instance, *S. pasteurianus* UPCC 78 metabolises Dextrin, D-Maltose, D-Fructose, D-Trehalose, D-Galactose, D-Cellobiose, Gentiobiose, N-Acetyl-D-Glucoside and Sucrose heavier than α -D-Glucose, making it suitable to ferment matrices rich in starch, cellulose or chitin, as well as milk, grain and fruit industries by-products. Another example is *S. pasteurianus* UPCC 104 which shows a stronger activity on α -D-Lactose, D-Mannose, D-Mannitol, β -Methyl-D-Glucoside, D-Galactose and N-Acetyl-D-Glucoside than on α -D-Glucose, therefore particularly adapt to metabolise milk by-products. Gluconic Acid and D-Turanose are preferred by *L. zeae* UPCC 2033, Pectin by *E. faecalis* UPCC 2443, and L-Rhamnose by *L. zeae* UPCC 2413 compared to α -D-Glucose.

The distinction between homofermentative and heterofermentative LAB is not only due to the different metabolic pathways activated for hexose degradation but also to the presence or absence of other metabolic paths for sugars and amino acids assimilation (Gänzle, 2015). Species of both physiological groups are widespread and their metabolic adaptation is especially based on the different phylogenetic groups (Zheng *et al.*, 2015). The coexistence of both in many habitats supports the theory that homofermentative and heterofermentative behaviours are not competitive but complementary. Furthermore, the variability of metabolic profiles and growth temperature can allow us to better understand niche adaptations (Duar *et al.*, 2017). The flexibility in substrate consumption allows a wider spectrum of applications, while metabolically limited microorganisms can be involved in adaptive evolution studies by betting on their strengths and heading towards targeted selection processes for industrial-specific uses. More in detail, only F-HetF bacteria can utilise D-Sorbitol (*L. rhamnosus*, *L. curvatus*, *L. plantarum*) and N-Acetyl-D-Galactosamine (*Lactocaseibacillus* spp.). D-Mannitol is mainly used by *Lactocaseibacillus* spp., but weakly also by some *Streptococcus* spp. (HomF): in this case, in terms of process optimization, strains metabolically more active will be chosen. On the other hand, F-HetF strains never use either Acetoacetic Acid or Sucrose. Additionally, Inosine is only metabolised by *L. fermentum* and *L. hilgardii* (O-HetF), and very weakly by

some *Enterococcus* spp. (HomF), while D-Fructose-6-PO₄ only by *Leuconostoc* spp. Furthermore, O-HetF strains never consume β -Methyl-D-Glucoside nor D-cellobiose, and very slightly D-salicin and Gentiobiose.

5. CONCLUSION

This work set up an approach to develop microbial platforms with the aim of supplying integrated information about microorganisms for industrial applications. It focuses on a microbial core properly selected to be taken as a model and uses indices to objectify the analysis of geno- and phenotypic biodiversity and simplify the selection processes. In particular, this paper highlights the importance of phenotyping as a powerful means for characterising microbial communities, enabling better comprehension and prediction of their behaviour. This knowledge can be leveraged to identify key players capable of driving specific fermentation, and consequently facilitate the design and optimization of targeted processes.

The understanding of microbial aptitudes also helps to deeply investigate the microorganisms' adaptation, still driving selection and optimization processes towards specific niches and/or specific biochemical production. Anticipating and manipulating the metabolic traits allows to maximize the functional potential of microorganisms and develop efficient and sustainable waste recovery strategies, which leads to the enhancement of waste conversion, the encouragement of circular economy practices, and the reduction of environmental impact.

FUNDING INFORMATION

This work was granted by the European Commission—NextGenerationEU, Project “Strengthening the MIRRI Italian Research Infrastructure for Sustainable Bioscience and Bioeconomy”, code n. IR00000005. V. Bonnici is partially supported by the University of Parma (Italy) project n. MUR_DM737_B_MAFI_BONNICI, and by the CINI InfoLife laboratory.

DATA AVAILABILITY

The datasets generated during and/or analysed during the current study are available from the corresponding author upon reasonable request.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Alessandra Masci for technical support in the maintenance and operation of the frozen cultures from the UPCC collection.

REFERENCES

- Acin-Albiac, M., Filannino, P., Gobetti, M., & Di Cagno, R. (2020). Microbial high throughput phenomics: The potential of an irreplaceable omics. *Computational and Structural Biotechnology Journal*, 18, 2290–2299. <https://doi.org/10.1016/j.csbj.2020.08.010>
- Adamiak, K., Lewandowska, K., & Sionkowska, A. (2021). The Influence of Salicin on Rheological and Film-Forming Properties of Collagen. *Molecules*, 26(6), 1661. <https://doi.org/10.3390/molecules26061661>
- Akbas, M. Y., & Stark, B. C. (2016). Recent trends in bioethanol production from food processing byproducts. *Journal of Industrial Microbiology and Biotechnology*, 43(11), 1593–1609. <https://doi.org/10.1007/s10295-016-1821-z>
- Arnold, B. J., Huang, I.-T., & Hanage, W. P. (2022). Horizontal gene transfer and adaptive evolution in bacteria. *Nature Reviews Microbiology*, 20(4), Article 4. <https://doi.org/10.1038/s41579-021-00650-4>
- Bancalari, E., D’Incecco, P., Savo Sardaro, M. L., Neviani, E., Pellegrino, L., & Gatti, M. (2019). Impedance microbiology to speed up the screening of lactic acid bacteria exopolysaccharide production. *International Journal of Food Microbiology*, 306, 108268. <https://doi.org/10.1016/j.ijfoodmicro.2019.108268>
- Bancalari, E., Savo Sardaro, M. L., Levante, A., Marseglia, A., Caligiani, A., Lazzi, C., Neviani, E., & Gatti, M. (2017). An integrated strategy to discover *Lactobacillus casei* group strains for their potential use as aromatic starters. *Food Research International*, 100, 682–690. <https://doi.org/10.1016/j.foodres.2017.07.066>
- Bertani, G., Savo Sardaro, M. L., Neviani, E., & Lazzi, C. (2019). AFLP protocol comparison for microbial diversity fingerprinting. *Journal of Applied Genetics*, 60(2), 217–223. <https://doi.org/10.1007/s13353-019-00492-0>
- Boone, D. R., Castenholz, R. W., & Garrity, G. M. (Eds.). (2001). *Bergey’s manual of systematic bacteriology* (2nd ed). Springer.
- Bosco, F., Cirrincione, S., Carletto, R., Marmo, L., Chiesa, F., Mazzoli, R., & Pessione, E. (2021). PHA Production from Cheese Whey and “Scotta”: Comparison between a Consortium and a Pure Culture of *Leuconostoc mesenteroides*. *Microorganisms*, 9(12), 2426. <https://doi.org/10.3390/microorganisms9122426>
- Buron-Moles, G., Chailyan, A., Dolejs, I., Forster, J., & Mikš, M. H. (2019). Uncovering carbohydrate metabolism through a genotype-phenotype association study of 56 lactic acid bacteria genomes. *Applied Microbiology and Biotechnology*, 103(7), 3135–3152. <https://doi.org/10.1007/s00253-019-09701-6>
- Cheba, B. A. (2011). Chitin and Chitosan: Marine Biopolymers with Unique Properties and Versatile Applications. *Global Journal of Biotechnology & Biochemistry*, 6, 149–153.
- Crini, G. (2020). Twenty years of dextrin research: A tribute to Professor Hans Pringsheim (1876–1940). *Journal of Inclusion Phenomena and Macrocyclic Chemistry*, 98(1), 11–27. <https://doi.org/10.1007/s10847-020-01013-x>
- Daly, A. J., Baetens, J. M., & De Baets, B. (2018). Ecological Diversity: Measuring the Unmeasurable. *Mathematics*, 6(7), Article 7. <https://doi.org/10.3390/math6070119>
- Dash, R. P., Srinivas, N. R., & Babu, R. J. (2019). Use of sorbitol as pharmaceutical excipient in the present day formulations – issues and challenges for drug absorption and

- bioavailability. *Drug Development and Industrial Pharmacy*, 45(9), 1421–1429. <https://doi.org/10.1080/03639045.2019.1640722>
- De Vero, L., Boniotti, M. B., Budroni, M., Buzzini, P., Cassanelli, S., Comunian, R., Gullo, M., Logrieco, A. F., Mannazzu, I., Musumeci, R., Perugini, I., Perrone, G., Pulvirenti, A., Romano, P., Turchetti, B., & Varese, G. C. (2019). Preservation, Characterization and Exploitation of Microbial Biodiversity: The Perspective of the Italian Network of Culture Collections. *Microorganisms*, 7(12), 685. <https://doi.org/10.3390/microorganisms7120685>
- Di Cagno, R., Coda, R., De Angelis, M., & Gobbetti, M. (2013). Exploitation of vegetables and fruits through lactic acid fermentation. *Food Microbiology*, 33(1), 1–10. <https://doi.org/10.1016/j.fm.2012.09.003>
- Duar, R. M., Lin, X. B., Zheng, J., Martino, M. E., Grenier, T., Pérez-Muñoz, M. E., Leulier, F., Gänzle, M., & Walter, J. (2017). Lifestyles in transition: Evolution and natural history of the genus *Lactobacillus*. *FEMS Microbiology Reviews*, 41(Supp_1), S27–S48. <https://doi.org/10.1093/femsre/fux030>
- Dubey, A., Malla, M. A., & Kumar, A. (2022). Taxonomical and functional bacterial community profiling in disease-resistant and disease-susceptible soybean cultivars. *Brazilian Journal of Microbiology*. <https://doi.org/10.1007/s42770-022-00746-w>
- Feeney, S., Gerlach, J. Q., Slattery, H., Kilcoyne, M., Hickey, R. M., & Joshi, L. (2019). Lectin microarray profiling and monosaccharide analysis of bovine milk immunoglobulin G oligosaccharides during the first 10 days of lactation. *Food Science & Nutrition*, 7(5), 1564–1572. <https://doi.org/10.1002/fsn3.950>
- Feigl, V., Ujaczki, É., Vaszita, E., & Molnár, M. (2017). Influence of red mud on soil microbial communities: Application and comprehensive evaluation of the Biolog EcoPlate approach as a tool in soil microbiological studies. *Science of The Total Environment*, 595, 903–911. <https://doi.org/10.1016/j.scitotenv.2017.03.266>
- F. Kornecki, J., Carballares, D., W. Tardioli, P., C. Rodrigues, R., Berenguer-Murcia, Á., R. Alcántara, A., & Fernandez-Lafuente, R. (2020). Enzyme production of D-gluconic acid and glucose oxidase: Successful tales of cascade reactions. *Catalysis Science & Technology*, 10(17), 5740–5771. <https://doi.org/10.1039/D0CY00819B>
- Flippen, J. L., & Gilardi, R. D. (n.d.). *Crystal Structure of α -Glucuronamide*. 3.
- Gallego-Lobillo, P., Doyagüez, E. G., Jimeno, M. L., Villamiel, M., & Hernandez-Hernandez, O. (2021). Enzymatic Synthesis and Structural Characterization of Novel Trehalose-Based Oligosaccharides. *Journal of Agricultural and Food Chemistry*, 69(42), 12541–12553. <https://doi.org/10.1021/acs.jafc.1c03768>
- Gänzle, M. G. (2015). Lactic metabolism revisited: Metabolism of lactic acid bacteria in food fermentations and food spoilage. *Current Opinion in Food Science*, 2, 106–117. <https://doi.org/10.1016/j.cofs.2015.03.001>
- Garcia, C. A., & Gardner, J. G. (2021). Bacterial α -diglucoside metabolism: Perspectives and potential for biotechnology and biomedicine. *Applied Microbiology and Biotechnology*, 105(10), 4033–4052. <https://doi.org/10.1007/s00253-021-11322-x>
- Ge, Z., Du, H., Gao, Y., & Qiu, W. (2018). Analysis on Metabolic Functions of Stored Rice Microbial Communities by BIOLOG ECO Microplates. *Frontiers in Microbiology*, 9, 1375. <https://doi.org/10.3389/fmicb.2018.01375>

- George, F., Daniel, C., Thomas, M., Singer, E., Guilbaud, A., Tessier, F. J., Revol-Junelles, A.-M., Borges, F., & Foligné, B. (2018). Occurrence and Dynamism of Lactic Acid Bacteria in Distinct Ecological Niches: A Multifaceted Functional Health Perspective. *Frontiers in Microbiology*, 9, 2899. <https://doi.org/10.3389/fmicb.2018.02899>
- Georgelis, N., Fencil, K., & Richael, C. M. (2018). Validation of a rapid and sensitive HPLC/MS method for measuring sucrose, fructose and glucose in plant tissues. *Food Chemistry*, 262, 191–198. <https://doi.org/10.1016/j.foodchem.2018.04.051>
- Gophen, M. (2018). Temperature Impact on the Shannon-Wiener Biodiversity Index (BDI) of Zooplankton in Lake Kinneret (Israel). *Open Journal of Modern Hydrology*, 8(2), Article 2. <https://doi.org/10.4236/ojmh.2018.82004>
- Granström, T. B., Takata, G., Tokuda, M., & Izumori, K. (2004). Izumoring: A novel and complete strategy for bioproduction of rare sugars. *Journal of Bioscience and Bioengineering*, 97(2), 89–94. [https://doi.org/10.1016/S1389-1723\(04\)70173-5](https://doi.org/10.1016/S1389-1723(04)70173-5)
- Hadj Saadoun, J., Levante, A., Ferrillo, A., Trapani, F., Bernini, V., Galaverna, G., Neviani, E., & Lazzi, C. (2022). Exploring the Potential of Lactic Acid Fermentation for the Recovery of Exhausted Vanilla Beans. *Frontiers in Nutrition*, 9, 858716. <https://doi.org/10.3389/fnut.2022.858716>
- Halasz, A. (n.d.). Lactic Acid Bacteria. *Lactic Acid Bacteria*, 6.
- Hayek, S. A., Gyawali, R., Aljaloud, S. O., Krastanov, A., & Ibrahim, S. A. (2019). Cultivation media for lactic acid bacteria used in dairy products. *Journal of Dairy Research*, 86(4), 490–502. <https://doi.org/10.1017/S002202991900075X>
- Houle, D., Govindaraju, D. R., & Omholt, S. (2010). Phenomics: The next challenge. *Nature Reviews Genetics*, 11(12), Article 12. <https://doi.org/10.1038/nrg2897>
- Jonathan, M. C., DeMartini, J., Van Stigt Thans, S., Hommes, R., & Kabel, M. A. (2017). Characterisation of non-degraded oligosaccharides in enzymatically hydrolysed and fermented, dilute ammonia-pretreated corn stover for ethanol production. *Biotechnology for Biofuels*, 10(1), 112. <https://doi.org/10.1186/s13068-017-0803-3>
- Kim, E., Bae, J., Lee, J., Shin, J.-H., Seok, P. R., Kim, Y., & Yoo, S.-H. (2019). Purification and characterization of turanose, a sucrose isomer and its anti-inflammatory effects in dextran sulfate sodium (DSS)-induced colitis model. *Journal of Functional Foods*, 63, 103570. <https://doi.org/10.1016/j.jff.2019.103570>
- Koch, H., Girmscheid, N., Freese, H. M., Noriega-Ortega, B., Lücking, D., Berger, M., Qiu, G., Marzinelli, E. M., Campbell, A. H., Steinberg, P. D., Overmann, J., Dittmar, T., Simon, M., & Wietz, M. (2020). Genomic, metabolic and phenotypic variability shapes ecological differentiation and intraspecies interactions of *Alteromonas macleodii*. *Scientific Reports*, 10(1), Article 1. <https://doi.org/10.1038/s41598-020-57526-5>
- Kuleaşan, H., & Cakmakçi, M. L. (2002). Effect of reuterin produced by *Lactobacillus reuteri* on the surface of sausages to inhibit the growth of *Listeria monocytogenes* and *Salmonella* spp. *Die Nahrung*, 46(6), 408–410. [https://doi.org/10.1002/1521-3803\(20021101\)46:6<408::AID-FOOD408>3.0.CO;2-T](https://doi.org/10.1002/1521-3803(20021101)46:6<408::AID-FOOD408>3.0.CO;2-T)
- Lazzi, C., Bove, C. G., Sgarbi, E., Monica, G., La Gioia, F., Sandra, T., & Neviani, E. (2009). Application of AFLP fingerprint analysis for studying the biodiversity of *Streptococcus thermophilus*. *Journal of Microbiological Methods*, 79(1), 48–54. <https://doi.org/10.1016/j.mimet.2009.07.021>

- Levante, A., Lazzi, C., Vatsellas, G., Chatzopoulos, D., Dionellis, V. S., Makrythanasis, P., Neviani, E., & Folli, C. (2021). Genome Sequencing of five *Lacticaseibacillus* Strains and Analysis of Type I and II Toxin-Antitoxin System Distribution. *Microorganisms*, 9(3), 648. <https://doi.org/10.3390/microorganisms9030648>
- Lorn, D., Nguyen, T.-K.-C., Ho, P.-H., Tan, R., Licandro, H., & Waché, Y. (2021). Screening of lactic acid bacteria for their potential use as aromatic starters in fermented vegetables. *International Journal of Food Microbiology*, 350, 109242. <https://doi.org/10.1016/j.ijfoodmicro.2021.109242>
- Mekonnen, E., Kebede, A., Tafesse, T., & Tafesse, M. (2019). Investigation of carbon substrate utilization patterns of three ureolytic bacteria. *Biocatalysis and Agricultural Biotechnology*, 22, 101429. <https://doi.org/10.1016/j.bcab.2019.101429>
- Mieszkin, S., Richet, P., Bach, C., Lambrot, C., Augusto, L., Buée, M., & Uroz, S. (2021). Oak decaying wood harbors taxonomically and functionally different bacterial communities in sapwood and heartwood. *Soil Biology and Biochemistry*, 155, 108160. <https://doi.org/10.1016/j.soilbio.2021.108160>
- O'Connell, L. M., Kelleher, P., van Rijswijk, I. M. H., de Waal, P., van Peij, N. N. M. E., Mahony, J., & van Sinderen, D. (2022). Natural Transformation in Gram-Positive Bacteria and Its Biotechnological Relevance to Lactic Acid Bacteria. *Annual Review of Food Science and Technology*, 13(1), 409–431. <https://doi.org/10.1146/annurev-food-052720-011445>
- Overmann, J. (2015). Significance and future role of microbial resource centers. *Systematic and Applied Microbiology*, 38(4), 258–265. <https://doi.org/10.1016/j.syapm.2015.02.008>
- Padmavathi, T., Bhargavi, R., Priyanka, P. R., Niranjana, N. R., & Pavitra, P. V. (2018). Screening of potential probiotic lactic acid bacteria and production of amylase and its partial purification. *Journal of Genetic Engineering and Biotechnology*, 16(2), 357–362. <https://doi.org/10.1016/j.jgeb.2018.03.005>
- Picot-Allain, M. C. N., Ramasawmy, B., & Emmambux, M. N. (2022). Extraction, Characterisation, and Application of Pectin from Tropical and Sub-Tropical Fruits: A Review. *Food Reviews International*, 38(3), 282–312. <https://doi.org/10.1080/87559129.2020.1733008>
- Qi, X., & Tester, R. F. (2020). Lactose, Maltose, and Sucrose in Health and Disease. *Molecular Nutrition & Food Research*, 64(8), 1901082. <https://doi.org/10.1002/mnfr.201901082>
- Ranganathan, S., Dutta, S., Moses, J. A., & Anandharamakrishnan, C. (2020). Utilization of food waste streams for the production of biopolymers. *Heliyon*, 6(9), e04891. <https://doi.org/10.1016/j.heliyon.2020.e04891>
- Reinhardt, A., Johnsen, U., & Schönheit, P. (2019). L-Rhamnose catabolism in archaea. *Molecular Microbiology*, 111(4), 1093–1108. <https://doi.org/10.1111/mmi.14213>
- Ricci, A., Bernini, V., Maoloni, A., Cirilini, M., Galaverna, G., Neviani, E., & Lazzi, C. (2019). Vegetable By-Product Lacto-Fermentation as a New Source of Antimicrobial Compounds. *Microorganisms*, 7(12), 607. <https://doi.org/10.3390/microorganisms7120607>
- Ricci, A., Cirilini, M., Calani, L., Bernini, V., Neviani, E., Del Rio, D., Galaverna, G., & Lazzi, C. (2019). In vitro metabolism of elderberry juice polyphenols by lactic acid bacteria. *Food Chemistry*, 276, 692–699. <https://doi.org/10.1016/j.foodchem.2018.10.046>

- Ricci, A., Cirlini, M., Levante, A., Dall'Asta, C., Galaverna, G., & Lazzi, C. (2018). Volatile profile of elderberry juice: Effect of lactic acid fermentation using *L. plantarum*, *L. rhamnosus* and *L. casei* strains. *Food Research International*, 105, 412–422. <https://doi.org/10.1016/j.foodres.2017.11.042>
- Riemann, L., & Azam, F. (2002). Widespread N-Acetyl-d-Glucosamine Uptake among Pelagic Marine Bacteria and Its Ecological Implications. *Applied and Environmental Microbiology*, 68(11), 5554–5562. <https://doi.org/10.1128/AEM.68.11.5554-5562.2002>
- Sáez, G. D., Flomenbaum, L., & Zárate, G. (2018). Lactic Acid Bacteria from Argentinean Fermented Foods: Isolation and Characterization for their Potential Use as Starters for Fermentation of Vegetables. *Food Technology and Biotechnology*, 56(3), 398–410. <https://doi.org/10.17113/ftb.56.03.18.5631>
- Santagata, R., Ripa, M., Genovese, A., & Ulgiati, S. (2021). Food waste recovery pathways: Challenges and opportunities for an emerging bio-based circular economy. A systematic review and an assessment. *Journal of Cleaner Production*, 286, 125490. <https://doi.org/10.1016/j.jclepro.2020.125490>
- Sasaki, M., Jojima, T., Inui, M., & Yukawa, H. (2008). Simultaneous utilization of d-cellobiose, d-glucose, and d-xylose by recombinant *Corynebacterium glutamicum* under oxygen-deprived conditions. *Applied Microbiology and Biotechnology*, 81(4), 691–699. <https://doi.org/10.1007/s00253-008-1703-z>
- Sgarbi, E., Bottari, B., Gatti, M., & Neviani, E. (2014). Investigation of the ability of dairy nonstarter lactic acid bacteria to grow using cell lysates of other lactic acid bacteria as the exclusive source of nutrients. *International Journal of Dairy Technology*, 67(3), 342–347. <https://doi.org/10.1111/1471-0307.12132>
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3), 379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Socas-Rodríguez, B., Álvarez-Rivera, G., Valdés, A., Ibáñez, E., & Cifuentes, A. (2021). Food by-products and food wastes: Are they safe enough for their valorization? *Trends in Food Science & Technology*, 114, 133–147. <https://doi.org/10.1016/j.tifs.2021.05.002>
- Steele, A. D., Teijaro, C. N., Yang, D., & Shen, B. (2019). Leveraging a large microbial strain collection for natural product discovery. *Journal of Biological Chemistry*, 294(45), 16567–16576. <https://doi.org/10.1074/jbc.REV119.006514>
- Tao, S., Zhou, D., Chen, H., Li, N., Zheng, L., Fang, Y., Xu, Y., Jiang, Q., & Liang, W. (2023). Analysis of genetic structure and function of clustered regularly interspaced short palindromic repeats loci in 110 *Enterococcus* strains. *Frontiers in Microbiology*, 14, 1177841. <https://doi.org/10.3389/fmicb.2023.1177841>
- Ucar, R. A., Pérez-Díaz, I. M., & Dean, L. L. (2020). Gentiobiose and cellobiose content in fresh and fermenting cucumbers and utilization of such disaccharides by lactic acid bacteria in fermented cucumber juice medium. *Food Science & Nutrition*, 8(11), 5798–5810. <https://doi.org/10.1002/fsn3.1830>
- Usmani, Z., Sharma, M., Awasthi, A. K., Sharma, G. D., Cysneiros, D., Nayak, S. C., Thakur, V. K., Naidu, R., Pandey, A., & Gupta, V. K. (2021). Minimizing hazardous impact of food waste in a circular economy – Advances in resource recovery through green strategies. *Journal of Hazardous Materials*, 416, 126154. <https://doi.org/10.1016/j.jhazmat.2021.126154>

- Venkata Mohan, S., Varjani, S., Pant, D., Sauer, M., & Chang, J.-S. (2020). Circular bioeconomy approaches for sustainability. *Bioresource Technology*, 318, 124084. <https://doi.org/10.1016/j.biortech.2020.124084>
- Wang, G., Sun, J., Yao, Y., An, X., Zhang, H., Chu, G., Jiang, S., Guo, Y., Sun, X., & Liu, Y. (2020). Detection of Inosine Monophosphate (IMP) in Meat Using Double-Enzyme Sensor. *Food Analytical Methods*, 13(2), 420–432. <https://doi.org/10.1007/s12161-019-01652-y>
- Włodarczyk, M., Śliżewska, K., Barczyńska, R., & Kapuśniak, J. (2022). Effects of Resistant Dextrin from Potato Starch on the Growth Dynamics of Selected Co-Cultured Strains of Gastrointestinal Bacteria and the Activity of Fecal Enzymes. *Nutrients*, 14(10), Article 10. <https://doi.org/10.3390/nu14102158>
- Wu, H., Zhang, W., & Mu, W. (2019). Recent studies on the biological production of D-mannose. *Applied Microbiology and Biotechnology*, 103(21), 8753–8761. <https://doi.org/10.1007/s00253-019-10151-3>
- Wu, L., Zhang, Y., Guo, X., Ning, D., Zhou, X., Feng, J., Yuan, M. M., Liu, S., Guo, J., Gao, Z., Ma, J., Kuang, J., Jian, S., Han, S., Yang, Z., Ouyang, Y., Fu, Y., Xiao, N., Liu, X., ... Zhou, J. (2022). Reduction of microbial diversity in grassland soil is driven by long-term climate warming. *Nature Microbiology*, 7(7), Article 7. <https://doi.org/10.1038/s41564-022-01147-3>
- Zheng, J., Ruan, L., Sun, M., & Gänzle, M. (2015). A Genomic View of Lactobacilli and Pediococci Demonstrates that Phylogeny Matches Ecology and Physiology. *Applied and Environmental Microbiology*, 81(20), 7233–7243. <https://doi.org/10.1128/AEM.02116-15>
- Zheng, J., Wittouck, S., Salvetti, E., Franz, C. M. A. P., Harris, H. M. B., Mattarelli, P., O'Toole, P. W., Pot, B., Vandamme, P., Walter, J., Watanabe, K., Wuyts, S., Felis, G. E., Gänzle, M. G., & Lebeer, S. (2020). A taxonomic note on the genus Lactobacillus: Description of 23 novel genera, emended description of the genus Lactobacillus Beijerinck 1901, and union of Lactobacillaceae and Leuconostocaceae. *International Journal of Systematic and Evolutionary Microbiology*, 70(4), 2782–2858. <https://doi.org/10.1099/ijsem.0.004107>

WEB REFERENCES

- <https://www.atcc.org/>
- <https://bacdive.dsmz.de/>
- <https://kegg.jp/>
- <https://hmdb.ca/>
- <https://blast.ncbi.nlm.nih.gov/>
- <http://insilico.ehu.eus/AFLP/>
- <https://www.mirri-it.it/>

Chapter 2

Understanding of lactic acid bacteria tolerance to extreme environments and antimicrobial compounds

Laura Troiani¹, Camilla Lazzi^{1*}, Alessia Levante¹, Vincenzo Bonnici^{2*}, Valentina Bernini¹,
Velia Pierdomenico², Alessandro Dal Palù², Erasmo Neviani¹

¹ Department of Food and Drug Sciences, Parma University, Parma, Italy

² Department of Mathematical, Physical and Computer Sciences, University of Parma, Parma, Italy

Manuscript in preparation

ABSTRACT

Lactic acid bacteria (LAB) are widespread in different niches, and this leads them to face several challenging growth conditions. Concurrently, this ubiquity allowed them to evolve stress response strategies that favour resistance to adverse abiotic factors. The purpose of the study was to screen a core of 150 strains (SMC) properly selected from the University of Parma's Culture Collection (UPCC) to understand their tolerance to different types of stress they may cope with, as well as their resistance to different antimicrobials. This was achieved by using phenotypic microarrays, with the perspective of facilitating the selection process and the design of ad hoc processes in those cases in which the substrates available do not match the physiological needs of the microorganisms involved.

Strains were grown in microplates under extreme levels of temperature, pH, and salt concentration to make them cope with the main stress they encounter in industrial processes. Results underline the phenotypic biodiversity within the SMC and the remarkable adaptability of *L. paracasei*, *L. rhamnosus*, *L. zae*, *E. faecalis* and *E. faecium*. Additionally, the chemical sensitivity of LAB was tested over 21 different inhibiting substrates and antibiotics using Biolog GENIII Microplates (BIOLOG Inc.©, USA). This assay highlighted the multi-resistance of strains to different compounds.

The knowledge gained from this study helps to identify strains tolerant to hostile environments, such as waste substrates, and specifically employ them thus overcoming time-consuming and expensive issues in the upstream process. Moreover, antibiotic resistance of screened LAB reminds the need to limit the use of antibiotic-resistant (AR) microorganisms in industrial processes to prevent the spreading of AR genes.

KEYWORDS

Phenotyping, Lactic acid bacteria, Stress tolerance, Chemical sensitivity, Microbial platform

ABBREVIATIONS

LAB = Lactic Acid Bacteria

SMC = Selected Microbial Core

UPCC = University of Parma's Culture Collection

OPT = optimal condition

NaCl = saline environment (NaCl = 6%)

pHmax = alkaline environment at pH = 9

pHmin = acidic environment at pH = 4

Tmax = high temperature (T = 50°C)

Tmin = low temperature (T = 15°C)

AR = Antibiotic Resistance

1. INTRODUCTION

In industrial applications, microorganisms are usually considered in their ideal laboratory conditions, but what if we pull them out of their comfort zone? Beyond the physiological status of the cells, microbial growth and survival are strongly affected by environmental factors (Tsakalidou & Papadimitriou, 2011).

Lactic acid bacteria (LAB) are present in any environment that is abundant in carbohydrates, including plants and animals. They face various extreme conditions, including high and low temperatures, acidic and alkaline environments, osmotic pressure, oxidative stress, and starvation, during food processing and in the gastrointestinal tract (Mbye et al., 2020). These stress significantly and differently affect the biological functions of microorganisms, forcing them to implement various strategies to survive such as cross-protection (Yang et al., 2021), and develop stress responses and species-specific adaptation mechanisms that safeguard them from adverse environmental conditions.

LAB are exposed to high solute concentrations in nature (e.g. salt), for instance when changes in water availability occur, or in the food industry during certain manufacturing processes, such as meat fermentation and cheese ripening (Tsakalidou & Papadimitriou, 2011). The osmotic pressure on the cells causes excessive dehydration reducing the survival rate and affecting various metabolic activities.

The acidic condition represents a self-inflicted stress since lactic acid is the primary end-product of LAB metabolism. Organic acids produced or positively charged molecules present in the environment can challenge microorganisms by modifying the cytoplasmic anion pool and the intracellular pH, influencing cell growth and essential enzyme activities (Zhu et al., 2022). The alkaline environment also causes the anion-proton's unbalance between the inside and outside of the cell.

The use of high and low temperatures is very common in the food industry sector to preserve food products and control microbial growth at safety levels, as well as for industrial purposes. Therefore, LAB employed in the manufacturing processes are subjected to various degrees of heat and cold stress. High temperatures cause the degradation and function loss of proteins and nucleic acids leading to the inhibition of metabolism, as well as the disruption of cell activities due to the increase of the membrane fluidity (Mbye et al., 2020). Low temperatures reduce the membrane fluidity altering the absorption of nutrients, elimination of metabolites, and secretion and transportation of proteins (Yang et al., 2022b). Moreover, cold stress affects the DNA or

RNA supercoiling, thus altering replication, transcription, and protein production (Moon et al., 2023).

Understanding the effects of the environment on microorganisms helps us to describe their distribution in nature, to control them, but also to promote their growth in unusual matrices. This work is focused on the latter advantage: if stress-tolerant bacteria are normally undesired because hard to control, we want to exploit this potential by using them in the fermentation of waste substrates, which provide a carbon-rich environment that is still usable but sometimes hostile to microbial growth.

Food waste and by-products are used as fermentation matrices to produce value-added compounds by LAB (Hadj Saadoun et al., 2022; Sabater et al., 2020) but most of the time they need pretreatments to be suitable for microorganisms or to recover as much as possible from the biomass. Based on the matrix and the substances to recover, pretreatments include physical, mechanical, chemical, and biological methods or combinations of them (Ojo & De Smidt, 2023; Rajesh Banu et al., 2020). These treatments modify the composition of the matrix and often release degradation compounds that can inhibit bacteria. Furthermore, the purification of treated matrices is costly and therefore increases the final product price (Gubelt et al., 2020). Other times, instead, changes in temperature, pH or other factors are implemented to reach and improve specific results. A recent study investigates the effects of combining high temperature with moderate low pH on lactic acid production, thus using less base for neutralization and facilitating the downstream processing (Yang et al., 2022a).

Nowadays, microorganisms have also to face other adverse factors like antimicrobials, raising concerns about the possibility of antimicrobial resistance traits diffusion in microbial ecosystems (Zambrano, 2023). In this case, knowing resistant strains is important to avoid the diffusion of antibiotic resistance (AR) genes.

The University of Parma owns a large collection of lactic acid bacteria (UPCC) that were screened for different features. This work wanted to investigate the tolerance of 150 lactic acid strains (SMC) to abiotic stress, as well as their resistance to different antimicrobials, through phenotypic microarrays.

From the perspective of research into industrial biochemical production from food waste and by-product fermentation, the purpose of this phenotyping was to understand the potential abilities of microorganisms under non-physiological conditions to develop ad hoc processes that recycle substrates otherwise destined for disposal.

2. MATERIALS AND METHODS

2.1. Preparation of the bacterial culture

A hundred fifty strains of lactic acid bacteria (LAB), isolated between 2012 and 2022 mostly from dairy products, were selected from the University of Parma's Culture Collection (UPCC). Cultures were preserved as frozen stocks at -80°C in their growth medium supplemented with 15% glycerol (w/v). The strains were then propagated three times using 3% (v/v) inoculum and incubated overnight under anaerobic conditions at the optimal temperature. The composition of the microbial core studied (SMC) is reported in Table 1.

Table 1 Composition of the Selected Microbial Core (SMC) object of the study. Reference: Boone et al., 2001; Zheng et al., 2020; <https://atcc.org/>; <https://bacdive.dsmz.de/>.

Genera	Species	No. strains	Medium ¹	Temperature
<i>Companilactobacillus</i>	<i>farciminis</i>	2	MRS	30°C
<i>Enterococcus</i>	<i>avium</i>	1	TSB	37°C
	<i>faecalis</i>	10		
	<i>faecium</i>	6		
<i>Lacticaseibacillus</i>	<i>casei</i>	2	MRS	37°C
	<i>paracasei</i>	14		
	<i>rhamnosus</i>	12		
	<i>zeae</i>	8		
<i>Lactiplantibacillus</i>	<i>plantarum</i>	5	MRS	37°C
<i>Lactobacillus</i>	<i>amyolyticus</i>	1	MRS	37°C
	<i>delbrueckii</i>	11		
	<i>gasseri</i>	1		
	<i>helveticus</i>	9		
<i>Lactococcus</i>	<i>lactis</i>	1	M17	30°C
<i>Latilactobacillus</i>	<i>curvatus</i>	10	MRS	30°C
<i>Lentilactobacillus</i>	<i>diolivorans</i>	1	MRS	30°C
	<i>hilgardii</i>	2		
	<i>otakiensis</i>	2		
<i>Leuconostoc</i>	<i>citreum</i>	8	MRS	30°C
	<i>lactis</i>	1		
	<i>mesenteroides</i>	6		
	<i>pseudomesenteroides</i>	1		
<i>Levilactobacillus</i>	<i>brevis</i>	8	MRS	30°C

<i>Limosilactobacillus</i>	<i>fermentum</i>	8	MRS	37°C
<i>Streptococcus</i>	<i>macedonicus</i>	2	M17	42°C
	<i>pasteurianus</i>	2	TSB	37°C
	<i>porcorum</i>	2		
	<i>thermophilus</i>	9		
	<i>uberis</i>	5		

1) MRS and TSB media were purchased from VWR International bvba (Leuven, Belgium), while M17 medium from Merck (Darmstadt, Germany).

2.2. Tolerance to extreme environmental conditions

The tolerance of strains to abiotic stress was investigated through phenotype microarrays, growing them in 96-well microplates on their optimal medium properly modified as reported in Table 2. The five conditions tested were chosen among the most distant points from the optimum recorded in the literature for LAB, to understand their ability to colonize extreme habitats.

Table 2 Environmental conditions tested through phenotype microarrays.

Stress	Temperature (°C)	pH	NaCl (%)
Osmotic pressure (NaCl)	30-37	6.5	6
Alkaline environment (pHmax)	30-37	9	0
Acidic environment (pHmin)	30-37	4	0
High temperature (pHmin)	50	6.5	0
Low temperature (Tmin)	15	6.5	0

Overnight precultures were used to inoculate (3% v/v) the 96-well microplates. The cultures were incubated until 72 h in aerobiosis, and the growth was determined through absorbance readings at 590 nm, at 9 different time points (0, 2, 4, 6, 8, 16, 24, 48, 72 h) by using the BIOLOG MicroReader Station (BIOLOG Inc.©, USA). Every bacterial strain culture was carried out in triplicate. Only cultures from strains showing positive growth results (n=133) were included in the analysis.

2.3. Chemical sensitivity assay by BIOLOG© Technology

To investigate the resistance to 23 chemicals, 10 to 12 columns of Biolog GENIII Microplates (BIOLOG Inc.©, USA) were employed. This technology consists of ready-to-use microplates whose wells are prefilled with the compounds to be tested and Tetrazolium Violet redox dye used as a colorimetric indicator of microbial respiration. A few modifications were applied to

the GENIII Microplates protocol. Biomass in 1 ml of 16 h old cultures was harvested by centrifugation at 13400 rpm for 10 min and washed twice in Ringer's solution (VWR International bvba, Leuven, Belgium). Cells were resuspended in 12 ml of inoculation fluid IF-A (BIOLOG Inc.©, USA) to a transmittance of 90-98%, corresponding in average to 10^5 CFU/ml. The inoculated IF-A was distributed into the 96 wells and the microplate was incubated at the strains' optimal temperature to allow the phenotypic fingerprint to form. Collected data refers to the absorbance of the reduced Tetrazolium Violet dye (purple coloured) read at 590 nm after 72 h by the BIOLOG MicroReader Station (BIOLOG Inc.©, USA).

2.4. Data Analysis

2.4.1. Tolerance to extreme environmental conditions

OD₅₉₀ values related to growth were referred to the blank and the general growth trend was observed.

The ability of growing under the adverse conditions reported in Table 2 was examined calculating the maximum growth increase with respect to the inoculum using the following equation (Eq. 1):

$$\text{Growth increase} = \frac{OD_{max}}{OD(t_0)} \quad (1)$$

Results were visualized by box and whiskers plot (R version 4.3.1, Package “ggplot2” version 3.4.2): species with growth increments different from zero were considered tolerant to the specific environmental factor.

To investigate the single strain's ability, raw data were elaborated by means of the T-SNE (Hinton & Roweis, 2002) to reduce the dimensionality of the dataset (more than 24300 OD₅₉₀ readings).

2.4.2. Chemical sensitivity of strains

The Biolog raw data were elaborated with Rstudio (R version 4.3.1). Having single end-point measurements, the Average Well Colour Development (AWCD) was calculated as suggested by previous studies (Acin-Albiac et al., 2020; Dubey et al., 2022). To reduce the noise levels, absorbance values were referred to half of the positive control well (A10) and subsequently divided by the respective AWCD (Tab. 3). Negative values were set to 0. Normalized data were used for statistical analysis (Package “corrplot” version 0.92) and data visualization (Package “pheatmap” version 1.0.12). P-values < 0.05 were considered significant.

Table 3 Normalization of the chemical sensitivity assay results.

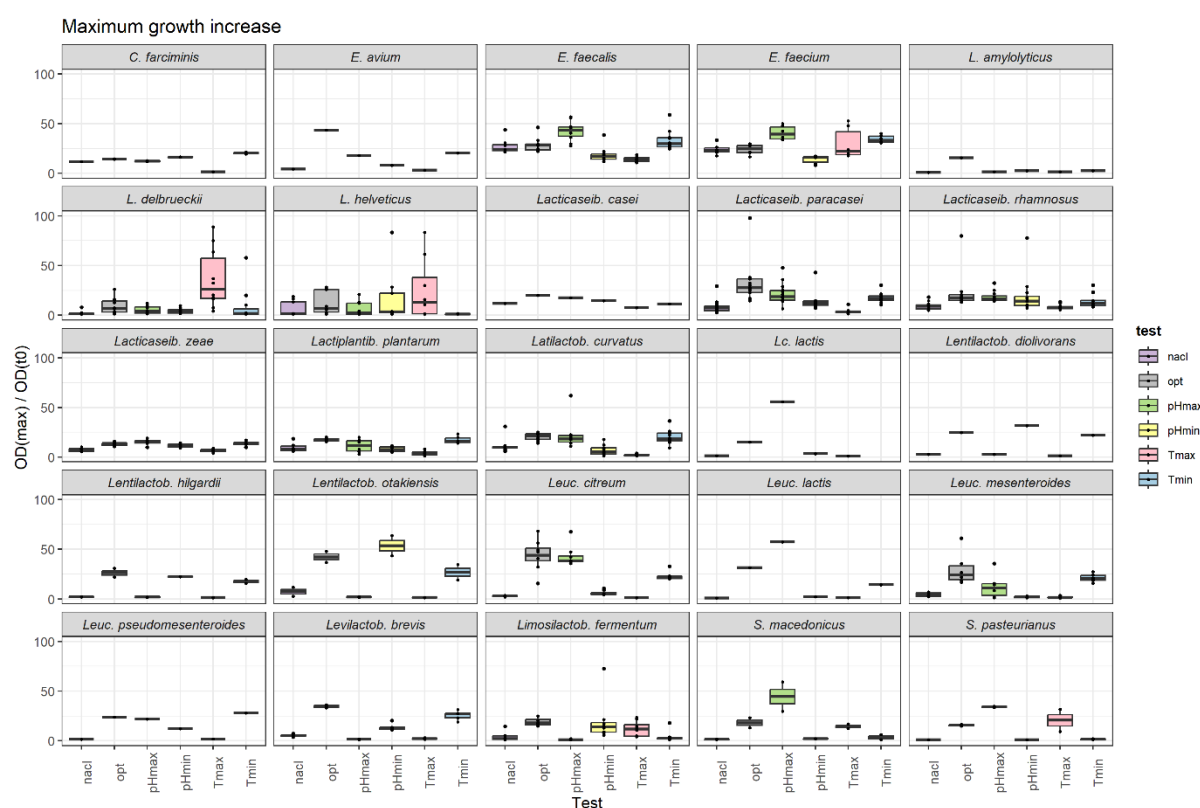
Estimation	Formula	Chemical sensitivity
Average Well Colour Development (AWCD)	$AWCD = \sum_{i=1}^n (ODi - R)/n$	ODi = raw OD_{590} value of each well from 10 to 12 columns R = half of the OD_{590} value of positive control well n = number of substrates analysed (23)
Normalization	$Rsi = (ODi - R)/AWCD$	

3. RESULTS

3.1. Tolerance to extreme environmental conditions

Looking at the general behaviour of the species (Fig. 1), the further away the growth conditions from the optimum, the longer the lag phase and the lower the growth increase. It emerges an inter-specific variability in adaptation and maximum OD₅₉₀ reached. However, also within the same species, it is noteworthy that strains can take a different time to reach the same OD₅₉₀ and the other way around. In Figure 1, the maximum growth increments of each species are reported, allowing us also to qualitatively compare their stress tolerance.

Figure 1 Growth increase represented by the maximum OD₅₉₀ reached by every strain during 72 h over the starting OD₅₉₀ value at time = 0 (see Eq. 1).



Applying an osmotic pressure (NaCl = 6%) on the cultures, *E. faecalis* and *E. faecium* particularly tolerate this condition, while most of the other species do not tolerate it or do not show a great increase in growth compared to the starting point.

Moving away from the neutrality, in an alkaline environment (pHmax = 9), a growth increment occurs for 17 species out of 25, especially in *E. faecalis*, *E. faecium*, *Lc. Lactis*, *Leuc. citreum*, *Leuc. lactis*, *S. macedonicus*, and *S. pasteurianus*.

On the opposite, cultures generally suffer acid environment ($\text{pH}_{\text{min}} = 4$), thus showing a low growth increase or no growth except for *L. otakiensis* which appears tolerant to low pH.

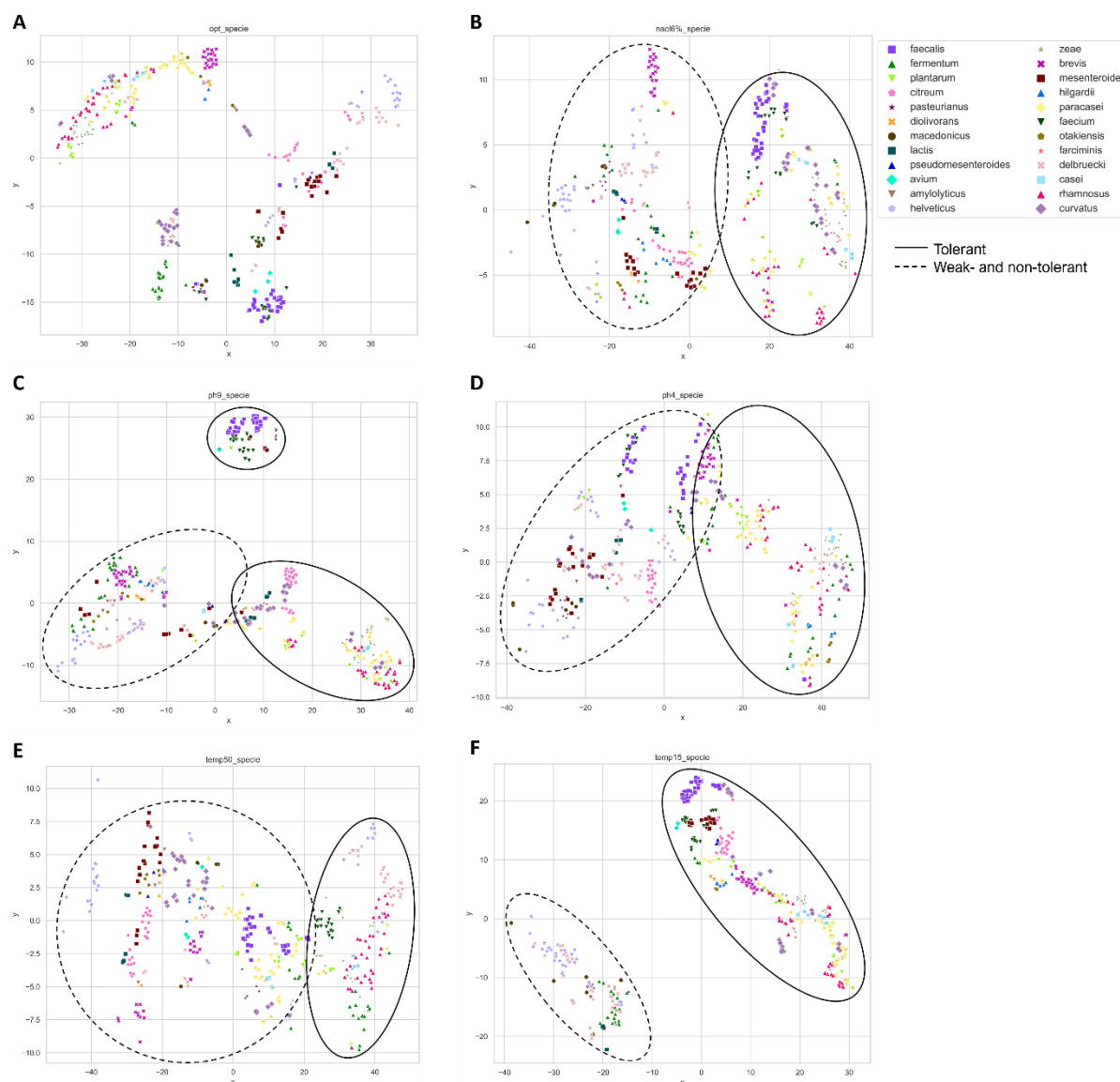
High temperature ($T_{\text{max}} = 50^{\circ}\text{C}$) is particularly tough for strains, however, *E. faecalis*, *E. faecium*, *L. delbrueckii*, *L. helveticus*, *L. fermentum*, *S. macedonicus*, and *S. pasteurianus* appear to grow under heat stress.

The cold stress ($T_{\text{min}} = 15^{\circ}\text{C}$) obviously causes an extension of the lag phase. However, the OD_{590} increase highlights that most of the species are psychrotolerant, while *Lactobacillus* spp., *Lc. lactis*, *L. fermentum*, and *Streptococcus* spp. are not tolerant to low temperatures.

Furthermore, it is interesting to note the phenotypic variance observed within the *L. helveticus* species in every condition.

3.1.1. T-SNE analysis

Going into more detail, the T-SNE analysis considers OD_{590} readings at every single time-point for every strain, thus summarizing the growth curves' trend and confirming the inter- and intra-specific variability. This elaboration (Fig. 2) clusters strains based on their similarities and/or differences: what is close in the multidimensional space is also close in the bi-dimensional space, and the other way around.

Figure 2 Visualization of microbial growth data elaborated through T-SNE analysis.

In the optimal condition (OPT), *Lacticaseibacillus* spp., *L. plantarum* and *L. brevis* reach the highest OD₅₉₀ of the screening and cluster together on the top-left side of Figure 2A. On the opposite, we see *Enterococcus* spp., *L. delbrueckii*, *L. helveticus* and *Leuconostoc* spp. that do not overcome OD₅₉₀ = 0.6 in 72 h of incubation. At the optimum, strains belonging to the same species generally are close to each other with some exceptional subset within some species (e.g. *Leuc. citreum*).

Inside the NaCl panel (Fig. 2B), *Enterococcus* spp., *Lacticaseibacillus* spp., *L. plantarum* and *L. curvatus* group on the right side, opposite to other species that do not tolerate saline environments. Under this condition, strains of *L. paracasei* show different behaviours that move them close to non-tolerant species, and the other way around for *L. fermentum* and *L. helveticus*.

Enterococcus spp. greatly tolerates alkaline environments (pH_{max}), indeed they are completely isolated from other species on the top of Figure 2C. *Lacticaseibacillus* spp., *L. plantarum*, *L. curvatus* and *Leuc. citreum* are still tolerant although some strains behave differently and so move to the dotted cluster. *Leuconostoc* spp. is halfway between tolerant and non-tolerant strains since their growth curves differ from the others for a rapid increase in the first 8 h, and after that, an extremely slow decay begins.

Lacticaseibacillus spp. and *L. fermentum*, although slightly, grow in the acidic environment (pH_{min}), therefore they are together in the bottom-right side of Figure 2D, differently from all other species that do not greatly tolerate this condition. However, some strains of *L. paracasei*, *L. zae* and *L. fermentum* move away from others, as well as within the dotted cluster the same happens for strains of *Enterococcus* spp.

At T_{max} (Fig. 2E), species which are more thermotolerant and show regular growth curves (*L. delbrueckii*, *L. helveticus* and *E. faecium*) are situated in the right part of the solid-lined cluster, while *L. rhamnosus*, *L. zae* and *L. fermentum* which reach a pick around 8 h but then decrease rapidly, still fall in the same cluster but their high dispersion brings them to join the other group. On the other cluster, we see subsets of *L. brevis*, and some *L. delbrueckii* and *L. helveticus* (particularly variable in this condition).

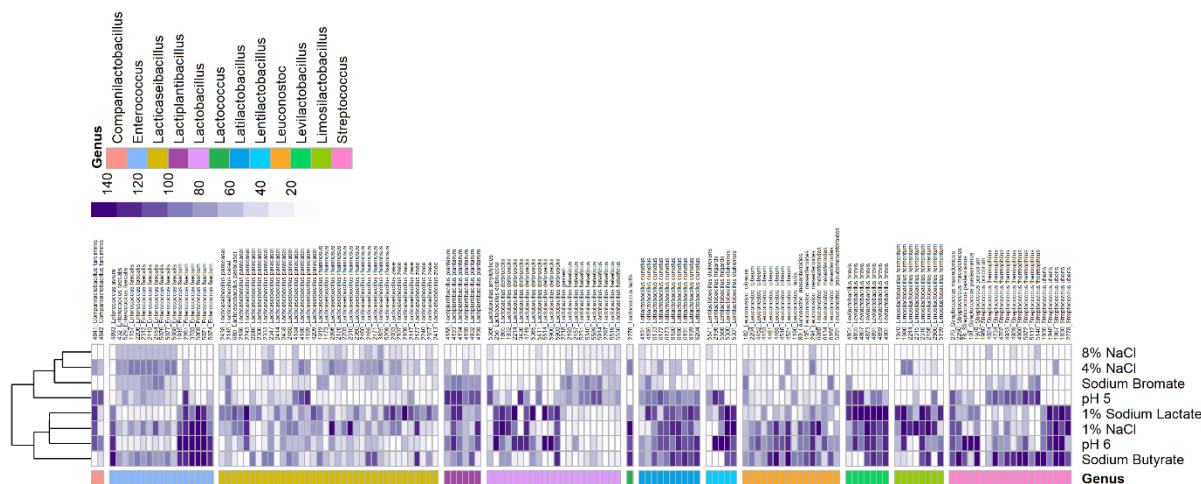
At T_{min} (Fig. 2F), there is a clear distinction between psychrotolerant microorganisms and non, with a little variability within species. *Lacticaseibacillus* spp. and *L. plantarum*, species that reach the highest OD, are grouped in the lower part of the solid-lined cluster, while *Enterococcus* spp., *Leuconostoc* spp. and *L. brevis* are moved to the top. All other species in the dotted cluster do not even arrive at OD₅₉₀ = 0.2 during the 72 h.

3.2. Chemical sensitivity assay

3.2.1. Sensitivity to different artificial environments found in the food industry

Setting 1 as the cut-off of normalized OD₅₉₀ values to establish the tolerance to man-made environments, the general trend emerging from Figure 3 highlights the resistance of more than half of the screened strains to 1% Sodium Lactate (114 strains), slightly acidic conditions such as pH = 6 (104 strains), Sodium Butyrate (94 strains) and 1% NaCl (77 strains). Only three strains can significantly endure Sodium Bromate (*L. plantarum* UPCC-4194, *Leuc. citreum* UPCC-182 and *S. thermophilus* UPCC-5140) and none strongly resists to 8% NaCl. *L. helveticus* appears to be the least resistant species of the screening, *E. faecalis* shows a high sensitivity common to all its strains, while *Lacticaseibacillus* spp., always very sensitive, displays a wide variability among strains.

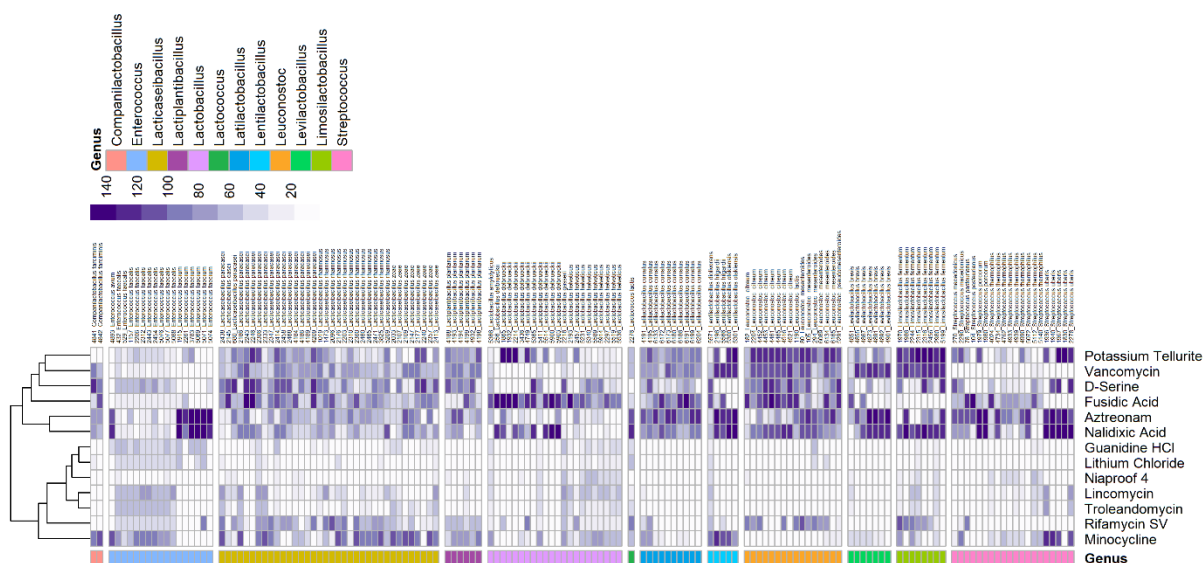
Figure 3 Sensitivity profiling of the SMC strains subjected to different bacterial-inhibiting substrates, usually present in man-made environments such as antimicrobials encountered in the food industry.



3.2.2. Sensitivity to different antibiotics and chemotherapics

In Figure 4 is reported a strong antibiotic resistance (AR) of more than half of all strains to Fusidic Acid (125 strains), Aztreonam (104 strains), Nalidixic Acid (101 strains), Vancomycin (88 strains), Potassium Tellurite (86 strains) and D-Serine (53 strains). However, some strain seems to slightly tolerate Rifamycin SV and Minocycline. As for the man-made environments, *L. helveticus* and *E. faecalis* are the least resistant species of the SMC, as well as *Lacticaseibacillus* spp. which show very different levels of sensitivity among its strains.

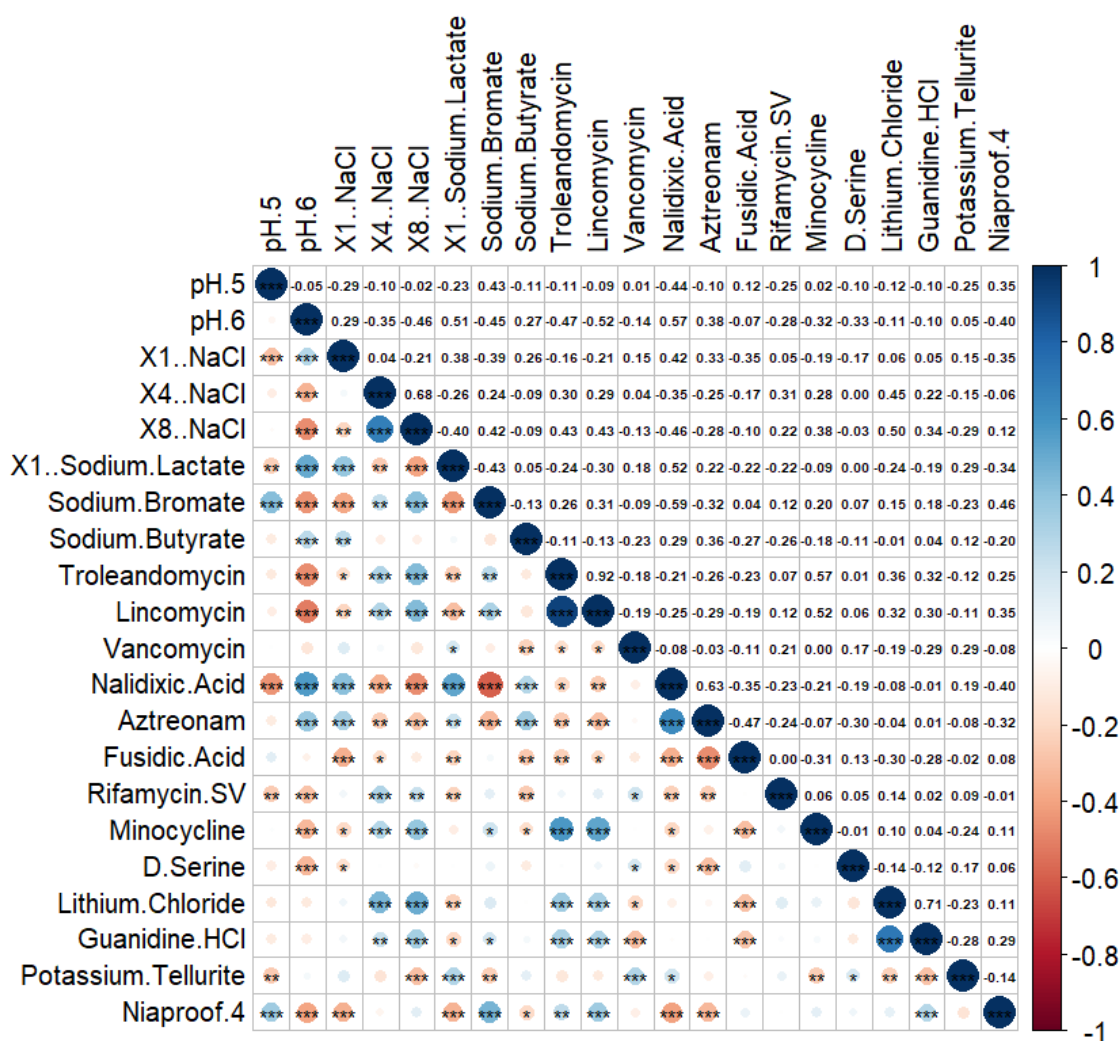
Figure 4 Profiling of SMC strains' sensitivity to different antibiotics and chemoterapics, used in several industrial applications.



3.2.3. Correlation among chemicals

Looking at Figure 5 it is possible to remark on many interesting correlations between antimicrobials and antibiotics. There is a strong and significative correlation (Pearson coefficient (ρ) = 0.92) between Troleandomycin and Lincomycin, and between Guanidine HCl and Lithium Chloride (ρ = 0.71). This means that strains resistant to one of the two antibiotics are also resistant to the other. Moreover, although moderate, there is a noteworthy direct correlation between some man-made environments and some antibiotics. For instance, 8% NaCl is correlated to both Troleandomycin (ρ = 0.43) and Lincomycin (ρ = 0.43), Nalidixic Acid to pH = 6 (ρ = 0.57), 1% NaCl (ρ = 0.42) and 1% Sodium Lactate (ρ = 0.52), Lithium Chloride to both 4% (ρ = 0.45) and 8% NaCl (ρ = 0.50), and still Sodium Bromate and Niaproof 4 (ρ = 0.46).

Figure 5 Plotted correlation matrix composed of Pearson correlation coefficients (ρ) between each pairwise combination of normalized OD₅₉₀ values obtained from the chemical sensitivity assay. The upper triangular matrix reports ρ comprised between 1 and -1: $\rho > |0.70|$ indicates a strong correlation between the two variables. In the lower triangular matrix, the Pearson correlation is represented by a solid circle integrated with P-values: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.05$.



4. DISCUSSIONS

This work aimed to evaluate the tolerance and resistance of the SMC strains to environmental stress and antimicrobial compounds, respectively. A phenotypic approach was applied to study the behaviour of microorganisms under non-optimal growth conditions with the final goal of identifying and easily selecting performing strains able to grow in hostile environments, such as waste substrates. This can help to overcome the strains adaptation process and minimize the biomass pre-treatments, which are time-consuming, costly, and often non-eco-friendly.

This study also highlights the phenotypic diversity among LAB and even the variability within the same species. The results have shown us that microorganisms can survive under extreme conditions, sometimes at levels comparable to or even higher than the optimal one, confirming what is reported in the literature (H. Yang et al., 2021; Mbye et al., 2020; Yin et al., 2019).

The osmotic stress heavily affects the growth, causing the total inhibition of 14 species out of 25 (Fig. 1). It always causes a decrease of the biomass production and stability in cell viability since for many strains the OD₅₉₀ collapses as soon as the growth peak is reached, suggesting the occurrence of a very short stationary phase. LAB activate specific mechanisms to keep the turgor pressure of the cell, thus preventing cellular death by osmotic shock (Mbye et al., 2020). To keep water within the cell, lactobacilli can accumulate specific solutes that do not affect cell physiology. For instance, *L. plantarum* synthesizes and accumulates quaternary ammonium compounds which allow it to resist longer to the inhibition (Papadimitriou et al., 2016). Within the purpose of this work, halophilic strains become interesting in industrial processes in which salt-rich substrates are involved. For instance, they can be used to recover seaweed biomass which is characterised by a high concentration of salt but still represents a promising alternative C source for fermentations (Papadopoulou et al., 2023), or even dairy by-products (Estrada et al., 2023).

Most of the screened strains (Fig. 2C) tolerate alkaline stress and likely this is due to the concomitance of growth and lactic acid production which acidifies the medium moderating the effect of high pH. Still, LAB can alkalinize the cytoplasm thus regulating their intra-cellular pH or use enzymes active at high pH (Mbye et al., 2020), and sodium- and potassium-proton antiporters demonstrated to be essential for microbial viability under alkaline conditions (Zhang et al., 2017). Additionally, the aggregation as biofilm and the production of ESP create a barrier against the external environment (Yin et al., 2019). In a fermentation, pH far from neutrality brings benefits to the process itself: neutral conditions facilitate contamination and sometimes lower the efficiency of fermentation. From the perspective of waste valorisation, it has been

demonstrated that alkaliphiles can carry out fermentations at high pH of lignocellulosic biomass obtained from banana peels (Abdel-Rahman et al., 2019) and municipal solid waste (Cheah et al., 2019) with interesting production of lactic acid and volatile fatty acid, respectively.

The acidic environment is challenging for bacteria, indeed only a few species involved in the study were able to grow at pH_{min} (Fig. 1). However, since they cope with acidity for their nature, they evolved strategies to survive it. Examples of defence mechanisms activated during acid stress are the exploitation of the amino acids metabolism to neutralize the cytoplasm, the production of EPS which isolate the cell cytoplasm from free H⁺ (Nguyen et al., 2020), and modifications of the cell membrane composition (Kathiriya et al., 2023). The tolerance to acid stress prevents the neutralization of acidic substrates such as those obtained after the pre-treatment of lignocellulosic biomass through acid hydrolysis (Chatgasem et al., 2023), which is one of the bottlenecks for a green process. Figuring acidophiles out is also useful in the food industry where LAB are essential for the manufacturing process and low pH is used as an antimicrobial strategy against pathogens and spoilage bacteria.

The minority of screened strains, 9 species out of 25, exhibited thermal adaptation properties (Fig. 2). Nevertheless, especially in the food industry, LAB often face heat stress. This leads them to stress responses such as changes in gene and protein expression (Palud et al., 2018), production of specific conserved proteins implicated in the repair mechanism of cells (Bommasamudram et al., 2022), regulation of metabolic activity by expression of toxin-antitoxin systems (Levante et al., 2019), or availability of saturated and straight-chain fatty acids that guarantee a proper membrane fluidity (Min et al., 2020). Both Chatgasem et al. (2023) and Yang et al. (2022) recently described the advantage of selecting thermophilic bacteria to increase the efficiency of lactic acid production from rice straw and food waste, respectively.

The cold stress causes an extension of the adaptation time, but SMC strains still tolerate a temperature of 15°C to finally reach OD₅₉₀ values often comparable to or even higher than those obtained at their optimal growth temperature. In this case, the positive reaction to stress could be due to the production of antifreeze and cold-shock proteins (CSPs) or cold-adapted enzymes (CAEs) as reported by Jin et al. (2022). The tolerance to low temperatures is particularly important for starters and probiotics management and production or preservation of microbial collections, in which bacteria are exposed to stabilization processes (Tsakalidou & Papadimitriou, 2011). Psychrotolerant microorganisms become interesting when cold is employed to preserve the chemical composition of the substrates to ferment or the organoleptic features of food products (Kazakos et al., 2020).

The knowledge of the physiological abilities of LAB helps both to select suitable strains based on fermentation requirements and avoid undesirable invaders. Summarizing, in terms of adaptability, *L. paracasei*, *L. rhamnosus*, *L. zeae*, *E. faecalis* and *E. faecium* appear the most promising species within the screening since they tolerate well the stress imposed, while others struggle more to get out of their comfort zone. Results confirm the fact that ubiquity is a sort of adaptation process: different niches and environmental changes provide a wide variability in lifestyle which helps to evolve mechanisms to more easily overcome physiological challenges (Tsakalidou & Papadimitriou, 2011).

Other adverse factors microorganisms must face are antimicrobial compounds and antibiotics. Here the issue, with global health and economic consequences, is the resistance of LAB to these chemicals and, accordingly, the transfer of AR genes to other microorganisms or animals and humans. Since LAB are inevitably associated with fermented food, the food industry can represent a starting environment for AR spreading (Nunziata et al., 2022).

In this paper, we dealt with the susceptibility of screened strains to different man-made environments, easily encountered in the food industry, and antibiotics. Our results interestingly highlight in the majority of strains a low sensitivity to 1% Sodium Lactate (Fig. 3), used in the food industry for its antimicrobial effect even against pathogens (Tian et al., 2022; de Wit & Rombouts, 1990).

Furthermore, it is particularly noteworthy as in a limited microbial core, like the SMC, there is a strong multi-resistance to different antibiotics (Fig. 4-5). The recognition of AR within a microbial collection is necessary to avoid the cross-dissemination of AR genes and helps to formulate selective media for pure cultures' studies. Concurrently, it reminds the conscientious and responsible use of LAB in the industry to avoid the over-exploitation of antibiotic-resistant microorganisms which increases the risk of AR gene transfer to other microorganisms, including pathogens (Arias & Murray, 2015). Our study also notes that, generally, strains showing high OD₅₉₀ values within this chemical sensitivity test appear to be more specifically tolerant for certain chemicals. On the other hand, strains reporting lower OD₅₉₀ values (*E. faecalis*, *Lacticaseibacillus* spp., *L. helveticus*) show a lower degree of resistance, but to many more substrates (Fig. 3-4). This reflects the scientific literature especially related to the *Enterococcus* genus, which is particularly widespread in several environments, including hospitals. This allowed bacteria to evolve an intrinsic resistance and the possibility to easily acquire and transfer AR genes, thus explaining also because many species of this genus are associated with infections (Golob et al., 2019; Nunziata et al., 2022; Tsakalidou & Papadimitriou, 2011).

5. CONCLUSIONS

This work collects information about the ability of LAB to tolerate hostile environmental conditions and resist antimicrobial compounds to boost the knowledge related to the UPCC isolates, with the purpose of building up a microbial platform for targeted industrial applications. It involves a microbial core properly selected and adopts a phenotyping approach to investigate the aptitude of strains. In particular, this study highlights the phenotypic biodiversity within the SMC, the higher adaptability of few species rather than others, and the multi-resistance to several chemicals.

The knowledge acquired can aid in the identification of microorganisms particularly capable of leading specific fermentation processes, thereby facilitating the targeted design and optimization of such processes. It also remarks the importance of responsible use of antibiotics and antibiotic-resistant microorganisms to avoid the dissemination of AR genes, which is rather uncontrolled already.

Understanding microbial capabilities also enables a comprehensive investigation into microorganisms' adaptation mechanisms. This allows to maximize the physiological potential of microorganisms by driving the selection and process optimization towards specific niches and/or targeted biochemical production.

FUNDING INFORMATION

This work was granted by the European Commission—NextGenerationEU, Project “Strengthening the MIRRI Italian Research Infrastructure for Sustainable Bioscience and Bioeconomy”, code n. IR00000005. V. Bonnici is partially supported by the University of Parma (Italy) protect n. MUR_DM737_B_MAFI_BONNICI, and by the CINI InfoLife laboratory.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Alessandra Masci for technical support in the maintenance and operation of the frozen cultures from the UPCC collection.

REFERENCES

- Abdel-Rahman, M. A., Hassan, S. E.-D., Roushdy, M. M., Azab, M. S., & Gaber, M. A. (2019). Free-nutrient supply and thermo-alkaline conditions for direct lactic acid production from mixed lignocellulosic and food waste materials. *Bioresource Technology Reports*, 7, 100256. <https://doi.org/10.1016/j.biteb.2019.100256>
- Acin-Albiac, M., Filannino, P., Gobetti, M., & Di Cagno, R. (2020). Microbial high throughput phenomics: The potential of an irreplaceable omics. *Computational and Structural Biotechnology Journal*, 18, 2290–2299. <https://doi.org/10.1016/j.csbj.2020.08.010>
- Arias, C. A., & Murray, B. E. (2015). A New Antibiotic and the Evolution of Resistance. *New England Journal of Medicine*, 372(12), 1168–1170. <https://doi.org/10.1056/NEJMcibr1500292>
- Bommasamudram, J., Kumar, P., Kapur, S., Sharma, D., & Devappa, S. (2022). Development of Thermotolerant Lactobacilli Cultures with Improved Probiotic Properties Using Adaptive Laboratory Evolution Method. *Probiotics and Antimicrobial Proteins*. <https://doi.org/10.1007/s12602-021-09892-3>
- Boone, D. R., Castenholz, R. W., & Garrity, G. M. (Eds.). (2001). *Bergey's manual of systematic bacteriology* (2nd ed). Springer.
- Chatgasem, C., Suwan, W., Attapong, M., Siripornadulsil, W., & Siripornadulsil, S. (2023). Single-step conversion of rice straw to lactic acid by thermotolerant cellulolytic lactic acid bacteria. *Biocatalysis and Agricultural Biotechnology*, 47, 102546. <https://doi.org/10.1016/j.bcab.2022.102546>
- Cheah, Y.-K., Vidal-Antich, C., Dosta, J., & Mata-Álvarez, J. (2019). Volatile fatty acid production from mesophilic acidogenic fermentation of organic fraction of municipal solid waste and food waste under acidic and alkaline pH. *Environmental Science and Pollution Research*, 26(35), 35509–35522. <https://doi.org/10.1007/s11356-019-05394-6>
- de Wit, J. C., & Rombouts, F. M. (1990). Antimicrobial activity of sodium lactate. *Food Microbiology*, 7(2), 113–120. [https://doi.org/10.1016/0740-0020\(90\)90017-C](https://doi.org/10.1016/0740-0020(90)90017-C)
- Dubey, A., Malla, M. A., & Kumar, A. (2022). Taxonomical and functional bacterial community profiling in disease-resistant and disease-susceptible soybean cultivars. *Brazilian Journal of Microbiology*. <https://doi.org/10.1007/s42770-022-00746-w>
- Estrada, M., Navarrete, C., Møller, S., Procentese, A., & Martínez, J. L. (2023). Utilization of salt-rich by-products from the dairy industry as feedstock for recombinant protein production by *Debaryomyces hansenii*. *Microbial Biotechnology*, 16(2), 404–417. <https://doi.org/10.1111/1751-7915.14179>
- Golob, M., Pate, M., Kušar, D., Dermota, U., Avberšek, J., Papić, B., & Zdovc, I. (2019). Antimicrobial Resistance and Virulence Genes in *Enterococcus faecium* and *Enterococcus faecalis* from Humans and Retail Red Meat. *BioMed Research International*, 2019, e2815279. <https://doi.org/10.1155/2019/2815279>
- Gubelt, A., Blaschke, L., Hahn, T., Rupp, S., Hirth, T., & Zibek, S. (2020). Comparison of Different Lactobacilli Regarding Substrate Utilization and Their Tolerance Towards Lignocellulose Degradation Products. *Current Microbiology*, 77(10), 3136–3146. <https://doi.org/10.1007/s00284-020-02131-y>

- Hadj Saadoun, J., Levante, A., Ferrillo, A., Trapani, F., Bernini, V., Galaverna, G., Neviani, E., & Lazzi, C. (2022). Exploring the Potential of Lactic Acid Fermentation for the Recovery of Exhausted Vanilla Beans. *Frontiers in Nutrition*, 9, 858716. <https://doi.org/10.3389/fnut.2022.858716>
- Hinton, G., & Roweis, S. (2002). Stochastic Neighbor Embedding. *Stochastic Neighbor Embedding, Advances in neural information processing systems*(15).
- Jin, S., Wang, Y., & Zhao, X. (2022). Cold-adaptive mechanism of psychrophilic bacteria in food and its application. *Microbial Pathogenesis*, 169, 105652. <https://doi.org/10.1016/j.micpath.2022.105652>
- Kathiriya, M. R., Vekariya, Y. V., & Hati, S. (2023). Understanding the Probiotic Bacterial Responses Against Various Stresses in Food Matrix and Gastrointestinal Tract: A Review. *Probiotics and Antimicrobial Proteins*. <https://doi.org/10.1007/s12602-023-10104-3>
- Kazakos, S., Mantzourani, I., & Plessas, S. (2020). Assessment of Pomegranate Juice as an Alternative “Substrate” for Probiotic Delivery. Recent Advances and Prospects. *Fermentation*, 6(1), Article 1. <https://doi.org/10.3390/fermentation6010024>
- Levante, A., Folli, C., Montanini, B., Ferrari, A., Neviani, E., & Lazzi, C. (2019). Expression of DinJ-YafQ System of *Lactobacillus casei* group strains in response to food processing stresses. *Microorganisms*, 7(10), 438.
- Mbye, M., Baig, M. A., AbuQamar, S. F., El-Tarabily, K. A., Obaid, R. S., Osaili, T. M., Al-Nabulsi, A. A., Turner, M. S., Shah, N. P., & Ayyash, M. M. (2020). Updates on understanding of probiotic lactic acid bacteria responses to environmental stresses and highlights on proteomic analyses. *Comprehensive Reviews in Food Science and Food Safety*, 19(3), 1110–1124. <https://doi.org/10.1111/1541-4337.12554>
- Min, B., Kim, K., Li, V., Cho, S., & Kim, H. (2020). Changes in Cell Membrane Fatty Acid Composition of *Streptococcus thermophilus* in Response to Gradually Increasing Heat Temperature. *Journal of Microbiology and Biotechnology*, 30(5), 739–748. <https://doi.org/10.4014/jmb.1912.12053>
- Moon, S., Ham, S., Jeong, J., Ku, H., Kim, H., & Lee, C. (2023). Temperature Matters: Bacterial Response to Temperature Change. *Journal of Microbiology*, 61(3), 343–357. <https://doi.org/10.1007/s12275-023-00031-x>
- Nguyen, P.-T., Nguyen, T.-T., Bui, D.-C., Hong, P.-T., Hoang, Q.-K., & Nguyen, H.-T. (2020). Exopolysaccharide production by lactic acid bacteria: The manipulation of environmental stresses for industrial applications. *AIMS Microbiology*, 6(4), 451–469. <https://doi.org/10.3934/microbiol.2020027>
- Nunziata, L., Brasca, M., Morandi, S., & Silvetti, T. (2022). Antibiotic resistance in wild and commercial non-enterococcal Lactic Acid Bacteria and Bifidobacteria strains of dairy origin: An update. *Food Microbiology*, 104, 103999. <https://doi.org/10.1016/j.fm.2022.103999>
- Ojo, A. O., & De Smidt, O. (2023). Lactic Acid: A Comprehensive Review of Production to Purification. *Processes*, 11(3), 688. <https://doi.org/10.3390/pr11030688>
- Palud, A., Scornec, H., Cavin, J.-F., & Licandro, H. (2018). New Genes Involved in Mild Stress Response Identified by Transposon Mutagenesis in *Lactobacillus paracasei*. *Frontiers in Microbiology*, 9. <https://www.frontiersin.org/articles/10.3389/fmicb.2018.00535>

- Papadimitriou, K., Alegría, Á., Bron, P. A., de Angelis, M., Gobbetti, M., Kleerebezem, M., Lemos, J. A., Linares, D. M., Ross, P., Stanton, C., Turróni, F., van Sinderen, D., Varmanen, P., Ventura, M., Zúñiga, M., Tsakalidou, E., & Kok, J. (2016). Stress Physiology of Lactic Acid Bacteria. *Microbiology and Molecular Biology Reviews*, 80(3), 837–890. <https://doi.org/10.1128/membr.00076-15>
- Papadopoulou, E., Rodriguez de Evgrafov, M. C., Kalea, A., Tsapekos, P., & Angelidaki, I. (2023). Adaptive laboratory evolution to hypersaline conditions of lactic acid bacteria isolated from seaweed. *New Biotechnology*, 75, 21–30. <https://doi.org/10.1016/j.nbt.2023.03.001>
- Rajesh Banu, J., Merrylin, J., Mohamed Usman, T. M., Yukesh Kannah, R., Gunasekaran, M., Kim, S.-H., & Kumar, G. (2020). Impact of pretreatment on food waste for biohydrogen production: A review. *International Journal of Hydrogen Energy*, 45(36), 18211–18225. <https://doi.org/10.1016/j.ijhydene.2019.09.176>
- Sabater, C., Ruiz, L., Delgado, S., Ruas-Madiedo, P., & Margolles, A. (2020). Valorization of Vegetable Food Waste and By-Products Through Fermentation Processes. *Frontiers in Microbiology*, 11. <https://www.frontiersin.org/articles/10.3389/fmicb.2020.581997>
- Tian, T., Liu, Y., & Wang, X. (2022). Shelf-life extension of chilled beef by sodium lactate enhanced with Natamycin against discoloration and spoilage. *Food Science and Technology*, 42, e30522. <https://doi.org/10.1590/fst.30522>
- Tsakalidou, E., & Papadimitriou, K. (Eds.). (2011). *Stress Responses of Lactic Acid Bacteria*. Springer US. <https://doi.org/10.1007/978-0-387-92771-8>
- Yang, H., He, M., & Wu, C. (2021). Cross protection of lactic acid bacteria during environmental stresses: Stress responses and underlying mechanisms. *LWT*, 144, 111203. <https://doi.org/10.1016/j.lwt.2021.111203>
- Yang, L., Chen, L., Li, H., Deng, Z., & Liu, J. (2022). Lactic acid production from mesophilic and thermophilic fermentation of food waste at different pH. *Journal of Environmental Management*, 304, 114312. <https://doi.org/10.1016/j.jenvman.2021.114312>
- Yang, Y., Wang, R., Yang, Y., E, J., & Wang, J. (2022). Effects of different pre-freezing temperatures on the freeze-drying survival rate and stability during room temperature storage of *Lactiplantibacillus plantarum* LIP-1. *Food Bioscience*, 50, 102087. <https://doi.org/10.1016/j.fbio.2022.102087>
- Yin, W., Wang, Y., Liu, L., & He, J. (2019). Biofilms: The Microbial “Protective Clothing” in Extreme Environments. *International Journal of Molecular Sciences*, 20(14), Article 14. <https://doi.org/10.3390/ijms20143423>
- Zambrano, M. M. (2023). Interplay Between Antimicrobial Resistance and Global Environmental Change. *Annual Review of Genetics*, 57(1), null. <https://doi.org/10.1146/annurev-genet-022123-113904>
- Zhang, W., Guo, H., Cao, C., Li, L., Kwok, L.-Y., Zhang, H., & Sun, Z. (2017). Adaptation of *Lactobacillus casei* Zhang to Gentamycin Involves an Alkaline Shock Protein. *Frontiers in Microbiology*, 8, 2316. <https://doi.org/10.3389/fmicb.2017.02316>
- Zheng, J., Wittouck, S., Salvetti, E., Franz, C. M. A. P., Harris, H. M. B., Mattarelli, P., O’Toole, P. W., Pot, B., Vandamme, P., Walter, J., Watanabe, K., Wuyts, S., Felis, G. E., Gänzle, M. G., & Lebeer, S. (2020). A taxonomic note on the genus *Lactobacillus*: Description of 23 novel genera, emended description of the genus *Lactobacillus* Beijerinck 1901, and union of *Lactobacillaceae* and *Leuconostocaceae*. *International*

Journal of Systematic and Evolutionary Microbiology, 70(4), 2782–2858.
<https://doi.org/10.1099/ijsem.0.004107>

Zhu, Z., Yang, P., Yang, J., & Zhang, J. (2022). Comparative transcriptome analysis reveals the contribution of membrane transporters to acid tolerance in *Lactococcus lactis*. *Journal of Biotechnology*, 357, 9–17. <https://doi.org/10.1016/j.jbiotec.2022.08.006>

WEB REFERENCES

<https://atcc.org/>

<https://bacdive.dsmz.de/>

Chapter 3

Machine Learning approach to explore the microorganisms' biodiversity: bridging experimental datasets and predictive models

Laura Troiani¹, Camilla Lazzi¹, Alessia Levante^{1*}, Vincenzo Bonnici^{2*}, Daniel Puttini², Velia Pierdomenico², Valentina Bernini¹, Erasmo Neviani¹, Alessandro Dal Palù²

¹ Department of Food and Drug Sciences, Parma University, Parma, Italy

² Department of Mathematical, Physical and Computer Sciences, University of Parma, Parma, Italy

Manuscript in preparation

ABSTRACT

Lactic acid bacteria (LAB) represent a diverse group of microorganisms found in various nutrient-rich environments, including food and humans. Their metabolic and physiological flexibility has made them valuable in biotechnology. Phenotypic variability arises from intricate interactions between genotype and environment thus building a “geno-phenotype” network, and the exhaustive understanding of phenotype lags behind the genome characterisation. Microbiology in the age of computational biology has witnessed remarkable advancements in recent years. Machine Learning (ML) is a broad field of advanced computational and statistical methods whose models can help to describe complex data relationships and facilitate the interpretation of large datasets in all sectors, including microbiology. This study aimed to explore the application of ML models to predict the metabolic capabilities of a selected microbial core (SMC) of LAB by exploiting the AFLP genotypic fingerprinting as a feature for computational predictive models.

Two different benchmarks were employed for metabolic activity prediction. The first one is composed of metabolic profiles (GENIII microarrays) and genotypic fingerprints (AFLP) obtained from experiments on 141 strains. The second benchmark was built by extracting publicly available API-50CH test data and *in silico* AFLP profiles related to 509 strains. Random Forest was chosen as ML model and a 5-fold cross-validation strategy was applied to calculate the error/correctness in prediction.

Within the GENIII benchmark, the model worked on 55 substrates out of 71, and the performances in terms of F1-score were directly related to the percentage of metabolically positive strains. Within the API-50CH benchmark, instead, predictions were possible on 28 substrates out of 49. Models relative to the last benchmark yielded good performances for substrates with a balanced distribution of metabolically positive strains. We obtained, on average, an accuracy of 0.81, an F1-score of 0.70 and a ROC-AUC of 0.84.

A combination of genotypic and phenotypic data provides insight into the metabolic capabilities of industrially relevant strains. The exploitation of ML to predict LAB's metabolic capabilities enhances our understanding of these valuable microorganisms and their potential applications in biotechnology. The application of Artificial Intelligence to investigate the biodiversity of microorganisms represents a powerful approach to expand the existing knowledge on bacterial physiology and can set a step in developing novel industrial applications of LAB.

KEYWORDS

Phenotyping, Lactic acid bacteria, Machine Learning, Microbial platform, Targeted fermentation

ABBREVIATIONS

LAB = Lactic Acid Bacteria

ML = Machine Learning

SMC = Selected Microbial Core

AFLP = Amplified Fragment Length Polymorphism

UPCC = University of Parma's Culture Collection

TP = True Positive

TN = True Negative

FP = False Positive

FN = False Negative

TPR = True Positive Rate

FPR = False Positive Rate

1. INTRODUCTION

The microbial group generally referred to as lactic acid bacteria (LAB) consists of a wide range of genera present in various nutrient-rich environments, thanks to their metabolic and physiological flexibility and adaptation. They are associated with food, animals, humans, and health and are largely employed in biotechnology. Their exploitation is hampered by metabolic diversity and confusing identity (Z. Sun et al., 2015; Zheng et al., 2020), thus underlying the complexity of studying a such valuable group of microorganisms.

Phenotypic variability arises from complex interactions between genotype and environment, forming a “geno-phenotype” network that remains inaccessible without comprehensive phenotypic data enabling the study of these interactions. The exhaustive understanding of phenotype lags behind the genome characterisation (Houle et al., 2010). Koch et al. (2020) found a correlation between phenotypic variability and competition in co-culture as well as geographic distribution, thus suggesting connections between intraspecific diversity, microbial interactions, and biogeography. The ecological distribution of microorganisms is strongly influenced by the variability in their genomic and metabolic characteristics which influence adaptive strategies.

Nowadays, revolutionary advances in genome sequencing, phenotyping technologies, and computational sciences are reshaping experimental design. Despite new technologies allowing us to explore the potential metabolic functionality of microbial genomes, the expression of these encoded traits into a metabolic phenotype must be proved experimentally, and experiments are expensive, time-consuming, and manually intensive to perform.

Microbiology in the era of computational biology has witnessed remarkable advances in recent years (Bolívar et al., 2021; Metz et al., 2018; Sharma et al., 2023; Tice & Farag, 2019), enabling researchers to delve deeper into the intricate world of microorganisms. This marriage between microbiology and computational techniques has paved the way for innovative approaches to address fundamental issues in the field. Within this context, Machine Learning (ML) assumes a crucial role in data management, mining and analysis, offering valuable insights for decision-making and facilitating the achievement of specific goals (Esposito et al., 2019; Steele et al., 2019). ML is a broad field of advanced computational and statistical methods whose models can help to describe complex data relationships and facilitate the interpretation of large sets of data in all fields, including microbiology (Bezbaruah et al., 2023; Özel Duygan & Van Der Meer, 2022), thus making easier the selection process and the design of targeted experiments to carry out.

The AFLP is a well-known PCR-based technique that generates genomic fingerprints (Ramadan, 2022). AFLP stands as a valuable technique for classifying bacteria at the species and strain level, demonstrating good discriminatory power and reliable reproducibility. Despite the outcome is not comparable to whole genome sequencing, AFLP fingerprinting offers several advantages over other techniques: a prior microorganism's genomic information is not required, it is possible to label AFLP alleles by fluorescence thus allowing simultaneous characterisation of several samples, it is a rapid technique to obtain accurate information and efficiently discriminate species and strains, at lower costs than genome sequencing and with outcomes easier to analyse. AFLP profiles serve as a fingerprint of the genetic makeup of the organism, capturing variations that can be attributed to evolutionary divergence and adaptation to different environments (M. Sun et al., 2017). The AFLP implementation led to important progress in the study of biodiversity and the taxonomy of bacteria (Bertani et al., 2019). The effectiveness of this technique is largely shown by wet lab analyses, but no computational model is currently available to harness its power.

Regarding phenotyping, different commercial kits allow the study of bacterial substrate utilisation profiles, which are cheaper, faster and easier to use than culture-dependent approaches (Mekonnen et al., 2019; Tsigkrmani et al., 2022). For instance, Biolog GENIII Microplates (BIOLOG Inc.©, USA) provide the parallel investigation of 71 different carbon sources assimilation, evaluating the metabolic aptitude of microorganisms (Khalifa, 2022; Luthfi et al., 2023; Mariz-Ponte et al., 2022). Another useful method to easily and simultaneously screen bacteria for different metabolic capabilities is the API-50CH test which allows the fermentation of 49 carbohydrates (Fossi et al., 2022; Illikoud et al., 2023).

This work focused on the application of Machine Learning models to manage biological data and predict the metabolic capabilities of a selected microbial core (SMC) of LAB, by using as input information derived solely from the AFLP fingerprints, and metabolic profiles both obtained from laboratory experiments and collected from publicly available database.

In the domain of predictive modelling, the choice of the appropriate model can significantly impact the quality and reliability of results. For our study, we have selected Random Forest because it offers several compelling advantages. Random Forest has garnered recognition for its robustness against overfitting, making it a reliable option even in the presence of noisy or limited datasets. Moreover, its remarkable ability to handle high-dimensional data, assess variable importance, and handle missing values simplifies the modelling process (Breiman, 2001).

2. MATERIALS AND METHODS

Two different benchmarks were employed in the predictive study.

2.1. Benchmark on own dataset

The first benchmark is composed of data related to 141 UPCC strains obtained from GENIII microassays and AFLP performed in the laboratory as described below.

2.1.1. Selection and cultivation of LAB strains

The University of Parma's Culture Collection (UPCC) was represented by a microbial core (SMC) of 141 LAB strains of food origin. The strains' selection criteria aimed to guarantee a great level of biodiversity: strains belong to 29 different species, and they were isolated over more than 10 years (2012-2022) from several food matrices, especially of dairy origin.

Cultures were stored as frozen stocks (-80°C) supplemented with 15% glycerol (w/v). LAB were propagated three times in fresh medium (inoculum = 3% v/v) and incubated for 16 h at their optimal temperature in anaerobic conditions.

2.1.2. Genotypic analysis

Strains of the SMC were genotypically characterised by Amplified Fragment Length Polymorphism (AFLP), a fingerprinting technique that provides resolution to the subspecies level. After DNA extraction by using the "Pure Food Pathogen" Kit (Maxwell® RSC Instrument, Promega, Milan, Italy), the AFLP analysis was performed. The composition of the Restriction-Ligation (RL) reaction is reported in Table 1.

Table 1 Restriction-Ligation.

Reagent	Volume (μl)
250 ng/ μl BSA ^a	10
1 $\times$ T4 DNA ligase buffer with 1 μl 1 mM ATP ^a	5
50 pmol/ μl MseI adaptor ^b	1
5 pmol/ μl EcoRI adaptor ^b	1
10 mM ATP ^c	1
10 U/ μl MseI ^a	0.5
20 U/ μl EcoRI ^a	0.25
200U/ μl of T4 DNA ligase ^a	0.1
DNA (500 ng)	15
Total volume	50

a) Reagents purchased from New England Biolabs (Ipswich, MA, USA)

b) Reagents purchased from BMR Genomics (Padova, Italy)

- c) Reagent purchased from Invitrogen s.r.l. (Milano, Italy)

The reaction was incubated at 37°C for 4 hours, and then two sequential amplifications were executed: a pre-selective PCR followed by a selective one (Tab. 2).

Table 2 Preselective and Selective PCR amplification.

Reagent	Volume (μl)
<i>I) Pre-selective PCR</i>	
GoTaq® Master Mix 2X ^a	10
10 μM EcoRI-0 Primer ^b (Seq. 5'-3' GACTGCGTACCAATTC (labelled FAM 5'))	0.6
10 μM MseI-0 Primer ^b (Seq. 5'-3' GATGAGTCCTGAGTAA)	0.6
digested-ligated DNA	2
Total volume	20
<i>II) Selective PCR</i>	
GoTaq® Master Mix 2X ^a	10
10 μM EcoRI-A Primer ^b (Seq. 5'-3' GACTGCGTACCAATTCA (labelled FAM 5'))	0.6
10 μM MseI-C Primer ^b (Seq. 5'-3' GATGAGTCCTGAGTAAC) or MseI-A Primer ^b (Seq. 5'-3' GATGAGTCCTGAGTAAA)	0.6
DNA diluted 1:10 in TE (10 mM TRIS HCl at pH=8, 0.1 mM EDTA at pH=8)	2
Total volume	20

- a) Reagents purchased from Promega (Madison, WI, USA)

- b) Reagents purchased from BMR Genomics (Padova, Italy)

All PCRs were performed in a GeneAmp® PCR System 2700 (Applied Biosystem, Foster City, CA, USA) by following the thermocycler program of Bertani et al. (2019). The amplifications were confirmed by electrophoresis on a 1.2% agarose gel in 1×TAE buffer at 90 V/cm for 20 min and visualized by UV transillumination after staining with RedSafe DNA Gel Stain (Invitrogen Life Technologies, Carlsbad, California, USA). An external service (BMR Genomics, Padova, Italy) analysed the AFLP products by capillary electrophoresis.

2.1.3. Phenotypic analysis

The metabolic capabilities of strains were tested against 71 different carbon sources (including sugars, amino acids, carboxylic acids, etc.) by using GENIII Microplates (BIOLOG Inc.®, USA). In each well of the microplate there is the dried substrate to be tested together with a patented Tetrazolium Violet dye used as a colorimetric indicator of microbial activity.

Overnight cultures were washed twice to remove any medium contamination and resuspended in 12 ml of inoculation fluid IF-A (BIOLOG Inc.©, USA) to a transmittance of 90-98%. The microplates were incubated at the optimal temperature until 72 h when the metabolic activity was detected on the BIOLOG MicroReader Station (BIOLOG Inc.©, USA) by absorbance reading at 590 nm. Raw data were referred against the negative control well (A1).

2.2. Benchmark of publicly available data

The second benchmark was built by extracting publicly available API-50CH test data from the Bacterial Diversity Metadatabase (BacDive) (<https://bacdive.dsmz.de>). We selected 509 strains belonging to several species by considering their positive metabolic activity on at least one of the 49 carbon sources in the API-50CH strip. Subsequently, in silico AFLP profiles were extracted through the AFLPInSilico software (Kobilhova et al. 2018). Such a tool provides a computational approach for simulating the run of the AFLP protocol on a given DNA sequence.

2.3. Computational modelling

We tested different recognized supervised Machine Learning (ML) models: XGBClassifier, Decision tree, Random Forest, Extra Tree, AdaBoots, Gaussian Naive Classifier, k-neighbours, MLP Classifier and SVM. However, preliminary results (not shown here) highlighted Random Forest as the best approach to explore the AFLP as feature within computational predictive models. The computational methodology used in this study has been implemented on top of the scikit-learn Python library (<https://scikit-learn.org>) which provides already developed ML tools.

In our approach, each strain is represented by means of its AFLP profile. For every strain, the AFLP process generates a list of fragments. A distribution of fragment lengths was built by counting the number of fragments having the same length. We considered only fragments having lengths from 50 to 500bp, both included. For both benchmarks, AFLP profiles were subjected to a multi-value discretization. Fragment lengths were represented by 4 values which are centred at the mean of the AFLP distribution. Given an AFLP distribution, with m the mean value of such a distribution, we mapped AFLP values lower than m to 1 and AFLP values greater or equal to m to 2. Furthermore, we set to 0 the original 0 values.

For a given metabolic activity, we trained a ML model by considering the AFLP profiles as input features of the model, and by setting a binary label classification based on the ability of a specific strain to either employ or not a substrate. To ensure the reliability and generalized extent of the predictive model, we employed the 5-folding evaluation procedure for evaluating

the performance of the model. This strategy consists of partitioning the dataset into five equally sized subsets or "folds." During each iteration, the model is trained on four of these folds and validated on the remaining fifth one. This process is repeated five times, with each fold serving as the validation set at once. The use of cross-validation helps both to guard against overfitting by assessing the model performance on multiple subsets and aid in fine-tuning the model's hyperparameters to identify the optimal configuration. This evaluation aims to calculate the error/correctness of the trained model in predicting the metabolic activity of strains in the validation set, without any information about their activity but only based on their AFLP profile. For each run, different performance measures were computed by averaging the values obtained for the 5 different runs. Furthermore, the 5-folding approach was also applied to achieve optimal parameterization. This involved the exploration of parameter combinations such as 'n_estimators' and 'max_depth', individually tailored to each distinct metabolic activity.

To estimate the effectiveness of the classifier, several measures were calculated as reported in Table 3. For every measure, the higher the value of the measure, the better the performance of the model. A score of 0.5 is usually considered as the result of a model which randomly predicts the outcome. Thus, values greater than 0.6 are usually considered good performance scores (see Russel & Norvig (2010) for further details about such performance measures).

A True Positive (TP) occurs when our model correctly identifies whether a strain can metabolise the substrate associated with the model. Symmetrically, a True Negative (TN) occurs when the model correctly identifies whether a strain does not metabolise the substrate associated with the model. A False Positive (FP) occurs when a strain does not metabolise the substrate, but the model predicts an opposite behaviour. A False Negative (FN) occurs when a strain can metabolise a given substrate, but the model predicts a non-utilisation of that substrate.

Table 3 Evaluation metrics to appraise the classification results of the model. TP = True Positive; TN = True Negative; FP = False Positive; FN = False Negative; TPR = True Positive Rate; FPR = False Positive Rate.

Metrics	Formula	Definition
Accuracy	$\frac{TP + TN}{TP + FN + TN + FP}$	Number of correct predictions with respect to the total number of predictions
Precision	$\frac{TP}{TP + FP}$	Number of correct positive predictions with respect to the total number of positive predictions
Recall	$\frac{TP}{TP + FN}$	Number of correct positive predictions with respect to the number of strains that effectively use a given substrate

F1-score	$\frac{2 * precision * recall}{precision + recall}$	Harmonic mean of precision and recall, meaning the ability of a model to find all the relevant cases (positive metabolic activity on a substrate) within a data set
ROC-AUC	$\int_0^1 TPR(FPR) dFPR$ with $TPR=TP/(TP+FP)$ and $FPR=FP/(TN+FN)$	Robustness of the model in correctly predicting the metabolically positive strains on varying an arbitrary threshold within the Random Forest model to decide whenever a strain uses or not a given substrate

Moreover, feature importance techniques were also applied to produce a selection of fragment lengths associated with the predictive power within a given group of genomes and/or for a specific metabolic activity. On top of this selection, the idea is to perform functional enrichment of the selected fragments by comparing their position within the genome with the coordinates of genes for which functional annotation is already available.

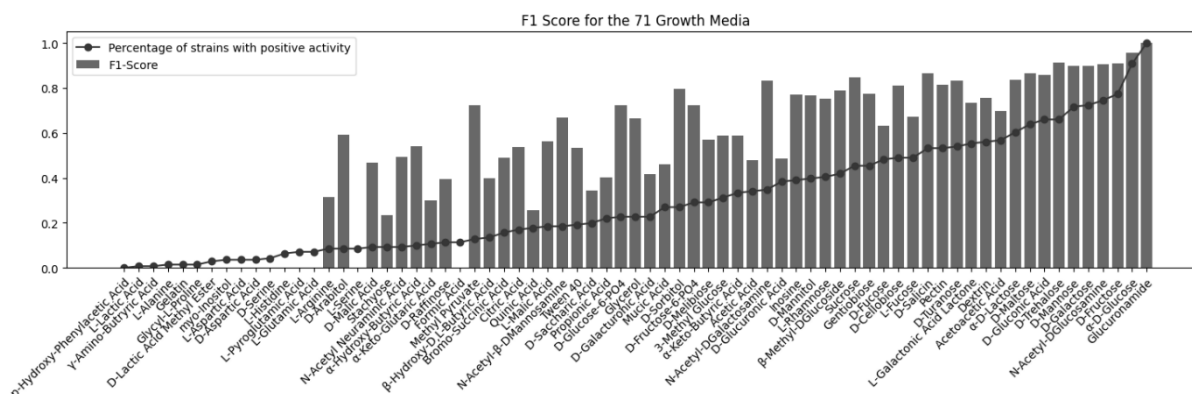
3. RESULTS

3.1. GENIII metabolic activity prediction

In pursuit of the research objective, the power of the well-established Random Forest Machine Learning algorithm was harnessed.

Within the GENIII benchmark, the model worked on 55 substrates out of 71, which exhibited remarkable heterogeneity in terms of utilisation by bacteria. This information is shown in Figure 1 where GENIII substrates were ordered by the percentage of strains that exhibit positive activity on a given substrate. p-hydroxy-phenylacetic acid and glucuronamide are two exceptional cases because we recorded 0% and 100% of positive strains. In these two cases, we were not able to train a computational model because at least one positive or negative element must be present in the input data set to perform the analysis. The performance of the trained model is directly related to the percentage of positive strains. Indeed, on average, the higher the percentage, the better the F1-score. However, there are some exceptions: the trained models for two substrates were not able to predict the metabolic activity. On the contrary, there are substrates with a few positive strains for which an F1-score greater than 0.6 was reached.

Figure 1 Percentage of UPCC strains (dots) with positive metabolic activity on a given substrate of the GENIII Microplate and F1-score (bars) of the respective computational model trained on that specific substrate.

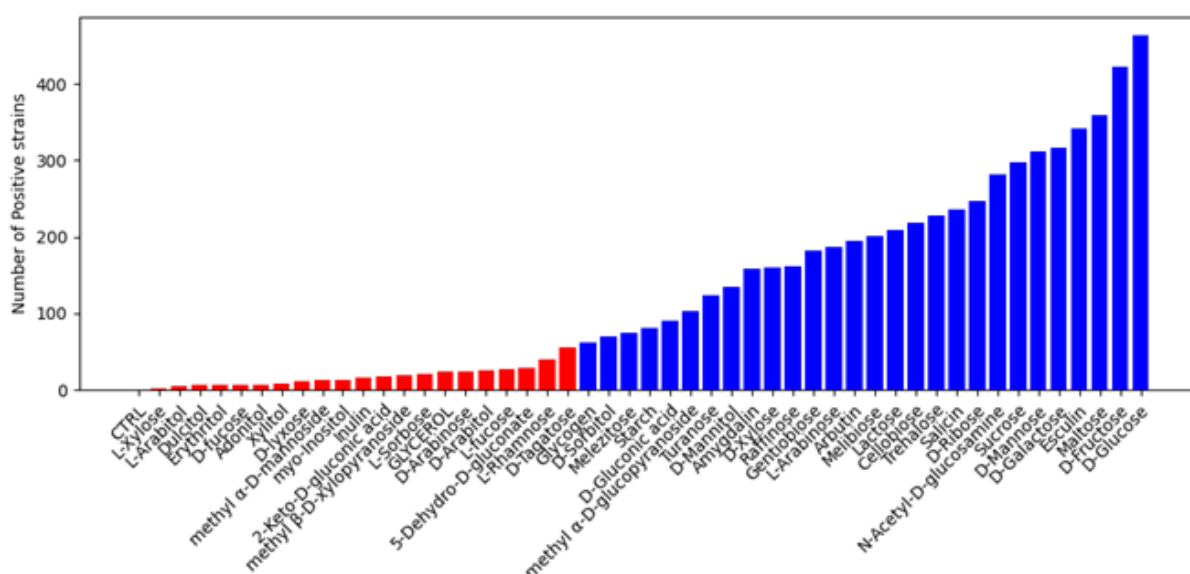


3.2. API-50CH metabolic activity prediction

The discretized AFLP profiles led to the construction of a data matrix representing the patterns of all bacteria. The distinguishing factor among experiments lies in the binary target vector associated with a specific metabolic activity. Each target vector signifies whether a given strain is engaged in the metabolic activity under consideration. Notably, within the scope of this

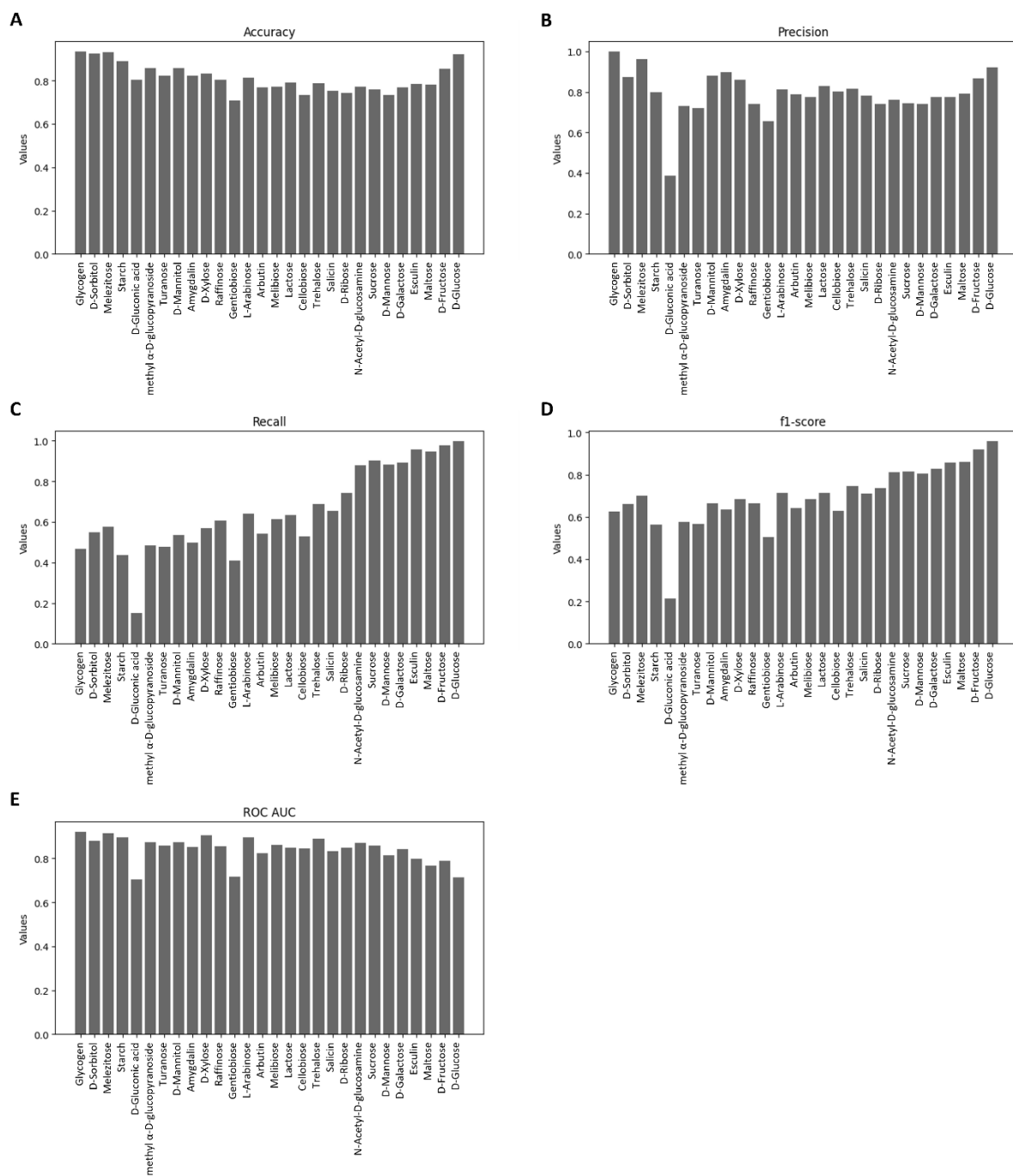
binary classification problem, 28 of the 49 API-50CH test substrates were selected for the model training (Fig. 2 reports the number of positive strains for each substrate). The selected substrates are those on which more than 10% (55 strains) of the total number of strains considered in this study showed a positive metabolic activity, thus solving imbalances in the metabolic activity representation.

Figure 2 Number of strains with positive metabolic activity on each substrate, displayed in ascending order. The threshold of the number of strains to use a specific substrate for model training was set at 10% of the total. Substrates on which less than 10% of the total strains showed metabolic activity were not considered for analysis (red bars).



The selected substrates (Fig. 2, blue bars), in ascending order by number of positive strains, are Glycogen, D-Sorbitol, Melezitose, Starch, D-Gluconic acid, methyl α -D-glucopyranoside, Turanose, D-Mannitol, Amygdalin, D-Xylose, Raffinose, Gentiobiose, L-Arabinose, Arbutin, Melibiose, Lactose, Cellobiose, Trehalose, Salicin, D-Ribose, N-Acetyl-D-glucosamine, Sucrose, D-Mannose, D-Galactose, Esculin, Maltose, D-Fructose, D-Glucose. These were the most frequently consumed C sources thus reflecting the results obtained from the GENIII microarrays.

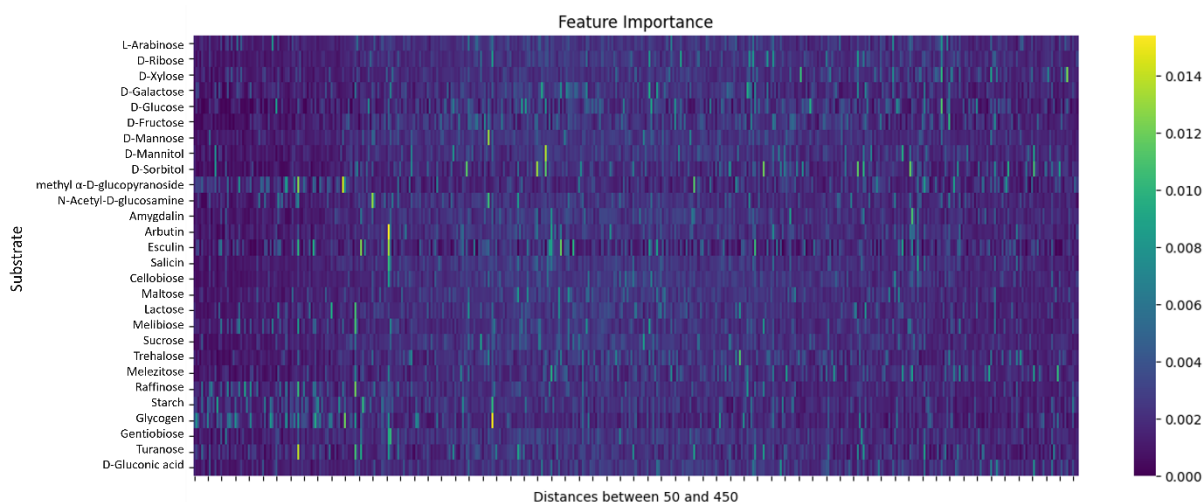
Figure 3 Evaluation metrics for the models trained on the 28 selected substrates: A) accuracy; B) prediction; C) recall; D) F1-score; E) ROC-AUC.



The average accuracy was 0.81, with the highest value (0.93) for both Glycogen and Melezitose, and the lowest (0.71) for Gentioibiose. The F1-score, which considers both precision and recall of the model, had an average of 0.70: the highest value was reached by the model predicting the metabolic activity on D-Glucose (0.96) and the lowest value on D-Gluconic acid (0.21). The ROC-AUC metric indicates how far the model is from random predictions. In this case, the average was 0.84, recording the highest value (0.92) for both Glycogen and Melezitose, and the

lowest for D-Gluconic acid (0.70). Further analysis to determine the optimal parameterization for each model, led to the selection of 45 features with the highest 'feature_importance' values, as derived from the Random Forest model (Fig. 4). The analysis showed that the set of AFLP fragment lengths that are important for predicting the metabolic activity related to a given substrates varies among the substrates.

Figure 4 Feature importance of each fragment length composing the AFLP profiles.



4. DISCUSSIONS

The central objective of this study was to harness the power of ML to predict whether a particular strain possesses the capacity to use a specific substrate, relying solely on the information obtainable through the AFLP technique.

Amplified Fragment Length Polymorphism is used for the selective amplification of a subset of digested DNA fragments to generate genomic fingerprints that can be exploited in genomic, transcriptomic and epigenetic studies of plants (Hong et al., 2021), and for investigation of microbial diversity (Bertani et al., 2019). Given two restriction enzymes, this technology aims to amplify the DNA fragments enclosed by their restriction sites and retrieve their length. Mathematically, the signal produced by the AFLP technology corresponds to a recurrence distance distribution which gives the number of times two specific k-mers recur at a given distance within a genomic string (Bonnici & Manca, 2015). It has been used for detecting CpG islands (Hackenberg et al., 2006) and for extracting keywords in natural languages and DNA strings (Astorino et al., 2023; Ortuño et al., 2002).

The 71 binary classifiers related to the GENIII benchmark were designed to decode the intricate relationship between AFLP and metabolic profiles: the revealed biodiversity underscores the microbial ecosystems' complexity and highlights the need for a predictive model to capture the most subtle aspects.

The choice of the F1-score (Fig. 1) for the preliminary evaluation of the model performances stems from its suitability in the case of unbalanced datasets, where the utilisation of specific substrates is significantly more prevalent than that of others. Results suggest that the achievement of robust predictive performance is contingent on the prevalence of positive instances (metabolically positive strains) within the binary target vectors (substrates). Accordingly, the presence of a sufficient number of positives in the target vectors facilitated the model learning process thus leading to promising performances. Moreover, the classification is made even harder by the relatively small size of the dataset. However, for those substrates with a balanced instances distribution, where approximately half of the strains have grown, our models yielded moderate performances (e.g. Fig. 1, D-Fucose).

To gain deeper insight into the complexity of microbial behaviour prediction and to explore the potential of our approach in handling unbalanced datasets more effectively, the analysis pipeline was further implemented on a larger dataset from the publicly available BacDive repository. This new dataset introduced additional complexities, thus providing a richer landscape for the examination of our predictive models' capabilities.

The pattern generated by counting the fragments that fall within lengths ranging from 50 to 500bp succinctly captures the distribution of fragment counts across different length intervals. It becomes a distinctive descriptor, thus portraying the unique AFLP profile of each strain within our computational framework. It is noteworthy that our approach exclusively relies on the information embedded within the counts of fragment lengths rather than the individual identities of the fragments. Accordingly, in our analysis, two fragments of equal length convey equivalent information, even if they differ in their composition. This strategy underscores our focus on the distribution of fragment length counts as the primary basis for bacterial representation.

The rationale behind our choice of discretization lies in the variability of absolute count values obtained from the AFLP fragment lengths, which inherently depend on the quantity of DNA present in the AFLP reaction. It is reasonable to hypothesize that, in general, bacteria with longer DNA would yield AFLP profiles comprising numerous fragments, resulting in higher count values for their associated fragment lengths. To accommodate bacterial species characterised by diverse DNA lengths, we employed the discretization approach. However, we intended to strike a balance between capturing this variability and avoiding an overly simplistic representation. Unlike a binary transformation, which would result in a stark division (values equal to 0 remain unchanged, while values greater than 0 become 1), our chosen discretization method ensures a nuanced portrayal of the underlying distribution while accommodating variations associated with DNA length. Disregarding this imbalance can significantly impact the interpretation of our data and the reliability of our findings. When metabolic activities exhibit substantial disparities in the number of positive (indicating engagement in the activity) and negative (indicating non-engagement) instances, ML models tend to favour the majority class. In the context of microbial metabolic studies, this could lead to an overemphasis on common activities at the expense of rarer ones. Consequently, the predictive models might perform well for prevalent metabolic activities but struggle to accurately identify and characterise the less common ones.

The 28 selected substrates (Fig. 2, blue bars) were used to train each respective model. This skewed focus on the majority classes can lead to several detrimental outcomes. Firstly, it can result in a high number of false negatives, where the model incorrectly predicts the absence of metabolic activities that are instead present but are less frequent in the dataset. This misclassification can obscure the true metabolic landscape, potentially causing researchers to overlook biologically significant activities. Conversely, the model might exhibit high precision for the majority class, leading to a high number of false positives for minority classes. This

scenario can inflate the apparent prevalence of rare metabolic activities, leading to erroneous conclusions regarding their significance in microbial processes. In essence, the careful management of dataset imbalance ensures a more balanced perspective on the various metabolic activities under investigation. It allows us to uncover both common and rare activities with greater accuracy, leading to a more nuanced and biologically relevant interpretation of microbial metabolic data. By addressing these imbalances, we not only improve the performance of our predictive models but also enhance the overall quality and validity of our scientific insights (He & Garcia, 2009).

Accuracy is a fundamental metric in the evaluation of binary classifiers, quantifying the proportion of correctly classified instances out of the total predictions. It provides a straightforward measure of the overall performance of a model, making it widely used in ML assessments. However, caution must be exercised when dealing with imbalanced datasets, where one class significantly outnumbers the other. In such a scenario, a classifier may achieve a deceptively high accuracy by predominantly predicting the majority class. Therefore, accuracy alone may not provide an accurate representation of a classifier's effectiveness, necessitating the exploration of alternative metrics like precision, recall and F1-score, which consider false positives and false negatives. A comprehensive evaluation, considering both accuracy and these alternative metrics, is crucial for assessing the real capabilities of a binary classifier (Koyejo et al., 2014).

Since the proportion of positive entities increases from 10% to 90%, we encounter a classic example of the 'U-shaped curve' phenomenon (Fig. 3A). Initially, with a low percentage of positives, the classifier's task is relatively straightforward, resulting in high accuracy. However, as the proportion of positives grows, the classifier faces a more challenging task in distinguishing the classes, leading to a temporary decrease in accuracy around the 30% to 70% range. As the class imbalance continues to decrease beyond 70%, the classifier's performance improves, ultimately reaching its peak. This U-shaped behaviour highlights the interplay between class distribution and classifier performance, highlighting the importance of considering the proportion of positive samples when evaluating the effectiveness of our model. Hence, it is crucial to complement accuracy with other evaluation metrics for a more comprehensive assessment.

The observed variation in precision values (Fig. 3B) closely aligns with the patterns we noticed in accuracy, by further underscoring the impact of class distribution on our model's performance. High precision values are evident at both the extreme ends of the positive entity

proportion spectrum, while a notable dip occurs in the middle range where the class balance is closer.

We assessed our model's performance using the F1-score, a metric that balances precision and recall in a binary classification context. The F1-score graph (Fig. 2C) provides a succinct overview of our model's accuracy for each metabolic activity, accounting for both true positives and true negatives, and offering insight into its discriminatory capabilities within the intricate biological landscape (Bishop & Nasrabadi, 2006).

The observed ROC-AUC values shed light on the model's ability to discriminate classes under varying proportions of positive entities within the target variable (Müller & Guido, 2016). From Figure 3D, it is evident that the model exhibits a commendable performance, with ROC curve values consistently exceeding 0.8 across most scenarios. Notably, as the proportion of positive entities reaches its zenith at 90%, the ROC-AUC values remain generally high, albeit with a slight, gradual decline towards 0.71 in the final instances. This trend suggests that the model excels in distinguishing classes when the class imbalance is at its peak. However, the decrease in ROC-AUC in these latter cases indicates that, while the model retains its discriminative power, it faces a nuanced challenge in these circumstances. Overall, the ROC-AUC values, when considered together with other metrics such as accuracy, precision, recall and F1-score, collectively reflect a solid performance by our model, especially when class imbalance is at its most extreme.

5. CONCLUSIONS

In conclusion, the achieved results support the adequacy of our model's performance. The binary classifiers, each tuned to a specific substrate, are designed to decode the intricate interplay between bacterial genetics and environmental factors, providing valuable insights into the adaptive strategies of microorganisms. The varying F1-scores across different metabolic activities highlight the versatility of our predictive approach. Notably, the slight alignment between F1-scores and the count of metabolically positive strains can be attributed to the small scale of the dataset, encompassing just over 500 strains. This insight underscores how dataset size impacts predictive accuracy.

The predictive approach allows us to design ad-hoc experimentation and have a preliminary idea of the feasibility and microbial performance to avoid waste of material, labour, and time in the laboratory.

A combination of genotypic and phenotypic data provides an insight into the metabolic capabilities of industrially relevant strains, disclosing their potential application for the biotransformation of various substrates, including wastes and by-products of food origin. The application of Artificial Intelligence to investigate the biodiversity of microorganisms represents a powerful approach to expand the existing knowledge on bacterial physiology and can set a step in developing novel industrial applications of LAB.

REFERENCES

- Astorino, S., Bonnici, V., & Franco, G. (2023). An Investigation to Test Spectral Segments as Bacterial Biomarkers. In D. Genova & J. Kari (Eds.), *Unconventional Computation and Natural Computation* (pp. 1–16). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-34034-5_1
- Bertani, G., Savo Sardaro, M. L., Neviani, E., & Lazzi, C. (2019). AFLP protocol comparison for microbial diversity fingerprinting. *Journal of Applied Genetics*, 60(2), 217–223. <https://doi.org/10.1007/s13353-019-00492-0>
- Bezbaruah, R., Ghosh, M., Kumari, S., Nongrang, L., Ali, S. R., Lahiri, M., Waris, H., & Kakoti, B. B. (2023). Role of AI and ML in Epidemics and Pandemics. In *Bioinformatics Tools for Pharmaceutical Drug Product Development* (pp. 345–369). Scopus. <https://doi.org/10.1002/9781119865728.ch15>
- Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern Recognition and Machine Learning* (1st ed., Vol. 4). New York: springer. <https://link.springer.com/book/9780387310732>
- Bolívar, A., Tarlak, F., Costa, J. C. C. P., Cejudo-Gómez, M., Bover-Cid, S., Zurera, G., & Pérez-Rodríguez, F. (2021). A new expanded modelling approach for investigating the bioprotective capacity of *Latilactobacillus sakei* CTC494 against *Listeria monocytogenes* in ready-to-eat fish products. *Food Research International*, 147, 110545. <https://doi.org/10.1016/j.foodres.2021.110545>
- Bonnici, V., & Manca, V. (2015). Recurrence Distance Distributions in Computational Genomics. *American Journal of Bioinformatics and Computational Biology*, 3. <https://doi.org/10.7726/ajbcb.2015.1002>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Esposito, S., Carputo, D., Cardi, T., & Tripodi, P. (2019). Applications and Trends of Machine Learning in Genomics and Phenomics for Next-Generation Breeding. *Plants*, 9(1), 34. <https://doi.org/10.3390/plants9010034>
- Fossi, B. T., Ekabe, D. E., Toukam, L. L., Tatsilong Pambou, H. O., Gagneux-Brunon, A., Nkenfou Nguéfeu, C., & Bongue, B. (2022). Probiotic lactic acid bacteria isolated from traditional cameroonian palm wine and corn beer exhibiting cholesterol lowering activity. *Heliyon*, 8(11), e11708. <https://doi.org/10.1016/j.heliyon.2022.e11708>
- Hackenberg, M., Previti, C., Luque-Escamilla, P. L., Carpena, P., Martínez-Aroza, J., & Oliver, J. L. (2006). CpGcluster: A distance-based algorithm for CpG-island detection. *BMC Bioinformatics*, 7(1), 446. <https://doi.org/10.1186/1471-2105-7-446>
- He, H., & Garcia, E. A. (2009). Learning from Imbalanced Data. *IEEE Transactions on Knowledge and Data Engineering*, 21(9), 1263–1284. <https://doi.org/10.1109/TKDE.2008.239>
- Hong, N., Chen, M., & Xu, J. (2021). Molecular Markers Reveal Epidemiological Patterns and Evolutionary Histories of the Human Pathogenic *Cryptococcus*. *Frontiers in Cellular and Infection Microbiology*, 11. <https://www.frontiersin.org/articles/10.3389/fcimb.2021.683670>
- Houle, D., Govindaraju, D. R., & Omholt, S. (2010). Phenomics: The next challenge. *Nature Reviews Genetics*, 11(12), 855–866. <https://doi.org/10.1038/nrg2897>

- Illikoud, N., Do Carmo, F. L. R., Daniel, N., Jan, G., & Gagnaire, V. (2023). Development of innovative fermented products by exploiting the diversity of immunomodulatory properties and fermentative activity of lactic and propionic acid bacteria. *Food Research International*, 166, 112557. <https://doi.org/10.1016/j.foodres.2023.112557>
- Khalifa, A. (2022). First Isolation and Characterization of *Serratia liquefaciens* Associated with Rot Disease of *Malus domestica* (Apple) Fruit and Its Inhibition by *Origanum vulgare* (Oregano) Oil. *Horticulturae*, 8(8), 752. <https://doi.org/10.3390/horticulturae8080752>
- Koch, H., Gernscheid, N., Freese, H. M., Noriega-Ortega, B., Lücking, D., Berger, M., Qiu, G., Marzinelli, E. M., Campbell, A. H., Steinberg, P. D., Overmann, J., Dittmar, T., Simon, M., & Wietz, M. (2020). Genomic, metabolic and phenotypic variability shapes ecological differentiation and intraspecies interactions of *Alteromonas macleodii*. *Scientific Reports*, 10(1), 809. <https://doi.org/10.1038/s41598-020-57526-5>
- Koyejo, O. O., Natarajan, N., Ravikumar, P. K., & Dhillon, I. S. (2014). Consistent Binary Classification with Generalized Performance Metrics. *Advances in Neural Information Processing Systems*, 27. https://proceedings.neurips.cc/paper_files/paper/2014/hash/30c8e1ca872524fbf7ea5c519ca397ee-Abstract.html
- Luthfi, A. A. I., Abdul, P. M., Jahim, J. M., Engliman, N. S., Jamali, N. S., Tan, J. P., Manaf, S. F. A., Sajab, M. S., & Bukhari, N. A. (2023). Isolation and Characterization of Biohydrogen-Producing Bacteria for Biohydrogen Fermentation Using Oil Palm Biomass-Based Carbon Source. *Applied Sciences*, 13(1), 656. <https://doi.org/10.3390/app13010656>
- Mariz-Ponte, N., Gimranov, E., Rego, R., Moura, L., Santos, C., & Tavares, F. (2022). Distinct phenotypic behaviours within a clonal population of *Pseudomonas syringae* pv. *Actinidiae*. *PLOS ONE*, 17(6), e0269343. <https://doi.org/10.1371/journal.pone.0269343>
- Mekonnen, E., Kebede, A., Tafesse, T., & Tafesse, M. (2019). Investigation of carbon substrate utilization patterns of three ureolytic bacteria. *Biocatalysis and Agricultural Biotechnology*, 22, 101429. <https://doi.org/10.1016/j.bcab.2019.101429>
- Metz, Z. P., Ding, T., & Baumler, D. J. (2018). Using genome-scale metabolic models to compare serovars of the foodborne pathogen *Listeria monocytogenes*. *PLoS ONE*, 13(6). Scopus. <https://doi.org/10.1371/journal.pone.0198584>
- Müller, A. C., & Guido, S. (2016). *Introduction to Machine Learning with Python: A Guide for Data Scientists*. O'Reilly Media, Inc.
- Naidu, G., Zuva, T., & Sibanda, E. M. (2023). A Review of Evaluation Metrics in Machine Learning Algorithms. In R. Silhavy & P. Silhavy (Eds.), *Artificial Intelligence Application in Networks and Systems* (pp. 15–25). Springer International Publishing. https://doi.org/10.1007/978-3-031-35314-7_2
- Ortuño, M., Carpena, P., Bernaola-Galván, P., Muñoz, E., & Somoza, A. M. (2002). Keyword detection in natural languages and DNA. *Europhysics Letters*, 57(5), 759. <https://doi.org/10.1209/epl/i2002-00528-3>
- Özel Duygan, B. D., & Van Der Meer, J. R. (2022). Recent advances in microbial community analysis from machine learning of multiparametric flow cytometry data. *Current Opinion in Biotechnology*, 75, 102688. <https://doi.org/10.1016/j.copbio.2022.102688>

- Russel, S. J., & Norvig, P. (2010). *Artificial Intelligence: A Modern Approach*. <https://www.pearson.com/en-us/subject-catalog/p/artificial-intelligence-a-modern-approach/P2000000003500/9780137505135>
- Sharma, P., Dahiya, S., Kaur, P., & Kapil, A. (2023). Computational biology: Role and scope in taming antimicrobial resistance. *Indian Journal of Medical Microbiology*, 41, 33–38. Scopus. <https://doi.org/10.1016/j.ijmmb.2022.12.005>
- Steele, A. D., Tejjaro, C. N., Yang, D., & Shen, B. (2019). Leveraging a large microbial strain collection for natural product discovery. *Journal of Biological Chemistry*, 294(45), 16567–16576. <https://doi.org/10.1074/jbc.REV119.006514>
- Sun, M., Zhang, C., Zhang, X., Fan, Y., Fu, K., Wu, W., Bai, S., Zhang, J., Peng, Y., Huang, L., Yan, Y., & Ma, X. (2017). AFLP assessment of genetic variability and relationships in an Asian wild germplasm collection of *Dactylis glomerata* L. *Comptes Rendus - Biologies*, 340(3), 145–155. Scopus. <https://doi.org/10.1016/j.crv.2016.12.003>
- Sun, Z., Harris, H. M. B., McCann, A., Guo, C., Argimón, S., Zhang, W., Yang, X., Jeffery, I. B., Cooney, J. C., Kagawa, T. F., Liu, W., Song, Y., Salvetti, E., Wrobel, A., Rasinkangas, P., Parkhill, J., Rea, M. C., O’Sullivan, O., Ritari, J., ... O’Toole, P. W. (2015). Expanding the biotechnology potential of lactobacilli through comparative genomics of 213 strains and associated genera. *Nature Communications*, 6(1), 8322. <https://doi.org/10.1038/ncomms9322>
- Tice, A. M., & Farag, H. A. (2019). Machine Learning in Microbiology: Finding the Signal in the Noise. *Clinical Microbiology Newsletter*, 41(14), 121–127. Scopus. <https://doi.org/10.1016/j.clinmicnews.2019.06.004>
- Tsigkrimini, M., Bakogianni, M., Paramithiotis, S., Bosnea, L., Pappa, E., Drosinos, E. H., Skandamis, P. N., & Mataragas, M. (2022). Microbial Ecology of Artisanal Feta and Kefalograviera Cheeses, Part I: Bacterial Community and Its Functional Characteristics with Focus on Lactic Acid Bacteria as Determined by Culture-Dependent Methods and Phenotype Microarrays. *Microorganisms*, 10(1), 161. <https://doi.org/10.3390/microorganisms10010161>
- Zheng, J., Wittouck, S., Salvetti, E., Franz, C. M. A. P., Harris, H. M. B., Mattarelli, P., O’Toole, P. W., Pot, B., Vandamme, P., Walter, J., Watanabe, K., Wuyts, S., Felis, G. E., Gänzle, M. G., & Lebeer, S. (2020). A taxonomic note on the genus *Lactobacillus*: Description of 23 novel genera, emended description of the genus *Lactobacillus* Beijerinck 1901, and union of *Lactobacillaceae* and *Leuconostocaceae*. *International Journal of Systematic and Evolutionary Microbiology*, 70(4), 2782–2858. <https://doi.org/10.1099/ijsem.0.004107>

WEB REFERENCES

<https://bacdiv.dsmz.de>

<https://scikit-learn.org>

Chapter 4

Optically pure D-lactic acid production: lactic acid bacteria as key players in the bioplastic industry

Laura Troiani¹, Alessia Levante¹, Hannes Russmayer^{2,3}, Hans Marx², Erasmo Neviani¹,
Valentina Bernini¹, Camilla Lazzi¹, Michael Sauer^{2,4}

¹ Department of Food and Drug Sciences, Parma University, Parma, Italy

² Institute of Microbiology and Microbial Biotechnology, Department of Biotechnology,
University of Natural Resources and Life Sciences, Vienna, Austria

³ FermX GmbH, Vienna, Austria

⁴ OMV AG, Innovation & Technology, Department of Biotechnology, Vienna, Austria

Manuscript in preparation

ABSTRACT

Biodegradable polymers are the green alternative to conventional oil-based plastics and have a key role in the achievement of the Sustainable Development Goals of Agenda 2030. Nowadays, polylactic acid (PLA) is one of the most common bioplastics present in the global market. The lactic acid (LA) polymerization process and the enantiomeric ratio affect both the chemical and physical properties of the polymer. It has recently been discovered that the stereo complex obtained from the combination of poly-L(+)-lactic acid (PLLA) and poly-D(-)-lactic acid (PDLA) has better physical and chemical properties than the two polymers separately. L-lactic acid (L-LA) is currently more commonly produced and diffused than D-lactic acid (D-LA). However, D-LA has now gained attention because of the superior stability of PDLA. This study aimed to select D-LA-producing lactic acid bacteria (LAB) strains from the University of Parma's Culture Collection (UPCC). Furthermore, fermentation in bioreactors is proposed to maximize microbial D-LA production.

An enzymatic method was employed to find D-LA producers among 150 strains. Successively, 20 strains were selected and cultivated on a small scale for 72 h and LA production was followed by HPLC analysis.

Based on the results of the small-scale screening, *Lactobacillus delbrueckii* UPCC 2214 and *Leuconostoc citreum* UPCC 4516 were selected for fed-batch cultivations using a DASGIP® parallel bioreactor system (Eppendorf International, Jülich, Germany). After optimization of fermentation parameters, *L. delbrueckii* UPCC 2214 was found to be the most efficient strain for D-LA production as it demonstrated the double $Y_{p/s}$ of *Leuc. citreum* UPCC 4516 and higher productivity.

The shift in bioreactors has helped to understand the microbial natural limits and boost their potentialities. This work supports further investigations and improvements in D-LA production processes to advance the field of biomaterials with benefits for both industry and the environment.

KEYWORDS

Lactic acid bacteria, Bioreactors, *Lactobacillus delbrueckii*, *Leuconostoc citreum*, D-lactic acid, PLA, Bioplastic Industry

ABBREVIATIONS

LA = Lactic Acid

PLA = Polylactic acid

PLLA = Poly-L(+)-lactic acid

PDLA = Poly-D(-)-lactic acid

PDLLA = Poly-D/L-lactic acid

LAB = Lactic Acid Bacteria

UPCC = University of Parma's Culture Collection

D-LDH = D(+)-Lactate Dehydrogenase

L-LDH = L(-)-Lactate Dehydrogenase

$Y_{p/s}$ = Yield of LA

VP = Volumetric Productivity

r = Glucose Consumption Rate

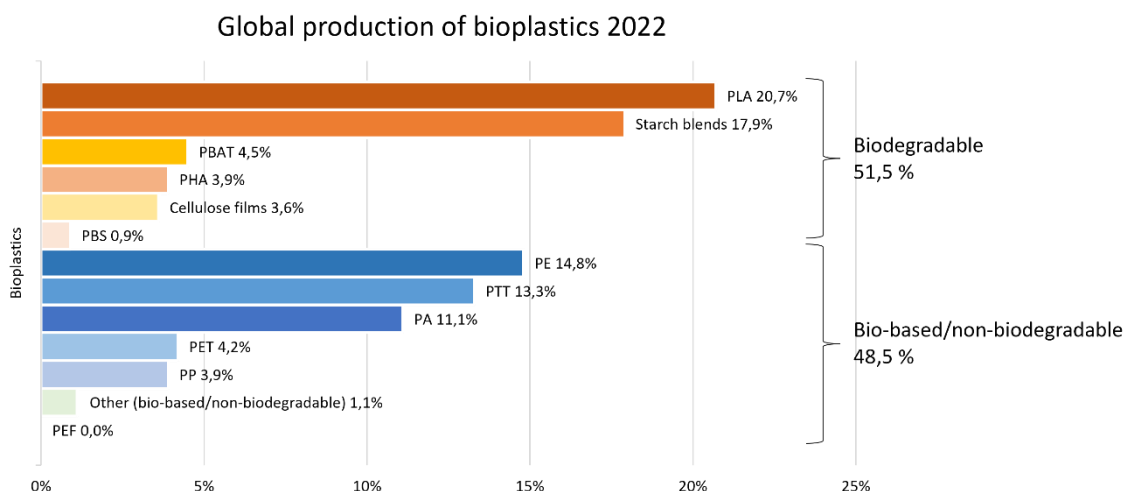
1. INTRODUCTION

In 2020, the European Commission published “*A new Circular Economy Action Plan for a cleaner and more competitive Europe*” which confirmed the importance of following a European strategy for plastics already included in the first Circular Economy Action Plan (2015) and adopted in January 2018. Among other actions provided by the strategy, there also is the support for innovation and investment in recyclable plastic materials to ensure that all plastic packaging is recyclable by 2030 (COM 28 final, 2018; COM 98 final, 2020). In this context, the EU Policy Framework for Bioplastics (2022) and the last revision of the EU Waste Framework Directive (2023) highlight the key role of biobased, biodegradable and compostable plastics to reach the Sustainable Development Goals 2030 (<https://www.european-bioplastics.org/>).

Biodegradable polymers are the green alternative to conventional oil-based plastics as they are supposed to turn into minerals at the end of their life cycle and can be absorbed by both microorganisms and the environment (Silva et al., 2023). The increasing awareness of plastic pollution has led to an escalation of bioplastic production, resulting in a collapse in prices (Haider et al., 2019).

Figure 1 displays the global industrial-scale production of biodegradable polymers. Market data from the European Bioplastics report a total bioplastic production of 2.22 million tons in 2022, of which 41.4% was produced in Asia, followed by Europe (26.5%). About 700,000 tons of bioplastic are destined for the flexible packaging market: together with the rigid packaging, this market segment represents 48% of the total bioplastic market. The global production consists of 51.1% biodegradable bioplastics and 48.5% bio-based and non-biodegradable bioplastics. The former is represented by polylactic acid (PLA), starch blends, poly(butylene adipate terephthalate) (PBAT, Ecoflex®), polyhydroxyalkanoates (PHA), regenerated cellulose films, and poly(butylene succinate) (PBS); while the latter include polyethylene (PE), poly(trimethylene terephthalate) (PTT), polyamides (PA), polyethylene terephthalate (PET), polypropylene (PP), and others. Moreover, industries are now developing Polyethylene Furanoate (PEF) that likely will be available on the market by 2023.

Figure 1 Global bioplastic production in 2022. Data source: <https://www.european-bioplastics.org/market>, accessed in October 2023.



Current statistics (<https://www.european-bioplastics.org/>, <https://renewable-carbon.eu/>, <https://www.databridgemarketresearch.com/reports/>, accessed in October 2023) clearly show the growing demand for PLA also due to its biodegradability, biocompatibility and processability features, advances and innovation in the production processes, and the numerous application it can find in many sectors (Alexandri et al., 2022; Balla et al., 2021; Djukić-Vuković et al., 2019; Huang et al., 2021; Jem & Tan, 2020; Maleki et al., 2022; Rajeshkumar et al., 2021; Singhvi et al., 2019). PLA represents the most important biodegradable plastic because of the similarity of its mechanical properties with oil-based plastics polystyrene (PS) and polyethylene terephthalate (PET) (Karamanlioglu et al., 2017).

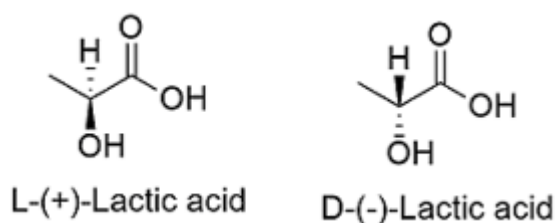
The total lactic acid (LA) produced annually worldwide is not only destined for polymerization (30% market share in 2022), but it is implicated in different economic sectors where it finds several applications (Fig. 2). In 2022, the global market of LA was valued at 3.46 billion US dollars, however, it is estimated to grow up to 8.70% by 2032 due to its use as a building block of polylactic acids (Abedi & Hashemi, 2020; Augustiniene et al., 2022; Ojo & De Smidt, 2023) (<https://www.precedenceresearch.com/>, accessed in October 2023). Corbion (Amsterdam, The Netherlands), Galactia (Celles, Belgium), and NatureWorks LLC (Minnetonka, MN, USA) are the global leader companies for lactic acid production (Alves de Oliveira et al., 2018).

Figure 2 All current applications of lactic acid. It is involved in many industries: food and beverages, chemicals, pharmaceuticals, cosmetics, biomedical, textiles, and bio-based materials. Figure adapted by Ojo & De Smidt (2023).



Lactic acid ($C_3H_6O_3$) exists in two isoforms (Fig. 3, L(+)-LA and D(-)-LA) or in a racemic mixture of them (D/L-LA) based on the production process (Jaramillo et al., 2021). LA can be produced both via chemical synthesis and microbial fermentation (Castillo Martinez et al., 2013; Sheeladevi & Ramanathan, 2010). The chemical synthesis mostly gives LA as a racemic mixture, whereas microbial fermentative processes guarantee optically pure enantiomers (Kim et al., 2018; Trontel & Batu, 2011).

Figure 3 The two enantiomers of lactic acid.



PLA can be obtained from the racemic mixture (PDLLA), or either from L-LA (PLLA) or D-LA (PDLA) (Swetha et al., 2023). The polymerization process and enantiomeric ratio affect both the chemical and physical properties of the final polymer (Klotz et al., 2016). PDLLA is thermally and mechanically unstable due to its amorphous form, while PLLA and PDLA have

better properties, but they are still unstable to heat. However, it was studied that the stereo complex obtained from the combination of PLLA and PDLA has enhanced mechanical performance, hydrolysis resistance, and thermal stability (Alexandri et al., 2022). This is due to the chemical stabilization via hydrogen and van der Waals bonding which allows to increase the melting point of the polymer from 60°C (PLLA) to 230°C (Alexandri et al., 2019).

L-LA is more commonly produced and diffused than D-LA for its major use in the food and pharmaceutical industry (Liu et al., 2022). D-LA, instead, stayed behind the scenes because of its toxicity (Pohanka, 2020), but it has now gained attention for the more significant stability of PDLA.

The production of pure enantiomers and the possibility of recovering waste substrates, thus implementing the circular economy, are some of the advantages of fermentation. This has led the fermentative processes to drive the LA production market, estimating that about 90% of the total LA is obtained by microorganisms (Alexandri et al., 2019). LA is the main end-product of lactic acid bacteria (LAB) metabolism, therefore they are doubtlessly involved in the LA market (Punia Bangar et al., 2022).

Among them, *Leuconostoc* spp. and *Lactobacillus delbrueckii* strains may produce D(-)-lactic acid with high optical purity and yield (Chemama et al., 2020; Karnaouri et al., 2020; Liu et al., 2022; Ma et al., 2022). Compared to L-LA-producing microorganisms, optically pure D-LA producers have lower fermentative performances. For this reason, genetic engineering is largely used to improve and boost the natural capabilities of bacteria. However, genetically modified microbes face restricted industrial use because of limitations posed by biosafety regulations and issues related to genetic instability. Therefore, industrial production facilities continue to favour the use of wild strains (Prasirtsak et al., 2017).

This work aimed to identify D-LA-producing strains within a microbial collection composed of lactic acid bacteria mainly of dairy origin. After the selection of the most promising strains, we wanted to scale fermentation in bioreactors to investigate both the natural capabilities and limitations of the strains and, through process optimization, boost their potential, setting a starting point for further exploration in the field of biomaterials.

2. MATERIAL AND METHODS

2.1. Primary screening

A set of 150 lactic acid bacteria (LAB) strains was selected from the University of Parma's Culture Collection (UPCC) by ensuring a large biodiversity of the sample. The screening involved species belonging to 12 different genera, previously isolated from several food matrices, especially of dairy origin, since 2012. Cultures were stored at $-80\text{ }^{\circ}\text{C}$ in their optimal growth medium (MRS, TSB or M17) containing 15% glycerol (w/v). Strains were propagated three times in a fresh medium and incubated overnight. 6 ml of MRS, TSB (VWR International bvba, Leuven, Belgium) or M17 (Merck, Darmstadt, Germany) were inoculated with 3% (v/v) overnight preculture and incubated at the optimal temperature of strains for 24 h in anaerobiosis (AnaeroGen, ThermoScientific, Landsmeer, The Netherlands). After incubation, 1 ml of culture was centrifugated at 13400 rpm for 10 min and the supernatant was analysed. L-LA and D-LA were separately detected by an enzymatic method (Megazyme®, Ireland). Samples were diluted to a detectable concentration (0.005-0.3 g/L for each isomer) and two sequential enzymatic reactions were performed by using D-LDH and L-LDH. The absorbance of every reaction was read at 340 nm and the concentration of the two isomers was calculated as follows:

$$c = \frac{V * MW}{\varepsilon * d * v} * \Delta A_{D-LA \text{ or } L-LA} \quad (1)$$

where c = concentration of D-LA or L-LA (g/L)

V = final volume (ml)

MW = molecular weight of lactic acid (g/mol)

ε = excitation coefficient of NADH at 340 nm

d = light path (cm)

v = sample volume (ml)

ΔA_{D-LA} = difference between the absorbance of the first enzymatic reaction (after addition of D-LDH) and that of the sample before the reaction

ΔA_{D-LA} = difference between the absorbance of the second enzymatic reaction (after addition of L-LDH) and that of the first enzymatic reaction

The supernatants were also analysed by HPLC (Shimadzu, Korneuburg, Austria) to accurately quantify the total lactate produced. The HPLC was equipped with an Aminex HPX-87H column (300 × 7.8 mm, Biorad) operated at a temperature of $60\text{ }^{\circ}\text{C}$, a flow of 0.6 ml/min and by using 4 mM H_2SO_4 as mobile phase. Samples and standards for HPLC analysis consisted of 900 μL of sample/standard added with 100 μL 40 mM H_2SO_4 . Subsequently, samples and standards

were filtered. For the detection of metabolites, a refraction index detector (RID-10A, Shimadzu, Korneuburg, Austria) was used.

2.2. Secondary screening and scaling-up

Strains producing 80% or more of D-LA were selected ($n = 20$). Cultures were repeated on a small scale on MRS (pH = 6.2) to further reduce the number of strains to cultivate in bioreactors. MRS medium contained per litre: 10 g casein peptone, 10 g meat extract, 5 g yeast extract, 1 g tween 80, 2 g K_2HPO_4 , 5 g sodium acetate, 2.60 g tri-sodium citrate· $2H_2O$, 1.17 g $NH_4(HPO_4)_2$, 0.20 g $MgSO_4 \cdot 7H_2O$, 0.05 g $MnSO_4 \cdot H_2O$ and 22 g glucose· H_2O . 6 ml MRS were inoculated to an $OD_{600} = 0.2$ and subsequently incubated at 30°C until 72 h in anaerobic conditions (AnaeroGen, ThermoScientific, Landsmeer, The Netherlands). The growth was followed every 24 h by OD_{600} measurement on a spectrophotometer (Hitachi U-1900). The supernatant was centrifugated (13400 rpm for 10 min), the pH decrease was monitored, and both lactate and glucose were quantified by HPLC analysis (Shimadzu, Korneuburg, Austria). Based on the highest amount and purity of lactate produced, two strains were selected for the scaling up (*Leuconostoc citreum* UPCC 4516 and *Lactobacillus delbrueckii* UPCC 2214).

2.3. Scaling-up and optimization

Pre-cultures were prepared to inoculate a batch with a specific OD_{600} . Fermentations in a DASGIP® parallel bioreactor system (Eppendorf International, Jülich, Germany) were set up as suggested by Russmayer et al. (2019). The batch process was carried out with an initial volume of 700 ml MRS with 22 g/L glucose monohydrate as a carbon source, inoculated to an OD_{600} of 0.2 with an overnight culture. After the total glucose depletion, we started the fed-batch process initially characterised by a constant feeding rate of 2.5 ml/h 50% (w/v) glucose. The pH was kept constant at 6.2 via the addition of 4M NaOH, temperature set at 30°C for *Leuc. citreum* UPCC 4516 and 37°C for *L. delbrueckii* UPCC 2214, anaerobiosis, agitation = 200 rpm. Fermentations were stopped at 74 h. This process was successively optimized: i) by starting the feeding as soon as the glucose was over (7 h for both strains); ii) by adjusting the feeding rate to 4 ml/h for *L. delbrueckii* UPCC 2214 and 6 ml/h for *Leuc. citreum* UPCC 4516; iii) by increasing the NaOH concentration to 8M; iv) by adding 10 times concentrated fresh medium without glucose as soon as the growth stopped (24 h) to provide new nutrients. Cultures were sampled periodically: the OD_{600} was measured to follow the growth, and glucose consumption, lactate and ethanol production were analysed through HPLC analysis.

Glucose consumption and lactate/ethanol production were reported in grams considering the fermentation volumes. The yield of LA ($Y_{p/s}$), the volumetric productivity (VP), and the glucose consumption rate (r) were calculated using the following equations:

$$Y_{p/s} = \frac{\text{lactate produced (g)}}{\text{glucose consumed (g)}} \quad (2)$$

$$VP = \frac{\text{total lactate concentration (g/L)}}{\text{fermentation time (h)}} \quad (3)$$

$$r = \frac{\text{glucose consumed (g)}}{\text{fermentation time (h)}} \quad (4)$$

3. RESULTS

3.1. Primary screening

Out of 150 strains, setting 80% as the threshold for isomeric purity of lactic acid, 74 produced L-LA, 20 produced D-LA and the rest (56 strains) produced a racemic mixture of both isomers (Fig. 4). The total lactate production was successively quantified by HPLC analysis: results related to the selected D-LA-producing strains are reported in Table 1.

Figure 4 Percentage of lactic acid isomers produced by screened LAB and quantified by enzymatic method. Strains producing more than 80% of D-LA were selected (right side panel).

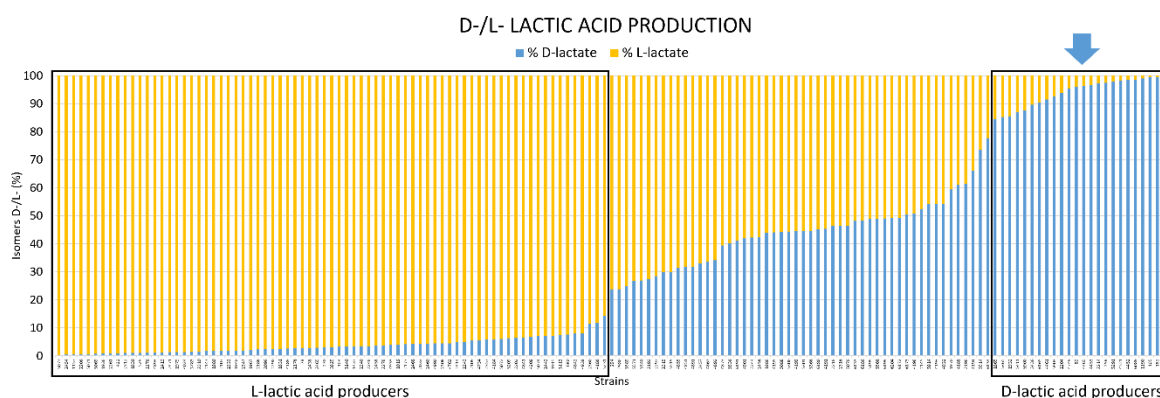


Table 1 Selected strains producers of optically pure D-lactic acid (> 80%). D-LA (%) was measured by enzymatic method and Lactate (g/L) was detected at 24 h by HPLC analysis.

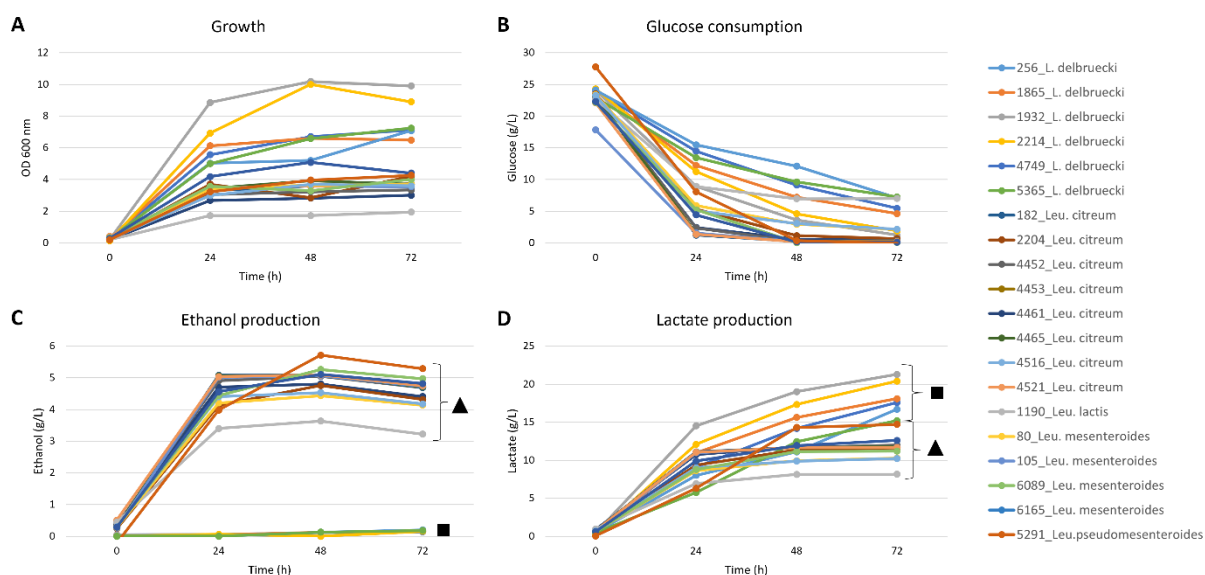
ID strain	Species	Lactate (g/L)	D-LA (%)
80	<i>Leuconostoc mesenteroides</i>	8.98	96.08
105	<i>Leuconostoc mesenteroides</i>	8.96	99.39
182	<i>Leuconostoc citreum</i>	8.84	99.46
256	<i>Lactobacillus delbrueckii</i>	13.55	97.40
1190	<i>Leuconostoc lactis</i>	7.72	99.03
1865	<i>Lactobacillus delbrueckii</i>	11.75	84.46
1932	<i>Lactobacillus delbrueckii</i>	17.24	85.56
2204	<i>Leuconostoc citreum</i>	8.26	93.87
2214	<i>Lactobacillus delbrueckii</i>	17.05	97.21
4452	<i>Leuconostoc citreum</i>	8.91	98.35
4453	<i>Leuconostoc citreum</i>	8.76	96.64
4461	<i>Leuconostoc citreum</i>	8.08	91.44
4465	<i>Leuconostoc citreum</i>	8.24	96.20
4516	<i>Leuconostoc citreum</i>	8.77	99.70
4521	<i>Leuconostoc citreum</i>	7.85	98.32

4749	<i>Lactobacillus delbrueckii</i>	8.95	95.43
5291	<i>Leuconostoc pseudomesenteroides</i>	8.51	97.87
5365	<i>Lactobacillus delbrueckii</i>	13.88	85.26
6089	<i>Leuconostoc mesenteroides</i>	8.02	98.42
6165	<i>Leuconostoc mesenteroides</i>	7.62	90.30

3.2. Secondary screening

With the secondary screening, as shown in Figure 5, LAB's growth (OD₆₀₀), glucose consumption, lactate and ethanol production were analysed over time (72 h). Compared to *Leuconostoc* spp. strains, *L. delbrueckii* ones showed a greater production of biomass (Fig. 5A), a slower consumption of glucose (Fig. 5B), and constant production of lactate until 72 h (Fig. 5D). *Leuconostoc* spp. strains finished all the sugar by 24 h and, accordingly, stopped producing lactate. Only *Leuconostoc* spp. strains produced ethanol because of the heterofermentative metabolism.

Figure 5 Small-scale fermentations: A) growth (OD₆₀₀); B) glucose consumption (g/L); C) ethanol production (g/L); D) lactate production (g/L). In panels C and D, the closed black square indicates *L. delbrueckii* strains, while the closed black triangle indicates *Leuconostoc* spp. strains.



Despite all strains of *L. delbrueckii* produced more lactate than *Leuconostoc* spp. by using the same amount of glucose, the final product is characterised by a lower optical purity. However, considering both the greatest amount of LA produced and the highest isomeric purity, *L.*

delbrueckii UPCC 2214 and *Leuc. citreum* UPCC 4516 were selected to have a comparison between two different species in the scaling process. At the end of the fermentation, *L. delbrueckii* UPCC 2214 and *Leuc. citreum* UPCC 4516 produced 20.40 g/L and 11.69 g/L lactate, respectively (Tab. 2).

Table 2 Selected strains from the secondary screening for the simultaneous high lactate titre and high optical purity.

ID strain	Species	Lactate 24 h (g/L)	Lactate 48 h (g/L)	Lactate 72 h (g/L)
2214	<i>Lactobacillus delbrueckii</i>	12.09	17.32	20.41
4516	<i>Leuconostoc citreum</i>	10.97	11.68	11.69

3.3. Fermentation in bioreactors

L. delbrueckii UPCC 2214 and *Leuc. citreum* UPCC 4516 were selected for bioreactor cultivations and, after optimization (Fig. 6, Tab. 3), the former turned out to be the most efficient.

The batch process lasted 7 h: *L. delbrueckii* UPCC 2214 and *Leuc. citreum* UPCC 4516 consumed 15.95 ± 0.43 g and 15.88 ± 0.45 g of glucose, respectively, and produced 13.03 ± 0.40 g and 7.24 ± 0.21 g of lactate, respectively. *L. delbrueckii* showed a yield_{p/s} ($Y_{p/s}$) of 0.82 ± 0.003 g/g and a volumetric productivity (VP) of 2.80 ± 0.08 g/L/h. *Leuc. citreum*, instead, had $Y_{p/s} = 0.46 \pm 0.0002$ g/g and $VP = 1.51 \pm 0.05$ g/L/h.

The last 2 hours of the batch phase allowed us to calculate the theoretical maximum glucose consumption rate (r) to subsequently set the optimal feeding rate: $r = 4.55 \pm 0.05$ g/h for *L. delbrueckii* UPCC 2214 and $r = 4.70 \pm 0.13$ g/h for *Leuc. citreum* UPCC 4516.

After 7 h, the fed-batch phase was started, and the process continued until 74 h. At the end of the fermentation, *Leuc. citreum* and *L. delbrueckii* produced approximately the same amount of lactate (63.08 ± 0.73 g and 55.18 ± 0.16 g, respectively). However, *Leuc. citreum* consumed the double amount of glucose (149.64 ± 0.81 g) more constantly until 48.5 h compared to *L. delbrueckii* (74.26 ± 0.35 g glucose consumed) that started accumulating the sugar after 24 h. At 74 h, *L. delbrueckii* reached twice the yield_{p/s} of *Leuc. citreum* (0.74 ± 0.001 g/g and 0.42 ± 0.003 g/g, respectively), although they present a similar final volumetric productivity (0.96 ± 0.009 g/L/h and 0.85 ± 0.01 g/L/h, respectively). *L. delbrueckii* UPCC 2214 was selected as the most promising strain of the screening for D-lactic acid production as it was the fastest and the least C-consuming strain ($Y_{p/s} = 0.74 \pm 0.001$).

Figure 6 Comparison between fermentation in bioreactors of *L. delbrueckii* UPCC 2214 and *Leuc. citreum* UPCC 4516. Data represent the average of duplicate experiments with experimental error bars. During 74 hours of fermentation, four parameters were monitored: ●) glucose consumption (g); ■) lactate production (g); □) ethanol production (g); ▲) cellular growth (OD₆₀₀). Ten times concentrated fresh medium was added at 30 h (grey arrow).

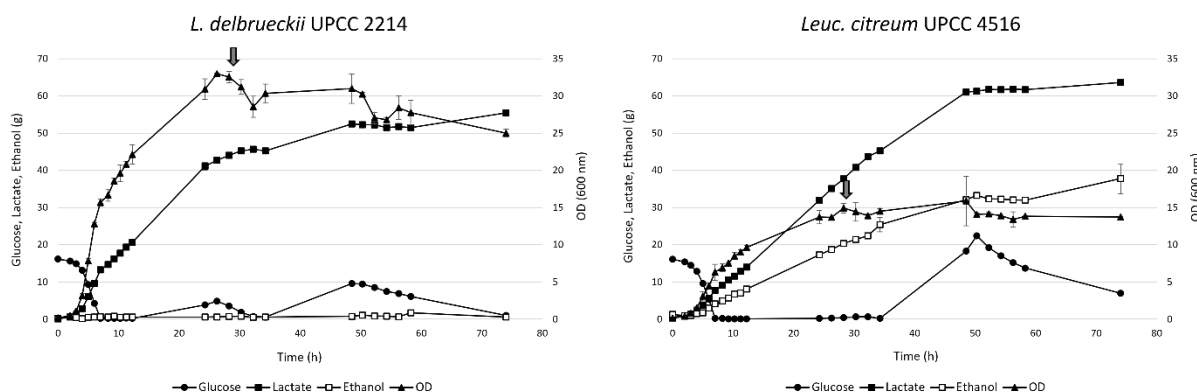


Table 3 Parameters monitored during different stages of cultivation in bioreactors.

Parameter	Batch		Fed-batch		Constant glucose consumption		End-point	
UPCC	2214	4516	2214	4516	2214	4516	2214	4516
Time (h)	7	7	67	67	24.25	48.50	74	74
Δ glucose (g)	15.95 ± 0.43	15.88 ± 0.45	58.31 ± 0.78	133.77 ± 0.36	49.65 ± 0.49	133.14 ± 0.06	74.26 ± 0.35	149.64 ± 0.81
Δ lactate (g)	13.04 ± 0.40	7.24 ± 0.21	42.15 ± 0.56	55.84 ± 0.52	40.90 ± 1.09	60.54 ± 0.53	55.18 ± 0.16	63.082 ± 0.73
Yield _{p/s} (g/g)	0.82 ± 0.003	0.46 ± 0.0002	0.72 ± 2.60E-05	0.42 ± 0.003	0.82 ± 0.01	0.45 ± 0.004	0.74 ± 0.001	0.42 ± 0.003
Volumetric productivity (g/L/h)	2.80 ± 0.08	1.51 ± 0.05	0.77 ± 0.02	0.78 ± 0.006	2.24 ± 0.04	1.18 ± 0.01	0.96 ± 0.009	0.85 ± 0.01
Glucose consumption rate (g/h)	2.28 ± 0.06	2.27 ± 0.06	0.87 ± 0.01	2.00 ± 0.005	2.05 ± 0.02	2.74 ± 0.001	1.00 ± 0.005	2.02 ± 0.01

4. DISCUSSIONS

This study aimed to identify optically pure D-LA producers among wild-type strains of lactic acid bacteria within the University of Parma's Culture Collection. Furthermore, we wanted to investigate both their natural capabilities and limitations through fermentation in bioreactors for further explorations in the field of biomaterials.

From the primary screening, *Lactobacillus delbrueckii* and *Leuconostoc* spp. strains were found to produce optically pure D-LA, thus confirming what was reported in the literature (Bustamante et al., 2020; de la Torre et al., 2018; Ju et al., 2016; Liang et al., 2020; Prasirtsak et al., 2017; Sae-Tang et al., 2023). After 24 hours, it was already possible to distinguish two different behaviours (Tab. 1): *L. delbrueckii* strains produced more than 10 g/L of lactate, while *Leuconostoc* spp. showed a lower titre. On the other hand, D-lactic acid obtained from *Leuconostoc* spp. has a higher purity than from *L. delbrueckii*.

The secondary screening confirmed the two different trends (Fig. 5) due to the type of metabolism characterising the strains. *L. delbrueckii* is a homofermentative bacteria that produces primarily lactic acid from sugars via the Embden-Meyerhof Pathway (EMP), while *Leuconostoc* strains are characterised by a heterofermentative metabolism, which addresses sugars for lactic acid, ethanol/acetic acid and carbon dioxide production via the Phosphoketolase Pathway (PP) (Duar et al., 2017; Halasz, 2009). In the first 24 hours of fermentation (Fig. 5B), *Leuconostoc* spp. strains finished all the glucose supplied, therefore they stopped producing LA before *L. delbrueckii* (Fig. 5D), which instead consumed the sugar slower and continued to produce LA until the end of fermentation. Both strains also showed a different biomass yield on glucose, which was higher for *L. delbrueckii* strains than *Leuconostoc* ones (Fig. 5A), suggesting a different carrying capacity of the two species.

Based on the highest amount and purity of D-lactic acid produced, two strains were selected for bioreactor cultivation: *L. delbrueckii* UPCC 2214 and *Leuc. citreum* UPCC 4516. With the same initial amount of sugar, after 72 h, they produced one the double quantity of lactate than the other (Tab. 2) since, as cited above, *Leuconostoc* addresses part of the carbon to produce ethanol (4.73 g/L at 72 h).

Moving on to a larger scale experiment, results still support the higher efficiency of *L. delbrueckii* UPCC 2214. It had twice the yield of *Leuc. citreum* UPCC 4516 that took longer and consumed more glucose to finally produce about the same amount of lactate.

The batch phase (first 7 h) highlights the natural capabilities of microorganisms, what they can do with any external action. Hence, the batch mode was used to characterise the strains. The

advantages of the batch are: less technical equipment is needed, easy management, high titre and yield, and minor risk of contamination. However, the accumulation of potential by-products is a disadvantage of the batch culture. Excessive concentration of the substrate can prolong the lag phase and cause cell lysis, thus reducing the sugar utilisation and fermentation rate. Moreover, the inhibitory effects of the end-product concentration may cause reduced productivity, while a limitation of nutrients results in a low concentration of cells (Rawoof et al., 2021). The two strains used the same amount of glucose and, accordingly, at the same rate. However, in this first phase, *L. delbrueckii* already shows a double yield_{p/s} and productivity than *Leuc. citreum*.

Once the fed-batch phase is started, microorganisms are subjected to an artificial manipulation, a sort of stressful process to push them to enhance their abilities by prolonging the exponential phase. The fed-batch mode gives low substrate inhibition by supplying calculated substrate concentration during the fermentation. This allows to maintain the cell growth and increase the productivity. However, bottlenecks are represented by the design and optimization of the feeding, and the inhibitory effect of the end-product accumulation (Ojo & De Smidt, 2023). Around 34 h, the lactate production curves intersected each other until *Leuc. citreum* UPCC 4516 overtook *L. delbrueckii* UPCC 2214 by using more than double glucose. In Figure 6 it is clear the great request of glucose by *Leuc. citreum* UPCC 4516 (high consumption rate) that kept constantly consuming it until 48.5 h, thus strongly affecting yield_{p/s} and productivity. Furthermore, *Leuc. citreum* UPCC 4516 produced LA and ethanol (36.51 ± 3.35 g at 74 h) maintaining a ratio of 1.73:1. Kim et al. (2018) explained that such a characteristic is genetically determined and sugar concentration independent.

As soon as the OD₆₀₀ stopped increasing, indicating the beginning of the stationary phase, a concentrated fresh medium without glucose was added to reintegrate missing nutrients. The lack of nutrients could lead to growth restriction, which also negatively affects lactate productivity. Results show that supplementation of new nutrients did not restart the cellular growth of any of the two bacteria.

Nevertheless, it is noteworthy to analyse the stationary phase of both strains. LA production curve of *L. delbrueckii* UPCC 2214 is paired with its cellular growth curve, right as the primary metabolite production trend should be (Lee et al., 2009): cells and metabolite are produced more or less simultaneously. Furthermore, we can suppose that cells do not react to the addition of new nutrients because they reached their physiological limit of lactic acid tolerance, thereby they start being inhibited by their metabolism product (Cubas-Cano et al., 2018; Peyer et al., 2017), or they achieve the carrying capacity of the environment (Dedrick et al., 2023).

Leuc. citreum UPCC 4516, instead, did not stop producing lactate when it reached the stationary phase. Within this phase, available carbon is still converted to metabolites, as observed for secondary metabolite biosynthesis. It leads to high glucose consumption, which seems to be entirely destined for lactate conversion. In this case, *Leuc. citreum* UPCC 4516 has probably reached its carrying capacity, whereas it no longer produces biomass, but its metabolism is still active.

Looking at Table 4, results obtained from this work are promising since, with minimal optimization, we reached values fully in line with the data present in the literature.

Table 4 Comparison of D-lactic acid production by *L. delbrueckii* using different C sources and processes in the literature.

Microorganism	Fermentation mode	C source	Productivity _{LA} (g/L/h)	Yield _{p/s} (g/g)	Reference
<i>L. delbrueckii</i> UPCC 2214	Batch	Glucose	2.80	0.82	This study
<i>L. delbrueckii</i> UPCC 2214	Batch + Constant feeding rate fed-batch (24 h)	Glucose	2.24	0.82	This study
<i>L. delbrueckii</i> sp. <i>lactis</i> QU 41	Batch	Glucose	0.52	n.d.	(Tashiro et al., 2011)
<i>L. delbrueckii</i> IFO 3202	Batch	Rice saccharified	0.65	n.d.	(Fukushima et al., 2004)
<i>L. delbrueckii</i> sp. <i>delbrueckii</i> CECT 286	Batch	Orange peel waste hydrolysate	2.35	0.92	(de la Torre et al., 2018)
<i>L. delbrueckii</i>	Batch	Molasses sugar	3.36	0.61	(Beitel et al., 2020)
<i>L. delbrueckii</i>	Constant feeding rate fed-batch	Molasses sugar	3.96	0.61	(Beitel et al., 2020)
<i>L. delbrueckii</i> sp. <i>delbrueckii</i> CECT 286	Batch	Orange peel waste	0.63	0.86	(Bustamante et al., 2020)
<i>L. delbrueckii</i> sp. <i>bulgaricus</i> CECT 5037	Batch	Orange peel waste	0.55	0.84	(Bustamante et al., 2020)
<i>L. delbrueckii</i> sp. <i>bulgaricus</i> WT	Batch	Glucose	n.d.	0.86	(Vishnu Prasad et al., 2020)

<i>Lb. delbrueckii</i>	Batch	Broken rice	3.58	0.81	(Abedi & Hashemi, 2020)
<i>Lb. delbrueckii</i> NCIM 2025	Batch SSF	Cassava bagasse	1.36	0.94	(Abdel-Rahman et al., 2013)

Bioreactor tests reflect what has already been seen on a small scale. However, in the small-scale screening, the fast pH drop caused a decrease in productivity, therefore cultivation in bioreactors was necessary to understand the natural limitations of the studied strains and try in future to boost their potential.

Future outlooks for this study are to investigate the lactic acid tolerance of strains and to push them beyond their limits by designing adaptive laboratory evolution experiments. Moreover, to further optimize the fermentative process in bioreactors, it is suggested to prolong the batch phase by supplying a higher concentration of sugar and increasing the inoculum, thus leveraging the natural capabilities of microorganisms. It is very interesting also to implement different process modes, such as repeated batch or continuous fermentation. The first one gives high growth rates and it is time and labour-saving, although it requires specific devices for cell concentration. The second one, instead, guarantees high productivity and controlled growth rates, but it implies carbon leftovers and a high risk of contamination (Ojo & De Smidt, 2023).

5. CONCLUSIONS

In conclusion, this investigation provides useful insights into the production of D-lactic acid, precursor of PDLA, by *Lactobacillus delbrueckii* and *Leuconostoc citreum*. The screening of strains and cultivation in bioreactors gave out several key findings on the fermentation processes.

The microbial strain selection confirmed the efficient capacity of the two species in D-LA biosynthesis. The regulation of fermentation conditions (pH, temperature, anaerobic condition, and substrate concentration), especially in the fed-batch mode, and the scaling in bioreactors are essential to maximize yield and productivity. A controlled environment allows to obtain reliable and reproducible results for potential industrial applications. This research is an additional piece of the puzzle for further exploration and optimization of D-LA production processes. Future studies may be addressed to different fermentation approaches or to the improvement of microbial capabilities.

The successful production of D-LA represents a step forward in PLA production, thus contributing to the development and expansion in the field of biomaterials with benefits for both industry and the environment.

REFERENCES

- Abdel-Rahman, M. A., Tashiro, Y., & Sonomoto, K. (2013). Recent advances in lactic acid production by microbial fermentation processes. *Biotechnology Advances*, 31(6), 877–902. <https://doi.org/10.1016/j.biotechadv.2013.04.002>
- Abedi, E., & Hashemi, S. M. B. (2020). Lactic acid production – producing microorganisms and substrates sources-state of art. *Heliyon*, 6(10), e04974. <https://doi.org/10.1016/j.heliyon.2020.e04974>
- Alexandri, M., Hübner, D., Schneider, R., Fröhling, A., & Venus, J. (2022). Towards efficient production of highly optically pure d-lactic acid from lignocellulosic hydrolysates using newly isolated lactic acid bacteria. *New Biotechnology*, 72, 1–10. <https://doi.org/10.1016/j.nbt.2022.08.003>
- Alexandri, M., Schneider, R., Mehlmann, K., & Venus, J. (2019). Recent Advances in D-Lactic Acid Production from Renewable Resources: Case Studies on Agro-Industrial Waste Streams. *Food Technology and Biotechnology*, 57(3), 293–304. <https://doi.org/10.17113/ftb.57.03.19.6023>
- Alves de Oliveira, R., Komesu, A., Vaz Rossell, C. E., & Maciel Filho, R. (2018). Challenges and opportunities in lactic acid bioprocess design—From economic to production aspects. *Biochemical Engineering Journal*, 133, 219–239. <https://doi.org/10.1016/j.bej.2018.03.003>
- Augustiniene, E., Valanciene, E., Matulis, P., Syrpas, M., Jonuskiene, I., & Malys, N. (2022). Bioproduction of L—And D -lactic acids: Advances and trends in microbial strain application and engineering. *Critical Reviews in Biotechnology*, 42(3), 342–360. <https://doi.org/10.1080/07388551.2021.1940088>
- Balla, E., Daniilidis, V., Karlioti, G., Kalamas, T., Stefanidou, M., Bikiaris, N. D., Vlachopoulos, A., Koumentakou, I., & Bikiaris, D. N. (2021). Poly(lactic Acid): A Versatile Biobased Polymer for the Future with Multifunctional Properties—From Monomer Synthesis, Polymerization Techniques and Molecular Weight Increase to PLA Applications. *Polymers*, 13(11), 1822. <https://doi.org/10.3390/polym13111822>
- Beitel, S. M., Coelho, L. F., & Contiero, J. (2020). Efficient Conversion of Agroindustrial Waste into D(-) Lactic Acid by *Lactobacillus delbrueckii* Using Fed-Batch Fermentation. *BioMed Research International*, 2020, 1–13. <https://doi.org/10.1155/2020/4194052>
- Bustamante, D., Tortajada, M., Ramón, D., & Rojas, A. (2020). Production of D-Lactic Acid by the Fermentation of Orange Peel Waste Hydrolysate by Lactic Acid Bacteria. *Fermentation*, 6(1), Article 1. <https://doi.org/10.3390/fermentation6010001>
- Castillo Martinez, F. A., Balciunas, E. M., Salgado, J. M., Domínguez González, J. M., Converti, A., & Oliveira, R. P. de S. (2013). Lactic acid properties, applications and production: A review. *Trends in Food Science & Technology*, 30(1), 70–83. <https://doi.org/10.1016/j.tifs.2012.11.007>
- Chemama, T., Pumas, C., Tragoolpua, Y., & Thongwai, N. (2020). Feasibility Study for D-Lactic Acid Production from Thai Rice by *Leuconostoc pseudomesenteroides* TC49 and D-Lactic Acid Purification. *Chiang Mai Journal of Science*, 47, 403–417.
- COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS A European Strategy for Plastics in a Circular

- Economy, (2018). <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1516265440535&uri=COM:2018:28:FIN>
- COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS A new Circular Economy Action Plan For a cleaner and more competitive Europe, (2020). <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1583933814386&uri=COM:2020:98:FIN>
- Cubas-Cano, E., González-Fernández, C., Ballesteros, M., & Tomás-Pejó, E. (2018). Biotechnological advances in lactic acid production by lactic acid bacteria: Lignocellulose as novel substrate: Lactic acid production from lignocellulose: Biotechnological advances. *Biofuels, Bioproducts and Biorefining*, 12(2), 290–303. <https://doi.org/10.1002/bbb.1852>
- de la Torre, I., Ladero, M., & Santos, V. E. (2018). Production of d-lactic acid by *Lactobacillus delbrueckii* ssp. *delbrueckii* from orange peel waste: Techno-economical assessment of nitrogen sources. *Applied Microbiology and Biotechnology*, 102(24), 10511–10521. <https://doi.org/10.1007/s00253-018-9432-4>
- Dedrick, S., Warriar, V., Lemon, K. P., & Momeni, B. (2023). When does a Lotka-Volterra model represent microbial interactions? Insights from in vitro nasal bacterial communities. *mSystems*, 8(3), e00757-22. <https://doi.org/10.1128/msystems.00757-22>
- Djukić-Vuković, A., Mladenović, D., Ivanović, J., Pejin, J., & Mojović, L. (2019). Towards sustainability of lactic acid and poly-lactic acid polymers production. *Renewable and Sustainable Energy Reviews*, 108, 238–252. <https://doi.org/10.1016/j.rser.2019.03.050>
- Duar, R. M., Lin, X. B., Zheng, J., Martino, M. E., Grenier, T., Pérez-Muñoz, M. E., Leulier, F., Gänzle, M., & Walter, J. (2017). Lifestyles in transition: Evolution and natural history of the genus *Lactobacillus*. *FEMS Microbiology Reviews*, 41(Supp_1), S27–S48. <https://doi.org/10.1093/femsre/fux030>
- Fukushima, K., Sogo, K., Miura, S., & Kimura, Y. (2004). Production of D-lactic acid by bacterial fermentation of rice starch. *Macromolecular Bioscience*, 4(11), 1021–1027. <https://doi.org/10.1002/mabi.200400080>
- Haider, T. P., Völker, C., Kramm, J., Landfester, K., & Wurm, F. R. (2019). Plastics of the Future? The Impact of Biodegradable Polymers on the Environment and on Society. *Angewandte Chemie International Edition*, 58(1), 50–62. <https://doi.org/10.1002/anie.201805766>
- Halasz, A. (n.d.). Lactic Acid Bacteria. *Lactic Acid Bacteria*, 6.
- Huang, S., Xue, Y., Yu, B., Wang, L., Zhou, C., & Ma, Y. (2021). A Review of the Recent Developments in the Bioproduction of Polylactic Acid and Its Precursors Optically Pure Lactic Acids. *Molecules*, 26(21), Article 21. <https://doi.org/10.3390/molecules26216446>
- Jaramillo, L., Santos, D., Guedes, D., Dias, D., Borges, E., & Pereira, N. (2021). Production of Lactic Acid Enantiomers by *Lactobacillus* Strains under Limited Dissolved Oxygen Conditions in the Presence of a Pentose Fraction. *Applied Biochemistry and Microbiology*, 57(4), 452–457. <https://doi.org/10.1134/S0003683821040050>
- Jem, K. J., & Tan, B. (2020). The development and challenges of poly (lactic acid) and poly (glycolic acid). *Advanced Industrial and Engineering Polymer Research*, 3(2), 60–70. <https://doi.org/10.1016/j.aiepr.2020.01.002>

- Ju, S. Y., Kim, J. H., & Lee, P. C. (2016). Long-term adaptive evolution of *Leuconostoc mesenteroides* for enhancement of lactic acid tolerance and production. *Biotechnology for Biofuels*, 9(1), 240. <https://doi.org/10.1186/s13068-016-0662-3>
- Karamanlioglu, M., Preziosi, R., & Robson, G. D. (2017). Abiotic and biotic environmental degradation of the bioplastic polymer poly(lactic acid): A review. *Polymer Degradation and Stability*, 137, 122–130. <https://doi.org/10.1016/j.polymdegradstab.2017.01.009>
- Karnaouri, A., Asimakopoulou, G., Kalogiannis, K. G., Lappas, A., & Topakas, E. (2020). Efficient d-lactic acid production by *Lactobacillus delbrueckii* subsp. *Bulgaricus* through conversion of organosolv pretreated lignocellulosic biomass. *Biomass and Bioenergy*, 140, 105672. <https://doi.org/10.1016/j.biombioe.2020.105672>
- Kim, J.-W., Choi, B.-H., Jung, J.-H., Yuan, X., Kim, J. M., & Lee, P. C. (2018). Genome resequencing and analysis of d-lactic acid fermentation ability of *Leuconostoc mesenteroides* subsp. *Mesenteroides* ATCC 8293. *Process Biochemistry*, 75, 83–88. <https://doi.org/10.1016/j.procbio.2018.09.009>
- Klotz, S., Kaufmann, N., Kuenz, A., & Prüße, U. (2016). Biotechnological production of enantiomerically pure d-lactic acid. *Applied Microbiology and Biotechnology*, 100(22), 9423–9437. <https://doi.org/10.1007/s00253-016-7843-7>
- Lee, J.-E., Hwang, G.-S., Lee, C.-H., & Hong, Y.-S. (2009). Metabolomics Reveals Alterations in Both Primary and Secondary Metabolites by Wine Bacteria. *Journal of Agricultural and Food Chemistry*, 57(22), 10772–10783. <https://doi.org/10.1021/jf9028442>
- Liang, S., Jiang, W., Song, Y., & Zhou, S.-F. (2020). Improvement and Metabolomics-Based Analysis of d-Lactic Acid Production from Agro-Industrial Wastes by *Lactobacillus delbrueckii* Submitted to Adaptive Laboratory Evolution. *Journal of Agricultural and Food Chemistry*, 68(29), 7660–7669. <https://doi.org/10.1021/acs.jafc.0c00259>
- Liu, J., Piao, H., Liu, C., Li, G., Cui, H., & Jin, Q. (2022). Characterization of Key Enzymes for D-lactic Acid Synthesis in *Leuconostoc citreum* KM20. *Biotechnology and Bioprocess Engineering*, 27(6), 921–929. <https://doi.org/10.1007/s12257-022-0110-0>
- Ma, K., Cui, Y., Zhao, K., Yang, Y., Wang, Y., Hu, G., & He, M. (2022). D-Lactic acid production from agricultural residues by membrane integrated continuous fermentation coupled with B vitamin supplementation. *Biotechnology for Biofuels and Bioproducts*, 15(1), 24. <https://doi.org/10.1186/s13068-022-02124-y>
- Maleki, H., Azimi, B., Ismaeilimoghadam, S., & Danti, S. (2022). Poly(lactic acid)-Based Electrospun Fibrous Structures for Biomedical Applications. *Applied Sciences*, 12(6), 3192. <https://doi.org/10.3390/app12063192>
- Ojo, A. O., & De Smidt, O. (2023). Lactic Acid: A Comprehensive Review of Production to Purification. *Processes*, 11(3), 688. <https://doi.org/10.3390/pr11030688>
- Peyer, L. C., Bellut, K., Lynch, K. M., Zarnkow, M., Jacob, F., De Schutter, D. P., & Arendt, E. K. (2017). Impact of buffering capacity on the acidification of wort by brewing-relevant lactic acid bacteria. *Journal of the Institute of Brewing*, 123(4), 497–505. <https://doi.org/10.1002/jib.447>
- Pohanka, M. (2020). D-Lactic Acid as a Metabolite: Toxicology, Diagnosis, and Detection. *BioMed Research International*, 2020, 1–9. <https://doi.org/10.1155/2020/3419034>
- Prasirtsak, B., Thitiprasert, S., Tolieng, V., Assabumrungrat, S., Tanasupawat, S., & Thongchul, N. (2017). Characterization of D-lactic acid, spore-forming bacteria and

- Terrilactibacillus laevilacticus SK5-6 as potential industrial strains. *Annals of Microbiology*, 67(11), Article 11. <https://doi.org/10.1007/s13213-017-1306-y>
- Punia Bangar, S., Suri, S., Trif, M., & Ozogul, F. (2022). Organic acids production from lactic acid bacteria: A preservation approach. *Food Bioscience*, 46, 101615. <https://doi.org/10.1016/j.fbio.2022.101615>
- Rajeshkumar, G., Arvinth Seshadri, S., Devnani, G. L., Sanjay, M. R., Siengchin, S., Prakash Maran, J., Al-Dhabi, N. A., Karuppiah, P., Mariadhas, V. A., Sivarajasekar, N., & Ronaldo Anuf, A. (2021). Environment friendly, renewable and sustainable poly lactic acid (PLA) based natural fiber reinforced composites – A comprehensive review. *Journal of Cleaner Production*, 310, 127483. <https://doi.org/10.1016/j.jclepro.2021.127483>
- Rawoof, S. A. A., Kumar, P. S., Vo, D.-V. N., Devaraj, K., Mani, Y., Devaraj, T., & Subramanian, S. (2021). Production of optically pure lactic acid by microbial fermentation: A review. *Environmental Chemistry Letters*, 19(1), 539–556. <https://doi.org/10.1007/s10311-020-01083-w>
- Russmayer, H., Marx, H., & Sauer, M. (2019). Microbial 2-butanol production with *Lactobacillus diolivorans*. *Biotechnology for Biofuels*, 12(1), 262. <https://doi.org/10.1186/s13068-019-1594-5>
- Sae-Tang, K., Bumrungham, P., Mhuanthong, W., Champreda, V., Tanapongpipat, S., Zhao, X.-Q., Liu, C.-G., & Runguphan, W. (2023). Engineering Flocculation for Improved Tolerance and Production of d-Lactic Acid in *Pichia pastoris*. *Journal of Fungi*, 9(4), 409. <https://doi.org/10.3390/jof9040409>
- Sheeladevi, A., & Ramanathan, N. (2010). *Lactic Acid Production Using Lactic Acid Bacteria under Optimized Conditions*. 2, 6.
- Silva, R. R. A., Marques, C. S., Arruda, T. R., Teixeira, S. C., & de Oliveira, T. V. (2023). Biodegradation of Polymers: Stages, Measurement, Standards and Prospects. *Macromol*, 3(2), Article 2. <https://doi.org/10.3390/macromol3020023>
- Singhvi, M. S., Zinjarde, S. S., & Gokhale, D. V. (2019). Polylactic acid: Synthesis and biomedical applications. *Journal of Applied Microbiology*, 127(6), 1612–1626. <https://doi.org/10.1111/jam.14290>
- Swetha, T. A., Bora, A., Mohanrasu, K., Balaji, P., Raja, R., Ponnuchamy, K., Muthusamy, G., & Arun, A. (2023). A comprehensive review on polylactic acid (PLA) – Synthesis, processing and application in food packaging. *International Journal of Biological Macromolecules*, 234, 123715. <https://doi.org/10.1016/j.ijbiomac.2023.123715>
- Tashiro, Y., Kaneko, W., Sun, Y., Shibata, K., Inokuma, K., Zendo, T., & Sonomoto, K. (2011). Continuous d-lactic acid production by a novelthermotolerant *Lactobacillus delbrueckii* subsp. *Lactis* QU 41. *Applied Microbiology and Biotechnology*, 89(6), 1741–1750. <https://doi.org/10.1007/s00253-010-3011-7>
- Trontel, A., & Batu, A. (2011). *Production of D- and L-Lactic Acid by Mono- and Mixed Cultures of Lactobacillus sp.* 9.
- Vishnu Prasad, J., Sahoo, T. K., Naveen, S., & Jayaraman, G. (2020). Evolutionary engineering of *Lactobacillus bulgaricus* reduces enzyme usage and enhances conversion of lignocellulosics to D-lactic acid by simultaneous saccharification and fermentation. *Biotechnology for Biofuels*, 13(1), 171. <https://doi.org/10.1186/s13068-020-01812-x>

WEB REFERENCES

<https://www.european-bioplastics.org/>

<https://renewable-carbon.eu/>

<https://www.databridgemarketresearch.com/reports>

<https://www.precedenceresearch.com/>

Chapter 5

Investigation of the capabilities of *Enterococcus faecalis* strains to use glycerol as alternative carbon source

Laura Troiani¹, Alessia Levante¹, Hannes Russmayer^{2,3}, Hans Marx², Erasmo Neviani¹,
Valentina Bernini¹, Camilla Lazzi¹, Michael Sauer^{2,4}

¹ Department of Food and Drug Sciences, Parma University, Parma, Italy

² Institute of Microbiology and Microbial Biotechnology, Department of Biotechnology,
University of Natural Resources and Life Sciences, Vienna, Austria

³ FermX GmbH, Vienna, Austria

⁴ OMV AG, Innovation & Technology, Department of Biotechnology, Vienna, Austria

Manuscript in preparation

ABSTRACT

The global shift towards a green economy, driven by environmental concerns, highlights the pillar role of biorefineries in sustainable development (SDG n.7 - 2030 Agenda). Biodiesel stands as an important eco-friendly alternative to fossil fuels, but its growing demand has led to a surplus of crude glycerol as a by-product. Glycerol holds great potential in various industries, albeit hindered by costly purification processes. Microbiology offers promising avenues for glycerol conversion. *Enterococcus faecalis*, among the lactic acid bacteria (LAB), has demonstrated the ability to metabolise glycerol into valuable products. Another by-product deriving from several technologies, that LAB are potentially able to use, is acetate.

This study aimed to identify glycerol and acetate utilisers among LAB strains from the University of Parma's Culture Collection (UPCC) and evaluate their potential on these alternative carbon sources aiming at their conversion into valuable chemicals, such as lactate, useful in the production of bio-based plastics in line with the principles of the circular economy. 150 LAB strains were metabolically characterised through Biolog GENIII Microplates. After sequential selections, 8 strains were selected for cultivation in a DASGIP® parallel bioreactor system (Eppendorf International, Jülich, Germany) by setting a batch process. C sources consumption and metabolites production were quantified over time by HPLC analysis. Furthermore, *E. faecalis* UPCC 2443 and UPCC 2454 were selected to further investigate the co-utilisation of glycerol and acetate.

E. faecalis strains turned out to be more efficient in glycerol consumption than *Lactocaseibacillus* spp. ones and the shift in bioreactors helped to boost their potential. *E. faecalis* consumed four times the amount of glycerol and about five times the amount of acetate compared to the small scale and yielded over 60% of lactate in aerobic conditions, thus supporting further investigations in the two by-products recovery by *E. faecalis*.

KEYWORDS

Lactic acid bacteria, Bioreactors, *Enterococcus faecalis*, Glycerol, Biodiesel by-product

ABBREVIATIONS

LAB = Lactic Acid Bacteria

UPCC = University of Parma's Culture Collection

GRAS = Generally Recognized As Safe

Gly = Glycerol

Acet = Acetate

LA = Lactic Acid

EtOH = Ethanol

dhaK = Dehydrogenation Pathway

glpK = Phosphorylation Pathway

Nox = NADH oxidase

DHAP = Dihydroxyacetone Phosphate

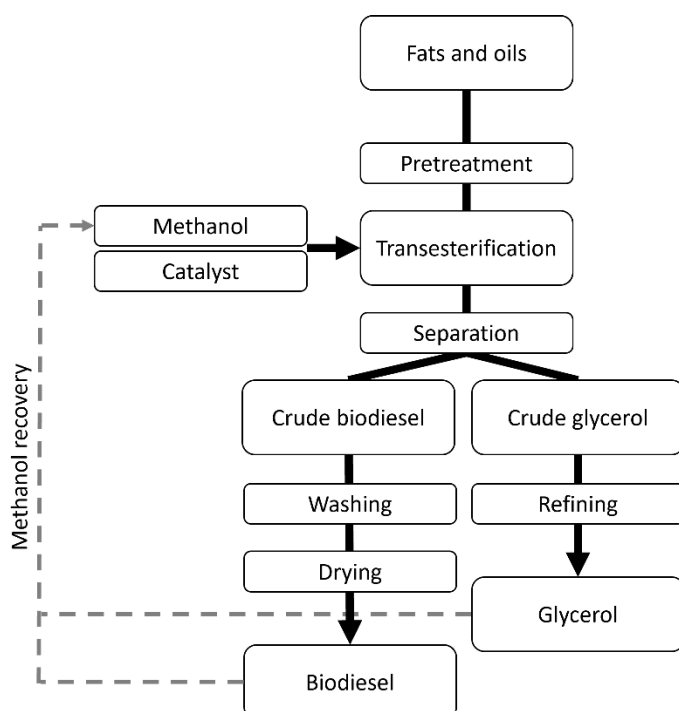
1. INTRODUCTION

Nowadays, the environmental impact of industrial processes on our lives and the transition to a green economy is a daily concern of governments, as well as of both the public and scientific community.

Biorefineries have become pillars of sustainable development: biofuels, with a global market size estimated at 116.46 billion US dollars in 2022 (<https://www.precedenceresearch.com/>, accessed in October 2023), are acquiring more and more importance until they will totally replace conventional fossil fuels (Goal n.7 of the Sustainable Development Goals Agenda 2030). Additionally, second-generation biofuels work as well as first-generation ones, but they derive from waste or non-edible feedstocks which are valorised without interfering with the human food supply chain (Bhuiya et al., 2016).

Biodiesel represents a valuable alternative to conventional oil-based fuels highly contributing to carbon dioxide emission reduction (Rodrigues et al., 2017). Its market size was evaluated at 39.25 billion US dollars in 2023 and is forecasted to increase up to 8.1% by 2032 (<https://www.precedenceresearch.com/>, accessed in October 2023). However, the biodiesel increasing demand has generated a glut of glycerol as a by-product (Fig. 1): the transesterification reaction yields approximately 10% (w/w) of crude glycerol, not easy to re-use as it is due to the low purity degree (Attarbach et al., 2023; Monteiro et al., 2018; Russmayer, Marx, et al., 2019).

Figure 1 Biodiesel production process from oleaginous waste.



The global glycerol market is forecasted to reach a volume of 3.3 billion by 2023. Only the EU, followed by the USA and Brazil, covers 26% of the global biodiesel market (Tabatabaei & Aghbashlo, 2018) and, in 2015, it modified the previous Directive EC/2009/28 to promote the use of energy from renewable sources (DIR. (EU) 2015/1513). Crude glycerol is second only to lignocellulosic biomass considering the global available amount. The global market value of glycerol reached 2.55 billion US dollars in 2021, expected to increase up to 6.45% by 2029 (<https://www.databridgemarketresearch.com/reports/>, accessed in October 2023).

Pure glycerol finds many applications in several industries (food, cosmetics, pharmaceuticals, etc.). However, the purification process is rather expensive (\$0.15/lb) with respect to the glycerol market price (\$0.05/lb), thus discouraging its recovery (Vivek et al., 2019). After several pre-treatments, through chemical, physical or even biological processes (Kaur et al., 2020), it can be addressed to biogas generation (He et al., 2017), livestock feeding (Lokesha et al., 2017; Martínez-Miró et al., 2021), production of high-value chemicals (Chen & Liu, 2016; Lima et al., 2022), discarded in landfills (Pitt et al., 2019) or incinerated. Researchers are investigating the improvement of purification procedures and the recovery of this valuable by-product for the production of value-added products, thus unlocking waste management and extending the ancillary market of biodiesel (Ben et al., 2022). The chemical industry is particularly engaged in the production of chemicals from renewable resources. In this context, microbiology and biotechnology offer interesting alternative solutions to chemical synthesis. Current advances in genomics and metabolic profiling can help to deeply investigate microbial metabolic pathways for the synthesis of new chemicals or the exploitation of specific metabolic mechanisms.

This study arose from the willingness to explore the biodiversity within the University of Parma's Culture Collection (UPCC) to find wild-type strains able to metabolise alternative C sources, thus leveraging microorganisms usually confined to food for practical and sustainable industrial purposes.

The UPCC is mainly composed of lactic acid bacteria (LAB), a well-known group of GRAS (Generally Recognized As Safe) microorganisms normally associated with food and feed products or human health, producing lactic acid as the main end-product of their metabolism. Lactic acid has gained great interest as a building block for polylactic acid (PLA), one of the most popular bioplastics currently available on the market (Balla et al., 2021). Approximately 90% of the global lactic acid is produced by biological method, and part of it is obtained from waste substrate, thus implementing the circular economy (Alexandri et al., 2019).

Glycerol, or Propane-1,2,3-triol (IUPAC), CAS number 56-81-5, is the trihydric alcohol backbone of lipids. It is a symmetrical, small, and uncharged molecule, with a high degree of reduction. Microorganisms can metabolise glycerol through a few pathways. During dissimilation, glycerol produces a surplus of reducing equivalents (e.g. NADH) compared to hexoses and pentoses, which is advantageous for reduced chemicals production like alcohols, used as biofuels and solvents (Russmayer, Egermeier, et al., 2019). How glycerol enters the cell depends on the microorganisms and the concentration of glycerol in the substrate. Then, it is used as a carbon source to grow through two metabolic pathways, depending on the environmental conditions: the phosphorylation pathway (*glpK*) is chosen in the presence of oxygen, while the dehydrogenation one (*dhaK*) is taken otherwise. Both pathways lead to the generation of dihydroxyacetone phosphate (DHAP) which gets into glycolysis (Vivek et al., 2019). Several microorganisms can use glycerol as the sole carbon feedstock (Costa-Gutierrez et al., 2021; Guerfali et al., 2020). Many of them even tolerate high concentrations of salt and alcohol thus becoming suitable to grow on crude glycerol, often rich in these impurities (Crosse et al., 2019). Some Enterobacteriaceae taxa are known to grow on glycerol and produce valuable metabolites from it (Vivek et al., 2019) since they possess both of the glycerol catabolic pathways mentioned above (Bizzini et al., 2010). In particular, *Enterococcus faecalis*, known as a member of human and animal gut microbiota, sometimes as an opportunistic pathogen (Doi, 2015; Kao & Kline, 2019), is one of the bacteria able to produce mainly lactic acid and ethanol from glycerol. However, only 142 articles were found searching on Scopus (<https://www.scopus.com/>, accessed in October 2023) by using (TITLE-ABS-KEY (enterococcus-faecalis) AND TITLE-ABS-KEY (glycerol)) AND (LIMIT-TO (DOCTYPE, "ar")) from 2000 to 2023, also considering that a good part of them are human studies, thus off-topic for this research. Moreover, there are even fewer papers on the use of glycerol under aerobic conditions, which means that this point was not really investigated.

Moreover, in the perspective of exploiting the metabolic potential of LAB's primary metabolism, we also investigated the utilisation of acetate. A previous study (data not published), indeed, interestingly showed metabolic activity of some *Enterococcus* spp. strains on acetate, in that case in the form of acetic acid. Acetate is the end-product or a by-product of several technologies (microbial gas fermentation, lignocellulose depolymerization, microbial photosynthesis, microbial electrosynthesis, or methane's anaerobic oxidation) and it could represent an alternative carbon source for industrial purposes. Similar to glycerol, acetate can be valorised through the conversion into value-added bioproducts, for instance in the bioplastic

field (PHAs) (Kiefer et al., 2021). However, the scientific literature about the utilisation of acetate by LAB is scarce, and an in-depth investigation is worthwhile.

The present work focused on the screening of a selected core of LAB to identify glycerol utilisers, with concerns also to acetate utilisation, and evaluate their potential on these alternative carbon sources with the final aim of producing lactic acid, as a valuable chemical with application in the bioplastic industry, thus implementing a circular economy model.

2. MATERIAL AND METHODS

2.1. Primary screening

The primary screening was conducted on 150 lactic acid bacteria (LAB) strains, belonging to 29 different species, previously selected from the University of Parma's Culture Collection (UPCC) to evaluate their metabolic capabilities by using the Biolog GENIII Microplates (BIOLOG Inc.©, USA). Strains maintained at -80°C in their optimal growth medium (MRS or TSB (VWR International bvba, Leuven, Belgium), or M17 (Merck, Darmstadt, Germany)) containing 15% glycerol (w/v), were propagated three times in a fresh medium prior use. For each overnight culture, 1 ml was aliquoted, and the biomass was harvested by centrifugation at 13400 rpm for 10 min. Cells were washed in Ringer's solution (VWR International bvba, Leuven, Belgium) twice, and suspended in 12 ml of inoculation fluid IF-A (BIOLOG Inc.©, USA) to a transmittance range of 90-98%. Each GENIII Microplate was filled out with the inoculation fluid and incubated at the optimal temperature of the strain. Before incubation and after 72 h, the metabolic fingerprint was determined on the BIOLOG MicroReader Station (BIOLOG Inc.©, USA) by absorbance reading at a wavelength of 590 nm. The absorbance is relative to the Tetrazolium Violet dye included in each well that changes colour (colourless to purple) when metabolic activity occurs. Data obtained were normalized by the Average Well Colour Development (Eq. 1, where OD_i = raw OD_{590} value of glycerol well, R = OD_{590} value of negative control well, n = number of substrates analysed (71)) (Acin-Albiac et al., 2020), and strains able to metabolise glycerol were selected for the secondary screening. The selection process also considered results from a previous study (data not published) that noticed even the utilisation of acetate by LAB.

$$AWCD = \sum_{i=1}^n (OD_i - R)/n \quad (1)$$

2.2. Secondary screening

Selected strains were cultivated on a small scale to confirm the ability to use glycerol. Precultures were performed on MRS medium (pH = 6.2) containing per litre: 10 g casein peptone, 10 g meat extract, 5 g yeast extract, 1 g tween 80, 2 g K_2HPO_4 , 5 g sodium acetate, 2.60 g tri-sodiumcitrate $\cdot 2\text{H}_2\text{O}$, 1.17 g $\text{NH}_4(\text{HPO}_4)_2$, 0.20 g $\text{MgSO}_4\cdot 7\text{H}_2\text{O}$, 0.05 g $\text{MnSO}_4\cdot \text{H}_2\text{O}$ and 22 g glucose $\cdot \text{H}_2\text{O}$. Subsequently, the biomass was harvested and washed twice in Ringer's solution to remove glucose traces. Cultures were carried out on MRS by replacing glucose with glycerol (22 g/L). 6 ml of broth were inoculated to an OD_{600} of 1 and incubated at 30°C for 72

h in both aerobic and anaerobic conditions (AnaeroGen, ThermoScientific, Landsmeer, The Netherlands). The OD₆₀₀ was measured every 24 hours on a spectrophotometer (Hitachi U-1900). Supernatants were centrifuged at 13400 rpm for 10 min and components (glycerol, acetate, and lactate) were quantified by HPLC analysis (Shimadzu, Korneuburg, Austria). The HPLC was equipped with an Aminex HPX-87H column (300 × 7.8 mm, Biorad) operated at a temperature of 60 °C, a flow of 0.6 ml/min and by using 4 mM H₂SO₄ as mobile phase. Samples and standards for HPLC analysis consisted of 900 µL of sample/standard added with 100 µL 40 mM H₂SO₄. Subsequently, samples and standards were filtered. For the detection of molecules, a refraction index detector (RID-10A, Shimadzu, Korneuburg, Austria) was used. 8 strains were selected for cultivation in bioreactors.

2.3. Scaling-up in bioreactors

Pre-cultures were prepared to inoculate a batch with a specific OD₆₀₀. Fermentations in a DASGIP® parallel bioreactor system (Eppendorf International, Jülich, Germany) were set up as suggested by Russmayer et al. (2019). A batch process was carried out: initial volume of 700 ml MRS containing 22 g/L glycerol and 5 g/L sodium acetate as carbon source (Medium A), inoculum of OD₆₀₀ = 0.2, constant pH = 6.2 (neutralization through addition of 4M NaOH), temperature at 30°C, aerobic atmosphere (O₂ flux = 2 sl/h, DO = 30%, [O₂] = 21%, [CO₂] = 0%), agitation at 200 rpm.

Based on the increased growth associated with the higher consumption of glycerol and acetate, two strains (*Enterococcus faecalis* UPCC 2443 and *Enterococcus faecalis* UPCC 2454) were selected for further investigation on the co-use of the two carbon sources. To estimate the capability of strains to use acetate simultaneously with glycerol, the glycerol-to-acetate molar ratio was decreased from 4:1 (Medium A) to 2:1 (Medium B) and 1:1 (Medium C), maintaining unchanged the total amount of carbon moles. Fermentations were repeated in the new media, microbial growth was measured by OD₆₀₀ readings, and glycerol, acetate, lactate, and ethanol were evaluated through HPLC analysis. The lactic acid yield (mol_{LA}/mol_{Gly}), the total yield (mol/ mol) and the balance of carbon (Cmol/Cmol) were calculated as follows (Murakami et al., 2016):

$$\text{Lactic acid yield} = (g_{LA}/90.08)/(g_{Gly}/92.09) \quad (2)$$

$$\text{Total yield} = (g_{LA}/90.08 + g_{etOH}/46.07)/(g_{Gly}/92.09 + g_{Acet}/60.05) \quad (3)$$

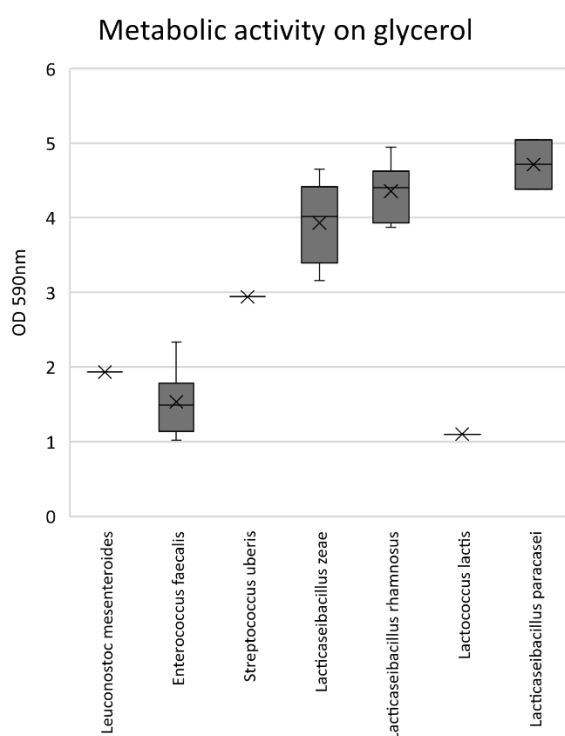
$$\text{C balance} = (g_{LA}/90.08*3 + g_{etOH}/46.07*2)/(g_{Gly}/92.09*3 + g_{Acet}/60.05*2) \quad (4)$$

3. RESULTS

3.1. Primary and secondary screening

From the phenotyping, 27 strains were able to metabolise glycerol (normalized $OD_{600} > 1$) after 72 h: 7 *Enterococcus faecalis*, 2 *Lacticaseibacillus paracasei*, 7 *Lcb. rhamnosus*, 8 *Lcb. zeae*, 1 *Lactococcus lactis*, 1 *Leuconostoc mesenteroides*, 1 *Streptococcus uberis* (Fig. 2).

Figure 2 Glycerol assimilation efficiency represented by the absorbance intensity (OD_{590}) of the Tetrazolium Violet dye included in the Biolog GENIII Microplates. Data normalized by AWCD.



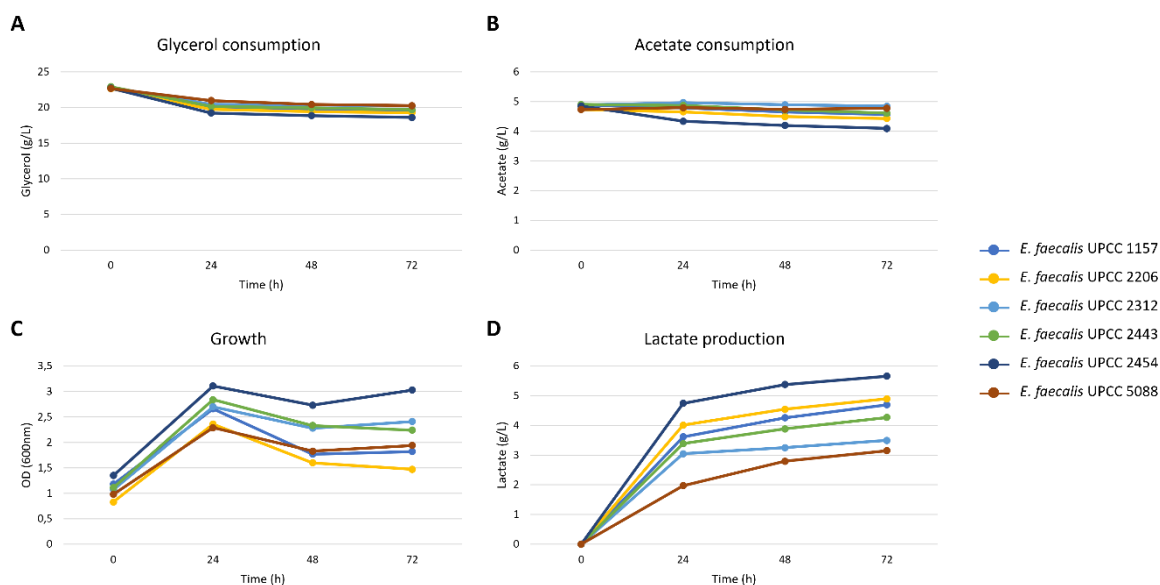
The medium A used in both tubes and bioreactor tests mainly contained carbon in the form of glycerol, but a small amount of it derived also from sodium acetate, thus obtaining a glycerol-to-acetate molar ratio of 4:1. From the Biolog assays, *Lacticaseibacillus* spp. strains reduced the dye into the glycerol's well stronger than other species, meaning a great metabolic activity. Nevertheless, moving cultures on tubes, among the 27 selected strains, 6 belonging to *E. faecalis* species showed to actively use both C sources (Tab. 1). Their behaviour during fermentation is displayed in Figure 3: glycerol and acetate consumption (g/L), growth (OD_{600nm}) and lactate production (g/L) were analysed for a further selection of strains.

Table 1 Best performers in the small-scale fermentations. The glycerol and acetate consumption (g/L) at the end of the fermentation (72h) is reported as the percentage of the initial amount supplied. The maximum growth (OD₆₀₀) was reached at 24 h.

ID strain	Species	Glycerol consumed at 72h (%)	Acetate consumed at 72h (%)	Max OD _{600nm} (24h)
1157	<i>Enterococcus faecalis</i>	14.43	6.52	2.66
2206	<i>Enterococcus faecalis</i>	14.98	6.42	2.36
2312	<i>Enterococcus faecalis</i>	11.49	0.84	2.70
2443	<i>Enterococcus faecalis</i>	14.06	5.96	2.84
2454	<i>Enterococcus faecalis</i>	18.07	15.53	3.11
5088	<i>Enterococcus faecalis</i>	10.66	0	2.29

After 72 h, *E. faecalis* UPCC 2454 and UPCC 2443 were the most promising strains as they grew better than others, reaching OD₆₀₀ of 3.11 and 2.84, respectively. Furthermore, the first one most efficiently used both glycerol and acetate (4.11 g/L and 0.75 g/L, respectively).

Figure 3 Small-scale fermentations. A) glycerol consumption (g/L), B) acetate consumption (g/L), C) growth (OD₆₀₀), and D) lactate production (g/L) were monitored until 72 h.

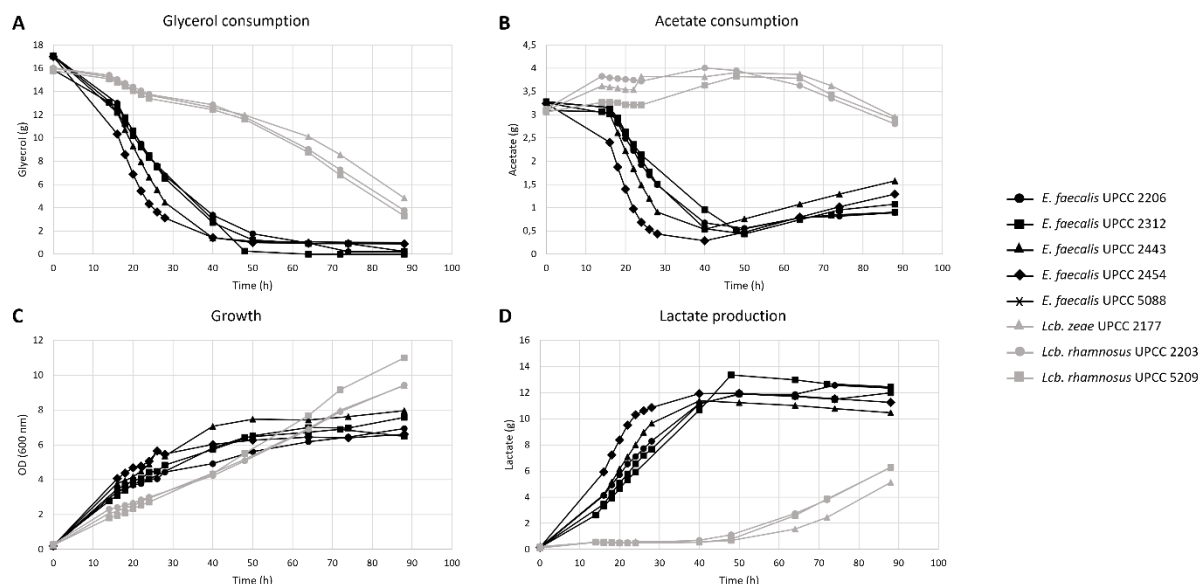


3.2. Scaling-up in bioreactors

Despite *Lacticaseibacillus* spp. strains did not have great performance in the small-scale tests, they were still grown in bioreactors to evaluate their reaction in a more controlled environment. Therefore, for the next step, 8 strains were selected: *Enterococcus faecalis* UPCC 2206, UPCC

2312, UPCC 2443, UPCC 2454, UPCC 5088, *Lacticaseibacillus zeae* UPCC 2177, *Lacticaseibacillus rhamnosus* UPCC 2203 and UPCC 5209 (Fig. 4).

Figure 4 Fermentations in bioreactors. During the process, four parameters were monitored: A) glycerol consumption (g); B) acetate consumption (g); C) cellular growth (OD₆₀₀); D) lactate production (g). Black lines, *E. faecalis* strains; grey lines, *Lacticaseibacillus* spp. strains.



E. faecalis strains started to consume glycerol (Fig. 4A) and acetate (Fig. 4B) earlier and faster than *Lacticaseibacillus* spp. ones, thus growing (Fig. 4C) and producing lactate (Fig. 4D) before the latter. The fermentation of *E. faecalis* took about 50 h, whereas *Lacticaseibacillus* spp. strains kept working over 88 h. Despite *E. faecalis* strains reached the stationary phase earlier than *Lacticaseibacillus* spp. ones, they consumed more glycerol and therefore produced more lactic acid.

Considering the fastest consumption of both C sources, *E. faecalis* UPCC 2443 and *E. faecalis* UPCC 2454 were newly selected for the next evaluation (Tab. 2).

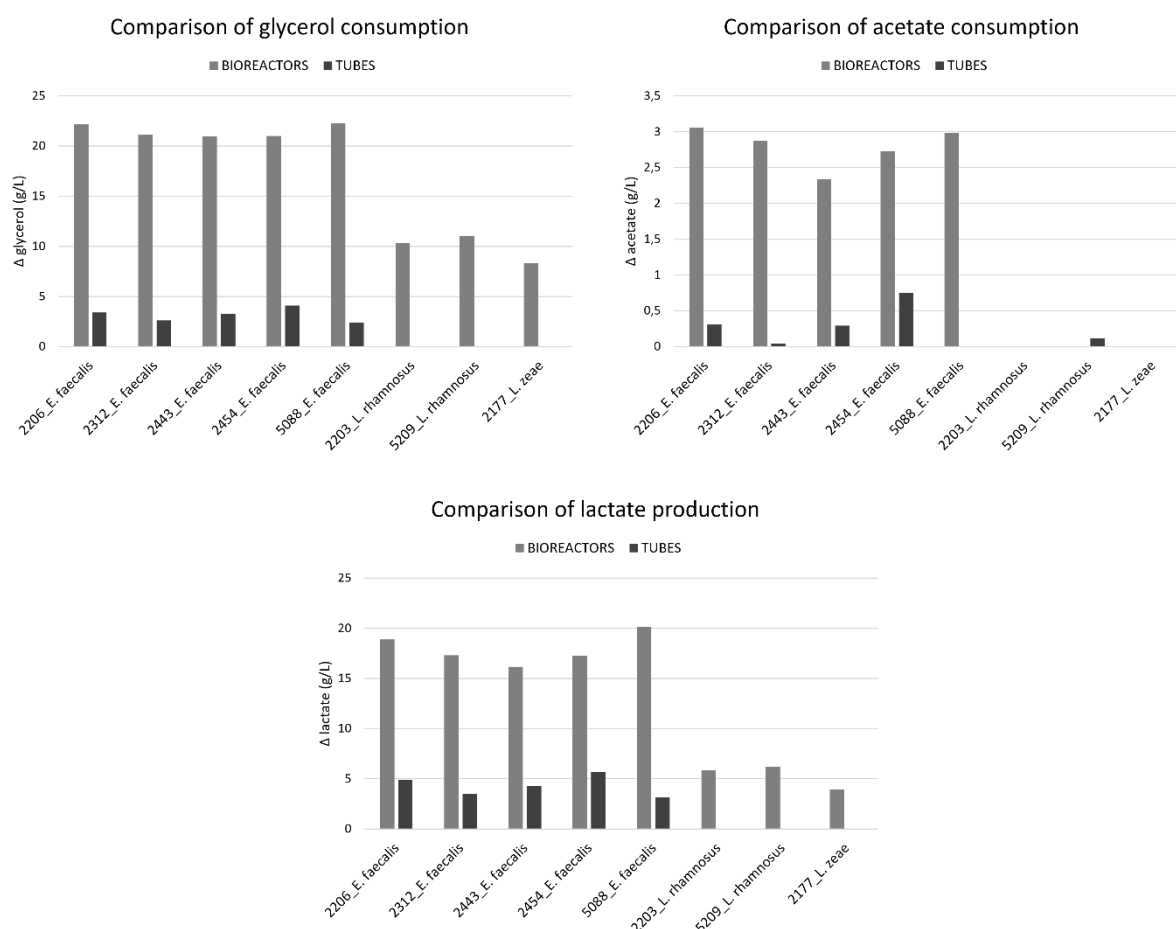
Table 2 Glycerol and acetate consumption (percentage and max rate) during the batch fermentation in bioreactors.

ID strain	Species	Glycerol consumed at 88h (%)	Max glycerol consumption rate (g/h)	Acetate consumed at 88h (%)	Max acetate consumption rate (g/h)
2206	<i>E. faecalis</i>	98.71	0.37	72.78	0.06
2312	<i>E. faecalis</i>	98.55	0.38	67.18	0.07
2443	<i>E. faecalis</i>	94.12	0.45	51.62	0.08

2454	<i>E. faecalis</i>	94.83	0.53	60.13	0.11
5088	<i>E. faecalis</i>	100	0.32	70.85	0.05
2177	<i>Lcb. zeae</i>	69.75	0.13	4.45	0.002
2203	<i>Lcb. rhamnosus</i>	76.55	0.14	10.30	0.004
5209	<i>Lcb. rhamnosus</i>	79.07	0.14	5.09	0.002

It is noteworthy that cultivation in bioreactors and controlled growth conditions helped anyhow all strains to make better use of the C sources provided. Looking at Figure 5, the difference in microbial performances between the two experimental scales is clear.

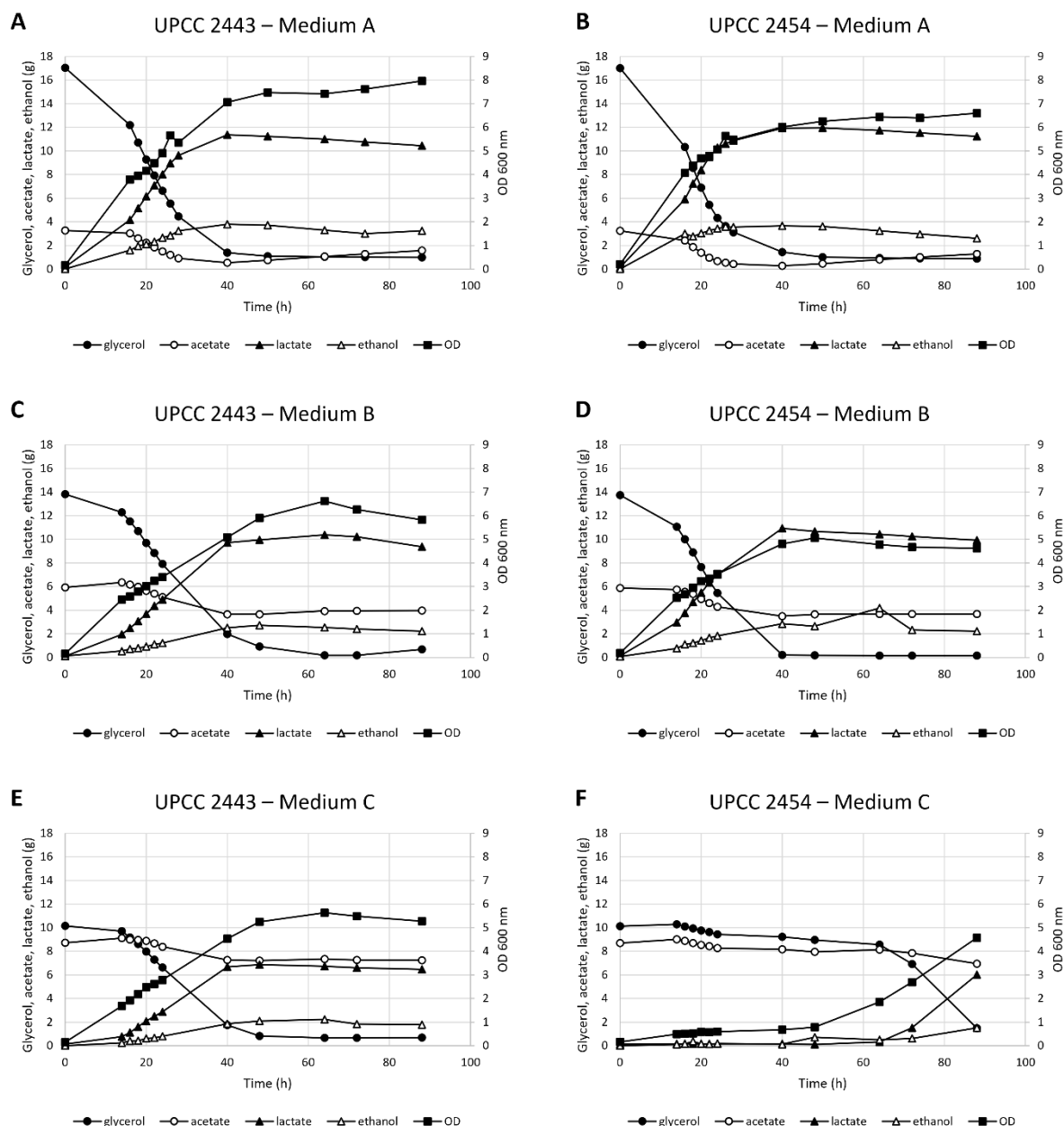
Figure 5 Comparison between small- and large-scale experiments in terms of glycerol and acetate consumption, and lactate production.



3.3. Investigation of the glycerol-acetate co-utilisation

The capability of strains to use acetate simultaneously with glycerol was evaluated by changing the carbon molar ratio of the two carbon sources in the medium (Fig. 6).

Figure 6 Batch process in bioreactors of *E. faecalis* UPCC 2443 and *E. faecalis* UPCC 2454. Three different media were employed: A-B) Medium A; C-D) Medium B; E-F) Medium C. During 88 hours of fermentation, four parameters were monitored: ●) glycerol consumption (g); ○) acetate consumption (g); ▲) lactate production (g); △) ethanol production (g); ■) cellular growth (OD₆₀₀).



Lowering the glycerol-to-acetate molar ratio, the biomass and lactate production and the general performance of strains decreased (Fig. 6E-F) despite the availability of the same total amount of C moles.

In all media, *E. faecalis* UPCC 2443 produced a higher amount of biomass than *E. faecalis* UPCC 2454. The maximum consumption rates of both C sources are reported in Table 3. *E. faecalis* UPCC 2443 reached its maximum glycerol and acetate consumption rate at 28 h in the Medium A, and at 40 h in the Medium B and C. On the other hand, *E. faecalis* UPCC 2454 showed its maximum glycerol and acetate consumption rate at 24 h in the Medium A and B, and at the end of the fermentation in the Medium C. In this latter case, it is interesting to observe the kinetic of *E. faecalis* UPCC 2454 (Fig. 6F): it took a long time to adapt to the Medium C, thus starting its exponential growth phase around 50 h and, accordingly, delaying its metabolic activity.

Table 3 Glycerol and acetate max consumption rates (g/h) and co-utilisation index calculated as $\text{Max } \Delta \text{ glycerol consumption} / \text{Max } \Delta \text{ acetate consumption}$.

Microorganism	Medium	Max glycerol consumption rate (g/h)	Max acetate consumption rate (g/h)	Co-utilisation index
<i>E. faecalis</i> UPCC 2443	A (4:1)	0.45	0.08	5.89
	B (2:1)	0.30	0.06	5.97
	C (1:1)	0.21	0.04	6.37
<i>E. faecalis</i> UPCC 2454	A (4:1)	0.53	0.11	5.46
	B (2:1)	0.34	0.07	5.71
	C (1:1)	0.10	0.02	5.52

Considering the glycerol and acetate maximum consumption during fermentation, the co-utilisation index is given by the ratio between glycerol and acetate consumption. The lower the co-utilisation index, the more efficient the simultaneous use of the two carbon sources, as recorded in the Medium A (Tab. 3). Both strains also utilised the C sources more rapidly in the Medium A than in the others.

Table 4 Glycerol-acetate fermentation yields: C sources consumption (g), metabolites production (g), lactic acid yield ($\text{mol}_{\text{LA}}/\text{mol}_{\text{Gly}}$), total yield (mol/mol), and balance of carbon (C mol/C mol) at 88 h.

Microorganism	Medium	Consumption (g)		Production (g)		LA yield	Total yield	C balance
		Gly	Acet	LA	EtOH			
<i>E. faecalis</i> UPCC 2443	A (4:1)	16.04	1.68	10.29	3.25	0.65	0.91	0.84
	B (2:1)	13.12	1.95	9.26	2.08	0.72	0.84	0.81
	C (1:1)	9.48	1.48	6.33	1.77	0.68	0.85	0.80
<i>E. faecalis</i> UPCC 2454	A (4:1)	16.13	1.95	11.10	2.61	0.70	0.87	0.82
	B (2:1)	13.57	2.20	9.78	2.14	0.74	0.84	0.81
	C (1:1)	8.63	1.75	5.89	1.51	0.70	0.80	0.79

At the end of the fermentation (88 h), the metabolism products were measured to evaluate the microorganisms' potentialities. As shown in Table 4, even increasing the acetate concentration in the medium, its utilisation stayed unvaried. The highest amount of both metabolites was found in the Medium A where the concentration of glycerol was higher, in accordance with the highest total yield. Accordingly, the lower the glycerol concentration, the lower both the lactate and ethanol production. The estimated carbon balance was equal in all three media.

4. DISCUSSIONS

This work focused on the capability of UPCC strains of assimilating alternative carbon sources to glucose, such as glycerol and acetate, to select promising LAB for the recovery of these by-products.

The primary screening notified 27 LAB strains showing metabolic activity on glycerol, with different efficiency levels. This selection also considered results from a previous study (not published) on the use of acetate by these strains. According to the method used, the assimilation efficiency of a given C source is represented by the degree of reduction of the Tetrazolium Violet dye included in the Biolog GENIII Microplate wells. The reduction of the dye is caused by a flux of electrons which is directly correlated to the metabolic activity of a microorganism. However, it is not always correlated to microbial growth: strains can survive on a medium, but not necessarily produce biomass (Mekonnen et al., 2019). It is important to point out that this assay is meant to be considered qualitative rather than quantitative: it drives the selection of microorganisms but validation on a larger scale is needed. The results support this theory: among the 27 strains, 17 belonging to the genus *Lacticaseibacillus* strongly reduced the dye showing an average normalized absorbance value of 4.20, while *E. faecalis* strains gave values equal to less than half the values of the former (average normalized OD₆₀₀ = 1.53) (Fig. 2).

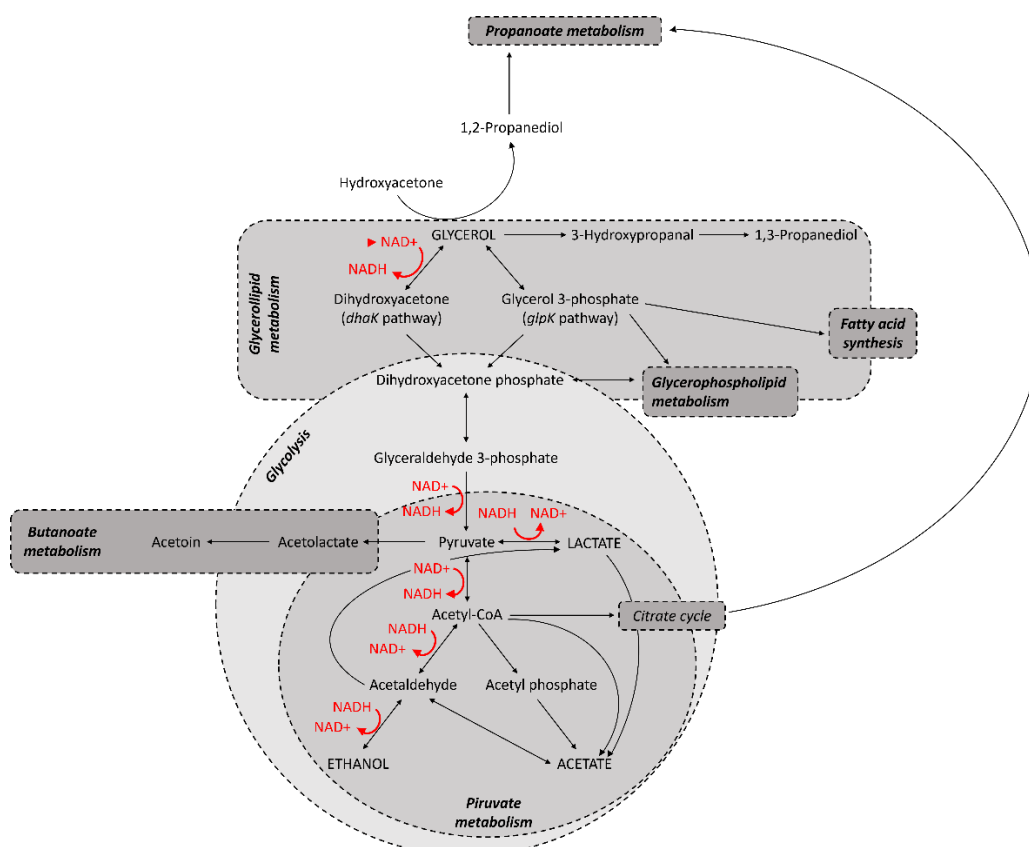
However, when cultivated in tubes using mostly glycerol as a carbon source, *Lacticaseibacillus* spp. strains did not grow well, and no glycerol decrease was observed in the medium. Contrarily, *E. faecalis* strains were metabolically active, producing biomass and lactate although using little amount of glycerol (Tab. 1; Fig. 3). In particular, *E. faecalis* UPCC 2454 was the most performing strain of the screening since it reached the highest OD₆₀₀ value (3.11) and the greatest lactic acid production (5.66 g/L after 72h) by metabolising 18.07% glycerol and 15.53% acetate.

To in-depth understand the aptitude of our LAB, 8 selected cultures were scaled up in bioreactors. The controlled environment of the batch process also favoured the growth of *Lacticaseibacillus* spp. strains (Fig. 4C). Nevertheless, although they reached higher OD₆₀₀ values at the end of the fermentation than *E. faecalis* strains, they were very slow to adapt. Hence, they consumed less glycerol and consequently produced less lactate (Fig. 4A, 4D). In bioreactors (Tab. 2), all *E. faecalis* strains consumed more than 90% of the glycerol supplied with consumption rates up to 0.53 g/h (UPCC 2454). They also increased the acetate assimilation, consuming an average of 64.5%.

The secondary screening has highlighted two opposite kinetics between the two taxonomic groups, both in tubes and bioreactors. *E. faecalis* strains were confirmed to be the most efficient in the utilisation of glycerol, therefore the study kept focusing on their behaviour. However, the scaling improved the overall performance of all strains analysed as shown in Figure 5. *E. faecalis* consumed four times the amount of glycerol and about five times the amount of acetate compared to the tube scale. They also produced more than three times the amount of lactic acid. Concerning *Lacticaseibacillus* spp. strains, despite they never achieved the performance of *E. faecalis*, cultivation in bioreactors allowed them to grow and produce lactate differently from what happened in tubes.

Murakami et al. (2016) investigated the glycerol metabolism in *E. faecalis* looking at the redox balance during the lactic acid fermentation. *E. faecalis* can metabolise glycerol through two different pathways, the phosphorylation pathway (*glpK*) and the dehydrogenation pathway (*dhaK*), both leading to the formation of dihydroxyacetone phosphate (DHAP) (Doi, 2018). Results from the literature (Bizzini et al., 2010; Sauvageot et al., 2012) strongly suggest that under aerobic conditions *E. faecalis* uses both pathways to catabolize glycerol, whereas under anaerobic conditions it only uses the *dhaK* pathway.

Figure 7 Different metabolic pathways involving glycerol. Figure adapted by (Doi, 2019; Murakami et al., 2016; <https://www.kegg.jp/pathway/>).



As shown in Figure 7, when *E. faecalis* uses glycerol under aerobic conditions, along the conversion to pyruvate, passing through the *dhaK* pathway, two NADH molecules are produced and then re-oxidized by NADH oxidase (Nox) consuming one oxygen molecule (Blanco & Blanco, 2017; Dobson et al., 2012; Doi, 2019; Murakami et al., 2016; Sauvageot et al., 2012; Vivek et al., 2019). Therefore, aerobiosis favours the expression of the Nox and thus a higher consumption of glycerol through the *dhaK* pathway. This is confirmed by our results since the secondary screening on the small scale. Indeed, all strains studied worked more efficiently in the presence of oxygen than in its absence, as proved by the highest production of lactate and ethanol. Accordingly, the following experiments were carried out in aerobic conditions, where yields were much higher. The type of metabolites derived from pyruvate depends on the intracellular NADH/NAD⁺ ratio, therefore, as results suggest, the lactate production, despite being the main metabolic pathway, can be attributed to other pathways when glycerol is metabolised in aerobic culture conditions (Doi, 2018). During lactate production, only one NADH molecule is oxidized. Thus, the excess NADH is then oxidized during the ethanol production to restore the redox balance. It is noteworthy that, during the fermentation (Fig. 4B), the acetate firstly decreases and then increases, thus lowering the consumption rate. An explanation can be given by Murakami et al. (2016): the high NADH/NAD⁺ ratio during glycerol metabolism activates also the acetic acid metabolism to maintain the redox balance. However, acetate is also involved in other intermediate reactions, and we believe that part of the acetyl-CoA generated from acetate to produce ethanol goes back to acetate itself.

The secondary screening ended with the selection of two strains (*E. faecalis* UPCC 2243 and *E. faecalis* UPCC 2454) to further investigate the occurrence of a glycerol-acetate co-utilisation. They were newly cultivated in media with different glycerol-to-acetate molar ratios. The higher the glycerol-to-acetate molar ratio (4:1), the better they consume acetate (co-utilisation index equal to 5.89 for UPCC 2243 and to 5.46 for UPCC 2454) (Tab. 3). The overall performance of both strains is better in the Medium A, probably because in the other media there is a partial inhibition of bacteria caused by the increasing concentration of sodium since acetate has been added to the media in the form of sodium acetate. Differences in metabolite production are particularly evident in the Medium C: the low availability of glycerol reduced its conversion into end-products. Glycerol and acetate are utilised simultaneously by strains (Fig. 6), but independently from each other. Cultures stop using acetate as soon as glycerol is consumed, even increasing the acetate availability. Additionally, glycerol and acetate are not proportionally consumed (Tab. 4): while the consumption of glycerol follows its decreasing availability, the acetate assimilation stays rather stable despite its increasing concentration from

medium A to C. Moreover, lactic acid and ethanol production follow the same proportional trend of glycerol. This, coupled with the stable LA yield among the three media, suggests that the fermentation is entirely guided by glycerol catabolism. The reduced performance observed in lowering the glycerol-to-acetate molar ratio is caused not only by the increasing concentration of sodium in the media as assumed above, but mainly by the lower glycerol concentration. However, the C balance (> 0.80) roughly calculated without considering the biomass or CO_2 produced, suggests an efficient utilisation of the carbon during the fermentation by limiting the carbon losses.

5. CONCLUSIONS

In conclusion, this work contributes to extending the scientific literature on glycerol and acetate utilisation by LAB, with a particular focus on *Enterococcus faecalis* strains.

Outcomes from the primary and secondary screening confirm the effective ability of this species to utilise glycerol and produce an acceptable amount of lactic acid under aerobic conditions. On the other hand, strains did not demonstrate to use acetate to produce lactate, even if cultivation in bioreactors improved its consumption.

The controlled environment of the fermenters (pH, temperature, atmosphere, and substrate concentration) allows us to obtain reliable and reproducible results for potential industrial applications. This research is a preliminary investigation of the UPCC strains' potentialities and further studies may be directed towards improving the ability of *Enterococcus faecalis* to efficiently use crude glycerol. The successful recovery of glycerol represents an opportunity to overcome waste management issues and produce high-value chemicals, thus widening the market of biodiesel and circularly closing its production process.

REFERENCES

- Alexandri, M., Schneider, R., Mehlmann, K., & Venus, J. (2019). Recent Advances in D-Lactic Acid Production from Renewable Resources: Case Studies on Agro-Industrial Waste Streams. *Food Technology and Biotechnology*, 57(3), 293–304. <https://doi.org/10.17113/ftb.57.03.19.6023>
- Attarbach, T., Kingsley, M. D., & Spallina, V. (2023). New trends on crude glycerol purification: A review. *Fuel*, 340, 127485. <https://doi.org/10.1016/j.fuel.2023.127485>
- Balla, E., Daniilidis, V., Karlioti, G., Kalamas, T., Stefanidou, M., Bikiaris, N. D., Vlachopoulos, A., Koumentakou, I., & Bikiaris, D. N. (2021). Poly(lactic Acid): A Versatile Biobased Polymer for the Future with Multifunctional Properties—From Monomer Synthesis, Polymerization Techniques and Molecular Weight Increase to PLA Applications. *Polymers*, 13(11), 1822. <https://doi.org/10.3390/polym13111822>
- Ben, Z. Y., Samsudin, H., & Yhaya, M. F. (2022). Glycerol: Its properties, polymer synthesis, and applications in starch based films. *European Polymer Journal*, 175, 111377. <https://doi.org/10.1016/j.eurpolymj.2022.111377>
- Bhuiya, M. M. K., Rasul, M. G., Khan, M. M. K., Ashwath, N., & Azad, A. K. (2016). Prospects of 2nd generation biodiesel as a sustainable fuel—Part: 1 selection of feedstocks, oil extraction techniques and conversion technologies. *Renewable and Sustainable Energy Reviews*, 55, 1109–1128. <https://doi.org/10.1016/j.rser.2015.04.163>
- Bizzini, A., Zhao, C., Budin-Verneuil, A., Sauvageot, N., Giard, J.-C., Auffray, Y., & Hartke, A. (2010). Glycerol Is Metabolized in a Complex and Strain-Dependent Manner in *Enterococcus faecalis*. *Journal of Bacteriology*, 192(3), 779–785. <https://doi.org/10.1128/JB.00959-09>
- Blanco, A., & Blanco, G. (2017). Chapter 14—Carbohydrate Metabolism. In A. Blanco & G. Blanco (Eds.), *Medical Biochemistry* (pp. 283–323). Academic Press. <https://doi.org/10.1016/B978-0-12-803550-4.00014-8>
- Chen, Z., & Liu, D. (2016). Toward glycerol biorefinery: Metabolic engineering for the production of biofuels and chemicals from glycerol. *Biotechnology for Biofuels*, 9(1), 205. <https://doi.org/10.1186/s13068-016-0625-8>
- Costa-Gutierrez, S. B., Saez, J. M., Aparicio, J. D., Raimondo, E. E., Benimeli, C. S., & Polti, M. A. (2021). Glycerol as a substrate for actinobacteria of biotechnological interest: Advantages and perspectives in circular economy systems. *Chemosphere*, 279, 130505. <https://doi.org/10.1016/j.chemosphere.2021.130505>
- Crosse, A. J., Brady, D., Zhou, N., & Rumbold, K. (2019). Biodiesel’s trash is a biorefineries’ treasure: The use of “dirty” glycerol as an industrial fermentation substrate. *World Journal of Microbiology and Biotechnology*, 36(1), 2. <https://doi.org/10.1007/s11274-019-2776-9>
- DIRECTIVE (EU) 2015/1513 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL - of 9 September 2015—Amending Directive 98/70/ EC relating to the quality of petrol and diesel fuels and amending Directive 2009/28/ EC on the promotion of the use of energy from renewable sources.* (n.d.).
- Dobson, R., Gray, V., & Rumbold, K. (2012). Microbial utilization of crude glycerol for the production of value-added products. *Journal of Industrial Microbiology & Biotechnology*, 39(2), 217–226. <https://doi.org/10.1007/s10295-011-1038-0>

- Doi, Y. (2018). Lactic acid fermentation is the main aerobic metabolic pathway in *Enterococcus faecalis* metabolizing a high concentration of glycerol. *Applied Microbiology and Biotechnology*, 102(23), 10183–10192. <https://doi.org/10.1007/s00253-018-9351-4>
- Doi, Y. (2019). Glycerol metabolism and its regulation in lactic acid bacteria. *Applied Microbiology and Biotechnology*, 103(13), 5079–5093. <https://doi.org/10.1007/s00253-019-09830-y>
- Doi, Y. (2015). *l-Lactate Production from Biodiesel-Derived Crude Glycerol by Metabolically Engineered Enterococcus faecalis: Cytotoxic Evaluation of Biodiesel Waste and Development of a Glycerol-Inducible Gene Expression System*. <https://doi.org/10.1128/AEM.03418-14>
- Guerfali, M., Ayadi, I., Sassi, H.-E., Belhassen, A., Gargouri, A., & Belghith, H. (2020). Biodiesel-derived crude glycerol as alternative feedstock for single cell oil production by the oleaginous yeast *Candida viswanathii* Y-E4. *Industrial Crops and Products*, 145, 112103. <https://doi.org/10.1016/j.indcrop.2020.112103>
- He, Q. (Sophia), McNutt, J., & Yang, J. (2017). Utilization of the residual glycerol from biodiesel production for renewable energy generation. *Renewable and Sustainable Energy Reviews*, 71, 63–76. <https://doi.org/10.1016/j.rser.2016.12.110>
- Kao, P. H. N., & Kline, K. A. (2019). Dr. Jekyll and Mr. Hide: How *Enterococcus faecalis* Subverts the Host Immune Response to Cause Infection. *Journal of Molecular Biology*, 431(16), 2932–2945. <https://doi.org/10.1016/j.jmb.2019.05.030>
- Kaur, J., Sarma, A. K., Jha, M. K., & Gera, P. (2020). Valorisation of crude glycerol to value-added products: Perspectives of process technology, economics and environmental issues. *Biotechnology Reports*, 27, e00487. <https://doi.org/10.1016/j.btre.2020.e00487>
- Kiefer, D., Merkel, M., Lilge, L., Henkel, M., & Hausmann, R. (2021). From Acetate to Bio-Based Products: Underexploited Potential for Industrial Biotechnology. *Trends in Biotechnology*, 39(4), 397–411. <https://doi.org/10.1016/j.tibtech.2020.09.004>
- Lima, P. J. M., da Silva, R. M., Neto, C. A. C. G., Gomes e Silva, N. C., Souza, J. E. da S., Nunes, Y. L., & Sousa dos Santos, J. C. (2022). An overview on the conversion of glycerol to value-added industrial products via chemical and biochemical routes. *Biotechnology and Applied Biochemistry*, 69(6), 2794–2818. <https://doi.org/10.1002/bab.2098>
- Lokesha, E., Marappan, G., Ramasamy, D. K., Brijesh, P., Banakar, P., Revanasiddu, D., Pruthviraj, D., Bhanuprakash, V., & Mahesh, M. (2017). Crude glycerol: By-product of biodiesel industries as an alternative energy source for livestock feeding. *Journal of Experimental Biology and Agricultural Sciences*, 5(6), 755–766. [https://doi.org/10.18006/2017.5\(6\).755.766](https://doi.org/10.18006/2017.5(6).755.766)
- Martínez-Miró, S., Madrid, J., López, M. J., Orengo, J., Sánchez, C. J., & Hernández, F. (2021). Feeding Crude Glycerin to Finishing Iberian Crossbred Pigs: Effects on Growth Performance, Nutrient Digestibility, and Blood Parameters. *Animals : An Open Access Journal from MDPI*, 11(8), 2181. <https://doi.org/10.3390/ani11082181>
- Mekonnen, E., Kebede, A., Tafesse, T., & Tafesse, M. (2019). Investigation of carbon substrate utilization patterns of three ureolytic bacteria. *Biocatalysis and Agricultural Biotechnology*, 22, 101429. <https://doi.org/10.1016/j.bcab.2019.101429>
- Monteiro, M. R., Kugelmeier, C. L., Pinheiro, R. S., Batalha, M. O., & da Silva César, A. (2018). Glycerol from biodiesel production: Technological paths for sustainability.

- Renewable and Sustainable Energy Reviews*, 88, 109–122. <https://doi.org/10.1016/j.rser.2018.02.019>
- Murakami, N., Oba, M., Iwamoto, M., Tashiro, Y., Noguchi, T., Bonkohara, K., Abdel-Rahman, M. A., Zendo, T., Shimoda, M., Sakai, K., & Sonomoto, K. (2016). L-Lactic acid production from glycerol coupled with acetic acid metabolism by *Enterococcus faecalis* without carbon loss. *Journal of Bioscience and Bioengineering*, 121(1), 89–95. <https://doi.org/10.1016/j.jbiosc.2015.05.009>
- Pitt, F. D., Domingos, A. M., & Barros, A. A. C. (2019). Purification of residual glycerol recovered from biodiesel production. *South African Journal of Chemical Engineering*, 29, 42–51. <https://doi.org/10.1016/j.sajce.2019.06.001>
- Rodrigues, A., Bordado, J. C., & Santos, R. G. dos. (2017). Upgrading the Glycerol from Biodiesel Production as a Source of Energy Carriers and Chemicals—A Technological Review for Three Chemical Pathways. *Energies*, 10(11), Article 11. <https://doi.org/10.3390/en10111817>
- Russmayer, H., Egermeier, M., Kalemasi, D., & Sauer, M. (2019). Spotlight on biodiversity of microbial cell factories for glycerol conversion. *Biotechnology Advances*, 37(6), 107395. <https://doi.org/10.1016/j.biotechadv.2019.05.001>
- Russmayer, H., Marx, H., & Sauer, M. (2019). Microbial 2-butanol production with *Lactobacillus diolivorans*. *Biotechnology for Biofuels*, 12(1), 262. <https://doi.org/10.1186/s13068-019-1594-5>
- Sauvageot, N., Ladjouzi, R., Benachour, A., Rincé, A., Deutscher, J., & Hartke, A. (2012). Aerobic glycerol dissimilation via the *Enterococcus faecalis* DhaK pathway depends on NADH oxidase and a phosphotransfer reaction from PEP to DhaK via EIIADha. *Microbiology*, 158(10), 2661–2666. <https://doi.org/10.1099/mic.0.061663-0>
- Tabatabaei, M., & Aghbashlo, M. (2018). *Biodiesel: From Production to Combustion* (Vol. 8). Springer. <https://doi.org/10.1080/01430750.2022.2098175>
- Vivek, N., Hazeena, S. H., Rajesh, R. O., Godan, T. K., Anjali, K. B., Nair, L. M., Mohan, B., Nair, S. C., Sindhu, R., Pandey, A., & Binod, P. (2019). Genomics of Lactic Acid Bacteria for Glycerol Dissimilation. *Molecular Biotechnology*, 61(8), 562–578. <https://doi.org/10.1007/s12033-019-00186-2>

WEB REFERENCES

- <https://www.precedenceresearch.com/>
- <https://www.databridgemarketresearch.com/reports/>
- <https://www.scopus.com/>
- <https://www.kegg.jp/pathway/map=map00561&keyword=glycerol>

Conclusions

This Ph.D. project is aimed at valorising the invaluable microbial diversity preserved within the University of Parma's Culture Collection (UPCC), also through the implementation of a microbial platform to be used for the exploitation of lactic acid bacteria (LAB) in targeted applications, with a particular focus on sustainability. The present research enriches our knowledge of UPCC strains with in-depth insights into the functional capabilities of the selected LAB.

This project is based on the set-up of a phenotypic approach for the metabolic and physiological characterisation of a selected microbial core (SMC), and on the use of Artificial Intelligence for the prediction of microbial performance to allow the design of ad hoc processes. In addition to the knowledge gained on the possibility of fermenting a wide variety of substrates, even in non-optimal conditions, this study investigates the potentially valuable and feasible employment of LAB in the bioplastic industry and for the recovery of industrial waste and by-products, thus offering significant economic benefits and promoting sustainable development, looking towards a circular economy.

The first part of this Ph.D. thesis deals with the phenotypic characterisation of the SMC. A good level of diversity within the SMC, guaranteed by the selection criteria, suggests the possibility of generalizing the findings of this study to the whole UPCC that the SMC represents. The genotypic clustering generally confirmed the belonging to the species clustering. However, metabolic profiling data indicated that the variability of a microbial community is better represented by phenotyping than genotyping, revealing a degree of metabolic biodiversity even between isolates of the same species. The evaluation through functional diversity indices objectifies the biodiversity description and helps in the strain selection processes. Results remarked the aptitude of LAB strains to metabolise several compounds, thus proving the ability to colonize and ferment several matrices of both animal and plant origin, even different from those of isolation. *Enterococcus* spp., *Lactocaseibacillus* spp. and *Streptococcus* spp. strains were found to use the greatest number of substrates, which are present in the composition of many agro-industrial wastes. This flexibility allows for a broader range of applications, while metabolically restricted microorganisms can be involved in adaptive evolutionary studies to strengthen their edges and direct them towards targeted selection processes for specific uses. Metabolic profiles revealed significant phenotypic diversity, even among closely related

strains, which means that unexpected correlations or metabolic functions may likely occur due to the evolution of biotypes driven by a multifaceted ecosystem. Therefore, the wideness of biodiversity in a species could be a measure of the species' ability to adapt and evolve (Chapter 1).

Physiological data confirmed the phenotypic diversity observed within the SMC, highlighting a remarkable tolerance to conditions far from the optimal ones, a higher adaptability of some species rather than others to environmental variations, and a multi-resistance to several chemicals. The ability to grow in hostile environments represents an opportunity to facilitate the adaptation process and avoid many pre-treatments on recovered biomass often required to make the substrates suitable for microorganisms. Ubiquity represents a form of adaptability: a broad range of lifestyles helps to trigger mechanisms to cope more effectively with physiological challenges. In addition, the diversity of metabolic profiles and resilience to adverse environmental conditions provides a better understanding of niche adaptation. The outcomes can help to identify microorganisms that are particularly capable of carrying out specific fermentation processes, thereby facilitating the targeted design and optimization of such processes. *Lactocaseibacillus paracasei*, *L. rhamnosus*, *L. zeae*, *Enterococcus faecalis* and *E. faecium* appeared to be the most adaptable species of the SMC, tolerating the applied stress better than others. The chemical sensitivity assays showed that LAB strains easily resist low concentrations of different salts used in the food industry to control microbial growth and several antibiotics. The observed multi-resistance to different antimicrobials pointed out the importance of recognizing antibiotic-resistant microorganisms within a microbial collection and of the responsible use of antibiotics to limit the spread of antibiotic-resistance genes, which is already rather uncontrolled (Chapter 2).

In summary, these studies have emphasized the importance of phenotyping as a powerful tool to characterise microbial communities, leading to a deeper understanding of microbial potentialities and allowing the prediction of their behaviour. The knowledge gained can help to identify promising strains and guide both selection and optimization processes towards specific features.

The second part of this Ph.D. project focuses on the development of a predictive computational model based on the information collected in the first part. A predictive approach gives us a preliminary idea of microbial behaviour and suitability for application to a specific experiment thus avoiding waste of material, labour, and time in the laboratory. The use of Artificial Intelligence to explore microorganism diversity represents a powerful approach to advancing

our current understanding of bacterial physiology, potentially paving the way for the development of new industrial applications for LAB. The integration of genotypic and phenotypic data provides a broader view of microbial capabilities. Results obtained in the present work support the adequacy of the performance of our model and underscore how the dataset size impacts predictive accuracy. Random Forest has been chosen as the ideal model to manage the complexity of the considered datasets. The use of a 5-fold cross-validation strategy increased the robustness of the findings and strengthened the validity of the outcomes gained from the analysis (Chapter 3).

The third part of this Ph.D. research investigates the feasibility of using LAB in industrial processes.

Screening strains for D-lactic acid production and their cultivation in bioreactors revealed several key findings. *Lactobacillus delbrueckii* and *Leuconostoc* spp. strains were confirmed to be some of the few D-lactic acid-producing wild-type microorganisms. Process scale-up and environmental conditions control have been essential to maximise yield and productivity, offering reliable and reproducible results for potential industrial applications. *L. delbrueckii* UPCC 2214 was selected as the most promising strain of the SMC, as it showed performance in line with the literature with minimal optimization. Every progress in D-lactic acid production represents an advance in PLA production, thus promoting the growth and improvement of environmentally friendly substitutes of conventional plastics with benefits for both industry and the environment (Chapter 4).

The last screening of this project identified glycerol utilisers within the UPCC, and the outcomes contribute to expanding the scientific literature on the consumption of glycerol by *Enterococcus faecalis* wild-type strains which is currently quite limited. Data obtained by Biolog GENIII Microplates (BIOLOG Inc.©) underlined the qualitative effectiveness of this high-throughput approach. However, a quantitative evaluation is necessary to validate the results. Although the co-utilisation of glycerol and acetate did not appear to occur, the presence of oxygen and acetate metabolism influenced the redox balance during glycerol catabolism, affecting lactate production and limiting carbon losses. This research serves as a first exploration of the UPCC strains' potential in the field of biodiesel. The successful recovery of glycerol, as a biodiesel by-product, provides an opportunity to address waste management issues and produce valuable chemicals. Consequently, it enhances the secondary market of biodiesel and contributes to closing its life cycle (Chapter 5).

Conclusions

The results achieved in this Ph.D. project underline the great versatility of lactic acid bacteria and how they can be harnessed for many different purposes. The surrounding environment, in terms of substrate availability and conditions under which microorganisms are exposed, affects the expression of specific genetic traits resulting in specific metabolic and physiological characteristics that could lead to new processes. This research focused on biodiversity, with the aim of achieving sustainability in the industry, and set up an approach to build not only a phenotypic database for the University of Parma's Culture, but also a microbial platform to predict targeted applications of bacteria.

ABOUT THE AUTHOR

Laura Troiani was born on 9th February 1994 in Norcia (PG), Italy. She got her bachelor's degree in Food Economy and Culture at the University of Perugia, where she pursued her studies with the master's degree programme in Food Technology and Biotechnology. During her last year, she started falling in love with microbiology and, under the supervision of Prof. Pietro Buzzini, in 2019 she discussed her master's thesis entitled "Production of non-conventional yeasts lipids from lignocellulosic biomass".

After graduation, she moved to the National Collection of Yeast Culture in Norwich, England, for an Erasmus Internship programme where she became passionate about research and decided to apply for a Ph.D. position. In November 2020, Laura passed the selection for the Doctoral School in Food Science at the University of Parma under the supervision of Prof. Erasmo Neviani, Prof. Camilla Lazzi and Prof. Alessia Levante.

During the three years of her Ph.D., she focused on the phenotyping of lactic acid bacteria to investigate how to drive targeted fermentation with particular concern for sustainability. During her last year, she spent six months at the University of Natural Resources and Life Sciences of Vienna, Austria, under the supervision of Prof. Michael Sauer to learn how to set up fermentations in bioreactors.

The main results achieved in the three years of her Ph.D. are described in this thesis.



SCIENTIFIC ACTIVITIES

❖ Oral presentation

Troiani L. Leveraging lactic acid bacteria for new sustainable processes. 27th Workshop on the Developments in the Italian Ph.D. Research on Food Science Technology and Biotechnology. University of Naples Federico II, 13-15 September 2023.

❖ Poster presentation

Troiani, L., Levante, A., Bonnici, V., Pierdomenico, V., Dal Palù, A., Bernini, V., Lazzi, C., Neviani, E. Metabolic profiling microarrays for targeted sustainable processes. 7th International Conference on Microbial Diversity “Agrifood microbiota as a tool for a sustainable future”. Parma, 26-29 September 2023.

Levante, A., Troiani, L., Pierdomenico, V., Dal Palù, A., Neviani, E., Bernini, V., Bonnici, V., Lazzi, C. Exploring the biodiversity of a selected microbial core of LABs: bridging experimental datasets and predictive models. Gordon Research Conference “Lactic Acid Bacteria Biology, Symbioses and Applications”. Ventura (CA, USA), 23-28 July 2023.

Troiani, L., Levante, A., Bernini, V., Neviani, E., Lazzi, C. Phenotypic studies for targeted processes in a circular economy model. 10th Congress of European Microbiologists. Hamburg, 9-13 July 2023.

Bonnici, V., Levante, A., Troiani, L., Pierdomenico, V., Puttini, D., Bernini, V., Lazzi, C., Neviani, E., Dal Palù, A. A machine learning approach to predict bacterial phenotypic patterns by means of fragment length distribution fingerprints. 19th Annual Meeting of the Bioinformatics Italian Society. Bari, 21-23 June 2023

Troiani, L. Leveraging lactic acid bacteria collection for new sustainable processes. First Virtual Workshop on the Developments in the Italian Ph.D. Research on Food Science Technology and Biotechnology. University of Palermo, 14-15 September 2021.