BABEŞ-BOLYAI UNIVERSITY
FACULTY OF BIOLOGY AND GEOLOGY
DOCTORAL SCHOOL OF INTEGRATIVE BIOLOGY

DOCTORAL THESIS

Summary

Scientific supervisor:
Prof. Dr. Horia-Leonard BANCIU

PhD student:
Dumitrana-Daniela IORDACHE

Cluj-Napoca,
2024

BABEŞ-BOLYAI UNIVERSITY
FACULTY OF BIOLOGY AND GEOLOGY
DOCTORAL SCHOOL OF INTEGRATIVE BIOLOGY

Characterisation and evolutionary impact of CRISPR-Cas systems in clinically relevant pathogenic bacteria

Summary

Scientific supervisor:
Prof. Dr. Horia-Leonard BANCIU

PhD student:
Dumitrana-Daniela IORDACHE

Cluj-Napoca,
2024

TABLE OF CONTENTS

OUTLINE OF THE THESIS

The CRISPR-Cas system is an adaptive defence mechanism found in bacteria and archaea. It comprises the CRISPR (clustered regularly interspaced short palindromic repeats) array and the CRISPR-associated (*cas*) genes and acts as a sophisticated machinery that incorporates foreign DNA fragments from viruses and mobile genetic elements (MGEs) into CRISPR arrays, and later uses them to recognise and destroy matching invaders.

The **main aim** of the present thesis was to characterise and explore the evolutionary implications of the CRISPR-Cas system in clinically relevant prokaryotic species. To achieve this aim, **three specific objectives** were pursued as follows:

- Characterise CRISPR-Cas systems in several pathogenic bacterial taxa and design PCR primers for the CRISPR loci, facilitating detection in future studies aimed at understanding the evolutionary pathways of CRISPR-Cas systems, monitoring the dissemination of bacterial pathogens, and potentially enabling the development of rapid diagnostic methods.
- Investigate correlations between CRISPR array characteristics (e.g., repeat sequences, spacer number) and geographical origin or collection source of the bacterial isolates.
- Examine the influence of different CRISPR-Cas system types on the acquisition of antimicrobial resistance genes (ARGs) in clinically relevant bacteria.

Chapter I reviews recent findings on the structure, ecology, and evolution of CRISPR-Cas systems, emphasising their role in modulating accessory genomes – the variable portion of a species' genome that often includes genes acquired through horizontal gene transfer – and impacting prokaryotic evolution and physiology.

Chapter II aimed to investigate the relationship between the characteristics of enterobacterial CRISPR loci, their geographical origin, and collection source, potentially providing insights into the evolution and spread of the CRISPR-Cas systems. Highly efficient PCR primers were developed for each locus enabling targeted amplification and analysis of specific CRISPR loci in diverse bacterial populations to facilitate future research.

Chapter III explored the impact of type I-E and I-F CRISPR-Cas systems on ARGs acquisition in two clinically relevant pathogenic bacteria, *Escherichia coli* and *Pseudomonas aeruginosa*. Both taxa possess genomes harbouring either type I-E or I-F systems, with varying proportions and rarely co-existing within the same genome. This characteristic presents a valuable

opportunity to dissect the independent functionalities of each system and how their bacterial regulatory mechanisms influence them in the context of ARG acquisition.

Chapter IV summarises the key findings and conclusions of this thesis, providing a comprehensive overview of the research contributions and outlining potential avenues for future investigation.

Overall, this doctoral thesis contributes to a deeper understanding of the CRISPR-Cas system's role in prokaryotic adaptation and evolution, with a focus on its impact on genomic diversity, geographical distribution, and antimicrobial resistance. This research facilitates the detection of these systems in future studies by characterising CRISPR-Cas systems in pathogenic bacteria and designing specific PCR primers. Furthermore, it explores correlations between CRISPR array characteristics and geographical origin, providing insights into the evolution and spread of this adaptive immune system. Novel findings highlight the complex relationship between different CRISPR-Cas system types and the acquisition of antimicrobial resistance genes in clinically relevant bacteria. This research contributes to the foundation for future investigations into developing innovative CRISPR-based strategies to combat antimicrobial resistance.

Chapter I: CRISPR-Cas system: the powerful modulator of accessory genomes in prokaryotes

Being frequently exposed to foreign nucleic acids, bacteria and archaea have developed an ingenious adaptive defence system called CRISPR-Cas. The system is composed of the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) array, together with CRISPR (*cas*)-associated genes. This system consists of a complex machinery that integrates fragments of foreign nucleic acids from viruses and mobile genetic elements (MGEs), into CRISPR arrays. The inserted segments (spacers) are transcribed and then used by Cas proteins as guide RNAs for recognition and inactivation of the targets. Different types and families of CRISPR-Cas systems consist of distinct adaptation and effector modules with evolutionary trajectories, partially independent. The origin of the effector modules and the mechanism of spacer integration/deletion is far less clear. A review of the most recent data regarding the structure, ecology, and evolution of CRISPR-Cas systems and their role in the modulation of accessory genomes in prokaryotes is proposed in this article. The CRISPR-Cas system's impact on the physiology and ecology of prokaryotes, modulation of horizontal gene transfer events, is also discussed here. This system gained popularity after it was proposed as a tool for plant and animal embryo editing, in cancer therapy, as antimicrobial against pathogenic bacteria and even in the fight against viruses.

Chapter II: Diversity and distribution of CRISPR-Cas systems in *Enterobacteriaceae*: implications for evolution and epidemiology

The successive addition of spacer sequences in the CRISPR array has made the system a valuable molecular marker, with multiple applications. Due to the high degree of polymorphism of the CRISPR loci, their comparison in bacteria from various sources may provide insights into the evolution and spread of the CRISPR-Cas systems. The aim of this study was to establish a correlation between the enterobacterial CRISPR loci, the sequence of direct repeats (DR), and the number of spacer units, along with the geographical origin and collection source. For this purpose, a total of 4401 genomes assigned to *Escherichia coli* (n=2133), *Salmonella enterica* (n=1297), and

Klebsiella pneumoniae (n=971) (*Enterobacteriaceae* family) were retrieved from CRISPRCas++ database (https://crisprcas.i2bc.paris-saclay.fr/MainDb/StrainList). The most prevalent was the I-E CRISPR-Cas system in all three studied taxa. *E. coli* also presents the I-F type, but in a much lesser percentage. The systems found in *K. pneumoniae* can be classified into I-E and I-E*. The I-E and I-F systems have two CRISPR loci, while I-E* has only one locus upstream of the Cas cluster.

Statistical analysis of the characterised CRISPR loci yielded significant findings. While distinct clustering was not evident, statistically significant relationships occurred between the different CRISPR loci and the number of spacer units: for each of the queried taxa, the number of spacers was significantly different ($p < 0.01$) by origin (Africa, Asia, Australia and Oceania, Europe, North America, and South America) but was not linked to the isolation source type (human, animal, plant, food, or laboratory strains). No association was found with the DR sequence, showcasing that the CRISPR-Cas variants are ancient and widespread.

Highly efficient PCR primers were developed for each locus, leveraging the conserved flanking sequences identified during CRISPR characterisation. These primers facilitate future research, enabling targeted amplification and analysis of specific CRISPR loci in diverse bacterial populations. A subset of these primers is already being utilised in our lab to investigate CRISPR-Cas systems in bacteria from various environments, showcasing their immediate practical value.

Chapter III: CRISPR-Cas systems and antimicrobial resistance: a paradoxical relationship in *E. coli* and *P. aeruginosa*

This study aimed to investigate the impact of type I-E and I-F CRISPR-Cas systems on the acquisition of antibiotic resistance genes (ARGs) in Escherichia coli and Pseudomonas aeruginosa, two clinically relevant bacterial pathogens. Genomic data from publicly available databases were analysed for the presence and type of CRISPR-Cas systems and ARGs. Spacer sequences were extracted, and their targets were identified using CRISPRTarget. Statistical tests were employed to compare ARG prevalence between genomes with and without CRISPR-Cas systems.

The prevalent CRISPR-Cas system types (I-E in *E. coli* and I-F in *P. aeruginosa*) did not impede ARG acquisition. Surprisingly, *E. coli* genomes with the I-E system exhibited a higher

prevalence of ARGs compared to those without CRISPR-Cas ($p<00001$). Many ARG variants were exclusively found in genomes possessing type I-E CRISPR-Cas. Conversely, the less common system types (I-F in *E. coli* and I-E in *P. aeruginosa*) were associated with reduced ARG prevalence ($p<00001$ in *E. coli* and $p=0.0122$ in *P. aeruginosa*). No spacer sequences were found to target ARGs or integron components directly.

Under antibiotic selective pressure, highly active CRISPR-Cas systems would likely impose a significant fitness cost on bacteria by preventing the acquisition of ARGs by targeting the mobile genetic elements (MGEs) that carry them. These findings challenge the assumption that CRISPR-Cas systems universally impede ARG acquisition, highlighting a complex interplay that may be influenced by fitness costs and alternative selective pressures. This research contributes to a deeper understanding of the multifaceted role of CRISPR-Cas systems in bacterial evolution and AMR.

Chapter IV: General conclusions and novelty of the results

IV.1. General conclusions

This PhD thesis has explored the multifaceted role of the CRISPR-Cas system and its evolutionary implications in clinically relevant prokaryotic species. Through a combination of literature review, bioinformatic analysis, and hypothesis-driven research, we have revealed several key findings:

- CRISPR-Cas systems shape prokaryotic evolution through adaptive immunity and genome modulation: Our comprehensive review underscores the pivotal role of CRISPR-Cas as an adaptive immune system in prokaryotes, constantly evolving in response to the dynamic landscape of MGEs. Beyond its role in adaptive defence, CRISPR-Cas systems also influence horizontal gene transfer, shaping the acquisition of both beneficial and detrimental genetic material. This interplay highlights the complexity of bacterial evolution and adaptation, with the CRISPR-Cas system serving as a key modulator of prokaryotic genomes.
- CRISPR-Cas systems are ancient, widespread, and diverse: Our work confirms the deep evolutionary roots of CRISPR-Cas immunity. The widespread distribution of these systems

and the diversity of types and subtypes highlight their importance in shaping prokaryotic genomes and adaptive responses.

- Core machinery is stable, spacer repertoire is dynamic: While the core CRISPR-Cas machinery (the *cas* genes and the DR sequence) exhibits remarkable conservation across taxa and time, the spacer repertoire undergoes rapid evolution in response to local environmental challenges. This balance of stability and adaptability allows prokaryotes to maintain an effective defence against diverse and ever-changing threats.
- CRISPR-Cas systems influence HGT through a complex interplay of factors: While CRISPR-Cas systems do not directly target ARGs or integron components, they can indirectly affect ARG acquisition by targeting plasmids and phages that often carry them. These plasmids may also harbour integrated prophages, which can be targeted by the CRISPR-Cas system, leading to the elimination of both the prophage and the plasmid carrying the ARGs. However, our work suggests that the relationship between CRISPR-Cas and HGT is complex and context-dependent, with some systems potentially facilitating the acquisition of beneficial genes under certain conditions.
- CRISPR-Cas and antimicrobial resistance (AMR): CRISPR-Cas systems do not necessarily prevent the acquisition of ARGs. For example, the prevalent CRISPR-Cas system types in both *E. coli* (I-E) and *P. aeruginosa* (I-F) did not impede ARG acquisition and, in the case of *E. coli*, were even associated with increased ARG prevalence. Conversely, less prevalent CRISPR-Cas types in both species were linked to reduced ARG presence. This paradoxical relationship underscores the urgent need for further research to understand the complex interplay between CRISPR-Cas, AMR, and other selective pressures.

Overall, this thesis demonstrates the significant impact of CRISPR-Cas systems on prokaryotic evolution, ecology, and physiology. By understanding the mechanisms and dynamics of these systems, we can gain valuable insights into the ongoing battle between bacteria and mobile genetic elements and develop innovative tools and strategies to address global health challenges such as AMR.

IV.2. Novelty of the results

This thesis provides several novel contributions to the field of CRISPR-Cas and bacterial research:

Analysis of CRISPR-Cas diversity in *Enterobacteriaceae*: Our bioinformatic analysis offers a comprehensive characterisation of CRISPR-Cas systems within three clinically important *Enterobacteriaceae* genera (*Salmonella*, *Escherichia*, and *Klebsiella*). Notably, we identified significant geographic variation in the number of spacers within CRISPR arrays. This suggests that local environmental pressures, such as exposure to diverse bacteriophages, may drive adaptation of the spacer repertoire in these pathogens. This finding provides valuable insights into the ecological and evolutionary forces shaping CRISPR-Cas diversity in these bacteria and highlights the importance of considering geographic context when studying CRISPR-Cas-mediated immunity.

Investigation of the relationship between CRISPR-Cas and ARGs: Our study addresses the complex interplay between CRISPR-Cas systems and the acquisition of ARGs, a topic of significant interest in the field. The novelty of our approach lies in the systematic comparison of two prevalent CRISPR-Cas systems (I-E and I-F) in both *E. coli* and *P. aeruginosa*. This comparative analysis allows us to dissect the unique impact each system has within different bacterial species, where they are subject to distinct regulatory mechanisms. Our findings challenge the conventional view that CRISPR-Cas universally acts as a barrier to ARG acquisition, highlighting the importance of considering the specific system type and ecological context when evaluating its impact on antimicrobial resistance.

Observations consistent with an ancient and globally distributed CRISPR-Cas backbone: Our findings, in conjunction with previous research, support the hypothesis of an ancient origin and global dissemination of repeat sequences within CRISPR-Cas systems. This reinforces the notion that the core CRISPR-Cas machinery is highly conserved and likely predates the divergence of many bacterial lineages. Our observations add to the growing body of evidence suggesting that while the spacer repertoire evolves rapidly in response to local pressures, the fundamental framework of the system remains remarkably stable over evolutionary time.

Development of CRISPR-specific PCR primers: We have designed and validated PCR primers for each CRISPR locus identified in our study. These primers can be used in future

research to investigate the prevalence and diversity of CRISPR-Cas systems in diverse environments and clinical settings.

Collectively, these contributions advance our understanding of the CRISPR-Cas system and its implications for prokaryotic evolution, ecology, and the ongoing battle against antimicrobial resistance.

LIST OF PUBLICATIONS INCLUDED IN THE THESIS AS CHAPTER

CHAPTER I

Butiuc-Keul, A., Farkas, A., Carpa, R. and **Iordache, D**. (2022). CRISPR-Cas system: the powerful modulator of accessory genomes in prokaryotes. *Microb Physiol* **32***:* 2–17. https://doi.org/10.1159/000516643

CHAPTER II

Iordache, D., Baci, G.M., Căpriță, O., Farkas, A., Lup, A., and Butiuc-Keul, A. (2022). Correlation between CRISPR loci diversity in three enterobacterial taxa. *Int J Mol Sci* **23:** 12766. https://doi.org/10.3390/IJMS232112766

LIST OF PUBLICATIONS NOT INCLUDED IN THE THESIS

[1] Costache, C., Colosi, I. A., Toc, D. A., Daian, K., Damacus, D., Botan, A., Toc, A. G., Pana, A., Panaitescu, P., Neculicioiu, V., Șchiopu, P., **Iordache, D**. and Butiuc-Keul, A. (2024) CRISPR-Cas system, antimicrobial resistance and *Enterococcus* genus – a complicated relationship. *Biomed* **12**: 1625. https://doi.org/10.3390/biomedicines12071625

[2] Butiuc-Keul, A., Farkas, A., Carpa, R., Dobrota, C. and **Iordache, D.** (2022) Development of smart fruit crops by genome editing. *Turkish J Agric For* **46**: 129–140. https://doi.org/10.3906/TAR-2012-104

[3] Toc, D. A., Butiuc-Keul, A. L., **Iordache, D.**, Botan, A., Mihaila, R. M., Costache, C. A., Colosi, I. A., Chiorean, C., Neagoe, D. S., Gheorghiu, L., and Junie, L. M. (2022) Descriptive analysis of circulating antimicrobial resistance genes in vancomycin-resistant

Enterococcus (VRE) during the COVID-19 pandemic. *Biomed* **10**: 1122. https://doi.org/10.3390/BIOMEDICINES10051122

[4] Totolici, A. I., Mitrea, S., Cioloca, A. T., Lupu, A., Móricz, P. M., Muntean, D., Negre, R., Topîrceanu, A., Țoc, M. **Iordache, D.** (2022) Breathing chemicals: a review of air pollution over the years. *Stud Univ Babeș-Bolyai Biol* **67**: 177–198. https://doi.org/10.24193/SUBBBIOL.2022.1.10

[5] Dobrescu, M. Ștefan, **Iordache, D.** and Butiuc-Keul, A. (2022) Revealing the CRISPR array in bacteria living in our organism. *Stud Univ Babeș-Bolyai Biol* **67**: 131–142. https://doi.org/10.24193/SUBBBIOL.2022.1.07

[6] Butiuc-Keul, A., Carpa, R., Podar, D., Szekeres, E., Muntean, V., **Iordache, D**. and Farkas, A. (2021) Antibiotic resistance in *Pseudomonas* spp. through the urban water cycle. *Curr Microbiol* **78**: 1227–1237. https://doi.org/10.1007/S00284-021-02389-W

[7] Toth, P., **Fit, D.**, Culda, C. A., Carpa, R., Butiuc-Keul, A., Roba, C. A. and Rosu, C. (2020) Quality assessment of drinking water from several private wells from Lazuri Village (Satu Mare County, Romania). *J Environ Prot Ecol* **21**: 106–115. https://hero.epa.gov/hero/index.cfm/reference/details/reference_id/7302539

[8] Butiuc-Keul, A., **Fiț, D**. and Farkas, A. (2019) Trends in molecular biology of several fruit trees. *JOJ Hort Arboricult* **2**: 63–65. https://doi.org/10.19080/JOJHA.2018.02.555589

[9] Butiuc-Keul, A., Goia, I., Cristea, V., **Fiț, D**., Șuteu, A. and Farkas, A. (2019) RAPD markers associated with linolenic acid synthesis in several *Boraginaceae* plant species. *Ann Oradea Univ Biol Fascicle* **26**: 62–66. https://bioresearch.ro/2019-1/062-066-AUOFB.26.1.2019-BUTIUC-KEUL.A.-RAPD.markers.associated.pdf

ATTENDANCES AT INTERNATIONAL CONFERENCES

[1] **Iordache, D**., Butiuc-Keul, A. (2024) The puzzling paradox of CRISPR-Cas: why common systems don't block antibiotic resistance genes.

- Oral presentation and abstract at the 2nd Conference of the Romanian Society of Bioinformatics - RoBioinfo 2024, 11-13 April 2024, Cluj-Napoca, Romania. *Accelerating human and environmental knowledge discovery through bioinformatics*

advances, Abstract book, ISBN: 978-606-28-1795-4, pp. 57. https://www.editurauniversitara.ro/medicina-33/accelerating-human-and-environmental-knowledge-discovery-through-bioinformatics-advances-2nd-robioinfo-conference-2024.html#resp-tab1

[2] **Iordache, D**., Dobrescu, M., Fita, A., Butiuc-Keul, A. (2023) The intricate relationship between the CRISPR-Cas system and antibiotic-resistance genes in *Klebsiella pneumoniae*.

- Abstract at the Annual International Conference of the *Romanian Society of Biochemistry and Molecular Biology*, September 13-15, 2023, Cluj-Napoca, Romania. *Studia Universitatis Babeş-Bolyai, Biology, 68(2), RSBMB Book of abstracts pp. 81,* https://biogeo.ubbcluj.ro/RSBMB/wp-content/uploads/2023/09/Abstract-book-RSBMB-2023_varianta-finala.pdf.

[3] **Fiț, D**., Farkas, A., Butiuc-Keul, A. (2019) CRISPR-CAS system in *Pseudomonas sp*. – structure and role in horizontal gene transfer of antimicrobial resistance genes.

- Abstract at the Annual International Conference of the *Romanian Society of Biochemistry and Molecular Biology*, September 26-27, Iași, Romania. *Journal of Experimental and Molecular Biology, 20(3), RSBMB Book of abstracts, pp. 25.* http://www.jemb.bio.uaic.ro/index.php/jemb/article/view/36/23.