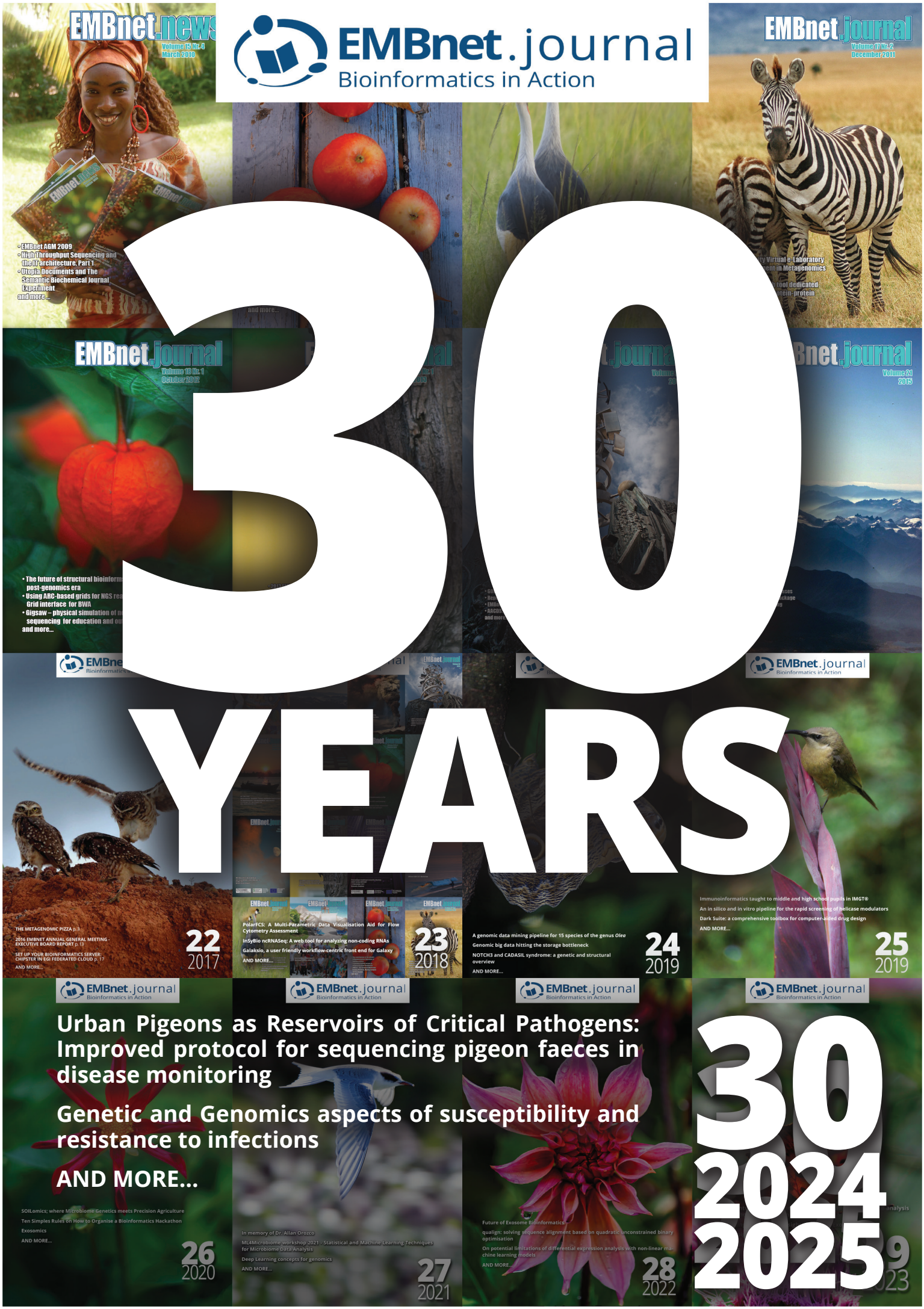

EMBnet.news
Volume 15 Nr. 4
March 2010
• EMBnet AGM 2009
• High Throughput Sequencing and the IT architecture, Part 1
• Utopia Documents and The Semantic Biochemical Journal Experiment
and more...
EMBnet.journal
Volume 17 Nr. 2
December 2011
EMBnet.journal
Volume 18 Nr. 1
October 2012
• The future of structural bioinformatics in the post-genomics era
• Using ARC-based grids for NGS read mapping: Grid interface for BWA
• Gigsaw – physical simulation of DNA sequencing for education and outreach
and more...
EMBnet.journal
Volume 21
2015
EMBnet.journal
Bioinformatics in Action
30
YEARS
THE METAGENOMIC PIZZA p. 3
2016 EMBNET ANNUAL GENERAL MEETING - EXECUTIVE BOARD REPORT p. 13
SET UP YOUR BIOINFORMATICS SERVER: CHIPSTER IN EGI FEDERATED CLOUD p. 17
AND MORE...
22
2017
PolarFCS: A Multi-Parametric Data Visualisation Aid for Flow Cytometry Assessment
InSyBio ncRNASeq: A web tool for analyzing non-coding RNAs
Galaksio, a user friendly workflow-centric front end for Galaxy
AND MORE...
23
2018
A genomic data mining pipeline for 15 species of the genus Olea
Genomic big data hitting the storage bottleneck
NOTCH3 and CADASIL syndrome: a genetic and structural overview
AND MORE...
24
2019
Immunoinformatics taught to middle and high school pupils in IMGT®
An in silico and in vitro pipeline for the rapid screening of helicase modulators
Dark Suite: a comprehensive toolbox for computer-aided drug design
AND MORE...
25
2019
EMBnet.journal
Bioinformatics in Action
Urban Pigeons as Reservoirs of Critical Pathogens: Improved protocol for sequencing pigeon faeces in disease monitoring
Genetic and Genomics aspects of susceptibility and resistance to infections
AND MORE...
SOILomics; where Microbiome Genetics meets Precision Agriculture
Ten Simples Rules on How to Organise a Bioinformatics Hackathon
Exosomics
AND MORE...
26
2020
In memory of Dr. Allan Orozco
ML4Microbiome workshop 2021 - Statistical and Machine Learning Techniques for Microbiome Data Analysis
Deep Learning concepts for genomics
AND MORE...
27
2021
Future of Exosome Bioinformatics
qualign: solving sequence alignment based on quadratic unconstrained binary optimisation
On potential limitations of differential expression analysis with non-linear machine learning models
AND MORE...
28
2022
30
2024
2025

EMBnet.journal
Bioinformatics in Action

Contents

Editorial 2

News

In Memoriam – Prof. Amos Bairoch - A visionary pioneer of bioinformatics and a giant of the Life Sciences
Laurent Falquet 3

Reports

Report on the Swiss-Colombian workshop 2024: "Metagenomics data analysis of mixed microbial communities."
Laurent Falquet, Jeferyd Yepes-Garcia, Iván Mateus Gonzalez, Emiliano Barreto Hernandez 6

Reviews

The Intersection of Artificial Intelligence and Precision Endocrinology
Dimitrios Vlachakis, Konstantina Dragoumani, Eleni Papakonstantinou, George Chrousos 9

Research Papers

Urban Pigeons as Reservoirs of Critical Pathogens: Improved protocol for sequencing pigeon faeces in disease monitoring
Elin Hermann, Renaud Van Damme, Erik Bongcam-Rudloff, Leila Nasirzadeh 12

Genetic and Genomics aspects of susceptibility and resistance to infections
Cesar Paz-y-Miño, Javier Gualdron, Ariane Paz-y-Miño .. 20

Protein Spotlight

Protein Spotlight 265 33

Protein Spotlight 267 35

Protein Spotlight 280 37

Protein Spotlight 282 39

Editorial

The latest issue of EMBnet.journal is released at a time of profound turmoil for global society. Across continents, scientific communities are navigating unprecedented challenges shaped by geopolitical instability, economic uncertainty, rapid technological transformation, and growing societal tensions. These political and societal shifts increasingly affect science itself, as a rising scepticism towards scientific knowledge can be observed in several countries, accompanied by significant reductions in public investment in research and education. In some cases, such as in Argentina, funding has been cut almost entirely, while many other nations are following a similar trajectory to varying degrees.

This trend stands in stark contradiction to well-established evidence. Countries that perform best economically today are precisely those that have invested heavily and consistently in education, research, and innovation over the past two to three decades. At the same time, we are witnessing the emergence of a profoundly new technological era driven by artificial intelligence. AI holds enormous promise but also raises legitimate concerns, and its impact will reshape societies, economies, and scientific practice in ways we are only beginning to understand. Without sustained investment in research and education, many countries risk entering even deeper and more prolonged economic and social recessions.

As a small but committed international network in bioinformatics, EMBnet can contribute—modestly yet meaningfully—by offering knowledge, collaboration, and hope. By bringing together researchers from diverse countries and backgrounds, EMBnet continues to do what it does best: educating and training students, researchers, and the broader public in bioinformatics and emerging biotechnologies that are transforming life sciences and society. We remain optimistic and firmly committed to publishing and promoting high-quality, accessible scientific work that helps demystify new technologies and highlights their potential to address global challenges.

With this issue, we conclude Volume 30 of EMBnet.journal, and we warmly invite our community to continue contributing. We encourage you not only to submit your own work but also to inspire colleagues and collaborators to publish with us, helping to strengthen an open, inclusive, and resilient scientific community.

Erik Bongcam-Rudloff
Editor-in-Chief
erik.bongcam@slu.se

http://dx.doi.org/10.14806/ej.30.0.1079

Cover picture: Selected cover pages of EMBnet.journal from past years

EMBnet.journal
Bioinformatics in Action

Falquet et al. (2026) EMBnet.journal 30, e1085
http://dx.doi.org/10.14806/ej.30.0.1085

In Memoriam – Prof. Amos Bairoch

A visionary pioneer of bioinformatics and a giant of the Life Sciences

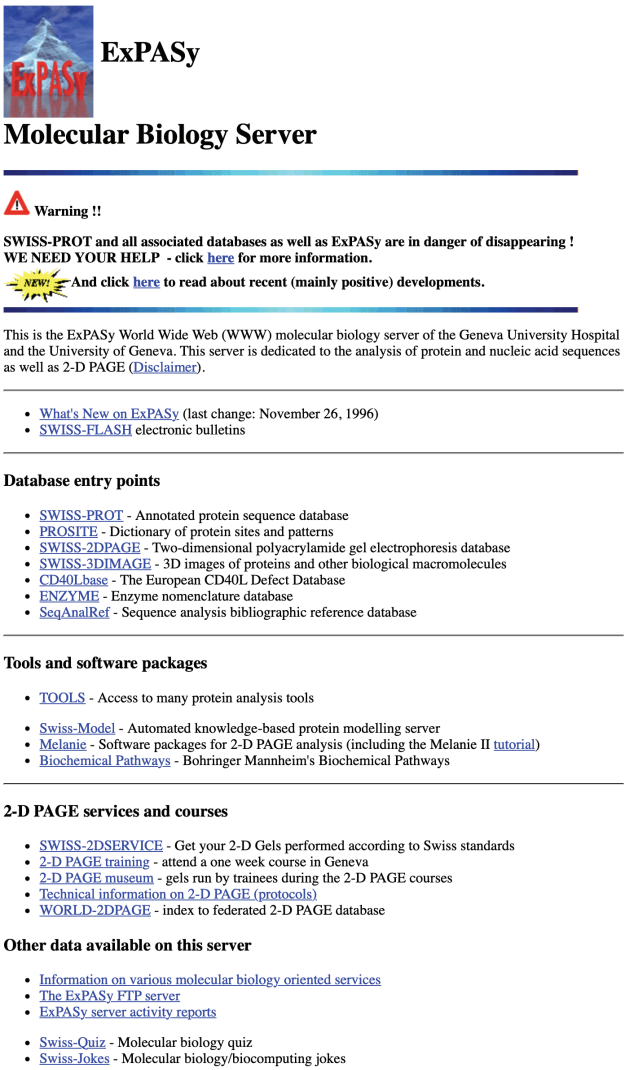
by Laurent Falquet & the EMBnet community

Prof. Amos Bairoch

Prof. Amos Bairoch (1957–2025) was one of the most influential and transformative figures in modern bioinformatics, a scientist whose vision reshaped how biological information is curated, accessed, and understood. As the creator of Swiss-Prot, UniProtKB/Swiss-Prot, PROSITE, ENZYME, and several other foundational databases, he defined the gold standard for accuracy, manual curation, and open accessibility in protein knowledgebases. His work enabled generations of researchers and laid the groundwork for breakthroughs across genomics, proteomics, biotechnology, precision medicine, and systems biology.

A professor at the University of Geneva and a driving force at the Swiss Institute of Bioinformatics (SIB), Prof. Bairoch dedicated his life to building high-quality, sustainable, community-oriented resources that serve millions of users worldwide. He was not only an exceptional innovator but also a generous mentor and a passionate advocate for open science, training and inspiring countless scientists who continue to carry forward his vision.

The EMBnet community mourns his sudden passing with profound sadness. Amos was a colleague, collaborator, and friend whose impact on our field—and on global Life Sciences—cannot be overstated. His



ExPASy

Molecular Biology Server

Warning !!

SWISS-PROT and all associated databases as well as ExPASy are in danger of disappearing !
WE NEED YOUR HELP - click here for more information.

NEW! And click here to read about recent (mainly positive) developments.

This is the ExPASy World Wide Web (WWW) molecular biology server of the Geneva University Hospital and the University of Geneva. This server is dedicated to the analysis of protein and nucleic acid sequences as well as 2-D PAGE (Disclaimer).

- What's New on ExPASy (last change: November 26, 1996)
- SWISS-FLASH electronic bulletins

Database entry points

- SWISS-PROT - Annotated protein sequence database
- PROSITE - Dictionary of protein sites and patterns
- SWISS-2DPAGE - Two-dimensional polyacrylamide gel electrophoresis database
- SWISS-3DIMAGE - 3D images of proteins and other biological macromolecules
- CD40Lbase - The European CD40L Defect Database
- ENZYME - Enzyme nomenclature database
- SeqAnalRef - Sequence analysis bibliographic reference database

Tools and software packages

- TOOLS - Access to many protein analysis tools
- Swiss-Model - Automated knowledge-based protein modelling server
- Melanie - Software packages for 2-D PAGE analysis (including the Melanie II tutorial)
- Biochemical Pathways - Bohringer Mannheim's Biochemical Pathways

2-D PAGE services and courses

- SWISS-2DSERVICE - Get your 2-D Gels performed according to Swiss standards
- 2-D PAGE training - attend a one week course in Geneva
- 2-D PAGE museum - gels run by trainees during the 2-D PAGE courses
- Technical information on 2-D PAGE (protocols)
- WORLD-2DPAGE - index to federated 2-D PAGE database

Other data available on this server

- Information on various molecular biology oriented services
- The ExPASy FTP server
- ExPASy server activity reports
- Swiss-Quiz - Molecular biology quiz
- Swiss-Jokes - Molecular biology/biocomputing jokes

Figure 1. Oldest screenshot of Expasy server on Wayback Machine (Nov 29, 1996), when Amos called for help to save SwissProt (web.archive.org), Wayback Machine started in 1996…

Article history
Received: 19 December 2025
Published: 09 January 2026

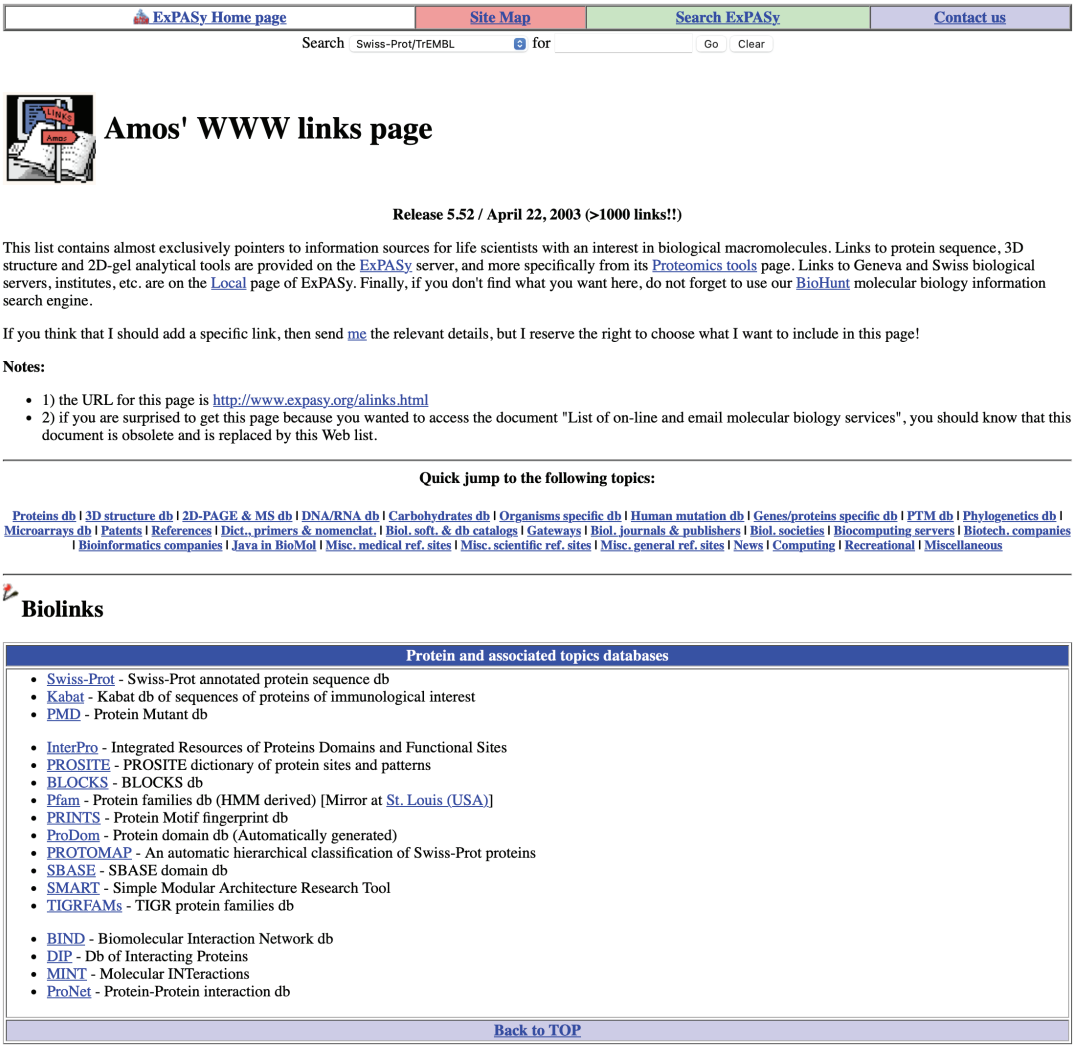


Figure 2. The famous Amos' WWW links page (ultimate version)

legacy endures in the resources he created, the standards he established, and the many people whose careers he touched. We honour his memory with deep respect, gratitude, and heartfelt sorrow.

I would like to pay tribute to him with a few additional personal memories.

1993 was an exciting year. In January, I started my thesis in biochemistry and Amos Bairoch was working on the same floor of the Department of Medical Biochemistry at the University of Geneva. In September 1993, he suddenly walked into my lab and frantically asked me, «Do you have Mosaic on your Mac? » (He was referring to NCSA Mosaic, one of the very first web browsers). I replied that I did, so he dragged me over to it and told me to type in «http://expasy.hcuge.ch», and suddenly I saw the very first page of the ExPASy server (Expert Protein Analysis System) (Figure 1) before my astonished eyes...

From there, I could click on SwissProt, ENZYME, or PROSITE to access the databases he was developing and lots of other links. It was revolutionary! That first page became later the famous «Amos' WWW links page» (Figure 2) a world reference containing more than 1000 links to bioinformatics tools and servers until 2005. ExPASy[1] is still now the entry page for the tools developed at the Swiss Institute of Bioinformatics.

In 1994, we were both at the first proteomics conference in Siena, Italy, entitled «From Genome to Proteome». He was sitting right in front of me with his laptop (the only one in the room) and throughout the conference he was constantly typing away on his keyboard, creating new SwissProt annotation records for the proteins that the speakers were presenting at that very moment. At the same time, he was able to ask the speaker relevant questions. I was amazed, especially when I learned later that Amos' laptop contained the only master copy of the SwissProt database until 2003 and the fusion into UniProtKB!

Amos was a strong supporter of the EMBnet organisation, he participated in many courses, workshops and conferences that EMBnet organised, one of them being the Bioinformatics conference in Torremolinos

[1]https://www.expasy.org

Falquet et al. (2026) EMBnet.journal 30, e1085
http://dx.doi.org/10.14806/ej.30.0.1085

Figure 3. Torremolinos EMBnet AGM2007: from left to right - Amos Bairoch, Francisco Melo, Laurent Falquet (© Erik Bongcam-Rudloff).

(Spain) in 2007 (Figure 3). He also managed, through the creation of the Swiss Institute of Bioinformatics in 1998, to raise the governmental funding to support SwissProt and the Swiss EMBnet node!

He received many awards during his career, the most recent one at the ISMB, and would certainly have deserved the Nobel Prize, but unfortunately his untimely death prevented him from receiving it. Despite all this, Amos kept his modesty. This summer, when he received the ISCB award, he simply introduced himself as a «biocurator», proud to explain the qualities of this profession. And of course, he kept his style of dress in all circumstances: Amos wore a polo shirt.

Thank you, Amos, for your exceptional scientific work and your friendship, we all miss you.

Dr Laurent Falquet
former Swiss EMBnet node manager 2002-2009.

Learn more about Amos personality and work, and deposit your remembrances on his SIB memorial website: https://www.sib.swiss/amos/

EMBnet.journal
Bioinformatics in Action

Falquet et al. (2024) EMBnet.journal 30, e1063
http://dx.doi.org/10.14806/ej.30.0.1063

Reports

Report on the Swiss-Colombian workshop 2024: "Metagenomics data analysis of mixed microbial communities."

Laurent Falquet[1,2] ✉, Jeferyd Yepes-Garcia[1,2], Iván Mateus Gonzalez[2], Emiliano Barreto Hernandez[3] ✉

[1] University of Fribourg, Fribourg, Switzerland

[2] Swiss Institute of Bioinformatics, Lausanne, Switzerland

[3] Instituto de Biotecnología, Universidad Nacional de Colombia, Edificio Manuel Ancizar, Bogotá, Colombia

Competing interests: LF is a member of EMBnet.journal Editorial Board; JYG none; IMG none, EBH is a member of EMBnet.journal Editorial Board

Introduction

This workshop was organised as part of the Swiss-Colombian project *"Improvement of rice straw management in Latin America using metagenomics and synthetic communities (SynComs) to reduce CO_2 impact"*. The project was financed by the Leading House for Latin America and managed by the University of St-Gallen, Switzerland. It aims at expanding preliminary work done to sample metagenomics sequences from various treatments applied to rice straw to speed up composting. To broaden the range of participants, the workshop topic was chosen to cover the world of mixed microbial communities.

The workshop announcement was carried out via several mailing lists in Colombia, via a web page[1], and on social networks. We gathered over 35 registrations and selected 27 participants, balancing their interest in the topic, academic level, and home university. They were distributed among students, researchers, and group leaders from various universities in Colombia: Ibagué, Bogotá, Cali, and Medellín. The gender balance was 30% male and 70% female.

Programme

The workshop was held from June 17th to June 21st, originally planned at the Biotechnology Institute of Universidad Nacional de Colombia (IBUN-UNAL). However, we had to move to a nearby hotel conference room due to a student's strike blocking the accesses to the University.

Participants were asked to have some prerequisite skills like UNIX and R basic knowledge. To ensure that all participants had a similar level from the beginning, the first day of the workshop was used to refresh the participants' programming-scripting competencies. The slides of the courses and the practical exercises were accessible through a Moodle page hosted by the University of Fribourg. Each half-day was split into theoretical lectures (60-90 minutes each), followed by hands-on practical sessions (150 minutes each). The detailed program is shown in Figure 1. For the practical sessions, we either used RStudio on the participants' laptops or a remote HPC cluster located at the University of Bern, Switzerland. Participants were allowed to connect remotely using dedicated access, and for the whole genome shotgun analysis, the participants worked in groups of four. To ensure the participants' motivation, we asked them to answer exercises at the end of each practical session. In addition, we asked them to prepare a presentation in groups for the final day.

On the final day, a mini-symposium was organised. We invited **Prof. Pilar Junier** as keynote speaker from the University of Neuchâtel, Switzerland. She gave a brilliant lecture on *"Fungal highway columns and interactions between soil fungi and bacteria"*, demonstrating the use of various 3D printed materials to study how the bacteria can travel using fungal hyphae as a road. Remotely via Zoom, we had the opportunity to listen to **Dr Vanessa Otero's** excellent lecture from the University of Idaho, USA. During her talk, Vanessa explained how *"Rice Straw recycling increases soil microbial functional diversity during rice straw decomposition"*. **Dr Ivan Mateus** from the University of Fribourg, Switzerland, gave a fascinating lecture on *"Non-canonical analyses of controlled small communities"*, where he demonstrated how one could study mycorrhizal and nodular symbiosis using bioinformatics analysis of Next Generation Sequence (NGS) data. We ended the mini-symposium with two

[1] http://bioinf.ibun.unal.edu.co/cursos/Metagenomics2024/

Article history
Received: 15 July 2024
Published: 11 October 2024

Falquet et al. (2024) EMBnet.journal 30, e1063
http://dx.doi.org/10.14806/ej.30.0.1063

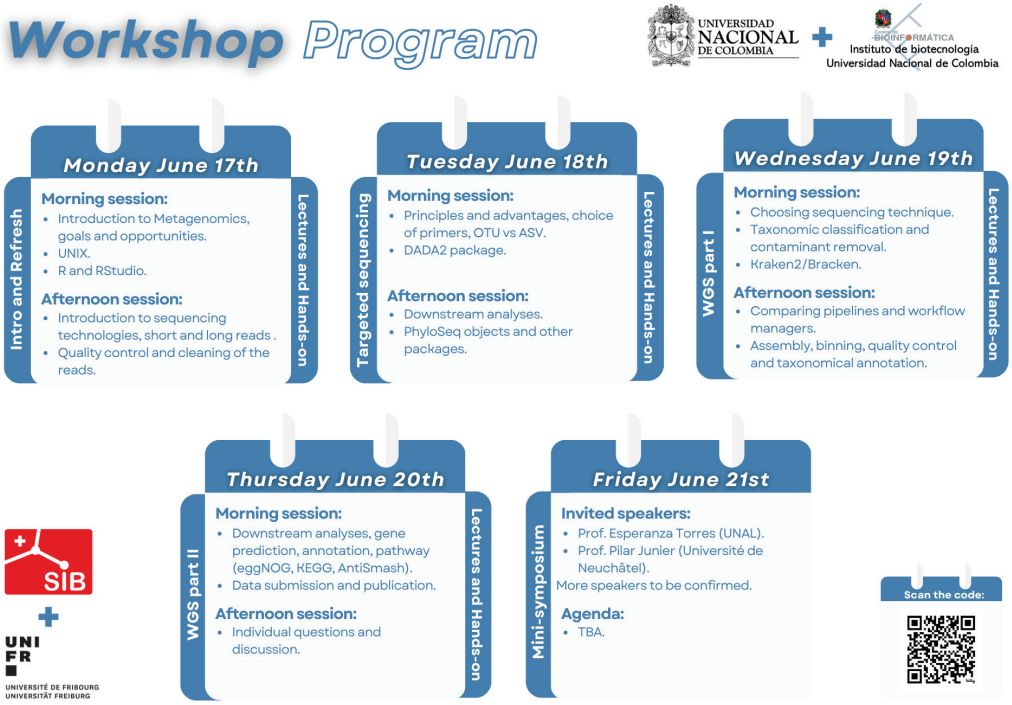


Figure 1. Detailed programme of the Swiss-Colombian workshop

lectures from MSc candidates, both from Universidad Nacional, **Nicolas Rodriguez-Romero**, *"Microbiome-based targeted bioprospecting to produce a nitrogen-fixing biofertilizer using waste-derived carbon sources"*, who reported on the use of pig manure to produce fixed nitrogen and short-chain fatty acids (SCFA) with bacterial communities, and **Nicolás Novoa**, *"Selection and perspectives on microbial consortium development: a*

Figure 2. Classroom of the workshop

Falquet et al. (2024) EMBnet.journal 30, e1063
http://dx.doi.org/10.14806/ej.30.0.1063

Figure 3. Group picture in front of the hotel

case of study on rice straw degradation" concluding with a demonstration of pairwise compatibility assessments of *Trichoderma* and *Bacillus* strains.

At the end of the workshop, the seven groups of participants presented the results obtained during the practical sessions.

Evaluation of the course

The participants were given an online evaluation form to evaluate the workshop. The form included an overall evaluation of content, objectives, methodology, and logistics. The participants from the different research institutions expressed their satisfaction with the high academic level of the course in general. They highly valued the knowledge shown by the trainers and the materials used in the lectures and practical exercises.

Some participants regretted the lack of time to finish the practical exercises and sometimes too few monitors to help them. They recommended improving this part in the future. The Wi-Fi connection issues on the first day were also mentioned as preventing a good understanding of the topic. Fortunately, the hotel staff solved these issues at the end of the day, and the rest of the week was perfect.

Other aspects, like motorbike noise and the "mooing cow" food truck in front of the hotel, partially disturbed the class. Luckily, these were counterbalanced by the delicious hotel coffee breaks and the social dinner organised by the teachers!

Before concluding, a little anecdote: the use of online presentations seems normal nowadays in 2024, but it wasn't in 2005 when we first experimented with it between Colombia and Switzerland for an EMBnet course with the now defunct Marratech system (see course report page 5[2]). It's this pioneering role that we like to play at EMBnet (Prof. Barreto is the EMBnet Node Manager for Colombia and Dr. Laurent Falquet is a former EMBnet Node Manager for Switzerland).

Conclusions

According to the comments of the attendees and the reported good feeling of the organisers, this workshop was very successful and led to potential collaborations. We will continue working on mixed microbial communities, looking to improve the knowledge on the interactions between fungi and bacteria. In particular, the management of rice straw using synthetic communities will remain our main interest in the future.

Acknowledgements

The organisers especially thank Prof. Pilar Junier (*University of Neuchâtel*), Prof. Daniel Uribe-Vélez (*Universidad Nacional de Colombia*), and MSc candidate Nicolás Novoa for their commitment to the workshop, contributing with all their expertise and endeavour to make it such a great successful event. This workshop was supported by the grant RPG2340 of the "Leading House for Latin America" and by the Universidad Nacional de Colombia. We thank Pierre Berthier for helping us with the HPC access and Claudia Parra, Yuly Escobar, and Carolina López (Instituto de Biotecnología, Universidad Nacional de Colombia) for supporting the administrative and logistics of the workshop.

[2]https://journal.embnet.org/index.php/embnetnews/article/download/94/110

EMBnet.journal
Bioinformatics in Action

Vlachakis et al. (2024) EMBnet.journal 30, e1052
http://dx.doi.org/10.14806/ej.30.0.1052

Reviews

The Intersection of Artificial Intelligence and Precision Endocrinology

Dimitrios Vlachakis[1,2,3]✉, **Konstantina Dragoumani**[1,2], **Eleni Papakonstantinou**[1,2], **George P Chrousos**[2]✉

[1]Laboratory of Genetics, Department of Biotechnology, School of Applied Biology and Biotechnology, Agricultural University of Athens, Athens, Greece

[2]University Research Institute of Maternal and Child Health & Precision Medicine, and UNESCO Chair on Adolescent Health Care, National and Kapodistrian University of Athens, "Aghia Sophia" Children's Hospital, Athens, Greece

[3]School of Informatics, Faculty of Natural & Mathematical Sciences, King's College London, London, U.K.

Competing interests: DV is a member of EMBnet.journal Editorial Board; KD none; EP none; GPC none

Abstract

Bioinformatics and artificial intelligence (AI) have emerged as transformative tools in modern medicine, revolutionising the landscape of medical diagnosis and treatment. Herein, we provide an overview of the synergistic relationship between bioinformatics and AI, elucidating their pivotal roles in deciphering complex biological data and advancing precision medicine and, in particular, endocrinology. We explore various applications of bioinformatics and AI in medical research, including genomic analysis, drug discovery, disease diagnosis, and personalised treatment strategies. Additionally, we discuss challenges and future directions in leveraging these technologies to enhance healthcare outcomes.

Introduction

The convergence of bioinformatics and artificial intelligence (AI) has catalysed a paradigm shift in medical science, offering unprecedented opportunities for understanding biological mechanisms, predicting disease outcomes, and tailoring therapies to individual patients. Bioinformatics, the interdisciplinary field that combines biology, computer science, and information technology, plays a crucial role in processing, analysing, and interpreting vast biological datasets, such as genomic sequences, proteomic profiles, and clinical records . Meanwhile, AI encompasses a spectrum of computational techniques that enable machines to learn from data, recognise patterns, and make predictions without explicit programming.

Artificial Intelligence (AI) has become an indispensable tool across various fields, revolutionizing processes and decision-making paradigms. In the realm of healthcare, AI's integration has significantly impacted diagnosis, treatment, and patient care. Endocrinology, a branch of medicine focusing on hormone-related disorders, stands to benefit greatly from AI's capabilities. Herein, we explore how AI is utilised in endocrinology, highlighting advancements, applications, and their implications for patient outcomes and healthcare systems.

In recent years, the integration of bioinformatics and AI has led to remarkable breakthroughs in biomedical research and clinical practice (Obermeyer and Emanuel, 2016). This article explores the multifaceted applications of these synergistic technologies in medicine, highlighting their contributions to genomic analysis, drug discovery, disease diagnosis, and personalised treatment strategies.

Genomic Analysis, Drug Design and Personalised Treatment and Diagnosis

One of the primary areas where bioinformatics and AI converge is in genomic analysis. The human genome comprises of billions of nucleotides, and deciphering its intricacies is essential for understanding genetic predispositions to diseases, identifying potential therapeutic targets, and developing personalised treatments. Bioinformatics tools facilitate the storage, retrieval, and analysis of genomic data, enabling researchers to explore the genetic basis of diseases and uncover novel insights into their molecular mechanisms.

AI algorithms, such as machine learning and deep learning, enhance the interpretation of genomic data by identifying patterns, predicting gene functions, and correlating genetic variations with clinical outcomes.

Article history
Received: 15 March 2024
Accepted: 07 April 2024
Published: 09 September 2024

For instance, predictive models trained on genomic data can stratify patients based on their risk of developing certain diseases, guiding preventive interventions and personalised healthcare plans. Moreover, AI-driven approaches accelerate the identification of disease-associated genetic variants, facilitating the discovery of biomarkers for early diagnosis and prognosis assessment.

Traditional drug discovery processes are costly, time-consuming, and often plagued by high failure rates. By leveraging bioinformatics and AI, researchers can streamline drug discovery pipelines, expedite target identification, and repurpose existing drugs for new indications. Bioinformatics tools enable the analysis of biological networks, protein structures, and drug interactions, facilitating the identification of druggable targets and the design of novel therapeutics.

AI algorithms enhance drug discovery efforts by predicting the binding affinity between drug molecules and target proteins, simulating molecular dynamics, and optimizing compound libraries for lead identification. Virtual screening methods, powered by machine learning models, prioritise potential drug candidates based on their structural properties, bioactivity profiles, and safety profiles, thereby accelerating the drug development process.

Furthermore, AI-driven approaches enable the exploration of vast chemical space, guiding the synthesis of small molecules with optimised pharmacological properties. By integrating genomic and clinical data with drug response profiles, researchers can develop predictive models for patient stratification and personalised treatment selection, fostering the era of precision pharmacotherapy (Ching *et al.*, 2018).

Early and accurate diagnosis is critical for improving patient outcomes and mitigating disease burden. Bioinformatics and AI play pivotal roles in disease diagnosis by integrating multi-omics data, clinical parameters, and imaging features to enhance diagnostic accuracy and prognostic prediction. Machine learning algorithms analyse complex datasets to identify disease-specific biomarkers, distinguish between benign and malignant conditions, and stratify patients into distinct subgroups based on their molecular profiles.

For example, in oncology, bioinformatics tools enable the characterization of tumor heterogeneity, the identification of driver mutations, and the prediction of treatment response. AI-based image analysis techniques enhance medical imaging modalities, such as magnetic resonance imaging (MRI) and computed tomography (CT), by automating lesion detection, segmentation, and classification. Integrating genomic data with imaging phenotypes enables radiogenomic analysis, which correlates imaging features with underlying molecular alterations, facilitating non-invasive diagnosis and treatment monitoring.

Personalised medicine aims to tailor medical interventions to individual patient characteristics, thereby optimizing therapeutic efficacy and minimizing adverse effects. Bioinformatics and AI empower personalised medicine by integrating genomic data, clinical information, and environmental factors to develop patient-specific treatment strategies. Pharmacogenomic analysis identifies genetic variants that influence drug metabolism, response, and toxicity, enabling clinicians to optimise drug selection, dosage, and administration regimens.

Moreover, AI-driven clinical decision support systems assist healthcare providers in interpreting complex data, generating treatment recommendations, and predicting patient outcomes. By leveraging electronic health records (EHRs) and real-world evidence, these systems facilitate evidence-based medicine and enable continuous learning from clinical practice. Patient stratification algorithms identify subpopulations with distinct disease phenotypes or treatment responses, enabling targeted interventions and precision healthcare delivery (Hripcsak and Albers, 2013).

AI in endocrinology

One of the primary applications of AI in endocrinology lies in diagnosis and risk prediction. AI algorithms, particularly machine learning models, can analyse vast amounts of patient's clinical data, including medical records, imaging studies, and genetic information, to identify patterns indicative of endocrine disorders. For instance, AI-powered systems can detect subtle anomalies in hormone levels, aiding in the early diagnosis of conditions such as diabetes, thyroid disorders, and adrenal diseases. Moreover, these algorithms can assess an individual's risk of developing endocrine disorders based on their genetic predispositions and lifestyle factors, enabling proactive interventions and personalised preventive care strategies.

AI plays a crucial role in advancing precision medicine approaches in endocrinology. By integrating patient-specific data with genomic information and biomarkers, AI algorithms can tailor treatment regimens to individuals' unique biological profiles. For instance, in the management of diabetes, AI-driven platforms can analyse glucose monitoring data to optimise insulin dosing schedules, minimizing the risk of hypoglycemia and hyperglycemia. Similarly, AI-based decision support systems assist healthcare providers in selecting the most effective medications and dosages for thyroid disorders, ensuring optimal therapeutic outcomes while minimizing adverse effects.

In endocrinology, imaging modalities such as ultrasound, MRI, and CT scans are essential for evaluating glandular structures and diagnosing tumors or abnormalities. AI-powered image analysis techniques have significantly enhanced the accuracy and efficiency of endocrine imaging interpretation. Deep learning algorithms can segment and classify anatomical structures, detect subtle lesions, and differentiate between benign and malignant nodules with high precision. This not only expedites diagnosis but also

Vlachakis et al. (2024) EMBnet.journal 30, e1052
http://dx.doi.org/10.14806/ej.30.0.1052

reduces the need for invasive procedures, leading to better patient experiences and outcomes.

The integration of AI with remote monitoring technologies has transformed the delivery of endocrine care, particularly in the context of telemedicine. Wearable devices equipped with AI algorithms enable continuous monitoring of glucose levels, hormone fluctuations, and medication adherence in patients with diabetes and other endocrine disorders. Real-time data analysis and predictive modeling empower healthcare providers to remotely assess patients' conditions, adjust treatment plans as needed, and intervene promptly in case of emergencies. This remote monitoring capability not only enhances patient convenience but also facilitates proactive disease management, ultimately improving clinical outcomes and reducing healthcare costs.

While the utilization of AI in endocrinology holds immense promise, it also presents various challenges and ethical considerations. Privacy and data security concerns arise due to the vast amounts of sensitive health information involved in AI-driven healthcare applications. Ensuring the confidentiality and integrity of patient data is paramount to maintain trust in AI systems. Additionally, there is a need for robust regulatory frameworks to govern the development, deployment, and validation of AI algorithms in clinical practice, safeguarding against biases, errors, and unintended consequences. Moreover, the integration of AI should be accompanied by comprehensive training programs for healthcare professionals to ensure proficiency in utilizing AI tools effectively and ethically.

Challenges and Future Directions

Despite the transformative potential of bioinformatics and AI in medicine, several challenges remain to be addressed to realise their full impact. Data integration and interoperability issues hinder the seamless exchange of information across disparate systems and platforms. Standardization of data formats, ontologies, and metadata is essential for facilitating data sharing and collaboration in biomedical research.

Moreover, ethical, legal, and regulatory considerations must be carefully navigated to ensure patient privacy, data security, and responsible use of AI-driven technologies. Transparent and interpretable AI models are needed to facilitate regulatory approval, clinical adoption, and physician trust. Furthermore, addressing biases in training data and algorithmic decision-making is crucial for mitigating disparities in healthcare delivery and ensuring equitable access to personalised medicine (Collins and Varmus, 2015).

Future research directions include the development of multimodal AI frameworks that integrate diverse data types, such as genomics, imaging, and clinical records, to provide comprehensive insights into disease mechanisms and treatment responses. Advances in learning and decentralised AI enable collaborative analysis of distributed datasets while preserving data privacy and security. Additionally, interdisciplinary collaborations between bioinformaticians, computer scientists, clinicians, and policymakers are essential for driving innovation and translating research findings into clinical practice.

All in all, the integration of bioinformatics and AI represents a transformative force in modern medicine, enabling personalised approaches to disease diagnosis, treatment, and prevention. From genomic analysis to drug discovery and clinical decision support, these synergistic technologies hold promise for revolutionizing healthcare delivery and improving patient outcomes. By addressing challenges related to data integration, ethics, and bias, we can harness the full potential of bioinformatics and AI to usher in a new era of precision medicine (Mamoshina *et al.*, 2016).

Key Points

- The convergence of bioinformatics and AI revolutionises medical science by facilitating the understanding of biological mechanisms, predicting disease outcomes, and personalising therapies.
- Synergistic applications in medicine include enhancing genomic analysis, drug discovery, disease diagnosis, and personalised treatment strategies through bioinformatics and AI integration.
- In endocrinology, AI aids in early diagnosis, risk prediction, treatment optimization, and accurate imaging interpretation, improving patient care and outcomes.
- Challenges include data integration, interoperability, ethical considerations, and addressing biases in algorithmic decision-making, while future directions involve developing multimodal AI frameworks and fostering interdisciplinary collaborations to advance precision medicine.

References

Ching T, Himmelstein DS, Beaulieu-Jones BK, Kalinin AA, Do BT et al. (2018) Opportunities and obstacles for deep learning in biology and medicine. J R Soc Interface **15**(141). http://dx.doi.org/10.1098/rsif.2017.0387

Collins FS and Varmus H (2015) A new initiative on precision medicine. N Engl J Med **372**(9), 793-795. http://dx.doi.org/10.1056/NEJMp1500523

Hripcsak G and Albers DJ (2013) Next-generation phenotyping of electronic health records. J Am Med Inform Assoc **20**(1), 117-121. http://dx.doi.org/10.1136/amiajnl-2012-001145

Mamoshina P, Vieira A, Putin E and Zhavoronkov A (2016) Applications of Deep Learning in Biomedicine. Mol Pharm **13**(5), 1445-1454. http://dx.doi.org/10.1021/acs.molpharmaceut.5b00982

Obermeyer Z and Emanuel EJ (2016) Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. N Engl J Med **375**(13), 1216-1219. http://dx.doi.org/10.1056/NEJMp1606181

EMBnet.journal
Bioinformatics in Action

Hermann et al. (2024) EMBnet.journal 30, e1059
http://dx.doi.org/10.14806/ej.30.0.1059

Research Papers

Urban Pigeons as Reservoirs of Critical Pathogens: Improved protocol for sequencing pigeon faeces in disease monitoring

Elin Hermann, Renaud Van Damme, Erik Bongcam-Rudloff✉, Leila Nasirzadeh

Swedish University of Agricultural Sciences, Uppsala, Sweden

Competing interests: EH none; RVD none; EBR is EMBnet.journal Editor; LN none

Abstract

The spread of pathogens by animals is a serious issue around the world that causes a severe threat to human health. Feral pigeons (*Columba livia forma domestica*) that live in urban areas are zoonotic carriers of various pathogens that can be transmitted to humans by faecal contamination. This study aimed to detect the presence of bacterial, viral, and specifically fungal pathogens in pigeon faeces based on the World Health Organization's priority pathogen list published in 2022. Fresh faecal samples were collected at Uppsala, Central Station, and Svandammen, the pigeons' most relevant gathering spots and feeding sites. Genomic DNA was directly extracted from these samples, and ScilifeLab performed High throughput sequencing through Oxford Nanopore Technologies (PromethION). Metagenomic analysis revealed that most of the critically prioritized viral and bacterial pathogens listed by World Health Organization were present in pigeon faeces. Regarding fungal pathogens, which were the main objective of this study, samples from both studied locations contained all critical, high and medium-important fungal pathogens published in the World Health Organization list, such as *Aspergillus fumigatus*, *Candida albicans*, *Candida auris*, *Cryptococcus neoformans*, *Nakaseomyces glabratus*, *Candida tropicalis*, and *Cryptococcus gattii*. These fungal pathogens pose the risk of invasive fungal diseases and severe infections in low-immunity individuals and vulnerable populations. The findings indicate the importance of conducting further research to comprehensively understand potential exposure to feral pigeons. Furthermore, keeping pigeons away from sensitive areas, such as hospitals, and implementing measures to control pigeon populations can significantly decrease the spread of pathogens.

Introduction

Microorganic spread is a public health issue as some microorganisms are health-hazardous to humans. Microorganisms transmitted by wild animals, *i.e.*, zoonotic spread, have caused pandemics such as the plague, where fleas associated with rodents spread bacteria, and the coronavirus pandemic that is assumed to have been spread by bats (Piret and Boivin, 2021). Zoonotic spread by birds is especially relevant as they can transport microorganisms between continents (Moschetti *et al.*, 2017; Briscoe *et al.*, 2021). Birds can live close to humans in cities, which can risk spreading pathogens through their littering (Glushakova *et al.*, 2021).

The feral pigeon (*Columba livia forma domestica*) is a bird species living close to humans and is globally common in urban areas (Haag-Wackernagel and Moch, 2004; Magnino *et al.*, 2009; de Vasconcellos *et al.*, 2022). In cities, feral pigeons have plenty of food, can build nests, have shelter and natural predators are absent. The pigeon is adapted to indoor nesting and can live in large colonies which can accumulate a lot of faeces (Moschetti *et al.*,2017; de Vasconcellos *et al.*, 2022). In Finland, it is reported that feral pigeon populations increase with human density, which is thought to be due to pigeons being omnivores (Jokimäki and Suhonen, 1998).

Pigeons in cities have a high pathogenic diversity. Transmission of microorganisms to birds as vectors is probable to occur through the environment and the food (Moschetti *et al.*, 2017). Feral pigeons can ingest trash, feed on landfills, and get fed by humans which could result in the spread of fungi, viruses, and bacteria, including antibiotic- and multidrug-resistant bacteria. Thus, human activities can promote the spread (Phan *et al.*, 2013; Moschetti *et al.*, 2017; Freire *et al.*, 2022). Pathogens can also colonize urban pigeons when drinking

Article history
Received: 13 June 2024
Accepted: 21 October 2024
Published: 04 November 2024

Hermann et al. (2024) EMBnet.journal 30, e1059
http://dx.doi.org/10.14806/ej.30.0.1059

contaminated water *e.g.*, with human faeces (Freire *et al.*, 2022). This was reported in a Brazilian study where human multi-drug resistant diarrheagenic *Escherichia coli* strains were identified as present in urban pigeons' cloaca (Borges *et al.*, 2017). Several different bacterial, viral, and pathogenic fungal species have been reported present in pigeon faeces (Catroxo *et al.*, 2011; Phan *et al.*, 2013; Medina *et al.*, 2017; Naz *et al.*, 2017; de Vasconcellos *et al.*, 2022; Kowalczyk and Wójcik-Fatla, 2022; Nualmalang *et al.*, 2023). This indicates that urban pigeons have the potential to transmit pathogens to the environment (Borges *et al.*, 2017).

Pigeons have, for a long time, been considered as pathogenic transmitters to humans. Haag-Wackernagel and Moch (2004) report almost 200 published papers between 1941-2003, which confirm cases of feral pigeon-human transmission of pathogens that resulted in disease. However, all cases are probably not documented. Pathogenic transmission to humans is common through inhalation (Haag-Wackernagel and Moch, 2004). Contaminated dust particles can get airborne and inhaled when pigeons beat their wings (Magnino *et al.*, 2009). Thus, there is a potential health hazard if pigeon nests are close to ventilation systems (Medina *et al.*, 2017). For instance, may the pathogenic fungi *Aspergillus fumigatus*, *Cryptococcus gattii*, *Mucorales*, *Cryptococcus neoformans*, and *Histoplasma spp.* be transmitted through inhalation of spores (World Health Organization, 2022). Airborne fungi spores have, for instance, been detected in urban pigeon-feeding areas in Karachi (Pakistan) which could result in these areas being health hazardous for humans (Naz *et al.*, 2017). In Baltimore (US), potentially viable spores of *Enterocytozoon bieneusi* were detected in both air and water samples. Therefore, it can also be a contamination risk after rainfall in water run-off systems (Graczyk *et al.*, 2007).

The fact that pigeons usually live near humans makes zoonotic disease transmission possible (Magnino *et al.*, 2009). It can result in unintentional and intentional encounters such as feeding pigeons or passing pigeon-dense areas (Haag-Wackernagel and Moch, 2004). Avoiding exposure to feral pigeons can be challenging as they can be close to hospitals and community transport systems (de Vasconcellos *et al.*, 2022). There is a potential health risk when pigeons live near hospitals, while immunocompromised people can be exposed to pigeon-spread pathogens. A study by de Vasconcellos and colleagues (2022) found pathogenic *Staphylococcus aureus* and *Enterococcus faecium* in pigeon faeces close to hospitals in Brazil. Moreover, transmission of pathogens to humans might occur when handling sick or dead pigeons (Magnino *et al.*, 2009).

Fungal pathogens may be an underestimated human health hazard (World Health Organization, 2022) since the risk of fungal infections is generally low. However, the risk of fungal infections is 1000 times higher for immunocompromised persons (Haag-Wackernagel and Moch, 2004). Further, fungi have the potential to cause even greater threats than viruses and bacteria since there are no vaccines and limited availability of antifungal medicine. In addition, there is an increase in invasive fungal diseases and fungal resistance is appearing. Invasive fungal diseases are mainly a threat to immunocompromised populations but can become a global health concern. For instance, comorbid invasive fungal infections with aspergillosis, mucormycosis, and candidaemia increased during the coronavirus pandemic, and infection range and dispersal of fungi could be increased with the climate change (Nnadi and Carter, 2021; World Health Organization, 2022). In 2022, World Health Organization (WHO) presented the first systematic global prioritization of 19 health-hazardous fungi to guide and encourage research and public health action. In that list, the fungi are ranked as critical, high, and medium risk.

To summarise, feral pigeons are a source of pathogenic spread to humans, which can cause disease and possibly even epidemic outbreaks (Haag-Wackernagel and Moch, 2004; Piret and Boivin, 2021). Thus, it is essential to monitor health-hazardous microorganisms to avoid outbreaks. Risk factors for pathogenic spread are increased human-animal interaction, urbanization, and globalization (Piret and Boivin, 2021). More specifically, fungal pathogens need to be more recognized (World Health Organization, 2022). Therefore, this study aimed to explore the prevalence of bacterial, viral, and especially fungal pathogens registered in the WHO's priority pathogen list (2022), in pigeon faeces in Uppsala.

Materials and methods

Sample collection

Four fresh stool samples from free-living feral pigeons were collected from two different urban locations of Uppsala, Central Station, and Svandammen, known to be the most relevant gathering spots and feeding sites for urban pigeons, respectively (Figure 1). Samplings were carried out in March 2023.

DNA extraction

Since DNA extraction from faecal samples were challenging, a combination of kit-based (specifically, the ZYMO RESEARCH Quick-DNA Faecal Microbe MiniPrep Kit) and non-kit-based protocols was employed to remove residual contaminations, proteins, and salts from the DNA samples, enhancing DNA concentration and minimizing fragmentation. DNA was extracted from 150 mg of faecal sample as recommended by the kit manufacturer. The extracted DNAs were eluted in 50 ul of elution buffer and assessed using Nanodrop, Qubit, and Tapestaion before being stored at -20°C. This resulted in four samples that were good enough for sequencing (Table 1).

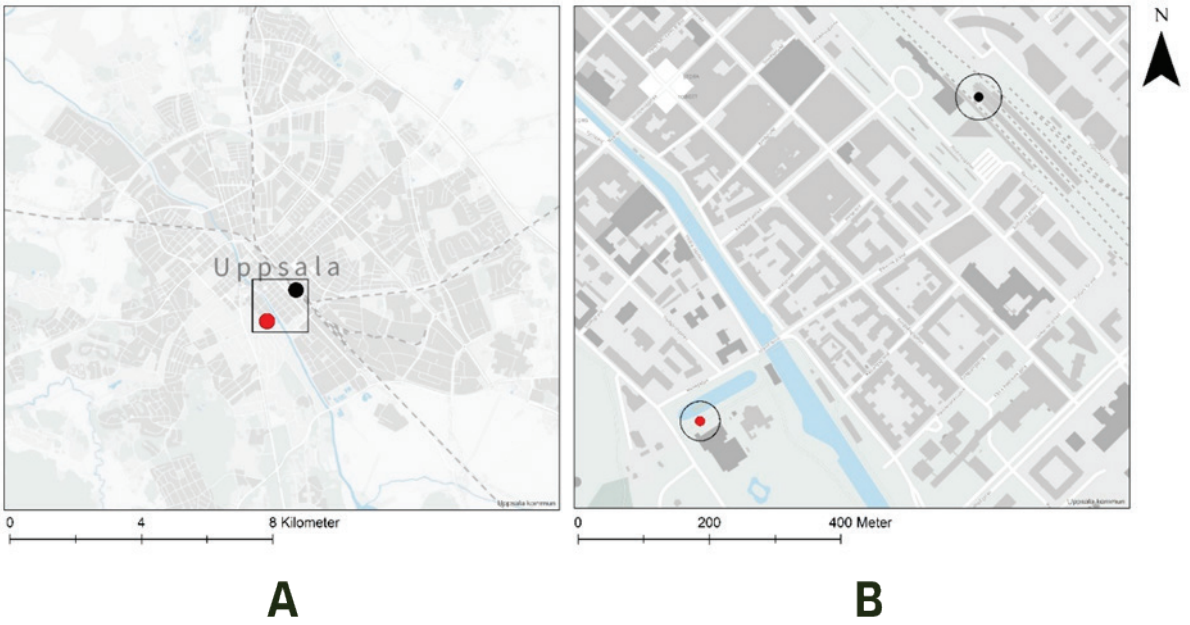


Figure 1. Collection sites of faecal samples, the train station (black bullet), and Svandammen (red bullet). In **A**, sites are presented in Uppsala where the square indicates **B**. In **B**, approximative collection ranges at each site are marked as circles. The figures are composed in ArcMap, the background map is from Uppsala Municipality. Each sample was collected using a wooden stick, placed into a 15ml falcon tube, and directly transported to the laboratory. Great care was taken to avoid the part of the droplet that had touched the ground.

Table 1. Sample ID, their location, the BioSample and Experiment Accession numbers.

Sample ID	Location	BioSample Accession	Experiment Accession
Sample1	Central station	SAMEA115741869	ERX12635354
Sample2	Central station	SAMEA115741870	ERX12635601
Sample3	Svandammen	SAMEA115741871	ERX12635579
Sample4	Svandammen	SAMEA115741872	ERX12635531

Sequencing

High throughput sequencing was performed by ScilifeLab through the Oxford Nanopore Technologies (ONT) method (PromethION) using SQK-NBD114-24 protocol. Reads were Quality controlled with third-party software pycoQC. The trimming of the adapter after sequencing was also done by the ScilifeLab.

High-Throughput Sequencing Analysis

Raw sequenced reads were quality-controlled with fastQC.Taxonomic classification of metagenomic reads at the species level was performed by aligning metagenomic reads against the "nr" database updated on 10-Mar-2022. This database is one of the databases provided by Kaiju[1] and contains all the sequences from the NCBI BLAST nr database belonging to Archaea, bacteria and viruses. For fungi detection, metagenome reads were aligned against the Fungi Database provided by Kaiju and updated on 29-Mar-2022 This database contains all the fungi sequences coming from the NBCI Refseq database.

Krona[2] was used to enhance the interpretability of the complex taxonomic data generated by Kaiju, global results of fungi, bacteria and viruses are presented in (Figure 2) (see all Krona results and the Kaiju read count for the 4 samples using NR and Fungi databases in Supplementary files[3] 1 to 40).

All the tools, their version and the parameters used are available in Supplementary file[3] 41.

Results

1. Detection of Viral Pathogens

Analysis showed that specific viruses that pose significant public health risks were discovered in both Uppsala

[1]https://github.com/bioinformatics-centre/kaiju

[2]https://docs.csc.fi/apps/krona/

[3]http://journal.embnet.org/index.php/embnetjournal/article/downloadSuppFile/1059/1059_supp_1

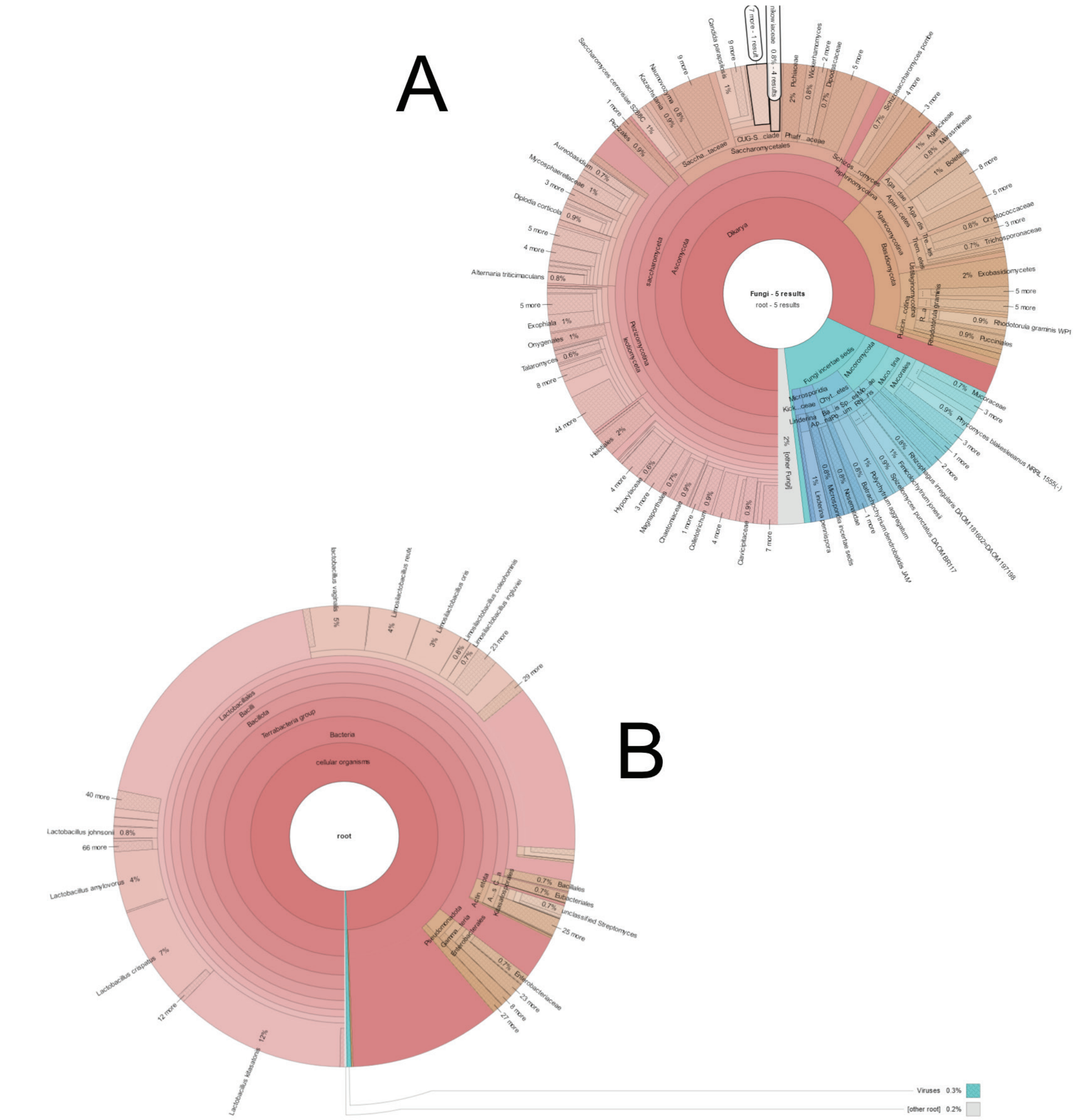


Figure 2. Krona representation of the global results using the Fungi (**A**) and NCBI nr (**B**) databases.

Central Station and Svandammen. These viruses have the potential to cause the diseases listed as priorities by the WHO because of their epidemic potential.

Identified Viral Pathogens

Several viruses were detected in both locations (Figure 3). All the samples from both locations contained the following viruses:

- *Crimean-Congo hemorrhagic fever virus* (CCHFV), responsible for causing Crimean-Congo hemorrhagic fever.
- *Lassa mammarenavirus*, which leads to Lassa fever.
- *SARS coronavirus* (SARS-CoV), causing Severe Acute Respiratory Syndrome (SARS) respiratory illness.
- *Monkeypox virus* (MPX), the causative agent of Monkeypox disease.
- *Alpha coronavirus*, associated with respiratory infections.
- *Beta coronavirus*, linked to respiratory infections.

2. Detection of Bacterial Pathogens

As bacteria develop resistance to antibiotics, they pose challenges in treating infections and increase the potential for diseases to spread. To address this, WHO has published a list of bacteria with critical, high and medium priorities based on the need for developing new antibiotics. The critical priority is linked to specifically threatening diseases within hospitals, while the high and medium priorities are associated with more common and severe illnesses.

Based on the Metagenomics analysis results, most of the bacteria mentioned in the WHO priority list were detected in fecal samples in both locations.

Identified Bacterial Pathogens

In the critical priority category, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and various

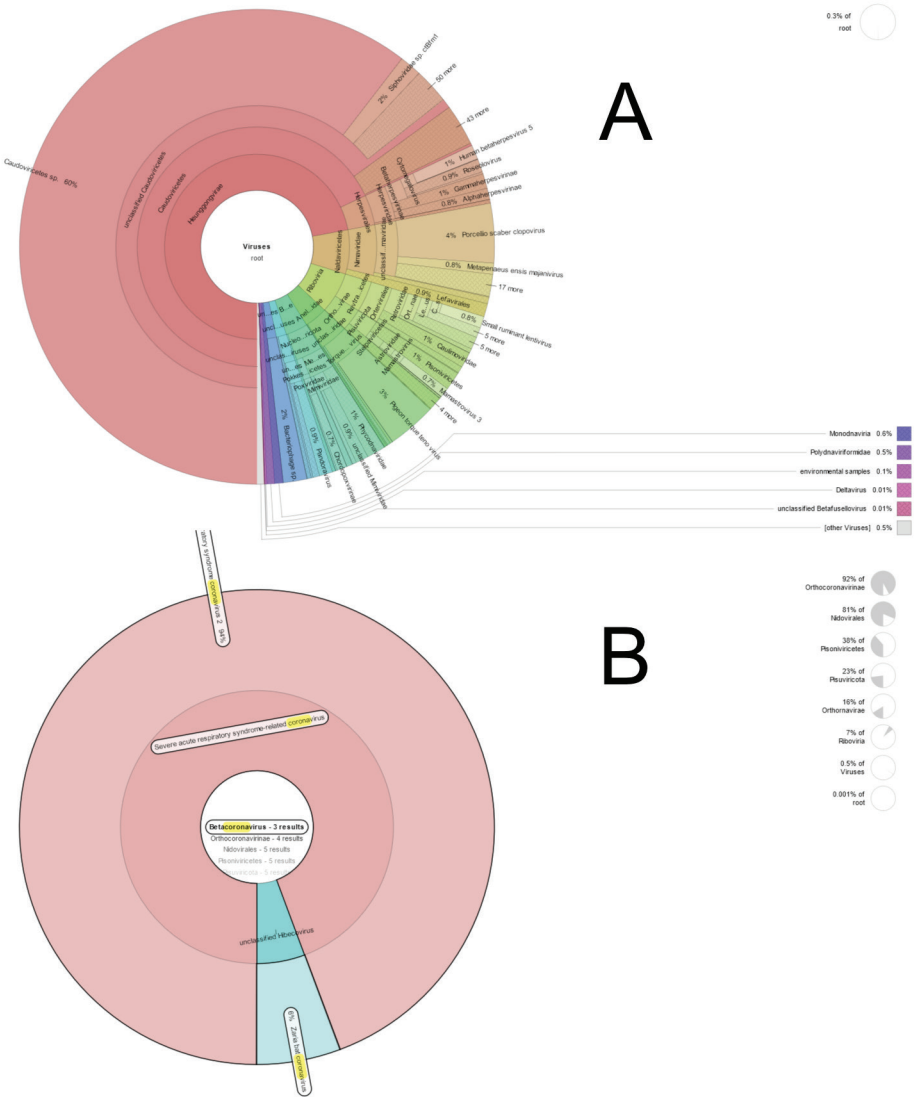


Figure 3. Krona representation of the Viral hit found in Kaiju with nr database. **A** represent all viral hits and **B** a magnification of the krona result to SARS-CoV.

enterobacteriaceae, including *Klebsiella*, *E. coli*, *Serratia*, and *Proteus*, were identified in all samples of both locations. These bacteria are carbapenem-resistant and can lead to severe and potentially fatal infections.

Within the high-priority category, we observed the presence of *Enterococcus faecium* (vancomycin-resistant), *Staphylococcus aureus* (methicillin-resistant, vancomycin-intermediate resistant), and *Campylobacter spp.* (fluoroquinolone-resistant) in all samples.

Based on the third priority category, all samples contained *Streptococcus pneumoniae* and *Shigella spp.* These bacteria are penicillin-non-susceptible and fluoroquinolone-resistant respectively.

In addition to the priority list, our samples also included *Clostridium difficile* and *Mycobacterium tuberculosis*. These bacteria, as indicated by the WHO, have demonstrated a growing resistance to traditional treatments in recent years. *Clostridioides difficile* is acknowledged as an emerging pathogen with the potential to cause zoonotic diseases in humans (Mitchell *et al.*, 2022).

3. Detection of Fungal Pathogens

This study focused on the detection of fungal pathogens in the collected samples because the latest report from the WHO (2022) shows a high increase in cases of invasive fungal diseases (IFDs) and countries are advised to enhance their capacity to manage fungal infections.

WHO has classified fungal pathogens into three priority groups — critical, high, and medium — based on drug resistance and challenges in treatment and management.

All fungi mentioned in the critical priority category, including *Cryptococcus neoformans*, *Candida auris*, *Aspergillus fumigatus*, and *Candida albicans*, were

Hermann et al. (2024) EMBnet.journal 30, e1059
http://dx.doi.org/10.14806/ej.30.0.1059

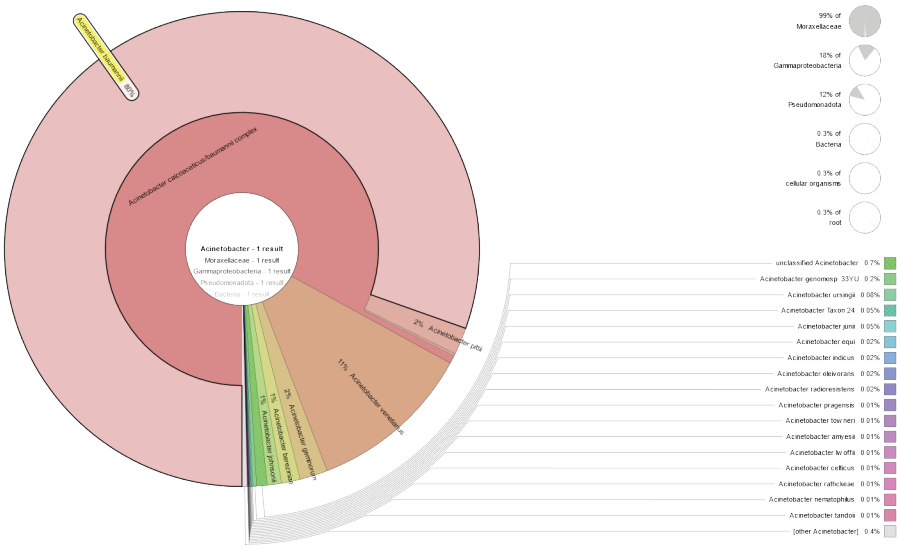


Figure 4. Krona representation of Acinetobacter baumannii. Other results are in Supplementary files[4].

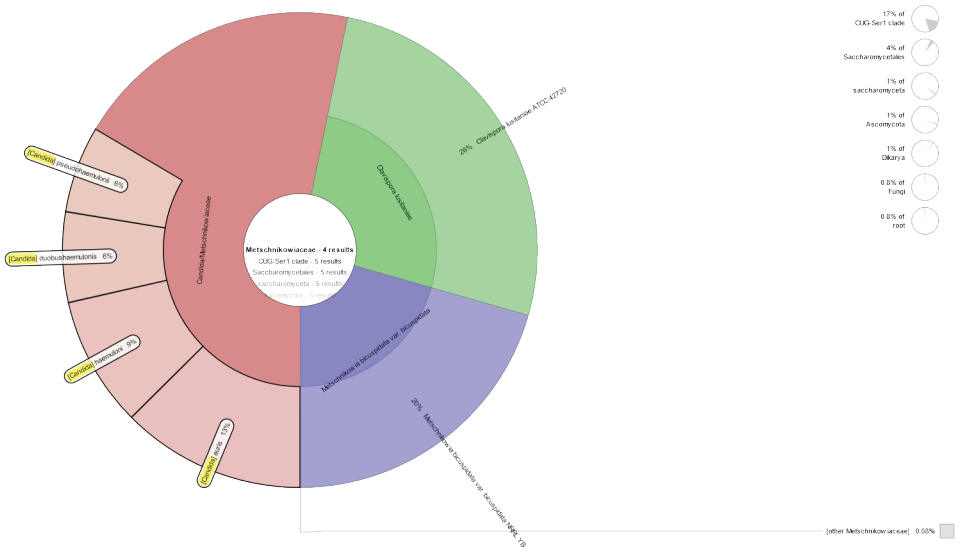


Figure 5. Krona representation of Candida auris.

present in samples from Uppsala Central Station and Svandammen.

Among seven fungal pathogens in the second priority group, six fungi, including *Nakaseomyces glabrata* (*Candida glabrata*), *Histoplasma spp.*, *Mucorales*, *Fusarium spp.*, *Candida tropicalis*, and *Candida parapsilosis*, were found in all samples.

From pathogens in the medium group, our samples contained nearly all fungi, including *Scedosporium spp.*, *Coccidioides spp.*, *Pichia kudriavzeveii* (*Candida krusei*), *Cryptococcus gattii*, *Talaromyces marneffei*, *Pneumocystis jirovecii* and *Paracoccidioides spp.*

Discussion

This study's extensive metagenomic analysis underscores the significant role that urban pigeons play in the epidemiology of critical pathogens. This includes the identification of viral, bacterial, and particularly fungal organisms, which have notable impact on public health, especially for immunocompromised populations. Our findings, which detail the detection of various pathogens in pigeon feces collected from Uppsala Central Station and Svandammen, reveal the diversity of the microbial

[4]http://journal.embnet.org/index.php/embnetjournal/article/downloadSuppFile/1059/1059_supp_1

community that urban pigeons can potentially harbor. This diversity is consistent with the WHO's priority pathogen list and suggests that pigeons could pose a risk as carriers of infectious diseases in areas where interactions between humans and pigeons are frequent. The identification of fungal pathogens such as *Aspergillus fumigatus*, *Candida auris* and *Cryptococcus neoformans* in all fecal samples from pigeons of both locations indicates an underestimated risk for public health. These fungi are known for their ability to cause invasive diseases, particularly in low immunity individuals, and could lead to outbreaks if not adequately monitored.

Similarly, the presence of bacterial pathogens, including those resistant to antibiotics, adds more concerns. The detection of carbapenem-resistant *Acinetobacter baumannii* among other bacterial species could indicate the threat of spreading antibiotic resistance which can significantly complicate both the treatment of the diseases caused by environmental exposure, and managing it in high populated areas.

Furthermore, presence of viral pathogens with outbreak potential, brings additional public health risk. The presence of viruses such as the *SARS coronavirus* and *Monkeypox virus* in pigeon feces suggests that pigeons could serve as carriers for viral diseases. This shows the requirement for more research on viral pathogen in urban regions.

The coexistence of these pathogens in urban pigeons faeces reveals the need for a multidimensional strategy. This strategy should aim not only to reduce human-pigeon interactions and manage pigeon populations in sensitive areas but also to increase monitoring of the presence of these pathogens in urban regions. Moreover, raising public awareness is important to inform urban residents about the potential health risks associated with pigeons and to reduce exposure to pathogens.

It is not entirely possible to discount the likelihood that some of the identified organisms might have originated from the soil surrounding the fecal samples. However, our testing approach aligns with standard procedures employed in collecting feral feces from various species. Therefore, we assert that our results are indicative of the microbial content typically present in feral pigeon feces.

Key Points

- Feral pigeons in urban areas can spread harmful pathogens to humans, primarily through fecal contamination.
- This study analyzed pigeon feces from Uppsala's central gathering spots to detect bacterial, viral, and fungal pathogens.
- Metagenomics analysis confirmed the presence of most high-priority viral and bacterial pathogens listed by the World Health Organization (WHO).
- All critical fungal pathogens identified by WHO, including **Aspergillus** and **Candida** species, were found in the samples, posing a risk to individuals with weakened immune system.
- Controlling pigeon populations and keeping them away from sensitive areas, like hospitals, could reduce the spread of these pathogens.

This study lays the foundation for future explorations into the transmission mechanisms and the formulation of strategies to hinder the spread of infections from pigeons to humans. These efforts are crucial in urban planning and public health policymaking, aimed at safeguarding community health, particularly in regions where human-pigeon interactions are commonplace.

Ethical consideration

Pigeons were not disturbed, captured, or harmed and fecal samples were collected after excretion. Therefore, ethical approval was not required.

Acknowledgments

We would like to express our deepest appreciation to the "Oxford Nanopore & NGI/SciLifeLab Project Competition" at the Long Read Conference in Uppsala (LRUA2022) for selecting us as one of the two winners, allowing us to have our samples sequenced for free at SciLife Lab. We also extend our sincere gratitude to the SLU Bioinformatics Infrastructure (SLUBI) for their invaluable support in the analysis of the data.

References

Catroxo, M.H.B., Martins, A.M.C.R.P.F., Petrella, S., Curi, N.A. and Melo, N.A. (2011). Research of Viral Agent in Free-living Pigeon Feces (*Columba livia*) in the City of São Paulo, SP, Brazil, for Transmission Electron Microscopy. International Journal of Morphology, **29**(2), pp.628–635. http://dx.doi.org/10.4067/s0717-95022011000200055.

Borges, C.A., Cardozo, M.V., Beraldo, L.G., Oliveira, E.S., Maluta, R.P., Barboza, K.B., Werther, K. and Ávila, F.A. (2017). Wild birds and urban pigeons as reservoirs for diarrheagenic *Escherichia coli* with zoonotic potential. Journal of Microbiology, **55**(5), pp.344–348. http://dx.doi.org/10.1007/s12275-017-6523-3.

Briscoe, A.G., Nichols, S., Hartikainen, H., Knipe, H., Foster, R., Green, A.J., Okamura, B. and Bass, D. (2021). High-throughput sequencing of faeces provides evidence for dispersal of parasites and pathogens by migratory waterbirds. Molecular Ecology Resources, **22**(4), pp. 1303–1318. http://dx.doi.org/10.1111/1755-0998.13548.

de Vasconcellos, H.V.G., Silva, K.F.B., Montenegro, H., Miguel, C.B., Tizioto, P., Agostinho, F., Araújo, M.C., Ribas, R.M., da Silva, M.V., Soares, S. de C., Rodrigues Júnior, V., Batistão, D.W. da F., Oliveira, C.J.F. and Rodrigues, W.F. (2022). *Staphylococcus aureus* and *Enterococcus faecium* isolated from pigeon droppings (*Columba livia*) in the external environment close to hospitals. Revista da Sociedade Brasileira de Medicina Tropical, **55**. http://dx.doi.org/10.1590/0037-8682-0353-2021.

Freire, S., Grilo, T., Poirel, L. and Aires-de-Sousa, M. (2022). Urban Pigeons (*Columba livia*) as a Source of Broad-Spectrum β-Lactamase-Producing *Escherichia* coli in Lisbon, Portugal. Antibiotics, **11**(10), p.1368. http://dx.doi.org/10.3390/antibiotics11101368.

Glushakova, A.M., Rodionova, E.N. and Kachalkin, A.V. (2021). Yeasts in Feces of Pigeons (*Columba livia*) in the City of Moscow. Current Microbiology, **78**(1), pp.238–243. http://dx.doi.org/10.1007/s00284-020-02251-5.

Graczyk, T.K., Sunderland, D., Rule, A.M., da Silva, A.J., Moura, I.N.S., Tamang, L., Girouard, A.S., Schwab, K.J. and Breysse, P.N. (2007). Urban Feral Pigeons (*Columba livia*) as a Source for Air- and Waterborne Contamination with *Enterocytozoon bieneusi* Spores.

Applied and Environmental Microbiology, **73**(13), pp.4357–4358. http://dx.doi.org/10.1128/aem.00202-07.

Haag-Wackernagel, D. and Moch, H. (2004). Health hazards posed by feral pigeons. Journal of Infection, **48**(4), pp.307–313. http://dx.doi.org/10.1016/j.jinf.2003.11.001.

Jokimäki, J. and Suhonen, J. (1998). Distribution and habitat selection of wintering birds in urban environments. Landscape and Urban Planning, **39**(4), pp.253–263. http://dx.doi.org/10.1016/s0169-2046(97)00089-3.

Kowalczyk, K. and Wójcik-Fatla, A. (2022). *Chlamydia psittaci* in Faecal Samples of Feral Pigeons (*Columba livia forma urbana*) in Urban Areas of Lublin city, Poland. Current Microbiology, **79**(12). http://dx.doi.org/10.1007/s00284-022-03072-4.

Magnino, S., Haag-Wackernagel, D., Geigenfeind, I., Helmecke, S., Dovč, A., Prukner-Radovčić, E., Residbegović, E., Ilieski, V., Laroucau, K., Donati, M., Martinov, S. and Kaleta, E.F. (2009). Chlamydial infections in feral pigeons in Europe: Review of data and focus on public health implications. Veterinary Microbiology, **135**(1–2), pp.54–67. http://dx.doi.org/10.1016/j.vetmic.2008.09.045.

Medina, I.R., Fuentes, L.R., Arteaga, M.B., Valcárcel, F.R., Arbelo, F.A., del Castillo, D.P., Suárez, S.D., Quintana, O.F., Gutiérrez, B.V., Sergent, F.S. and Acosta-Hernández, B. (2017). Pigeons and their droppings as reservoirs of *Candida* and other zoonotic yeasts. Revista Iberoamericana de Micología, **34**(4), pp.211–214. http://dx.doi.org/10.1016/j.riam.2017.03.001.

Mitchell M, Nguyen SV, Macori G, Bolton D, McMullan G, Drudy D, Fanning S. *Clostridioides difficile* as a Potential Pathogen of Importance to One Health: A Review. Foodborne Pathog Dis. 2022 Dec;**19**(12):806-816. http://dx.doi.org/10.1089/fpd.2022.0037.

Moschetti, G., Alfonzo, A. and Francesca, N. (2017). Yeasts in Birds. Yeasts in Natural Ecosystems: Diversity, pp.435–454. http://dx.doi.org/10.1007/978-3-319-62683-3_14.

Naz, S.A., Yaseen, M., Jabeen, N. and Shafique, M. (2017). Isolation of potentially pathogenic fungi from selected pigeons' feeding sites in Karachi: A new dimension to health hazard. JPMA. The Journal of the Pakistan Medical Association, [online] **67**(6), pp.901–906. Available at: https://pubmed.ncbi.nlm.nih.gov/28585590/ [Accessed 7 Dec. 2023].

Nnadi, N.E. and Carter, D.A. (2021). Climate change and the emergence of fungal pathogens. PLOS Pathogens, **17**(4). http://dx.doi.org/10.1371/journal.ppat.1009503.

Nualmalang, R., Thanomsridetchai, N., Teethaisong, Y., Sukphopetch, P. and Tangwattanachuleeporn, M. (2023). Identification of Pathogenic and Opportunistic Yeasts in Pigeon Excreta by MALDI-TOF Mass Spectrometry and Their Prevalence in Chon Buri Province, Thailand. International Journal of Environmental Research and Public Health, **20**(4), p.3191. http://dx.doi.org/10.3390/ijerph20043191.

Phan, T.G., Vo, N.P., Boros, Á., Pankovics, P., Reuter, G., Li, O.T.W., Wang, C., Deng, X., Poon, L.L.M. and Delwart, E. (2013). The Viruses of Wild Pigeon Droppings. PLoS ONE, **8**(9), p.e72787. http://dx.doi.org/10.1371/journal.pone.0072787.

Piret, J. and Boivin, G. (2021). Pandemics Throughout History. Frontiers in Microbiology, [online] **11**(1). http://dx.doi.org/10.3389/fmicb.2020.631736.

World Health Organization. (2022). WHO fungal priority pathogens list to guide research, development and public health action. [online] Available at: https://www.who.int/publications/i/item/9789240060241 [Accessed 7 Dec. 2023].

EMBnet.journal
Bioinformatics in Action

Paz-y-Miño et al. (2025) EMBnet.journal 30, e1073
http://dx.doi.org/10.14806/ej.30.0.1073

Research Papers

Genetic and Genomics aspects of susceptibility and resistance to infections

Cesar Paz-y-Miño[1]✉, Javier Gualdron[2], Ariane Paz-y-Miño[3]

[1] Equinoccial Technological University - Faculty of Health Sciences «Eugenio Espejo», Ecuador
[2] National blood center of the Ecuadorian Red Cross-Laboratory of Molecular Biology, Ecuador
[3] Mass General Brigham Salem Hospital, United States

Competing interests: CPM none; JG none; APM none

Abstract

This article addresses the genetic and genomic determinants that influence susceptibility or resistance to human infections, whether caused by bacteria, viruses, fungi, parasites or even prions. A large number of genes are involved in susceptibility or resistance, and this is because *Homo sapiens sapiens* evolutionarily acquired key DNA modifications adaptable to defense as a result of the interaction of infectious agents with their genetic properties, natural selection and mutations. This article provides a description of the methods used in the investigation of this evolution and mutual host-infectious agent adaptation. The genes involved in susceptibility and resistance are presented, as well as the polymorphic variants that provide greater or lesser reaction capacity. The high number of genes, which produce specific defense proteins, have intricate relationships among them. Relationships among these genes have been analysed. A list of variants, and even mutations of the 70 to 153 proteins/genes that are involved in the immune response to infection diseases has been compiled. Additionally, we designed a network of physical interactions as predictable, with BioGrid data and enrichment data obtained by the g:Profiler platform, finding 57 of the 70 genes with protein-to-protein interactions. The network detected is composed of 1,049 interacting genes (in total 1,106 genes and 1,910 interactions). This analysis shows the importance of immunity genes in the defense against infectious agents, as well as the effects of the genes involved such as *HLA*, immunoglobulins, interleukins, immune cells, among others. Based on the functional analysis of genes involved in susceptibility and resistance to infections, we compiled a list of genetic disorders that increase an individual's risk of developing infectious diseases. It is known that infectious diseases manifest general clinical characteristics, which are a good guide to suspect a disease. Furthermore, we describe general clinical signs that may suggest underlying genetic immunodeficiency, and highlight specific infectious diseases—such as HIV/AIDS, amebiasis, malaria, Chagas disease, tuberculosis, and COVID-19—where genetic susceptibility factors have been well characterized.

Introduction

Infectious diseases continue to be one of the leading causes of morbidity and mortality worldwide and the variability in susceptibility or resistance of individuals to these infections highlights the role of genetic factors. The ability of the immune system to respond to infectious agents such as viruses, bacteria, fungi, parasites and even prions is influenced by genetic variants, which may predispose to develop infections or, conversely, confer resistance to these (Bos *et al.*, 2019; Quintana-Murci, 2019).

The genetics of human immunity has been shaped throughout evolution, in a process of natural selection and adaptation to diverse environments. This process has resulted in key modifications to the DNA of *H. sapiens*, allowing specific genetic variants to optimise the immune defense against infectious agents (Quintana-Murci, 2019). In addition, the environment in which individuals live, such as tropical or temperate zones and interactions with other species have influenced the genetic diversity of immunity. Thus, coevolution with pathogens has imposed significant selective pressure on certain genes, favoring the permanence of variants that increase survival in contexts of high infectious exposure (Vasseur and Quintana-Murci, 2013; Quintana-Murci, 2019).

Several genes influence susceptibility and resistance to AIDS-causing HIV. The absence of TNPO3 confers

Article history
Received: 24 April 2025
Accepted: 06 July 2025
Published: 07 October 2025

Paz-y-Miño et al. (2025) EMBnet.journal 30, e1073
http://dx.doi.org/10.14806/ej.30.0.1073

absolute resistance, while other genes such as CPSF6 and SC35 are also involved. Mutations in CCR5, CXCR4 and SDF1 receptors increase susceptibility in homozygosis. In Ecuador, resistance mutations are present in 0.5% (CCR5), 16% (CCR2) and 48% (SDF1), but only in heterozygosis, which does not guarantee total resistance (Paz-y-Mino Cepeda *et al.*, 2005). In Europe, genetic resistance is 26%, while in Latin America it is only 4% (Chapman and Hill, 2012; McLaren *et al.*, 2015).

Amoeba infections are a public health problem, influenced by innate and adaptive immunity genes. Mutations in IL8 and TNF-α increase the risk of *Entamoeba histolytica* infection; variations in MHC affect susceptibility, since this gene regulates the immune response. On the other hand, inflammatory genes such as IL1β and TLR2 are also associated with increased risk and mutations in epithelial barrier genes, such as filaggrin and TLR4, favor infection (Haque *et al.*, 2002; Guo *et al.*, 2011; Robledo *et al.*, 2018). Whether the host genome influences the localization of infection, along with parasite factors, is debated.

Malaria, caused by *Plasmodium* and transmitted by *Anopheles*, is influenced by genetic and environmental factors. In Ecuador, 1,946 cases were reported in 2020, mainly due to *P. vivax* (Muñoz Cabas *et al.*, 2023). The mutation in the HBB gene (HbS) confers resistance to *P. falciparum* in heterozygotes, but increases the risk in homozygotes (Hill *et al.*, 1991; Malaria Genomic Epidemiology Network *et al.*, 2019; Muñoz Cabas *et al.*, 2023). In Esmeraldas, 24.3% of the population has hemoglobin variants, with Hb.AS as the most common, other mutations in HBC, HBE and G6PD also influence resistance, with a G6PD prevalence of 10% in endemic areas. *Plasmodium vivax*, is found in 11% of the Afro-Ecuadorian population and genes such as TLR1, IL10 and PfEMP1, together with MHC and TNF, also affect susceptibility and severity of malaria (Auburn and Barry, 2017; Muñoz Cabas *et al.*, 2023).

Chagas disease, caused by *Trypanosoma cruzi*, affects about 8 million people worldwide. In Ecuador, 113 cases were reported in 2020, mostly chronic (Dumonteil *et al.*, 2016). Genetic factors influence susceptibility and severity, including immune system genes, MHC, IL-10, IL-6 and SOD, which affect the inflammatory response and oxidative stress (Dumonteil *et al.*, 2016). Variants in POLD1 and repair genes may contribute to cell damage and a study of SNPs found no association with cardiomyopathy, but identified 44 SNPs associated with various disease traits (Vasconcelos *et al.*, 2012; Deng *et al.*, 2013; Frade *et al.*, 2013).

Tuberculosis is a multifactorial disease influenced by genetic and environmental factors. In Ecuador, in 2020, the incidence was 24 cases per 100,000 inhabitants, with an increase of 34% in deaths and 9.4% in cases (Anon, 2018). Genes such as NRAMP1 and IFNG are associated with increased susceptibility, affecting the immune response(McNicholl *et al.*, 2000; Davila *et al.*, 2008). TNF-α, TLRs, CISH and TYK2 also play a role, the latter with variants that have reduced incidence in the last 4,000 years, and VDR, IL10, IL12A, IL12B, IL6, IL17A and IL17F impact immunity and risk of pulmonary tuberculosis (Hawn *et al.*, 2007; African TB Genetics Consortium *et al.*, 2010; Cholo *et al.*, 2015; Curtis *et al.*, 2015).

Susceptibility to COVID-19 is influenced by genetic factors such as blood type, ethnicity, autoimmunity genes, HLA, and heart or kidney failure genes. The ACE2 receptor, located on chromosome Xp22.2, is key in the entry of SARS-CoV-2 into lung cells, likewise, a study identified 45 proteins associated with susceptibility, highlighting 11 with strong interactions, such as ACE2, AGT, AGTR1, REN and DPP4, related to blood pressure regulation and hormone metabolism (Debnath *et al.*, 2020; Paz-y-Miño *et al.*, 2021; Gupta *et al.*, 2022; Saengsiwaritt *et al.*, 2022; Pecoraro *et al.*, 2023). These interactions explain the clinical heterogeneity of the disease and its impact on different body systems.

The study of the genetic determinants associated with the immune response to infections not only provides a basis for a better understanding of the pathogenesis of these diseases, but also opens the possibility of developing personalised therapies. Several genes and their variants are associated with susceptibility or resistance to specific infections. For example, major histocompatibility complex (HLA) genes, Toll-like receptors (TLRs) and cytokines play an essential role in the innate and adaptive response, affecting the immune system's ability to recognize and destroy pathogens (Merker *et al.*, 2020; Martins *et al.*, 2023; Paz y Miño Cepeda, 2024).

This study analyzes the key genes involved in the immune response to infections and describes how the construction of protein-protein interaction networks, or interactomes, helps to understand the functional relationships between them. By analyzing these networks, we identify genes that act as "hubs" or central nodes in the immune response, revealing their role in resistance or susceptibility to infections such as HIV, tuberculosis, malaria, and COVID-19 (Casadevall and Pirofski, 2000; Rast and Litman, 2010).

Materials and methods

Gene Selection

To identify genes related to resistance and susceptibility to infections we considered studies describing genes and genetic variants associated with changes in gene expression, alterations in RNA structure, modifications in immunoglobulin conformation, modulation of the immune response, abnormal interactions with other genes and the influence of environmental factors (Janeway and Medzhitov, 2002; Khor and Hibberd, 2012; Rolland *et al.*, 2014; Mozzi *et al.*, 2018; Luck *et al.*, 2020; Nahon and Cobat, 2020).

A systematic review of scientific literature was carried out in databases such as PubMed and Scopus, using key terms such as immunity genes, genetic susceptibility and infections. The selection criteria

included: (i) functional association with infections, (ii) evidence of relevant genetic polymorphisms, and (iii) involvement in immune processes. The 70 selected genes were categorised into innate, adaptive, inflammatory and antiviral immune responses. Table S1a[1] presents some groups of genes relevant in this context, giving an overview of their contribution to resistance and susceptibility to infections, while Table S2a[2] presents the genes selected for this study.

To solve the problem related to genes codifying for different proteins but having the same names, we have consulted GeneCards (gencard.com)(Stelzer *et al.*, 2016). Here it is possible to locate and designate each protein with its proper gene using its accession number in the UniProt[3] database (The UniProt Consortium *et al.*, 2023; in Supplementary Statistical Material, Table S1[4]).

Construction of Protein-Protein Interaction (PPI) Networks

Experimental and predicted data from BioGrid (60) and STRING (Oughtred *et al.*, 2021; Szklarczyk *et al.*, 2023) were used to map protein-protein interactions. PPI networks were generated using the Cytoscape software (Doncheva *et al.*, 2019), following the steps:

- Nodes represented genes/proteins and edges indicated physical or functional interactions.
- The MCODE algorithm was used to identify functional clusters in the network.
- Interactions were filtered for reliability according to the scores provided by STRING.

Functional Enrichment

Functional analysis was performed with Gprofiler (Kolberg *et al.*, 2023), considering statistically significant terms ($p<0.05$) from Gene Ontology (GO) (in Supplementary Statistical Material, Table S3[5]), KEGG (in Supplementary Statistical Material, Table S4[6]) and Reactome (in Supplementary Statistical Material, Table S5[7]) (Ashburner *et al.*, 2000; Kanehisa, 2000; Rothfels *et al.*, 2023). Pathways related to biological processes, such as cytokine activation, immune signaling and inflammatory regulation were evaluated. Enriched terms were visualised using Manhattan diagrams generated in Cytoscape (Doncheva *et al.*, 2019).

Topological Analysis

Key topological metrics of PPI networks were analysed to identify critical nodes ("hubs") using Cytoscape (in Supplementary Statistical Material, Table S2[8]) (Doncheva *et al.*, 2019).

- Degree of connection: nodes with the highest number of interactions.
- Intermediation centrality: nodes that acted as control points between interactions.
- Identification of key genes such as TNF, HBB and CD79B, prioritised for their impact on immunity.

Results and discussions

Susceptibility- and resistance-related proteins listed in Table S2a[9] were subjected to a bioinformatics analysis of protein-protein interactions (PPIs), defined as highly specific physical contacts between two or more proteins, driven by electrostatic forces, hydrogen bonds, or hydrophobic effects. These interactions were used to construct a protein interaction network, or interactome (Figure 1)(Rolland *et al.*, 2014; Luck *et al.*, 2020). The interactome was generated using the STRING database, and may serve as a reference model for understanding complex protein interactions involved in the immune response to infections (Szklarczyk *et al.*, 2023).

Related pathologies

The PPI network in Figure 1 reveals that among all the proteins analysed (n=70) there are 624 interactions, with a total correlation score of 0.669. Such correlation indicating that the proteins are at least partially biologically connected, as a group. Each protein involved in the interactome has a known function and these functions are observed in different pathological manifestations (Stelzer *et al.*, 2016; Szklarczyk *et al.*, 2023; The UniProt Consortium *et al.*, 2023). Table 1 shows the diseases related to their anomalous interactions, therefore, product of mutations or genetic variants, resulting from the analysis of the proteins/genes interactome produced.

Additionally, the ontological analysis carried out by the String Software's of the 70 genes analised, yields at least 17 functional nodes congruent with different functional pathways such as: Immune receptor activation, transmembrane signaling receptor activations, virus receptor activation, 2-5-oligoadenylate synthetase activation, activation of regulators of molecular

[1]http://journal.embnet.org/index.php/embnetjournal/article/downloadSuppFile/1073/1073_supp_S1a

[2]http://journal.embnet.org/index.php/embnetjournal/article/downloadSuppFile/1073/1073_supp_S2a

[3]www.uniprot.org

[4]http://journal.embnet.org/index.php/embnetjournal/article/downloadSuppFile/1073/1073_supp_1

[5]http://journal.embnet.org/index.php/embnetjournal/article/downloadSuppFile/1073/1073_supp_1

[6]http://journal.embnet.org/index.php/embnetjournal/article/downloadSuppFile/1073/1073_supp_1

[7]http://journal.embnet.org/index.php/embnetjournal/article/downloadSuppFile/1073/1073_supp_1

[8]http://journal.embnet.org/index.php/embnetjournal/article/downloadSuppFile/1073/1073_supp_1

[9]http://journal.embnet.org/index.php/embnetjournal/article/downloadSuppFile/1073/1073_supp_S2a

Paz-y-Miño et al. (2025) EMBnet.journal 30, e1073
http://dx.doi.org/10.14806/ej.30.0.1073

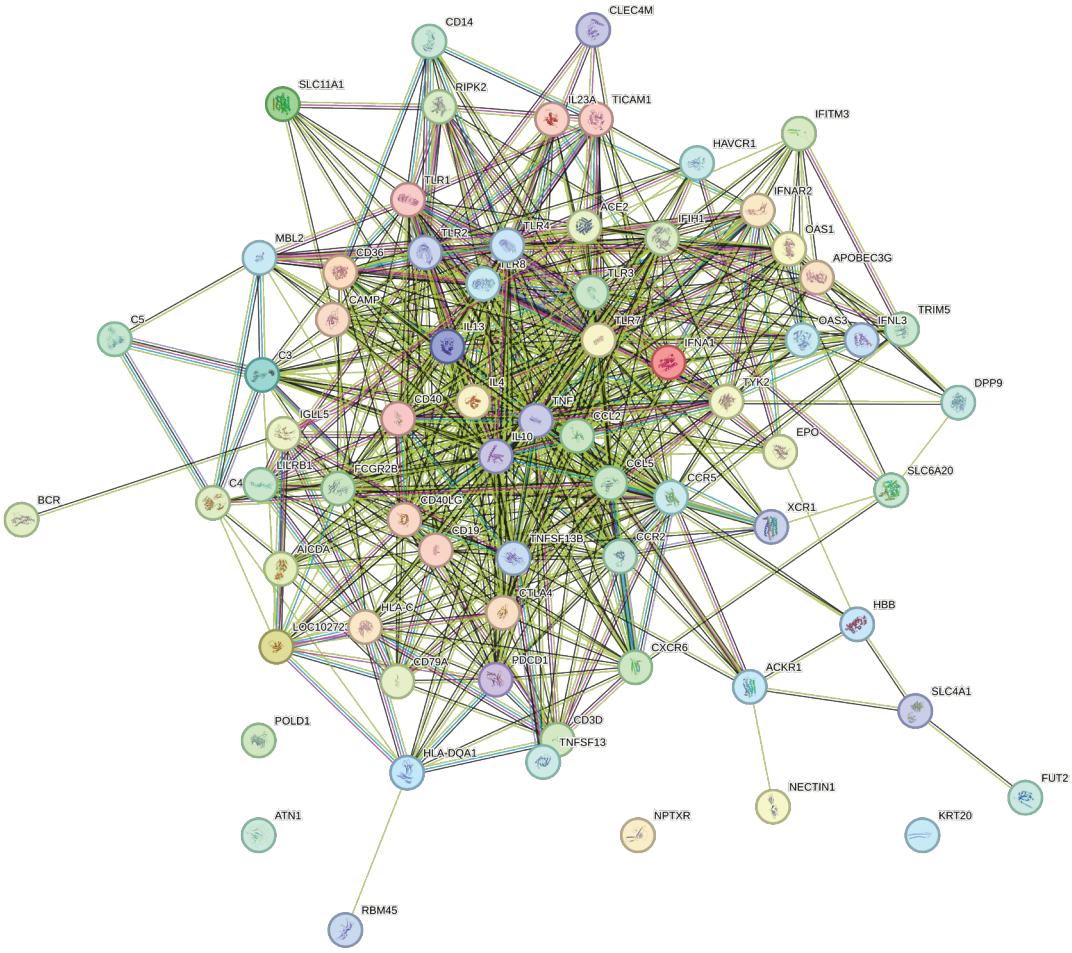


Figure 1. Interactome of proteins related to resistance and susceptibility to infections.

functions, activation of NAD(P) nucleosidase and cyclic ADP-ribose generating NAD nucleotidase, cytokine receptor binding and cytokine activation, signaling receptor binding and receptor activation, protein binding and identical protein binding, antigen binding, tumor necrosis factor receptor binding, peptide and peptide antigen binding, lipopeptide binding, lipopolysaccharide binding and immune receptor activity, CC chemokine binding and receptor activation, beta amyloid binding and CCR2 chemokine receptor binding. This means that genes and their proteins are interrelated as common triggering groups of various functional pathways of efficient or inefficient defense against infectious agents (Rolland *et al.*, 2014; Luck *et al.*, 2020; Szklarczyk *et al.*, 2023).

Enrichment analysis and network evaluation

Subsequently, all interactions described for this dataset were analysed using the BioGrid database and all interaction information, both physical and predicted, was downloaded. An enrichment analysis was applied to evaluate both functional and signaling pathway concepts (Oughtred *et al.*, 2021) identifying 57 of the 70 genes described in Table S2a[10]. The analysis in BioGrid added 1,049 genes (nodes) that interact with those identified by us, allowing to construct a network of 1,106 genes and 1,910 interactions. The integration and analysis of this information was performed using Cytoscape and applying a topological analysis to identify key structural features of the network. In addition, we used the g:Profiler platform to interpret overrepresented biological pathways. This analysis considered statistically significant enrichment terms focused on biological processes and metabolic pathways from the KEGG (Kyoto Encyclopedia of Genes and Genomes) and Reactome databases (Figure 2). The interaction network enrichment analysis statistically emphasizes immune-related processes and signaling pathways of the mainly innate immune system, underlining the important role of the network in mediating immune responses and interactions with biotic stimulus (Kanehisa, 2000; Doncheva *et al.*, 2019; Oughtred *et al.*, 2021; Kolberg *et al.*, 2023).

The topological analysis of the protein interaction network provides information on the structural and functional organisation of the network (Figure 2 A). Of

[10]http://journal.embnet.org/index.php/embnetjournal/article/downloadSuppFile/1073/1073_supp_S2a

Table 1. Diseases related to abnormal protein-gene interactions and infection predisposition genes

Disease	Monogenic (M) and/or Polygenic (P) Component*
X-linked Hyper IgM Syndrome	M
Immunodeficiency with Hyper IgM Type 3	M
Common Cold	P
Pulmonary Eosinophilia	P
Adult Respiratory Distress Syndrome	P
Stevens-Johnson Syndrome	M
Rosacea	P
Influenza	P
Membranous Glomerulonephritis	P
Intrinsic Asthma	P
Behçet's Disease	M
Allergic Asthma	P
Myocarditis	P
Crohn's Disease	P
Infectious Disease Due to Parasites and Protozoa	P
Glomerulonephritis	P
Hypersensitivity Disease and Type IV Reaction	P
Pneumonia	P
Helminth Disease	P
Bronchial Disease	P
Primary Bacterial Infection	P
Infectious Agent Disease	P
Upper Respiratory Tract Disease	P
Bacterial Infections	P
Allergic Diseases	M, P
Lower Respiratory Tract Diseases	P
Immune System Diseases	M, P
Urinary System Diseases	M, P
Intestinal Diseases	M, P
Connective System Disea;ses	M
Diseases of the Gastrointestinal System	M, P

Disease	Monogenic (M) and/or Polygenic (P) Component*
Guillain-Barré Syndrome	M, P
Genital Herpes	P
Cutaneous Leishmaniasis	P
Visceral Leishmaniasis	P
Hyperimmunoglobulin Syndrome	M
Mantle Cell Lymphoma	M, P
Acquired Immune Deficiency Syndrome	P
Conjunctivitis	P
Lymphocytopenia	P
Leprosy	P
COVID-19	P
Allergic Rhinitis	M, P
Encephalomyelitis	P
Vasculitis	P
B-cell Deficiency	M
Leukopenia	P
Viral Infections	P
Parasitic Infections	P
Asthma	P
Leukocyte Disease	M, P
Liver Disease	M, P
Primary Immunodeficiency	M
Pulmonary Disease	P
Dermatitis	P
Renal Disease	M, P
Skin Diseases	M, P
Vascular Diseases	M, P
Carcinoma	M, P
Various Genetic Diseases	M
Human Immunodeficiency Virus	M, P

*Each disease listed is associated with either monogenic (M) or polygenic (P) mutations that influence susceptibility to infections. These genetic variations can impact disease severity and symptomatology. For example, the CCR5 gene provides resistance to HIV infection in homozygous individuals but not in heterozygous ones, illustrating a monogenic effect. However, other genes also contribute to susceptibility or resistance, demonstrating the polygenic nature of many conditions.

all the metrics yielded by the analysis, we used the degree of interaction of the nodes to define the "centers" or Hubs of connection in the network and the "Intermediation Centrality" as a measure of the frequency with which a node appears in the shortest paths between other nodes (Figure 2 B). High values in this metric define critical control points or bottlenecks in the network (Wu *et al.*, 2008; Koh *et al.*, 2012). We found that CD79B, TNF, TYK2, POLD1, BCR and HBB genes are the "hubs" with the highest degree of connection in the analysed network. The identification of these "hub nodes" underlines their possible roles in immune response and pathogen defense mechanisms. CD79B and BCR are integral parts of the B-cell receptor complex and influence adaptive immune responses (Pleiman *et al.*, 1994; Tanaka and Baba, 2020), while TNF is a fundamental cytokine involved in systemic inflammation (Bulló *et al.*, 2003), crucial for infection control, but also implicated in inflammatory diseases (Rickert, 2013). TYK2, part of the JAK-STAT signaling pathway, is key to the signaling of several type I interferons and cytokines, influencing both innate and adaptive immunity (Villarino *et al.*, 2017). POLD1, although primarily associated with DNA replication and repair (Nicolas *et al.*, 2016), may have functions in the immune system that are less studied, but potentially significant, particularly in terms of genomic stability in rapidly proliferating immune cells (Nichols-Vinueza *et al.*, 2021). Finally, HBB, part of the hemoglobin complex, is critical in oxygen transport, but also plays a role in modulating the response to oxidative stress during infections (Ma, 2013; Ramezani *et al.*, 2018). These "hubs", being highly interconnected within the

Paz-y-Miño et al. (2025) EMBnet.journal 30, e1073
http://dx.doi.org/10.14806/ej.30.0.1073

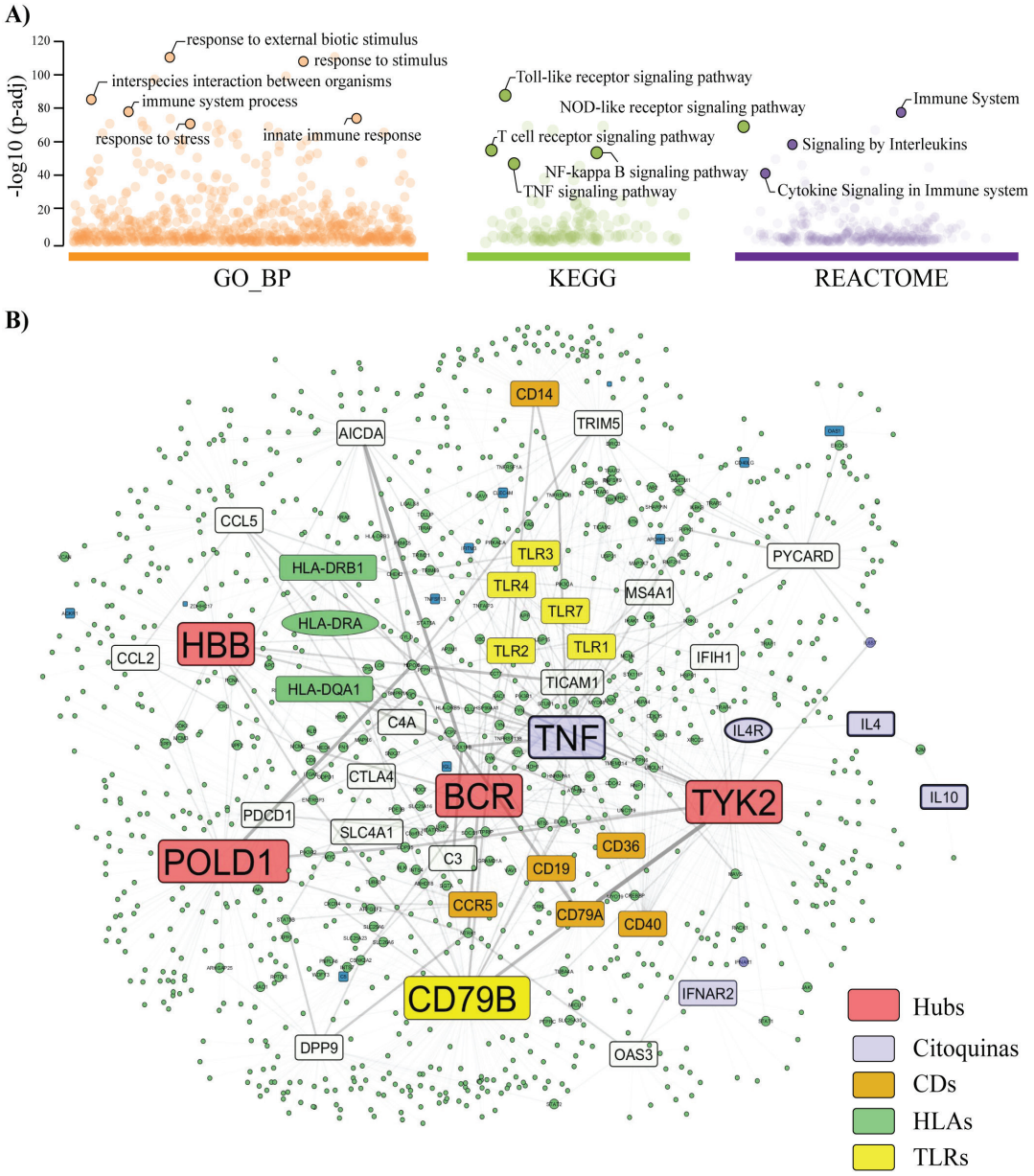


Figure 2. The topological analysis of the protein interaction network.

Note* Enrichment analysis and protein-protein interaction network associated with susceptibility to infection. A) Manhattan-type graphical representation showing the most significant ontology terms derived from the enrichment analysis. Terms are color-coded according to their origin: molecular function terms from the Gene Ontology (GO) in red, signaling pathways from the KEGG database in green and pathways from the REACTOME database in purple. B) Protein-protein interaction network constructed from BioGrid database interaction data. The topology of the network highlights the central nodes (red boxes) according to their degree of interaction with the proteins studied and their association with different cytokines (purple), clusters of differentiation (CDs) (red), major histocompatibility complex (HLA) molecules (green), and Toll-like receptors (TLRs) (yellow). The size of the nodes reflects the "Centrality of Intermediation" index within the network.

network, suggest that they not only play individual roles in the response to infection, but may also interact synergistically, offering potential strategic targets for improving resistance to infection and understanding mechanisms of susceptibility.

Immune cells, such as lymphocytes, macrophages and dendritic cells, detect and destroy invading pathogens. Immune system proteins, such as immunoglobulins, interferons and cytokines, help regulate and coordinate the immune response. Failures in the genetic construction of these elements are associated with the development of disease (Janeway and Medzhitov, 2002; Khor and Hibberd, 2012; Mozzi *et al.*, 2018; Rotival, 2019; Nahon and Cobat, 2020). Interestingly, protein interaction network analysis significantly weights cytokines, human leukocyte antigen (HLA) genes and Toll-like receptors (TLRs) as genes associated with fundamental mechanisms of resistance or susceptibility to infection (Szklarczyk *et al.*, 2023). Each group of proteins contributes uniquely to the

immune system's ability to defend against pathogens are presented in Table S3a[11].

Cytokine are signaling molecules that mediate and regulate immunity, inflammation and hematopoiesis (Kany *et al.*, 2019). Genetic variants in this group of proteins, such as in the gene encoding the interleukin-23 (IL-23) receptor, have been linked to susceptibility to inflammatory bowel disease and fungal infections. Likewise, genetic variants in interleukin-6 (IL-6) and tumor necrosis factor-alpha (TNF-α) are associated with increased disease severity in patients with COVID-19 (Khor and Hibberd, 2012; Mozzi *et al.*, 2018; Horwood *et al.*, 2019; Masin *et al.*, 2022). The presence of IFNAR1, IFNAR2, IL10, IL4, IL4R, IL6ST and TNF in our interaction network indicates their role in orchestrating immune responses. Cytokines such as TNF and IL6 participate in the acute inflammatory response, which is crucial for controlling the spread of infection (Rankin, 2004). In contrast, regulatory cytokines such as IL10 and IL4 can modulate the immune response to prevent over activation, which can lead to tissue damage or autoimmune disorders (Couper *et al.*, 2008; Junttila, 2018). The balance in cytokine signaling may dictate the outcome of an infection, so the study of the aforementioned proteins may improve the description of effective pathogen clearance processes or susceptibility due to an uncontrolled inflammatory response.

Major histocompatibility complex (MHC) proteins, also known as the human leukocyte antigen (HLA) group, play a crucial role in the immune response by facilitating the recognition of foreign molecules by the immune system. These proteins are essential for the presentation of antigens to T cells, enabling the identification and elimination of pathogens (Chaplin, 2010). In the context of Lyme disease for example, a tick-bite-borne infection caused by the bacterium *Borrelia burgdorferi*, certain genetic variants in the HLA genes, specifically HLA-DRB1, HLA-DQA1 and HLA-DQB1, have been found to be linked to an increased susceptibility to this disease (Mozzi *et al.*, 2018; Rotival, 2019). Molecular variants in these genes significantly influence the host's ability to respond to infection, highlighting the importance of genetic polymorphisms in the pathogenesis of infectious diseases.

On the other hand, Toll-like receptors (TLRs) are fundamental components of the innate immune system, acting as essential molecular sensors in the detection of pathogens. These receptors identify highly conserved structures called pathogen-associated molecular patterns (PAMPs), characteristics of invading microorganisms. When activated, TLRs initiate signaling cascades that elicit inflammatory responses and trigger crucial defense mechanisms, including cytokines production and effector cells activation (Kumar *et al.*, 2009; Kawai and Akira, 2010, 2011). This rapid and widespread responsiveness is vital for immediate defense against infection and facilitates the transition between innate and adaptive immunity by influencing the activation and maturation of cells of the adaptive immune system (O'Neill *et al.*, 2013). Genetic variations in the genes encoding for TLRs can significantly modify the efficacy of the innate immune system in recognising and responding appropriately to PAMPs (Vijay, 2018). Such variations can result in attenuated or over-activated immune responses, increasing susceptibility to infections, as in the case of tuberculosis or predisposing to chronic inflammatory responses (Khor and Hibberd, 2012; Walker *et al.*, 2015; Mozzi *et al.*, 2018; Paz-y-Miño *et al.*, 2021). Our network analysis identifies TLR1, TLR2, TLR3, TLR4 and TLR7 as central nodes in pathogen susceptibility. These TLRs are crucial not only for the recognition of pathogen-associated molecular patterns, but also for triggering innate immune responses and subsequent modular adaptive responses. Their central position in the network underscores their pivotal role in the initial detection of and response to infectious agents (Janeway and Medzhitov, 2002; Khor and Hibberd, 2012; Masin *et al.*, 2022).

Weighing these groups in network analysis provides deep insight into the collective and individual contributions to immune system functionality, highlighting how imbalances or deficiencies in any of these components could lead to increased susceptibility to infections, while optimal functioning and interaction confer resistance. Understanding these dynamics is essential for developing interventions that aim to stimulate immune responses and effectively manage infection-related diseases.

Identification of disease clusters

Central clusters within the network were identified using Mcode and possible disease associations with the analysed network were explored. The analysis of clusters and diseases with high significance in the analysed network is presented in Table 2.

Many of the diseases presented in Table 2 are caused by microorganisms, and genetic variability among them can also influence their aggressiveness or resistance to antimicrobials. In the adaptive game of evolution, infectious agents do their best to escape the immune system (selective pressure), and to do so they frequently change their genomes(Khor and Hibberd, 2012; Mozzi *et al.*, 2018; Masin *et al.*, 2022). This is precisely the reason for the difficulty of effective treatments, new resistances and the impossibility of designing vaccines.

Immune cells, such as lymphocytes, macrophages and dendritic cells, detect and destroy invading pathogens. Immune system proteins, such as immunoglobulins, interferons and cytokines, help regulate and coordinate the immune response. Failures in the genetic construction of these elements are associated with the development of disease (Janeway and Medzhitov, 2002; Khor and Hibberd, 2012; Mozzi *et al.*, 2018). Genetic and functional determinants of defence

[11]http://journal.embnet.org/index.php/embnetjournal/article/downloadSuppFile/1073/1073_supp_S3a

Paz-y-Miño et al. (2025) EMBnet.journal 30, e1073
http://dx.doi.org/10.14806/ej.30.0.1073

Table 2. Diseases associated with gene network clusters.

GO/KEGG Term	Adjusted p-value
Hepatitis B	1.14E-30
Epstein-Barr virus infection	2.35E-29
Lipids and atherosclerosis	5.92E-26
Measles	2.94E-24
Kaposi's sarcoma-associated herpesvirus infection	9.60E-22
Influenza A	5.77E-21
Hepatitis C	9.77E-21
Shigellosis	1.48E-19
Toxoplasmosis	4.14E-19
Human cytomegalovirus infection	9.70E-19
Yersinia infection	1.20E-17
Viral carcinogenesis	1.96E-16
Tuberculosis	5.02E-16
Human T-cell leukemia virus infection	6.80E-16
Chronic myeloid leukemia	2.21E-15
Human immunodeficiency virus infection	9.08E-15
Coronavirus disease - COVID-19	1.47E-14
Salmonella infection	1.73E-13
Human papillomavirus infection	4.66E-12
Necroptosis	9.04E-12
Chagas disease	9.31E-12
Alcoholic liver disease	3.38E-10
Pancreatic cancer	5.44E-10
Leishmaniasis	8.10E-10
Shear stress and atherosclerosis	2.57E-09
Prostate cancer	2.85E-09
Small cell lung cancer	3.40E-09
Renal cell carcinoma	4.82E-05
Malaria	2,56E-01
Acute myeloid leukemia	3,69E-01
Glioma	6,18E-01
Non-small cell lung cancer	0.00013277

mechanisms allow most people to recover fully from infections, although there are cases in which an infection can have serious or even fatal consequences. The groups of genes with a role in the resistance or susceptibility to infections are presented in Table S1a[12].

The genetic variants involved in the impairment of the immune response to infections include changes in the immunoglobulin chain configuration, B cell behavior and elevation, HLA system organisation, participation in the inflammatory response, complement system activities, cytokine, receptor and enzymes involved in immunoglobulin maturation molecular conformation (Janeway and Medzhitov, 2002; Khor and Hibberd, 2012; Rotival, 2019).

Susceptibility to infectious diseases may be due to mutations in genes encoding defense proteins. Some clear examples such as mutations in the gene encoding the interleukin-12 receptor, interleukin-1 and variants in the TLR2 gene, which encodes the Toll-like receptor 2, have been associated with susceptibility to tuberculosis infection in at-risk populations (McNicholl *et al.*, 2000; Hawn *et al.*, 2007; Davila *et al.*, 2008; Curtis *et al.*, 2015). The gene encoding the interleukin-23 receptor (IL23R) is implicated in susceptibility to inflammatory bowel disease and fungal infections. Additionally, IL23R is a cytokine that stimulates the production of inflammatory cytokines and activates immune cells, which help fight fungal infections (Junttila, 2018; Kany *et al.*, 2019).

The FUT2 gene, which encodes an enzyme involved in the synthesis of cell surface antigens, is associated with increased susceptibility to norovirus and rotavirus infections. Individuals with mutations in the FUT2 gene cannot produce certain antigens that fight these viral infections (Mozzi *et al.*, 2018; Masin *et al.*, 2022).

The IFITM3 gene encodes a protein that helps prevent entry of the influenza virus into cells, individuals who carry one copy of a gene variant have an increased risk of infection and hospitalisation for influenza (Mozzi *et al.*, 2018; Horwood *et al.*, 2019; Rasch, 2019).

In Lyme disease caused by the tick-borne bacterium *Borrelia burgdorferi*, genetic variants in the HLA-DRB1, HLA-DQA1 and HLA-DQB1 genes, which code for major histocompatibility complex (MHC) molecules, are known to be associated with increased susceptibility to this disease (De Los Rios *et al.*, 2015; Mozzi *et al.*, 2018).

Genetic variants in the genes encoding for the cytokines IL6 and TNF-α are associated with increased disease severity in patients with COVID-19, so identification of these genetic variants may help to better understand the pathogenesis of the disease and develop new therapies (Khor and Hibberd, 2012; Walker *et al.*, 2015; Mozzi *et al.*, 2018; Debnath *et al.*, 2020; Masin *et al.*, 2022).

Evolutionarily, genetic variation involves the suitable or unsuitable production of immunoglobulins, resulting in greater antibody diversity, greater or lesser affinity and specificity of these antibodies and individual immune response (Bachman and LeBar, 2018; Nielsen and Boyd, 2018). Table S4a[13] shows some diseases of genetic origin and their international classification (OMIM) (Amberger and Hamosh, 2017), which confer greater predisposition to infections.

The common symptoms and signs that raise suspicion of a genetic deficiency or variant predisposing to infections, or a genetic disease with a risk of infections, are shown in Table S5a[14] (Nielsen and Boyd, 2018; Carroll and Pfaller, 2023).

[12]http://journal.embnet.org/index.php/embnetjournal/article/downloadSuppFile/1073/1073_supp_S1a

[13]http://journal.embnet.org/index.php/embnetjournal/article/downloadSuppFile/1073/1073_supp_S4a

[14]http://journal.embnet.org/index.php/embnetjournal/article/downloadSuppFile/1073/1073_supp_S5a

Conclusions

This study analysed the genetic basis of human susceptibility and resistance to infections by identifying key genes, their polymorphic variants, and the protein–protein interaction networks shaping the immune interactome. The aim was to understand how genetic architecture influences immune defense mechanisms and to provide novel interpretative approaches through bioinformatic and topological analyses.

Notably, we constructed an extensive interaction network comprising 1,106 genes and 1,910 interactions, identifying central hub genes such as CD79B, TNF, TYK2, POLD1, and HBB. These genes, although known individually, had not previously been jointly prioritised as immune control nodes in the literature. Moreover, the integration of databases such as BioGrid and g:Profiler enabled functional enrichment analysis, highlighting immune-related pathways including cytokine signaling, Toll-like receptor activation, and the MHC complex.

These findings provide a comprehensive framework for future investigations into the molecular mechanisms of immune response to infection. They may contribute to the development of personalised therapies, identification of genetic susceptibility biomarkers, and a better understanding of infectious diseases through the lens of evolutionary biology and precision medicine.

Key Points

- Human susceptibility or resistance to infections is strongly influenced by genetic variants that modulate immune response.
- An immune interactome was constructed with 1,106 genes and 1,910 interactions, identifying key hub genes such as TNF, CD79B, and TYK2.
- Functional pathway analysis revealed enrichment in immune-related processes including cytokine activation, TLR signaling, and the MHC complex.
- Specific genetic variants are associated with diseases such as HIV/AIDS, tuberculosis, malaria, and COVID-19.
- Understanding these genetic interactions provides valuable insights into personalised medicine in the context of infectious diseases.

The dataset and prioritised genes proposed here may serve as a valuable resource for translational and clinical research in immunogenetics.

References

African TB Genetics Consortium, The Wellcome Trust Case Control Consortium, Thye T, Vannberg FO, Wong SH, Owusu-Dabo E, et al. Genome-wide association analyses identifies a susceptibility locus for tuberculosis on chromosome 18q11.2. Nat Genet. 2010;42:739–741. http://dx.doi.org/10.1038/ng.639.

Akbay B, Germini D, Bissenbaev AK, Musinova YR, Sheval EV, Vassetzky Y, et al. HIV-1 Tat Activates Akt/mTORC1 Pathway and AICDA Expression by Downregulating Its Transcriptional Inhibitors in B Cells. Int J Mol Sci. 2021;22:1588. http://dx.doi.org/10.3390/ijms22041588.

Amato KR, Jeyakumar T, Poinar H, Gros P. Shifting Climates, Foods, and Diseases: The Human Microbiome through Evolution. BioEssays. 2019;41:1900034. http://dx.doi.org/10.1002/bies.201900034.

Amberger JS, Hamosh A. Searching Online Mendelian Inheritance in Man (OMIM): A Knowledgebase of Human Genes and Genetic Phenotypes. Curr Protoc Bioinforma [Internet]. 2017 [cited 2024 Nov 5];58. http://dx.doi.org/10.1002/cpbi.27.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. Nat Genet. 2000;25:25–29. http://dx.doi.org/10.1038/75556.

Auburn S, Barry AE. Dissecting malaria biology and epidemiology using population genetics and genomics. Int J Parasitol. 2017;47:77–85. http://dx.doi.org/10.1016/j.ijpara.2016.08.006.

Bachman MA, LeBar WD. First Principles of Clinical Microbiology: Collection, Handling, and Diagnostics. Ref Module Biomed Sci [Internet]. Elsevier; 2018 [cited 2024 Nov 5]. p. B9780128012383661160. Available from: https://linkinghub.elsevier.com/retrieve/pii/B9780128012383661160.

Baranova A, Cao H, Chen J, Zhang F. Causal Association and Shared Genetics Between Asthma and COVID-19. Front Immunol. 2022;13:705379. http://dx.doi.org/10.3389/fimmu.2022.705379.

Baranova A, Cao H, Zhang F. Unraveling Risk Genes of COVID-19 by Multi-Omics Integrative Analyses. Front Med. 2021;8:738687. http://dx.doi.org/10.3389/fmed.2021.738687.

Baxter VK, Heise MT. Immunopathogenesis of alphaviruses. Adv Virus Res [Internet]. Elsevier; 2020 [cited 2024 Dec 16]. p. 315–382. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0065352720300208.

Beigier-Bompadre M, Montagna GN, Kühl AA, Lozza L, Weiner J, Kupz A, et al. Mycobacterium tuberculosis infection modulates adipose tissue biology. Behr MA, editor. PLOS Pathog. 2017;13:e1006676. http://dx.doi.org/10.1371/journal.ppat.1006676.

Boletín anual de Tuberculosis [Internet]. Ministerio de Salud Pública del Ecuador; 2018. Available from: www. https://www.salud.gob.ec/gacetas-tuberculosis/.

Bos KI, Kühnert D, Herbig A, Esquivel-Gomez LR, Andrades Valtueña A, Barquera R, et al. Paleomicrobiology: Diagnosis and Evolution of Ancient Pathogens. Annu Rev Microbiol. 2019;73:639–666. http://dx.doi.org/10.1146/annurev-micro-090817-062436.

Bulló M, García-Lorda P, Megias I, Salas-Salvadó J. Systemic Inflammation, Adipose Tissue Tumor Necrosis Factor, and Leptin Expression. Obes Res. 2003;11:525–531. http://dx.doi.org/10.1038/oby.2003.74.

Butler-Laporte G, Kreuzer D, Nakanishi T, Harroud A, Forgetta V, Richards JB. Genetic Determinants of Antibody-Mediated Immune Responses to Infectious Diseases Agents: A Genome-Wide and HLA Association Study. Open Forum Infect Dis. 2020;7:ofaa450. http://dx.doi.org/10.1093/ofid/ofaa450.

Carroll KC, Pfaller MA, editors. Manual of clinical microbiology. Washington, DC: ASM Press; 2023.

Casadevall A, Pirofski L. Host-Pathogen Interactions: Basic Concepts of Microbial Commensalism, Colonization, Infection, and Disease. Portnoy DA, editor. Infect Immun. 2000;68:6511–6518. http://dx.doi.org/10.1128/IAI.68.12.6511-6518.2000.

Chaplin DD. Overview of the immune response. J Allergy Clin Immunol. 2010;125:S3–S23. http://dx.doi.org/10.1016/j.jaci.2009.12.980.

Chapman SJ, Hill AVS. Human genetic susceptibility to infectious disease. Nat Rev Genet. 2012;13:175–188. http://dx.doi.org/10.1038/nrg3114.

Chen L, Hua J, He X. Bioinformatics analysis identifies a key gene HLA_DPA1 in severe influenza-associated immune infiltration. BMC Genomics. 2024;25:257. http://dx.doi.org/10.1186/s12864-024-10184-7.

Chen L-Y, Hwang A-C, Huang C-Y, Chen L-K, Wang F-D, Chan Y-J. CMV infection, CD19+ B cell depletion, and Lymphopenia as predictors for unexpected admission in the institutionalized elderly. Immun Ageing. 2021;18:21. http://dx.doi.org/10.1186/s12979-021-00233-0.

Chen Z, Chen R, Ou Y, Lu J, Jiang Q, Liu G, et al. Construction of an HLA Classifier for Early Diagnosis, Prognosis, and Recognition of Immunosuppression in Sepsis by Multiple Transcriptome Datasets. Front Physiol. 2022;13:870657. http://dx.doi.org/10.3389/fphys.2022.870657.

Choi H, Lee JY, Yoo H, Jeon K. Bioinformatics Analysis of Gene Expression Profiles for Diagnosing Sepsis and Risk Prediction in Patients with Sepsis. Int J Mol Sci. 2023;24:9362. http://dx.doi.org/10.3390/ijms24119362.

Cholo MC, Van Rensburg EJ, Osman AG, Anderson R. Expression of the Genes Encoding the Trk and Kdp Potassium Transport Systems of Mycobacterium tuberculosis during Growth In Vitro. BioMed Res Int. 2015;2015:1–11. http://dx.doi.org/10.1155/2015/608682.

Costafreda MI, Kaplan G. HAVCR1 (CD365) and Its Mouse Ortholog Are Functional Hepatitis A Virus (HAV) Cellular Receptors That Mediate HAV Infection. Pfeiffer JK, editor. J Virol. 2018;92:e02065-17. http://dx.doi.org/10.1128/JVI.02065-17.

Couper KN, Blount DG, Riley EM. IL-10: The Master Regulator of Immunity to Infection. J Immunol. 2008;180:5771–5777. http://dx.doi.org/10.4049/jimmunol.180.9.5771.

Cui Y, Keles S, Charbonnier L-M, Julé AM, Henderson L, Celik SC, et al. Combined immunodeficiency caused by a loss-of-function mutation in DNA polymerase delta 1. J Allergy Clin Immunol. 2020;145:391-401.e8. http://dx.doi.org/10.1016/j.jaci.2019.10.004.

Curtis J, Luo Y, Zenner HL, Cuchet-Lourenço D, Wu C, Lo K, et al. Susceptibility to tuberculosis is associated with variants in the ASAP1 gene encoding a regulator of dendritic cell migration. Nat Genet. 2015;47:523–527. http://dx.doi.org/10.1038/ng.3248.

Da Silva FAR, Pascoal LB, Dotti I, Setsuko Ayrizono MDL, Aguilar D, Rodrigues BL, et al. Whole transcriptional analysis identifies markers of B, T and plasma cell signaling pathways in the mesenteric adipose tissue associated with Crohn's disease. J Transl Med. 2020;18:44. http://dx.doi.org/10.1186/s12967-020-02220-3.

Davila S, Hibberd ML, Hari Dass R, Wong HEE, Sahiratmadja E, Bonnard C, et al. Genetic Association and Expression Studies Indicate a Role of Toll-Like Receptor 8 in Pulmonary Tuberculosis. Gojobori T, editor. PLoS Genet. 2008;4:e1000218. http://dx.doi.org/10.1371/journal.pgen.1000218.

De Los Rios M, Criscitiello MF, Smider VV. Structural and genetic diversity in antibody repertoires from diverse species. Curr Opin Struct Biol. 2015;33:27–41. http://dx.doi.org/10.1016/j.sbi.2015.06.002.

Debnath M, Banerjee M, Berk M. Genetic gateways to COVID-19 infection: Implications for risk, severity, and outcomes. FASEB J. 2020;34:8787–8795. http://dx.doi.org/10.1096/fj.202001115R.

Deng X, Sabino EC, Cunha-Neto E, Ribeiro AL, Ianni B, Mady C, et al. Genome Wide Association Study (GWAS) of Chagas Cardiomyopathy in Trypanosoma cruzi Seropositive Subjects. Tanowitz HB, editor. PLoS ONE. 2013;8:e79629. http://dx.doi.org/10.1371/journal.pone.0079629.

Desai JV, Kumar D, Freiwald T, Chauss D, Johnson MD, Abers MS, et al. C5a-licensed phagocytes drive sterilizing immunity during systemic fungal infection. Cell. 2023;186:2802-2822.e22. http://dx.doi.org/10.1016/j.cell.2023.04.031.

Doncheva NT, Morris JH, Gorodkin J, Jensen LJ. Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data. J Proteome Res. 2019;18:623–632. http://dx.doi.org/10.1021/acs.jproteome.8b00702.

Druszczynska M, Seweryn M, Wawrocki S, Kowalewska-Pietrzak M, Pankowska A, Rudnicka W. Cytokine Biosignature of Active and Latent Mycobacterium Tuberculosis Infection in Children. Pathogens. 2021;10:517. http://dx.doi.org/10.3390/pathogens10050517.

Duan T, Du Y, Xing C, Wang HY, Wang R-F. Toll-Like Receptor Signaling and Its Role in Cell-Mediated Immunity. Front Immunol. 2022;13:812774. http://dx.doi.org/10.3389/fimmu.2022.812774.

Dumonteil E, Herrera C, Martini L, Grijalva MJ, Guevara AG, Costales JA, et al. Chagas Disease Has Not Been Controlled in Ecuador. Tanowitz HB, editor. PLOS ONE. 2016;11:e0158145. http://dx.doi.org/10.1371/journal.pone.0158145.

Fan H, Lu B, Cao C, Li H, Yang D, Huang L, et al. Plasma TNFSF13B and TNFSF14 Function as Inflammatory Indicators of Severe Adenovirus Pneumonia in Pediatric Patients. Front Immunol. 2021;11:614781. http://dx.doi.org/10.3389/fimmu.2020.614781.

Frade AF, Pissetti CW, Ianni BM, Saba B, Lin-Wang HT, Nogueira LG, et al. Genetic susceptibility to Chagas disease cardiomyopathy: involvement of several genes of the innate immunity and chemokine-dependent migration pathways. BMC Infect Dis. 2013;13:587. http://dx.doi.org/10.1186/1471-2334-13-587.

Ghiglione Y, Falivene J, Ruiz MJ, Laufer N, Socías ME, Cahn P, et al. Early Skewed Distribution of Total and HIV-Specific CD8+ T-Cell Memory Phenotypes during Primary HIV Infection Is Related to Reduced Antiviral Activity and Faster Disease Progression. Boasso A, editor. PLoS ONE. 2014;9:e104235. http://dx.doi.org/10.1371/journal.pone.0104235.

Giraldo MI, Xia H, Aguilera-Aguirre L, Hage A, Van Tol S, Shan C, et al. Envelope protein ubiquitination drives entry and pathogenesis of Zika virus. Nature. 2020;585:414–419. http://dx.doi.org/10.1038/s41586-020-2457-8.

Gómez I, Thomas MC, Palacios G, Egui A, Carrilero B, Simón M, et al. Differential Expression of Immune Response Genes in Asymptomatic Chronic Chagas Disease Patients Versus Healthy Subjects. Front Cell Infect Microbiol. 2021;11:722984. http://dx.doi.org/10.3389/fcimb.2021.722984.

Goretzki A, Zimmermann J, Lin Y-J, Schülke S. Immune Metabolism–An Opportunity to Better Understand Allergic Pathology and Improve Treatment of Allergic Diseases? Front Allergy. 2022;3:825931. http://dx.doi.org/10.3389/falgy.2022.825931.

Grant M, Larijani M. Evasion of adaptive immunity by HIV through the action of host APOBEC3G/F enzymes. AIDS Res Ther. 2017;14:44. http://dx.doi.org/10.1186/s12981-017-0173-8.

Guo X, Roberts MR, Becker SM, Podd B, Zhang Y, Chua SC, et al. Leptin signaling in intestinal epithelium mediates resistance to enteric infection by Entamoeba histolytica. Mucosal Immunol. 2011;4:294–303. http://dx.doi.org/10.1038/mi.2010.76.

Gupta K, Kaur G, Pathak T, Banerjee I. Systematic review and meta-analysis of human genetic variants contributing to COVID-19 susceptibility and severity. Gene. 2022;844:146790. http://dx.doi.org/10.1016/j.gene.2022.146790.

Haque R, Duggal P, Ali IM, Hossain MB, Mondal D, Sack RB, et al. Innate and Acquired Resistance to Amebiasis in Bangladeshi Children. J Infect Dis. 2002;186:547–552. http://dx.doi.org/10.1086/341566.

Hartman ZC, Kiang A, Everett RS, Serra D, Yang XY, Clay TM, Amalfitano A. Adenovirus Infection Triggers a Rapid, MyD88-Regulated Transcriptome Response Critical to Acute-Phase and Adaptive Immune Responses In Vivo. J Virol. 2007;81:1796–1812. http://dx.doi.org/10.1128/JVI.01936-06.

Hawn TR, Misch EA, Dunstan SJ, Thwaites GE, Lan NTN, Quy HT, et al. A common human TLR1 polymorphism regulates the innate immune response to lipopeptides. Eur J Immunol. 2007;37:2280–2289. http://dx.doi.org/10.1002/eji.200737034.

Heuberger J, Trimpert J, Vladimirova D, Goosmann C, Lin M, Schmuck R, et al. Epithelial response to IFN-γ promotes SARS-CoV-2 infection. EMBO Mol Med. 2021;13:e13191. http://dx.doi.org/10.15252/emmm.202013191.

Hill AVS, Allsopp CEM, Kwiatkowski D, Anstey NM, Twumasi P, Rowe PA, et al. Common West African HLA antigens are associated with protection from severe malaria. Nature. 1991;352:595–600. http://dx.doi.org/10.1038/352595a0.

Horwood PF, Tarantola A, Goarant C, Matsui M, Klement E, Umezaki M, et al. Health Challenges of the Pacific Region: Insights From History, Geography, Social Determinants, Genetics, and

the Microbiome. Front Immunol. 2019;10:2184. http://dx.doi.org/10.3389/fimmu.2019.02184.

Ishibashi M, Morita N, Nomura-Kawaguchi C, Shimizu Y, Wakita T, Esumi M. CLEC4M-positive and CD81-negative Huh7 cells are not susceptible to JFH-1 HCVcc infection but mediate transinfection. Arch Virol. 2014;159:2949–2955. http://dx.doi.org/10.1007/s00705-014-2150-z.

Ismailova A, White JH. Vitamin D, infections and immunity. Rev Endocr Metab Disord. 2022;23:265–277. http://dx.doi.org/10.1007/s11154-021-09679-5.

Janeway CA, Medzhitov R. Innate Immune Recognition. Annu Rev Immunol. 2002;20:197–216. http://dx.doi.org/10.1146/annurev.immunol.20.083001.084359.

Josyula JVN, Talari P, Pillai AKB, Mutheneni SR. Analysis of gene expression profile for identification of novel gene signatures during dengue infection. Infect Med. 2023;2:19–30. http://dx.doi.org/10.1016/j.imj.2023.02.002.

Junttila IS. Tuning the Cytokine Responses: An Update on Interleukin (IL)-4 and IL-13 Receptor Complexes. Front Immunol. 2018;9:888. http://dx.doi.org/10.3389/fimmu.2018.00888.

Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 2000;28:27–30. http://dx.doi.org/10.1093/nar/28.1.27.

Kany S, Vollrath JT, Relja B. Cytokines in Inflammatory Disease. Int J Mol Sci. 2019;20:6008. http://dx.doi.org/10.3390/ijms20236008.

Kawai T, Akira S. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. Nat Immunol. 2010;11:373–384. http://dx.doi.org/10.1038/ni.1863.

Kawai T, Akira S. Toll-like Receptors and Their Crosstalk with Other Innate Receptors in Infection and Immunity. Immunity. 2011;34:637–650. http://dx.doi.org/10.1016/j.immuni.2011.05.006.

Khasnis S, Veenstra H, McClellan MJ, Ojeniyi O, Wood CD, West MJ. Regulation of B cell receptor signalling by Epstein–Barr virus nuclear antigens. Biochem J. 2022;479:2395–2417. http://dx.doi.org/10.1042/BCJ20220417.

Khor CC, Hibberd ML. Host–pathogen interactions revealed by human genome-wide surveys. Trends Genet. 2012;28:233–243. http://dx.doi.org/10.1016/j.tig.2012.02.001.

Kitada S, Kayama H, Okuzaki D, Koga R, Kobayashi M, Arima Y, et al. BATF2 inhibits immunopathological Th17 responses by suppressing Il23a expression during Trypanosoma cruzi infection. J Exp Med. 2017;214:1313–1331. http://dx.doi.org/10.1084/jem.20161076.

Koh GCKW, Porras P, Aranda B, Hermjakob H, Orchard SE. Analyzing Protein–Protein Interaction Networks. J Proteome Res. 2012;11:2014–2031. http://dx.doi.org/10.1021/pr201211w.

Kolberg L, Raudvere U, Kuzmin I, Adler P, Vilo J, Peterson H. g:Profiler—interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). Nucleic Acids Res. 2023;51:W207–W212. http://dx.doi.org/10.1093/nar/gkad347.

Kulkarni DH, Starick M, Aponte Alburquerque R, Kulkarni HS. Local complement activation and modulation in mucosal immunity. Mucosal Immunol. 2024;17:739–751. http://dx.doi.org/10.1016/j.mucimm.2024.05.006.

Kumar H, Kawai T, Akira S. Toll-like receptors and innate immunity. Biochem Biophys Res Commun. 2009;388:621–625. http://dx.doi.org/10.1016/j.bbrc.2009.08.062.

Kwissa M, Nakaya HI, Onlamoon N, Wrammert J, Villinger F, Perng GC, et al. Dengue Virus Infection Induces Expansion of a CD14+CD16+ Monocyte Population that Stimulates Plasmablast Differentiation. Cell Host Microbe. 2014;16:115–127. http://dx.doi.org/10.1016/j.chom.2014.06.001.

Lang B, Cao C, Zhao X, Wang Y, Cao Y, Zhou X, et al. Correction: Genomic alterations related to HPV infection status in a cohort of Chinese prostate cancer patients. Eur J Med Res. 2024;29:348. http://dx.doi.org/10.1186/s40001-024-01935-z.

Li X, Yang Y, Zhou F, Zhang Y, Lu H, Jin Q, et al. SLC11A1 (NRAMP1) Polymorphisms and Tuberculosis Susceptibility: Updated Systematic Review and Meta-Analysis. Herrmann JL, editor. PLoS ONE. 2011;6:e15831. http://dx.doi.org/10.1371/journal.pone.0015831.

Lio CT, Dhanda SK, Bose T. Cluster Analysis of Dry Eye Disease Models Based on Immune Cell Parameters – New Insight Into Therapeutic Perspective. Front Immunol. 2020;11:1930. http://dx.doi.org/10.3389/fimmu.2020.01930.

Loterio RK, Zamboni DS, Newton HJ. Keeping the host alive – lessons from obligate intracellular bacterial pathogens. Pathog Dis. 2021;79:ftab052. http://dx.doi.org/10.1093/femspd/ftab052.

Luck K, Kim D-K, Lambourne L, Spirohn K, Begg BE, Bian W, et al. A reference map of the human binary protein interactome. Nature. 2020;580:402–408. http://dx.doi.org/10.1038/s41586-020-2188-x.

Ma H, Liu J, Li Z, Xiong H, Zhang Y, Song Y, et al. Expression profile analysis reveals hub genes that are associated with immune system dysregulation in primary myelofibrosis. Hematology. 2021;26:478–490. http://dx.doi.org/10.1080/16078454.2021.1945237.

Ma Q. Role of Nrf2 in Oxidative Stress and Toxicity. Annu Rev Pharmacol Toxicol. 2013;53:401–426. http://dx.doi.org/10.1146/annurev-pharmtox-011112-140320.

Malaria Genomic Epidemiology Network, Band G, Le QS, Clarke GM, Kivinen K, Hubbart C, et al. Insights into malaria susceptibility using genome-wide data on 17,000 individuals from Africa, Asia and Oceania. Nat Commun. 2019;10:5732. http://dx.doi.org/10.1038/s41467-019-13480-z.

Malaria Genomic Epidemiology Network, Rockett KA, Clarke GM, Fitzpatrick K, Hubbart C, Jeffreys AE, et al. Reappraisal of known malaria resistance loci in a large multicenter study. Nat Genet. 2014;46:1197–1204. http://dx.doi.org/10.1038/ng.3107.

Malhotra S, Yen JY, Honko AN, Garamszegi S, Caballero IS, Johnson JC, et al. Transcriptional Profiling of the Circulating Immune Response to Lassa Virus in an Aerosol Model of Exposure. Geisbert T, editor. PLoS Negl Trop Dis. 2013;7:e2171. http://dx.doi.org/10.1371/journal.pntd.0002171.

Martins YC, Ribeiro-Gomes FL, Daniel-Ribeiro CT. A short history of innate immunity. Mem Inst Oswaldo Cruz. 2023;118:e230023. http://dx.doi.org/10.1590/0074-02760230023.

Masin PS, Visentin HA, Elpidio LNS, Sell AM, Visentainer L, Lima Neto QAD, et al. Genetic polymorphisms of toll-like receptors in leprosy patients from southern Brazil. Front Genet. 2022;13:952219. http://dx.doi.org/10.3389/fgene.2022.952219.

Mathew NR, Jayanthan JK, Smirnov IV, Robinson JL, Axelsson H, Nakka SS, et al. Single-cell BCR and transcriptome analysis after influenza infection reveals spatiotemporal dynamics of antigen-specific B cells. Cell Rep. 2021;35:109286. http://dx.doi.org/10.1016/j.celrep.2021.109286.

McLaren PJ, Coulonges C, Bartha I, Lenz TL, Deutsch AJ, Bashirova A, et al. Polymorphisms of large effect explain the majority of the host genetic contribution to variation of HIV-1 virus load. Proc Natl Acad Sci. 2015;112:14658–14663. http://dx.doi.org/10.1073/pnas.1514867112.

McNicholl JM, Downer MV, Udhayakumar V, Alper CA, Swerdlow DL. Host-Pathogen Interactions in Emerging and Re-emerging Infectious Diseases: A Genomic Perspective of Tuberculosis, Malaria, Human Immunodeficiency Virus Infection, Hepatitis B, and Cholera. Annu Rev Public Health. 2000;21:15–46. http://dx.doi.org/10.1146/annurev.publhealth.21.1.15.

Merker M, Tueffers L, Vallier M, Groth EE, Sonnenkalb L, Unterweger D, et al. Evolutionary Approaches to Combat Antibiotic Resistance: Opportunities and Challenges for Precision Medicine. Front Immunol. 2020;11:1938. http://dx.doi.org/10.3389/fimmu.2020.01938.

Moroney JB, Chupp DP, Xu Z, Zan H, Casali P. Epigenetics of the antibody and autoantibody response. Curr Opin Immunol. 2020;67:75–86. http://dx.doi.org/10.1016/j.coi.2020.09.004.

Mozzi A, Pontremoli C, Sironi M. Genetic susceptibility to infectious diseases: Current status and future perspectives from genome-wide approaches. Infect Genet Evol. 2018;66:286–307. http://dx.doi.org/10.1016/j.meegid.2017.09.028.

Muñoz Cabas D, Cando Caluña W, Cando Herrera F, Cando Herrera J. Epidemiología de la Malaria en Ecuador de 2018–2022. RECIAMUC. 2023;7:151–162. http://dx.doi.org/https://doi.org/10.26820/reciamuc/7.(1).enero.2023.151-162.

Na K, Oh B-C, Jung Y. Multifaceted role of CD14 in innate immunity and tissue homeostasis. Cytokine Growth Factor Rev. 2023;74:100–107. http://dx.doi.org/10.1016/j.cytogfr.2023.08.008.

Nahon P, Cobat A. Human genetics of HCV infection phenotypes in the era of direct-acting antivirals. Hum Genet. 2020;139:855–863. http://dx.doi.org/10.1007/s00439-020-02136-4.

Nichols-Vinueza DX, Delmonte OM, Bundy V, Bosticardo M, Zimmermann MT, Dsouza NR, et al. POLD1 Deficiency Reveals a Role for POLD1 in DNA Repair and T and B Cell Development. J Clin Immunol. 2021;41:270–273. http://dx.doi.org/10.1007/s10875-020-00903-6.

Nicolas E, Golemis EA, Arora S. POLD1: Central mediator of DNA replication and repair, and implication in cancer and other pathologies. Gene. 2016;590:128–141. http://dx.doi.org/10.1016/j.gene.2016.06.031.

Nielsen SCA, Boyd SD. Human adaptive immune receptor repertoire analysis—Past, present, and future. Immunol Rev. 2018;284:9–23. http://dx.doi.org/10.1111/imr.12667.

O'Connor KS, Ahlenstiel G, Suppiah V, Schibeci S, Ong A, Leung R, et al. IFNL3 mediates interaction between innate immune cells: Implications for hepatitis C virus pathogenesis. Innate Immun. 2014;20:598–605. http://dx.doi.org/10.1177/1753425913503385.

O'Neill LAJ, Golenbock D, Bowie AG. The history of Toll-like receptors — redefining innate immunity. Nat Rev Immunol. 2013;13:453–460. http://dx.doi.org/10.1038/nri3446.

Oughtred R, Rust J, Chang C, Breitkreutz B, Stark C, Willems A, et al. The BIOGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. Protein Sci. 2021;30:187–200. http://dx.doi.org/10.1002/pro.3978.

Paz y Miño Cepeda C. Rompecabezas evolutivo. Baltimore, Maryland: Project Muse; 2024.

Paz-y-Mino Cepeda C, Morillo SA, Celi AP, Witte T, Munoz Guijarro MJ, Collantes JC, et al. CCR5Δ, CCR2-64I, and SDF1-3'A Polymorphisms Related to Resistance to HIV-1 Infection and Disease in the Ecuadoran Population. Hum Biol. 2005;77:521–526. http://dx.doi.org/10.1353/hub.2005.0068.

Paz-y-Miño C, Zambrano AK, Leone PE. Interactoma de predisposición y resistencia a SARS-CoV-2. Proteínas, genes y funciones. Bionatura. 2021;6:1555–1562. http://dx.doi.org/10.21931/RB/2021.06.01.17.

Pecoraro V, Cuccorese M, Trenti T. Genetic polymorphisms of ACE1, ACE2, IFTM3, TMPRSS2 and TNFα genes associated with susceptibility and severity of SARS-CoV-2 infection: a systematic review and meta-analysis. Clin Exp Med. 2023;23:3251–3264. http://dx.doi.org/10.1007/s10238-023-01038-9.

Peng B, Kong G, Yang C, Ming Y. Erythropoietin and its derivatives: from tissue protection to immune regulation. Cell Death Dis. 2020;11:79. http://dx.doi.org/10.1038/s41419-020-2276-8.

Pleiman CM, D'Ambrosio D, Cambier JC. The B-cell antigen receptor complex: structure and signal transduction. Immunol Today. 1994;15:393–399. http://dx.doi.org/10.1016/0167-5699(94)90267-4.

Qi S, Sun C, Wang J, Wo L, Li Y, Wang C, et al. Identification of NECTIN1 as a novel restriction factor for flavivirus infection. Meng X-J, Weger-Lucarelli J, editors. mBio. 2024;15:e02708-24. http://dx.doi.org/10.1128/mbio.02708-24.

Quintana-Murci L. Human Immunology through the Lens of Evolutionary Genetics. Cell. 2019;177:184–199. http://dx.doi.org/10.1016/j.cell.2019.02.033.

Rajbhandari L, Shukla P, Jagdish B, Mandalla A, Li Q, Ali MA, et al. Nectin-1 Is an Entry Mediator for Varicella-Zoster Virus Infection of Human Neurons. Goodrum F, editor. J Virol. 2021;95:e01227-21. http://dx.doi.org/10.1128/JVI.01227-21.

Ramezani A, Nahad MP, Faghihloo E. The role of Nrf2 transcription factor in viral infection. J Cell Biochem. 2018;119:6366–6382. http://dx.doi.org/10.1002/jcb.26897.

Ramos-Casals M, Brahmer JR, Callahan MK, Flores-Chávez A, Keegan N, Khamashta MA, et al. Immune-related adverse events of checkpoint inhibitors. Nat Rev Dis Primer. 2020;6:38. http://dx.doi.org/10.1038/s41572-020-0160-6.

Rankin JA. Biological Mediators of Acute Inflammation: AACN Clin Issues Adv Pract Acute Crit Care. 2004;15:3–17. http://dx.doi.org/10.1097/00044067-200401000-00002.

Rasch RFR. Ancient History and New Frontiers: Infectious Diseases. Nurs Clin North Am. 2019;54:xv–xvi. http://dx.doi.org/10.1016/j.cnur.2019.03.001.

Rast J, Litman G. Evolution of innate and adaptive immune recognition structures. Semin Immunol. 2010;22:1–2. http://dx.doi.org/10.1016/j.smim.2010.01.001.

Redondo-García S, Barritt C, Papagregoriou C, Yeboah M, Frendeus B, Cragg MS, et al. Human leukocyte immunoglobulin-like receptors in health and disease. Front Immunol. 2023;14:1282874. http://dx.doi.org/10.3389/fimmu.2023.1282874.

Rickert RC. New insights into pre-BCR and BCR signalling with relevance to B cell malignancies. Nat Rev Immunol. 2013;13:578–591. http://dx.doi.org/10.1038/nri3487.

Ricklin D, Reis ES, Mastellos DC, Gros P, Lambris JD. Complement component C3 – The "Swiss Army Knife" of innate immunity and host defense. Immunol Rev. 2016;274:33–58. http://dx.doi.org/10.1111/imr.12500.

Robledo D, Matika O, Hamilton A, Houston RD. Genome-Wide Association and Genomic Selection for Resistance to Amoebic Gill Disease in Atlantic Salmon. G3 GenesGenomesGenetics. 2018;8:1195–1203. http://dx.doi.org/10.1534/g3.118.200075.

Rolland T, Taşan M, Charloteaux B, Pevzner SJ, Zhong Q, Sahni N, et al. A Proteome-Scale Map of the Human Interactome Network. Cell. 2014;159:1212–1226. http://dx.doi.org/10.1016/j.cell.2014.10.050.

Rothfels K, Milacic M, Matthews L, Haw R, Sevilla C, Gillespie M, et al. Using the Reactome Database. Curr Protoc. 2023;3:e722. http://dx.doi.org/10.1002/cpz1.722.

Rotival M. Characterising the genetic basis of immune response variation to identify causal mechanisms underlying disease susceptibility. HLA. 2019;94:275–284. http://dx.doi.org/10.1111/tan.13598.

Saengsiwaritt W, Jittikoon J, Chaikledkaew U, Udomsinprasert W. Genetic polymorphisms of ACE1 , ACE2 , and TMPRSS2 associated with COVID-19 severity: A systematic review with meta-analysis. Rev Med Virol. 2022;32:e2323. http://dx.doi.org/10.1002/rmv.2323.

Santiago VF, Dombrowski JG, Kawahara R, Rosa-Fernandes L, Mule SN, Murillo O, et al. Complement System Activation Is a Plasma Biomarker Signature during Malaria in Pregnancy. Genes. 2023;14:1624. http://dx.doi.org/10.3390/genes14081624.

Setliff I, Shiakolas AR, Pilewski KA, Murji AA, Mapengo RE, Janowska K, et al. High-Throughput Mapping of B Cell Receptor Sequences to Antigen Specificity. Cell. 2019;179:1636-1646.e15. http://dx.doi.org/10.1016/j.cell.2019.11.003.

Sharygin D, Koniaris LG, Wells C, Zimmers TA, Hamidi T. Role of CD14 in human disease. Immunology. 2023;169:260–270. http://dx.doi.org/10.1111/imm.13634.

Shi Y, Su J, Chen R, Wei W, Yuan Z, Chen X, et al. The Role of Innate Immunity in Natural Elite Controllers of HIV-1 Infection. Front Immunol. 2022;13:780922. http://dx.doi.org/10.3389/fimmu.2022.780922.

Silva AMVD, Alvarado-Arnez LE, Azamor T, Batista-Silva LR, Leal-Calvo T, Bezerra OCDL, et al. Interferon-lambda 3 and 4

Polymorphisms Increase Sustained Virological Responses and Regulate Innate Immunity in Antiviral Therapy With Pegylated Interferon-Alpha. Front Cell Infect Microbiol. 2021;11:656393. http://dx.doi.org/10.3389/fcimb.2021.656393.

Smatti MK, Alkhatib HA, Al Thani AA, Yassine HM. Will Host Genetics Affect the Response to SARS-CoV-2 Vaccines? Historical Precedents. Front Med. 2022;9:802312. http://dx.doi.org/10.3389/fmed.2022.802312.

Stefani C, Sangalli A, Locatelli E, Federico T, Malerba G, Romanelli MG, et al. Increased Prevalence of Unstable HLA-C Variants in HIV-1 Rapid-Progressor Patients. Int J Mol Sci. 2022;23:14852. http://dx.doi.org/10.3390/ijms232314852.

Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, et al. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. Curr Protoc Bioinforma [Internet]. 2016 [cited 2024 Nov 5];54. http://dx.doi.org/10.1002/cpbi.5.

Stertz S, Hale BG. Interferon system deficiencies exacerbating severe pandemic virus infections. Trends Microbiol. 2021;29:973–982. http://dx.doi.org/10.1016/j.tim.2021.03.001.

Su N, Hu C, Shao T, Jin C, Li H, Ji J, et al. Functional role of CD40 and CD154 costimulatory signals in IgZ-mediated immunity against bacterial infection. Fish Shellfish Immunol Rep. 2021;2:100038. http://dx.doi.org/10.1016/j.fsirep.2021.100038.

Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, et al. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 2023;51:D638–D646. http://dx.doi.org/10.1093/nar/gkac1000.

Tanaka S, Baba Y. B Cell Receptor Signaling. In: Wang J-Y, editor. B Cells Immun Toler [Internet]. Singapore: Springer Singapore; 2020 [cited 2024 Nov 5]. p. 23–36. Available from: http://link.springer.com/10.1007/978-981-15-3532-1_2.

Tarris G, De Rougemont A, Charkaoui M, Michiels C, Martin L, Belliot G. Enteric Viruses and Inflammatory Bowel Disease. Viruses. 2021;13:104. http://dx.doi.org/10.3390/v13010104.

The UniProt Consortium, Bateman A, Martin M-J, Orchard S, Magrane M, Ahmad S, et al. UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Res. 2023;51:D523–D531. http://dx.doi.org/10.1093/nar/gkac1052.

Vasconcelos RHT, Montenegro SML, Azevedo EAN, Gomes YM, Morais CNL. Genetic susceptibility to chronic Chagas disease: An overview of single nucleotide polymorphisms of cytokine genes. Cytokine. 2012;59:203–208. http://dx.doi.org/10.1016/j.cyto.2012.04.035.

Vasseur E, Quintana-Murci L. The impact of natural selection on health and disease: uses of the population genetics approach in humans. Evol Appl. 2013;6:596–607. http://dx.doi.org/10.1111/eva.12045.

Velastegui E, Vera E, Vanden Berghe W, Muñoz MS, Orellana-Manzano A. "HLA-C: evolution, epigenetics, and pathological implications in the major histocompatibility complex." Front Genet. 2023;14:1206034. http://dx.doi.org/10.3389/fgene.2023.1206034.

Velavan TP, Pallerla SR, Rüter J, Augustin Y, Kremsner PG, Krishna S, et al. Host genetic factors determining COVID-19 susceptibility and severity. eBioMedicine. 2021;72:103629. http://dx.doi.org/10.1016/j.ebiom.2021.103629.

Vijay K. Toll-like receptors in immunity and inflammatory diseases: Past, present, and future. Int Immunopharmacol. 2018;59:391–412. http://dx.doi.org/10.1016/j.intimp.2018.03.002.

Villarino AV, Kanno Y, O'Shea JJ. Mechanisms and consequences of Jak–STAT signaling in the immune system. Nat Immunol. 2017;18:374–384. http://dx.doi.org/10.1038/ni.3691.

Walker TM, Kohl TA, Omar SV, Hedge J, Del Ojo Elias C, Bradley P, et al. Whole-genome sequencing for prediction of Mycobacterium tuberculosis drug susceptibility and resistance: a retrospective cohort study. Lancet Infect Dis. 2015;15:1193–1202. http://dx.doi.org/10.1016/S1473-3099(15)00062-6.

Wang H, Liu M. Complement C4, Infections, and Autoimmune Diseases. Front Immunol. 2021;12:694928. http://dx.doi.org/10.3389/fimmu.2021.694928.

Wang J, Chai Q, Lei Z, Wang Y, He J, Ge P, et al. LILRB1-HLA-G axis defines a checkpoint driving natural killer cell exhaustion in tuberculosis. EMBO Mol Med. 2024;16:1755–1790. http://dx.doi.org/10.1038/s44321-024-00106-1.

Wang J, Ganaie SS, Cheng F, Xu P, Ning K, Wang X, et al. RNA Binding Motif Protein RBM45 Regulates Expression of the 11-Kilodalton Protein of Parvovirus B19 through Binding to Novel Intron Splicing Enhancers. Meng X-J, editor. mBio. 2020;11:e00192-20. http://dx.doi.org/10.1128/mBio.00192-20.

Wang Y, Qian G, Zhu L, Zhao Z, Liu Y, Han W, et al. HIV-1 Vif suppresses antiviral immunity by targeting STING. Cell Mol Immunol. 2022;19:108–121. http://dx.doi.org/10.1038/s41423-021-00802-9.

Wang Y, Sun Q, Zhang Y, Li X, Liang Q, Guo R, et al. Systemic immune dysregulation in severe tuberculosis patients revealed by a single-cell transcriptome atlas. J Infect. 2023;86:421–438. http://dx.doi.org/10.1016/j.jinf.2023.03.020.

Wang Y-T, Ji S-X, Xia X-Y, Wan T-J, Zou J-X, Li B-X, et al. Proteomic investigation and biomarker identification of lung and spleen deficiency syndrome in HIV/AIDS immunological nonresponders. J Thorac Dis. 2023;15:1460–1472. http://dx.doi.org/10.21037/jtd-23-322.

Wen C, Dong Z, Wang Y, Ye G, Ma Y, Yi X, et al. CTLA4+CD4+CXCR5−FOXP3+ T cells associate with unfavorable outcome in patients with chronic HBV infection. BMC Immunol. 2023;24:3. http://dx.doi.org/10.1186/s12865-022-00537-w.

Werner M, Schefczyk S, Trippler M, Treckmann JW, Baba HA, Gerken G, et al. Antiviral Toll-like Receptor Signaling in Non-Parenchymal Liver Cells Is Restricted to TLR3. Viruses. 2022;14:218. http://dx.doi.org/10.3390/v14020218.

Wu J, Tan Y, Deng H, Zhu D. Relationship between degree–rank function and degree distribution of protein–protein interaction networks. Comput Biol Chem. 2008;32:1–4. http://dx.doi.org/10.1016/j.compbiolchem.2007.08.008.

Wu Y-H, Chang W-T, Hsu C-L, Lu Y-F, Wang J-T, Tzeng S-J. FcγRIIB modulates splenic germinal center response against immune subversion during acute influenza A virus infection. J Microbiol Immunol Infect. 2024;57:64–75. http://dx.doi.org/10.1016/j.jmii.2023.11.007.

Wu Z, Mehrabi Nasab E, Arora P, Athari SS. Study effect of probiotics and prebiotics on treatment of OVA-LPS-induced of allergic asthma inflammation and pneumonia by regulating the TLR4/NF-kB signaling pathway. J Transl Med. 2022;20:130. http://dx.doi.org/10.1186/s12967-022-03337-3.

Wudhikarn K, Perales M-A. Infectious complications, immune reconstitution, and infection prophylaxis after CD19 chimeric antigen receptor T-cell therapy. Bone Marrow Transplant. 2022;57:1477–1488. http://dx.doi.org/10.1038/s41409-022-01756-w.

Xu Z, Li D, Qu W, Yin Y, Qiao S, Zhu Y, et al. Card9 protects sepsis by regulating Ripk2-mediated activation of NLRP3 inflammasome in macrophages. Cell Death Dis. 2022;13:502. http://dx.doi.org/10.1038/s41419-022-04938-y.

Yao P-A, Sun H-J, Li X-Y. Identification of key genes in late-onset major depressive disorder through a co-expression network module. Front Genet. 2022;13:1048761. http://dx.doi.org/10.3389/fgene.2022.1048761.

Zhu M, Ma X, Huang J, Lu F, Chen Y, Hu J, et al. Extracellular vesicle-derived miR-1249–5p regulates influenza A virus-induced acute lung injury in RAW246.7 cells through targeting SLC4A1. Microbes Infect. 2022;24:104998. http://dx.doi.org/10.1016/j.micinf.2022.104998.

rupture

Vivienne Baillie Gerritsen

Who has not been accosted by someone they would prefer to avoid? You greet them meekly, exchange a few polite words while wondering how to cut the exchange short without being disrespectful. If being disrespectful is not an issue, however, things become gloriously easy. You could tell them that their company bores you, that you have more important things to tend to or interrupt the chat with a brisk shake of the hand and move on. If you were part of a cartoon strip, you could burst into flames or simply disintegrate. It so happens that cells can actually opt to disintegrate when dealing with something that has become toxic to their environment, or at any rate redundant. This can be a virus, a bacterium, or perhaps simply age. Such an option is generally called programmed cell death, or apoptosis. And as there are many ways of being disrespectful to your acquaintance, there are many ways a cell can choose to disappear. One is by generating fatal rips in its own plasma membrane. Though rips such as these have been observed by scientists for many years, plasma membrane rupture was long believed to be a passive event. Until a protein known as ninjurin-1 was discovered.

Manawatu Gorge

by Timon Maxey, New Zealand
instagram: @timonmaxey

courtesy of the artist

Though a subtle difference between programmed cell death (PCD) and apoptosis is under debate, today, both words are usually used to describe the same thing by the great majority of biologists. That is to say, the intrinsic 'choice' a cell makes to end its own life. This is opposed to 'necrosis', the term used when a cell dies due to external factors, such as a cardiac arrest or poisoning. The word 'apoptosis' has Greek roots and literally relates to the periodic shedding of leaves in autumn. Hippocrates used it to describe gangrene, perhaps because he saw it as a kind of shedding of the skin. Almost two thousand years later, during the 18th and the 19th centuries, scientists began to take an interest in cell death which, in those days, was referred to as 'natural cell death'.

Until the early 20th century, the causes of natural cell death were attributed to mechanical changes, such as the swelling and subsequent bursting of a cell. However, the Scottish microbiologist Alexander Fleming (1881-1955) suggested that cell death could also be caused by chemical changes, such as the disintegration of the cell's chromatin which he had observed under the microscope. The term 'programmed cell death' only appeared in the early 1960s. From a purely philosophical point of view, the notion that a cell could choose to die – albeit under given circumstances – seemed almost unethical. However, the 1970s and the advent of molecular biology changed such views which, all of a sudden, became attractive: a cell could plan its own demise.

A cell can choose to put an end to its life in several ways. Frequently, more than one strategy is used: two strategies can occur in unison for instance, or one can trigger off a second. Sometimes, too, a cell can decide to delay its death to arrange a few things beforehand – this has been described in a previous article*. Which

strategies does a cell choose? It can form pores in its plasma membrane, for example, fragment its DNA, or slow down its mitochondria. Each of these strategies is fatal, by letting vital constituents leave the cell, inhibiting novel protein synthesis or failing to provide ATP, respectively. If performed simultaneously, little hope is left for the cell. The upside? No harm is caused to neighbouring cells – as opposed to necrosis.

Rips, or tears, on a cell's plasma membrane had been observed for centuries and scientists thought they were passive, purely due to mechanical forces – following the swelling of a cell for instance, much like stretch marks that appear on skin when weight has been put on abruptly. However, it turns out that rips, or plasma membrane ruptures, are quite intentional. This became obvious when researchers discovered the protein ninjurin-1. Ninjurin-1 is widely expressed, especially in tissues of epithelial origin, where it seems to be involved in cell adhesion. Structurally, ninjurin-1 looks like a hairpin. The curved bend is formed by a pair of alpha-helices, as in a classic transmembrane protein embedded in the cellular membrane. The N- and C-terminal extremities jut out to bathe in the extracellular medium, and the N terminus carries 2 alpha-helices (alpha1 and alpha2) separated by a kink.

Scattered across a cell's plasma membrane, monomeric ninjurin-1 is inactive. But during PCD, researchers noticed that ninjurin-1 monomers polymerise to form filaments of varied length. Structurally, each filament is built in the style of a classical wooden fence. The transmembrane region remains embedded in the cellular membrane while the N-terminus plunges into it. Alpha2 positions itself parallel to the sides of the hairpin while, thanks to the kink, alpha1 juts out at 90 degrees, reaching out to bind to a neighbouring monomer. The height of each filament, or 'ninjurin-fence', thus spans the width of the cellular membrane. In this way, filaments form linear tears in the membrane, or their ends meet to form pore-like structures. Either way, vital cytosolic constituents are free to spill out into the extracellular medium.

PCD was observed in multicellular organisms to begin with. In fact, for a long time, it was thought that the notion of PCD could only apply to multicellular organisms, or at least to cells where some kind of cooperation, or sort of cellular society, exists. But towards the end of the 20th century, it became apparent that PCD also applied to unicellular organisms. In fact, researchers now believe that PCD actually originated in unicellular organisms – although, admittedly, the existing theories do point, one way or another, towards multicellular cooperation. As an example, in a colony of bacteria, an infected cell could decide to give up its life to put a halt to viral replication and, hence, the further spread of infection.

How is ninjurin-1 activated? Is it the target of another factor triggered off by PCD? Does it react to a molecular change in the plasma membrane? Or to a conformational change such as membrane swelling? Ninjurin-1 may respond to both kinds of cue – chemical and mechanical. So far, no one knows. Some helices, such as those found in ninjurin-1, are known to sense lipid-packing defects or inflections of the membrane, both of which could be the result of cell swelling. Despite its seemingly tragic outcome, PCD is an elegant biophysical mechanism used by cells to create fatal tears in their own membrane without causing damage to their neighbours. In fact, frequently, it is precisely to preserve them from harm. A mechanism such as this could open opportunities for healing cancers, infection and inflammatory diseases by forcing cells that have become harmful to their environment to commit suicide.

**** Protein Spotlight Issue 252***

Cross-references to UniProt

Ninjurin-1, *Homo sapiens* (Human): Q92982
Ninjurin-1, *Mus musculus* (Mouse): O70131

References

1. Durand P.M., Ramsey G.
The concepts and origins of cell mortality
History and Philosophy of the Life Sciences 46 (2023)
PMID: 37289372

2. Degen M., Santos J.C., Pluhackova K., et al.
Structural basis of NINJ1-mediated plasma membrane rupture in cell death
Nature 618 (2023)
PMID: 37198476

proteinspotlight

Protein Spotlight (ISSN 1424-4721) is a montly review written by the **Swiss-Prot** team of the **SIB Swiss Institute of Bioinformatics**. Spotlight articles describe a specific protein or family of proteins on an informal tone. http://web.expasy.org/spotlight/

seizure

Vivienne Baillie Gerritsen

Many years ago, I was sitting opposite a man whose body suddenly froze. His eyes seemed to be staring at something on the wall behind me while his left hand drew small circles in the air, repetitively. I had no idea what was happening to him until he came back to his senses and told me that he had just had an epileptic fit. Deeply embarrassed, he got up and left the room. Many of us will have witnessed a close relative, a friend, an acquaintance or perhaps simply a passerby, under the influence of an epileptic seizure – which are frequently more violent and alarming than the one I experienced that day. Epilepsy affects millions of people worldwide. However, what is happening on the molecular scale remains elusive. What we do know is that an epileptic fit is caused by neural activity that has suddenly gone out of control. In this light, researchers discovered two tarantula venom peptides – Aa1a and Ap1a – which inhibit channels that are used to relay signals in our central nervous system. Peptides such as these could perhaps be used to keep abnormal neural activity at bay in people suffering from epilepsy.

Komposition 1939

Otto Freundlich (1878-1943)

Until at least the 17th century, epileptic fits were thought to have a divine origin, or be caused by evil spirits. Even though, two thousand years earlier, the Greek physician Hippocrates had rejected the idea that epilepsy had anything to do with spirits but was a problem that stemmed directly from the brain and could be medically treated. Nonetheless, for a very long time after his death, the only kind of healing offered was spiritual, and many suffering from epilepsy were shunned by society, if not interned. Still today, there are societies where epilepsy is believed to be associated with evil spirits, witchcraft or poisoning – and even sometimes considered contagious.

Defined as a disease of the brain, epilepsy seems to have a strong genetic predisposition although it can also occur in patients who have suffered brain trauma for example. When witnessing an epileptic fit, the first thing that comes to mind is that something has gone very wrong with the control of our body. In fact, we have no control over it anymore, as though the system for transmitting signals to and from our brain has suddenly gone haywire – which is exactly the case. Epilepsy is the result of a gross imbalance between neural activation and deactivation. Something our body is unable to cope with. So, it temporarily loses hold.

Neural activation and deactivation depend on the seamless coordination of the opening and closing of channels scattered throughout the central nervous system. The sum of concerted channel activation and deactivation allows us to remain in relative control of our mobility – but also of many other less tangible functions such as consciousness for instance. When channel concertation fails, there is an abnormal surge of neural excitation resulting in an epileptic seizure. A voltage-gated potassium

channel known as hEAG is important for human cognitive development. Scientists know this because mutated forms of hEAG give rise to syndromes known as Temple-Baraitser and Zimmermann-Laband, both of which result in mental retardation. Patients also happen to suffer from epilepsy – and this gave researchers the opportunity to understand the matter better.

With this in mind, animal venom is constituted of toxins, several of which specifically target channels involved in neural activation and deactivation. What better way to neutralize predators or prey than to meddle with their central nervous system by disrupting essential metabolic pathways that cause temporary paralysis – as the aggressor makes a run for it – or perhaps even death. For decades now, scientists have been singling out venom peptides that could have some kind of therapeutic potential. In the case of epilepsy, which is caused by the irrational and uncontrolled opening and closing of channels, a well-chosen toxin could perhaps counter this by guaranteeing some kind of regulation. So scientists scanned the venom of tarantulas, namely: *Avicularia aurantiaca* and *Avicularia purpurea.* The choice was far from innocent since the venom of both tarantulas is known to inhibit hEAG. Two peptides were extracted – kappa-theraphotoxin-Aa1a (Aa1a) from *A.aurantiaca*, and mu/kappa-theraphotoxin-Ap1a (Ap1a) from *A.purpurea* – both of which turned out to be potent hEAG inhibitors.

Aa1a and Ap1a are 81% identical, consisting of 36 residues with an amidated C-terminus and three disulfide bridges. Their singularity lies in the structure formed by the three bridges: a cystine knot. Imagine a loop formed by two of the bridges, and the third slides through it. This forms a cystine knot – in our case, an inhibitor cystine knot. Cystine knots are relatively common in venom toxins because they confer chemical stability and resistance to enzymatic degradation, which means they can survive for a long time inside the victim. It is perhaps one of the rare times in life when the formation of a knot – whatever its nature – is not deemed a nuisance. Besides the cystine knot, there is another intriguing formation: a sort of ladder whose rungs are formed by the stacking of hydrophobic patches on one side of each inhibitor peptide. These molecular rungs may be necessary to form interactions with the target channel as well as the lipid membrane of brain cells. Certainly, molecular ladders such as these are frequently found in spider toxins whose role is to modify voltage-gated channels.

It is likely that Aa1a and Ap1a act by binding to the extracellular regions of hEAG where they cause a depolarising shift in the cell's membrane, reducing the probability that the channel opens by as much as 50%. In short, Aa1a and Ap1a do not deactivate the channels by blocking the central pore, for example, but inhibit them by exerting pressure, so to speak, on the pore domain. Of the two inhibitor peptides, Ap1a is the most potent inhibitor of hEAG, which makes it a potential candidate for the development of anti-epileptic drugs. Indeed, epilepsy affects as many as 70 million people worldwide – of all ages. Although seizures are always short-lived, there are invisible side-effects such as neurobiological and cognitive consequences but also psychological and sociological repercussions. Naturally, there are already many anti-epileptic drugs on the market. However, they are ineffective in as much as one third of persons suffering from recurrent bouts of epilepsy. The more scientists understand about the channels that are responsible for this widespread affliction, the better their knowledge will be to design drugs that will help everyone.

Cross-references to UniProt

Potassium voltage-gated channel subfamily H member 1, *Homo sapiens* (Human): O95259
Kappa-theraphotoxin-Aa1a, *Avicularia aurantiaca* (Yellow-banded pinktoe tarantula): A0A3F2YLP5
Mu/kappa-theraphotoxin-Ap1a, *Avicularia purpurea* (Ecuadorian purple pinktoe tarantula): P0DQJ6

References

1. Ma L., Chin Y.K.Y., Dekan Z. et al.
Novel venom-derived inhibitors of the human EAG channel, a putative antiepileptic drug target
Biochemical Pharmacology 158:60-72(2018)
PMID: 30149017

2. Cázares-Ordoñez V., Pardo L.A.
Kv10.1 potassium channel: from the brain to the tumors
Biochemistry and Cell Biology 95:531-536(2017)
PMID: 28708947

proteinspotlight

Protein Spotlight (ISSN 1424-4721) is a montly review written by the **Swiss-Prot** team of the **SIB Swiss Institute of Bioinformatics**. Spotlight articles describe a specific protein or family of proteins on an informal tone. http://web.expasy.org/spotlight/

SIB Swiss Institute of Bioinformatics

a chromosome's glue

Vivienne Baillie Gerritsen

We all begin with one cell, which divides into two – and so on. It sounds straightforward but a cell has various components (nucleus, mitochondria, Golgi apparatus…) each of which carries out vital activities. If two daughter cells are to survive, they must receive a copy of each component from the mother cell. A mother cell cannot just split in two, pour half of its contents into one cell and tilt the rest in the second. That would be like producing two cars of the same make where one is built with no engine and the other with no wheels. Every part of a cell has a specific and an essential role, which is why each part has to be inherited by progeny. Among these essential components daughter cells must receive a copy of their mother's DNA. The only way to do this is for the mother cell to double its DNA and then distribute it in such a way that the DNA in each daughter cell is identical in quantity and nature. This can occur thanks to a mechanism known as mitosis. During mitosis, a dividing cell's chromosomes (its DNA) alternate between two opposing states: individualized and clustered. It turns out that a protein – already known to scientists – is directly involved in the making of these two chromosomal states. Its name? Ki-67.

"Chromosomes Chromatids" by Jenny Gray

Courtesy of the artist

Mitosis is the process by which cells divide while distributing their contents – in particular their DNA – in a balanced manner to the two daughter cells. Mitosis is billions of years old since the first form of eukaryotic life used it. As a consequence, the process has had plenty of time for refinement, and the intricacies and beauty of its various stages is awe-inspiring. Names have been given to each stage: prophase, metaphase, anaphase and telophase – with an in-between stage called interphase when not much happens. Without going into any detail whatsoever, when a cell is not dividing (interphase), its chromosomes are kept (and protected) within an organelle known as the nucleus. When a cell is dividing (prophase to telophase), its nuclear envelope disassembles thus freeing the chromosomes (that have just been doubled), which are then dispatched to the daughter cells who rapidly reform their own nucleus to protect their own batch of chromosomes.

During the various stages of mitosis, chromosomes adopt two major conformations: loose and tight. Before a cell divides, you cannot make out individual chromosomes in the nucleus because they are lank and gathered into an indistinctive clump. When a cell is about to divide, however, the lank chromosomes double their content and tighten up so as to adopt a more rigid shape while breaking away from one another. In this individualized, more rigid conformation, the mother cell can distribute them far more easily, and correctly, to the daughter cells. This is carried out by a structure known as the mitotic spindle – a sort of wonderfully evolved multichord mechanism, which can be compared to the chords of a violin gathered at each end. Each (doubled) chromosome is attached to one chord. As the mother

cell halves the chromosomes halve too, and each half is gently pulled into a nascent daughter cell.

How do chromosomes switch between their lank formless state and a more rigid condensed one? This puzzled scientists for years until they came up with a very elegant model that involves a protein known as Ki-67. Ki-67 is a large protein, with a high net electrical charge. It seems to have no particular 3D structure and spends most of its time unfolded like spaghetti. However, unlike spaghetti, Ki-67 has an amphiphilic molecular structure: its C-terminus is highly attracted to chromatin (what chromosomes are made of) while its N-terminus prefers the cytoplasm. The body of Ki-67 is made up of repeats that carry over 100 potential phosphorylation sites. Their phosphorylation causes Ki-67 to unfurl into a sort of tail, while dephosphorylation causes the tail to collapse. All in all, the structure of Ki-67 is very similar to that of surfactants, agents that are found at the boundaries of different phases, such as solid and liquid. Does Ki-67 actually behave like a surfactant? At the boundary of chromatin and cytoplasm?

This is the model that has been proposed. Ki-67 has no role in the internal structure of chromosomes, that is to say in their condensation as a cell is about to divide for instance. Ki-67's role is simply to keep condensed chromosomes apart during mitosis. It does this by forming a sort of repellent on the chromosomal surface. How? The C-terminus of Ki-67 binds to the chromosome while its extended N-terminus juts out into the cytoplasm. Consider the fact that an estimated 270,000 Ki-67 molecules bind to the surface of a mitotic (condensed) chromosome with an average spacing of about 69nm. Visually, this would look like a very hairy chromosome, whose whole surface is covered with a sort of brush-like arrangement of Ki-67. Since the body of each Ki-67 has a high – and identical – electrical charge, like magnets showing the same poles, two neighbouring chromosomes are repelled. In other words, they'll be kept separated from one another. This is how surfactants behave too.

Now, once the individual chromosomes have been dispersed in the daughter cells, they readopt their clumped state in a newly formed nucleus. What happens to the brush-like repellent on their surface? And to Ki-67? It would be natural to assume that since Ki-67 is not needed anymore, it is probably degraded. Scientists suggest something else. In its 'repellent' state, Ki-67 is highly phosphorylated which lends the protein a highly negative charge. At the end of mitosis, the body of Ki-67 is stripped naked as it is dephosphorylated thus giving it a highly positive charge. This might attract negatively-charged stretches of nucleotides known as ribosomal RNA (rRNA) that are found in a cell's nucleus. rRNA acts like a kind of glue as it binds to the 'collapsed tails' of Ki-67, concomitantly bridging neighbouring chromosomes. This is how the temporarily individualized chromosomes stick to one another and cluster in the newly formed nucleus.

So phosphorylation and dephosphorylation of Ki-67 would be the key to chromosome individualization and clustering, respectively. However, the role of phosphorylation in this process still has to be demonstrated. Likewise, there is currently no evidence of rRNA actually binding to Ki-67. For years, Ki-67 has been used both as a marker for cell proliferation and to assess the growth of tumour cells in cancer diagnostics. This was until researchers realised that the protein's role was probably less in cell proliferation *per se* than in the formation of two opposing structural states of chromosomes as a cell divides. Until evidence proves otherwise, what has been described above remains a model – but it certainly is a very elegant one, and demonstrates how powerful computational models can be, and how they are a researcher's precious ally.

Cross-references to UniProt

Proliferation marker protein Ki-67, *Homo sapiens* (Human) : P46013

References

1. Hernandez-Armendariz A., Sorichetti V., Hayashi Y., *et al.*
A liquid-like coat mediates chromosome clustering during mitotic exit
Molecular Cell 84: 3254-3270(2024)
PMID: 39153474

2. Cuylen S., Blaukopf C., Politi A.Z., *et al.*
Ki-67 acts as a biological surfactant to disperse mitotic chromosomes
Nature 535: 308-312(2016)
PMID: 27362226

proteinspotlight

Protein Spotlight (ISSN 1424-4721) is a montly review written by the **Swiss-Prot** team of the **SIB Swiss Institute of Bioinformatics**. Spotlight articles describe a specific protein or family of proteins on an informal tone. http://web.expasy.org/spotlight/

SIB Swiss Institute of Bioinformatics

ramifications

Vivienne Baillie Gerritsen

There is much talk about plastic these days. And with reason. Besides depending on fossil fuels, plastic is infesting every nook and cranny of our planet because there is simply too much of it. Since the invention of bakelite in 1907, human dependency on the astonishingly varied properties of plastics has, understandably, never ceased to grow. Up popped the polyester polyethylene terephthalate, or PET, in the 1940s and an inventor's idea to use it to bottle soft drinks in the 1970s – which marked the beginning of a catastrophe. Today, we are desperate to find ways of recycling plastics and degrading them in eco-friendly ways. We have already discovered bacteria that are happy to eat PET for dinner – although not fast or efficiently enough for our liking. Lately, scientists came across a similar process that occurs in Nature when certain fungi, such as *Aspergillus oryzae*, invade plants. Plant cells in contact with the air are protected not only by a film of wax but also by cutin polyesters which are similar in structure to PET. Fungi have to degrade the cutin polyesters in order to reach the plants' flesh. They do this with the help of an intriguing little protein known as hydrophobin which, it turns out, can also be used to stimulate the degradation of PET.

"10 Green Bottles" by Alison Deegan

Courtesy of the artist

PET – a synthetic polymer consisting of ($C_{10}H_8O_4$) units – was discovered in the 1940s in the hunt for new man-made fibres following the revolutionary discovery of nylon. Natural polymers such as resin, rubber and cellulose have been used throughout history by humans but, as the industrial revolution kicked in, scientists began to look for materials whose properties were less limited. Bakelite, the first fully synthetic plastic, had already emerged in the 1920s and was used for all sorts of electrical and household appliances, automobile parts and even jewellery.

However, when PET made its appearance with its stash of superior physical and chemical properties, bakelite could not compete. PET became widely used in fibres for clothing, in containers for foods and liquids and in all kinds of manufacturing thanks to its thermoforming characteristics. In the 1980s, production technology had reached such a level of sophistication that PET bottles fast became the primary material for beverage packaging. The polymer is light, durable and can be blow-moulded into shape. Almost half a century later, though, we do not know how to deal with all the PET we are producing.

Hydrophobins may turn out to be part of the solution. Hydrophobins are small proteins, all of which carry eight conserved cysteine residues and four disulfide bonds. They are found exclusively in filamentous fungi, notably in *Aspergillus oryzae* – which sports 50 different kinds. Secreted onto the surface of fungal hyphae, hydrophobins self-assemble into rodlike formations – or rodlets – to form a hydrophobic protective sheath which has its

role in fungal growth, development and dispersal. Hydrophobins also seem to contribute to host infection, and it is this particular trait that turned out to be, in a meandering kind of way, comparable to what is involved in PET degradation.

The outer tissues of plants – and insects for that matter – are covered with a protective hydrophobic sheath, the cuticle which, among other things, prevents dehydration while acting as a barrier to bacterial or fungal infection. Cuticles are composed of various waxes and polyesters known as cutins. One characteristic of hyphal hydrophobins is their capacity to adhere to hydrophobic surfaces such as those formed by cutins. Hydrophobin adhesion is thought to stimulate the fungal expression of cutinases, enzymes that are able to degrade cutin. This would then create vulnerable areas and points of entry for fungal infection.

Surprisingly, how scientists currently understand the mode of action of hydrophobins on cuticles was discovered while studying the interaction of *A.oryzae* with a biodegradable plastic known as polybutylene succinate-coadipate, or PBSA. The team found that one hydrophobin, called rolA, adheres specifically to the hydrophobic PBSA surface. This particular adhesion causes a conformational change in rolA which goes on to stimulate the expression of a fungal cutinase, cutL1. RolA and cutL1 have an affinity for one another. Even more so, it seems, if rolA has adhered to a hydrophobic surface. Once the two are paired, the cutinase then proceeds to degrade the plastic whose carbon atoms are used as a source of energy for fungal growth. In a way, *A.oryzae* doesn't really mind where it gets its food from – synthetic plastic or plants – as long as it gets it. And this form of fungal indifference can be put to use by scientists. It seems, too, that as long as rolA is not bound to cutL1, the hydrophobin can move laterally on the plastic surface. Lateral mobility ceases, however, the moment the two are paired, which probably is just a way of giving cutL1 a better chance to degrade the plastic.

This was observed with PBSA. Could rolA and cutL1 behave the same way with the synthetic plastic PET? The answer is yes. A team of scientists presented rolA to some PET while adding a pinch of an enzyme they knew could degrade PET: an enzyme of bacterial origin, coined PETase. Like with PBSA, hydrophobin rolA adhered to the hydrophobic plastic PET surface. As expected, adhesion caused a conformational change in rolA which increased the degradative action of PETase.

This is very good news. The degradation of PET via PETase is not only environment friendly but occurs at relatively low temperatures, i.e. 30°C. PET plastic recycling by any current method requires high temperatures and therefore consumes a lot of energy – besides generating other, perhaps even toxic, substances and just a different kind of pollution. The rolA/PETase system still remains slow, however. Its mechanism needs to be better understood so that it can be fine-tuned to speed up the reaction. Especially as more than half of the world's beverages continue to be stored in PET bottles not to mention many fibres that continue to be made of polyester – and there is little sign that production will decrease.

Cross-references to UniProt

Class I hydrophobin A, *Aspergillus oryzae* (strain ATCC 42149 / RIB 40) (Yellow koji mold) : Q2U3W7
Cutinase I, *Aspergillus oryzae* (strain ATCC 42149 / RIB 40) (Yellow koji mold) : P52956

References

1. Puspitasari N., Tsai S.-L., Lee C.-K.
Fungal hydrophobin rolA enhanced PETase hydrolysis of polyethylene terephthalate
Applied Biochemistry and Biotechnology 193: 1284-1295 (2021)
PMID: 32506284

2. Takahashi T., Maeda H., Yoneda S. et al.
The fungal hydrophobin RolA recruits polyesterase and laterally moves on hydrophobic surfaces
Molecular Microbiology 57:1780-1796(2025)
PMID: 16135240

proteinspotlight
Swiss Institute of Bioinformatics
Protein Spotlight (ISSN 1424-4721) is a montly review written by the **Swiss-Prot** team of the **SIB Swiss Institute of Bioinformatics**. Spotlight articles describe a specific protein or family of proteins on an informal tone. http://web.expasy.org/spotlight/

EMBnet.journal
Bioinformatics in Action

ISSN 1023-4144

DEAR READER,

If you have any comments or suggestions regarding this journal we would be very glad to hear from you. If you have a tip you feel we can publish then please let us know. Before submitting your contribution read the "Authors guidelines"[1] and send your manuscript and supplementary files using our on-line submission system[2].

Past issues are available as PDF files from the web archive[3].

Visit EMBnet website for more information: www.journal.embnet.org

EMBNET.JOURNAL EXECUTIVE EDITORIAL BOARD

[1]http://journal.embnet.org/index.php/embnetjournal/about/submissions#authorGuidelines
[2]http://journal.embnet.org/index.php/embnetjournal/author/submit
[3]http://journal.embnet.org/index.php/embnetjournal/issue/archive

PUBLISHER

EMBnet Stichting p/a
CMBI Radboud University
Nijmegen Medical Centre
6581 GB Nijmegen
Th e Netherlands

Email: erik.bongcam@slu.se
Tel: +46-18-67 21 21