



ABRomics Analysis Platform:

A service for analyzing microbial antibiotic resistance
within a “One Health” framework

User Charter

10/01/2025

Version 1.0

ABRomics Web Portal: www.abromics.fr

ABRomics Platform Web Application: <https://analysis.abromics.fr/>

Table of contents

Key Points of the User Charter	3
Objectives of the ABRomics Platform	4
Who are the users of the ABRomics platform?	4
User Registration and Management	5
Creating a User Account	5
Closing a User Account	6
Creating Projects, Uploading Data, and Collaborating Among ABRomics Users	6
Sample-associated data managed by ABRomics	6
Data provided by the user	6
Data from the ABRomics Platform analysis pipelines	7
Data Accessibility and Confidentiality	8
Accessibility Status of the Sequence Data and Associated Metadata	8
Making Certain Metadata and Analysis Results Public	8
How ABRomics Works (Services Offered by the platform)	9
Organization and Administration of the ABRomics Platform	9
Data security	9
Data Backup	10
Privacy of Personal Data (GDPR)	10
Responsibilities of the ABRomics Platform	10
Responsibilities of the Sequence Submitter	11
Scientific Acknowledgments	11

Key Points of the User Charter

ABRomics is a multi-omics digital platform dedicated to the **storage, sharing, and analysis** of bacterial (meta)genomic data for research and surveillance of antibiotic resistance. It is intended for researchers and clinicians in the fields of human, animal, and environmental health. Version 1.0 of the charter covers genomic data.

To **create a user account**, you must have a French institutional email address. If you do not have a valid institutional email address, please contact the ABRomics support team at abromics-support@groupe.france-bioinformatique.fr.

Submission: When the ABRomics platform opened in early January 2025 (version 1.0), sequences must be submitted in paired-read FASTQ format. In a future version of the platform (v1.x), submissions in FASTA format will be accepted under certain conditions. These FASTQ files are accompanied by metadata that corresponds, in particular, to the data required for submission to international repositories (ENA, EMBL, and NCBI). No patient data is required. After the submitted data has been verified for consistency, the sequences are automatically analyzed using a suite of tools organized into a single pipeline. The results of these analyses are stored in the database and accessible via the application's visualization interface and in various formats.

Concept of Project: Data is submitted as part of a project that can be shared among different users with an ABRomics account. The project creator (or owner) assigns different levels of permissions to project members.

Data confidentiality: After submission and the initial default analysis, the data owner (project supervisor) can choose to make the data public within ABRomics (disabled in version 1.0) or to keep it confidential for a period of one year, renewable twice (3 years of confidentiality), and in exceptional cases up to four times (5 years of confidentiality). When data is made public within ABRomics, the owner will have the option to request that it be uploaded to the ENA (European Nucleotide Archive, EMBL) and to the BigsDB or Enterobase typing databases (not functional in ABRomics v1).

Data Sharing: To promote collaboration among teams and identify epidemiological links, certain data on strains and the results of computational analyses are accessible to all users of the ABRomics Platform (see the "Data Accessibility and Confidentiality" section). Sequence data for non-public strains are not accessible. Searchable strain data remains the property of the owner and may not be used in a publication without the owner's written permission.

Objectives of the ABRomics Platform

ABRomics is a cross-sectoral (One Health) digital platform, available online, hosted in France, secure, and dedicated to antimicrobial resistance (AMR). It facilitates access to (meta)genomic data on bacterial infectious diseases and their metadata. The ABRomics consortium is partnered with a broad community of scientists, epidemiologists, and clinical microbiologists.

The platform has three main objectives:

- To establish a repository of structured, interoperable, standardized, and well-annotated multi-omic microbiological data from human, animal, and environmental sources.
- To provide a set of standard analysis pipelines for the typing and prediction of resistance genes in bacterial strains, as well as mathematical and bioinformatics tools, which will enable the addressing of research questions related to antibiotic resistance that could not be tackled without such a platform.
- To establish a shared platform to facilitate epidemiological surveillance of antibiotic resistance in human and veterinary medicine, as well as in environmental and food isolates: ABRomics must enable the surveillance of the transmission and outbreaks of multidrug-resistant bacteria in near real time, delivering results that can also be utilized by public health authorities. FAIR (Findable, Accessible, Interoperable, and Reusable) data management procedures will enable retrospective studies.

Who are the users of the ABRomics platform?

- Microbiologists who sequence “natural” strains (*i.e.*, not derived from evolutionary or synthetic biology experiments) and submit data related to antimicrobial resistance (regardless of whether it is human, veterinary, or environmental) to the ABRomics database, for example:
 - Clinical microbiologist
 - Researcher
 - Clinical microbiologist and researcher
- Epidemiologist / public health professional seeking access to data for surveillance purposes.
- Researchers seeking access to data without submitting their own for research purposes. In this case, open-access publications using open data from the ABRomics platform must explicitly cite the project and its funding (see the “Scientific Impact” section).

Users from private companies (laboratories and industrial firms) who wish to use the platform for the purpose of analyzing and/or utilizing the platform’s database are invited to contact the platform’s support team (abromics-support@groupes.france-bioinformatique.fr).

The submission of sequences and/or strains as part of research projects is not subject to a project selection process. For projects requiring significant storage space and, consequently, substantial computing time (*e.g.*, projects analyzing several thousand genomes or metagenomics projects), a feasibility study will be conducted with the ABRomics Platform

administrators, and the project's acceptance will then be approved at a Steering Committee meeting.

Regardless of the objectives for implementing the ABRomics platform—that is:

- **create and/or use specific project spaces** in ABRomics, upload sequences, and run analyses
- **explore/download public data and certain “open” data** (see below) from the ABRomics-DB database, without uploading sequences

users must first have a secure personal account and have accepted this charter.

The ABRomics web portal (www.abromics.fr) provides a summary of the data managed in ABRomics (in text or graphical form) for site visitors who do not have a secure account. In addition, the ABRomics Platform user manual is also available on the web portal:

<https://www.abromics.fr/abromics-platform/user-tutorial/>

This user charter is intended for individuals who have a secure account on the ABRomics platform.

User Registration and Management

The ABRomics application's web interface (in English to prepare for future international use) is available at this address:

<https://analysis.abromics.fr/>

Creating a User Account

Registering with the ABRomics Platform allows users to enter the following information:

- Last Name, First Name
- Email: The entered value must be on a whitelist of domain names that covers all French research organizations and institutes. Requests to add a new domain can be sent to the platform administrators.
- ORCID (optional)
- Institution (optional)
- Password to be created (minimum 12 characters, at least 1 uppercase letter, at least 1 number, and at least 1 special character)

After submitting the form, the user receives an email inviting them to activate their account. The ABRomics Platform username is the user's email address. Accounts are personal and must not under any circumstances be transferred to a third party. The resources made available to a user are intended to be used solely for the purpose of carrying out their professional activities.

Initially, the ABRomics Platform is open to foreign teams only as part of a consortium under the umbrella of a user from a French research organization. Access to the ABRomics Platform for private partners requires an explicit request, which will be reviewed by the steering committee members.

Users can edit and/or delete their profile information at any time, except for their email address. If their email address changes, they must contact the ABRomics administrators to have the change made.

Closing a User Account

- When initiated by a user, deleting an account automatically deletes data from private projects for which the user is the supervisor (FASTQ files, results, analyses). Projects that have already been made public are retained.
- When initiated by ABRomics administrators, if a user account has remained *inactive for four calendar years*.

Creating Projects, Uploading Data, and Collaborating Among ABRomics Users

La création d'un compte utilisateur et la signature de la charte des utilisateur.trice.s sont **obligatoires pour créer des projets, déposer des séquences, faire des analyses** dans la PF ABRomics et/ou **explorer les données** de la base ABRomics.

Creating a user account and signing the user charter are required to create projects, upload sequences, perform analyses in the ABRomics Platform, and/or **explore data** in the ABRomics database.

When a project is created within the ABRomics application's web interface, **the user automatically becomes the supervisor** of that project space. The user can then:

- upload (meta)genomic sequence data (and/or MAGs), whether private or public,
- upload metadata files using the ABRomics metadata repository (genomic and metagenomic),
- run ABRomics analysis workflows,
- invite collaborators, who must first have requested a secure personal account on ABRomics. The supervisor defines access rights associated with the project and the sequences uploaded to the project space for their collaborators.

A collaborator:

- must have a secure ABRomics user account
- can run ABRomics analysis workflows
- cannot invite other collaborators to this project space (they must go through the supervisor)

Sample-associated data managed by ABRomics

Data provided by the user

The ABRomics Platform database and file storage space allow for the storage of data associated with various types of samples:

- data provided by "depositors," *i.e.*, genomic sequence files (raw sequences) and metadata associated with biological samples and sequences.
- reference genome sequence data for ABRomics (already published) along with their metadata (when available in public databases).

ABRomics has a **metadata repository** that allows users to enter biological data for the projects they create (fields marked with an asterisk (*) are required):

- **Sample metadata:**
 - Sample type (human, animal, environment)*
 - Host species (animal species)*
 - Sample source (for human : stool, urine, blood, lung etc.)*

- Location (Country Region Place)*
- Collected date (at least the year)*
- Sample ID (Sample name for the user)
- Travel countries (Countries visited in the three months prior to sampling)
- Sample comment (Free-form comment on the sample)
- Microorganism scientific name (Bacterial species)*
- Strain ID or name (may be identical to sample ID)*
- Accession number (Accession numbers associated with the sample if the sequence has already been submitted to public sequence database)
- **Technical data:**
 - Instrument model (Illumina HiSeq 2000, MinION, etc...)*
- **Sequence data:**
 - Raw read data: This data is uploaded to the ABRomics Platform file repository in the form of short reads (Illumina) or long reads (Nanopore or PacBio)

A detailed metadata table is available on the ABRomics web portal:

<https://www.abromics.fr/abromics-platform/metadata-referential/>

With the exception of the (meta)genomic sequences of the analyzed strains, these data (biological, temporal, location-related, etc.) are integrated into the ABRomics database (in several relational tables). For each sample integrated into ABRomics (metadata and sequences), an identifier is automatically generated and remains internal to the database. Data related to biological samples and technical data can be used in database queries: for example, “number of strains per department,” “number of strains exhibiting resistance X and originating from sample Y,” etc.

Data from the ABRomics Platform analysis pipelines

After the quality of the FASTQ sequences has been verified, they are first assembled into longer DNA sequences in a FASTA-format file. Uploading sequences in FASTA format (i.e., assembled outside the ABRomics Platform) will be available in version v1.x of the ABRomics Platform. These sequences will be analyzed only if they pass the quality check (contamination then assembly).

The sequences are analyzed using pipelines developed as part of the ABRomics project to obtain information such as:

- Species identification
- Identification of antibiotic resistance genes
- Identification of mutations responsible for antibiotic resistance
- Identification of virulence genes
- Molecular typing using multi-locus sequence typing (MLST) and core-genome MLST (cgMLST) to identify relationships between strains.

The results of the analysis pipelines are stored in relational tables in the ABRomics database.

Data Accessibility and Confidentiality

The ABRomics platform is based on the infrastructure of the French Institute of Bioinformatics (IFB) and the procedures established as part of data lifecycle management and Open Science. These procedures are based on the principle of **“as open as possible, as closed as necessary.”**

All data and associated metadata from public resources integrated into ABRomics are open and therefore freely accessible.

Accessibility Status of the Sequence Data and Associated Metadata

The supervisor defines one of the following:

- **Public access:** either upon submission to ABRomics or at any time before the end of the embargo period (see below). The data (*i.e.*, metadata and analysis results) are open and accessible to all users registered on the ABRomics platform (secure professional account with a password)
- **Private access:** sensitive data (raw sequence files and certain metadata, see below) will be accessible only to collaborators designated by the supervisor of a project space. The private status of the data is specified by the supervisor after the sequence data and metadata have been submitted.
 - The data remains confidential for one year, renewable up to two times. An email is sent to the submitter one month, then one week, before the status change deadline.
 - **This embargo period may not exceed three years.**
 - However, **extending this embargo period to five years** may be discussed at a Steering Committee meeting.

At the end of the embargo period, a sequence that was previously restricted must become freely accessible. ABRomics will offer the submitter the option of automatic deposit into public databases (EBI's ENA). The ENA identifiers are stored in the ABRomics database.

Making Certain Metadata and Analysis Results Public

Certain information stored in the ABRomics database is extremely useful to the microbiology community because it not only facilitates joint analyses and collaborations but also provides valuable insights in the event of emerging drug resistance.

Therefore, regardless of the status of the sequences and metadata deposited in ABRomics, **certain data (listed below) will remain freely accessible.**

The following rules have therefore been established in agreement with the entire ABRomics consortium:

- Assembled sequences (FASTA) and raw sequence reads (FASTQ) under embargo are accessible only to the depositor and their partners, in accordance with the permissions defined by the project supervisor.
- **Certain metadata for sequences under embargo will be shared** with users who have an ABRomics account (without any specific rights to these sequences):
 - origin of the sample: human, animal, environmental (other)
 - scientific name of the host species
 - scientific name of the strain (genomic data)
 - location (country, city, department, etc.)
- **Certain analysis results for sequences under embargo will be shared** with users who

have an ABRomics account (without any specific rights to these sequences):

- MLST and cgMLST taxonomy
- AMR—predicted resistance genes
- AMR—predicted mutations and antibiotic families targeted by the predicted resistance

How ABRomics Works (Services Offered by the platform)

- Platform User:
 - Interface for creating user accounts and managing user data;
 - Interface for creating a project space and inviting collaborators with specific permissions for that project space;
 - Submission of sequences and metadata entered in the ABRomics metadata repository form or entered directly into the application's web interface;
 - Analysis workflows can be run on data to which the user has access;
 - Current workflows (and future versions) have been optimally configured (documentation describes how the workflows function and the choice of default settings).
- ABRomics provides:
 - Specific project spaces, which will be accessible only to individuals authorized by the project space supervisor;
 - Tools and an environment for performing analyses: there is no concept of memory allocation for running pipelines, nor of disk space per project for storing sequences and results;
 - Automated analysis results (data validation, assembly, ARG, ST, cgMLST, etc.);
 - Tools for visualizing and querying analysis results—integrating both open-access data and private data from the database for which the user has access rights;
 - Data management tools;
 - Tools and interfaces for curating resistance genes over time;
 - An environment to facilitate collaboration and data exchange through access.

Organization and Administration of the ABRomics Platform

Data security

- a. The ABRomics application is hosted on the national computing infrastructure for the life sciences, the IFB Core Cluster. This High-Performance Computing (HPC) cluster is itself hosted within the Data Center (DC) of the *Institut du Développement et des Ressources en Informatique Scientifique* (IDRIS - CNRS). The institute and its building are located in a Restricted Access Zone (ZRR). As such, access is authorized only to individuals with ZRR clearance. In addition to strict access controls, external visitors are accompanied by DC staff.
- b. Within the IFB Core Cluster, access to data is limited to a small number of administrators (4 people) and the developers of the ABRomics application. Administrators adhere to the ethical guidelines inherent to this role and access data only after receiving explicit approval from the data supervisor.
- c. The ABRomics analytics platform and all its software layers (API, Galaxy framework, web interface) are themselves secured by individual usernames and passwords.

Data Backup

- a. The virtual machine hosting the application is backed up to internal storage within the IFB Core Cluster (IDRIS) according to the following schedule:
 - one backup per day, retaining 7 backups
 - one backup per week, retaining 4 backups
 - one backup per month, retaining 3 backups
- b. A daily backup of the database (in dump format) is encrypted and uploaded to a remote site within the IFB platform network (7 copies retained).
- c. A daily backup of file-based data (raw sequences, analysis results) is encrypted and uploaded to a remote site within the IFB platform network (3 copies retained).

Privacy of Personal Data (GDPR)

- a. Opening an account to access the IFB Core Cluster, to use one of the web instances (in this case, the ABRomics application website), or one of the databases hosted by the IFB infrastructure (in this case, ABRomics-DB), requires the collection of the following personal information: last name, first name, affiliation, email address, and ORCID identifier.
- b. This data, as well as any data stored in the user's designated storage space, is not disclosed or shared with third parties under any circumstances.
- c. Connection logs on the IFB servers are retained for one year for legal purposes.

Responsibilities of the ABRomics Platform

- The developers and project leaders commit to monitoring and applying, at a minimum, security updates for the various software layers implemented (stability and confidentiality)
- The administrators of the ABRomics Platform ensure that project data is backed up on-site (see the section "Organization and Administration of the ABRomics Platform"), but they are not responsible for archiving raw data. Users must retain a copy of their data until it is published in one of the international repositories (*i.e.*, the ENA).
- ABRomics Platform developers verify the quality of the sequences provided by the user before running the analysis workflows: the ABRomics Platform issues a quality report, the outcome of which determines how the results of the analysis pipeline are presented (*i.e.*, ABRomics will flag results obtained from sequences of average or poor quality with "warnings").
- The ABRomics Platform undertakes not to disclose "confidential" data or results without the authorization of the data owners through access controls.
- The ABRomics Platform commits to regularly updating the reference databases used by the analysis pipelines.
- The ABRomics Platform is committed to retaining raw sequence data (FASTQ format) on the hosting infrastructure for a **minimum of three years** after initial upload to the ABRomics platform or after submission to EBI-ENA. Sequence files in FASTA format, however, will be retained on the hosting infrastructure's servers indefinitely.
- The developers of the ABRomics Platform commit to submitting the FASTQ files to EBI-ENA, subject to the depositor's consent.

Responsibilities of the Sequence Submitter

- Ensure ownership of the submitted material
- Ensure the validity of the metadata
- Ensure the quality of the submitted sequences: ABRomics is not responsible for incorrect results due to poor sequence quality
- Manage access rights to their data for project collaborators
- Agree to the rules governing data availability
- Maintain the confidentiality of access to the ABRomics platform (username and password)
- For metagenomes, verify that they do not contain sequences from humans or other animal species.

Scientific Acknowledgments

Please acknowledge the ABRomics Platform in all oral and written communications using the following statement:

“The present study used the ABRomics platform which is financially supported by the French Priority Research Programme on Antimicrobial Resistance (PPR Antibiorésistance), coordinated by Inserm and funded by the Secrétariat Général Pour l'Investissement (SGPI).”