

CHARTER OF THE MICROSCOPE PLATFORM

Presentation of the platform

The MicroScope platform is an informatics infrastructure dedicated to the annotation and comparative analysis of microbial genomes and metagenomes. It is located at the Centre National de Sequencage (CEA/Genoscope, Evry), and developed in the « Laboratoire d'Analyses Bioinformatiques pour la Genomique et le Metabolisme » ([LABGeM](#)). The three main components of MicroScope are:

- A complete pipeline for syntactic (i.e. prediction of genes and signals on nucleic sequences), functional (i.e. prediction of biological function of protein-coding genes) and relational (i.e. prediction of synteny conservation and metabolic networks) annotation, detection of SNPS/indels in evolved strains, and differential expression analysis of genes using RNA-Seq data
- A relational database (PkgDB, Prokaryotic Genome Database) which enables organization and management of sequence data, the results of analysis tools, the expert annotations performed by the users of the platform, and a metabolic database (MicroCyc)
- A user-friendly graphic Web interface (MaGe, Magnifying Genome) in which the numerous functionalities enable exploration and analysis of data, and correction/enrichment of expert annotations of the genes

The integration of a large ensemble of bioinformatics tools and the wealth of the PkgDB database in terms of expert functional annotation offer to the Microscope users a unique environment for an efficient exploration of the genomes of the studied organisms. Furthermore, in the context of the explosive expansion of sequencing technologies, the platform is regularly enriched with new functionalities which aim to guide biologists in the analysis of high throughput genomic data, and thus to respond more efficiently to their research problems. This constant evolution of the platform is today enhanced by the integration of MicroScope into the National Infrastructure « [France Genomique](#) » and the « [French Bioinformatics Institute](#) ».

Location and organization of the platform

The MicroScope platform is located on the premises of Genoscope (2 rue Gaston Cremieux 91057 Evry Cedex France). The LABGeM team (directed by Claudine Medigue) develops and maintains the platform, processes the large set of projects submitted to MicroScope and ensure user support. The IT team of the « Laboratoire d'Informatique Scientifique » directed by Claude Scarpelli, supports the LABGeM team in these tasks. The informatics infrastructure of Genoscope in terms of calculation and storage power evolves continuously (>1000 cores and >1 petabytes). The scientific directors of the platform are David Vallenet and Claudine Medigue.

Conditions of access to the platform

The platform is accessible to any public research laboratory which makes a request (both on-site and exterior groups) and to private research entities. MicroScope is open to National, European and

International groups.

Pricing:

- No cost at the moment for academic laboratories
- Private entities should make a specific request to the platform to establish a contract for services/industrial collaboration: contact mage@genoscope.cns.fr. In this case a price quotation is established before any services are provided by the platform to the customer

It is, however, highly recommended to follow the [training course](#) which presents the ensemble of functionalities available on the Web interface: this enables users to be much more efficient in the expert annotation task and in the exploration of the data available in MicroScope.

MicroScope services

The MicroScope platform services are detailed on the « [Service Offer](#) » webpage. All requests should be made through [the forms available on the Microscope Website](#). Users are guided in the different steps of the process and requirements about file formats, integration of public data, etc. are clearly described in specific fields. Once sequence data are uploaded, the user is asked to read and sign our MicroScope « Charter ». Then, a feasibility study of the submitted request follows.

The managers of the platform reserve the right to reject it in the following conditions:

- The submitted data does not fit the requirements (erroneous data format/data content)
- The platform load is such that the time before initiation of the project risks exceeding one month

In the later case the request should be renewed later.

As a guide, once the request has been accepted and taking the load of the Genoscope servers into account, the processing time of a request is:

- roughly 4 weeks for Genome/Metagenome projects
- roughly 2 weeks for RNA-Seq projects

These figures are only an estimate and depend on the computational charge of our cluster and the number of projects processed by the MicroScope platform in the same timeframe.

Once the request has been accepted, a more precise estimate of the delay is given.

While public genomic data is available to all users, access to user submitted data is managed by the administrator(s) (in general the submitter) who may grant rights (consultation and annotation) to other users at any moment. Both the user and the administrator are notified by email of the modifications.

Responsibilities of the platform

- Security of data: This is ensured by regular (differential or total) backups which ensure a primary security of the data. The LABGeM team cannot be held responsible for any damages on the data hosted on its servers.
- Follow-up of services: Users are informed regularly about the progress of the treatment of their

data and the possible problems encountered with the analysis.

- Confidentiality of the results: The platform agrees to neither transmit, distribute nor transfer to third parties the data that have been entrusted to MicroScope without prior authorization from the users who submitted the request. This data is used only in the context of the requested service or to perform internal tests and global analysis (e.g. protein families, genome clusters, pangenome graphs).
- Results continuity and data storage: The platform agrees to keep continuity of the different available functionalities by updating software, databanks and computations. However, the LABGeM team reserves the right to modify or suppress a functionality for maintenance or scientific reasons. Moreover, the LABGeM ensures storage of user submitted data during a minimal period of 6 months. At the end of this duration and after having warned the collaborator, LABGeM reserves the right to delete this data. Finally, in case of refusal of the request service, all submitted data files will be deleted from our servers, whereas service request entries will be kept within our database.
- Quality of the results: Note that the quality of the results depends on the quality of the primary data provided. Therefore the LABGeM can't be held responsible for the poor quality of results.

Responsibilities of the collaborators/users to the platform

- For all requests, contact the engineers and scientists of the platform at the following email address mage@genoscope.cns.fr and provide the « identification code » (i.e. #A4K8P7) of your Delivery of Service.
- Give a precise and sufficient information about the technical characteristics of the sequencing experiment and provide the data in the requested file format (i.e, the web forms should be carefully filled in for initiation of a service, following the requirements listed in the fields).
- The users agree that they are responsible for the initial quality of biological data uploaded on the platform.
- Each user agrees to preserve the confidentiality of his/her password for access to the MicroScope platform and must not give it to any other person, including collaborators.
- The users agree that they retain the intellectual property rights of the data transmitted to LABGeM for provision of services.

Scientific valorization

Scientific valorization of the results must follow the normal rules of signature and acknowledgments.

- In the context of a standard « Delivery of Service », the applicant agrees to cite the reference paper describing the MicroScope platform and to acknowledge the platform for the equipment made available and the persons from the platform for their participation in the analysis.
- However, the applicant agrees that the person(s) from the LABGeM who perform some specific analysis will co-sign the article if the results produced by the platform are essential to the article or the communication.

In both cases, the applicant should also acknowledge the National Infrastructures « [France Genomique](#) » and « [French Bioinformatics Institute](#) ». Please use the following sentence for acknowledgement: "The LABGeM (CEA/Genoscope & CNRS UMR8030), the France Génomique and

French Bioinformatics Institute national infrastructures (funded as part of Investissement d'Avenir program managed by Agence Nationale pour la Recherche, contracts ANR-10-INBS-09 and ANR-11-INBS-0013) are acknowledged for support within the MicroScope annotation platform."