GM - Génomique métabolique

Rapport Hcéres

agence d'évaluation de la recherche
et de l'enseignement supérieur

Section des Unités de recherche

Evaluation report

Research unit :

Genomique Metabolique

University of Evry

March 2009

agence d'évaluation de la recherche
et de l'enseignement supérieur

Section des Unités de recherche

Evaluation report

Research unit :

Genomique Metabolique

University of Evry

Le Président
de l'AERES

Jean-François Dhainaut

Section des unités
de recherche

Le Directeur

Pierre Glorieux

March 2009

Evaluation report

The research unit :

Name of the research unit : Genomique Metabolique

Requested label : UMR

N° in case of renewal : UMR 8030

Head of the research unit : M Jean Weissenbach

University or school:

University of Evry

Other institutions and research organization:

CEA, CNRS

Date of the visit :

November 19, 2008

Members of the visiting committee

Chairman of the commitee :

M. Jean-Michel Claverie, IGS Information génomique et structurale - UPR2589 , Marseille.

Other committee members :

M. Athanasios Lykidis, Walnut Creek, DOE Joint Genome institute , USA.

M. Krishna Mahadevan (University of Toronto)

M. Ruan Yijun (University of Singapour)

M. Hans Lehrach, Berlin, Vertebrate Genomics ,Allemagne

M. Antoine Danchin, CNRS URA2171, Paris

Expert(s) représentant des comités d'évaluation des personnels (CNU, CoNRS, CSS INSERM, représentant INRA, INRIA, IRD.....) :

M. Laurent Duret, CNRS representative.

Observers

AERES scientific representative:

M. Philippe Bouvet

University or school representative:

Ms Jeanine Tortajada, University of Evry

Ms. Anne Fleury-Herard, CEA

Research organization representatives :

Ms. Martine Defais, CNRS

Ms. Anne Fleury-Herard, CEA

Evaluation report

1 • Short presentation of the research unit

- Numbers researchers with teaching duties : 3
- Number of full time researchers: 25
- Number of engineers : 16
- Number of technicians: 5
- Numbers of HDR : 9
- Numbers of students who have obtained their PhD during the past 4 years: 6
- Aerage lenght of a PhD during the past 4 years: 4 years
- Numbers of PhD students: 4, all with a fellowship
- Numbers of lab members who have been granted a PEDR : 1
- Numbers of "publishing" lab members: 26

2 • Preparation and execution of the visit

The visit of UMR8030 (CEA, CNRS, Evry University) took place on Nov 19, 2008. The site visit was nicely prepared, with most of the personnel attending the various oral presentations, and available for answering questions. The written reports provided beforehand were of good quality. Additional fliers were distributed by the 5 individual teams in support of their oral presentation. Following a closed door meeting of the committee members, then joined by the Laboratory director, the first presentation summarized the past activities of the whole laboratory, and was followed by an outlook of the project for the next 4 years by the proposed future director. The rest of the day was then devoted to 5 presentations of the individual team activities, a meeting with PhD students and postdoctoral fellows, a meeting with engineers, technicians and administrative assistants, and a meeting with researchers with permanent position. The site visit ended by a close-door meeting with the Evry University and the CNRS and CEA representatives, then joined by the laboratory Director. For the last two hours, the committee members exchanged their opinions and reached a consensus on the overall notations. Reports were then drafted later on by separate members (2 for each team) and synthesized by the chairman.

3 • Overall appreciation of the activity of the research unit, of its links with local, national and international partners

Historically, the activity of research unit was closely linked with the activity of the Genoscope, the French National Sequencing Center. The Genoscope is now part of the "Genomic Institute" of the French Atomic Energy Commission (CEA).

The main part of the activity, which is ongoing, involves the analysis of the sequence of eukaryotic and prokaryotic genomes using bioinformatics methods. This annotation of sequences in silico is mainly carried out within the framework of the sequencing projects of Genoscope, that are mainly collaborative projects, and performed as a service to the community. As these analyses are leading to original observations on the structure and evolution of genomes with decreasing frequency today.

The director of the research unit wished to enlarge the scope of analysis of sequence data to include the experimental identification of functions. The research on new enzymatic activities involves a collaboration between most of the research unit five research teams, as its overall goal is 1) to obtain an integrated picture of a bacterial organism and, (2) to complete the construction of unknown bacterial metabolic pathways, especially those active under anaerobic conditions. These two objectives are complemented by metagenomics approaches, in particular on the flora of sewage treatment plants.

To continue the collaborative research linked to the Genoscope sequencing projects, and to tackle the above new goals, the research unit was structured into five groups:

- Atelier de génomique comparative ;
- Laboratoire de bioinformatique et d'analyse des séquences ;
- Laboratoire de bioinformatique des réseaux ;
- Laboratoire de génomique et biochimie du métabolisme ;
- Laboratoire de métagénomique des procaryotes.

The first three of these are mainly or entirely focused on bioinformatics whereas the latter two use essentially experimental "wet lab" approaches.

The research unit has produced 116 publications (with review) to be attributed to 25 full-time permanent scientists (CEA and CNRS), 3 professors (University of Evry), 18 engineers and technicians, 2 postdoctoral fellows and 7 PhD students. A majority of these publications (as well as those in the highest ranking journals) originates from collaborations with the Genoscope sequencing activity. It is thus very difficult to evaluate the research unit independently from the participation of some of its personnel to the Genoscope activities. The Genoscope+UMR8030, as a whole site, is very well known and connected nationally and internationally, and its director is recognized as a pioneer and leader in genomic research by the international community. However, it is clear (and very honestly mentioned in the director's report) that the credibility of the UMR8030, a new comer in the field of metabolic biochemistry, is not established, and that the new research direction, as presented by the future director, does not take advantage of the experience of the laboratory in robotics and large-scale biology. Also, the attractiveness of this research topic for PhD students and Post-doc, an important aspect given the relative geographical isolation of the laboratory, might not be sufficient to ensure its proper development in the future years. As the laboratory is now taking some risk in its new direction, it is clearly unfortunate that the current director had to step down as head of the laboratory, and will not even be part of the research group that is the most involved in metabolic biochemistry.

4 • Specific appreciation team by team and/or project by project

- Comparative Genomics

This team includes 4 researchers with or without teaching duties, 1 postdoctoral fellow, 1 PhD student, 3 permanent engineers and technicians and 2 engineers on short term contract.

The "Atelier of Comparative Genomics" has been successful in contributing to microbial genome analysis and comparative genomics by developing computational tools that enable efficient analysis of sequence data. They should continue to maintain and develop state-of-the-art computational approaches, visualization tools, databases and platforms to make sequence data useful and accessible. The development of programs, such as the Microscope platform, is appropriate and needed and their development should be further encouraged.

Strengths: This is clearly one of the best bacterial genomics/bioinformatics research group in France. The "Microscope" database system has no equivalent

Weaknesses: Some progresses have to be made in term of publication of the software tools, and in an increased participation to international genomic/bioinformatic conferences.

Nom de l'équipe : Laboratoire de génomique comparative

Note de l'équipe	Qualité scientifique et production	Rayonnement et attractivité, intégration dans l'environnement	Stratégie, gouvernance et vie du laboratoire	Appréciation du projet
A	A	A	A	A

- Bioinformatic and Sequence Analysis

This team includes 9 researchers with or without teaching duties, 1 postdoctoral fellow, 0 PhD student, 5 permanent engineers and technicians and 1 short term engineer.

This is the key team in bioinformatics in the research unit. It is involved in three main research themes: nucleotide sequence analysis: from assembly to gene annotation; eukaryote genome analysis: evolutive and comparative genomics; and metagenomics: bioinformatics methods for exploring genetic and functional diversity.

Strength:

Excellent collaborations with applied groups with very good publication record;

State-of-the-art tools and methods, mainly adapted from external developments;

Experience with next generation sequencing and other cutting edge technologies;

Very good international collaborations in relevant EU bioinformatics networks (Biosapiens, ENFIN9);

Interesting projects in genome annotation, whole genome duplications and metagenomics.

Weaknesses

Mainly a service group;

Low number of own publications (first, last authorship);

Own developments seem limited (or not published). The work is based mostly on adaptations of existing software.

Nom de l'équipe : Laboratoire d'analyses bioinformatiques des séquences

Note de l'équipe	Qualité scientifique et production	Rayonnement et attractivité, intégration dans l'environnement	Stratégie, gouvernance et vie du laboratoire	Appréciation du projet
A	A	A	A	A

- Bioinformatic of Networks

This team includes 2 researchers with or without teaching duties, 3 postdoctoral fellows, 2 PhD students, technicians and 1 engineer.

This team is the most recent research team in the UMR8030. The objectives were to use bioinformatic techniques for topological and dynamical properties of biological networks with focus on metabolic networks.

The team in addition has also been involved in several projects related to development of software tools for the efficient development and analysis of metabolic networks. In addition, the group is involved in additional projects including dynamic modeling of galactose assimilation in yeast, genetic analysis of E. coli strains, modeling biofilms in chemostats and many other scattered projects. After a careful evaluation of the report and the future project, the following assessment about the strengths and weaknesses as well as some feedback is included. Biological relevance: The topological analysis presented did not convey a clear sense of integration with experimental data and thus had limited biological relevance. The group is asked to focus more on projects which involve biological data and collaborators so that the physiological relevance is clearly present. Abstract projects based purely on computational methods will have to be carefully reviewed for their relevance. In fact the group would substantially benefit from recruiting a microbiologist rather than a computational person.

The current research group has presented less than a dozen conferences, of which only very few are international conferences. Presentations at international conferences are critical in increasing the visibility of the research group and perhaps attracting new sets of students for future research projects. The group has also established strong collaborations nationally with other experimental groups that can increase the biological relevance of the proposed projects. The laboratory has established a core of computationally oriented researchers with excellent skills and who can communicate with the scientists in the institute working on the sequencing and sequence analysis projects

Strengths: This area of computational systems biology is extremely critical for making sense of the large amounts of sequence and other biochemical data that will be generated at the Genoscope. Hence, it is highly important that an integrated experimental and computational research group be present at Genoscope.

The establishment of this research unit hence is very timely and in a short span of time, the unit has adapted to the research demands in this area. They have already gained expertise in metabolic modeling and sequence analysis that will be critical to the future research direction of the group.

Weaknesses: There are several research projects and given the size of the group (which at this point, is fairly small), the presence of many research projects can lead to a lowered research intensity. While such side projects are important, it significantly distracts from the focus that is required during the initial stages of the development of a new research unit. It is highly recommended that the group focuses on a few projects (perhaps no more than three including the A. bayli analysis, yeast metabolome and the microme project), while allowing the rest of the projects to grow organically. It is also recommended that the research does not include too many high risk projects.

Nom de l'équipe : Laboratoire de bioinformatique des réseaux

Note de l'équipe	Qualité scientifique et production	Rayonnement et attractivité, intégration dans l'environnement	Stratégie, gouvernance et vie du laboratoire	Appréciation du projet
C	C	B	B	B

- Genomic and biochemistry of metabolism

This team includes 6 researchers with or without teaching duties, 0 postdoctoral fellow, 1 PhD student, 5 engineers and technicians.

This is a key group in the research unit. It is involved with the biochemical characterization and the discovery of novel metabolic pathways using high-through biochemical assays. So far, they have characterized the lysine fermentation pathways, glucarate degradation pathways, and the ornithine degradation pathways.

Strengths:

The availability of a high-throughput biochemical characterization platform and biochemistry skills along with access to genome sequence from various projects, most notably, the Cloaca maxima project, which is highly interesting. The characterization of the metabolic pathways associated with this pathway could lead to several novel biochemical reactions that can have practical applications.

The construction and expertise associated with the genome wide deletion array for A. baylyi which is critical in evaluating biochemical pathways. The choice of A. baylyi as a model organism for aerobic metabolism can be quite insightful and perhaps components unique to aerobic metabolism could be effectively investigated through the study of this organism.

The proposed transverse research project is highly interesting as it involves the exploration of biochemical potential of the pathways discovered. In addition, the proposed focus on the metabolomics and transcriptomics is appropriate, although the transcriptome can sometimes not provide as much information as expected (if it is not carefully planned). Ideally, the physiology and the omics data are from the same physical samples. The availability of a metabolic model should be leveraged as a data analysis and visualization platform at this point for the integration of the omics data.

Weaknesses:

Collaboration with researchers on the applied side (engineers), who can utilize some of the novel biochemical pathways are missing. This component is quite critical to translate some of the knowledge obtained from biochemical characterization into practical applications. Such collaborations could also point to potential pathways that are not well characterized, the study of which can lead to additional novel pathways.

A lesser involvement of the original team leader (and former laboratory director) in a research topic of which he was clearly the main motor.

The group is highly encouraged to present their research at international conferences to increase visibility as well as make connections with international peers that could lead to seamless scientific communication and collaboration. There were only two oral communications for a group this size, which is certainly a weakness that can be easily overcome.

Nom de l'équipe : Laboratoire de génomique et biochimie du métabolisme

Note de l'équipe	Qualité scientifique et production	Rayonnement et attractivité, intégration dans l'environnement	Stratégie, gouvernance et vie du laboratoire	Appréciation du projet
A	A	B	A	A

- Procaryotic metagenomic

This team includes 6 researchers with or without teaching duties, 0 postdoctoral fellow, 1 PhD student, 2 engineers and technicians.

This team is engaged in cutting edge research in the novel field of metagenomics. They have achieved the complete reconstruction of the genome of an uncultivable bactertium which belongs to an unexplored phylogenetic division. The activities of the group are at the forefront of contemporary genomics research. Further research along these lines should be encouraged and supported.

Strengths: This is probably one of the best research group in bacterial metagenomics, and the one benefiting the best to be linked with a national sequencing center. The group has a very strong publication record.

Weaknesses: The group is understaffed, particularly in term of post-docs and PhD students.

Nom de l'équipe : Laboratoire de métagénomique des procaryotes

Note de l'équipe	Qualité scientifique et production	Rayonnement et attractivité, intégration dans l'environnement	Stratégie, gouvernance et vie du laboratoire	Appréciation du projet
A	A+	A+	A	A

5 • Appreciation of resources and of the life of the research unit

Management :

With the exception of the Director (forced to step down for administrative reasons), the proposed laboratory structure is stable with the same 5 research teams already identified, with the planned development of a small "Chimie organique biologique" group. The laboratory as a whole is very coherent, multidisciplinary, and there is a good collaboration between the research teams as proven by many cross-authored articles. The committee met with PhD students and postdoctoral fellows. Overall they are very satisfied with the working conditions. However, some students have expressed the feeling that inter-team communication is limited. Moreover, discussion with students revealed that it would be useful to set up PhD committees with external advisors to follow the thesis work of students and thus detect and discuss problems that may arise during the course of a PhD project.

Human ressources:

As pointed out by the Director, the recruitment of new personnel is difficult, in part due to the relative isolation of the Evry campus, and more generally due to the bad context of scientific employment in France. There is a particular problem with post-docs (4 of the 5 teams had none during the period of reference), as well as with PhD students (6 during the period of reference). In this context, the emergence of new teams or research areas, as well as of the competitive recruitment of new group leaders from elsewhere in France, or from abroad, is also difficult. Even if the University of Evry appears willing to support the laboratory development, its capacity in term of available positions, or as student provider, remains quite limited due to its small size. It is a pity that a world famous genomic center such as the Genoscope cannot take advantage of the much larger reservoir of the main Paris Universities.

En termes de communication :

Although the laboratory does participate to scientific meetings, and to the diffusion of scientific information to the public, this involvement is far from been equally shared among all research groups and group leaders. These differences in communication skill were also apparent from the oral presentation.

6 • Recommendations and advice

— Strong points :

The UMR8030 + Genoscope center is the leading genomic research platform in France. It enjoys an international recognition and, combined with the collaborative projects it centralizes, has a highly visible publication record (with many articles in the top journals Nature & Science).

— Weak points :

The recruitment of students and post-docs is unsatisfactory, as well as the perspective for attracting permanent scientists. The exclusive association with the small Evry University is probably detrimental to the future of the UMR8030. The priority given to the identification of new enzymatic functions and pathways for the next 4 years is a gambling, that might not take a full advantage of the scientific expertise and

technological experience in large-scale biology acquired by the UMR8030/Genoscope. It is also unfortunate that the past director has to leave the helm of the laboratory, as it is changing its priority from bioinformatics to wet lab research.

— Recommendations :

The committee insisted that evaluating the research unit independently of the activity and R&D of the Genoscope (French National Sequencing Center) is nonsense, and recommended that future evaluations by the AERES should include both.

While maintaining its link with the local Evry-Val d'Essonne university, the UMR/Genoscope should seek ways to become associated with one or several of the much larger and mature universities of the Paris area, and become an active training center in sequencing technology and large-scale biology.

Note de l'unité	Qualité scientifique et production	Rayonnement et attractivité, intégration dans l'environnement	Stratégie, gouvernance et vie du laboratoire	Appréciation du projet
A	A	A	B	A

Génomique métabolique UMR8030

Comments on the report from the visiting AERES committee

We thank the visiting committee of the research unit Génomique métabolique (UMR 8030) for the report that evaluates our activity and for the useful comments and suggestions. We feel however that several of the points raised may result form a misunderstanding that is quite understandable given the limited time (1 day) that was devoted to review the entire Unit. We wish to re-address these points.

We would first like to comment the following sentence of the overall appreciation: *"it is clear (and very honestly mentioned in the director's report) that the credibility of the UMR8030, a newcomer in the field of metabolic biochemistry, is not established, and that the new research direction, as presented by the future director, does not take advantage of the experience of the laboratory in robotics and large-scale biology"*.

Although we are certainly newcomers in the field, our new orientation (re-visiting metabolic biochemistry using genomic approaches) took effect several years ago as was already made fully clear in the previous report. Since the previous evaluation, the UMR teams have established large scale biological resources in the area of functional genomics and acquired new theoretical and experimental expertise and skills in prokaryotic genetics, biochemistry, analytical chemistry and bioinformatics. This reorientation started in early 2000 with the setting up of the Cloaca maxima project, followed by the creation of the "laboratoire de genomique et biochimie du métabolisme" in 2004. The bioinformatics teams also followed this new direction with the exploitation of "omics" data to explore biological processes such as metabolic networks. Our new commitment in metabolic biochemistry has been progressive and a deeper involvement of all UMR teams is still ongoing (e.g. the recent arrival of a chemistry professor). In a next phase we intend to integrate more tightly the various activities through a transversal project aimed at tackling the function of large protein families and that will require the competence of all research groups. Our expertise that now extends from genomics to chemistry (skills that are only rarely found in a single location) constitutes a unique context to make in

depth contributions that should permit the establishment of the unit as an internationally recognized player in this scientific field.

The report mentions in several instances that the new research direction does not "take advantage of the experience of the laboratory in robotics and large-scale biology". We would just like to mention a few facts that may have been overlooked by the committee. During the last few years we carried out

- Construction of a genome-wide deletion array of *Acinetobacter baylyi* ADP1 (~2600 deletion strains). This collection is one of the very few true gene replacement collections described so far. The scarcity of such collections can clearly be explained by the fact that their establishment necessitates a large scale organization and set-up.
- Growth phenotyping of this mutant collection on a large number on different carbon sources (40 000 data points). Initial phenotyping of the collection has been reported two years after the publication of KEIO collection on *E coli*, that originates from an institute that has a much larger *ad hoc* infrastructure. We feel, however, that our contribution in this area compares with the most advanced similar large efforts abroad (usually supported at much higher levels).
- Construction of an ORFeome of A. baylyi that is being used for large scale enzymatic screenings (~2300 ORFs)
- The setting up of a high-throughput cloning and screening platform which is currently used to discover new enzymatic activities for various kinds of purposes.

Furthermore as mentioned in the project part we intend to undertake large-scale metabolomics and transcriptomics studies on the mutant collection

Does the committee believe that the construction of the collections, their phenotyping and the planned large scale experiments can be performed without robotics?

The criticism is difficult to understand, but the committee may not have had the time to really perceive the large scale commitment of the unit.

It has also been indicated that the Bioinformatics and Sequence Analysis group is a service group that has a low number of its own publications. This comment prompts two remarks. First, we agree that the group is first of all a service group: all the CEA staff (i.e. 90% this group) and notably its leader have been recruited with the mission to contribute to the bioinformatics analyses of the sequencing projects carried out at Genoscope. We think however that such a service remains in many instances a true research endeavour and the group is therefore included in the UMR in order to participate in its scientific life. Because of its service duties it would however be unfair to consider the group as a pure basic research group.

Second, members of the group are first authors in three Nature papers and last author in two of these. These two latter papers represent important steps in the analysis of eukaryote genomes and their evolution and are based on thorough data analyses which take time (usually a year and a half or so).

The group has played the leading role in these milestone analyses and has been invited to international conferences. For a "service group" this seems to be of excellent standard.

Jean Weissenbach
Director
UMR8030

Marcel Salanoubat
Proposed director