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

Section des Unités de recherche

AERES report on the research unit

Information Génomique et Structurale

From the

CNRS

February 2011

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

Section des Unités de recherche

AERES report on the research unit

Information Génomique et Structurale

From the

CNRS

Le Président de l'AERES

Didier Houssin

Section des unités de recherche

Le Directeur

Pierre Glorieux

February 2011

Research Unit

Name of the research unit: Information Génomique et structurale

Requested label : UPR CNRS

N° in the case of renewal : UPR2589

Name of the director: Mr Jean Michel CLAVERIE

Members of the review committee

Committee chairman

Mr Patrick FORTERRE, Institut Pasteur, Paris, France

Other committee members

Mr Frédéric DARDEL, University of Paris-Descartes, Paris, France

Mr Patrice GOUET, University of Lyon, Lyon, France (CoNRS representative)

Mr Curtis SUTTLE, University of British Columbia, Vancouver, Canada

Observers

AERES scientific advisor

Mr Pierre LEGRAIN

University, School and Research Organization representatives

Ms Florence NOBLE, CNRS

Mr Pierre CHIAPPETTA, Aix-Marseille University

Report

1 • Introduction

- Date and execution of the visit

The visit took place on February 1st, 2011. In the morning, a general presentation of the scientific activities at IGS by the head of the unit was followed by presentations of the four main research topics by the responsible senior scientists. In the afternoon, discussions with all laboratory members were organized around posters, illustrating various aspects of research at IGS. Finally, three parallel discussions were organized between visiting committee members and different groups of staff members (scientists with permanent positions, engineers and technicians, post-doctoral and PhD students). The visit ended with a closed-door committee meeting.

- History and geographical localization of the research unit, and brief presentation of its field and scientific activities

The IGS unit is an UPR "unité propre du CNRS" located in Marseille on the Luminy campus of the University of Méditerrannée. The IGS is part of the Institut Fédératif de Recherche Mediterranean Institute of Microbiology (IFR88). The group leader (Professor of Medicine) has led the unit since 2006. The IGS combines expertise in genomics, bioinformatics, structural biology and cellular biology. Research activities focus on studying various aspects of microorganism physiology with genomic and post-genomic tools (including structural biology and metagenomics). The IGS has four main research topics: 1) microbial genomics, 2) structural and cellular biology of mimivirus, 3) biodiversity and ecology of giant viruses, 4) applied bioinformatics. In recent years, more emphasis has been put on the study of the giant mimivirus. The unit has been at the forefront of research in this area. The IGS has several partnerships with industries which benefit from their expertise, provide funding and eventually new research topics. The IGS also provides user-friendly tools for genomic analyses through a public internet portal.

- Management team

IGS has no fixed internal structure but is organized around projects that are defined according to scientific interest, funding and manpower. Senior scientists are responsible for a particular project but can be involved in several projects at once. The projects are discussed between the head of the unit and scientists with permanent positions on a rather informal basis.

- Staff members

N1: Number of researchers with teaching duties (Form 2.1 of the application file)	2
N2: Number of full time researchers from research organizations (Form 2.3 of the application file)	6
N3: Number of other researchers including postdoctoral fellows (Form 2.2 and 2.4 of the application file)	1
N4: Number of engineers, technicians and administrative staff with a tenured position (Form 2.5 of the application file)	9
N5: Number engineers, technicians and administrative staff without a tenured position (Form 2.6 of the application file)	
N6: Number of Ph.D. students (Form 2.7 of the application file)	
N7: Number of staff members with a HDR or a similar grade	3

2 • Overall appreciation on the research unit

- Summary

The IGS is a very dynamic unit at the forefront of research on the biology of mimivirus. The track record of the head of the unit is outstanding as well as his international visibility. The research performed at IGS has a strong impact in the debate on the origin and evolution of viruses. The projects on mimivirus host interactions and biodiversity of giant viruses in the environment are challenging but could be highly rewarding.

- Strengths and opportunities

IGS is a world leader in the study of mimivirus; these results have had an international impact with publications of the IGS in highly ranked journals (e.g. Nature, Science, Genome Res., J. Virol.); the objectives of the project are clear and the IGS has excellent expertise in bioinformatics, genetics, molecular biology, biochemistry and structural biology. The study of mimivirus and the search for new giant viruses are exciting topics with the potential for dramatic discoveries.

The interest of the scientific community for giant viruses and the expertise in IGS raises many opportunities of collaboration with other laboratories. The projects of IGS on mimivirus host interactions and giant-virus biodiversity are ambitious and very promising. The proposed projects will open many opportunities to recruit additional staff members

IGS has thought more deeply about giant viruses, and has more expertise on the subject than any other group in the world. There is undoubtedly a huge wealth of undiscovered giant viruses in the oceans, and by bringing the expertise of the IGS to bear on this system promises to yield many exciting and new findings.

The IGS has several partnerships with industry, which benefit from their expertise and provides funding and new research topics.

Other strengths are the access of the IGS to the IFR88 platforms, the on-going national and international collaborations, and the arrival of new researchers in the laboratory.

- Weaknesses and threats

Some parts of the project such as the mapping of the entire interactome are extremely ambitious, and will probably take more than 4 years to complete. There is a lack of expertise in cell biology, which IGS is aware of and for which recruitment is recommended (via CNRS or University).

There are other groups internationally that have more experience than the IGS at tackling questions of diversity in viruses and other microbes in the oceans. This provides some challenges to ensure that projects do not substantially overlap. The emphasis by the IGS on giant viruses provides an angle that is less explored by other groups, which should minimize this potential threat.

A possible weakness is that much of these analyses are computationally intensive, and many of the advances have been the result of development of new analytical tools. Historically, IGS has had considerable strength in this area, but this expertise has been reduced by the loss of key personnel. This is also an area in which it is also extremely hard to recruit, and often involves moving teams of people. It will be important to develop strong collaborations with groups that are developing analytical methods for working with large environmental metagenomic datasets.

The future of the unit, UPR or UMR, and its location (remaining at Luminy or moving to the CNRS campus at Joseph Aiguier) is unclear. The present situation has some advantages. The IGS is located in a convenient building that was set up according to the wishes of the IGS staff. However, the building is now rather full, limiting the possibility of further expansion.

Although the IGS is located on a campus, there is little contact with the university and the numbers of PhD students and post-doctoral associates needs to be increased.

- Recommendations

The team should continue its focus on giant viruses and their biodiversity and evolution; however, the metagenomic data may provide many challenges because of the lack of sequence coverage for each given sample. Consequently, it is also likely a good strategy to take a PCR-based approach to examine diversity by targeting genes that are specifically associated with giant viruses.

Research at the IGS relies on a dynamic approach with flexible project teams. This management has proven to be efficient, but the mimivirus project is extremely diverse and ambitious. It likely requires more defined teams lead by PIs working on specific aspects of the viral cycle.

CNRS and University should rapidly find an agreement for the future of the IGS (i.e. whether they should stay in Luminy or move to Joseph Aiguier)". Any move in terms of statute or location should be made very cautiously in order to maintain the high productivity of IGS. The priority should be to help IGS to conserve its excellence and cutting-edge international reputation.

The international visibility of the Head of unit is very high. The IGS should now strive to increase the visibility of other staff members through promotion, HDR defense and by promoting their participation at meetings (including post-docs and PhD students). At the end of the period it will also be important to carefully consider the future of the unit, knowing that the leader will be approaching retirement age.

The IGS should try to increase the number of post-docs and PhD students.

- Production results

A1: Number of permanent researchers with teaching duties (recorded in N1) who are active in research	2
A2: Number of permanent researchers without teaching duties (recorded in N2) who are active in research	6
A3: Ratio of members who are active in research among staff members [(A1 + A2)/(N1 + N2)]	1
A4: Number of HDR granted during the past 4 years	0
A5: Number of PhD granted during the past 4 years	1

3 • Specific comments

- Appreciation on the results

The research at the IGS has focused on the exploitation of genomic and post-genomic data to understand the biology of various types of microorganisms (bacteria and viruses). During the period 2006-2010, the unit has progressively shifted its main interest from the study of intracellular bacteria of the order Rickettsiae to the study of Mimivirus and more recently to other giant viruses of the NCLDV superfamily.

Mimivirus infects amoebae and raises fundamental questions about the role of viruses in evolution. This is the largest virus ever found with a genome of 1.2 Mb and a virion whose size is comparable to those of small bacteria. The discovery of mimivirus has revived the debate about the origin and nature of viruses and the IGS unit has played an important role in this discussion, bringing both new experimental data and new hypotheses. The IGS has revealed remarkable structural and functional characteristics of mimivirus such as the presence of active viral aminoacyl tRNA synthetases.

Recently, the IGS has developed its own expertise in cell biology in order to initiate new projects on the interactions between the mimivirus and its Amoeba host. For that purpose, the IGS has implemented Acanthamoeba room cultures and is now autonomous to study the replication cycle of mimivirus, and other NCLDVs. The IGS team has already managed to characterize the transcriptome of mimivirus at different times post-infection. This new orientation in cell biology is essential to the success of the project that also relies on the use of the platforms of the joined IFR88 IMM Institute (proteomics, cell imaging…).

The committee notes that the study of mimivirus can lead to unexpected results with great scientific impact. The first image with an X-ray laser of a non-crystalline biological system was obtained in collaboration with the IGS from a single mimivirus particle (Nature, 2011). This remarkable result can reveal the structure of the entire virus (exterior and interior) at the nanometer scale and be combined with crystallographic studies of single proteins (current IGS study on capsid proteins for example). The expertise of the IGS in structural biology is further strengthened by its platform of structural genomics. The performance of the platform was (for example) validated by a study performed with Sanofi-Aventis on lysozyme inhibitors and was published in PNAS in 2007.

A major impact of IGS research has been the discovery of environmental sequences related to mimivirus sequences in metagenomic databases. IGS researchers have published a number of papers that clearly show that relatives of mimivirus are not only widespread in the environment, but that they are also diverse. The researchers at the IGS are now making inroads in this area of research. The IGS is certainly the perceived leader in investigations on the molecular diversity of genetic relatives of viruses related to Mimivirus.

Recently, the IGS has also initiated new projects on the genomic and post-genomic of microalgae and their viruses and on Deinococcaccaea. These are other promising area of research with potential impact in biofuel technology.

For the reporting period (2006-2010) the unit has been highly productive in term of publications. One of the more remarkable aspects of these communications has been the breadth of topics that have been addressed. They range from reports on specific aspects of the nucleic-acid composition of giant DNA viruses (e.g. BMC Genomics; Genome Res.) to analysis of metagenomic data to infer the genetic diversity of giant viruses in nature (Genome Biology; Virology Journal).

The articles on the structure and cell biology of Mimivirus have fuelled this debate on the origin of life. They were published in high-impact journals (structural and functional study of unique viral proteins, J. Virol. 2006, 2007 and 2009, unique mechanism polyadenylation of mimivirus transcripts, Genome Res, 2009, transcriptomic analysis of Mimivirus in Acanthamoeba castellani, Genomic Res, 2010). The desire of the laboratory to promote its research at the international level has also resulted in the publication of significant review articles and a book chapter (Encyclopedia of Virology, 2008).

A number of the papers have been quite provocative for the community of virologists. These include arguments that viruses should not be confused with their virions and that focus should be on the viral factory (Genome Biology) and that viruses should be included in the Tree of Life. Although these papers are largely conceptual, and the ideas might be considered on the fringe of current scientific thinking, they have stimulated a vigorous and healthy debate. This is another example of the strength of the group. It is their willingness to challenge dogma. This has greatly helped to increase the visibility, not only of the IGS team, but has also helped to increase the overall profile of research in to the diversity and evolution of giant viruses.

The work on viral biodiversity has resulted in the publication of a dozen papers related to the diversity of mimiviruses and their relatives. In all but one of these manuscripts, researchers 2009, from the IGS were the primary authors on the manuscript. Typically, these papers have been published in excellent journals, and two shorter commentary papers have been published in some of the top journals in terms of impact (i.e. Science and Nature Reviews Microbiology).

- Appreciation on the impact, the attractiveness of the research unit and of the quality of its links with international, national and local partners

The IGS has established several new long-term collaborations with French and foreign laboratories and with industrial partners. Some of these collaborations have already produced remarkable publications. The recent Nature paper in which IGS is involved further shows the ability of this unit to promote its projects at the international level. This asset will certainly contribute to the continuous funding of this project, which is supported by European and ANR grants until 2013 (3 PIs from the IGS are involved in these grants).

The recent involvement of the IGS in the fields of metagenomics has resulted in the initiation of collaborations with top research groups internationally that are studying giant viruses in the sea, including well-known groups in Norway, and in Japan. The result has been access by the IGS team to other samples for analyses, as well as expanding their knowledge of protocols for collecting samples from the field. As a consequence, the IGS team has initiated new and exciting areas of work such as partnering with the TARA global sampling expedition to begin tackling the challenge of exploring the diversity of potential hosts in the ocean, and isolating new uncharacterized giant viruses.

Another promising collaboration has been set up in the fields of genomic and post-genomic of microalgae with several academic and industrial partners, a LABEX project has been proposed. Finally, a partnerships has been set up with the company DEINOVE for the genomic and post-genomic studies of a private collection of Deinococcus related strains.

In contrast to this high scientific productivity, the number of granted PhD between 2006 and 2010 has been quite low considering the number of permanent staff at the IGS. This problem should be addressed and the number of HDRs should increase during this 4-year period.

- Appreciation on the management and life of the research unit

The head of IGS, who also heads the IFR188, is an outstanding scientist with a high publication record. Although not explicitly addressed in the documentation, members of the IGS have been regular participants at international meetings where they have continued to increase the profile of their work, as well as emphasize the scientific importance of exploring questions on the diversity and evolution of giant viruses.

Several highly productive staff members have left the unit during the last period. This has been compensated for by the recruitment of two young researchers (CR2 CNRS) and the arrival of a senior researcher (CR1 CNRS) with strong expertise in protein and nucleic-acid complexes. The unit has attracted a significant number of grants from ANR and from industrial partners to support their different activities

IGS has no fixed internal structure but is organized around projects that are defined according to scientific interest, funding and manpower. Senior scientists are responsible for a particular project, but can be involved in several projects at once. The projects are discussed between the head of the unit and scientists with permanent positions on a rather informal basis. This management has proved to be efficient, but given that the Mimivirus project is extremely diverse and ambitious, more defined teams lead by PIs working on specific aspects of the projects should emerge in the future.

- Appreciation on the scientific strategy and the project

A major project of IGS in the next four years will be the study of mimivirus-host relationships. With this ambitious project, IGS aims to extend its international connections to capitalize on the pioneering work that it has undertaken since 2003 on mimivirus. The team is particularly interested in studying mimivirus/mimivirus and mimivirus/amoeba interactomes. Biological complexes will first be identified through a double-hybrid approach. The project on the viral transcription complex is well defined and continues previous studies. The large-scale study of the entire interactome of mimivirus (about 1000 ORFs) seems difficult to achieve in the four-year period. IGS does not have enough staff for this task and knows the risk. Therefore, collaboration has been undertaken with a laboratory at Harvard. The project on the development of a minimum in vitro system for replication and transcription is exciting and reflects the relative autonomy of NCLDVs; it can show that mimivirus depends on its host for translation only. The development of new in cellulo techniques is needed for the entire project (eg. transfection) and will require the recruitment of new staff members with a strong expertise in cell biology.

It can be highlighted that focused projects like those on sumo protease, transcription or nucleotide metabolism are well defined and fit well with the expertise of the team (researchers, engineers, technicians). The study of intrinsically unstructured proteins from mimivirus should also be developed considering the expertise of the IGS in bioinformatics and X-ray crystallography/structural biology.

Another major project is the study of the diversity and evolution of giant viruses and their hosts The oceans are vast and largely unexplored with respect to the genetic diversity of protists and the viruses that infect them. The previous work on the diversity and evolution of giant viruses has been instrumental in steering the IGS group into new and exciting areas of work.

There will be significant challenges as well as opportunities as the group moves to encompass more environmental diversity research, as has been outlined in the documentation provided. For example, they will undoubtedly be able to recover copious amounts of data on the number of different taxonomic groups of protists in different areas of the ocean. However, because of bias in data generated by amplification using the polymerase chain reaction (PCR), and by next-generation sequencing technologies, it will be very difficult, if not impossible to generate accurate frequency distribution data for the different taxa. Consequently, it may make it impossible to infer any effects of environmental changes on the composition of protistan communities. Similarly, because of the highly dynamic physical nature of marine systems, and the small-scale changes in the temporal and spatial distribution of organisms within them, teasing apart the environmental drivers and interrelationships within and among protists and giant viruses may not be tractable. However, the team will gather vast quantities of new sequence information that can be used to examine relative abundances of different taxonomic groups across regions. This may reveal interesting and unexpected relationships amongst variables.

Another area that will undoubtedly be challenging is the metagenomic and metatranscriptomic analyses of the samples. The diversity within the viral size fraction is so enormous that even with ultra-deep metagenomic sequencing it is likely that it will be impossible to assemble a significant number of reads. This is compounded by the problem that most virus sequences have no homologues in databases, making it very difficult to assign a functional role. It may very well be that the vast majority of reads will be orphans, because of the lack of coverage, making comparisons among samples of limited value. Finally, because of the overlap in size between cells and viruses, it can be very difficult to determine what is definitely viral and what is definitely cellular. There are a number of well-established and excellent groups, primarily in the United States, that are actively working on these types of questions, in some cases with very similar samples. On the positive side, there are more and more viral metagenomic data bases coming online, which should help greatly to begin closing the sequence space for marine viruses. However, it also means that others will also be addressing questions of sequence diversity in viruses and cells, and what it means. Moreover, any successes by the IGS group in isolating and sequencing new giant viruses will be very helpful for providing template on which to scaffold the sequences. One of the key advantages of the IGS team is that to date the other groups have not focused on giant viruses.

In conclusion, IGS contains a very active and capable group of researchers that has a wealth of expertise that can be used to address questions on the biology and biodiversity of giant viruses and their putative hosts in marine systems.

Intitulé UR / équipe	C1	C2	C3	C4	Note globale
UPR2589 - INFORMATION GÉNOMIQUE ET STRUCTURALE	A+	A+	A+	A+	A+

C1 Qualité scientifique et production

C2 Rayonnement et attractivité, intégration dans l'environnement

C3 Gouvernance et vie du laboratoire

C4 Stratégie et projet scientifique

Statistiques de notes globales par domaines scientifiques

(État au 06/05/2011)

Sciences du Vivant et Environnement

Note globale	SVE1_LS1_LS2	SVE1_LS3	SVE1_LS4	SVE1_LS5	SVE1_LS6	SVE1_LS7	SVE2 _LS3 *	SVE2_LS8 *	SVE2_LS9 *	Total
A+	7	3	1	4	7	6		2		30
A	27	1	13	20	21	26	2	12	23	145
B	6	1	6	2	8	23	3	3	6	58
C	1					4				5
Non noté	1									1
Total	**42**	**5**	**20**	**26**	**36**	**59**	**5**	**17**	**29**	**239**
A+	16,7%	60,0%	5,0%	15,4%	19,4%	10,2%		11,8%		12,6%
A	64,3%	20,0%	65,0%	76,9%	58,3%	44,1%	40,0%	70,6%	79,3%	60,7%
B	14,3%	20,0%	30,0%	7,7%	22,2%	39,0%	60,0%	17,6%	20,7%	24,3%
C	2,4%					6,8%				2,1%
Non noté	2,4%									0,4%
Total	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

*** les résultats SVE2 ne sont pas définitifs au 06/05/2011.**

Intitulés des domaines scientifiques

Sciences du Vivant et Environnement

• SVE1 Biologie, santé
- **SVE1_LS1 Biologie moléculaire, Biologie structurale, Biochimie**
- **SVE1_LS2 Génétique, Génomique, Bioinformatique, Biologie des systèmes**
- **SVE1_LS3 Biologie cellulaire, Biologie du développement animal**
- **SVE1_LS4 Physiologie, Physiopathologie, Endocrinologie**
- **SVE1_LS5 Neurosciences**
- **SVE1_LS6 Immunologie, Infectiologie**
- **SVE1_LS7 Recherche clinique, Santé publique**

• SVE2 Ecologie, environnement
- **SVE2_LS8 Evolution, Ecologie, Biologie de l'environnement**
- **SVE2_LS9 Sciences et technologies du vivant, Biotechnologie**
- **SVE2_LS3 Biologie cellulaire, Biologie du développement végétal**

CENTRE NATIONAL
DE LA RECHERCHE
SCIENTIFIQUE

Information Génomique et Structurale

Information Génomique & Structurale, UPR 2589 CNRS
Institut de Microbiologie de la Méditerranée – IFR88
163 Avenue de Luminy, Case 934, 13288 Marseille cedex 9
Tel : (+33) 4 91 82 54 20 - Fax : (+33) 4 91 82 54 21
www.igs.cnrs-mrs.fr

avril 6, 2011

The Chairman,
Jury AERES
Response to the report on the research unit « IGS », UPR2589

Sir:

In the part of the report listing "**weaknesses and threats**", the jury indicated that: "There are other groups internationally that have more experience than IGS at tackling questions of diversity in virus and microbes in the oceans".
This is perfectly true. However, the IGS laboratory is not addressing these questions in isolation but within several large international projects (e.g. Tara-Oceans, BioMarks) within which the traditional expertise in marine microbiology and ecology is well represented. In those projects, IGS is involved in tasks corresponding to its expertise such as sequence analysis, phylogenetics, and databanking. Moreover, the participation of IGS to these projects is very useful to establish new collaborations within the community of scientists working on marine protists and their viruses, hence gaining access to samples from which to characterize new giant viruses.

In the next paragraph of the same section, the jury indicated a "possible weakness in performing computationally intensive analyses".
If it is true that IGS lost some personnel in Bioinformatics. However its expertise in handling a large volume of NGS data remained intact, as indicated by its IBISA platform status centered on this topic. The IGS laboratory was among the first in France to perform microbial genomics using 454 reads (e.g. as soon as 2009 on Brucella), then transcriptomic studies combining the 454 and Solid approaches, then more recently Illumina reads to survey a large number of deinococci genomes.
The IGS is well experienced in the use of large compute farms, and is now routinely using several grids (Luminy-Grid and PACA-Grid). Again, IGS is not isolated in this area, as the various large scale projects in which we are involved, are done in collaboration with other experienced groups such as the one of Peer Bork at EMBL or the Genoscope. This being said, we do not dispute that recruiting additional scientists/engineers in this area would be helpful. However, we tend to believe that reinforcing IGS in virology, or cellular and structural biology is probably more urgent and indispensable to maintain our leadership and originality in the detailed structural and physiological characterization of newly discovered giant viruses.

Jean-Michel Claverie,
IGS head