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

Department for the evaluation of research units

AERES report on unit:

Cancer and Genome: Bioinformatics, Biostatistics and epidemiology of a complex system

Under the supervision of the following institutions and research bodies:

Institut Curie

MINES ParisTech

Institut national de la santé et de la recherche médicale

January 2013

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

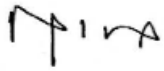
Research Units Department

President of AERES

Didier Houssin

Research Units Department

Department Head



Pierre Glaudes

The president of AERES "signs[...], the evaluation reports, [...]countersigned for each evaluation department by the relevant head of department (Article 9, paragraph 3, of french decree n°2006-1334 dated 3 november 2006, revised).

Grading

Once the visits for the 2012-2013 evaluation campaign had been completed, the chairpersons of the expert committees, who met per disciplinary group, proceeded to attribute a score to the research units in their group (and, when necessary, for these units' in-house teams).

This score (A+, A, B, C) concerned each of the six criteria defined by the AERES.

NN (not-scored) attached to a criteria indicate that this one was not applicable to the particular case of this research unit or this team.

Criterion 1 - C1: Scientific outputs and quality;
Criterion 2 - C2: Academic reputation and appeal;
Criterion 3 - C3: Interactions with the social, economic and cultural environment;
Criterion 4 - C4: Organisation and life of the institution (or of the team);
Criterion 5 - C5: Involvement in training through research;
Criterion 6 - C6: Strategy and five-year plan.

With respect to this score, the research unit concerned by this report and its in-house teams received the following grades:

- Grading table of the unit: Cancer and Genome: Bioinformatics, Biostatistics and epidemiology of a complex system

C1	C2	C3	C4	C5	C6
A+	A+	A	A+	A+	A+

- Grading table of the team: Genetic Epidemiology

C1	C2	C3	C4	C5	C6
A	A	A	NN	A	A+

- Grading table of the team: Computational Systems Biology of Cancer

C1	C2	C3	C4	C5	C6
A+	A	A	NN	A+	A+

- Grading table of the team: Statistical Learning and Modeling of Biological Systems

C1	C2	C3	C4	C5	C6
A+	A+	A	NN	A+	A+

- Grading table of the team: Clinical Biostatistics

C1	C2	C3	C4	C5	C6
A	A	A	NN	A+	A+

Evaluation report

Unit name:	Cancer and Genome : Bioinformatics, Biostatistics and epidemiology of a complex system
Unit acronym:	
Label requested:	INSERM
Present no.:	U900
Name of Director (2012-2013):	Mr Emmanuel BARILLOT
Name of Project Leader (2014-2018):	Mr Emmanuel BARILLOT , Curie Institute and INSERM Mr Jean-Philippe VERT, MINES ParisTech

Expert committee members

Chair:	Ms Hélène TOUZET, LIFL, Lille
Experts:	Mr Jenny CHANG-CLAUDE, DKFZ, Heidelberg, Germany
	Ms Florence D'ALCHE-BUC, IBISC, Université d'Evry-Val d'Essonne
	Mr Christophe JAGLA (Representative of CSS INSERM)
	Mr Stefan MICHIELS, Institut Gustave Roussy
	Ms Marie-France SAGOT, INRIA Rhones Alpes, Lyon
	Ms Isabelle STUCKER (Representative of CSS INSERM)
	Mr Alfonso VALENCIA, CNIO, Madrid, Spain
	Mr Lodewyk WESSELS, Netherlands Cancer Institute, Amsterdam

Scientific delegate representing the AERES:

Mr Jacques HAIECH

Representative(s) of the unit's supervising institutions and bodies:

Mr Daniel LOUVARD, Institut Curie

Ms Marie-Pascale MARTEL, INSERM

Mr Romain SOUBEYRAN, MINES ParisTech

1 • Introduction

History and geographical location of the unit

U900 is part of the Institut Curie (IC), that hosts a major multidisciplinary research center for cancer. The unit originates from a bioinformatics platform that has been founded in 2003 on the Paris Campus of IC. From the start, it was clear that the platform was also supposed to develop research projects. The unit was created in 2008 (INSERM unit U900). A consortium agreement was negociated between IC, INSERM and MINES ParisTech and Armines. The unit is in fact similar to a federation of two research units, one supported by IC and INSERM under the direction of Mr Emmanuel BARILLOT and one mainly funded by MINES ParisTech under the direction of Mr Jean-Philippe VERT. The bioinformatics platform is linked to the unit directed by Mr Emmanuel BARILLOT.

The whole unit is now composed of four teams, in addition to the bioinformatics platform:

- Team 1 : Genetic Epidemiology (Ms Nadine ANDRIEU)
- Team 2 : Computational Systems Biology of Cancer (Mr Emmanuel BARILLOT)
- Team 3 : Statistical Learning and Modeling of Biological Systems (Mr Jean-Philippe VERT)
- Team 4 : Clinical Biostatistics (Mr Bernard ASSELAIN)

The unit is located on three sites across Paris (Pôle de Biologie du Développement et Cancer of IC, Mouffetard street and Fossés-Saint-Jacques street) and one site at Fontainebleau (campus MINES ParisTech).

Management team

The unit is in fact composed of two units, the first one involves a partnership between IC and INSERM and is directed by Mr Emmanuel BARILLOT, the second one involves a partnership between IC and MINES ParisTech and is directed by Mr Jean-Philippe VERT.The two units constitute a unique entity similar to a research federation.The consortium agreement between IC, MINES ParisTech, Armines and INSERM defines the rights and duties of each partner.The two units are independent as far as administration and finances are concerned.The first unit comprises Team 1,2,4 and the bioinformatics platform.The second unit is composed of Team 3.

AERES nomenclature

SVE1_LS2

Unit workforce

Unit workforce	Number as at 30/06/2012	Number as at 01/01/2014	2014-2018 Number of project producers
N1: Permanent professors and similar positions	1	1	1
N2: Permanent researchers from Institutions and similar positions	9.2	11.35	10.5
N3: Other permanent staff (without research duties)	8.05	8.05	2
N4: Other professors (Emeritus Professor, on-contract Professor, etc.)			
N5: Other researchers from Institutions (Emeritus Research Director, Postdoctoral students, visitors, etc.)	9	9	8
N6: Other contractual staff (without research duties)	9	9	
TOTAL N1 to N6	**36.25**	**38.40**	**21.5**

Percentage of producers	***100 %***

Unit workforce	Number as at 30/06/2012	Number as at 01/01/2014
Doctoral students	12	
Theses defended	12	
Postdoctoral students having spent at least 12 months in the unit*	8	
Number of Research Supervisor Qualifications (HDR) taken		
Qualified research supervisors (with an HDR) or similar positions	6	7

2 • Assessment of the unit

Strengths and opportunities

- Since its foundation in 2008, the unit has reached a critical mass (from 50 people in 2008 to 80 now) and succeeded in creating a positive dynamic that is reflected by the significant scientific achievements;
- The unit is embedded in a clinical and biological environment, and is highly multidisciplinary. It has attracted biologists, mathematicians and clinicians, all focused on computational cancer research. It provides a continuum of expertise from epidemiology, modelling to clinical trials and has strong interactions with other departments of IC. This combination is unique in the French national research landscape;
- The two service platforms hosted by the unit (bioinformatics and investigation platforms) have developed a sought-after expertise that is of great value for the IC community. They also play a major role in the research activity of the unit, by allowing for cross-fertilization.

Weaknesses and threats

- Some of the teams are understaffed and could benefit from more senior researchers;
- The software production for the period is good (23 packages). However it still lacks of visibility;
- Some teams suffer from a lack of space which makes it difficult to deploy new activities;
- Several full-time researchers hold an engineer's position, which may not conform to their missions and may compromise their career progression.

Recommendations

- Pursuing and strengthening the collaborations between the different teams;
- Maintaining a good balance between methodology-driven projects and applicative biology-driven projects. A way to achieve this is to prioritize new partnerships;
- Acquiring more skills in software packaging and software development best practices, and establishing a common policy for software diffusion. This effort could be initiated by the bioinformatics platform, and then disseminated to all teams;
- Continuing the organizing efforts to better delineate the missions of the service platforms and the research groups;
- Developing specific management methods for short-term workers (especially on service platforms) both to enhance the sustainability of projects, and to support people in their professional development;
- It might be beneficial to add an international expert in biostatistical research to the Scientific Advisory Board of the unit.

3 • Detailed assessments

Assessment of scientific quality and outputs

The scientific production of the unit is outstanding. It covers more than 300 publications for the period 2008-2012, the majority of them being in high-standard international journals or conferences. The members of the unit have obtained several key results on cutting-edge topics: epidemiologic studies in familial breast cancer, tools for inferring and analyzing biological networks, genomic and transcriptomic signatures in several forms and stages of cancer. Many of these contributions are important both on the methodological side and in terms of applicative impact.

Assessment of the unit's academic reputation and appeal

The unit has gained high recognition at the national and international levels.

- It is involved in 50 grants, among them 8 European projects, 1 ERC grant, 3 “investissements d’avenir” (1 coord), 4 private grants;
- It took part in the organization of 17 international scientific events (conferences or workshops);
- It established several strategic international partnerships in Europe (EMBL, EORTC, Karolinska, MRC, Leicester...), Japan (Kyoto...), USA (Berkeley...);
- It is successful in attracting new researchers (4 positions), doctoral and postdoctoral fellows;
- The members of the unit have participated as invited speakers in more than 120 international conferences and workshops.

Assessment of the unit's interaction with the social, economic and cultural environment

- The activity of the unit has direct impact on translational research. It has four registered patents and several active collaborations with pharmaceutical and biotechnological companies;
- It participated in several TV and radio interviews for the general audience.

Assessment of the unit's organisation and life

- The unit organizes a general scientific meeting every two weeks, as well as a retreat every two-years. In addition, there are several journal clubs and book reading groups on well-oriented topics;
- The unit has three laboratory councils per year which deal with the administrative and financial life;
- The unit is setting up a unified computational environment, with shared computing facilities;
- In general, the unit and the teams seem extremely dynamic and well organized, with all the members, senior and junior, permanent and non-permanent, demonstrating a passion for the work they are conducting;
- The synergy between the teams also finds its expression in several transverse collaborative projects (SHIVA, Genesis, ABS4NGS).

Assessment of the unit's involvement in training through research

- The members of the unit have a strong commitment to training;
- They participate in numerous courses including biostastics, epidemiology and computational biology for graduate students;
- They have co-organized 10 INSERM workshops, and have been regularly involved to several EBI courses since 2009;
- They regularly supervise PhD students (12 PhD theses defended, 12 on-going).

Assessment of the five-year plan and strategy

The scientific project for the next five years mainly relies on two pillars: high-throughput sequencing and personalized medicine in cancer research. This is an important and topical issue, with intense international competition. The unit is in excellent position to take up such a challenge because of its affiliation to Curie Institute and its previous experience.

4 • Team-by-team analysis

Team 1 : Genetic Epidemiology

Name of team leader: Ms Nadine ANDRIEU

Workforce

Team workforce	Number as at 30/06/2012	Number as at 01/01/2014	2014-2018 Number of project producers
N1: Permanent professors and similar positions			
N2: Permanent EPST or EPIC researchers and similar positions	1	2	2
N3: Other permanent staff (without research duties)	3.25	3.25	
N4: Other professors (PREM, ECC, etc.)			
N5: Other EPST or EPIC researchers (DREM, Postdoctoral students, visitors, etc.)			
N6: Other contractual staff (without research duties)	5	5	
TOTAL N1 to N6	**9.25**	**10.25**	**2**

Team workforce	Number as at 30/06/2012	Number as at 01/01/2014
Doctoral students	2	
Theses defended		
Postdoctoral students having spent at least 12 months in the unit		
Number of Research Supervisor Qualifications (HDR) taken		
Qualified research supervisors (with an HDR) or similar positions	1	2

• Detailed assessments

Assessment of scientific quality and outputs

The "Genetic Epidemiology" team led by Ms Nadine ANDRIEU plays an important role in this unit in bridging epidemiological and clinical research. The main goal of the team is to identify the role of genetic factors and their interactions with non-genetic determinants of the etiology of cancer, and thus provide the basic knowledge required for the development of effective strategies for primary and secondary prevention. Compatible with the large programme of clinical research in breast cancer at the Curie Institute, the team has a research focus on breast cancer and has successfully established nationally epidemiologic studies of populations at high genetic risk for breast cancer.

As the head and sole senior scientist in the team, Ms Nadine ANDRIEU is an internationally well-recognised genetic epidemiologist in the field of breast cancer genetics. She has an excellent publication list with papers in leading journals of the field (*British Medical Journal-BMJ-IF 14, Journal of Clinical oncology –JCO IF ~19, Cancer Epidemiology Biomarkers & Prevention-CEBP IF ~5, Breast cancer research- BCR IF~5.25)*. Many of her publications involve international collaboration and provide clear evidence of her international standing.

Ms Nadine ANDRIEU has been instrumental in the establishment of important epidemiologic studies of familial breast cancer, such as GENEPSO and GENESIS, and of AT family members in CoF-AT. Following her previous work in collaboration with the International BRCA1/2 mutation Carrier Cohort Study, she recently made important observations regarding variation in risk according to radiation exposure, life-style risk factors and location of mutation based on the French national cohort GENEPSO. If these findings are validated in further, particularly prospective, cohorts, they will have very important clinical implications for the medical care of mutation carriers. She has also always pursued methodological research directly related to her field of research, for example in estimating penetrance in BRCA1/2 families, which is clinically highly relevant.

Assessment of the team's academic reputation and appeal

Both the rarity of hereditary cancer forms and the polygenic contribution to cancer susceptibility have necessitated multidisciplinary collaborations to collate large numbers of families, patient cohorts, or case-control samples for informative research in genetic epidemiology. The team has established multiple long-standing collaborations with all the major players in the field of genetic epidemiology of breast cancer at the international level, both for the thematic and methodological research. Ms Nadine ANDRIEU is also well established nationally as principal investigator of several large multicenter studies GENEPSO, CoF-AT and GENESIS. This involves close collaboration with clinical colleagues for patient recruitment and follow-up as well as with research scientists performing the molecular and genetic analysis of the biological samples collected. She was recipient of the prestigious Prize of Ligue contre le Cancer Yvelines in 2009.

Recently, a new research scientist who has an excellent publication record, was recruited to join the group in 2013.

Assessment of the team's interaction with the social, economic and cultural environment

In response to expert advice on the French oncogenetic activity requested by INCa, the team has developed a new scoring system for the diagnosis of BRCA1/2 associated breast-ovarian cancer predisposition. The scoring system could lead to more efficient genetic diagnosis when validated using real data from the French oncogenetic centers (which will be carried out in the next research period).

Ms Nadine ANDRIEU is active also in strengthening methodological research and organized a methodological workshop at INSERM in 2009 to introduce new developments in the design of epidemiological surveys.

Assessment of the team's involvement in training through research

Ms Nadine ANDRIEU is involved in teaching for international and national masters programmes and training of young scientists, including bachelor, masters and PhD students. Two PhD students have successfully completed their dissertation in the review period.

Assessment of the five-year plan and strategy

The establishment of the Cancer Genetic Epidemiological Investigation (CGEI) platform is timely. It is excellent that such a supportive platform now exists to serve the design and implementation of genetic epidemiological projects at the IC. This platform is expected to facilitate multidisciplinary translational research. It will be important to observe the demand for services and collaboration of the CGEI platform and provide appropriate additional personnel support.

The team will naturally build in part upon national resources as well as international collaborations already established for some of its research activities in the next 5 years. Both the validation of the modification in risk of breast cancer in BRCA1/2 mutation carriers by life-style factors, such as radiation exposure, as well as the new scoring system for prediction of mutation carriage are extremely important before the translation of this knowledge into clinical practice. The GENESIS study, which will be completed soon, provides an excellent and unique resource for identifying and characterizing new inherited susceptibility loci for breast cancer. The team has already obtained funding to genotype this population with the iCOGS chip, which is enriched for susceptibility loci of hormone related cancers. The newly recruited research scientist has an active programme to carry out complementary research activities using exome or targeted sequencing, with a focus on identifying uncommon or rare genetic variants associated with moderate penetrance. Although this is an internationally highly competitive research area, this new recruited scientist can draw upon the unique GENESIS study as well as her own international and national collaborations also for melanoma. The in-depth mining of data generated from the different studies will profit from the timely establishment of a Systemic Epidemiology Group involving this team together with researchers from both the MINES ParisTech team and the Computational Systems Biology team of the unit.

Conclusion

- Strengths and opportunities:

The team is well placed in U900 to carry out more internationally competitive cancer research. They have strengthened their profile at Curie Institute by establishing the Cancer Genetic Epidemiological Investigation (CGEI) platform, which should also foster clinical genetic epidemiological research at Curie Institute. Their research agenda has broadened to other cancer sites, such as melanoma and neuroblastoma, and will also involve new technologies, such as next generation sequencing. The intramural collaboration between the groups, made possible through a Systemic Epidemiology Group, is a positive development towards optimal utilization of a unique continuum of expertise in the unit. The interdisciplinary work of basic scientists and clinicians will likely lead to innovative approaches in the field of genetic epidemiology and to novel findings.

- Weaknesses and threats:

Ms Nadine ANDRIEU was the sole senior research scientist in the team in the past research period, thus limiting the recruitment and training of young scientists. The situation has improved with the recruitment of an additional research scientist. Expansion of research in the area of clinical genetic epidemiology at Curie Institute will however require more research personnel.

The physical location of the team puts it somewhat at a disadvantage for encouraging colleagues at Curie Institute to seek collaboration as well as for collaborative work with the other teams of the unit.

- Recommendations:

With the completion and established resources of several large genetic epidemiologic studies, the team scientists are in an excellent position to initiate and pursue innovative research questions, which can be followed up in international collaborations already existing, and are therefore recommended to give this greater priority.

In parallel to the research activities aimed at assessing the relevance of low-dose radiation exposure in modifying breast cancer risk in BRCA1/2 mutation carriers, the team could consider stimulating collaboration between lab-based and epidemiological studies to carry out further research on this topic using more (system) biological approaches. This seems particularly relevant to Curie Institute with its history evolving from the discovery of radioactivity and the strong presence of research in radiobiology.

The team is strongly encouraged to expand the research personnel in order to strengthen their research agenda, in particular by recruiting talented post-doctoral fellows. This would require the provision of adequate working space.

Team 2 : Computational Systems Biology of Cancer

Name of team leader: Mr Emmanuel BARILLOT

Workforce

Team workforce	Number as at 30/06/2012	Number as at 01/01/2014	2014-2018 Number of project producers
N1: Permanent professors and similar positions			
N2: Permanent EPST or EPIC researchers and similar positions	2	2	2
N3: Other permanent staff (without research duties)	2	2	2
N4: Other professors (PREM, ECC, etc.)			
N5: Other EPST or EPIC researchers (DREM, Postdoctoral students, visitors, etc.)	8	8	8
N6: Other contractual staff (without research duties)	1	1	
TOTAL N1 to N6	**13**	**13**	**12**

Team workforce	Number as at 30/06/2012	Number as at 01/01/2014
Doctoral students	2	
Theses defended	6	
Postdoctoral students having spent at least 12 months in the unit	8	
Number of Research Supervisor Qualifications (HDR) taken		
Qualified research supervisors (with an HDR) or similar positions	2	2

Detailed assessments

Assessment of scientific quality and outputs

Over the period 2008-2012, Team 2 has had a very strong scientific production, both in terms of publications (65 journal papers, in some of the best bioinformatics venues, including a *Mol. Syst. Biol.*, plus 2 books – one of which was co-authored by members of Team 3 – and 13 book chapters) and of software development (11 publicly available in GNU GPL, 7 for NGS (Next Generation Sequencing) analysis and 4 for network analysis – this number includes notably a Cytoscape plugin for the analysis of biological networks, BiNoM, and a software for the navigation and curation of molecular interaction maps, NaviCell, both available through a web interface). The team has also already a great part of and intends to complete an Atlas of Cancer Signalling Networks that certainly will be extremely useful to the community. All of this represents a quite substantial output, specially for a team that is relatively small (5 permanent persons) in view of its production and scientific engagements.

Assessment of the team's academic reputation and appeal

Team 2 has an extremely good academic reputation in France as well as internationally as is testified by the fact that its permanent researchers have been invited speakers in important international conferences, either on systems biology or more specialized in cancer research (such as the EMBL Conference on Cancer Proteomics). The team is also partner in the first European integrated project on systems biology of cancer (APO-SYS), as coordinator of the computational part, and in general participates in or coordinates an impressive number of national and international scientific projects.

The software publicy available, and specially the Atlas of Cancer Signalling Networks that the team maintains and proposes to extend, increases the appeal of the team, and would deserve to be very well supported in terms of manpower (meaning also working space).

Assessment of the team's interaction with the social, economic and cultural environment

There is clearly a strong impact of Team 2 on the economic environment through all the software and databases developed and made available publicly as is standard practice in the field of bioinformatics, as well as on the social and economic environment through their potential and actual impact on health-related issues, notably through a collaboration with the Servier pharmaceutical company for finding therapeutic targets in breast cancer. The same applies at the international level through the team's participation in EU projects that enable to link research being conducted in France / Institut Curie with on-going large scale efforts.

Assessment of the team's involvement in training through research

There has been an extremely good involvement of Team 2 in training through research with 6 PhD students in the period 2007-2012 and a considerable amount of organization or participation to schools and workshops at both the national and international (e.g. EBI) levels, testifying once more to a very good international visibility of the team. Indeed, the team has been actively involved in many of the more than 80 courses organized by the bioinformatics platform since 2008, including internal, PhD level, and international courses.

Assessment of the five-year plan and strategy

The five-year plan of Team 2 appears very ambitious, with 4 axes, each of which represents a substantial amount of work on problems such as: modelling of signalling networks for fundamental and practical issues such as personalized medicine, atlas of such networks in cancer, tools for network analysis, and data integration. The topics are well within the expertise of the team and extend the work they have already been conducting, with some newer parts proposed, notably on NGS data analysis, that follow the recommendations of a previous assessment. All are extremely pertinent, and well supported by national and international funding. The work on NGS notably is supported financially by a large national project (ABS4NGS investissements d'avenir bioinformatics project, 2012-2016) that involves 6 other partners and is coordinated by Team 2.

There will be a need to ensure there is enough human support to conduct all these ambitious projects, and this may represent one of the threats to some of them given that the team will not be able to expand much more in terms of number of non permanent, and specially of permanent (and senior) researchers. This is a problem that will need to be addressed and concerns all teams in the Unit. In some cases, such as software development, part of the work could be subcontracted, and the fact that the team appears in very good health in terms of funding means this will be possible for at least the next five years. Additionally some of the proposed research topics would be better addressed by scientific collaborations with groups in and outside the unit. In that sense also, the team appears very healthy.

Overall, the general impression of the team and of its plan for the next five years is therefore an extremely positive one.

Conclusion

- Strengths and opportunities

Team 2 is specialized in the analysis of high throughput data and in the modeling of signaling networks involved in tumorigenesis and tumor progression using such data.It proposes in particular to address the question of personalized medicine and to continue developing an atlas of cancer signaling networks, a huge task which will be extremely valuable to the community. The team responded quite well to a recommendation of the previous evaluation by developing not only methods for network analysis (structure and dynamics based) which were already part of its strength, but also methods to analyse NGS data which was an aspect somewhat less addressed by this and the other teams of the Unit U900.The level and quality of the publications of the team are very good.Team 2 is thus developing a highly important research, at the level of both methods and models, in a way that is very fruitful to the interests of the Institute and of the whole community in general.

- Weaknesses and threats

Although Team 2 develops important methods, it would need to invest more resources for software development and packaging that would in particular help increase even further its already strong international visibility and competitiveness. This is a relative weakness that the team (together with the unit in general) acknowledges in the report and which it proposes to address in the next five years. The fact that the team (and the unit more in general) cannot increase too much its size may be a potential threat to this ambition.This threat is addressed for now by subcontracting some of the work involved in software development and packaging, and by collaborating with other groups through the numerous projects the team has.

- Recommendations

Team 2 is doing an excellent work, both at the methodological and modeling levels in particular as concerns the cancer signaling networks atlas, and our main recommendation would therefore be to continue doing their work with the enthusiasm and dynamism that they have been already demonstrating.Some attention should be paid to increasing even further the international visibility that the team already has, in particular by investing more – as proposed – on software development and packaging for the methodology it elaborates using a mixture of in-house human investment and subcontracting. It is worth also to consider increasing the interaction with Team 3 for the mutual benefit of both teams, and thus of the unit in general. Team 2 could profit from the expertise in the more theoretical aspects of machine learning methods of Team 3, and Team 3 from the expertise in bioinformatics / biology of Team 2. The same type of comment can be made in relation to Teams 1 and 4. In particular, there are great opportunities of more intense collaboration with Team 4, notably on the SHIVA clinical trials as well as other trial set-ups, by for example introducing network analysis in their pipelines.

Team 3 : Statistical Learning and Modeling of Biological Systems

Name of team leader: Mr Jean-Philippe VERT

Workforce

Team workforce	Number as at 30/06/2012	Number as at 01/01/2014	2014-2018 Number of project producers
N1: Permanent professors and similar positions	1	1	1
N2: Permanent EPST or EPIC researchers and similar positions	2.5	3.5	3.5
N3: Other permanent staff (without research duties)	1	1	
N4: Other professors (PREM, ECC, etc.)			
N5: Other EPST or EPIC researchers (DREM, Postdoctoral students, visitors, etc.)	1	1	
N6: Other contractual staff (without research duties)			
TOTAL N1 to N6	**5.5**	**6.5**	**4.5**

Team workforce	Number as at 30/06/2012	Number as at 01/01/2014
Doctoral students	6	
Theses defended	6	
Postdoctoral students having spent at least 12 months in the unit		
Number of Research Supervisor Qualifications (HDR) taken		
Qualified research supervisors (with an HDR) or similar positions	2	2

• Detailed assessments

Assessment of scientific quality and outputs

Team 3 is a laboratory of MINES ParisTech. It develops original and excellent research essentially in machine learning and statistics dedicated to computational biology. First of all, the team interacts with several biology groups of IC, providing them with dedicated algorithms based on the expertise of the team. Consequently, the research activities cover a large set of biological topics brought by the biologists. However, the methods such as penalized regression and kernel-based modelling are common features of these applications. Other collaborations deal with Pharma labs especially in chemoinformatics. The main achievements of Team 3 concern systems biology, drug design and tools for personalized medicine. In systems biology, new learning methods have been proposed to infer gene regulatory networks either de novo (TIGRESS) or using labelled data (SIRENE). Relevant results have been also obtained in gene prioritization. In drug design, new methods for *in silico* virtual screening, chemo-genomics and side-effect prediction were developed with applications. For personalized medicine, tools for classification and prediction of molecular signatures have been built. The publication record of team 3 in peer-reviewed international journals is very impressive in the last 5 years: 43 papers. Publications alternate application-based work in response to specific problems brought by the biologists of the Institute and more fundamental papers about new developments in machine learning.

Assessment of the team's academic reputation and appeal

The team is highly respected in the community: the competences in kernel methods and computational biology are internationally renowned and were recently rewarded by an ERC-grant. Work on inference of biological networks and chemo-genomics have reached a very high visibility in the international community and are considered as reference works in their respective domains.
International collaborations are also developed with other machine learning and statistics teams in UC Berkeley and Washington University.

Assessment of the team's interaction with the social, economic and cultural environment

Team 3 interacts with the social and economic environment in different ways. Collaborations with several Pharma labs: Sanofi-Aventis, Pierre Fabre and Biomérieux were initiated and led to research contracts and one CIFRE PhD fellowship. Some of them are conducted within “Pôle de compétitivité”, a French mechanism that funds common projects between small and big companies and research labs. Two patents were recently filed at the European level: one on the assessment of sensitivity of a subject to breast cancer treatment and the other on virtual screening for identifying proteases inhibitors.

Assessment of the team's involvement in training through research

Team 3, as a Mines-ParisTech lab, is involved in teaching at Ecole Nationale Supérieure des Mines and is responsible in this school of all the courses related to biostatistics and machine learning. The team also participates to Master MVA (ENS Cachan) and welcomes several Master Students per year. 6 PhD students are currently on-going and 6 have defended during the past period. The team also participates to the coordination and the organization of the SMILE seminar in Paris about Machine Learning, especially intended for PhD students but also researchers of the community.

Assessment of the five-year plan and strategy

Team 3 has slightly reduced its broad range of themes into 3 main axes, which is an improvement upon the previous plan but remains still very ambitious regarding the size of the group. It covers the three main levels of biological organization (molecular, cellular and organism levels). On the first axis, devoted to genetic and epigenetic regulation of expression, emphasis has been put on NGS data, which are now a major source of information for gene expression analysis. Team 3 has recently begun to be involved in this area and wishes to continue to be active, especially through the "investissement d'avenir" project called ABS4NGS where the unit plays a leading role and also a EC-funded project called RADIANT. This activity on NGS data is strongly linked to the network inference and data integration subprojects. At the cell-level, a new field concerning biomages and the quantification of complex cellular phenotypes with time-lapse microscopy will be investigated based on the arrival of a new researcher. On one hand, this project has a high relevance since cells imaging will allow a better modelling of cell population and corresponding phenotypes. It will be interesting for team 3 to interact with team 2 on dynamical modelling and to cross-fertilize their expertise to achieve better modelling useful in cancer tissue. There seems to be a golden opportunity here as Team 2 is focusing on a knowledge based approach for pathway reconstruction while Team 3 is following more data-driven approaches. An integration of these complementary efforts may result in very good quality models. On the other hand, the relatively small size of the group with a lot of different ambitious goals may be an issue. A new direction, very promising, concerns the prediction of cell response to perturbation which will be addressed extending the quantitative structure-activity relationship paradigm to a quantitative structure-phenotype relationship. Activities on virtual screening, very well known to the community, will be developed using active learning. The third axis concerns predictive models for personalized medicine where a significant activity is planned. Many collaborations with attached contracts will support this research, for instance with Institut Gustave Roussy, teams from Curie Institute and UC Berkeley. Overall, the biological applications targeted in the project will benefit from the methodological background developed in the ERC grant SMAC that covers the 2012-17 period. In terms of funding's, the team has reached a high level of coverage, allowing for devoting most of the time to research and dissemination.

Conclusion

- Strengths and opportunities

Team 3 has developed a strong expertise in machine learning and statistics for computational biology and chemoinformatics. Its situation in the unit allows for rich and efficient interactions between the technological platform and biology groups of Institute Curie. Working already in harmony with the other teams, it faces the opportunity to contribute to major discoveries in cancer biology.

- Weaknesses and threats

The large spectrum of biological topics addressed through collaborations may be seen as a weakness. Although very efficient, the group may be slowed down by too many requests from biologists on a too wide range of subjects.

- Recommendations

As the group has acquired a very good reputation and has proven its ability to solve in a fast and efficient way questions from different groups of biologists, it could be wise to put more focus on one or two biological questions, in order to get a chance to contribute to a major discovery in biology. Collaborations with other teams (Team 2 and Team 4) should also be emphasized. Team 2 and Team 3 are closely connected in terms of topics as proven by the common writing of a book about systems biology in cancer. As already mentioned in the report for Team 2, more intense collaborations between those two groups should provide excellence results for the mutual benefit of both teams.

Team 4 : Clinical Biostatistics

Name of team leader: Mr Bernard ASSELAIN

Workforce

Team workforce	Number as at 30/06/2012	Number as at 01/01/2014	2014-2018 Number of project producers
N1: Permanent professors and similar positions			
N2: Permanent EPST or EPIC researchers and similar positions	3.70	3.85	3
N3: Other permanent staff (without research duties)	1.8	1.8	
N4: Other professors (PREM, ECC, etc.)			
N5: Other EPST or EPIC researchers (DREM, Postdoctoral students, visitors, etc.)			
N6: Other contractual staff (without research duties)	3	3	
TOTAL N1 to N6	**8.5**	**8.65**	**3**

Team workforce	Number as at 30/06/2012	Number as at 01/01/2014
Doctoral students	3	
Theses defended		
Postdoctoral students having spent at least 12 months in the unit		
Number of Research Supervisor Qualifications (HDR) taken		
Qualified research supervisors (with an HDR) or similar positions	2	2

• Detailed assessments

Assessment of scientific quality and outputs

The team is constituted of 3 medical practitioners and 7 engineers, most of them are part-time attached to the team. It is embedded in a Biostatistics Department that provides a mixture between high quality service platform for the IC (trial design, data management, analysis and reporting of clinical trials) and the research activity of the team.

The team has an excellent and longstanding publication track, 129 papers published from 2008 to June 2012. Due to the nature of the Curie Institute environment and the interaction with clinical investigators, most of the publications are in medium to high medical journals but very importantly there are also 30% of methodological publications, among which several first and last author publications in important journals such as Statistics in Medicine. There is a good balance between papers with lead positions and joint collaborative papers with other scientists.

The 3 main research axes for 2014-2018 “Methods for identification of optimal dose”, “Identification and integration of potential predictive factors of response to treatment in the early evaluation of treatments” and “Evaluation of surrogate endpoints” are a result of the close collaboration with clinicians and are up to date with the needs in the oncology field. The SHIVA trial is one of the first clinical trials ever to randomize targeted agents based on molecular profiling versus conventional therapy.

Assessment of the team's academic reputation and appeal

Some of the team members participate to an impressively large number of national scientific instances in oncology (at ANSM, ARC and INCA to name a few) and one has been instrumental in the French Society of Biometry. The team has participated in the organisation of several national events and is leading national (e.g. Phase I Inca Hub) and participating to international networks (ITCC) in oncology.

The team has been very successful in obtaining grants at the national level (PHRC, INCA etc).

The teams’ research is extremely well known at the French level, little less known at the international level, with the exception of the phase I methodological research work.

There is one postdoc student from Japan currently visiting the team.

Assessment of the team's interaction with the social, economic and cultural environment

The members of the team are very active in the INSERM project on long-term survival after cancer and consequences for insurance, loan and mortgage (French agreement AERAS), and the return to work of cancer survivors project.

Two members of the team are members of local ethics committees.

Through the clinical trial activities there are several collaborations with pharmaceutical companies.

Assessment of the team's involvement in training through research

The team has a very important implication in the master program of Public Health at University Paris 11 (Survival and clinical trials module), the STARC-CESAM program (DIU Paris 6, 7, Aix Marseille 2) and at the statistics school ENSAI in Rennes, and others, for a total of 200h a year.

There are currently 3 PhD students and every year one or two M2 students, and also a Public Health medical doctor intern.

Assessment of the five-year plan and strategy

The research topics are realistic and in response to real needs in the medical oncology field. The SHIVA trial is probably the most risk-taking due to the many methodological and logistic challenges of implementing next-generation sequencing or other molecular techniques in the design of clinical trials.

Conclusion

- Strengths and opportunities

It is important that the team has its own identity and research agenda, embedded within the unit. The 5-year plan provides an excellent opportunity to perform cutting-edge biostatistical research according to the methodological problems embedded within the 3 main projects. The SHIVA trial will provide a rich source for developments, and strengthen collaboration with the other teams in the unit. It opens many new perspectives for the team and for the whole unit.

The team's collaboration to the long-term follow-up of cancer survivors and their social integration is to be applauded.

The recent addition of two researchers further strengthens the team and broadens its scope.

- Weaknesses and threats

The team is still of relatively small size as compared to the other teams.

- Recommendations

The team could develop more international collaborations with methodological biostatistical units e.g. through application for European funding with external partners.

At least one additional HDR should to be obtained in biostatistics.

5 • Conduct of the visit

Visit dates:

Start: 17 January 2013 at 8:30 a.m.

End: 18 January2013 at 1:30 p.m.

Visit site:

Institution: Institut Curie

Address: Batiment L'homond, 12 rue L'homond, 75005 Paris

Conduct or programme of visit:

Day 1 – 17 / 01/ 2013

8:30 a.m.	Welcome and coffee
9:00 a.m.	Private preparatory meeting of the Committee
10:00 a.m.	Introduction by Curie Institute Director and MINES ParisTech Director (10 mn)
	General presentation of U900 (E. BARILLOT, 1h including 30 mn for questions)
11:10 a.m.	Break with posters
11:40 a.m.	Presentation of team 1 (45 mn including 25 mn for questions)
12:25 p.m.	Working lunch of the Committee with Institute Directors
1:45 p.m.	Presentation of team 2 (45 mn including 25 mn for questions)
2:30 p.m.	Presentation of team 3 (45 mn including 25 mn for questions)
3:15 p.m.	Break with posters
4:00 p.m.	Presentation of team 4 (45 mn including 25 mn for questions)
5:00 p.m.	Private meeting of the Committee
8:00 p.m.	End of Day 1

Day 2 - 18 / 01 / 2013

9:00 a.m.	Parallel meetings:
	with researchers and post-docs (without group leaders)
	with engineers and technicians
	with PhD students
10:00 a.m.	Break
10:30 a.m.	Visit of the bioinformatics platform: demo, posters
12:00 p.m.	Working lunch: Private final meeting of the Committee
1:30 p.m.	End of Day 2

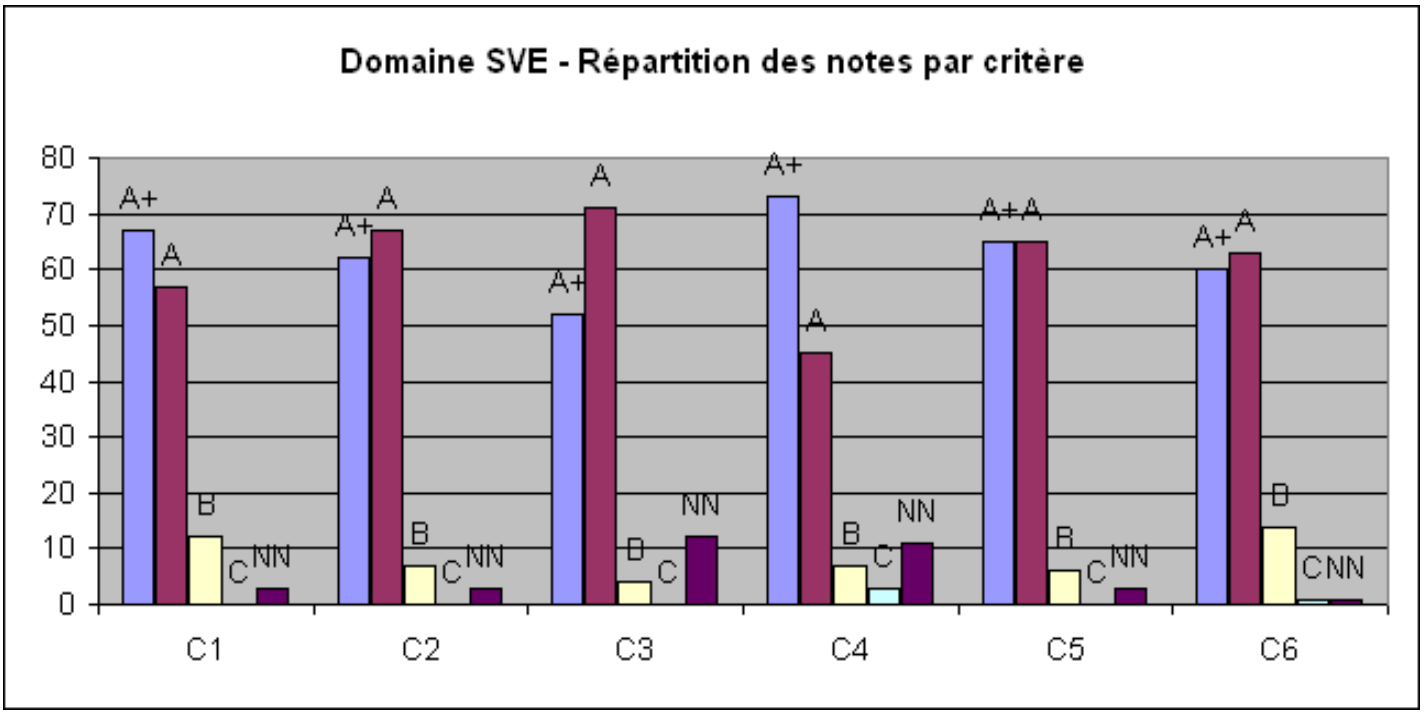
6 • Statistiques par domaine : SVE au 10/06/2013

Notes

Critères	C1 Qualité scientifique et production	C2 Rayonnement et attractivité académiques	C3 Relations avec l'environnement social, économique et culturel	C4 Organisation et vie de l'entité	C5 Implication dans la formation par la recherche	C6 Stratégie et projet à cinq ans
A+	67	62	52	73	65	60
A	57	67	71	45	65	63
B	12	7	4	7	6	14
C	0	0	0	3	0	1
Non Noté	3	3	12	11	3	1

Pourcentages

Critères	C1 Qualité scientifique et production	C2 Rayonnement et attractivité académiques	C3 Relations avec l'environnement social, économique et culturel	C4 Organisation et vie de l'entité	C5 Implication dans la formation par la recherche	C6 Stratégie et projet à cinq ans
A+	48%	45%	37%	53%	47%	43%
A	41%	48%	51%	32%	47%	45%
B	9%	5%	3%	5%	4%	10%
C	0%	0%	0%	2%	0%	1%
Non Noté	2%	2%	9%	8%	2%	1%



7 • Supervising bodies' general comments

CENTRE DE RECHERCHE

institutCurie
Ensemble, prenons le cancer de vitesse.

AERES
Section des Unités
20, rue Vivienne
75002 PARIS

Paris, le 19 avril 2013

Concerne : Rapport S2PUR140006131 - Cancer et Génome : Bioinformatique, Biostatistiques et Epidémiologie d'un Système Complexe - 0753172R
Directeur : Emmanuel Barillot

Chers Collègues,

En tant qu'organisme hébergeur et déposant unique des rapports des unités de recherche du site de Paris de l'Institut Curie – Vague D, je vous informe avoir bien reçu en date du 29 Mars 2013, le rapport d'évaluation de l'AERES sur l'unité indiquée en rubrique.

J'ai lu ce document avec attention et vous informe être d'accord avec les constats et les recommandations qui y sont formulés, que je m'emploierai à mettre en oeuvre. Nous tenons à saluer le travail réalisé par les experts.

Concernant la recommandation du comité de renforcer les collaborations entre équipes de l'unité, nous nous appuyons notamment sur des projets transversaux (comme l'essai clinique de médecine personnalisée SHIVA, ou le projet investissement d'avenir ABS4NGS) et les thématiques transversales (réparation de l'ADN, analyse de séquence ...).

Nous notons que l'équilibre trouvé entre projets méthodologiques et projets biologiques a été apprécié par le comité de visite, et nous nous attacherons à le préserver, comme recommandé. Enfin, les conseils portant sur le développement logiciel et la gestion des personnels temporaires seront intégrés dans la démarche d'amélioration de la qualité déjà en œuvre au sein de l'unité.

Afin d'assurer le succès continu de cette unité, j'ai bien noté les recommandations du comité pour appuyer ce plan en tenant compte de l'évolution de cette unité et tous les efforts seront faits en coordination avec les tutelles : l'Institut Curie, l'INSERM et notre partenaire Mines ParisTech, pour assurer les soutiens nécessaires.

26 rue d'Ulm - 75248 Paris Cedex 05
Tél. 33 (0)1 56 24 55 00 - www.curie.fr

Je tiens à exprimer tous mes remerciements aux membres du comité d'évaluation pour leurs commentaires et recommandations très pertinents qui sont basés sur un travail d'analyse approfondie. Je remercie également l'équipe de l'AERES qui a soutenu la mise en oeuvre de l'ensemble de cette évaluation.

Je vous prie d'accepter, Chers Collègues, mes plus cordiales salutations.

Daniel LOUVARD
Directeur de la Section de Recherche
INSTITUT CURIE

LE DIRECTEUR

AERES
Pierre GLAUDES
Directeur de la section
des unités de recherche
20 rue Vivienne
75002 PARIS

Paris, le 22 avril 2013

Objet: Rapport d'évaluation AERES **S2PUR140006131 - Cancer et Génome : Bioinformatique, Biostatistiques et Epidémiologie d'un Système Complexe - 0753172R**

Monsieur,

J'ai pris connaissance du rapport établi par les experts et les remercie pour la qualité de leur travail. Je me félicite de ses conclusions très positives qui soulignent la qualité des travaux menés dans cette unité et confirme notre avis sur la pertinence du partenariat mis en place entre MINES ParisTech, l'Instut Curie et l'INSERM.

MINES ParisTech continue de soutenir activement ce partenariat. Avec le recrutement récent de deux chercheurs, l'équipe 3 sera mieux armée pour développer de nouvelles interactions avec les autres équipes et faire face à la diversité des thématiques de son projet de recherche, soulignée par les experts. Elle pourra de plus s'appuyer sur de nombreuses synergies avec d'autres équipes de MINES ParisTech et de l'Institut Curie pour développer la thématique nouvelle de "bioimage analysis", qui ouvre de nombreuses opportunités scientifiques.

Je vous prie de croire, Monsieur, en l'assurance de mes sentiments les meilleurs.

Romain SOUBEYRAN
Directeur

CC : D. Goetz/J.P. Vert (MINES ParisTech)

www.mines-paristech.fr
www.mines-paristech.eu

MINES ParisTech
60, boulevard Saint-Michel - 75272 Paris Cedex 06 - France
romain.soubeyran@mines-paristech.fr
T. (+33) 1 40 51 90 14 / 94 43 - F. (+33) 1 1 40 51 90 25