

On the organization of bioinformatics core services in biology-based research institutes

With the growth of genomics, research institutes are increasingly confronted with the task of providing computational support to biologists and the provision of computational facilities, both for laboratory-based bioinformaticians and bioinformatics research groups. The optimal organization of bioinformatics support units to meet these needs is a common topic of discussion and consideration.

During the recent evaluation of a bioinformatics core facility, we ended up discussing—in general terms—what we consider basic guidelines for the organization of bioinformatics core units in large research institutes, based on the experience and developments in our own institutes.

We thought that part of our discussion could be useful for others, keeping in mind that every organization has specific needs and requirements, and that bioinformatics support is a fast moving area that requires continuous adaptation. In the following, we summarized what we consider some key recommendations:

- (1) Bioinformatics departments have to clearly separate their service units from their research laboratories. This contributes to the transparency of the organization, especially regarding the allocation of funds.
- (2) The organization of bioinformatics service unit should be clear with defined missions for each component. It is particularly important to separate tasks by topic. For example: support of database development and maintenance, statistical analysis of high-throughput data, automatic image acquisition and analysis of e.g. light microscopy data and analysis of next-generation sequencing data.
- (3) The unit should install a 'users committee' that includes users and specialists of the unit. This committee would be responsible for establishing clear priorities on what technical platforms are supported or aborted, and defining the rules applied to prioritize projects for bioinformatics support.
- (4) It is important that the services are provided in a highly transparent manner, including publicly accessible information on statistics of users, project progress, tools and use of resources.
- (5) Particularly important is the interface between the bioinformatics unit and the medical activities in an institute. This requires the integration of biobanking, medical information (records) and the corresponding genomic information. This activity is key for translational medicine and it requires efficient, multi-disciplinary interaction between clinical and computational experts. This process can be greatly enhanced by a seamless integration of the relevant data streams.
- (6) One of the key missions of the bioinformatics is to provide training to biologists at a basic level. It is important to also incorporate separate advanced training

in new technical developments for 'hybrid users' or for expert bioinformaticians. This should preferably be done in collaboration with the bioinformatics research laboratories in the institute.

- (7) It is useful to nominate a bioinformatics support person, whose task is to guide users in the use of public bioinformatics tools and databases as well as tools and methods developed within the institute. The goal is to facilitate the use of bioinformatics by the biologists using state of the art tools.

Finally, there is always the very difficult issue regarding the optimal size of the bioinformatics support units. This depends on many factors such as the diversity of the topics being covered in the institute and the level of the computational skills of the biologists. We will, therefore, not give an exact figure, but rather provide figures from our own institutions which are very comparable in this regard. In our own institutions, NKI-AVL, CNIO and FIMM, we have approximately 1 full-time, institutionally funded bioinformatician to support 100 scientists. In addition, the research groups themselves have up to 5–10 times this number of bioinformaticians working in their own research groups. This number varies, of course, depending on the specific topics researched by a particular group. These so-called 'embedded' bioinformatics researchers are involved in more research-oriented work and are largely funded by outside grants. In our opinion it is very important to develop structures that allow the 'embedded' bioinformaticians to meet and communicate.

In conclusion, we can state that we favor a model where a small core group of bioinformaticians provide transparently organized support on more general, institute-wide research problems, while the majority of the bioinformaticians are embedded in the research groups. Embedding ensures more direct interaction between the biologist and the bioinformatician, providing both researchers with a sense of ownership of the project. Not only does this elevate the skills level of both parties, but it also greatly enhances productivity.

Olli Kallioniemi

Institute for Molecular Medicine Finland FIMM
Nordic EMBL Partnership for Molecular Medicine
Biomedicum Helsinki, Finland

Lodewyk Wessels

Division of Molecular Biology
Netherlands Cancer Institute
1066 CX Amsterdam, The Netherlands

Alfonso Valencia

Structural Biology and BioComputing Programme
Spanish National Cancer Research Centre (CNIO)
Madrid, Spain