

**Institute for Integrative Genome Biology (IIGB)
University of California, Riverside**

Scientific Advisory Board Meeting

November 2, 2007

Executive Summary

The IIGB has established a very good-to-excellent set of core facilities, with state-of-the-art equipment in the areas of genomics, bioinformatics, microscopy and proteomics, creating an integrated enabling nucleus that stimulates and attracts innovative research and technology development at UCR. The core facilities address many of the needs of the UCR research community, and have been instrumental in facilitating the recruitment of young talented faculty. Students and postdoctoral fellows greatly benefit from having access to the various core facilities, and from the expertise of the respective Academic Coordinators and Academic Administrator. In general, these facilities provide a very solid foundation upon which additional services to the campus at large could be built. There is a need to expand the services in a few key areas, such as high-throughput next generation sequencing and the required bioinformatics support. In addition, there are concerns about the ability of sustaining these facilities in the future. The momentum established by the IIGB facilities should facilitate future growth in biological and medical research at UCR, but it will require continued support and expansion of resources.

Overview

The first Scientific Advisory Board (SAB) was established on 2007 with the role of evaluating the IIGB Core Suite of Instrumentation Facilities at Keen Hall and providing advice and scientific expertise to the IIGB Director. The IIGB SAB had the opportunity to tour the core facilities of the IIGB (Microscopy, Genomics, Bioinformatics, and Proteomics) on November 2, 2007. Each SAB Member had also previously received a booklet summarizing core services, personnel, training, revenues allocations, etc., over the past two fiscal years, to provide more statistical background information for this evaluation. The Core Academic Coordinators and Academic Administrator provided an overview of the instrumentation and capabilities that each of the facilities has available, and explained how the facilities operate. The SAB also met with student-users of the facilities, and listened to presentations by a number of UCR faculty, who described how their research has been enabled by the core units.

The overall impression formed by the SAB is that the IIGB has established a very good-to-excellent set of core facilities, with state-of-the-art equipment in many areas, which are directed by highly qualified scientists who are pro-active in reaching out and helping the users. This combination of instrumentation and personnel has enabled the IIGB to effectively address many of the needs of the UCR research community. Numerous

research programs on campus are now heavily dependent on having access to the technologies and instruments provided by the facilities. However, the SAB also noted the need to expand the services in a few key areas, such as high-throughput next generation sequencing and the required bioinformatics support. In particular, the SAB feels that IIGB provides a very solid foundation upon which additional services to the campus at large could be built.

It became obvious during the faculty presentations that the IIGB core units have been instrumental in facilitating the recruitment of top-notch young faculty members (examples include Venu Reddy, Sean Cutler, Monica Carson, and David Lo, among others), and that the research programs of these new recruits are critically dependent on the core facilities. The IIGB should continue to be a prime incentive to recruit world-class researchers, as UCR expands programs in planned growth areas, such as stem cell research and medical-related initiatives.

Students and postdoctoral fellows greatly benefit from having access to the various core facilities, and from the expertise of the respective coordinators. It is the impression of the SAB that student training has been significantly enhanced by access to state-of-the-art equipment to carry out their individual research projects. In particular, the core units are facilitating the recruitment of students for the Integrative Graduate Education and Research Trainee (IGERT) Program and, conversely, the IGERT program is making available students of an ideal profile for the research that is facilitated by the core units.

The staff of the facilities is very dedicated. When key individuals have left, the Academic Administrator has stepped in to carry out technical functions, and as a result has compromised his own research or other activities. The SAB feels that additional FTEs are urgently needed to relieve this situation. In addition, bioinformatics is a critical need for each of the cores. Each core unit would benefit from having a dedicated bioinformatician attached to it, focused on the needs of that specific unit, even though these persons may be under the umbrella of the Bioinformatics core.

The SAB perceived serious limitations in the ability to sustain these cores in the future, given the new campus' policy of substantially reducing the carry-forward of funds, which will in turn jeopardize the investigators' ability to access the instruments. Currently, key instrumentation has been maintained by carry-over funds. With the new caps on carry-forward funds, there is an unacceptably high likelihood that normal failure of key parts, such as lasers and control electronics, will place one or more of these instruments out of commission due to a lack of funds to replace those parts. As these are established and not new instruments, it is not economically feasible to place all of these instruments under service contracts, and an alternative maintenance plan is needed. In addition, two important areas that may be compromised by the limitations to carry-forward funds are: (1) a robust seminar series, and (2) the research and training awards (which are an outstanding idea, to enable innovation with a relatively small amount of money).

It should be noted that the cost recovery system that is in place has been a significant factor in encouraging new investigators to access and effectively use the research

facilities, which in turn has resulted in these new investigators bringing a much more substantial stream of research funds to the University (estimated to be at least a 4X increase in funds from research grants since the inception of the IIGB cores).

The new IIGB Director has excelled at effectively using the Institute's resources and at allocating them to the critical areas.

Bioinformatics

Current Status: The Bioinformatics Core is led by Dr. Thomas Girke, who also holds an Assistant Professor appointment. His background in biology and knowledge of bio/informatics have allowed him to establish an excellent core facility that has managed to provide IIGB researchers with a comprehensive open-source, open-access software infrastructure. Dr. Girke has been very active in introducing and training students, postdoctoral fellows, and researchers in the use of that infrastructure through various tutorials, workshops, and manuals. In addition, he has developed research databases that have been extensively used by other faculty on campus, as the committee members had the opportunity to observe during the faculty presentations. The core is currently well-equipped in terms of hardware.

Future Prospects: Despite its excellence, some areas of concern for the future of the bioinformatics core could be detected:

- 1) It is apparent that the effectiveness of the facility is limited by its personnel constraints (the facility currently consists of Dr. Girke and a Systems Administrator; additional undergraduate and graduate students work on the research and database projects of the bioinformatics lab). The facility is not capable of effectively fulfilling all the bioinformatics needs of the research community that it sustains. Whereas it should not be expected that the facility addresses all the IT and bioinformatics needs of all the IIGB-associated groups, it is clear that it would greatly benefit from having additional staff members at its disposal.
- 2) The current recharging system recovers only a very small fraction of the facility's operational cost – yet the SAB feels that it is crucial, for the IIGB mission to be fulfilled, that the Bioinformatics Core be expanded. Stable sources of funds for additional staff personnel are urgently needed.
- 3) An increased integration between the Bioinformatics Core and the other cores (Genomics, Microscopy, Proteomics) would be highly beneficial. This would facilitate the implementation in the other cores of more sophisticated data back-up systems, as well as increased capabilities for data QC and first pass analysis. It is the opinion of the SAB that this enhanced integration and coordination would be required if additional instruments (such as a high-throughput sequencer) are acquired and hosted at the IIGB. For example, a minimal custom LIMS, shared between Genomics and Bioinformatics, would be highly desirable in such a scenario.

4) The acquisition of a high-throughput sequencer (which this SAB recommends, see below) would put a heavy tax on the Bioinformatics core, and it will not be possible to successfully bring on-line such an instrument without additional personnel resources in the facility.

5) The core has excelled at using open-source software resources, and this is an ideal solution that fosters student training and enables researchers to be more independent. However, open-source software can be more taxing in terms of user training and support. It is possible that consideration of an increased use of commercial software in specific instances, as a way to alleviate the current general heavy taxing on the core personnel's time, would be appropriate.

Genomics

Current Status: The Genomics core facility provides a wide variety of state-of-the-art services including FACs sorting, DNA sequencing, qPCR, DNA fragment analysis, clone arraying services, DNA mini-prepping and also provides access to instrumentation for both DNA and RNA quantification. In addition, it provides a variety of state-of-the-art microarray services that include both full service Affymetrix chip and tiling array hybridization/wash/scanning.

Of all the IIGB facilities, the Genomics core is the most heavily utilized by the campus faculty; virtually all departments associated with the Institute have utilized this facility. Both the number of departments and percentage of time they have accessed the Genomics core services have grown each year. Genomics core services are also utilized on a fee for service basis by outside institutions, thus maximizing the usage of the equipment to both inside and outside investigators, and leveraging the UCR investment by recovering a significant level of revenues. One of the impressive aspects of Genomics and the other Cores is the extremely modest cost of services provided to UCR faculty. The overall costs of DNA sequencing, microarrays and FACs services as well as training (free) is very reasonable relative to other institutions. In the fast moving area of genome biology, it is fully appropriate to fully reinvest these revenues (carry forward) for the purchase of new state-of-the-art equipment or to add additional services, either outright or to provide matching funds for government equipment grant proposals.

Future Prospects: Under the Academic Administrator's leadership, the Genomics core has been proactive in identifying potential future needs of the campus Faculty in the area of genomic services. One example is the fast evolving area of next generation DNA sequencing technology. The Genomics core has organized seminars/presentations by the three major next generation DNA sequencing instrumentation companies (Roche/454, Illumina/Solexa and ABI-Solid; an approximate cost for the instrument is \$500,000). Evaluation of a possible purchase of one or more of these platforms is in progress. Based on presentations of several of the UCR faculty genomics core users, and from the perspective of maintaining supporting UCR faculty working on highly competitive areas such as epigenetics/genomics, the SAB recommends a rapid decision be made on the purchase of the most appropriate next generation DNA sequencing system. This may require involvement by the University due to the typically long cycle time of granting

agencies and fierce competition in this particular area. Overall the Genomics core facility was viewed very positively by the SAB as providing a high quality and cost effective service for a large and very growing fraction of UCR laboratories.

Microscopy

Current Status: The optical microscopy core is excellent, consisting of several advanced instruments that support a wide range of experiments. The facility benefits greatly from its Academic Coordinator, Dr. David Carter, who not only has the expertise to maintain these instruments and to train users, but also has the necessary training to understand the important problems in biology that microscopy is required to address. He has real knowledge of optics and he has professional experience in building instrumentation. The benefits of the Academic Coordinator's suite of expertise and active engagement in the research being conducted in the facility are manifest in such projects as the medium-throughput screening of mutants and responses to small molecules, where the director has been an active collaborator and developed customized solutions to handle biological specimens and to automate aspects of these screens.

Future Prospects: While having a solid foundation, there are a few areas where the facility could be developed further in the future:

1) The demand for the Leica SP2 microscope is starting to exceed instrument availability. Some of this demand might be picked up by the other instruments, such as the Zeiss 510, but the SP2 is preferred by users because of greater ease of use and higher sensitivity. At the same time, another research group on campus, focused on stem cells, is planning the purchase of a new confocal instrument and will inevitably rely on Dr. Carter to help maintain the instrument. It seems that a reasonable solution, and a more effective use of university resources, would be to determine if a new Leica SP5 system could serve the needs of the stem cell group and then to consider housing the new instrument in the IIGB core. This could help to alleviate the over-booking of the SP2 and also allow for more effective maintenance and training on the new instrument, as it would be placed under immediate supervision of Dr. Carter.

2) Users have expressed an interest in FLIM (Fluorescence Lifetime Imaging) and multi-photon imaging. Multi-photon imaging has superior ability to image deeply into thick tissue due to the longer wavelength of the excitation energy. It also excites a much smaller volume of the cell and thus can reduce photobleaching and phototoxicity, which is particularly important for live cell imaging. This imaging modality is particularly popular in thick animal tissues such as brain slices where tissue penetration is critical and photosensitivity is high. The Zeiss 510 in the facility can be upgraded to multi-photon imaging. This is an expensive upgrade primarily due to the approximately \$200,000 tunable femto- or pico-second laser that is needed. There are also set-up costs and the addition of an optional non-descan detector, which is highly recommended to increase instrument sensitivity. Communications of SAB member David Ehrhardt with facilities managers at Stanford and at the Max Plank Institute indicate that the Zeiss multi-photon system typically requires significant levels of maintenance and attention, so there are also increased administrative costs to consider.

FLIM exploits the fact that fluorescence emission carries information not only in frequency (color) and intensity, but also in time. Fluorescence lifetime can be used to probe changes in the environment of a chromophore, including the proximity of other chromophores that can act as acceptors for the transfer of fluorescence excitation. FLIM can be a superior method for probing molecular proximity, as it is less affected by certain important experimental variables such as chromophore concentration. FLIM comes in two basic flavors, frequency domain and time domain. Time domain FLIM requires pricey pulsed lasers and specialized detection electronics. It is typically set up on point scanning confocal systems such as the Zeiss 510. The same lasers that are used for multi-photon excitation can also be used for FLIM, so it is possible to upgrade to multi-photon imaging modality when upgrading to FLIM. Time domain FLIM is a major upgrade as it requires the costs for basic multi-photon imaging plus costs for the detection hardware and software.

Frequency domain FLIM also requires a device to modulate the laser source and specialized detector electronics, but both of these components are significantly cheaper. This type of FLIM detection can be added to a spinning disk confocal or to a conventional epifluorescence microscope. Time domain FLIM has theoretically superior sensitivity to low levels of signal, but requires long integration times to collect an image, sometimes several minutes. Frequency domain FLIM is capable of building a usable FLIM image more rapidly, if signal intensity is adequate. The cost of adding frequency domain FLIM to an existing spinning disk system is approximately \$120,000 (3i, Boulder).

Before investing in either multi-photon microscopy or FLIM, it is recommended that a user survey be performed to evaluate demand and to determine the types of experiments that would be performed.

3) The facility is primarily an optical microscope facility, but it recently acquired a benchtop environmental SEM (Hitachi TM 1000). This was an excellent addition, but currently, there is no TEM for high-resolution ultrastructural studies in the facility. It is likely that correlated light and electron microscopy will become an increasingly popular research tool and new technologies are being developed to bring molecular specificity to TEM imaging. As life science research at the University grows, it would be wise to plan on consolidating EM imaging at the IIGB core to facilitate the integration of light microscopy and high-resolution electron microscopy.

4) As imaging technology becomes increasing sophisticated, experiments are often less limited by the ability to acquire raw data as they are by the ability to process these datasets to extract the desired information. The current core facility has excellent image acquisition tools and the start of a good collection of image analysis software, but is short on computational resources, and more importantly, the computational/bioinformatics personnel to help users apply sophisticated image analysis methods and to develop customized analysis tools.

5) The facility Academic Coordinator is doing a good job keeping up with instrument maintenance, user training, and billing, but the facility is heavily used and administrative demands are starting to limit the time the Academic Coordinator has to apply his expertise and creativity to develop the facility and to help users with their experiments.

Proteomics

Current Status: The Proteomics Core is capably led by Dr. Songqin Pan. His expertise in proteomics and bioanalytical mass spectrometry and his leadership have built a solid foundation for the Core to grow in future years. The Core offers capabilities for sensitive protein identification, proteome profiling, protein post-translational modifications, and quantitative protein measurements with high quality tools. Two quadrupole time-of-flight mass spectrometers are used for protein analysis; one is dedicated for on-line LC-tandem mass spectrometry (MS/MS), and the other is dedicated for open access, user-operated MALDI-MS/MS measurements. A Biacore instrument has been purchased to add capabilities for protein-protein and protein-ligand interaction measurements. The Core's interactions with many excellent collaborators have resulted in several high quality publications in high impact journals.

Most of the Core's direct activities utilize the LC-MS/MS system, as this methodology best addresses projects requiring protein identification and proteome profiling. Utilization of LC-MS/MS is near capacity for this single platform. The MALDI-QTOF (MS/MS) instrument, although used from day to day and greatly appreciated by its current users as a user-based platform, is under utilized in terms of machine hours and has capacity for future growth. In addition, Dr. Pan is essentially the only person in the Core, although he has aid from some students and postdoctoral fellows from research laboratories. Dr. Pan is responsible for all Core day-to-day activities, including sample preparation, instrument operation, training, workshops, instrument maintenance, data processing and archiving, consulting with PIs, developing experimental strategies, and reporting.

The Chemistry department offers services and capabilities for analytical mass spectrometry, as discussed by Drs. Larive and Wang (Chemistry) at the SAB meeting. There is a mutual understanding between the two mass spectrometry laboratories that Chemistry focuses exclusively on small molecular analysis and the IIGB Proteomics Core focus is on peptide and protein analysis. Together, they can network effectively to plan for future growth in mass spectrometry services for the greater UCR community.

Future Prospects: With current resources available, including personnel, recharge income, and instrumentation utilization, a slight growth in MALDI-QTOF utilization could be anticipated. However, with limited additional personnel, the overall Proteomics Core is not expected to grow significantly. Projects requiring extensive protein identification from whole organisms, organelles, network-wide post-translational modifications, and pathway comparison/analysis that require "shotgun" proteomics strategies could not be served effectively without a significant negative impact to other smaller, more targeted projects, unless large multi-PI grants are obtained. These more complex projects would also require more bioinformatics support, which is not available

to the Core currently. It is difficult for the Core to be proactive to develop new methodologies that could be valuable for a number of projects (e.g., phosphopeptide enrichment strategies) without additional personnel. Dr. Pan only has the time to address the current projects and workload. He does not have the time to expand the capabilities of the Core by bringing in new, cutting-edge techniques. It is unlikely that many additional projects that might be more medically-relevant could be effectively addressed by the Core without additional resource allocation.

1) Dr. Pan and the Core would be well-served by the addition of junior personnel. This would allow the Core to expand with new techniques and tools that would be attractive to many researchers, including those involved in medically-relevant work.

2) Resources for bioinformatics support should be added to the Core. Routine data archiving and more sophisticated protein identification processing (e.g., de novo sequencing, protein quantification, etc) should be a relatively high priority for the Core.

3) Plans for adding a new mass spectrometer for LC-MS/MS should be developed, as the Core is near capacity for this service currently. Agencies that fund shared instrumentation grants (e.g., NIH, NSF) should be targeted, in addition to other resources that might be available (e.g., foundations, private donors, etc). Instruments to target include a high duty-cycle linear ion trap, another QTOF, or perhaps a high resolution OrbiTrap. Perhaps a more creative addition would be a MALDI-TOF/TOF. This would free the current MALDI-QTOF for its conversion to an LC-QTOF-MS/MS system. An alternative strategy to consider would be to add LC-MS/MS capabilities to the existing MALDI-QSTAR system. MALDI-QSTAR experiments would need to be scheduled, for example, every first week of the month (as switching between LC-MS and MALDI requires time to adjust). This option could be considered only if the service contract for the QSTAR is continued. (Changing sources on this system strains the current turbomolecular pumps, and often requires frequent replacement.) Additional personnel would be required to support an additional instrument.

4) The addition of new protein separation technologies is highly recommended. Current proteomics strategies often employ more extensive protein fractionation prior to the analytical LC-MS/MS measurements to allow deeper proteome profiling (i.e., identification of more proteins from complex mixtures). A PF-2D chromatofocusing/LC system is being considered by the Core. This is not recommended because it would greatly tax the existing resources. Although the PF-2D system has been demonstrated by some laboratories to greatly enhance protein identification, it results in greatly increasing the complexity of the overall workload. A single sample would be fractionated into several hundred fractions, each requiring a separate LC-MS/MS measurement. A more economical and perhaps equally effective means for fractionation would include the Zoom-IEF fractionation system (Invitrogen) and the Off-Gel system (Agilent Technologies). Both are under \$15K and both offer capabilities for fractionating complex mixtures into 6-24 fractions.

5) The Proteomics Core (in addition to the other Cores of the IIGB) should have regular contact with internal (at UCR) and/or external advisors that have expertise in the relevant area. For example, Professor Yinsheng Wang (UCR Chemistry) would be an obvious choice to discuss issues regarding mass spectrometry and protein analysis, and planning for future expansion of capabilities.

Overall Recommendations

Sustaining and expanding the state-of-the-art Core Facilities will facilitate the hiring of additional excellent investigators working on model organisms relevant to medicine. These additional faculty hirings will in turn further leverage the existing infrastructure, thus allowing a broader research mission and increasing grant research funds to the university. The following summarizes our particular recommendations:

- Allocation of additional FTE's to support the ever increasing bioinformatics requirements of each of the cores (two FTE's, half per core).
- The cores would benefit from additional FTE's: ½ FTE in microscopy, 1 FTE in genomics and microarray, 1 FTE in proteomics, 1-2 in bioinformatics.
- A plan for maintaining key instrumentation not under service contract is critically needed. An ability to increase the carry-forward funds, combined with creation of a capital reserve for instrument repairs and purchases, would allow these critical instruments to be maintained as well as for a more effective use of funds by the IIGB facilities.
- The availability of annual funds for maintaining a robust seminar program across the IIGB and continuation of research and training awards that emphasize interdisciplinary approaches is vital for faculty and students and for IIGB's ability to continue to serve as a flagship for facilitating biological and medical sciences at UCR. It also promotes continued future use of the facilities. The seminars by world-class researchers provide additional training and it helps promote the overall visibility of the IIGB to the research community.
- Given the requirement by the research of many of the junior and senior key faculty members of a next-generation sequencer, and in order to maintain IIGB's international status on small RNA, epigenomics, etc., having availability to one of those instruments on campus (with the required personnel to support it, both technically and bioinformatically) is an obvious urgent need.
- Access to the facilities should be dependent upon annual summary of what the facilities allowed the investigators (grants, papers, talks, posters, abstracts). It is suggested that the IIGB creates a "wall-of-fame" where publications related to the use of the cores are displayed prominently.

- A regularly scheduled survey (e.g., annual) should be distributed to users of the Cores to gauge the quality of the services provided and to solicit input for improvement and expansion.

Scientific Advisory Board Members

Joseph R. Ecker – The Salk Institute

David W. Ehrhardt – Carnegie Institution of Science, Stanford University

Joseph A. Loo – University of California, Los Angeles

José Luis Riechmann – California Institute of Technology