

# **University of Hawaii**

## **Center for Genomics, Proteomics, and Bioinformatics Research Initiative (CGBPRI)**

### **Annual Report 2004**

#### **Executive Summary**

In 2001 Chancellor Peter Englert initiated the formation of a Center for Genomics, Proteomics, and Bioinformatics at the University of Hawaii. The long-term goal of the Initiative is to become a center of excellence and a nucleus of multidisciplinary research efforts. During the initial planning stages, researchers with an interest in these areas of endeavor were polled, in order to identify areas of existing expertise and specific demands.

Existing administrative and scientific interactions with the Maui High Performance Computing Center (MHPCC) facilitated the creation of a collaborative alliance centered around computational proteomics. MHPCC will be the center for bioinformatics and will continue to build the computational infrastructure through funding from the USAMRMC/TATRC. The bioinformatics infrastructure has been expanded to include a high-performance server on the UH campus, sequence databases, and web-based analytical software to permit researchers and educators to perform high-speed sequence analyses, classroom demonstrations, and molecular simulations.

With initial funding from the Chancellors' office the infrastructure for a high-throughput genomic analysis facility was created. Today, this facility has the capacity to sequence ~ 6.7 million bases per day at a cost of \$ 2.50 per sequencing reaction for UH faculty.

Because of the high cost of analytical instrumentation for the high-throughput analysis of proteins, a research alliance with the Windber Research Institute, PA was formed, opening the door for high-throughput proteomics access for UH faculty.

Two main projects have been initiated by the Center Initiative to date. One concerns the analysis of extremophile microbial genomes, such as from organisms found near undersea volcanoes; the other involves the papaya genome sequence. A sequencing consortium spearheaded by CGPBRI was formed in 2004, involving scientists from Federal and State research organizations, as well as Nankai University, China. The completion of this project will not only represents a scientific milestone, but will further solidify the status of the University of Hawaii as an important resource for an emerging biotechnology industry.

The Center initiative has already assisted 5 UH faculty from a number of Departments with projects involving Bioinformatics and Genomics and has supported 5 grant proposals involving proteomics analyses. 42 UH faculty were involved in the beta testing of the Center's Bioinformatic website tools. Additional research collaborations have been established with Pacific Telehealth & Technology Hui, Hawaii Biotech, and IBM.

The Center Initiative also represents a training and educational resource center to contribute to the contemporary scientific education and training for students, researchers, and the community. Three undergraduate and graduate courses in Genomics and Bioinformatics are being offered by CGPBRI. The Center is collaborating with 2 newly hired bioinformatic faculty from ICS department.

Funding received by CGPBRI from the University has been matched by RCUH, Federal funds, NSF grants, and State of Hawaii funding has been applied for. As a result of this initial investment from RCUH and UH; the UH CGPBRI received a \$2.5M award for its Computational Proteomics Program from the US Army Medical Research and Materials Command (USAMRMC)/ TATRC in December 2004. Fiscal year 2005 funding for \$2.2M has also been committed. The UH CGPBRI has also received formal notification of forthcoming NSF award of \$420K. This combined funding support of \$5.12M represents a more than 3-fold return of UH/RCUH's initial investment of \$1.425M.

## **Introduction**

### **Genomics, Proteomics and Bioinformatics – Transcending the traditional boundaries of the life sciences**

Contemporary biological research has led to the rapid development of new areas of inquiry, including functional genomics, proteomics, and bioinformatics (GPB). Genomics and proteomics research are creating enormous amounts of quantitative laboratory data, which can only be organized and understood with high performance computing tools and resources through bioinformatics, the study of the inherent structure of biological information. The University of Hawaii's aspiration to become a leader in the life and medical sciences requires the creation of both the intellectual and technological infrastructure to support these endeavors and to become the nucleus for interdisciplinary collaboration.

The potential outcomes and ramifications of GPB research for economic growth and the improvement of quality of life cannot be underestimated. The rapid development of automated nucleotide sequencing technologies has enabled the analysis of entire genomes of agriculturally important organisms and some of their pests and has laid down the foundation to improve the health, yields, and nutritional value of crops. By the same token, large numbers of tissues and organismic samples can now be rapidly analyzed in a quantitative fashion, giving ecologists and population biologists new valuable tools relevant to biological conservation. The same methodology enables epidemiologists to quickly identify and to track pathogens. Genome and proteome databases form the basis of rapid protein identification from small peptide fragments by mass spectroscopy, supported by powerful computing tools and software development

These technologies, which have experienced exponential increases in growth in the past decade, also pose a formidable organizational challenge. The three interrelated technologies of genomics, proteomics, and bioinformatics must be amalgamated to develop into a coherent world-class resource that will allow UH researchers to compete nationally and internationally. Excellence in these fields is critical to the future of UH as an intellectual and technological resource for agriculture and a growing biotechnology industry in Hawai'i. The agricultural, life and medical sciences are fundamental to the education and training of Hawaii's students and are vital to the economy of the State of Hawaii, as compellingly outlined in a report commissioned by the Hawaii Department of Business, Economic Development, and Tourism (DBEDT) and the office of the Governor in 1999 "Biotechnology in Hawaii: A Blueprint for Growth" (PMP Public Affairs Consulting, Inc., Chapel Hill, NC).

This GPB revolution has also led to dramatic changes in the sociology and culture of biological research, as individual researchers find it essential to develop collaborative networks that transcend the arbitrary boundaries between traditional academic disciplines to solve organizationally and computationally demanding problems. Increasingly, they require integrated knowledge of entire biological systems rather than highly specialized expertise and projects demand problem solving skills guided by objectives in addition to

hypothesis-driven experimental approaches. The silo effect of vertical integration is no longer effective in projects of the anticipated magnitude, and an interdisciplinary approach will become the conduit of future breakthrough research.

In 2002, Chancellor Peter Englert charged a committee chaired by the Dean of Natural Sciences, Professor Chuck Hayes, to create a position paper on genomics, proteomics, and bioinformatics for the University of Hawaii. The committee identified a significant interest in these areas by numerous faculty from diverse departments and support for the formation of a central core facility to gain access to these technologies. Subsequently, an initiative to create a Center for Genomics, Proteomics and Bioinformatics Research (CGPBRI) was formed.

## **Mission and Objectives**

The Center for Genomics, Proteomics, and Bioinformatics Research Initiative (CGPBRI) serves the University of Hawai'i research community, medical, agricultural, and other life sciences research institutes and the community of the State of Hawaii, thereby providing leadership and support for the University of Hawaii's broader mission and goals.

- The aim of the Center Initiative is to provide intellectual and technological infrastructure and facilitate access to advanced high throughput genomics, proteomics, and bioinformatic technologies.
- The Initiative will assist researchers with the genomics, bioinformatics and proteomics research. The Center Initiative will seek extramural support to achieve financial sustainability.
- The long-term goal of the Initiative is to become a center of excellence and to develop a nucleus of multidisciplinary research efforts; thereby attracting top-tier researchers to the University of Hawai'i.
- The Center Initiative also represents a training and educational resource center to contribute to the contemporary scientific education and training for students, researchers, and the community.

## **Objectives**

Our objective is to create a UH Center for Genomics, Bioinformatics, and Proteomics to bring together multidisciplinary life and medical science teams to develop and apply high throughput genomic/proteomic-scale technologies with bioinformatic tools.

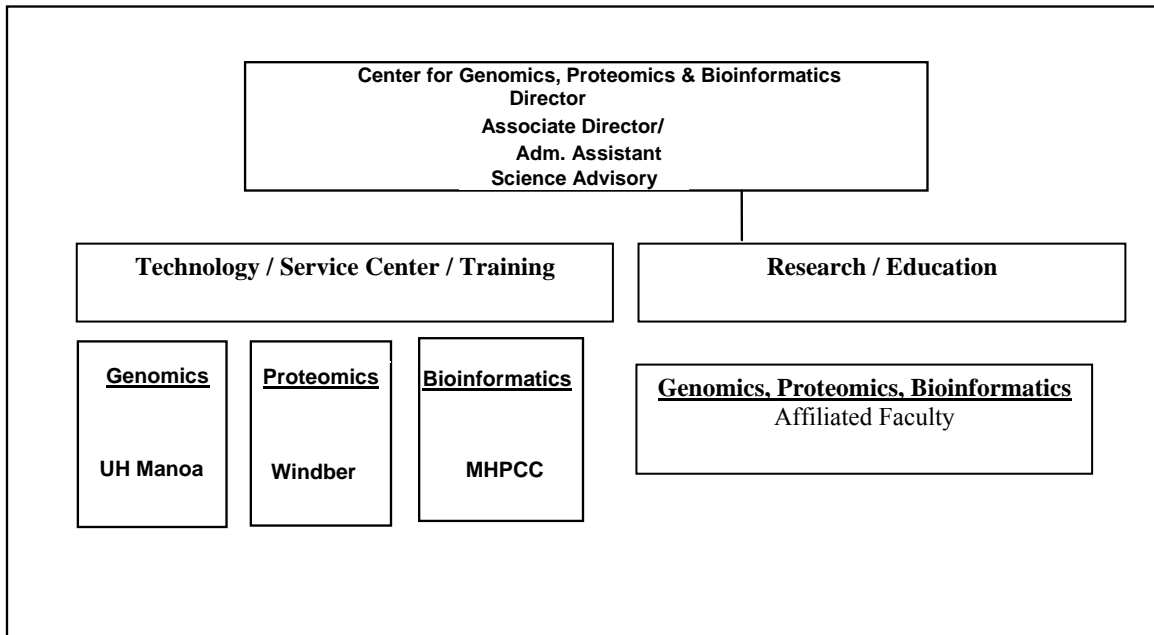
The current approach utilizes existing strengths and expertise in these areas, whilst taking a focused and sequential approach to building the necessary infrastructure.

**Develop technology and enhance access to advanced technologies in genomics core facilities** including high throughput sequencing, robotics facilities to spot clone libraries, high throughput PCR, gene synthesizers, and related technologies.

**Provide or enhance access to advanced technologies in proteomics Core facilities:** Several proteomic-related initiatives already exist in different UH Manoa units. However, no facilities exist relating to high throughput and high-resolution mass spectroscopy on campus. In the initial phase of the development of the CGPBRI, proteomics projects will be performed in collaboration with Windber Research Institute (WRI), PA., under a MOU signed in early 2004.

**Promote computational biology capabilities for related genomics/proteomic technologies and provide access to high performance computational facilities:** The University of Hawaii took over the management of the Maui High Performance Computing Center (MHPCC) in October 2001. MHPCC has since become the major site for high throughput bioinformatics, including genome assembly, annotation, and 3D modeling. Major DNA and protein databases, as well as various open source and commercial programs have been installed and are accessible to UH faculty and students.

## Organizational Structure and Staff



### **Interim Director** - Maqsudul Alam, Professor of Microbiology

Dr. Maqsudul Alam received his Ph. D. in Microbiology from the Moscow State University, Russia, and a Ph. D. in Biochemistry from the Max-Planck-Institute Martinsried in Germany. He was the recipient of the prestigious Shannon Award from the National Institutes of Health in 1996 and the UH Regents' Medal for Excellence in Research in 2002. His research team is actively involved in extremophile genomics and bioinformatics research at the Maui High Performance Computing Center. Recently his research team discovered human myoglobin-like proteins in extremophile microbes, the putative ancestral prototypes of hemoglobins. The National Science Foundation (NSF) and the National Institutes of Health (NIH) fund his research projects.

### **Associate Director, Administration and Chief of Staff** - Larry Sakima

As the budget and the complexity of fiscal matters of CGPBRI increased, the need for the expertise of a senior administrator with relevant experience became more pressing. Larry Sakima joined the Center Initiative in late 2004 and brought with him the experience of the long-time Assistant Director of Institute for Astronomy, and various senior finance management positions at UH and the Maui High Performance Computing Center.

**Associate Director, External Affairs** - Tak Sugimura, Ph. D. Maui High Performance Computing Center

Dr. Sugimura is a graduate of Columbia University (1962) and the University of California, Los Angeles (MS 1964, PhD 1967) and has focused on research in computational physics. Before joining the University of Hawaii managed Maui High Performance Computing Center (MHPCC) as Business Development Manager in 2001, he was the Director of the Maui Research and Technology Center where he developed and directed the technology incubation facility and broadband communications and telemedicine projects. He is currently Adjunct Professor of Information & Computer Sciences, University of Hawaii Manoa, Co-chair for the Digital Medical Arts Committee for the Hawaii Life Sciences Consortium, advisory board member of the National Center of Excellence for High Performance Computing Technology (Maui Community College) and member of the Kula Elementary School Community Council.

### **Associate Director – Research Coordination**

Heinz Gert de Couet, Ph.D. Professor of Zoology

Dr. de Couet is a graduate of the Technische Hochschule Darmstadt, Germany. Before joining the University of Hawaii in 1990 he was a Research Fellow at the Australian National University in Canberra and a lecturer in genetics at the University of New South Wales in Sydney. Trained as a cell biologist, de Couet first participated in a small-scale functional genomics project at the Australian National University. Research focus in cell biology, development and genome evolution, using a variety of organisms as models. The responsibility of the associate director for research is the liaison between UH faculty and both private and state-operated research institutions to facilitate access to high-throughput genomics. Other functions include the management of the CGPBRI equipment facilities on the UH campus and the creation of a focus group of researchers interested in genomics, proteomics, and bioinformatics. HG de Couet is a former member of the UH faculty senate.

## **High throughput Genome sequencing Facility**

Snyder Hall 312  
Phone (808) 956-8039

### **Genome Leader**

Shaobin Hou, Ph.D., Microbiologist

Dr. Hou received a BS degree in Microbiology from Nankai University, China. He received both his MS (1996) and Ph.D. (2000) in Microbiology from the University of Hawaii. He subsequently began to work as a Junior Researcher in Dr. Alam's laboratory. His research has focused on signal transduction of *Archaea* and *Bacteria*, and expression/purification of recombinant proteins. In 2000 he became the team leader for

microbial genome sequencing projects and is the lead scientists for the shotgun papaya genome project conducted by CGPBRI. He is lead author on several papers on the genome of extremophile microorganisms and co-discoverer of the first heme-based aerotaxis transducer in Archaea.

## **Office**

Keller Hall 319  
Phone 956-0924  
Fax 956-0925

### **Administrative Officer**

Margy Yuen

### **Webmaster CGPBRI website and Software Manager**

Alex Yu

Web programmer and designer for CGPBRI who maintains the website at the software end. He received his M.S in Microbiology and his B.S. in Computer Science. He handles the day-to-day operations of the site, including site layout, tool coding and optimization, database, and back-end support.

# Projects and Accomplishments

## Core Programs and Services

The Center Initiative core facility presently has six ABI 3730XL sequencers at its disposal. Our current throughput capacity approximates 2000-2400 samples per day, which we intend to boost to a total of 9,000 sequence runs, pending additional student help hires. With an average read length of around 800bp the daily sequence output will ultimately be in the order of 6.75 Mbp per day. Assuming a pass rate of 90%, the actual sequence output will be greater than 6 Mb. Below is a list of major equipment acquired for CGPBRI.

No.	Instrument	Manufacturer/Vendor
6	96 Cap DNA Analyzer	ABI 3730XL
1	DNA Shearing Device	Genemachines HYDROSHEAR
1	Pulsed Field Electrophoresis system	Biorad Chef-DRIII
1	Biomed FX-96	Beckman Coulter
1	Centrifuge	VWR International EPP581OR
1	Freezer	Revco ULT5030A18
1	Refrigerator	Revco REL5004A20
1	Hydra 96 PP	Art Robbins Instruments
1	Upgraded Hydra 384UG	Art Robbins Instruments
3	Thermal Cycler (dual 96-well)	ABI GENEAMP PCR 9700
5	Thermal Cycler (dual 384-well)	ABI GeneAmp PCR System 9700

## COMPUTING

The Maui High Performance Computing Center (MHPCC), established in September 1993, is an Air Force Research Laboratory (AFRL) Center managed by the University of Hawaii. A leader in scalable parallel computing technologies, MHPCC is a well-established member of the Department of Defense High Performance Computing (HPC) organization and an important resource for Hawaii's growing science and technology community, participating in collaborations and partnerships with academic institutions and private industry. A significant upgrade to its computational infrastructure is planned for FY 06 at an estimated cost of ~ \$29 M to \$ 31 M. MHPCC offers an innovative environment for High Performance Computing (HPC) applications including:

**Computational Resources:** Stable and secure parallel computing platforms for prototyping, benchmarking, and testing applications. MHPCC is ranked as one of the top HPC centers in the DoD in terms of computational capabilities.

**High-Speed Communications Infrastructure:** OC12 connections (Hawaii), offering 620 megabit per second (Mbps) capacity, provided direct access to MHPCC resources – over the Defense Research and Engineering Network (DREN) and the Hawaii Intranet Consortium (HIC).

**Support Services:** An expert staff provides MHPCC users with systems, network and applications support in addition to assistance with code porting, optimization, and application development.

An educational partnership agreement (EPA) was executed in 2004 between AFRL and the University of Hawaii to share the resources of the MHPCC for programs of mutual interest. Computational biology is an area of mutual interest for both the DoD and UH and has been authorized under the EPA.

### **CGPBRI website (<http://cgpbr.hawaii.edu>)**

The CGPBRI website (<http://cgpbr.hawaii.edu>) was launched in June of 2004 to provide easy access to numerous international genomics centers and web-based bioinformatics tools with high computational capacity. CGPBRI provides on-site accessibility to major NCBI databases, including all completed bacterial and eukaryotic genomes, conserved domain and protein structural databases (CDD, COG, SMART, PFAM) and enables rapid batch analysis of large sequence collections using the BLAST, ClustalW and Pfam algorithms. In addition, a large collection of smaller analysis tools (EMBOSS, JEMBOSS, Primer3, etc.) is accessible. An HP ALTIX 350 processor with 16 nodes performs computations. Database access is provided by SCSI RAID5 drives or directly from 40 GB of memory.

Initially, 42 UH faculty were involved in the beta testing of the site and the associated web tools. The website is currently accessible to UHM students and faculty only. The site receives between 10 and 50 hits per day, which reflects the popularity and usefulness, considering it is not a public site. A formal public release is planned for the first half of 2005.

## **Education and Training**

In order to fulfill the educational mission of CGPBRI, curricular activities for different educational backgrounds and setting were initiated in 2004.

One undergraduate and two graduate courses relating to bioinformatics and high-throughput genome analysis were designed and submitted for the Department of Microbiology. These courses will integrate with new course offerings in computational biology and bioinformatics offered by the Department of Information and Computer Sciences.

**MICR 361    Introductory Bioinformatics**

**MICR 630    Microbial Genome: High Throughput Sequence to Metabolic Pathway Discovery**

**MICR 661    Bioinformatics and Comparative Genomics**

## **K-12 Bioinformatics Initiative**

The Interim Director of the Center Initiative met with high school teachers participating in the Bioinformatics Teacher Training Program Initiative from two local high schools, to assist with involving students in grades 9-12 in applications of bioinformatic tools in contemporary biological problems.

The teachers are being trained by CGPBRI staff now and will continue their interactions with the center initiative during summer 2005 to create a high school-based bioinformatic curriculum. It is also planned to submit a NSF proposal to provide financial support for the Bioinformatic Teacher Training Program Initiative. The proposal will request one-month salary for the participating teachers in the program and travel to one national conference.

A separate K-12 web section will be created within the CGPBRI Bioinformatic Website by fall 2005. Participating schools will be able to access selected bioinformatics tools and databases by logging on to the K-12 web section. A demonstration of the Viz3D visualization system in the participating teachers' classrooms at Leilehua and Kaimuki High Schools is also being considered for fall 2005.

## **Research and Development**

### **The Hawaii Papaya Genome Project**

Early in 2004, a landmark study on the evolution and structure of the Y-chromosome of the papaya was published in the renowned journal NATURE by Hawaii Agricultural Research Center (HARC) scientist Ray Ming and collaborators. The study represented the first extensive sequence study of a papaya chromosome and elegantly related fundamental research questions to agriculturally important and applied issues. It was immediately realized that a combined effort of UH/CGPBRI and the state- and federal agricultural research centers had the capacity to grow into a multi-institutional effort with the potential to make major inroads into the understanding of the papaya genome with a focus in Hawaii. An informal consortium of scientists was formed in 2004 under the leadership of Dr. Dennis Gonsalves, Director of the USDA Pacific Basin Agricultural Research Center (PBARC) in Hilo, who was instrumental in developing the first pathogen-resistant papaya variety through genetic enhancement. The consortium includes the UH CGPBRI, MHPCC, HARC, USDA, Pacific Telehealth & Technology (PTT) Hui and Nankai University, China. A strategic plan was developed to establish the division of labor between the different institutions, and to identify funding sources for the project.

The papaya industry may benefit from the results of such a project in a number of ways. For example, the cultivation of papaya trees would become far more cost-effective if a routine way was found to generate true-breeding hermaphrodite plants, which are used for commercial fruit production. Papaya farmers have also been plagued by a devastating

disease caused by the papaya ringspot virus (PRSV), which devastated the entire papaya industry on the big island of Hawaii in the early 90's. A transgenic papaya resistant to the virus was developed by Dr. Dennis Gonsalves, then at Cornell University, and University of Hawaii researchers, which was released in 1998 following a lengthy deregulation process which included field trials. The transgenic "Rainbow" strain is the only commercially available transgenic papaya in the world and comprises approximately 40% of papayas grown in Hawaii today.

Currently, the transgenic papaya can be marketed to Canada and the mainland USA. However, it cannot be marketed to Japan, a major market for the Hawaiian non-transgenic papaya, which can be expanded if transgenic papaya were allowed. Efforts to get the transgenic papaya deregulated in Japan were taken up by the Hawaii papaya industry, USDA, and the State of Hawaii with the aid of local scientists that had developed the transgenic papaya. Although progress has been made in obtaining technical information to get the transgenic papaya deregulated in Japan, some key areas in understanding the genetic constitution of the transgenic papaya are required, and need to be completed in a timely manner. The ready accessibility of large-scale sequence information on these strains provided by CGPBRI will significantly aid in the deregulation process of transgenic papaya varieties and thus provide the local papaya industry with an important marketing tool.

After the approval of the transgenic papaya in Japan, the Hawaii papaya industry and the USDA will open other international markets for the transgenic papaya. Mainland China has already expressed keen interest in importing the Hawaiian transgenic papaya, however, like Japan; the transgenic papaya needs to be deregulated. The necessary technical information to get the transgenic papaya deregulated in Mainland China will be produced by the Hawaiian consortium of scientists. Korea and the European Union are also potential markets for the Hawaiian transgenic papaya.

The transgenic papaya is a major focus of interest to the USDA in marketing 'minor' transgenic products outside of the US. In many ways, it is the perfect test case for the US since the transgenic has been consumed in the US and Canada for over six years.

Papaya is also used in a variety of medical, biotechnological and cosmetic applications that could be further developed into marketable products once a complete picture of the genomic sequence is obtained.

Today, the Hawaii Papaya Genome Project is well on its way. With a genome size of 375 Mbp the papaya genome is only about one-tenth the size of the human genome. A project of this magnitude poses formidable technical, computational, financial, and organizational challenges. The HPGP will be the first of its' kind to be completed in Hawaii, with local resources, and by local scientists, and it will symbolize Hawaii's successful transition into the age of biotechnology. The completion of this project will not only represent a technological and scientific milestone, but it will further solidify the status of the University of Hawaii as a conduit between basic research and industry, and as an important resource for an emerging biotechnology industry. The Hawaii Papaya

Genome project will serve both as a proof-of-concept endeavor, give the facility visibility in the community, and build the scientific recognition needed for sustained financial support.

To complete the sequencing phase of the project, a two-pronged approach has been taken. UH-CGPBRI will pursue a high throughput whole-genome shotgun approach (WGS), which will be complemented by sequencing of a bacterial artificial chromosome library (BAC) in collaboration with HARC. At maximum capacity, the sequencing aspect of the project will take approximately one year to complete. At the time of print, 25% of the total papaya genome has already been sequenced.

During the final phase of the project, the information content of the assembled sequences will be analyzed and annotated using MHPCC's computational facilities. Collaborative arrangements with UH bioinformaticians have been formed, as well as with experienced genome researchers at the National Institutes of Health (NIH).

## **Extremophile Microbial Genomes**

Sequencing of whole genomes is reshaping all of the life sciences from microbiology to medical science. The majority of the biomass on earth is microbial. Many of these organisms offer considerable potential for commercial or industrial exploitation. Extremophile genomics holds the promise to develop creative, innovative approaches to exploiting the rich microbial diversity for environmental, energy, or commercial applications in Hawaii. Genomics will address (1) microbial diversity, (2) novel, more efficient strategies for determining the sequence of microbes for which the complete sequence of a very closely related organism is known, and (3) novel strategies and tools for characterizing, manipulating, and modeling entire reaction pathways or regulatory networks of microbes, or groups of microbes.

As part of the CGPBRI and computational proteomics strategic plan, we have completed and published the complete genome sequence of the deep-sea  $\gamma$ -*Proteobacterium*, *Idiomarina loihiensis* sp. nov., isolated from a hydrothermal vent at 1300 m depth on the Lō'ihi submarine volcano, Hawai'i. The *I. loihiensis* genome comprises a single chromosome of 2,839,318 base pairs, encoding 2,640 proteins, four rRNA operons and 56 tRNA genes. This suggests that *I. loihiensis* relies primarily on amino acid catabolism, rather than on sugar fermentation, for carbon and energy, revealing an integrated mechanism of metabolic adaptation to the constantly changing deep-sea hydrothermal ecosystem.

## **Publications**

Hou, S., Saw, J.H., Lee, K.S., Freitas, T.A., Belisle, C., Kawarabayasi, Y., Donachie, S.P., Galperin, M.Y., Koonin, E.V., Makarova, K.S., Omelchenko, M.V., Sorokin, A., Wolf, Y.I., Li, Q., Keum, Y.S., Campbell, S., Denery, J., Aizawa, S.-I., Shibata, S., Malahoff, A. and M. Alam (2004) Genome Sequence of the deep-sea *g*-proteobacterium *Idiomarina loihiensis* reveals amino acid fermentation as a source of carbon and energy. **PROC. NATL. ACAD. SCI. USA. 101(52):18036-18041.**

Editorial comment in: **NATURE Reviews Microbiology (2005) 3: 196-7**

## **UH Project support**

### **Gernot Presting (Department for Molecular Biosciences and Bioengineering)**

Since his appointment in 2003, Asst. Prof. Gernot Presting has become an important consultant to CGPBRI. His background in bioinformatics and expertise as a co-investigator of the rice genome project both in an academic and in a corporate setting made him an important collaborator on the papaya genome project, for which he is one of the bioinformatics lead scientist.

### **Thomas Hemscheidt (Dept. of Chemistry)**

The research group of Prof. Thomas Hemscheidt has been investigating and characterizing anti-tumor compounds from cyanobacteria. The organisms *Scytonema* and *Planktothrix* synthesize such anti-cancer and a tumor-promoting compound, respectively. CGPBRI has assisted in the sequence analysis of two cosmid clones derived from cyanobacterial DNA, encoding elements of the biosynthetic pathways creating these compounds.

### **Mark Martindale (Kewalo Marine Laboratory, PBRC)**

In the context of investigating the evolution of developmental mechanism, the Martindale laboratory is involved in sequencing several large cDNA collections from a number of invertebrates, including (*Capitella*, *Cerebratulus*, *Mnemiopsis*, Myzostomid, Phoronid, *Themistes*, and *Xenoturbella*). This effort is carried out under a collaborative NSF award to assemble the basic backbone of the protostome “Tree of Life” using broad-based genomic, anatomical, and developmental data.

CGPBRI assisted in this project by providing bioinformatic assistance related to the preparation of raw EST sequence data for assembly, and reformatting for homology analyses.

## **Coordination with External Organizations**

### **Windber Institute**

Recognizing the disparate infrastructure capacity of UH for proteomic research on one side, and bioinformatics on the other, and the significant cost involved with acquiring the necessary instrumentation, UH sought a research alliance with another research institution to facilitate access of UH researchers to these technologies. A reciprocal collaborative arrangement between UH and the Windber Research Institute (WRI) in Pennsylvania and UH was formalized in 2004 through a memorandum of understanding, which provides UH researchers and CGPBR access to WRI's high throughput proteomics facilities. Conversely, WRI will have access to the high performance computing facilities at the MHPCC. This relationship has been extended by Windber's focus on breast cancer treatment and the UH Cancer Research Center. Three extramural grant proposals from UH involving a significant proteomics component have since been supported with written statements from CGPBRI.

### **IBM**

A joint study agreement between UH/ MHPCC and IBM was signed in 2003, formalizing an ongoing collaboration in life sciences research. IBM subsequently donated a P SERIES 690 (P4) Computer valued at \$1,000,000 to UH-MHPCC under IBM's Shared University Research (SUR) Grant Program. The computer will be used for the assembly and annotation of extremophile microbial genomes and for the identification of metabolic pathways of these organisms. Furthermore, the computer will be used for the prediction of three-dimensional structures of identified proteins and for molecular dynamic simulations. In addition, IBM provides UH faculty and students access to IBM's bioinformatics web page, bioinformatics software, and software updates. The collaboration continues as IBM Life Sciences supports the creation of major databases for UH/MHPCC research on infectious diseases, cancer and extremophiles.

### **Hawaii Agriculture Research Center (HARC)**

Founded in 1895, the Hawaiian Sugar Planters' Association (HSPA) has become an internationally recognized research center. Originally dedicated to improving the sugar industry in Hawaii it changed its name in 1996 to Hawaii Agriculture Research Center (HARC), reflecting its expanding scope to encompass research in forestry, coffee, forage, vegetable crops, tropical fruits, and many other diversified crops in addition to sugarcane.

HARC's biotechnology team utilizes techniques such as gene transformation through particle bombardment and/or *agrobacterium* to achieve insect and disease resistance to solve production problems and genome mapping to enhance breeding programs.

Currently, genetic transformation projects are on sugarcane, pineapple, coffee, and papaya. This year, genome-mapping research has expanded from sugarcane to other tropical fruits such as papaya, coffee, macadamia nuts, and pineapple.

### **Hawaii Biotech**

CGPGRI, Computational Proteomics Program at MHPCC, and Hawaii Biotech have developed a joint research proposal to identify antiviral drugs active against dengue virus. The project uses the X-ray crystal structure of the Den 2 envelope protein (solved by Hawaii Biotech and their collaborator, Steve Harrison at Harvard University) as a basis for a drug discovery approach.

The implementation of this project will involve a combination of molecular modeling, simulated molecular dynamics, *in silico* screening of chemical libraries, and testing of candidate compounds in biochemical and biological assays for antiviral activity. The parallel development of the molecular modeling and the biological assays is critical to the success of the project as it is essential to have verification of “hits” identified *in silico* early in the process in order to validate the approach. The project can be divided into five key steps, which are applied in an iterative manner to identify potential drug leads: 1) Library design; 2) Docking; 3) Visual inspection; 4) Biological assay 5) Rational improvement.

### **Pacific Telehealth and Technology Hui**

In December 1999, a Memorandum of Understanding (MOU) was executed between the U.S. Army's Pacific Regional Medical Command, Tripler Army Medical Center (TAMC) and the Veterans Affairs Medical and Regional Office Center, Honolulu (VAMROC Honolulu) to establish a partnership, the Pacific Telehealth and Technology Hui. The Hui's purpose is to leverage their mutual strengths and resources to improve the quality, accessibility, patient satisfaction, and cost-effectiveness of healthcare services provided to beneficiaries through the use of emerging and existing telehealth technologies.

In 2003 the Hui received \$1 M for FY04 for bioinformatics and genomics research to be undertaken in collaboration with the University of Hawaii and Maui High Performance Computing Center (MHPCC). In 2004 the Hui provided \$ 250,000 to the UH Center for Genomics, Proteomics and Bioinformatics Research Initiative to support the Papaya Genome Project.

### **Institute for Geological and Nuclear Sciences, New Zealand**

The INSTITUTE FOR GEOLOGICAL & NUCLEAR SCIENCES (GNS) is one of nine CROWN RESEARCH INSTITUTES in NZ, which performs functions similar to the research divisions of the US DEPARTMENT OF INTERIOR AND ENERGY. CGPBRI and GNS jointly developed a new initiative in geomicrobiology with the objective to find new ways in which microorganisms living in extreme geophysical environments can benefit humans from

industrial applications to the development of pharmaceuticals. As a first step GNS has built a geomicrobiology research facility at their Wairakei center in the heart of New Zealand's geothermal and volcanic area where a variety of extremophilic organisms live.

### **Michigan Center for Biological Information Collaboration**

The Michigan Center for Biological Information (MCBI) and the University of Hawaii have developed a partnership to collaborate on a variety of projects involving Bioinformatics, Genomics, Proteomics, Structural Biology and Animal Models. Both institutions are collaborating with IBM Life Sciences on a variety of high performance computing initiatives and Bio-IT middleware. This collaboration envisages joint grant proposal applications, including those to TELEMEDICINE AND ADVANCED TECHNOLOGIES RESEARCH CENTER (TATRC), involving bioterrorism and other homeland security and defense initiatives that leverage the existing programs and initiatives. It is anticipated that in the near future joint efforts will be undertaken in technology development for online biomedical informatics, visible/digital human education/training, including telemedicine, and in the development of collaborative projects centered around the unique Hawaii gene pool and ecosystem.

### **Eugene V. Koonin, Michael Y. Galperin (National Institutes of Health)**

Eugene V. Koonin, Michael Y. Galperin and CGPBRI have agreed to collaborate on issues of microbial genome annotation. This collaboration has already produced a manuscript (see publications).

### **Funding and Budget**

The University of Hawaii assumed management of the Maui High Performance Computing Center (MHPCC) in October 2001. The development of a bioinformatics program was identified as a major focus area for the MHPCC to utilize the existing high performance computing resources. The Research Corporation of the University of Hawaii (RCUH) provided \$200,000 in 2002 to begin the collaboration in bioinformatics between MHPCC and the UH Manoa campus. The initiative was continued in 2003 with a \$150K RTRF award from David Lassner, PI MHPCC, and an additional \$75K support from RCUH. In October 2003 UHM Chancellor Englert created the Center for Genomics, Proteomics and Bioinformatics Research Initiative and provided \$1M to establish the center with commitments of an additional \$1M each year for FY 05 and FY 06.

As a result of this initial investment from RCUH and UH; the UH CGPBRI received a \$2.5M award for its Computational Proteomics Program from the US Army Medical Research and Materials Command (USAMRMC)/ TATRC in December 2004. Fiscal year 05 funding for \$2.2M has also been committed. The UH CGPBRI has also received formal notification of forthcoming NSF award of \$420K. This combined funding support of \$5.12M represents more than a 3-fold return of UH/RCUH's initial investment of \$1.425M.

The UH CGPBRI is also seeking funding for continuation of the Computational Proteomics Program in FY06 of \$4 M from the USAMRMC/TATRC, \$870K from the 2005 Hawaii State Legislature in support of the Papaya Genome Project, and \$200K initial funding support from the USAMRMC/TATRC for MHPCC's Maui Institute of Molecular Medicine (MIMM).

The Papaya Genome Project is an international collaboration, which is supported and funded through many organizations including the NSF. The UH efforts are focused on the sequencing resources developed at the UH CGPBRI and the computing resources at the MHPCC. The HARC has invested more than \$1.5 M to study specific regions of the papaya genome over of the previous six years under the leadership of Dr. Ray Ming. Dr. Ming has recently received a four-year \$2.9 M National Science Foundation (NSF) award to continue the physical mapping of the papaya genome.

## **Corporate support**

### **Pipelinefx, Honolulu**

PIPELINEFX provided a co-investment with UHCGPBR in the form of a donation of server farm management software and technical support services valued at \$109,000, in conjunction with the purchase of the new SGI Altix 350 computation server.

### **Core Systems Hawaii**

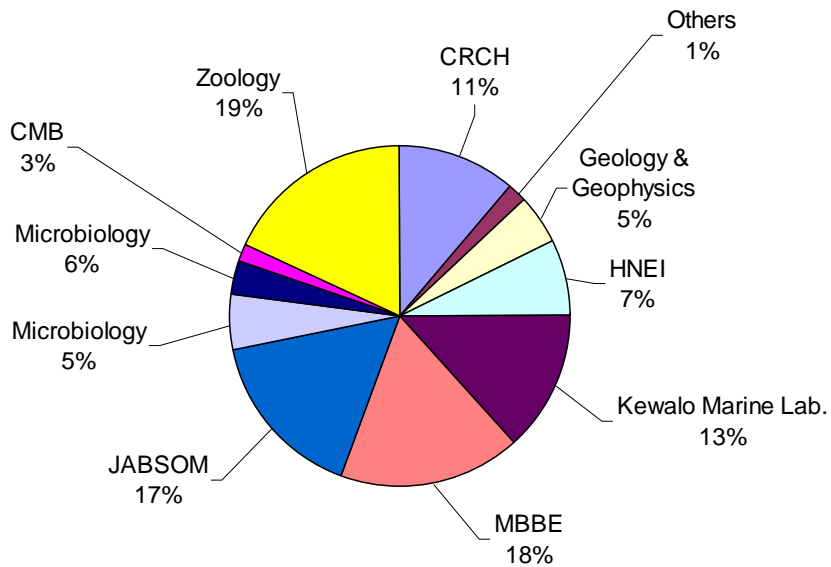
CORESystems Hawaii provided a co-investment with UH CGPBR in the form of a donation of products and services valued at \$17,854 in conjunction with the purchase of the SGI Altix computation server. This support follows previous donations of ten SGI Octane desktop systems, 10 20" color monitors and one 16 CPU Origin2000 server with six months of maintenance.

# Financial Statements

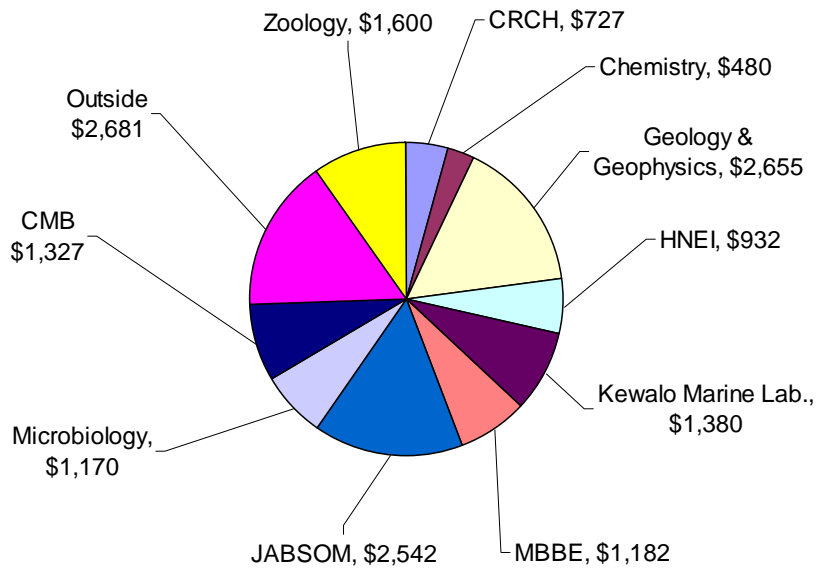
## FINANCIAL REPORT AS OF MARCH 31, 2005

<b>EXPENSE DESCRIPTION</b>	<b>EXPENSES</b>
SALARIES & FRINGE BENEFITS	\$225,332.74
CONTROLLED PROPERTY	\$31,867.00
MATERIALS & SUPPLIES	\$253,151.85
DOMESTIC TRAVEL	\$24,645.56
CONSULTANT/FEE FOR SERVICES	\$167,841.97
COMMUNICATIONS	\$439.76
EQUIPMENT	\$1,561,451.78
REPAIRS & MAINTENANCE	\$2,520.67
OTHER CURRENT EXPENDITURES	\$5,753.57
<b>TOTAL CUMULATIVE EXPENSES</b>	<b>\$2,273,004.90</b>

### Number of Orders by Department



### Charges by Department



## SURVEY LETTER

Dear Dr.

I am writing to you because you or someone in your lab has recently used or inquired about the services of our high throughput DNA sequencing facility within CGPBRI.

We strive to provide the UH research community and researchers state-wide with a competitively priced facility for high throughput sequencing with a fast turnaround time, and to offer superior quality sequence data. We believe that we now have both the equipment infrastructure and expertise to handle entire genome projects, opening the door for multidisciplinary and highly collaborative projects that may attract large extramural grants to our university (as part of our research program on Papaya Genome we have just completed 2X coverage of the genome, more than 1,200, 000 sequencing reactions with total raw data of 750 million bp).

We can only continue to provide these and additional services as long as you, the customer, are satisfied with our work. I therefore respectfully ask you to provide us with your candid opinion of the quality and pricing structure of our DNA sequencing service, and with any comments you may have regarding other services which we may use to direct our planning efforts.

Please send your response to [margy@hawaii.edu](mailto:margy@hawaii.edu) and is greatly appreciated.

Thank you for your support.

Heinz Gert de Couet  
Professor  
Department of Zoology  
Associate Director  
Center for Genomics, Proteomics, and Bioinformatics Research Initiative  
Keller Hall #319  
Honolulu, HI 96822  
808 9560924  
808 9560925 Fax  
<http://cgpbr.hawaii.edu/>

1.----- Original Message -----

From: "Eric Gaidos" <gaidos@hawaii.edu>  
To: "Shaobin Hou" <shaobin@hawaii.edu>  
Sent: Friday, June 03, 2005 8:23 AM  
Subject: Sequencing Service feedback

Dear Shaobin,

The initial experience of my lab using the sequencing services of the Center for Genomics, Proteomics, and Bioinformatics has been very positive. The last sequencing run was of very high quality, with extremely clear peaks, and 1000-basepair reads, and it was turned around in less than a day. More importantly, you worked with us on some of the earlier issues we encountered in terms of template quality and the matching of forward and reverse reads. With the level of quality we are experiencing now, we would expect to greatly expand the use of your sequencing service in the near future as we engage in some larger-scale genomics projects.

Sincerely yours,  
Eric Gaidos  
Assistant Professor of Geobiology  
Department of Geology & Geophysics

2--- Original Message -----

From: "kvtbwp" <kvtbwp@yahoo.com>  
To: "Jimmy Saw" <jimmy@hawaii.edu>  
Sent: Monday, May 09, 2005 10:10 PM  
Subject: Sequence

Hi Jimmy,

Thank you very much!

Both sequences are perfect and for the first time I had a good read over 800bps. My lab mates are eager to use this service in future. I do have a template of the form and shall forward it to them.

Thank you once again for the trial run and good sequence.

Thara

3. From: "Guangyi Wang" <guangyi@hawaii.edu>  
To: "Shaobin Hou" <shaobin@hawaii.edu>  
Cc: <quanzi@hawaii.edu>  
Sent: Wednesday, June 01, 2005 10:11 AM  
Subject: Re: Sequencing results

Hi Shaobin,

You did a great job! More samples will be sent to you next week. Also, since we are internal funding transfer, could you check if there is other way to pay the service fee? Thanks.

Guangyi

4. From: "Quanzi Li" <quanzi@hawaii.edu>  
To: "Shaobin Hou" <shaobin@hawaii.edu>  
Sent: Thursday, June 02, 2005 5:15 PM  
Subject: Re: Sequencing data

Hi shaobin,

I think your sequencing results is so good. Before I send our samples to othe two places, one place is very slow to finish the sequecning. One place is in california, but it will spend one day for shipping. it is acceptable to get the results around 3 days after we send out the sample.

The quality is good. This time I have 94 samples, 77 samples have whole sequence, Even the samples with bad results maybe is due to the sample template problem.

Compared with other sequence company, the price is so cheap. Usually they charge at least \$6 or 7. Is the method you used is different from the other company, or just use the different machine?

quanzi

5. From Peter B Erickson <petere@hawaii.edu>  
Date Fri, 03 Jun 2005 11:01:44 -1000  
To CGPBRI Center for Genomics Proteomics and Bioinformatic Research Initiative <sequence@hawaii.edu>  
Subject Re: DNA Sequencing service

---

Hi,

I was pleased with my sequencing results, especially at such an economical price. It did seem that the beginning of the sequence was messy for longer than I've seen before, i.e. the sequence wasnt readable until about 80 bp in, but I have fairly limited sequencing experience. Overall, I was pleased, and intend to send you many more samples. Thanks.

Bjorn

6. From Gernot Presting <gernot@hawaii.edu>  
Date Thu, 02 Jun 2005 15:00:10 -1000  
To CGPBRI Center for Genomics Proteomics and Bioinformatic Research Initiative <sequence@hawaii.edu>  
Subject Re: Feedback for Sequencing service

---

Shaobin,

thanks for your message. I will be glad to give feedback, but haven't had a chance to evaluate sequence quality quite yet.

However, I CAN already say that the turnaround time was fantastic and the staff (meaning you) are wonderful to interact with.

I am working on getting a couple of previous projects finished, hopefully by this weekend. Early next week I want to take a look at the sequences and will give you more feedback then.

Best regards,

Gernot.

7.----- Original Message -----

**From** Yan Chen <cheny@hawaii.edu

**Date** Thu, 02 Jun 2005 17:51:44 -1000

**To** CGPBRI Center for Genomics Proteomics and Bioinformatic Research Initiative <sequence@hawaii.edu

**Subject** Re: Sequencing Service

Hi, Dr. Hou:

Thank you for asking me opinions. I am very satisfied your services, the price is cheaper and the quality of sequencing data is much better than the one I used. The turnaround time is quicker, also. I would like to introduce my colleagues to sequence DNA in your center.

Best regards,

Yan Chen, Postdoctoral  
Molecular Biosciences and Bioengineering  
University of Hawaii  
Honolulu  
Hawaii, 96822, U.S.A  
Telephone: (808) 782-2754;  
E-mail: [cheny@hawaii.edu](mailto:cheny@hawaii.edu);

8. --- Original Message -----

**From** Sara Okamura <SOkamura@crch.hawaii.edu

**Date** Thu, 02 Jun 2005 16:33:07 -1000

**To** sequence@hawaii.edu

**Subject** Re: Feedback

---

Hi,

Dr. Lorenzo and I are both satisfied with your DNA sequencing service so far. Turnaround time was pretty quick and pricing is very competitive. Also, the length of accurate sequencing for several reactions was very good - rather impressive.

We are happy to have your facility as a new option for convenient and affordable sequencing, and we plan to continue to use your services in the future. Thank you for your interest in feedback.

Regards,

Sara

sokamura@crch.hawaii.edu

Sara M. Okamura, Ph.D.  
Laboratory of Dr. Patricia Lorenzo  
University of Hawaii  
Cancer Research Center of Hawaii  
1236 Lauhala Street, Room 315  
Honolulu, Hawaii 96813  
phone: (808)-440-5227  
FAX: (808)-586-2970

9.

From: "Tom Ramsey" <ramsey@math.hawaii.edu>  
To: <shaobin@hawaii.edu>  
Sent: Friday, June 03, 2005 4:02 PM  
Subject: new genomics services

Congratulations on being up and running! I was especially pleased to see the pricing chart offering services to non-campus researchers. That's entrepreneurship!

I'm a member of the SEC (Senate Executive Committee, UH Manoa Faculty Senate). A couple of SEC members were wondering where to put the center in a "org chart" (it is a rather bureaucratic thought). I'm more interested in how you set yourself up, obtained start-up funding, etc.

Again, the best of success to your endeavors!

Tom Ramsey

10.

From: C L <lippe@hawaii.edu>  
To: Shaobin Hou <shaobinhou@gmail.com>  
Date: Jun 13, 2005 1:43 PM  
Subject: Re : Sequencing data

[Reply](#) | [Reply to all](#) | [Forward](#) | [Print](#) | [Add sender to Contacts list](#) | [Trash this message](#) | [Report phishing](#)

Thanks for the sequences, they are great! We are really happy about the service you're offering, without talking about the price you ask for it. It's really good to get sequences going up to 800-900 bp. We should bring you some more this week, probably tomorrow in fact. Those sequences are really important for the project we're working on at the moment, so it would be great to get them in 2-3 days (maybe around thursday if possible?).  
Thanks again for your service,

Catherine

11.

From	Tricia Murata <m2sequences@yahoo.com <a href="#">Add Sender</a>
Sent	Friday, June 24, 2005 8:35 am
To	CGPBRI Center for Genomics Proteomics and Bioinformatic Research Initiative <sequence@hawaii.edu
Subject	Re: Sequencing data

Hi Shaobin!

the sequences were perfect and fast. When do I receive the new sequences of the probes I gave to you?

Best,

Andi

\*\*\*\*\*

12.

Hello,

We have been using your sequencing service with great success. I'm looking for a laboratory (preferably on Oahu) that is setup to do high-throughput SNP detection. Such as SSCP analysis or Fluorescent Allele Specific PCR. I thought you guys might be able to help me out with this.

Thanks,  
Dave

David B. Carlon  
Assistant Professor  
Department of Zoology  
2538 McCarthy Mall, Edmondson 152  
University of Hawaii  
Honolulu, HI 96822

13.

Michael John Melzer  
<melzer@hawaii.edu  
u [Add Sender](#)

Sent	Thursday, August 11, 2005 11:28 am
To	Margy J K Yuen <margy@hawaii.edu
Subject	Re: High Throughput DNA Sequencing Service at CGPBRI

Margy,

I think your facility is doing a great job. The price and data

quality (read lengths) are fantastic, however, the quick turn around time is by far what I appreciate the most about your facility. This 1- 2 day turnaround (compared to 6-8 days for Gilmore) is very helpful for my experiments.

Unfortunately, my experiments do not require high-throughput - I usually only have 2-15 samples per experiment, so I was very disappointed to learn you are soon going to only accept 96 well plates, with Gilmore taking in the 'loose' samples.

To me it would seem logical for people like me to continue to drop off their 'loose' samples at Gilmore, and have their staff combine them onto a single plate, and have that sequenced at your facility, even if that means an extra 'handling' charge to be paid by the user to the Gilmore folks. That way users would still get quicker turnaround, and the overworked Gilmore staff would get a break and focus more on their other core services.

Just an idea!

Take care,  
Mike.

14.

▶ Zheng Lan  
<ZLan@crch.hawaii.ed  
u [Add Sender](#)

Sent	Thursday, August 11, 2005 2:14 pm
To	<a href="mailto:margy@hawaii.edu">margy@hawaii.edu</a>
Subject	Sequencing

Dear Dr. de Couet,

We are very satisfied about the quality and pricing of your sequencing and it is a big help to our research. We will continue to send more samples to your center for sequencing.

Many thanks!

Zheng

Zheng Lan, Ph.D.  
Cancer Research Center of Hawaii  
1236 Lauhala Street #306  
Honolulu, HI 96813

Phone: (808) 440-5219  
Fax: (808) 586-2970  
Email: [zlan@crch.hawaii.edu](mailto:zlan@crch.hawaii.edu)

15.

▶ Lori Mieko  
Kaneshige  
<lorikane@hawaii.ed  
u [Add Sender](#)

Sent	Thursday, August 11, 2005 7:18 pm
To	Margy J K Yuen <margy@hawaii.edu
Subject	Re: High Throughput DNA Sequencing Service at CGPBRI

Dear Margy,

Yes, I have been using the sequencing facility at CGPBRI, and I am very pleased with the service. The price is very reasonable, and the turn-around time is quite fast.

The only suggestion/complaint I have would be to include N's for nucleotides that are ambiguous in the sequence read out.

Thank you for providing the service,

Lori Kaneshige  
Kewalo Marine Laboratory  
PBRC/University of Hawaii  
41 Ahui St.  
Honolulu, HI 96813

(of) 808-539-7323  
(FAX) 808-599-4817

lorikane@hawaii.edu

16.

Dear Ms. Yuen,

I can only say that I am absolutely thrilled with your service. The reads we have obtained are better than I have ever seen before (I am averaging something over 800 bases/sequence on plasmids), and the turnover time has been at most 2 days. This is so much better than I am used to. In addition, I find the price to be very acceptable.

Yours truly,

David C. Fritzinger, Ph.D.  
Associate Specialist  
Natural Products and Cancer Biology Program  
Cancer Research Center of Hawaii  
1236 Lauhala St.  
Honolulu, HI 96813  
(808) 586-2966

(808) 586-2970 (Fax)  
(808) 282-8843 (Cell)

---

**17.**

Dear CGPBRI -

I have used your service and was satisfied with the results. Thus far I have only used the service for one 96 well plate of purified PCR products, all of which turned out satisfactorily.

I will say however that one of my graduate student's has had some problems in terms of the data output. Specifically his text files lack any ambiguous base calls (N's), leading to the conclusion that the data are clean and sequences worked perfectly. However, when he tried to align these sequence fragments, certain portions of contigs were substantially different where they should overlap (F and R), leading him to take a look at the trace files (electropherograms). There we found a disturbing trend where there are instances where dye blobs, salts, or other contaminants have resulted in secondary, or even triple peaks but the software has read through the sequence in an apparently arbitrary way, without calling any "N"s, resulting in mistaken calls. This tendency has resulted in a requirement of the user (us) to check all trace files to be sure that miss-called bases do not throw off the results. This is not usually the case, as we can easily detect the rough spots in a fragment by the presence of N's with every other service we have used.

This is our only concern. Again in terms of price and turnaround time we are satisfied, and once this base-calling issue is resolved we will have confidence in the quality of CGPBRI's service .

Sincerely,

-Brenden Holland

---

**18.**

Hello,

I have used the services twice and had been very satisfied with the results and with the turn around time. I would love for you guys to add GeneScan fragment analysis services.

Thank you,  
Alexandra Gurary

---

**19.**

Hi,

So far, I have only needed your service for fairly small-scale sequencing, but I have been pleased with it, particularly the price and the turnaround time. I will continue to use your service for all my sequencing. Thanks.

Bjorn\_\_\_\_\_

**20.**

Hi,

I have been very satisfied with the sequencing results I've been receiving from the high-throughput DNA sequencing service at CGPBRI. Turnaround time is fast and reads are long. I've also gotten good feedback regarding samples that didn't work very well.

I just noticed today that I can no longer find the 1-30 sample order form on the website. Do we have to now submit only in 96-well plates? We do not actually have high-throughput needs, but we appreciate the speed and quality of your service and would like to continue to use your facility for our small scale sequencing. But have you done away with the small batch sequencing to be able to focus on higher throughput?

Thank you for your interest in our feedback.

Regards,  
Sara Okamura

**21.**

----- Original Message -----

From: Michele Champagne <saracrispy@earthlink.net>  
Date: Saturday, August 13, 2005 7:04 am  
Subject: High Throughput Sequencing Service at CGPBRI

Dear Margy,

I received an email requesting feedback regarding the service I have used at CGPBRI. I have limited experience with this type of work and can provide only a few comments:

The speed with which our order was processed was excellent. In less than one week we obtained data and were able to move our research efforts forward quickly because of this.

Dr. Shaobin Hou has been very helpful and pleasant to work with in all regards. He answered my questions quickly and has encouraged me to contact him if I have any other needs or questions.

As regards pricing, we have no objections to the fees we have been charged for your service. They are in line with what I expect a private company should pay for this service. If we were doing a larger sequencing project, we might request a small discount.

One service we would be interested in but do not know if it is available: Assistance with contig assembly and annotation. We are currently using Vector NTI software on our laboratory computers, and, while it's sufficient, we are not proficient at operating it quickly. If CGPBRI has available a person for hire to assist us with this

part of the data analysis, using either our software or a program operated by CGPBRI, we would be very interested in making us of that service immediately. Our research budget is well-funded at this time, and we would like to advance our work as quickly as possible.

Overall, we are very happy to work with the staff and facility at CGPBRI.  
Sincerely,  
Michele Champagne

Michele Champagne, Ph.D.  
Research Scientist  
Kuehnle AgroSystems, Co LLC  
Manoa Innovation Center  
2800 Woodlawn Drive  
Suite 281  
Honolulu, Hawaii 96822-1843  
808-753-2693 phone  
808-988-3170 fax  
[kasllc@hawaii.rr.com](mailto:kasllc@hawaii.rr.com)

22.

**From** Margaret A Onyango <onyango@hawaii.edu>

**Date** Mon, 15 Aug 2005 09:55:05 -1000

**To** Margy J K Yuen <margy@hawaii.edu>

**Subject** Re: High Throughput DNA Sequencing Service at CGPBRI

---

Thank you!

I will be needing more of your services beginning Spring 2006. I must say, so far you have done an excellent job for me with the few samples I sent you.

I wonder if you could do sequencing for direct pcr products without me cloning first? It would greatly reduce my cost because of sequencing if it is not necessary to clone before sequencing.

Thank you.

sincerely  
margaret

University of Hawaii at Manoa  
Tropical Plants and Soil Science  
3190 Maile way, St John room 102  
Honolulu, Hawaii HI 96822.  
808-956-2153

**23.**

HIRO  
<kazuhiro@hawaii.edu  
> [Add Sender](#)

Sent	Monday, August 15, 2005 11:57 pm
To	'Margy J K Yuen' <margy@hawaii.edu>
Subject	RE: High Throughput DNA Sequencing Service at CGPBRI

We have been using your DNA sequencing service since this July. We are much satisfied with the results and the quickness.

The prices are also good deals. But the cheaper the more helpful.

We want you to keep providing this service.

Hiro

Kazuhiro (Hiro) Iwahashi

Institute for Biogenesis Research

John A. Burns School of Medicine

University of Hawaii

1960 East-West Road

Honolulu, HI 96822

Tel: 808-956-5475

FAX: 808-956-7316

E-Mail: kazuhiro@hawaii.edu

**24.**

C L  
<lippe@hawaii.edu  
> [Add Sender](#)

Sent	Wednesday, August 17, 2005 1:30 pm
To	Margy J K Yuen <margy@hawaii.edu>
Subject	Re : High Throughput DNA Sequencing Service at CGPBRI

Dear Margy,

I talked to you on the phone this morning, and I'm really sorry I didn't answer before today.

I have been using your sequencing facilities a lot lately, and I appreciate the service you are offering. The sequences coming back to me are really clean and long, which

is great. The price is also extremely competitive, and makes sequencing a lot more affordable, as it should be. People as Shaobin are really helpful and ready to take some of their time to give suggestions about our results. Usually the results come back after 2 days, which I think is a good turnaround time.

I still have some suggestions for you and the sequencing team to make it even better. I know that you guys are mainly working on big projects as genome mapping, but I am working with about 20 samples for each morphs of a specific species. Precision is thus extremely important, and it happened 3-4 times that sequences were a bit mixed up (names, F/R, even species). For a project as mine, it is crucial that identification is perfect, so maybe a suggestion would be to take extra-care when loading and identifying the samples.

That's it! I really hope you're going to keep this service available to us, as it is good. Have a good day!

Catherine Lippé  
Lab technician  
University of Hawaii at Manoa

**25**

Ying Huang  
<yinghuan@hawaii.edu  
> [Add Sender](#)

Sent	Wednesday, August 17, 2005 3:26 pm
To	Margy J K Yuen <margy@hawaii.edu>
Subject	Re: High Throughput DNA Sequencing Service at CGPBRI

Hi

Since maybe my product has some problem, i didn't get very good sequencing result. So i don't know the quality of sequencing service. But i heard from my labmates the quality of sequencing is good. I want to mention i am very satisfied with the service your guys provided. Although i didn't get good results, shaobing also gave me good suggestions about the technique to help me improve my results.

Ying

**26.**

Madhusudhan Rapolu  
<rapolu@hawaii.edu  
> [Add Sender](#)

Sent	Thursday, August 18, 2005 7:14 am
To	Margy J K Yuen <margy@hawaii.edu>
Subject	Re: High Throughput DNA Sequencing Service at CGPBRI

Dear Prof. Heinz Gert de Couet and CGPBRI Team,

Sequencing a DNA of interest is one of the first steps in DNA research. I have been using your Service regularly. And, I have no complaint. Sequence quality is impressive, and the turn-around time can hardly be shorter.

Going through your website, I could not find a list of common primers (M13, T7 promoter etc.) that your Service has for sequencing.

Thanking you,

Yours sincerely  
Madhusudhan Rapolu  
Postdoctoral Researcher  
Molecular Biosciences and Bioengineering Dept.  
1955 East-West Road #218  
Honolulu, HI 96822

Tel: 808-956-9597  
Email: rapolu@hawaii.edu

**27.**

▶ Noemi Polgar  
<polgar@hawaii.edu  
> [Add Sender](#)

Sent	Thursday, August 18, 2005 9:39 am
To	Margy J K Yuen <margy@hawaii.edu>
Subject	Re: High Throughput DNA Sequencing Service at CGPBRI

Dear Margy,

I have a good experience with your sequencing core. The quality of the work is good and your prices are reasonable. I am satisfied with your services. Thank you for asking my opinion.

Noemi Polgar  
Ph.D. student  
Cardiovascular Research Center  
John A. Burns School of Medicine,  
University of Hawaii at Manoa  
1960 East-West Road T311  
Honolulu, HI 96822  
Tel: 808/ 956 6856  
E-mail: polgar@hawaii.edu

28.

peckham@hawaii.edu u <a href="#">Add Sender</a>	
Sent	Thursday, August 18, 2005 10:09 am
To	Margy J K Yuen <margy@hawaii.edu>
Subject	Re: High Throughput DNA Sequencing Service at CGPBRI

I have found your facility very accomodating. Your staff is always friendly and helpful to me. I think you guys provide excellent service. Our lab was having primer problems some time back and your facility helped us figure out the problem by providing T7 primers. Thanks, and keep up the good work!  
Gabriel

29.

Mayee Wong <mayee@hawaii.edu> > <a href="#">Add Sender</a>	
Sent	Thursday, August 18, 2005 12:28 pm
To	margy@hawaii.edu
Subject	CGPBRI seq facility

Hi Margy,

The following are my comments about the service I've received from the CGPBRI sequencing facility:

PRICES - Sequence prices are very reasonable  
TURN AROUND TIME - I usually receive my sequences 2-3 days from the time I submit my samples. I consider this a reasonable turn around time.  
QUALITY OF SERVICE - I have found the staff (Shaobin, Margy, etc) to be very courteous and easy to work with.  
AREAS FOR IMPROVEMENT - It would be nice to have an estimate of when I can expect to get my sequences back. I usually forget to ask when I turn in my samples, so it'd be nice to have staff be asked to remember to give a turnaround estimate to clients everytime they drop off samples.

---

Margy, my apologies for not replying to your original request in a timely manner. Feel free to let me know if you need more info.

-Mayee

30.

Yusuke Marikawa <marikawa@hawaii.edu> > <a href="#">Add Sender</a>	
Sent	Friday, August 19, 2005 9:22 am
To	'Margy J K Yuen' <margy@hawaii.edu>
Subject	RE: High Throughput DNA Sequencing Service at CGPBRI

Dear Margy,

I have been using the CGPBRI sequencing service since June 2005, and so far had more than 500 samples analyzed. The service is excellent. The quality and price are very good, and I am totally happy with your service.

Thank very much for helping our research.

Yusuke

31.

Ken Hayes  
<[khayes@hawaii.edu](mailto:khayes@hawaii.edu)  
> [Add Sender](#)

Sent	Friday, August 19, 2005 12:13 pm
To	<a href="mailto:margy@hawaii.edu">margy@hawaii.edu</a>
Subject	

Dear Margy Yuen,

I apologize for taking so long to reply to your request for feedback, however, I've been off island for the last few weeks and I'm just now getting caught up on the backlog of email.

I am happy with the pricing structure of the sequencing services. I do have a few problems with the data that I have received. Before I describe the problem I want to tell you that I am very happy with the speed at which Shaobin and his staff have produced results and their willingness to work through any problems. With that said I would like to bring a few problems to your attention that you may or may not already be aware. First, when our sequence data has been sent to us it contains only the .scf file without any of the information that is normally in the .abi file (i.e. signal strength, confidence values etc). Also, the base caller software is set to call all bases regardless of ambiguity, which results in sequences with many miscalls and hidden ambiguity. This combined with the lack of signal strength and confidence value data makes it difficult to evaluate the quality of the sequences as well as trouble shoot any potential problems. This is a problem that should be easy to solve. I'm not real familiar with the new ABI software, but I'm almost certain that these problems are just the result of improper or default settings.

Based on our limited usage of the facility to date I would have to say that we are for the most part happy with the service. I think once the above mentioned problems are corrected we will continue using the facility. I'd also like to suggest that you begin offering genotyping services as well. There are a number of labs including ours that utilize microsatellites and it would be convenient to have these services performed on campus.

If you have any other questions or need further feedback please don't hesitate to contact me.

Thank you,

Ken Hayes

32.

Wendy Kurata  
<WendyK@crch.hawaii.edu  
> [Add Sender](#)

Sent	Monday, August 22, 2005 11:24 am
To	margy@hawaii.edu
Subject	Re: High Throughput DNA Sequencing Service at CGPBRI

Hi Margy,

We found out and started using your facility just recently (May, I believe) and have sent a few samples so far.

We found the facility to be very efficient with a fast turnaround time and very reproducible sample runs. (We were having issues with our current facility and so this was a very welcomed find).

The pricing is very good.. although we were not paying for our facility (as our CORE grant helped to fund the facility), but at the time it was the cheapest one on campus and so we can't complain. Since then, I have heard that the other facilities are also charging comparable prices to your facility.. this is good for us too! Thank you.

We do not do a lot of sequencing continually.. we sequence in spurts (only to clone and mutate our genes of interest) but have turned on others at the center to the facility and they have been very happy with their results too.

As we are off campus, the only suggestion I can think of to make OUR life easier .. would be a courier service to pick up samples. We did have such a thing previously, although as I mentioned before, we were probably paying for it through a core grant.. and not individually. But we have students and post docs willing to drive there, so its not a really big issue for us... one our PI would definitely not want to pay for (at least.)

I hope this is sufficient for what you need. Let me know if you have any direct questions or comments you would like addressed.

Thank you.

wendy

33.

"Ivo M. Krab"  
<ivo@hawaii.edu  
> [Add Sender](#)

Sent	Monday, August 22, 2005 11:27 am
To	Margy J K Yuen <margy@hawaii.edu>
Cc	"Ivo M. Krab" <ivo@hawaii.edu>
Subject	Re: High Throughput DNA Sequencing Service at CGPBRI

Hi Margy,

I have actually been very happy with your service. The quality of the sequences is excellent, much better than the results I had with the core facility at Gilmore. The only time the sequencing failed was when I had a DNA rearrangement that eliminated the primer binding site :-(

Your turn-around time is also much better. Before I had to wait a week or so for results, now it's rarely more than 3 days. Of course if you can still improve that, it would be highly appreciated ;-)

Price is not my direct concern, but my PI doesn't complain 8-»

In other words, thanks, and keep up the good work....

Ivo

PS, have you read that science paper about non-electrophoretic genomic DNA sequencing? (Science Express PDF: Shendure et al., Accurate Multiplex Polony Sequencing of an Evolved Bacterial Genome, Science 2005 0: 11173891)

VERY interesting.....

34.

Clay Wakano  
<cwakano@queens.org  
> [Add Sender](#)

Sent	Tuesday, August 23, 2005 3:29 pm
To	margy@hawaii.edu
Subject	sequencing opinion

Dear Dr. de Couet,

I am very pleased with the service that was provided by the sequencing facility within CGPBRI. The speed at which we received the sequencing data from our samples was rapid. The accuracy of the sequences were good and the price of the sequencing was excellent. The service provided by Dr. Shaobin Hou was excellent. He was very helpful and approachable. He answered all technical questions asked. He responded to emails and phone calls in a timely manner. His service and help was great and much appreciated. Thank you.

Clay Wakano

35.

Dawn M Reding  
<reding@hawaii.edu  
> [Add Sender](#)

Sent	Tuesday, August 23, 2005 4:32 pm
To	margy@hawaii.edu

Subject sequencing at CGPBRI

Hi,

Just responding to your email about my opinion of the sequencing services offered at CGPBRI. I have only sent samples twice so far, but have had reasonably good results. The samples were a little noisy, but it may be b/c of things on my end. The price is very competitive (it's half of what I was paying before), and it is very convenient b/c I can walk the samples down to the lab instead of having to send them in the mail. The turn-around time is also quick, which is greatly appreciated. I plan to send more samples to CGPBRI for sequencing and may be able to offer a more detailed evaluation later, but I'm satisfied with the results and pricing so far.

Thanks,

Dawn

Dawn Reding  
Graduate Student  
University of Hawaii, Manoa  
Department of Zoology  
2538 McCarthy Mall, Edmondson 152  
Honolulu, HI 96822  
Office: 808-956-4717  
Email: reding@hawaii.edu

36.

Peter Racz  
<petiracz@yahoo.com  
> [Add Sender](#)

Sent	Wednesday, August 24, 2005 12:28 pm
To	Margy J K Yuen <margy@hawaii.edu>
Subject	Re: High Throughput DNA Sequencing Service at CGPBRI

Dear Margy K Yuen,

I`m sorry that I`ve answered so late, but I was on holiday!  
We are very pleased with your DNA sequence serrvice .  
You have a reasonable price and your service is really fast !

Thank you

Peter Racz

37.

From	Sharon E Mozley-Standridge <mozley@hawaii.edu> <a href="#">Add Sender</a>
Sent	Thursday, August 25, 2005 1:48 pm
To	Margy J K Yuen <margy@hawaii.edu>
Subject	Re: High Throughput DNA Sequencing Service at CGPBRI

Well I've only used your service once. The price is competitive but the sequence files are in a format that I am not familiar with. Do you know how to convert scf files into abi format?  
Sharon

Dr. Sharon E. Mozley-Standridge  
Post-Doctoral Fellow  
Hawaii Natural Energy Institute  
Department of Oceanography,  
University of Hawai`i at Manoa,  
1680 East-West Road, POST 109  
Honolulu, HI 96822  
Lab Phone: 808-956-8627  
Cell: 706-424-2220  
mozley@hawaii.edu

38.

	ketil@soest.hawaii.edu <a href="#">Add Sender</a>
Sent	Thursday, August 25, 2005 4:04 pm
To	Margy J K Yuen <margy@hawaii.edu>
Subject	Re: High Throughput DNA Sequencing Service at CGPBRI

Dear Margy

During the last semester, I have used the Center for Genomics, Proteomics, and Bioinformatics on several occasions for sequencing of environmental clone libraries. I am greatly satisfied with the quality of the results and the level of service provided by the staff. The prices are lower and the turn-around time is faster than any other commercial facility I am aware of.

Until now, we have been providing miniprep-products ready for sequencing. However, if the prices remain competitive we will outsource the miniprepping to the Center as well in the future.

Cheers,  
Ketil Sorensen, post doc