24 September 2007

Dear Colleagues,

My appointment to a tenure track faculty position at Baylor University represents my first appointment of this type in any university system. I am keenly aware of the faith that Baylor has demonstrated in my talents as a researcher and educator to further its 2012 vision. This faith is continuously demonstrated by the wonderful support and encouragement I have received from the Baylor administration and my senior colleagues, providing the necessary resources for successfully accomplishing the three primary goals of a new faculty member in the department of Computer Science specializing in bioinformatics. One goal is to establish a nationally recognized undergraduate curriculum in bioinformatics and to achieve excellence in the didactic of appropriate course material. A second goal is to establish an independent, self-sustaining extramurally funded research program in computational biology and/or bioinformatics. And, thirdly, I have a responsibility to uphold, promote and represent the ideals and morals implicit of Baylor’s Christian identity and focus in advanced research.

Previously, as a post-doctoral fellow in the Genome Analysis and System Modeling group at Oak Ridge National Laboratories, my efforts were divided 85% to 15% for research and teaching, respectively. While I am continuing my research efforts at Baylor with all possible intensity, I am likewise investing a considerable measure of energy into enhancing my thoroughness and effectiveness as a teacher. These efforts include self-education to expand my knowledge base to include more overlapping areas of interest with other faculty members and students, the creation of a curriculum tailored to the specific needs of bioinformatics, the implementation of new courses designed to meet the contemporary needs in the field, and an iterative approach of review and modification to enhance my effectiveness as an educator.

I believe that effective teaching begins outside of the classroom and to this end I have become fully engaged in the significant task of modifying the current bioinformatics curriculum and establishing new courses. Indeed, the lack of a standardized curriculum for undergraduate students majoring in bioinformatics has required an extensive national survey to find an optimal combination of form
and content. During this course of action, I have met with numerous colleagues from other universities, sponsored a session that I chaired at the 2003 O'Reilly Conference on Bioinformatics in San Diego, served on a panel for bioinformatics education at the 2003 Big XII research conference in Kansas City, participated in bioinformatics education workshops for the International Society for Computational Biology (ISCB), and submitted a peer-reviewed publication concerning bioinformatics education. One goal of these external conversations is to reinforce the need for communication in areas related to computational science education and to promote virtual communities such as genepedia.org to act as forums for researching bioinformatics curriculums, educational materials and future needs. Beginning in the fall of 2003 two new courses, BINF 3360, Introduction to Computational Biology, and BINF 4360, Bioinformatics Data Base Design, were added to an adjusted curriculum as a direct result of peer feedback. Starting in the spring of 2006, another new course, BINF 3350, Genomics and Bioinformatics, was added to the curriculum. This course is cross-listed with the Department of Biology as BIO 3350, and provides biology majors access to core bioinformatics concepts. A fourth course, BINF 3396, was added to the curriculum in order to allow students to gain credit for summer internships. The BINF course classification, implemented in the spring of 2006, has been extremely important in establishing bioinformatics as a unique discipline separate from that of life science or computer science.

In the fall of 2004 I proposed a Mater's of Bioinformatics degree program in the Department of Computer Science. While the department unanimously supported this proposal, faculty and logistical limitations has hindered its implementation. It is my hope that such a program will exist in the near term to increase the quality of students attracted to our university and provide the foundation for a larger diversity of graduate coursework and research.

As a co-investigator in a NSF-NIH pilot project in Bioinformatics and Education at the University of Tennessee, Knoxville, I had the tremendous opportunity during the summers of 2002 - 2005 to further enhance curriculum refinement. This institute is funded, in part, as a long-term study for understanding bioinformatics education needs, and consequently stipulates the creation of a Summer Institute for Advanced Bioinformatics Education. This is a collection of short bioinformatics educational workshops centered on discrete experimental modules. As the curriculum coordinator for the Summer Institute I have been able to leverage these modules for incorporation into the research experience here at Baylor; they have proven to be extremely portable and adaptable for undergraduate study.

Inside the classroom I am thrilled to say that my effectiveness as a teacher continues to improve. I am able to measure my progress with excellent teaching evaluations in the vast majority of my classes and I continually encourage all of my students to submit constructive criticism of both my teaching style and material. This on-going conversation with my students is possible due to the
comfortable environment that I intentionally promote in the classroom. Small class sizes aid in the effectiveness of this approach. To illustrate the impact that student feedback has on my self-evaluation and resulting modifications, I explain to every class that students influence the effectiveness of a teacher, and in particular the learning curve of a new instructor, through thoughtful suggestions and comments. As the result of both student and peer-review feedback, I have developed more hands-on tutorial packages that assist students in learning how to use the vast array of available bioinformatics tools. I have also limited the number of text books students are required to purchase in lieu of providing more classroom study material. In addition to standardized faculty evaluations, I have accepted course feedback through processes that I developed as part of my SACS participation. In this respect, I have been leading the effort to ensure that the B.S. in Informatics program meets and exceeds SACS standardization.

The didactic that I have gravitated towards includes one of lecture supplemented with small-scale experimental design and implementation. In the case of BINF 4360, a bioinformatics capstone course, the small implementations culminate in a large final project. During these projects, students have created local implementations of NCBI databases, created laboratory information management systems for biology researchers on campus, and have experimented with various models for genome viewing and browsing. A recent project, called Genepedia, is providing one data resource in my research efforts. One measure of the success of this approach can be found in standardized student surveys that reflect almost unanimous strong agreement that students learn a ‘great deal’ in class. Perhaps a better measure is the success that students are experiencing upon graduation. Out of 12 students in my class for the fall of 2002, for example, 5 applied for, and were accepted into, graduate school in my field of research. In the spring of 2005 nearly every graduating senior in my special topics class applied to graduate school in either medicine, biology, or bioinformatics and I have been honored to write numerous recommendation letters on their behalf. These factors, along with my increased ‘confidence of experience’ indicate that my quest to raise my effectiveness as a teacher has been successful.

Concurrent with education is my task to expand our fundamental understanding of God’s universe through disciplined research. To this end, I have maintained a very valuable and productive collaboration with the United States Department of Energy facilities at Oak Ridge, TN, the University of Tennessee in Knoxville, TN, the University of Tennessee Health Science Center in Memphis, TN, and Vanderbilt Medical School. These collaborations represent research into areas of mouse genetics, collaborative bioinformatics, and comparative genomics. Significant efforts with these and other collaborators since August of 2002 have resulted in eight peer-reviewed presentations, a first-author full-length manuscript in Leukemia Research and two in BMC Bioinformatics. On-going efforts with these collaborations have subsequently laid the foundation for peer-reviewed review articles on bioinformatics theory and the submission of seven extramural funding proposals.
Another powerful collaboration is between the University of Tennessee’s Department of Computer Science and Baylor’s bioinformatics lab. In conjunction with several graduate students, we have been utilizing concepts developed for logistical networking to assist in the scale-free distribution of biological databases across the internet. Mr. Ravi Kosuri, a former master’s student in my lab, presented this work in the Proceedings from the 12th International Conference on Intelligent Systems in Molecular Biology. Another former master's student, Mr. Nick Linn, furthered this research by designing and implementing a distributed FASTA program. The results of his significant contributions have been prepared as a full-length research manuscript and submitted for publication.

Perhaps more important are the collaborations that I have been able to build with faculty members here at Baylor. I am participating with Drs. Robert Adams, Myeongwoo Lee and Chris Kearney in the Department of Biology on separate but equally promising research areas into the comparative genomics of three taxonomically distinct families. Dr. Adams and I are interested in developing software that will identify genetic fingerprints of near-relatives in botanical families and developing a *poplar* and *juniperus*-specific genomic database. We currently have a small amount of funding to support this work, which has resulted in a publicly available online tool called Primer Analysis of Arabidopsis (PAA). Dr. Lee and I are utilizing undergraduate research assistants to isolate RNAI-dependant clustering of the *C. elegans* genome. This work forms the central part of a dissertation for a B.M.S. Ph.D. candidate named Mr. Sushil Batra. Sushil is the first Ph.D. candidate that I have mentored and the experience has been tremendous.

During the fall of 2004, I sought out and was appointed to an adjunct faculty position in Baylor’s Institute for Biomedical Studies (BMS). This allows me to participate in the curriculum develop of the program while simultaneously providing access to BMS Ph.D. students and post-doctoral fellows. For example, I am currently the direct mentor of two master’s candidates and one Ph.D. candidate. Overall, I have been involved in three Ph.D. committees and five master’s thesis committees. Working in this interdisciplinary area at Baylor has been truly rewarding.

I am also assisting Baylor’s forensic science program with an effort to curate DNA from unidentified individuals. The initial program, called *Reuniting Families*, was operational from the fall of 2002 through the spring of 2006. It garnered international attention because it was the first attempt to locate, identify and repatriate the remains of immigrants who have perished while trying to cross illegally into the United States. The on-line database, www.reunitingfamilies.org, provided an international service as well as a research reservoir for samples to study Latin-American population genetics, a traditional gap in the literature. In addition, this project has promoted the bioinformatics program at Baylor through intense media coverage in newspapers, such as *The Washington Post*, the *LA*
Among them, the Chicago Tribune, and USA Today, among others, and has been the subject of NPR’s Day to Day program and a special episode of National Geographic Extreme Explorer in March, 2004. A peer-reviewed presentation of the data base, co-authored with Dr. Lori Baker, has been published in the Journal of Forensic Science and American Association of Physical Anthropology, and presented at their respective national meetings in 2004. A full-length manuscript representing the data-matching algorithm is currently in press.

Early success with this program engendered the submission of a USAID/Mexico proposal for the creation of a joint U.S.-Mexican Forensic academy and database. This proposal, submitted in the summer of 2005, resulted in funding to Baylor’s Department of Anthropology, but the database section was declined. As a result, the Reuniting Families database was incorporated into the Mexican database for missing individuals, called SIRLI, during the spring of 2006.

Concomitant with research is the need for proper funding. Towards this end, I have been successful in obtaining several promising extramural awards. For the 2002-2005 academic years, I received a contract as co-investigator from an NIH-funded Center in Bioinformatics award totaling approximately $276,492.00. I have also received funding from the DOE program “Bringing Genomes to Life”, to create a distributed Shewanella data mining environment with colleges at ORNL. This award, tentatively scheduled for 2005-2008, will provide approximately $180,000.00 to Baylor. Also awarded recently through the NIH’s Integrative Neuroscience Initiative on Alcoholism (INIA) is a grant for the development of an Ontological Discovery System. This award provides $73,420.00 to Baylor from 2007-2009 to investigate novel bioinformatics approaches to studying phenotype and gene networks. To support the preliminary research needed to justify this research approach, I was awarded a $15,000 Baylor Faculty Research Improvement award in November, 2004. Finally, in conjunction with Dr. Keith Young of the Waco V.A. and Texas A&M, I am a contributing investigator in the discovery of SNPs (Single Nuclear Polymorphisms) in patients and military personal suffering from Post-traumatic Stress Disorder. This $3 million research effort is funded by the U.S. Congress and will begin in earnest during the spring of 2007, with the bioinformatics core beginning in the summer of 2008. The $60,000 awarded to Baylor from this award will be directed to the funding of a post-doctoral researcher and graduate students.

While not all of my grantsmanship efforts have been successful, I have spent a significant amount of time participating in this very critical area of research and collaboration. For example, while receiving positive reviews from the NSF, grants for the creation of a Summer Institute for Bioinformatics Education at Baylor and the NSF Young Investigator CAREER Award for 2002 were both declined. Also declined were a secondary proposal to the Young Investigator Award, 2003, and a proposal for the Alfred P. Sloan Fellowship. During this same time period one additional proposal to the NIH and two proposals to the NSF also failed to receive funding. The latest declined proposal to the NSF, however, received
rather high scores and was re-submitted during the summer of 2007. Regardless of these proposal results, I am fully committed to the appropriation of external funds to enhance Baylor’s research-orientated goals.

In order to become fully engaged in the many opportunities for service at Baylor, I have participated in several meaningful ways as a part of, and advocate for, the Baylor family. I served as an inaugural member of Baylor University’s Institutional Biosafety Committee, am currently chairing the departmental subcommittee on bioinformatics curriculum development, serving on the departmental committees for graduate education and student retention, and serving as bioinformatician for Baylor’s DNA sequencing center. In this later role, I am responsible for coordinating software that analyzes output from Baylor’s automated sequencer and, as a direct result, have received permission to establish a bioinformatics presence in the new science building. In addition to research and training, the science building-based bioinformatics lab serves as a community resource for Baylor researchers and researchers-in-training. We offer our expertise as a service for the Baylor community.

I am also on the Departmental Search Committee and the University Research Committee. The later committee oversees the funding of university grants and requires a substantial amount of time.

I have had the great pleasure of meeting parents during the parent-faculty coffee, participating as a marshal at most commencements, lecturing at premiers, and participating in the new faculty mentor program. During the spring of 2005, I also had the honor of giving the Breakfast at Baylor seminar along with my wife, Dr. Lori Baker. I was honored to be profiled as a featured faculty member in Baylor honor’s college student magazine, The Pulse. During 2003 I became the inaugural student organization co-advisor for the Baylor Pre-Physician Assistant Society and in 2004 became the inaugural advisor for the Baylor Society for Bioinformatics and Bio-Technology. This society brings me into contact with undergraduate bioinformatics majors on a regular basis, promoting student retention by creating a forum where freshmen and sophomores may hear guest speakers, discuss curriculum issues, explore the employment landscape and exchange ideas about our profession.

I have also attempted to extend working collaborations with other colleagues in Texas, including Southwest Medical Center, Baylor College of Medicine, and the San Antonio Health Science Center. It is my hope that this communication along with my professional service that includes session moderating at national meetings, presenting papers in front of the academic community, like the Tennessee Mouse Genome Consortium in Nashville, TN, and attending symposia on cutting edge research (Keck/GCC 2002 and 2005 Bioinformatics Symposium, 2005 UGA Symposium on Computational Biology) will build the fruitful collaborations necessary to support a self-sustaining research initiative, ultimately serving the whole of the Baylor community.
As further service to the academic community, I was asked to serve on an NIH-NIAID review panel for a national grant applications program empanelled in December, 2003, titled “Administrative Resource for Biodefense Proteomic Centers”. During this time I have also served on bioinformatics program committees for two national meetings (IEEE AINA-07, ICCS). These responsibilities represent a significant time and effort commitment to service as well as a positive step in advancement of academic qualifications.

Sincerely,

Erich J. Baker