Final Report for Period: 09/2008 - 08/2009
Submitted on: 12/01/2009
Principal Investigator: Borodovsky, Mark .
Award ID: 0531908
Organization: GA Tech Res Corp - GIT
Submitted By:
Borodovsky, Mark - Principal Investigator
Title:
UBM: Quantitative Systems Biology

Project Participants

Senior Personnel

Name: Borodovsky, Mark
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Choi, Jung
Worked for more than 160 Hours: No
Contribution to Project:

Name: Bunimovich, Leonid
Worked for more than 160 Hours: No
Contribution to Project:

Name: Goldstein, Gulliermo
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Goldsztein, Guillermo
Worked for more than 160 Hours: Yes
Contribution to Project:

Post-doc

Graduate Student

Undergraduate Student

Name: Kuhns, Sarah
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Saxton, Carina
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Galvin, Michael
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Dias, Alan

Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Lee, Byron
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Warden, Charles
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Park, Katherine
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Bolduc, Aaron
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Portz, Brent
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Sedlack, Rachel
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Bucharin, Vasiliy
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Wilson, Christine
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: May, Corwin
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Taing, Lee
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Coch, Chris
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Bruce, Marc
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Friend, AJ
Worked for more than 160 Hours: Yes
Contribution to Project:
conducted research on 'Copy Number Effects in Gene Regulatory Network Motifs' with Prof J. Weitz

Name: Fisher, Sofia
Worked for more than 160 Hours: Yes
Contribution to Project:
conducted research on 'Theorizing Ahead of the Data: Comparing Archaeal Virus Models' with Prof J. Weitz

Name: Lim-Hing, Krista
Worked for more than 160 Hours: Yes
Contribution to Project:
conducted research on 'Nuclear Magnetic Resonance Based Whale Shark Metabolomics' with Prof. Julia Kubanek

Name: Santomauro, Giles
Worked for more than 160 Hours: Yes
Contribution to Project:
conducted research on 'Decoding Algorithm Analysis of Alpha-helical Transmembrane Protein Topology Prediction Models' with Prof. Gueliermo Goldzstein

Name: Spears, Kevin
Worked for more than 160 Hours: Yes
Contribution to Project:
conducted research on 'A One Dimensional Model of Krill Swimming' with Profs. Silas Alben and Jeannette Yen

Name: Nowack, Tim
Worked for more than 160 Hours: Yes
Contribution to Project:
conducted research on 'Transposition of mtDNA prtein sequences into mammalian genomes' with Prof. Soojin Yi

Name: Patidar, Vainiya
Worked for more than 160 Hours: Yes
Contribution to Project:
conducted research on 'Bile Acids Regulate Cortisol Production in Adrenocortical Cells' (Why Eating Does Late at Night Lead to Increased Weight Gain?) with Prof. Marion Sewer

Name: Vanderbleek, Sandy
Worked for more than 160 Hours: Yes
Contribution to Project:
conducted research on 'Developing an HMM based algorithm for finding frameshifts in DNA sequences' with Prof. Mark Borodovsky

Name: Tolson, Chanin
Worked for more than 160 Hours: Yes
Contribution to Project:

Technician, Programmer

Other Participant

Research Experience for Undergraduates

Organizational Partners

Moscow State University

As an exchange with the foreign undergraduate program we sent one
undergraduate student (Michael Galvin, Math major) to Department of Bioinformatics at Moscow State University, Moscow, Russia for two months in summer of 2007, advised jointly by Georgia Tech and MSU professors. This student also attended International Conference on Bioinformatics in Moscow in July 2007.

J. Craig Venter Institute

hands-on experience visit by Chanin Tolson

Other Collaborators or Contacts

Activities and Findings

Research and Education Activities:

In this grant period we have had just $5,000.00 left which was used
for the summer research by Chanin Tolson the 2nd undergraduate student in life sciences.
Her research was supervised by Dr. Borodovsky and Dr. Lobachev

During summer 2009 Chanin Tolson also visited J.Craig Venter Institute of the hands-on experience in genomics.

Findings:

Chanin was working on a project of modeling molecular evolution, which temporal dynamics defines the similarity functions for orthologous proteins. Having better understood the concepts of similarity search between DNA and protein sequences, she used bioinformatics similarity search tools to identify the proteins that would have a resolvase function in eukaryotic species.

Training and Development:

Conducting interdisiplinary research in Summer of 2009 was a valuable experience for our student, Chanin Tolson. We saw her growing in understanding on how research projects in genomics are performed, and working hard on the assignments.

Another student earlier working in the program, Marc Bruce, has prepared a manuscript based on his NSF supported research: 'Gibbs sampler for finding functional sites in inhomogeneous backgrounds' The manuscript is to be submitted for publication this month.

Marc Bruce was accepted for PhD studies at Stanford University.
Another former student at the UBM program - Mike Galvin,
the Iraq war veteran, was accepted to the graduate program at Computational Science and Engineering.

Outreach Activities:

To increase the visibility of the program
we have set up the web pages devoted to the poster sessions

http://opal.biology.gatech.edu/GeneMark/math_bio_06/
http://opal.biology.gatech.edu/GeneMark/math_bio_07/
http://opal.biology.gatech.edu/GeneMark/math_bio_08/

Journal Publications

Kim, SH; Elango, N; Warden, C; Vigoda, E; Yi, SV, "Heterogeneous genomic molecular clocks in primates", PLOS GENETICS, p. 1527, vol. 2, (2006). Published, 10.1371/journal.pgen.002016

Books or Other One-time Publications

Web/Internet Site

URL(s):

http://opal.biology.gatech.edu/GeneMark/math_bio_08/

Description:

The site shows photos of the poster session, the final event of the Sumeer 2008 research activities

Other Specific Products

Product Type:

Software (or netware)

Product Description:

Marc Bruce presented both a talk and a poster ?Finding Biological Signals in Inhomogeneous DNA Sequences by a Gibbs Sampler Algorithm? at the 2008 SIAM Conference on the Life Sciences, August 4-7, 2008, Montreal, Canada.

Marc developed a software for the enhanced Gibbs Sampler algorithm that is able to align inhomogeneous nucleotide sequences that carry evolutionary conserved signals (such as branch point sites)

Sharing Information:

The software will be freely available for academic researchers
after publishing the journal paper.

Contributions

Contributions within Discipline:

Research experience of undergraduate students rarely generates noticable contributions in any scientific field including the field of quantitative biology. Still we had a case of a significant contribution reported in the 2007 (Charles Warden, co-author in the paper published in PLoS Biology)

Marc Bruce has prepared a manuscript for publication 'Gibbs sampler for finding functional sites in inhomogeneous backgrounds'
to be submitted to Nucleic Acids Research

Contributions to Other Disciplines:

Contributions to Human Resource Development:

All Biology major undergraduate students at Georgia Tech have studied the newly designed Calculus course (2 semesters) with a focus on biological examples.

The Summer undergraduate program in 2006-2009 provided extensive research experience for twenty five best Biology and Math Undergraduates; their interdisciplinary research was crossing the lines of Biology and Math

Contributions to Resources for Research and Education:

Contributions Beyond Science and Engineering:

Conference Proceedings

Categories for which nothing is reported:

Any Book
Contributions: To Any Other Disciplines
Contributions: To Any Resources for Research and Education
Contributions: To Any Beyond Science and Engineering
Any Conference

Progress report (2006-2007)

ID 0531908 2005-2008
Undergraduate Biology-Mathematics: Quantitative systems biology. (QSB)
PI: Borodovsky

Undergraduate students involved in summer research have completes the following projects:

Computer modeling of krill schooling behavior (**Sarah Kuhns** Prof. Jeannette Yen), and
Effects of oceanic turbulence on deep chlorophyll maxima (**Carina Saxton** Prof. Emmanuele Di Lorenzo)
Yeast two-hybrid analysis of the Che3 signal transduction system in Myxococcus xanthis (**Aaron Bolduc**, Prof. John Kirbi),
Survey of various spider venoms to test for sphingomyelinase D activity simial to Loxsceles recluse (**Katherine Park**, Prof. Al Merrill)
Analysis of heterogeneity of evolution of mammalian genomes (**Charles Warden**, Prof. Soojin Yi),
Improvement of Gibbs sampling simulated annealing step (**Michael Galvin**, Prof. Mark Borodovsky)
Visualizing of pericellular coat of living cells (**Lee Taing** Prof. Jennifer Curtis),
Linking leaf allometry to vascular network geometry (**Christine Wilson**, Prof. Joshua Weitz),
Finding biological signals by Gibbs sampler algorithm with multiple backgrounds (**Marc Bruce**, Prof. Mark Borodovsky)
Microarray analysis application in sphingolipid methabolic pathways (**Brent Portz**, Prof. Al Merrill)
Examining the isotope fractionation of stable nitrogene and carbon isotopes through digestion in cichlilds (**Rachel Sedlack**, Prof. Josheph Montoya)
Phage derived alternatives to antibiotica (**Corwin May**, Prof. Joshua Weitz)
Flow field analysis of Euchaeta elongate using particle image velocimetry (**Chris Coch**, Prof. Jannette Yen)
Novel Analysis of CFTR Single Channel Kinetics (**Vasiliy Buharin**, Prof. Nael McCarty)

The work on the project Analysis of heterogeneity of evolution of mammalian genomes - by Charles Warden made a critical contribution to a high profile publication in PLOS Genetics

Charles Warden worked on the following question: what is the effect of mutations at CpG dinucleotides on mammalian molecular clock? Because CpG dinucleotides are targets of DNA methylation in mammalian genomes, they are subject to methylation-origin mutations. Such mutations are known to occur much more frequently than other mutations (Our lab has shown that in primate genomes, they occur over one order of magnitude more frequent than others). Therefore, genomic regions with more CpG dinucleotides may evolve faster than other genomic regions.

Previous studies have analyzed molecular evolution of protein-coding genes in humans and Old World monkeys, and have shown that genes in Old World monkeys generally evolve faster. The explanation for this is the following: because Old World monkeys undergoes more generation times than humans in a given amount of evolutionary time, there are more mutations in Old World monkeys. This is referred to as 'generation time effect'. Interestingly, studies in our laboratory have found that mutations at CpG dinucleotides often do not show such generation time effect. This is expected since DNA methylation does not occur dependent on generation times.

Thus, Charles hypothesized that if we remove CpG dinucleotides from protein-coding genes, we will see more pronounced generation time effect. He has compiled over 70 genes from human, an Old World monkey (baboon or rhesus macaque was used), and a New World monkey (to be used as an outgroup). Then, he compared generation time effect when CpG dinucleotides are not removed, and when they are removed. He found that the former had about 10% rate difference, and in the latter comparison the rate difference was about 30%. Therefore, his hypothesis turned out to be correct.

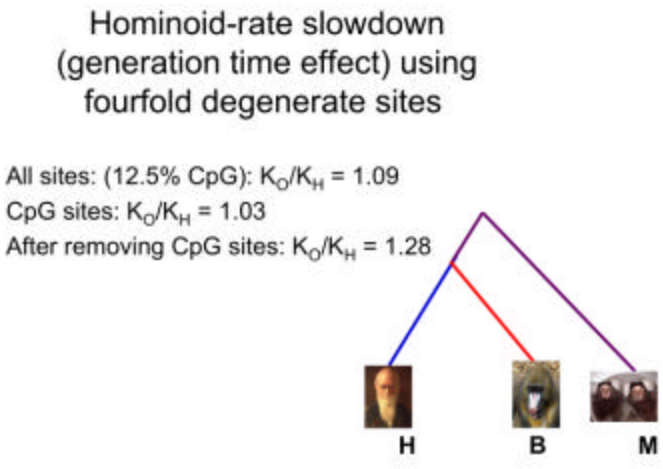

Figure1. A summary of Charles Warden's finding. The rate difference between humans and Old World monkeys is also referred to as 'hominoid-rate slowdown'. He has shown that sites including CpG dinucleotides show different degree of hominoid rate slowdown compared to the CpG dinucleotides only, or when CpG dinucleotides are removed.

Amazingly, this result is in an excellent accord with a theoretical analysis of the effect of mutations at CpG dinucleotides and molecular clock. My laboratory has calculated, based upon estimated parameters, that if the true generation time effect causes about 40% difference between humans and Old World monkeys, then sites in protein-coding genes will show approximately 10% rate difference (shown in the Figure). This is exactly what he found. This result is included in the Table 3 of the resulting paper, "Heterogeneous Genomic Molecular Clocks in Primates", which was published in the journal PLoS Genetics in the year 2006.
The analyses by Charles had added empirical support to the hypothesis that CpG dinucleotides follow relatively constant molecular clock. This has shed important lights on the controversy

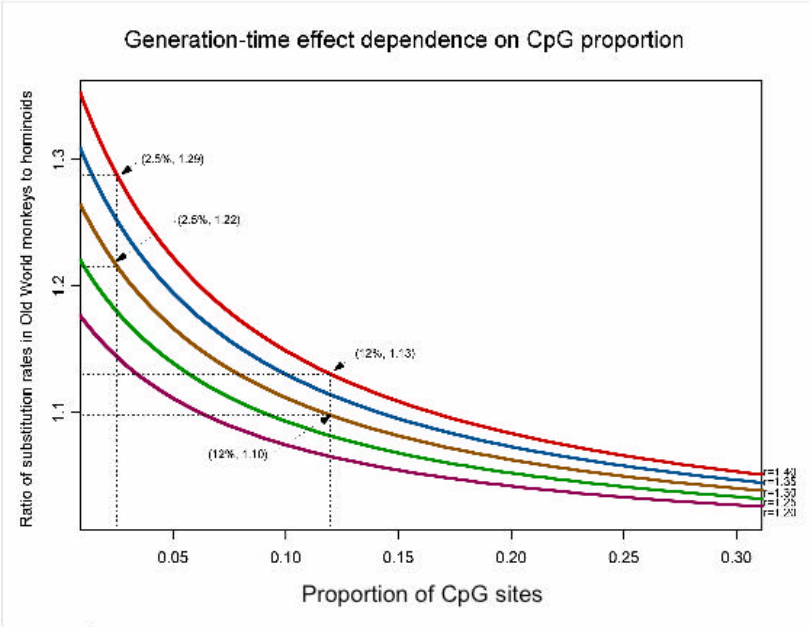

Figure 4.

surrounding whether genomic molecular clocks are constant or generation time dependent. This article has been cited several times in database, and several researchers in the field has expressed that this work made a substantial contribution to the field. Certainly, the support from the NSF allowed Charles and my laboratory to make an important discovery.

Other projects were successful on their own, but have not produced publications yet, as can be seen from the feedbacks from faculty.

Lee Taing's work has been presented on a poster at the 2008 Biophysical Society in Long Beach, CA. (her name was on the abstract), and a small portion of it will be included in a talk that I will give at a symposium in Paris next week called 'Optical Microscopy in Good Shape'.
http://www.biomedicale.univ-paris5.fr/neurophysiologie/omgs/program.html#session4

The work has been since continued by a graduate student and another undergraduate student (trained by Lee). In another 6 months it should lead to a contribution to a peer reviewed journal article. Lee has already done enough to merit her name on the publication.

Byron's Lee research contribution was acknowledged on the "Evidence for stage-specific modulation of microRNAs and miRNA processing components in zygotic embryo of loblolly pine" T. Oh, R.M. Wartell, J. Cairney, and G. Pullman New Phytologist 179,67,2008

Christina Wilson work contributed to "Scaling of leaf network features", manuscript in preparation. Christina will be acknowledged, or may be a co-author depending on the way the manuscript shapes up.

Similar level contributions were made by all other students

Additionally, the following questions are addressed below.

1. How are the activities of your project preparing undergraduate mathematics and biology students for careers and further academic work in fields at the intersection of the biological and mathematical sciences?

The new two semester Calculus course for Biology undergraduates includes biological objects into direct consideration. The course emphasizes connections between quantitative biology and mathematical approaches. The course introduces biologists to mathematical thinking in the easier context than it was done by a general Calculus course.

Summer research activity has already included 24 undergraduate students who have conducted research with Georgia Tech professors at the interface of Biology and Mathematics.

2. What successes have you experienced? What are the challenges you have faced when implementing your UBM program and what actions have you taken to meet these challenges? Have you made any significant changes or deviations from the planned grant activities? Why?

The Calculus course was developed and already taught for two Fall and two Spring semesters (Prof. Goldztein). Major and quite anticipated challenge we experienced in recruiting students for Summer research. We have much smaller number of Math students than Biology students. We found that Math students are less interested in interdisciplinary research than Biology students.

3. How are you recruiting and selecting your UBM students? Are you getting the number, quality, and diversity (ethnic, gender, and major) of students you want? When in their undergraduate course of study (e.g., freshmen, sophomore, junior, senior) do you recruit UBM students and when does their research experience occur?

Recruiting students goes for about two months in each Spring semester. The program publicity is given much attention. We do not make special efforts to increase diversity however the program usually has about equal representation of both genders and if we see an opportunity to attract ethnic minority we try our best to make it happen. Still we give priority to academic credentials in the student admission. We also give preference to sophomore and junior students, who already have sufficient background and who will spend yet some more time at Georgia Tech.

4. How are you selecting the research projects and mentors? What are your successes and challenges?

Most of the time we allowed the accepted students to contact professors that do the type of research that students are interested in and if the contact ended up in a mutual interest in doing a joint research work in the Summer we approved such an plan.

5. What courses and other activities are you requiring/recommending/advising students to take to prepare them for the UBM supported research experience? Has this changed from your original plans and if so, why? Specifically, how are you preparing (a) mathematics and (b) biology students to work on joint research projects?

Since selection of the students takes place in the Spring semester preceding the Summer research, we do not have much room for special preparation for the summer research in terms of recommendations for particular courses. However, the Calculus for Biology course is a requirement for all the Biology undergrads and it certainly makes a difference for students who are interested in mathematical biology.

6. Specifically, what are the contributions of the (a) mathematics and (b) biology students to the research projects? What does each type student contribute to and get from their work on the project?

Biology students are capable of doing experiments, however they are interested also in theoretical part and data processing part. Math students rarely get to the lab work, they become exposed to work with real data and models of biological objects which is a fascinating experience.

7. What cohort and other activities are your students engaged in while they are working on the research projects? How are your students writing up and where are they presenting their research project activities and results?

In addition to research work, the students attend a seminar where they either present their work at difference stages or present papers related to their research. This seminar help bring them together and keep as a group. It help develop their presentation skills. In the final week of summer research the students prepare posters and present them at the special open to the whole university poster session.

8. What measures and metrics are you using to monitor and assess the progress of your project? How many students have been impacted by each of the activities in your project? What is the nature of the impact? What are the demographics (ethnicity, gender, and major) of the participants?

The calculus course in each year is taken by 60-70 biology undergraduates. The summer research is conducted by 8 students each year. The letter grades are standard measures of success in the calculus class. The summer research is stimulated and motivated by the peer group as well as the success of the final poster presentation.

9. Have the UBM program activities been integrated within your department(s) and institution? How have the activities influenced the direction of your academic programs for a broad range of students? Are there similarities and differences between biology and mathematics departments in challenges, approaches, and outcomes in attempting institutional changes? Do the UBM classes and research experiences count towards the majors of mathematics and biology students?

The Calculus for Biologists course has been integrated into the Institute curriculum and is a requirement for Biology Major at the School of Biology.

10. Has this program had an impact beyond the intended project goals, for example on other students, faculty, departments, or institutions? What actions have you taken to address the sustainability of the project beyond the grant period?

While the summer research program helped undergrad students test their interest and abilities for conducting research it is difficult to assess its impact to other students. However, in one instance we had an undergraduate student, Michael Galvin, a Math major, and former marine, attending Moscow State University, Department of Bioengineering and Bioinformatics working on a problem of finding signals in biological sequences. Michael as an ambassador of Georgia Tech influenced academically outstanding students from MSU to apply for PhD studies at Georgia Tech.

The Calculus course will stay at Georgia Tech beyond the grant period. However, the sensitive issue of undergraduate research funding is difficult to solve with private sponsors, thus, we will apply for the grant extension by the next year deadline.

I should say that making teams of math and biology students was difficult. Creation of such a team requires an advisor or two advisors that have synergy and a right project at a right time. A concept of an ideal team of bright biology and math students working with a professor with biology and math expertise or two professors one from math and another from biology and able to speak the same language assumes that many happy coincidences will happen. This concept, so far did not materialized in reality of our campus. Still, we tried our best to start moving to this model.