Khuri-Bortolazzo

Undergraduate Research Grants Application Form

Due May 7, 2012

1. Student Information

First Name: Anthony | Last Name: Bortolazzo

Major: Biological Sciences | Student ID#: [redacted]

Email Address: [redacted] | Phone: [redacted]

Class Standing: ○ Freshman ○ Sophomore ◉ Junior ○ Senior

GPA [redacted] Have you been the recipient of this award before? ○ Yes ◉ No

2. Faculty Information

First Name: Sami | Last Name: Khuri

Email Address: sami.khuri@sjsu.edu | Phone: 924-5081

Department: Computer Science | College/School: Science

3. Project Information

Title of Project: Predicting Cryptic Splice Sites in Human Genes

Project Timeframe: Start Date 09/01/12 End Date 05/12/13

Will this project use (check all that apply)

- ☐ Animal Subjects
- ☐ Biohazards/Human Blood
- ☐ Human Subjects
- ☐ Recombinant DNA
- ☐ Radiation/Isotopes/Lasers
- ☐ Controlled Substances

TO BE COMPLETED JOINTLY BY THE STUDENT AND THE FACULTY MENTOR:

In a few sentences, describe the goals of the research or creative project you will be working on together.

This project has 2 main goals: a scientific goal as well as an outreach goal.

1) Scientific goal:
We will study point mutations of the beta globin gene that give rise to cryptic splice sites creation and study the consequences these sites have on protein synthesis. We concentrate on point mutations of the beta globin gene that are responsible for beta Thalassemia, a genetic blood disorder that causes severe anemia that requires routine blood transfusions. The goal is that to develop a statistical model, position weight matrix, that can predict cryptic splice sites in the beta globin gene. The model will be trained on known point mutations that lead to cryptic splice sites. We will use our model to predict cryptic splice sites in other annotated human genes. This is possible because the splicing machinery, known as the splicesome, recognizes sequences found at intron boundaries that are conserved across species.
2) Outreach goal:
Tony and I will prepare a lecture and hands-on exercises for high school students that use the beta globin gene and some of the mutations that lead to beta Thalassemia as examples of how mutations in the DNA (genotype) can lead to severe consequences (phenotype). The study of the various thalassemia syndromes can serve as a paradigm for gaining insights into the factors that can regulate or disrupt normal gene expression. Beta Thalassemia is a good example since it is a monogenic disease caused by the beta globin gene which is relatively small in size.

Additionally, we are going to create a website containing the background describing the Beta Thalassemia disease, the model for predicting the cryptic splice sites and all the educational materials mentioned above.

Enumerate the general activities the **student** will perform as part of their participation in this project. Provide a rough timeline.

Tony will start by reading several book chapters and papers on hemoglobinopathies since beta Thalassemia belongs to that group of diseases. He is well prepared for the task at hand since he has taken most of the genetic courses, namely Biol115 and Biol117. He has also taken the Bioinformatics I course, Biology/CS 123A, in Fall 2011.
He will then shift his attention to computational models and tools for representing and analyzing cryptic splice sites. He will become proficient in multiple publicly available splice site analysis tools: SplicePort, and GeneSplicer . More precisely:
First month:
a) Chapter 3 of "Variant Haemoglobins: A Guide to Identification" entitled: "thalassaemias and related conditions". The book has four chapters. The second book, "Hematology in Clinical Practice" will be used as a reference.
b) Papers on cryptic splice sites. These are recent articles that I have compiled over the last 8 years.
c) Two articles on position weight matrices that are commonly used in bioinformatics to represent patterns (and profiles) found in DNA and proteins. In our case, it will be DNA (or RNA) cryptic splice sites, i.e., splice sites that are not normally used by the spliceosome due to a point-mutation in the DNA.
Second month:
a) Compiling all the beta globin gene mutations (leading to beta Thalassemia) that create cryptic splice sites and modelling them with a position weight matrix (PWM).
b) Sending a summary of our findings to the CSUPERB annual conference (Jan 2013) for possible poster presentation.
Remaining months:
Perfecting the PWM model by using pseudocounts and log-odds scores and extrapolating our method to predict cryptic splice sites with other, annotated genes that lead to diseases.

Enumerate the general activities the **faculty mentor** will perform as they supervise or guide the student throughout this project.

I plan to have biweekly meetings with Tony. He will be part of the research group of students (most of whom are graduate students working on their Master's Thesis under me) I meet with every week. I will ask Tony to attend every other week and give 30 minute presentations on his readings. Namely:
a) "Thalassaemias and related conditions": Chapter 3 from "Variant Haemoglobins: A Guide to Identification" (mentioned above).
b) Two articles on cryptic splice sites.
c) Two articles on position weight matrices that are commonly used in bioinformatics to represent patterns (and profiles) found in DNA and proteins.

I will set up meetings with the Dr. Vichinsky in Oakland and with Dr. Kan and Dr. Weinkam from UCSF to discuss various aspects of our research.

Work on the paper we will submit to CSU Program for Education and Research in Biotechnology (CSUPERB) annual symposium (January 2013) for possible poster presentation.

I will then assign additional computational papers to understand how the PWM model can be enhanced by using pseudocounts and log-odds scores. We will meet several times (one on one) for that purpose. The ultimate goal is to have a publication (or submission to a refereed conference) by May 2013 (the end of the grant period).

How will the student's participation in this project contribute to their educational and/or career goals?

The undergraduate student, Anthony Bortolazzo, is hard-set in continuing on to graduate school to earn his doctorate in Biology. His research interests are of gene regulation – more specifically, the role of RNA in regulating gene expression.
Aside from Tony's research and course studies, he's also becoming more active in chemistry and biology extracurricular activities, such as public outreach. This research project is perfect for Tony for it will benefit his continuing education in the following manner:

- This project closely follows and researches the genetic basis for β-Thalassemia, a disease whose basis couldn't be more related to Tony's research interests.
- Bioinformatics tools are very important for biological research. Tony will deepen his Bioinformatics and statistics skills during this project.
- By interacting with experts in the study and/or treatment of β-Thalassemia, this project will give Tony a wider aspect of human genetics, outside of his studies at San Jose State University. More importantly, Tony will be given the opportunity to cross this genetic research with β-Thalassemia's real life implications.
- Tony will present his work at the 2013 CSUPERB Symposium (January 2013). This will be an invaluable opportunity for him to interact with faculty and fellow students.
- This project also gives Tony the opportunity to conduct reach-out activity in the awareness of β-Thalassemia and its genetic causes.

4. **Budget and Justification** (Budget itself to be submitted on Budget Form)

Explain how the funds requested for this project will be expended. Grant funds can be used to cover project related software, equipment, materials, supplies, travel, data collection and analysis, and other project related expenses. Up to one half the grant can be used as student assistant funds. Funds may not be used for faculty compensation or travel or to purchase food.

Grant funds will be used in the following way:

1) The student will get a stipend of $500 for his work on this project.

2) We shall purchase two books that will be used in the project:
Book One: Hematology in Clinical Practice, 5th Edition, by Robert S. Hillman, Kenneth A. Ault, and Henry Rinder [softcover] 2010. Pages: 560. ISBN: 9780071626996.
Book Two: Variant Haemoglobins: A Guide to Identification, 1st Edition, by Barbara J. Bain, Lorraine A. Phelan, Adrian D. Stephens, Barbara J. Wild [hardcover] 2011. Pages: 254. Publisher: Blackwell Pub ISBN: 9781405167154.

3) We plan to take day trips to consult with:
* Dr. Elliott Vichinsky, Head of the The Hemoglobinopathy Center at the Children's Hospital Oakland Research Institute.
* Dr. Y.W. Kan, MD, Departments of Medicine and Laboratory Medicine, UCSF
* Dr. Patrick Weinkam, post doc in Andrej Sali lab at UCSF.

5. **Signatures**

Student Signature [redacted] Date 05/07/12

Faculty Signature [redacted] Date 05/07/12

Department Chair Signature _____ Date _____ SK for chair

Could not get electronic signature

Save & Print Form

Submit completed application with supporting documents to the Center for Faculty Development (IRC 213) or email to cfd@sjsu.edu no later than **5:00 p.m. Monday, May 7, 2012**.

Undergraduate Research Grant Budget Form

Project Title
Predicting Cryptic Splice Sites in Human Genes

Budget item	Quantity	Unit Cost	Total Cost
Supplies			
			$0.00
			$0.00
			$0.00
			$0.00
			$0.00
			$0.00
Travel			
Oakland	2	$50.00	$100.00
UCSF	4	$50.00	$200.00
			$0.00
			$0.00
Other			
Hematology in Clinical Practice	1	$70.00	$70.00
Variant Haemoglobins: A Guide to Identification	1	$130.00	$130.00
Student Assistant Funds Specify number of hours and hourly rate	50	$10.00	$500.00
Total Costs (may not exceed $1000)			$1,000.00

Reset Form Save & Print Form

Include separate lines for software, equipment, materials, supplies, travel, data collection and analysis. You may also use the "other " line to insert additional lines as needed.