Preparing Underrepresented and First-Generation Students for Careers in Biomedical Big Data Science

Judith Canner, California State University, Monterey Bay
Dr. Carla Fresquez, California State University, Monterey Bay

Available at: https://works.bepress.com/judith-canner/1/
Preparing Underrepresented and First-Generation Students for Careers in Biomedical Big Data Science

Dr. Judith E. Canner (PI), Dr. Carla Fresquez

The BD2K Scholars Program @ CSUMB

Our first summer research program recruited students who were already "prepared" for the research program. Over the next year we will be developing a more intensive recruiting program starting with students in their first year at CSUMB (Table 1).

- 2016 Program dates were June 19–August 12 (8 weeks)
- Stipend, travel, and housing support is covered by the program.
- Program targets junior-level (rising senior) CSUMB students or students with 1 year remaining in their undergraduate education.
  - 3 Computer Science (2 Females, 2 URM)
  - 2 Mathematics (1 URM and First-Generation)
  - 3 Biology (2 URM, 1 First-Generation)

Scholars participated in a professional development seminar, applied for conference presentations (3 students accepted to SACNAS), and participated in training workshops. All students presented their research at the UCSC Summer Research Symposium. Scholars will apply to additional professional conferences, prepare/apply for graduate school, and four students will continue their research experiences during the 2016-2017 academic year.

In keeping with the CSU Vision, the majority of our program and course development focuses on undergraduate-level curriculum. The courses below are new and supported by the BD2K initiative and supplement already existing programs in fields related to Biomedical Data Science.

**New Courses Offered 2015-2016**
- BIO 380: Practical Computing for Scientists
- BIO 410: Bioinformatics
- CST 463: Data Mining
- STAT 210: Statistical Computing Seminar

**New Courses Planned 2016-2017**
- BIO 472: Comparative Genomics
- STAT 3XX: Statistical Simulations
- CST 4XX: Machine Learning
- PHY 1XX: Introduction to Inquiry

**New Programs 2015-2017**
- Data Science Concentration (Computer Science B.S.)
- Statistics Major (B.S.)
- Bioinformatics Concentration (Biology B.S.)
- Data Science Integrated Minor

---

### Table 1: Example Timeline of BD2K Scholar Training, Research, and Professional Development Experiences

<table>
<thead>
<tr>
<th>Year</th>
<th>Fall</th>
<th>Spring</th>
<th>Summer</th>
<th>Fall</th>
<th>Spring</th>
<th>Summer</th>
<th>Fall</th>
<th>Spring</th>
<th>Summer</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
</tr>
<tr>
<td></td>
<td>BSU Graduate Data Science</td>
<td>R25MD010391 Co-PI</td>
<td>Pre-Doctoral Summer Research Program</td>
<td>RSCH 200</td>
<td>RSCH 200</td>
<td>Pre-Doctoral Summer Research Program</td>
<td>RSCH 300</td>
<td>RSCH 300</td>
<td>Pre-Doctoral Summer Research Program</td>
</tr>
<tr>
<td></td>
<td>Python, R, MATLAB</td>
<td>Machine Learning</td>
<td>Data Mining, Bioinformatics</td>
<td>Data Mining, Bioinformatics</td>
<td>Genomic Data Analysis</td>
<td>Genomic Data Analysis</td>
<td>Genomic Data Analysis</td>
<td>Genomic Data Analysis</td>
<td>Genomic Data Analysis</td>
</tr>
</tbody>
</table>

### Examples

- Dr. Cheryl Logan (Biography) participated in an online training program to teach undergraduates advanced sequencing methods.
- Dr. Nite Jue (Biography) is pursuing research collaborations in evolutionary biology and cancer research in genomics.
- Dr. Judith Canner (Statistics) participated in training on supervised and unsupervised statistical learning methods and is pursuing research opportunities for her upcoming sabbatical.

### Upcoming Events

- Data Carpentry Workshop
- Genome Browser Workshop
- BD2K Speaker Series
- Summer Research Program Cohort 2 Community College and High School Recruitment Events

### Acknowledgements

Funding for the BD2K Biomedical Data Science program at CSUMB is supported by the Office Of The Director (OD) of the National Institutes of Health (NIH) under grant number R25MD010391. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

---

Category: Research and Professional Development Experiences

**Kara Spencer**, Computer Science B.S.

**Juan Cerda**, Biology B.S.

**Daniel Meza-Cortez**, Mathematics B.S.

---

California State University, Monterey Bay

CSUMB is envisioned as a comprehensive state university which values service through high quality education. The campus will be distinctive in serving the diverse people of California, especially the working class and historically undereducated and low-income populations.

- Four-year public university, a part of the California State University System
- Located in Monterey County, California
- 7280 students
- Hispanic Serving Institution (37%)
- 62% Women, 38% Men
- 55% First-Generation
- 35% Low-Income
- 47% Unders-Represented Minorities

Program Goals

The goals of the Biomedical Data Science Program at CSUMB, through funding from the NIH, are

- To establish a summer research program for CSUMB students at the Center for Big Data in Translational Genomics at UC Santa Cruz. The center works to help the biomedical community use genomic information to better understand human health and disease. Visiting CSUMB students will spend the summer working side-by-side with UCSC scientists and data specialists, learning research skills to manage and interpret genomic data.
- To develop new programs, such as an interdisciplinary statistics major, that will include math, statistics, biology, behavioral and computer sciences to prepare students for graduate school and careers in research or industry.
- To create opportunities to extend CSUMB faculty training and research in biomedical data science in collaboration with UCSC faculty members and researchers.

---

California State University MONTEREY BAY