UC San Diego
UC San Diego News Center

August 17, 2015 | By Doug Ramsey

Bioinformatics Pioneers Launch First Online Bioinformatics Specialization on Coursera

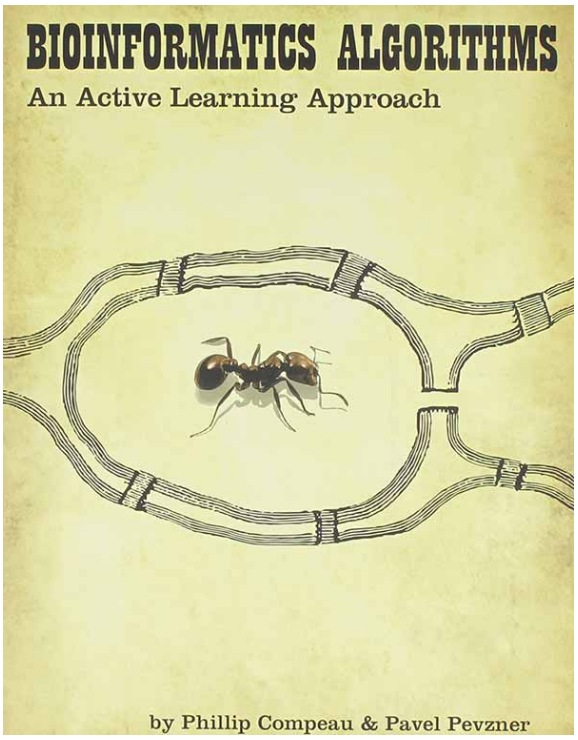
CSE Professor Pavel Pevzner (left) and former postdoc Phillip Campeau introduce the Wild West of bioinformatics in seven-part series of courses on Coursera.

Learners around the world will have the opportunity to enroll in a series of courses designed for biologists eager to gain computational skills and for computer scientists who want to explore the frontier of bioinformatics. UC San Diego will launch its six-course Specialization in Bioinformatics on Coursera, which culminates in a Capstone Project using software tools and big data provided by Illumina, a leading company in genome sequencing and the emerging field of personalized medicine.

The new Bioinformatics Specialization will allow learners to earn a Specialization Certificate that serves as a mini-degree in this fast-growing, cutting-edge field. “Our online courses are identical to a core class in the Bioinformatics and Systems Biology Program at the University of California, San Diego, one of the top programs in the world. Actually, they have even more content,” said Pavel Pevzner, a professor of computer science and engineering at UC San Diego, who co-developed the Specialization with longtime colleague Phillip Compeau, who this month joined the computational biology faculty at Carnegie Mellon University. “In fact, the Specialization will cover twice as much material as we teach in our UC San Diego course, so online learners can acquire world-class skills, even if they don’t know anything yet about biology – or computer programming.”

The Bioinformatics Specialization gives learners the option of participating in one of two separate tracks: one for students who already have programming skills, the other designed largely for biologists who don’t code but do want to learn how to use popular bioinformatics tools to solve practical problems.

“Biologists use bioinformatics tools such as BLAST in their daily lives,” said Compeau. “BLAST is like the Google of biology: everybody uses it, even if they don’t know how it works. But it’s important for a biologist to know how BLAST works to avoid pitfalls, so we explain how these tools work even if the learner doesn’t know how to program.” Learners who already know how to program will take a “hacker track” that will automatically test their programming skills using over 100 algorithmic puzzles motivated by modern biology.



Cover of textbook created for the online Bioinformatics courses, now used for online as well as offline classroom teaching.

Each of the Bioinformatics courses will run for four weeks, and the starter course in the series, **Finding Hidden Messages in DNA**, will begin August 31 and replay every six weeks. Subsequent courses include: **Genome Sequencing**; **Comparing Genes, Proteins, and Genomes**; **Deciphering Molecular Evolution**; **Genomic Data Science and Clustering**; and **Finding Mutations in DNA and Proteins**. These courses are followed by a Capstone Project, **Big Data in Biology**, which includes challenges in bioinformatics and personalized medicine developed jointly with scientists at Illumina. “In the Capstone, students will face the same kind of challenges that researchers in the biotech and pharmaceutical industry face,” said Compeau. Pevzner added, “Each section of the Capstone will include a motivating example illustrating how the emerging field of personalized medicine has contributed to decoding the causes of mysterious diseases that traditional approaches failed to diagnose.”

All sections of the Capstone have been developed jointly with scientists led by Semyon Kruglyak, the Senior Director of Informatics Research at Illumina. “Illumina cares about education. We offer continued education to our own scientists, and we have the BaseSpace cloud platform that thousands of biologists around the world use. We are making our data sets and analysis on BaseSpace available to people taking these courses,” said Kruglyak. “Illumina is most interested in educating biologists in bioinformatics because bioinformatics plays an important role in experimental design and data interpretation, but the subject is largely missing from even some of the best biology programs. This course seems like an ideal way to close that gap quickly.”

Plus, noted Kruglyak, “success in this Specialization could lead to Illumina job opportunities, because the company is looking for employees who can tackle biological Big Data.”

Pevzner and Compeau are not newcomers to the world of massive open online courses (MOOCs). The predecessor of this Specialization was the first bioinformatics MOOC, “Bioinformatics Algorithms”, which launched in 2013 on Coursera with more than 35,000 learners registered. The course was so successful that students at the University of Maryland and other top schools now take the course in a “flipped” classroom, watching videos and going through the interactive text before they arrive in the classroom. The curriculum has also been adopted in two high schools in Ohio and California.

The schools also use the textbook that Compeau and Pevzner put together. “A MOOC is the perfect vehicle for writing a textbook,” explained Pevzner. “From the very beginning, we invited learners to ask questions and make comments in response to MOOC materials, and as a result, we have compiled over 10,000 comments and used these comments to radically modify the content of the textbook.”

“Scouring the 10,000 comments resulted in 60 pages of new content for the textbook,” said Compeau. “As a result, students have far fewer questions because the content is adjusted to answer those questions before they are asked.”

“It’s a new modality for writing textbooks,” added Pevzner. “You revise the textbook continually, so the courses get better and better over time, and the content doesn’t become stale.” The 2nd edition of *Bioinformatics Algorithms: An Active-Learning Approach*, by Compeau and Pevzner, is being published this month.

The instructors anticipate a 50-50 balance between people who enroll in the Specialization, half of them from biology backgrounds, the other half with computational backgrounds and programming skills. Among learners who have taken Pevzner and Compeau’s previous online bioinformatics courses, 40 percent had at least a Master’s degree, and 56 percent were working full-time while taking the course.

Looking to the future, Pavel Pevzner thinks that the revolution in adaptive learning holds the key to the future of online education and believes in running online courses based on interactive texts rather than on videos, which have served as the workhorse of most MOOCs. “We have been unhappy with the term ‘MOOC’ because courses based on video lectures hardly ever change,” he said. “In contrast, our course, based on the interactive text, is very fluid, and its content is continually being improved. Our hope is that online courses will eventually move toward Massive Adaptive Interactive Texts, or MAITs, and when that happens, the online

revolution in education will gradually change the way courses are taught *offline*. This is already happening with our own interactive text, as many universities have adopted it in their offline classrooms."

This course is an initiative under the UC San Diego Office for Online and Technology Enhanced Education, which coordinates campus education efforts in the online arena. The office, part of the Teaching and Learning Commons, is closely aligned with the Commons' goals of supporting both instructors and students.

MEDIA CONTACT

Doug Ramsey, , dramsey@ucsd.edu

UC San Diego's Studio Ten 300 offers radio and television connections for media interviews with our faculty, which can be coordinated via studio@ucsd.edu. To connect with a UC San Diego faculty expert on relevant issues and trending news stories, visit https://ucsdnews.ucsd.edu/media-resources/faculty-experts.