

UC San Diego

UC San Diego News Center

September 30, 2013 | By Doug Ramsey

Is Massive Open Online Research the Next Frontier for Education?

A team from UC San Diego is launching a new course on the Coursera online learning network that breaks ground on several fronts.

In “Bioinformatics Algorithms – Part 1,” UC San Diego computer science and engineering professor Pavel Pevzner and his graduate students are offering a course that incorporates a substantial research component for the first time.



“To our knowledge, this is the first major online course that prominently features massive open online research, or MOOR, rather than just regular coursework” said Pevzner. “All students who sign up for the course will be given an opportunity to work on specific research projects under

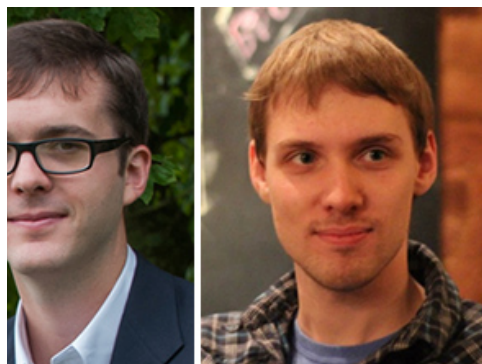
the leadership of prominent bioinformatics scientists from different countries, who have agreed to interact and mentor their respective teams.”

“What sets us apart is combining research with a MOOC,” said Ph.D. student Phillip Compeau, who helped develop the online course. “The natural progression of education is for people to make a transition from learning to research, which is a huge jump for many students, and essentially impossible for students in isolated areas. By integrating the research with an interactive text and a MOOC, it creates a pipeline to streamline this transition.”

Bioinformatics Algorithms (Part I) will run for eight weeks starting October 21, and students are now able to sign up and download some of the course materials.

Another unique feature of the online course: Pevzner and Compeau have developed a book for supporting the course that will soon be available in e-book form, while Pevzner’s colleagues in Russia developed a content delivery system that integrates the e-book with hundreds of quizzes and dozens of homework problems.

The U.S.-Russian team, led by Pevzner’s foreign student Nikolay Vyahhi, also implemented the electronic course using the beta version of Stepic, a new, fully integrated educational platform and startup developed by Vyahhi. (Stepic derives its name from the “step-by-step, epic” solution its developers delivered for electronic publishing.)



Compeau (left) and Nikolay Vyahhi are from Pevzner’s bioinformatics laboratories in San Diego and St. Petersburg, respectively.

The Bioinformatics Algorithms course also provides access to Rosalind, a free, online resource for learning bioinformatics through problem solving. Rosalind was developed by Pevzner’s students and colleagues in San Diego and St. Petersburg with funding from the Howard Hughes Medical Institute, the Russian Ministry of Education, and Russian Internet billionaires Yuri Milner and Pavel Durov through their “Start Fellows” award. Rosalind already has over 10,000 active users worldwide.

“We have recruited a very qualified and hard-working teaching assistant for this class, and her name is

Rosalind” said Pevzner in a tongue-in-cheek promotional video for the course now streaming on Coursera. “She has kindly agreed to instantly grade your programming assignments 24 hours a day, seven days a week.”

Rosalind – named in honor of British scientist Rosalind Franklin, whose X-ray crystallography with Raymond Gosling facilitated the discovery of the DNA double helix by Watson and Crick – will grade the programming assignments. They come in the form of bioinformatics problems of growing complexity as the course progresses.

We developed Rosalind to inspire both biologists and computer science students,” said Rosalind principal developer Vyahhi, who worked with Pevzner during the latter’s sabbatical in Russia. “The platform allows biologists to develop vital programming skills for bioinformatics at their own pace, and Rosalind can also appeal to programmers who have never been exposed to some of the exciting computational problems generated by molecular biology.”

“We have already started testing the first modules of the Bioinformatics Algorithms course in top Russian universities,” said Pevzner. “We are getting good feedback, and we hope to reach out to early students registering for the Coursera course so we can perform a ‘stress’ test – just to make sure that we will be able to deliver on the promise of massive open online research projects, whether it’s hundreds or thousands of students who enroll.”

The UC San Diego course, produced in Calit2’s Qualcomm Institute at UC San Diego, will cover many algorithms underlying fundamental topics in bioinformatics and will try to answer questions like, “What DNA Patterns Play the Role of Molecular Clocks Inside Cells,” or, “How Have Humans Populated the Earth?” To answer these questions, Pevzner and his team will introduce fundamental computational concepts such as greedy and randomized algorithms, graph theory and combinatorics, dynamic programming and combinatorial pattern matching, and many others.

“We say that you should know the basics of programming in whichever computer language you studied,” said Rosalind co-founder and course co-developer Phillip Compeau, who is completing his Ph.D. on combinatorics of genome rearrangements. “But if you don’t already know a language, you can probably start from ground zero by working on introductory problems on Rosalind, which will give you the basics in Python after about 10 hours of problem-solving. So students without computer science backgrounds should not shy away from the course.”

“Whatever you are studying now – computer science, biology, or mathematics – we want you to switch to bioinformatics,” exhorted Pevzner in the promotional video. “Not because these disciplines are not important, quite the opposite; but because bioinformatics is so cool, and because it is built on all of these disciplines!”

Pevzner was working on his Ph.D. in discrete mathematics when he had his first encounter with bioinformatics, and he says that he started over in the “futuristic new discipline – and never looked back.” He is now a professor of Computer Science and Engineering at UC San Diego, where he holds the Ronald R. Taylor Chair.

In 2006, Pevzner was named a Howard Hughes Medical Institute Professor. His research has focused on the creation of bioinformatics algorithms for analyzing genome rearrangements, DNA sequencing, and computational proteomics. He is a Fellow of both the Association for Computing Machinery (ACM) and the International Society for Computational Biology (ISCB).

In 2011, Pevzner founded the Algorithmic Biology Laboratory in St. Petersburg, Russia, which develops the Rosalind bioinformatics platform.

Topping a long list of publications, Pevzner authored *Computational Molecular Biology* (The MIT Press, 2000), co-authored (jointly with Neil Jones) *An Introduction to Bioinformatics Algorithms* (The MIT Press, 2004), and co-edited (with Ron Shamir) *Bioinformatics for Biologists* (Cambridge University Press, 2011).

But asked about which book makes him most proud – and Pevzner is unequivocal. “Absolutely this new book for the Bioinformatics Algorithms course,” he said. “Thanks to Coursera and the online learning explosion, we have the opportunity to reach many more students than would be possible at one university or even one country. There are students everywhere who could use this as a stepping stone to a career in a field that is going to change the world as we know it. Who wouldn’t want to be a part of that?”

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>.