BIG DATA UNITES ALL DISCIPLINES IN THE LIFE SCIENCES

ANALYSING VAST DATA SETS is baked into modern biomedical research, and academic programmes increasingly reflect this.

From vaccine development to personalized medicine, the influence of bioinformatics runs deep through the life sciences. Highly digitized labs rely on the ability to pull the right drop from the data lake to answer questions. Bioinformatics, an interdisciplinary branch of research that combines biology, chemistry, computer science and statistics, has become an essential tool to analyse big data sets — as demonstrated during light-speed vaccine development in response to the COVID-19 pandemic. As the field's prominence grows, so does its place in academia.

Brandeis University in Massachusetts offers multiple avenues to learn bioinformatics. In life sciences PhD programmes such as neuroscience or molecular and cell biology, students learn the tools of bioinformatics in their research, says Wendy Cadge, dean of the Graduate School of Arts & Sciences. The university's online master's degree in bioinformatics, meanwhile, has been growing fast.

Bioinformatics has allowed clinicians and scientists to identify cancer-causing cell mutations and provide targeted, personalized drug therapies, says Alan Cheng, chair of the bioinformatics graduate programme at Brandeis. "We now understand much more about the mechanisms of cancer; that's partially through bioinformatics. Cancer patients are sequencing their tumours before treatment to identify the

▲ Bioinformatics has helped inform the development of targeted, personalized cancer therapies.

subtype, enabling scientists and physicians to identify personalized drug regimens and therapies specific to the patient."

SEQUENCE AND STRUCTURE
Brandeis teaches sequence, functional, and structural bioinformatics. Sequence bioinformatics, along with functional genomics, is about how the gene sequence gets transcribed, regulated and expressed, as well as how genes interact with other genes, Cheng says. In structural bioinformatics, researchers identify the sequence, structure and function of biomolecules

and tie all that together. Another way to distinguish the two is that structural bioinformatics involves a 3D, high-resolution description of the protein, biological molecule or biological complex, while sequence bioinformatics is 1D or 2D.

"There are a lot of newer methods to go from DNA sequences, translate them to proteins and then predict their 3D structures," he says. "There have been fantastic AI deep-learning advances in the last couple of years, an example being AlphaFold, which allows you to predict 3D protein structure from sequence.

Using that, we can now do 3D prediction on a lot more proteins and biomolecules and use that to better understand how mutations affect function."

Bioinformatics is central to a variety of roles. Students may specialize in the biology side, the computational side, the information technology side, or at the interface between them. In the online master's degree programme, students are taught to code in Python and R, an essential skill for scientists, Cheng says. "That's a key part of the programme, giving you the tools early on and helping you apply them to different problems, like personalized medicine, and more generally analysing sequences, structures, and proteins."

Graduate students drawn to bioinformatics are a diverse bunch. At Brandeis, they include people working in biology, chemistry, computer science and environmental science, as well as tenured professors and finance professionals. And for a discipline that cuts across academic departments, joining a university like Brandeis — one of the smallest R1 research institutions in the United States — has its advantages.

"Pretty much everyone works across labs, across departments," says Cadge. "The faculty are there and very hands-on." ■

ArtemisDiana/Getty Images

Brandeis UNIVERSITY