

Broad Institute Launches Collaborative Genomics “Cloud” Tool for Scientists ,

In an effort to harness and allow sharing of exponentially-developing genetic data, the Broad Institute will launch [“Genome Space,”](#)—a co-operative Web based tool aimed at “frictionless” data transfer, later this week.

So said Jill Merisov, PhD, the Broad’s associate director and Chief Informatics officer, in a keynote speech at the opening of the Bio-IT World conference yesterday, in Boston. The Broad is a Harvard-MIT research center located in Kendall Square, Cambridge.

In her talk, Merisov pointed out that just ten years ago, scientists announced that they had identified all of the genes present in human beings. Since then, researchers have discovered 30 million genetic variations among 1000 different individuals, 3000 genetically-related disease traits, and a multitude of cancer types. Such findings are now being used to determine the genetic bases of many diseases, to develop treatments for those diseases, and to determine for which patients particular treatments are likely to be effective. In another ten years, she said, such “personalized medicine” will be commonly used by doctors, in clinics.

These advances are due in large part to less expensive, increasingly sophisticated and sensitive computer technologies that have led to an “explosion” of data , to less “noisy” data, and to new, international ways of reviewing the data, Merisov explained. Scientists can now buy the technology and carry out sequencing in their own labs and “”computing is now integral to every aspect of biomedical research.”

But these developments also mean that there are now seven-to-

ten thousand bioinformatics tools available for download on the Web and five thousand databases—many of which are “out of reach” for scientists who do not have sophisticated programming skills.

The new tool “bridges the gaps between bioinformatics tools, making it possible for [scientists] to move data smoothly between these tools, leveraging the available analyses and visualizations in each of these tools,” according to the [Genome Space Web](#) site.

Genome Space also allows for data storage in the Amazon cloud [a computing platform of Amazon.com] and “provides necessary file format transformations whenever a scientists selects an analysis or visualization within one of the tools.

The GenomeSpace project is a collaboration of the Mesirov and Regev laboratories at the Broad Institute; the Chang laboratory at Stanford University; the Ideker laboratory at the University of California, San Diego; the Nekrutenko laboratory at Pennsylvania State University; the Segal laboratory at the Weizmann Institute of Science; and the Haussler and Kent laboratories at the University of California, Santa Cruz. GenomeSpace is funded by the National Human Genome Research Institute, with additional support from Amazon Web Services, according to the Genome Space Web site.

The [Bio-IT Conference Expo](#) 2012 goes through April 26.

—Anita M. Harris