

## Microbiology

# The cutting EDGE: Bringing genomics to everyone

by **Patrick Chain**<sup>1</sup> | Professor

<sup>1</sup>:Los Alamos National Laboratory, U.S. Dept. of Energy's NNSA, USA

This Break was edited by Massimo Caine, *Editor-in-chief* - TheScienceBreaker

Diabetes, infertility, cancer, allergies, Alzheimer's disease - the key to one day preventing or even curing such afflictions and diseases (both infectious and genetically driven) may be locked in our own genetic code and the code of [microorganisms that inhabit our bodies](#). Genomics, the study of this DNA code, has recently become much more promising as a result of two things: (1) vast improvements in quality and quantity of [DNA-sequencing](#), and (2) an exponential decrease in the cost of such DNA-sequencing. For example, what originally cost \$3 billion for the [human genome sequence](#) is now less than \$1000 to resequence it. However, two key problems cause a significant bottleneck for data interpretation: (1) the sheer volume of genomic data and (2) the shortage of expertise necessary to understand and act it.

To help address both problems, we have created [EDGE \(Empowering the Development of Genomics Expertise\) Bioinformatics](#). EDGE includes a variety of innovative and integrated bioinformatic tools, housed within a user-friendly, web-based platform that requires little input and is mostly point-and-click. Preconfigured workflows can be quickly and easily selected to address a wide variety of goals and projects, including analyzing a genome or a much more complex metagenomic (microbiome) or clinical sample.

While the human genome contains approximately 6 billion nucleotides ("letters" of DNA in each human cell), it has been estimated that the human body also harbors at least as many bacterial cells as human cells. Analyzing a microbial community or a microbiome in a clinical sample may reveal thousands of different

bacterial and viral species, residing in a complex mixture with human cells. Taking into account the microbes inhabiting the human body, researchers have estimated that the gene content of the typical microbiome is over 300 times larger than the human genome. Characterizing the volume of data from such a sample is a daunting challenge.

Realistically, few biologists have the computational expertise or resources to accomplish such types of analyses, thus requiring the aid of experts to develop custom computer code to process genomic data. EDGE Bioinformatics holds promise to revolutionize the way scientists analyze genomic data by making sophisticated tools available via intuitive and easy-to-use web-based interfaces. Using a few mouse clicks, EDGE receives raw data from sequencers and creates reports and graphics based on the data - usually in minutes or hours instead of days or weeks of wrangling with sequence data. EDGE Bioinformatics truly makes it possible for nearly *any* biologist or physician with access to genomic data to use this technology and to derive meaning from it. The software is entirely open source, and can be downloaded at no cost and used locally or remotely via the web.

With EDGE on a local computer server or in the cloud, it is now possible to bring the power of complex, sequencing data analytics to smaller facilities, including clinics, hospitals, and university laboratories. Scientists are beginning to use genomics to assess human genome mutations contributing to cancer, and human microbiome shifts associated with Crohn's disease, irritable bowel syndrome, allergies, Alzheimer's disease, and countless other areas

of research, like algal biofuels and climate effects. In the field of infectious disease, using sequencing data and EDGE, it will be possible in the future for a nurse to swab your saliva and later tell you if your symptoms are more likely caused by a viral or bacterial infection. This could reduce the unnecessary and problematic use of antibiotics (read also the Break: [Collateral damage: antibiotics disrupt the balance in the gut](#)). It is also envisioned that hospitals could use EDGE to ascertain which pathogens inhabit the hospital environment, implement countermeasures, and improve infection-control procedures.

Important societal issues such as climate change and eradicating the spread of disease require an in-depth understanding of how microorganisms function (read also the Break: [Cold adaptation: gut bacteria can make the difference](#)). The key to that understanding is locked in the genetic code. Sequencing technology has become readily available to tackle such scientific challenges, and EDGE now makes the tools for understanding genomic data readily available as well.