

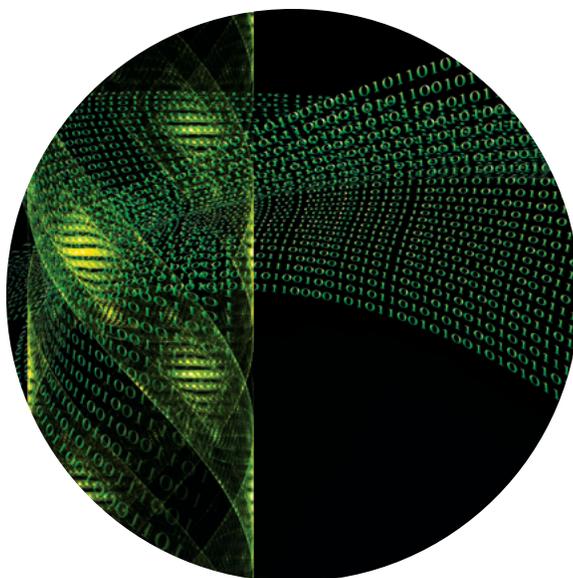
KEY TECHNOLOGIES

Bioinformatics

Acquiring DNA sequence has now become routine, and new technologies can sequence a bacterial genome in a single day. Similarly, microarray experiments shed light on the RNA levels produced by tens of thousands of genes. Current analysis platforms are capable of generating terabytes of data in a single run. For reference, one terabyte is equal to 1,000 gigabytes

Understanding the meaning of all that information is a daunting challenge. Deciphering the data requires a biological knowledge of what to look for, algorithms (computer programs) capable of detecting interesting features and computers powerful enough to perform complex analyses efficiently and rapidly. Fortunately, advances in all three areas have kept pace, and the resulting field of bioinformatics seeks to characterize functional sequences in genes and genomes through computational models. In addition, the data must be managed – stored in a form that is useful to the researcher and readily accessible. This has led to the development of many databases that store and provide data and analytical tools for researchers. The primary mission of all these databases is to provide unlimited free access to anyone, including Alabama students, interested in studying genomic sequences. It is no exaggeration to say that these databases and the immediate access to them through the Internet have changed the way that nearly all biological research is done.

Many bioinformatics experts, particularly in the early days of the genome sequencing efforts, were computer scientists who formed partnerships with biologists. With the growth of the field of genomics, it is not unusual today for a student to be trained in a truly interdisciplinary way by developing deep expertise in both biology and computational science.



The concept of bioinformatics is a critical component to understanding modern genomic discoveries. It provides software tools capable of exploring the structure of chromosomes and predicting the likelihood of a genetic match in a forensics case. Bioinformatics databases also manage, search and store the data produced by the human genome project and more recent large-scale studies such as ENCODE and the 1,000 Genome Project (Biology Standard 3b). This topic could be incorporated in an AP Biology class under Enduring Understanding 1. A and to support the Science Practices of data analysis and mathematical modeling. Lastly, the creation, management and utilization of bioinformatics databases can be incorporated into the Technology Education course (COS objectives 26 and 27).



The HudsonAlpha-developed high school lab activity Genes & ConSEQUENCES® connects information produced by a DNA sequencing system to genes, mutations and human disease. The activity incorporates biological databases used by genetic researchers on a daily basis. Students access a portion of the NCBI (National Center for Biotechnology Information) database known as BLAST. This program compares student entered sequence data to known sequences from a number of organisms, including humans, to identify genetic matches. The dataset allows students to determine chromosomal location of key genes, investigate their role in disease, and compare sequence in healthy individuals to patients experiencing symptoms. The activity has been incorporated into the Laying the Foundation curriculum for A+ College Ready and is available through the AMSTI/ASIM high school program across the state. The kit is available for purchase through a partnership with Carolina Biological Supply. More information can be found at www.hudsonalpha.org/available-educational-kits.

