THE REVOLUTION OF BIOINFORMATICS

Over the past few decades, major advances in the field of molecular biology have led to an explosive growth in the biological information generated by the scientific community.

Bioinformatics is the field of science in which biology, computer science, and information technology merge to form a single discipline. The ultimate goal of the field is to enable the discovery of new biological insights as well as to create a global perspective from which unifying principles in biology can be discerned. At the beginning of the «genomic revolution», a bioinformatics concern was the creation and maintenance of a database to store biological information, such as nucleotide and amino acid sequences.

A biological database is a large, organized body of persistent data, usually associated with computerized software designed to update, query, and retrieve components of the data stored within the system. A simple database might be a single file containing many records, each of which includes the same set of information. For example, a record associated with a nucleotide sequence database typically contains information such as contact name, the input sequence with a description of the type of molecule, the scientific name of the source organism from which it was isolated, and often, literature citations associated with the sequence.

A few years ago, a researcher wanting to localize a gene, or nucleotide sequence, was forced to manually map the genomic region of interest, a time-consuming process. Today, thanks to new technologies and the influx of sequence data, a number of high-quality, genome-wide maps are available to the scientific community for use in their research. Computerized maps make gene hunting faster, cheaper, and more practical for almost any scientist. In short, scientists would first use a genetic map to assign a gene to a relatively small area of a chromosome. They would then use a physical map to examine the region of interest close up, to determine a gene's precise location.

Ultimately, however, all of this information must be combined to form a comprehensive picture of normal cellular activities so that researchers may study how these activities are altered in different disease states. Therefore, the field of bioinformatics has evolved such that the most pressing task now involves the analysis and interpretation of various types of data, including nucleotide and amino acid sequences, protein domains, and protein structures. The actual process of analyzing and interpreting data is referred to as computational biology. Important sub-disciplines within bioinformatics and computational biology include the development of new algorithms (mathematical formulas) and statistics with which to assess relationships among members of large data sets, such as methods to locate a gene within a sequence, predict protein structure and/or function, and cluster protein sequences into families of related sequences.

Adapted from «Just the Facts: A Basic Introduction to the Science Underlying NCBI Resources- Bionformatics», *National Center for Biotechnology Information (NCBI)*, 29 March 2004.

Answer the following questions.

A) What is bionformatics?

B) What was the main focus of bionformatics, at the beginning of the «genomic revolution»?