

---

**Shimadzu**

---

**Developments**

---

**in Biotechnology**

---

# The Real Age of Biotechnology is Only Just Beginning

Shimadzu Contributes to the Development of Biotechnology by Developing the World's Fastest Sequencer and DNA Microchip

**T**he human genome contains all human genetic information. The Human Genome Project, which aims to determine all the DNA sequences in the human genome, is now entering its final stages. However, completion of the Human Genome Project will be no more than a single milestone along the road to the development of biotechnology, one of the key technologies in the 21st Century. Shimadzu is contributing to the improvement of human society through the development of instruments and systems that aid in analyzing the genomes of humans and other organisms and in the analysis of proteins.

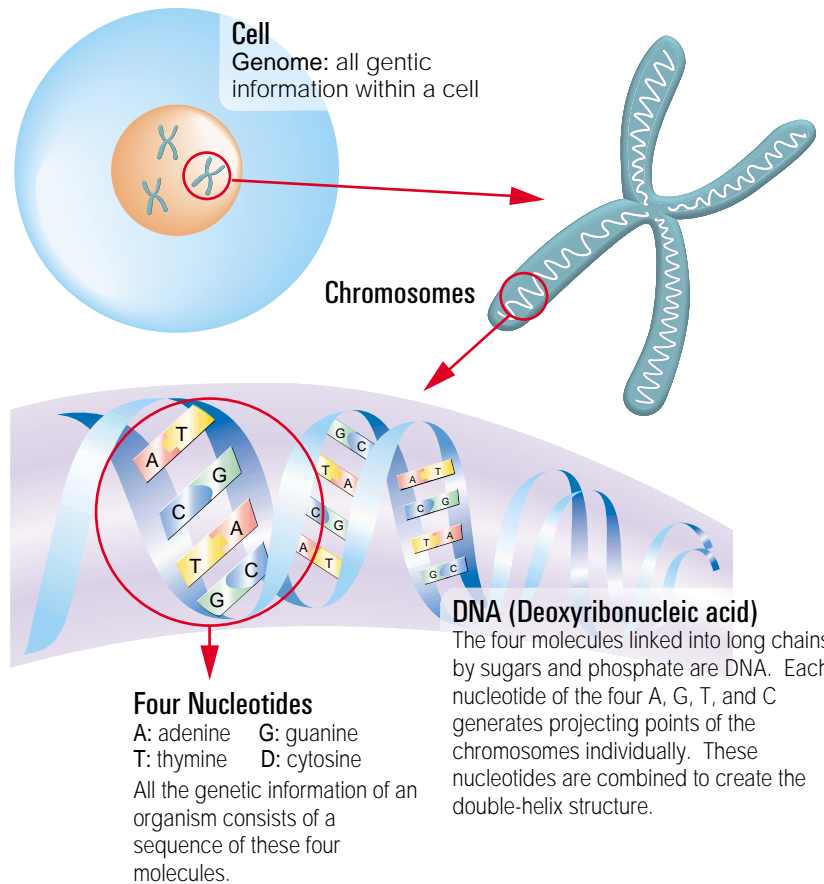
## Genetic Information is Expressed as a Sequence of Four Nucleotides

Be it humans or other organisms, all gene information is expressed as a series of just four substances called nucleotides – adenine, guanine, thymine, cytosine – arranged in different sequences. These four molecules linked into long chains by sugars and phosphate are DNA. Groups of DNA form chromosomes, and all the chromosomal information in a cell makes up a gene. It is incredible that all the genetic information of a complex entity such as a human comprises a sequence of just four molecules. Still, if we consider that today's computer culture is built on just 1's and 0's, it's not hard to imagine the huge amount of information that can be expressed by a sequence of these four molecules.

The total number of nucleotides in the human genome is estimated as some three billion base pairs. DNA has the famous double-helix structure in which the nucleotides form pairs, so it is more accurate to state that DNA has three billion base pairs. Determining the sequence of the three billion base pairs is the research object of the Human Genome Project, which has the cooperation of researchers from countries all over the world. The Human Genome Project began in the late 80's and has proceeded significantly faster than originally predicted, so that a draft base sequence is expected to be revealed during 2000. So, just what will this achieve?

Proteins are the most important constituent of all organisms, including humans. For example, the hemoglobin that transports oxygen in red blood cells is a protein, and the enzymes essential for pro-

## DNA Structure



sequence does not clarify the mechanisms of human life. Why? Because all we will know in 2000 is the DNA sequence itself, but decoding the meaning and function of the base sequence remains a topic for the future. Of course, the functions are being decoded in parallel with identification of the DNA sequence, but we still understand very little. For example, we often do not even know where genetic information starts and ends along the enormously long DNA sequence. In practice, the number of human genes determined by the sequence of three billion base pairs is assumed to be approximately 100,000, but as research proceeds, some calculations estimate the actual number as several ten thousand genes.

Also, the Human Genome Project will determine the DNA sequence for some so-called "standard" human. But real people come in all types. Skin and eye color vary according to race, and other physical features, traits, and physiological functions also differ. On top of these are differences from individual to individual. In reality, no "standard" human exists, and the variety of individual humans derives from differences in genetic information; that is, from differences in the DNA sequence. Diseases such as cancer caused by abnormalities in the gene information ultimately result from errors in

moting chemical reactions in living organisms, such as digesting nutrients and generating energy, are also proteins. The number of proteins is enormous. There are 2000 human enzymes alone. Proteins normally comprise a series of at least 100 of 20 types of amino acids. Assuming 100 amino acids, the number of combinations is 20 to the power of 100 – a number that defies comprehension. Whether or not they all actually exist, the potentially huge number of proteins permits creation of the various constituent structures and many enzymes comprising numerous organisms.

## Decoding the Meaning and Function of the DNA Sequence is a Topic for the Future

The commands to make these proteins are given by gene information. That is, the sequence of DNA in the gene defines the sequence of amino acids; the sequence of amino acids defines specific proteins; and these define the form and function of the organism. This is the method by which genetic information is transmitted and is the mechanism that imparts the inherent characteristics of all organisms, including humans. However, knowing the DNA

## Possible solutions to society's problems promised by biotechnology

### Treatments for cancer, strokes, and other life-threatening diseases

Approaching cancer, strokes, life-threatening disease, and genetic handicaps at the gene level promises to lead to the development of effective new drugs and treatments.

### Effective measures against food shortages

To counter food shortages, the food supply could be increased through biotechnological measures, such as transplantation of fertilized eggs in livestock and gene recombination in plants.

### 1 Genome science

Across the globe work is under way in areas such as the "Human Genome Project" (people: 3 billion base pairs, 100,000 genes) that is hoped to advance diagnosis methods and therapy for illnesses as well as the "Rice Genome Project" (rice: 433 million base pairs, some 30 or 40,000 genes) that is expected to lead to the development of new products.

### 2 SNPs (Single Nucleotide Polymorphisms)

This is an important pointer in distinguishing drug effects on individual patients as genes (a gene is made up of several hundred DNA, and just one DNA may be different) vary minutely from person to person. This is a fundamental element that will help to bring about tailor-made therapy.

### 3 i-biotechnology

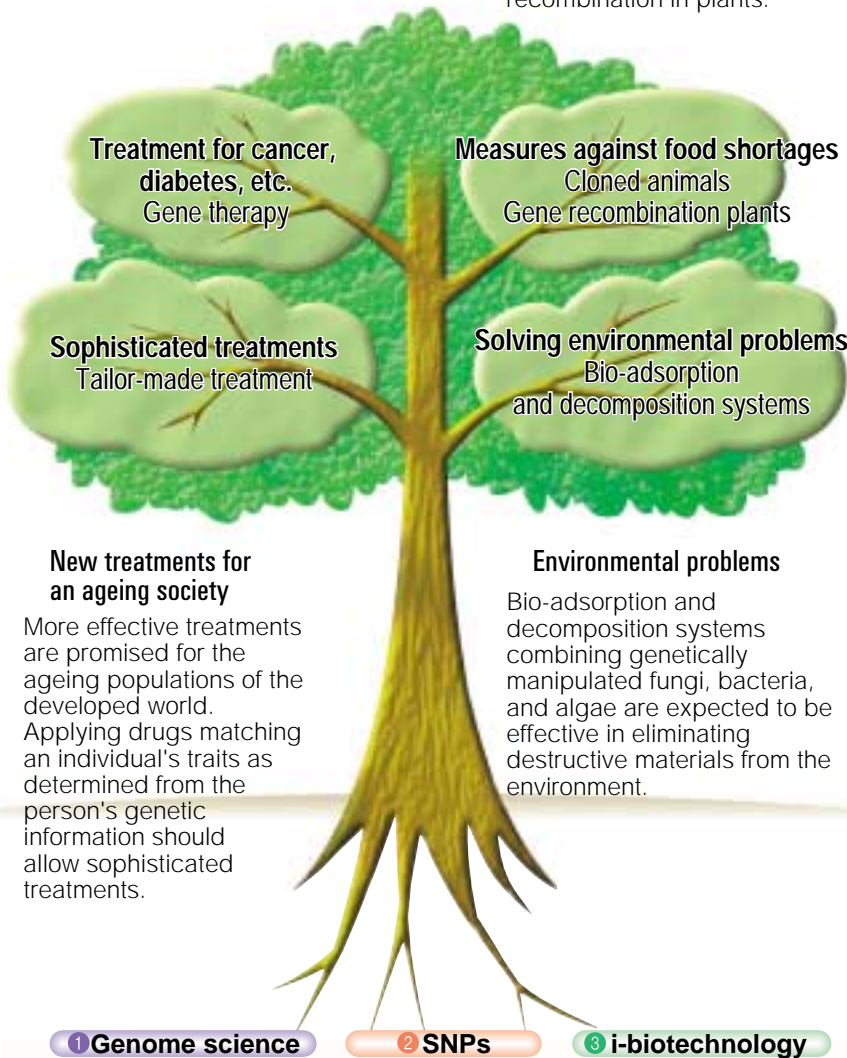
Bioinformatics and DNA chips are two examples of the merge between biotechnology and information technology that are providing advancements in computer-based bio research.

### 4 Protein engineering

This involves structure changing, function changing or the adding of new functions for proteins such as enzymes and antibodies. It is considered to be third biotechnology followed by fermentation technology and gene recombination technology, and will undoubtedly have applications in medicines and electronic sensors, etc.

### 5 DNA chips

Until now DNA analysis required test tubes, flasks and various other analysis apparatus. Now these are reproduced as channels and wells engraved onto glass plates mere centimeters in size. High expectations are held for this technology as tests can be completed with the merest micro amount of expensive test chemical while sequencing time also can be drastically reduced.



### New treatments for an ageing society

More effective treatments are promised for the ageing populations of the developed world. Applying drugs matching an individual's traits as determined from the person's genetic information should allow sophisticated treatments.

### Environmental problems

Bio-adsorption and decomposition systems combining genetically manipulated fungi, bacteria, and algae are expected to be effective in eliminating destructive materials from the environment.

1 Genome science

2 SNPs

3 i-biotechnology

4 Protein engineering

5 DNA chips





*Large-scale gene analysis using Shimadzu's world-fastest DNA sequencers, RISA-384 at the Institute of Physical and Chemical Research.*

the base sequence. Research to determine what differences in an organism or what diseases result from an error in a particular part of the DNA sequence is only just beginning

## Analytical Instruments are Essential for Genetic Research

Determining the DNA sequence of the genes of cancer cells in the case of skin cancer, for example, and determining the differences with the genes of normal cells would lead to a complete understanding of skin cancer, exact diagnosis and effective treatments. In an SNPs = Single Nucleotide Polymorphisms, a single mismatched nucleotide is known to cause cancer, but an overall understanding is needed to identify this one base.

And applications aren't just

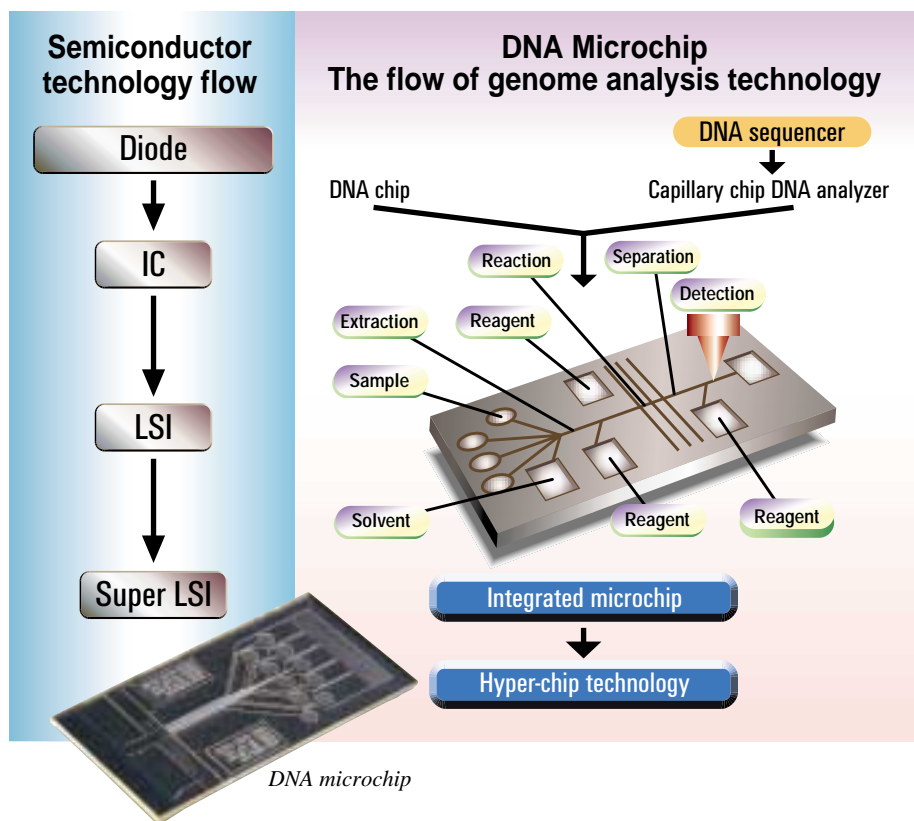
limited to humans. For example, an understanding of the mechanisms of life at the gene level is effective in the breeding of high-yield, disease-resistant rice, but rice and other cereals have a genome as long as that of humans. The scope of genome analysis work still to be challenged in the future is truly enormous.

This work involves manipulating the bases, sugars, phosphates, and other chemical compounds of DNA at a molecular level, then decoding the vast number of codes represented by these chemicals. This task would be impossible without the aid of automated instruments and computers. Chromosomes are extracted from cells. They are then cleaved into fragments of about 500 to 1000 bases because a sequence of several

tens of millions bases long cannot be read simultaneously. Before the analysis of samples for the automatic DNA sequencer, a large number of copies of these fragments is generated and subjected to a special pre-treatment. The information analyzed is then computer processed and recombined to correspond to the length of the original chromosome. This is roughly the procedure for analyzing a genome, but it requires many instruments, including automatic DNA sequence analyzers, called "sequencers;" ultrasonication for fragmentation of the chromosomes; thermal cyclers to generate many copies of the chromosome fragments in a short time; and automatic Sanger reaction devices for pre-treatment.

## Capacity of Automatic DNA Sequencers Leaps 400%

Shimadzu has delivered a variety of products since starting full-scale development of instruments for research in the biotechnology field in 1984. Each product was highly regarded for radical contributions to progressing research. For example, the sequencer RISA-384, an automatic DNA sequence analyzer developed in 1999. If the three billion base pairs are cleaved into fragment length of



500 base pairs for analysis, a total of six million fragments must be analyzed. Investigating these fragments one at a time would be so slow that research would not progress. A sequencer is able to simultaneously analyze multiple fragments. Previous sequencers could only analyze 96 fragments but the RISA-384 dramatically increased the capacity to 384 fragments in the same time, achieving significant reductions in analysis time. The DNA microchip currently under development could be called an ultra-small automatic DNA analyzer. It is a micro laboratory offering automatic DNA analysis functions on a glass

substrate of several square centimeters. While not a general-purpose device for all DNA analysis, a DNA microchip loaded with known research results as software can easily conduct on-the-spot investigation of a specific single base mutation, for example. This device will be useful in diagnosis of specific genetically caused diseases or in the determination of the existence of certain genetic traits as easily as measuring body temperature.

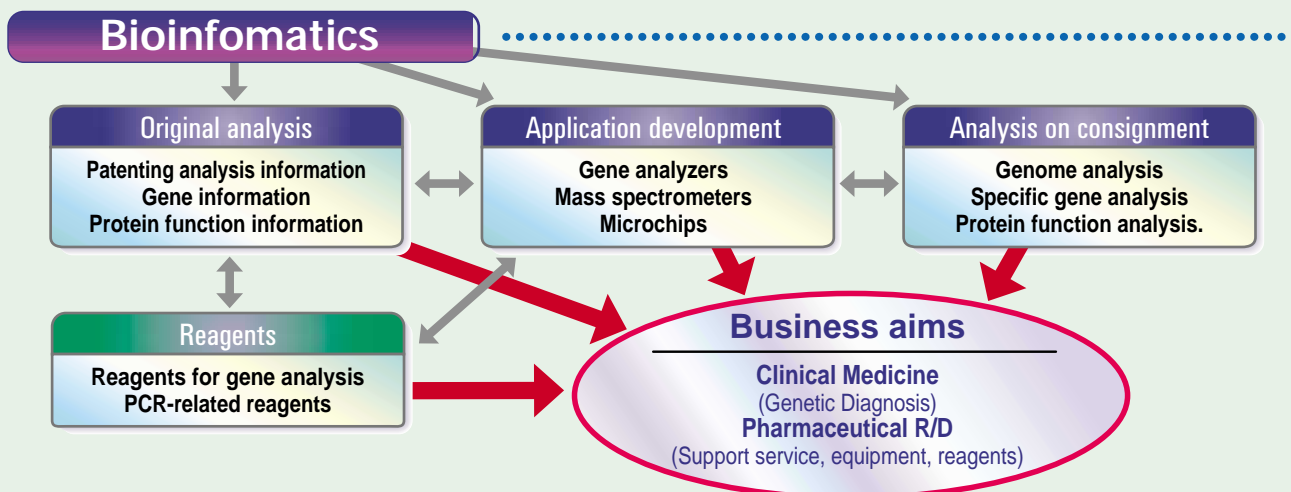
### Genome Analysis Service Supports Researchers

As stated earlier, proteins constitute all organisms, including humans, and play the most direct-

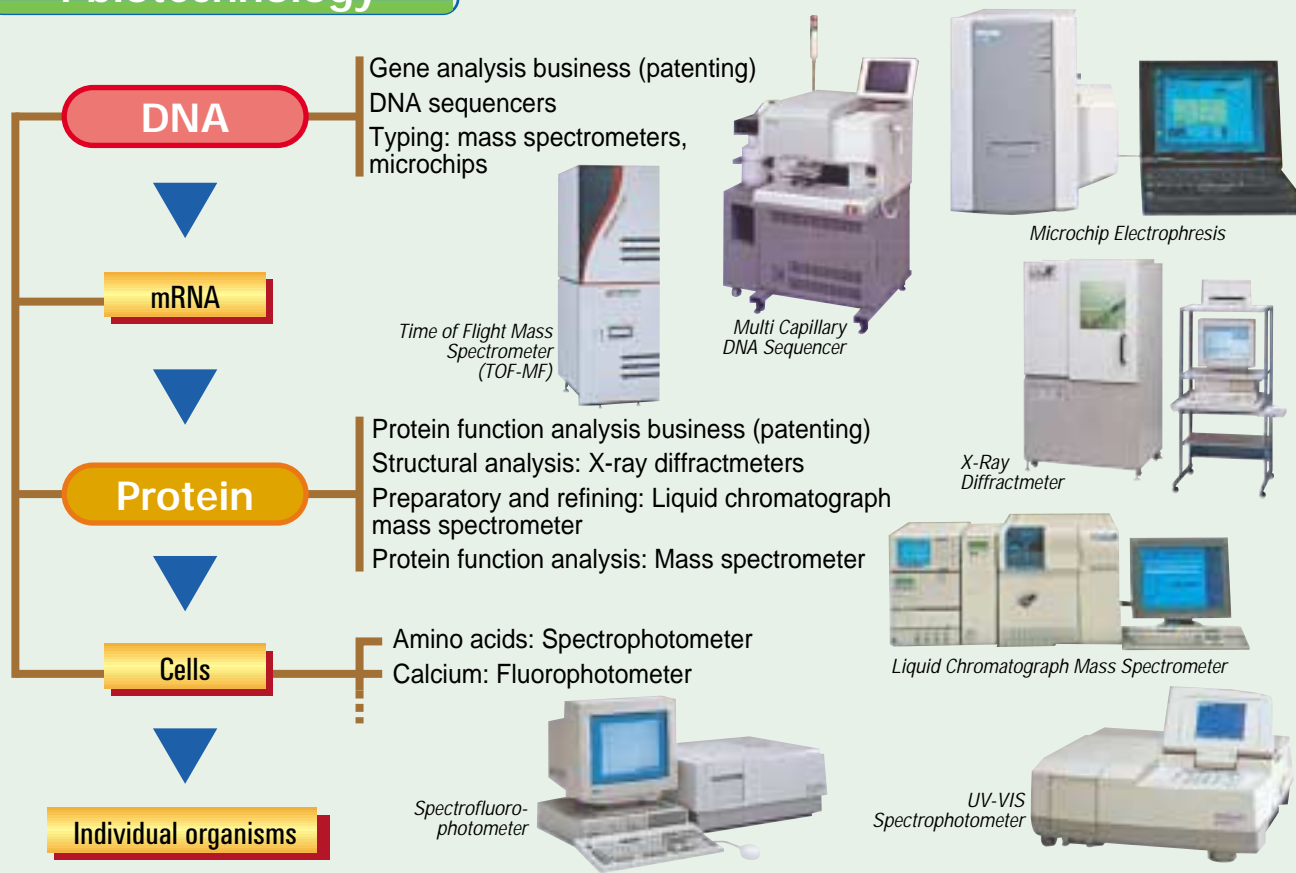
ly essential roles in maintaining the organism's functions. Research into the identification of the various proteins and their actions must be carried out subsequent to, or in parallel with, the work to determine the gene information. Paired with the Human Genome Project is the **Proteome** Project: a project that has already begun to analyze all the proteins produced inside an organism. Shimadzu has developed many sophisticated products, including mass spectrometers, that are essential for the identification of proteins. The company is continuing research that should lead to more sophisticated technologies and mechanisms for protein analysis systems. But Shimadzu does not only develop instruments. Under commission from researchers at universities and national research organizations, Shimadzu established the Genomic Research Center in April 2000 to provide a genome analysis service. Equipped with many leading-edge analytical devices, including RISA-384, the center offers a faster, cheaper, and more accurate analysis service that supports researchers at the forefront of their field. The Genomic Research Center is conducting independent analysis of micro-organisms and bacteria useful for brewing and other useful plants and animals and plans to release this knowledge and information to society in the future.



## Potential Project Areas in Biotechnology



## i-biotechnology



**mRNA:** Nucleic acid that copies the DNA deoxyribonucleic acid genetic information and synthesizes proteins based on this information with the ribosomes. Ribonucleic acid.