Genome technology and electronics

Yasuo Miyamoto

In recent years the deciphering of the blueprint of humans, the human genome, has progressed. Because of this progress, various changes are due to take place in our lives and in industry's structure[1]. The deciphering of the human genome has provided huge contributions to the development of the information technology and a variety of analysis equipment. It is expected that in the future the genome technology's relationship with electronics will further deepen [2]. This paper provides a general description of genome-related technology and industry trends from the viewpoint of its involvement with electronics, along with recent topics.

Summary of genome technology and the genome industry

The genome is the complete set of genetic information in living organisms, which means it is the necessary blueprint for sustaining life, facilitating growth, and procreation of individuals. The genome for humans is called the "human genome". The genome entity is a set of macromolecules known as the DNA (deoxyribonucleic acid), which is composed of four types of chemical compounds linked in a corded fashion. It is possible to represent these compounds with characters, which means the genome can be interpreted as a string of characters written with four types of characters. Basically the genome information can further be described as all the information encompassing the entire life phenomena, related to genes and proteins, which are created based on the genes described by the character string (genome sequence). Proteins are constituent components of the human body, such as skin or muscle and internal organs, as well as functional components to maintain actual life activities, such as enzymes or hormones that regulate the workings of a person. Further, genes are the components of the genome sequence that provide information necessary to create proteins. It is estimated that the human genome is a character string composed of approximately three billion characters, in which thirty to seventy thousand genes exist and from these, one hundred to three hundred thousand types of proteins can be generated.

The genome technologies related to electronics are summarized in Fig. 1, as well as products and services related to genome technologies. Although genome technologies are those related to the engineering of genes (such as technologies used to cut out certain segments of a DNA sequence, to multiply a small sample of DNA, or gene modification, which are all critically important), we shall concentrate on technologies related to electronics.

The technologies related to electronics, can be classified into the categories of measurement and analysis technology, bioinformatics (computational biology) and genome chip technology. Of all these technologies the measurement and analysis technology has the longest history while being instrumental in the invention of various types of equipment. These included equipment known as a sequencer (used to investigate the sequence of DNA), a mass spectrometer (used to investigate the structure of proteins, made famous last year when Koichi Tanaka received the Nobel Prize for Chemistry), and x-ray analysis equipment (it led James Watson and Francis Crick to the discovery of the double helix structure of DNA). Descriptions concerning bioinformatics (that are closely related to information processing technology), as well as genome chip technology (closely related to the semiconductor device technology), are provided in another chapter.

From a genome technology and industry perspective, the era preceding the deciphering of the human genome was an era wherein the deciphering of genome sequences and genes were the central activities in this field. Since the year 2000, however, the following of the declaration concerning the deciphering of the human genome, is known as the post-genome era. The term "post-genome" is an abbreviation of "post-genome sequence", representing an era wherein investigations of genes and protein functions are based on the sequence of the genome, which leads to the development of new drugs (genome-based drug discovery) and technologies leading to the industrialization of genome information, became the core [3]. The deciphering of various living organisms, including those that are not human, are under way, thanks to the establishment of technologies capable of expedient investigations of the genome sequence. For example, deciphering the genomes of mice or chimpanzees has proven to be an aid for better understanding human beings and can be helpful in understanding mechanisms of disease and illness. Further, deciphering the genomes of plants like rice, can lead to the development of harvests with special features, such as larger yields or hypoallergenic characteristics, which will contribute to food and agricultural industries. It is also believed that the deciphering of various microorganisms can bring about large contributions to environmental technologies and the chemical industry by making it possible to design and use microorganisms with genes, for example, to create petroleum or break down toxic substances, such as dioxins.

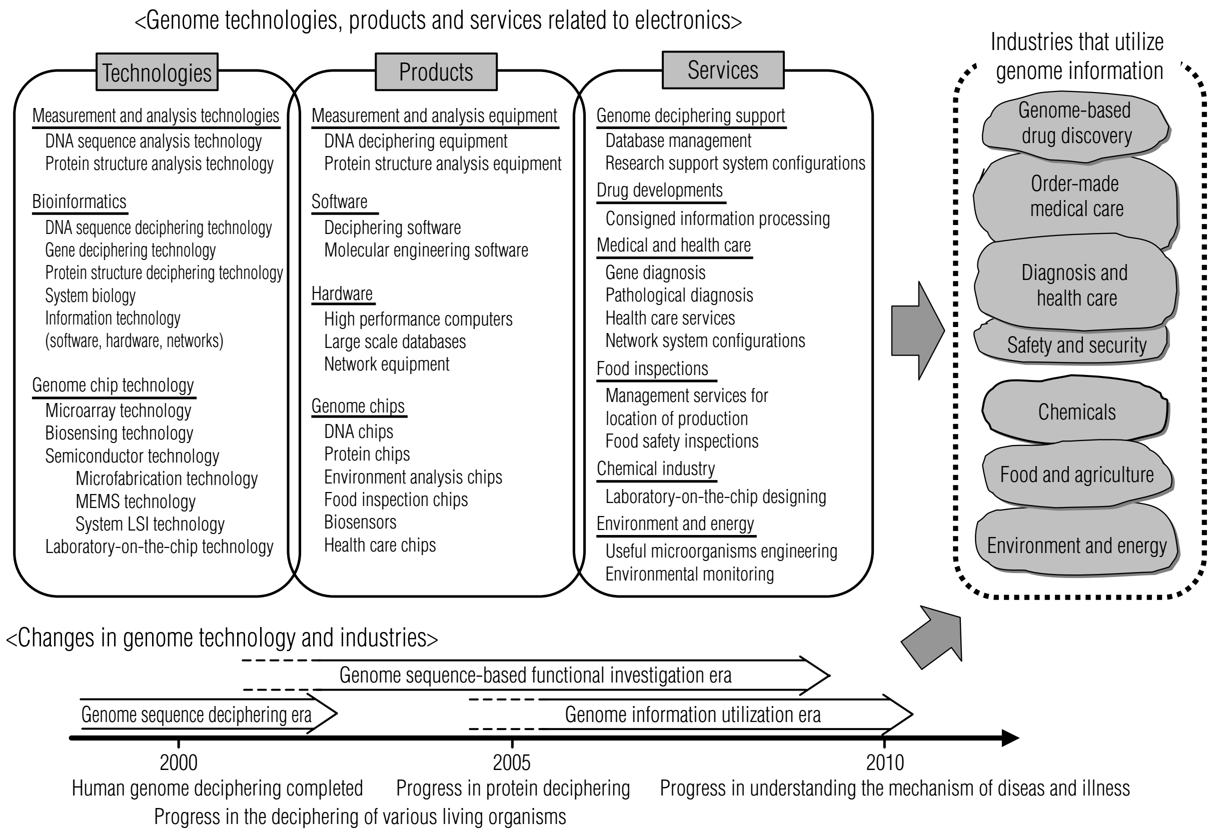

Fig. 1 Summary of genome technologies and genome industries related to electronics

Human genome deciphering gave us the understanding of the genome sequence that is common to all of mankind. The genome sequence, unique to an individual person, constitutes only approximately 0.1% of their entire genome sequence. Work to clarify this segment is currently under way. Order-made medical care (also known as "tailor-made medical care”), which provides prescription drugs with fewer side effects or medical treatment administered to counter diseases with a genetically higher probability of incidence, is becoming a more feasible practice [4].

Further, because the genome sequence of an individual person is unique and specific to an individual, its application is under consideration for security use, such as DNA identification and personal authentication. Specific application examples, currently used in trials, include a commodity authentication system using a “DNA Ink” (ID Technica Co., Ltd. and others), which is based on the DNA information provided by customers and the “DNA registered seal IC card” (NTT Data Corporation), which is based on individual DNA information. Since unique genome sequences also exist in other living organisms, it can be applied to the safety management of food as well. In Europe, for example, a practical source identifying certification system for beef, based on DNA identification, is now in place, following the BSE panic that swept across Europe in 1996. The technology, utilized in this system, was developed by a company in Ireland called IdentiGEN.

Information technology and genome (bioinformatics)

Bioinformatics is a study conducted for understanding, the life phenomena, that is, genome information, from an information science standpoint. Three stages of development exist for bioinformatics with respect to the human genome.

The first stage is the information processing technology related to reading the actual three billion characters of a genome sequence character string. The length of the segment of a character string, which can be deciphered by analysis equipment, is between several hundred to several thousand characters. It is, therefore, necessary to correctly reconnect the data that represents small fragments of the sequence, on a computer. An example of the methods used in information processing for such instances, would be the method for searching the optimal solution, developed in the field of artificial intelligence. Next, a prediction of the segment for the genes is made in the deciphered genome sequence. This involves the deciphering of the grammar used to express the genes and the understanding of the language used to express life. A mathematical model similar to that of natural language recognition, used in speech recognition technology, is applied.

The second stage is bioinformatics that respond to the post-genome era. Investigations into how proteins are made and the kind of proteins that are made from genes as well as the type of protein structure (the three-dimensional coordinate structure of atoms that form proteins) these proteins have. A specific example is the international collaborative project that was launched last year to investigate all human proteins. This project includes a plan to investigate, within five years, approximately ten thousand different types of proteins believed to be deeply connected to the onset of disease. To this end, the Institute of Physical and Chemical Research is playing a central role in Japan and analysis work has begun through the implementation of analysis equipment, such as cutting-edge nuclear magnetic resonance equipment and x-ray analysis equipment. Bioinformatics is utilized to predict the function of proteins, based on the protein structure calculated from the analysis of data provided by this equipment.

Pharmaceutical companies are participating in this research because, if the structures of proteins related to disease become known, then it will be possible to engineer new drugs to act on such disease. Their aim, to ensure an advantageous position in their business, is to decipher the protein structure at the earliest possible time, determine the functions of such structures and apply for a patent.

Bioinformatics of the third stage is known as the "system biology" and is intended for understanding life systems. Through genome research, information is about to become known concerning genes related to the growth occurrence of an individual, their metabolism and immune system, as well as brain functions. The research of system biology integrates all such information on the computer and targets the clarification of the mechanism that triggers the onset of disease. An example is the attempt to create a virtual cell on the computer. Prof. Masaru Tomita of the Keio University and his associates, are developing a cell simulation system called "E-Cell". They have been successful in simulating the metabolism of a single-cell organism with 127 genes, a most simple genome. Humans, however, are multicellular organisms with approximately 60 trillion cells, which makes the size of the genome large also. When stretched out the DNA included in each cell, it measures about one meter in length. This means that for a single human being an enormous amount of information is contained in the DNA, which has a total length of 60 trillion meters, a distance roughly equivalent to the diameter of our solar system. Although all cellular information is not directly related to life activities, a new concept of the information processing technology is needed in order to analyze such a complex system. Furthermore, in terms of hardware, development for large-scale database technology and high-speed computation technology is vital, as these are both technologies essential for supporting bioinformatics.

The genome research proceeded very early on through the utilization of the Internet. For this reason, numerous databases of genome deciphering results, by publicly funded research organizations including those open to public access, have been disclosed on the Internet. As for the details of bioinformatics, descriptions can be found on the page in the Genome Net, listed as the fifth reference [5] in this paper, offering a study that starts from a very basic level using an actual database. Further, the sixth reference [6], a published book, covers the same information while the seventh reference [7] provides examples of practical activities being conducted by electronics manufacturers.

Semiconductor technology and genome (genome chip)

Device technology, represented by the semiconductor process technology, is also a field where the relationship with the genome technology will continue to develop further in the future with the development of an industry expected. The market scale forecast [8] of post-genome-related technologies in individual industrial segments, indicates that the domestic market which combines the DNA chips, protein chips [9] and biosensors [10] will greatly expand from 49 billion yen in the year 2000, to 2.3 trillion yen in 2010. The main reason behind this is due to the chip developed for research, which is expected to spread into practical applications in clinical laboratories and for diagnostic use. The current annual medical care expenditure in Japan is approximately 30 trillion yen and this expenditure is expected to increase as society's declining birth rate and aging population continues, which makes this a very urgent issue with some type of inhibiting measure necessary to limit the cost of medical care. Substantiated results, indicating a reduced medical care expenditure per person, exists in municipalities where tests are intended for the early detection of disease arising from life styles. As the examples show, measures for the future are believed to shift toward medical care with an emphasis on early detection, such as measures for cancer detected through tests and diagnostics, rather than symptomatic treatment that is sought after a person becomes ill.

Genome chips are considered to include particular chips that detect substances directly related to genome information, such as DNA chips and protein chips, but also environment analysis chips and food inspection chips, as well as biosensors that take the measurement of blood sugar levels and pathogens. Because life's activities are conducted based on information that is statically encoded in genome, an entity receives various influences from many sources, such as the environment, in which the entity actually lives, or the food that is eaten as well as stress derived from sociogroups to which the entity belongs. Such information is conveyed to the genome and proteins are produced, as required. Such a dynamic facet is also considered to be important.

So how will the development of such genome chips proceed? In all cases, the subject of detection will be the DNA and proteins, environmental hormones, dioxins, other biomolecules and chemical substances. Currently, the precision analysis equipment that detects these substances is both expensive and large. Although the miniaturization of DNA chips is progressing, the series of actual processes, such as the extraction of necessary components from collected cells, or the preprocessing for taking measurements, require individual equipment and no single piece of equipment has been developed to complete the entire set of processes. Under these circumstances, an approach for the development of the next generation genome chips, is believed to be promising and thus drawing attention is the microchip technology called the "laboratory-on-the-chip" [11]. This technology was derived from research conducted in the field of chemistry, which involved, for example, the creation of a microscopic channel (also known as microchannel) on a glass plate to perform chemosynthesis and analysis, which had previously been conducted in flasks and test tubes. Highly efficient and high-speed chemosynthesis has become possible through the appropriate design of microchannel structures that control the flow of fluids. In the field of silicon semiconductors, on the other hand, the development of a technology for manufacturing a microscopic electromechanical system, called MEMS (Micro Electro Mechanical System), is progressing which will make it possible to form mechanical components, such as pumps and valves, on silicon substrates.

By fusing these technologies, it will be possible to detect various values related to health. The collection of a sample of blood, for example, is performed by suction through a microscopic needle, which causes no pain. From the sample taken, the necessary components are extracted and separated and various sensors, integrated on the same silicon substrate, are utilized to obtain measurements for blood sugar levels, cholesterol counts, or proteins used as the index for liver functions. When such chips evolve even further, a health care chip [12] is expected to be developed that will conduct information processed with the system LSI and access a database on the network by utilizing its low power wireless communication capability. Such devices may in the future be used for diagnostic and health care systems for clinical and home use.

Conclusion

The genome technology, which is experiencing quantum leaps, will continue to fuse even more with information technology and semiconductor technology, as well as nano technology in genome technology, opening up brand new technologies and industries [13]. An example of such a development and the talk of the time, is the fusion with the ubiquitous technology, with the concept of network access anywhere any time. Currently this involves information derived by physical sensors, such as locations, temperatures, sounds and images. Once the sensing capabilities for biomolecules and chemical substances are given to the ubiquitous device, our surrounding environment and living organisms, such as people like us, will become fused together with computers and networks. There are expectations that such technology will make it possible to obtain personal health care and security management, an emerging new communication method, as well as have an impact environmental protection by being able to offer extremely precise monitoring of the environment [14].

Dr. James Lovelock once said, “our planet is a biological organism, not merely a rock”. It would not be an overstatement to say that by utilizing new technologies that emerge from the fusion of genome technology and electronics, we can create and evolve the “genome” of the earth as a living organism, with “genes” to resolve problems such as international peace or global environmental protection.

References

1) http://www.meti.go.jp/policy/bio (a web page concerning bio-policies of the Ministry of Economy, Trade and Industry).
2) Hiroki Yomogita et al: “Bio premonitions - Enlisting flourishing biotechnology to electronics”, Nikkei Electronics, pp. 73-83, January 7 2002.
3) Editors of Nikkei Science: “Medical innovations in the post-genome era”, Extra Issue of Nikkei Science No. 139, 2002.
4) Toshiyuki Gotoh, et al. editors: Special issue on “Science and technology that support the medical care and welfare of the 21st Century”, Journal of the Institute of Electronics, Information and Communication Engineers, Vol. 84, No. 5, pp.281-286, 341-367, 2001.
5) http://www.genome.ad.jp/Japanese/about_J.html (the GenomeNet)
6) Toshihisa Takagi, et al. editors: “Bioinformatics and Information Biology (post sequence genome science, vol. 6)”, Nakayama Shoten publication, 2000.
7) Hiroya Madarame et al.: “Special issue : life science”, FUJITSU, Vol. 53, No. 5, September 2002.
8) http://www.jpo.go.jp/shiryou/toushin/chousa/pdf/genome.pdf (Report on the investigative analysis of trends of patent application technologies, Japan Patent Office)
9) Yosuke Kinouchi et al. editors: A feature on “Getting ready for the blossoming of the microtechnology”, Journal of the Institute of Electronics, Information and Communication Engineers, Vol. 85, No. 10, pp.715-727, 2002.
10) Isao Karube et al.: “Current status and future development of biosensors”, Oyo Buturi, Vol. 70, No. 12, pp.1407-1413, 2001.
11) Yoshinobu Baba: “Micro-and nanobiochip technology”, Oyo Buturi, Vol. 71, No. 12, pp.1481-1487, 2002.
12) Yasuhiro Horiike: Development of superior microfabrication technology for medical fields”, Nikkei Sentangijutsu (Nikkei Leading Edge Technology), No. 9, pp.35-38, March 11 2002.
13) Tomoji Kawai: “Japan can be the leaders of nano biotechnology - incorporating the strengths in mechanical and electrical industries”, Nikkei Bio Business (Nikkei Biotechnology & Business), pp.80-83, December 2002.
14) Kiyoshi Itao: “Microsystems technology for wearable information systems”, Oyo Buturi, Vol. 70, No. 8, pp.932-940, 2001.

(All Refs. are in Japanese)

Authors

Yasuo Miyamoto: Research and Development Center, Business Research Laboratory