Review Article

OVERVIEW OF THE HUMAN GENOME PROJECT IN JAPAN

Yoshiyuki SAKAKI

*Human Genome Center, Institute of Medical Science, The University of Tokyo,
4-6-1 Shiroganedai, Minato-ku, Tokyo 108, Japan*

More than five years have passed since the Human Genome Project (HGP) started in Japan. Aiming to “read and understand” all the genetic information written in the human genome, the HGP in Japan has established the infrastructures including several research centers and computer networks and has developed various types of researches. Here, I will summarize various activities and scientific accomplishments which have been done for these five years in Japan, and also discuss about the new direction/targets for the coming five years.

Organization

Human Genome Project (HGP) in Japan has been supported by the Science Technology Agency (STA), the Ministry of Education, Science, Sports and Culture (MESSC, Monbu-sho), and also the Ministry of Health and Welfare (MHW). Roughly speaking, Monbu-sho project covered wide-range of the HGP, that is, the mapping, cDNA analysis, technology development and informatics, but STA focused on the map construction and MHW on the disease genes. In addition, the Ministry of Agriculture, Forestry and Fisheries (MAFF) has developed the “Rice Genome Project” and Ministry of International Trade and Industry (MITI) has recently started the analysis of the microorganism genomes useful from the industrial viewpoints. Each Ministry or Agency has promoted the Project by research grants or funds to individual and group researchers and also by the activities of the Research Centers. Total budget for these activities is 5101M yen (about 50M USD) in 1995FY. These genome activities are coordinated by the Human Genome Committee under the Science Council Japan. The Committee consists of the leaders of each project and the officials of each Ministry and Agency, and is now chaired by Dr. Wataru Mori, the member of the Science Council Japan.

Scientific Accomplishments in Phase I

Japan has taken several different types of approaches for the HGP, which seem

similar to those taken by USA, UK and other countries: 1) Genetical and physical analysis of the human genome, 2) Functional analysis of the human genome, 3) Analysis of the non-human genomes as a model of the human genome, 4) Development of the technology for the genome analysis, and 5) Development of the informatics (computer science) for the genome analysis. Through those approaches, significant progress and development have been made for these five years to establish the basis of the HGP in Japan.

1) The HGP in each country has focused on the construction of genetic and physical maps of the human genome. Among the maps, the fine genetic map using thousands of microsatellites and the backbone YAC contig (physical) map constructed by CEPF/Genethone group are most remarkable ones. In Japan, Ohki and his colleagues constructed *NotI* restriction maps of the chromosome 21q and 11q (Ichikawa *et al.*, 1993). Those are the first chromosome-wide restriction maps and has been recognized as a solid reference for various types of physical maps including YAC contig map. The maps have been also applied for the analysis of the translocations associated with leukemia/cancer. Indeed, Ohki and his colleagues succeeded in identifying a new gene AML-1 for acute myeloid leukemia (Miyoshi *et al.*, 1991). Nakamura and his colleagues have mapped thousands of cosmid clones by FISH on chromosome 3, 6, 8, 11, 17 and so forth (for example, Nakamura *et al.*, 1994), which have provided a strong basis for positional cloning of the disease genes. By using these maps, more than 10 disease genes have been mapped and identified, including those for cancers (APC, AML-1), neuronal diseases (DRPLA, Machado-Joseph disease) and so forth (Nishisho *et al.*, 1991; Nagafuchi *et al.*, 1994; Koide *et al.*, 1994; Hirotsume *et al.*, 1995).

2) Another remarkable progress in the HGP is a very large-scale collection of cDNA sequences. For example, a group represented by Craig Venter has established a database of more than 170,000 (partial) cDNA sequences and a group of Washington University supported by Merck has registered 220,000 or more (partial) cDNA sequences termed EST (expressed sequence tag) in the public database (GenBank). From the very beginning of HGP in Japan, particular emphasis has been placed on the functional aspects of genome analysis. A group was organized for the production of high-quality cDNA libraries and the systematic collection of information on expression patterns of individual genes. Among the many achievements of this group, the body-mapping approach by Matsubara, Okubo and their colleagues (Okubo *et al.*, 1992, 1995) which catalogues genes according to their expression extent in individual tissues or cells, has a profound impact on the fields of genome/cDNA research. Also, the development of unique techniques for the construction of specialized cDNA libraries such as full-length clone libraries (Maruyama and Sugano, 1994) is expected to accelerate the functional analysis of novel genes uncovered by the advent of genome analysis.

3) Human is not an experimental organism, so that for understanding the human genome it seems essential to combine the knowledge obtained from

experimental organisms with those from the human beings. The genomes of various bacteria, yeast, nematode, fly and mouse have been extensively studied. Particularly, the complete nucleotide sequence analysis of *Saccharomyces cerevisiae* is the most remarkable achievement for these five years. In Japan, the genomes from several organisms have been extensively analyzed, including *Bacillus subtilis* (Ogasawara *et al.*, 1994, 1995), *S. cerevisiae* (Murakami *et al.*, 1995), *Caenorhabditis elegans* and rice.

4) Various new, powerful technologies for the genome analysis have been developed to accelerate the HGP. Those developed by Japanese groups include the genome-wide scanning of the structural alterations, the novel strategies for large-scale sequencing and high throughput scanning of transcripts (Hatada *et al.*, 1991; Ito *et al.*, 1994).

5) The importance of informatics for the success of genome projects is well recognized and has been separately funded to efficiently start up the novel discipline of genome informatics. Japanese group has focused not only on the development of database, tools and softwares for various aspects of genome research but also on leading studies for the integration of genome data with various aspects of biological knowledge ranging from cellular signal transduction mechanisms to the life cycle of an organism and the origin and evolution of species. It should be emphasized that this group has also made a profound contribution to the facilitation of collaborations between biologists and computational scientists and between computational and experimental laboratories by providing various databases and computational services over the computer network, Genome Net.

From Phase I to Phase II

In 1994, The Human Genome Committee prepared the Report for the Promotion of the second phase of the HGP in Japan. The Report emphasized the following points: 1) large-scale sequencing of the human genome including the construction of high resolution physical maps for the sequencing, 2) functional analysis of the genome including cDNA analysis and the analysis of model organism genomes, 3) development of new technology for the genome analysis including the computer technology for the genome analysis.

According to the Report, each Ministry or Agency is now taking actions to promote the Phase II of HGP in Japan. STA through JICST (Japan Information Center for Science and Technology) launched a three-year grant for sequencing the human genome in a megabase scale, and also RIKEN will start a new project focusing on the functional analysis of the (mouse) genome. Monbu-sho (MESSC) started a new five year grant for the Human Genome Analysis (1996–2000) which is focusing on 1) sequence-based analysis of genome structure, 2) functional analysis of the genomes, and 3) bioinformatics. Monbu-sho also opened two new laboratories for genome sequencing in the Human Genome Center at Institute of Medical Science, The University of Tokyo in 1996. Certainly “sequencing” is a key

word of the Phase II of the HGP in Japan.

REFERENCES

- Hatada I, Hayashizaki Y, Hirotsune S, Komatsubara K, Mukai T (1991): A genomic scanning method for higher organisms using restriction sites as landmarks on the genome. *Proc Natl Acad Sci USA* **88**: 9523-9527
- Hirotsune S, Takahara T, Sasaki N, Hirose K, Yoshiki A, Ohashi A, Kusakabe M, Murakami Y, Muramatsu M, Watanabe S, Nakao K, Katsuki M, Hayashizaki Y (1995): The reeler gene encodes a protein with an EGF-like motif expressed by pioneer neurons. *Nature Genet* **10**: 77-83
- Ichikawa H, Hosoda F, Arai Y, Shimizu K, Ohira M, Ohki M (1993): A *NotI* restriction map of the entire long arm of human chromosome 21. *Nature Genet* **4**: 361-366
- Ito T, Kito K, Adati N, Mitsui Y, Hagiwara H, Sakaki Y (1994): Fluorescent differential display: arbitrarily primer RT-PCR fingerprinting on an automated DNA sequencer. *FEBS Lett* **351**: 231-236
- Koide R, Ikeuchi T, Onodera O, Tanaka H, Igarashi S, Endo K, Takahashi H, Kondo R, Ishikawa A, Hayashi T, Saito M, Tomoda A, Miike T, Naito H, Ikuta F, Tsuji S (1994): Unstable expansion of CAG repeat in hereditary dentatorubral-pallidoluysian atrophy (DRPLA). *Nature Genet* **6**: 9-13
- Maruyama K, Sugano S (1994): Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. *Gene* **138**: 171-174
- Miyoshi H, Shimizu K, Kozu T, Maseki N, Kaneko Y, Ohki M (1991): t(8; 21) breakpoints on chromosome 21 in acute myeloid leukemia are clustered within a limited region of a single gene, AML1. *Proc Natl Acad Sci USA* **88**: 10431-10434
- Murakami Y, Naitou M, Hagiwara H, Shibata T, Ozawa M, Sasanuma S, Sasanuma M, Tsuchiya Y, Soeda E, Yokoyama K, Yamazaki M, Tashiro H, Eki T (1995): Analysis of the nucleotide sequence of *Saccharomyces cerevisiae* chromosome VI. *Nature Genet* **10**: 261-268
- Nagafuchi S, Yanagisawa H, Sato K, Shirayama T, Ohsaki E, Bundo M, Takeda T, Tadokoro K, Kondo I, Murayama N, Tanaka Y, Kikushima H, Umino K, Kurosawa H, Furukawa T, Nihei K, Inoue T, Sano A, Komure O, Takahashi M, Yoshizawa T, Kanazawa I, Yamada M (1994): Dentatorubral and pallidoluysian atrophy expansion of an unstable CAG trinucleotide on chromosome 12p. *Nature Genet* **6**: 14-18
- Nakamura Y, Okui K, Takahashi E, Koyama K (1994): Isolation and mapping of 328 new cosmid markers on human chromosome 8: construction of a high resolution cytogenetic map of chromosome 8 with 416 markers. *Cytogenet Cell Genet* **65**: 115-118
- Nishisho I, Nakamura Y, Miyoshi Y, Miki Y, Ando H, Horii A, Koyama K, Utsunomiya J, Baba S, Hedge P, Markham A, Krush AJ, Petersen G, Hamilton SR, Nilbert MC, Levy DB, Bryan TM, Preisinger AC, Smith KJ, Su L, Kinzler KW, Vogelstein B (1991): Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. *Science* **253**: 665-669
- Ogasawara N, Nakai S, Yoshikawa H (1994): Systematic sequencing of the 180 kilo bases region of the *Bacillus subtilis* chromosome containing the replication origin. *DNA Res* **1**: 1-14
- Ogasawara N, Fujita Y, Kobayashi Y, Sadaie Y, Tanaka T, Takahashi H, Yamane K, Yoshikawa H (1995): Systematic sequencing of *Bacillus subtilis* genome: progress report of the Japanese group. *Microbiology* **141**: 257-259
- Okubo K, Hori N, Matoba R, Niiyama T, Fukushima A, Kojima Y, Matsubara K (1992): Large scale cDNA sequencing for analysis of quantitative and qualitative aspects of gene expression. *Nature Genet* **2**: 173-179
- Okubo K, Itoh K, Fukushima A, Yoshii J, Matsubara K (1995): Monitoring cell physiology by expression profiles and discovering cell type-specific genes by compiled expression profiles. *Genomics* **30**: 178-186