Review

A Japanese history of the Human Genome Project

By Yoshiyuki SAKAKI*1,*2,*3,†
(Edited by Takao SEKIYA, M.J.A.)

Abstract: The Human Genome Project (HGP) is one of the most important international achievements in life sciences, to which Japanese scientists made remarkable contributions. In the early 1980s, Akiyoshi Wada pioneered the first project for the automation of DNA sequencing technology. Ken-ichi Matsubara exhibited exceptional leadership to launch the comprehensive human genome program in Japan. Hideki Kambara made a major contribution by developing a key device for high-speed DNA sequencers, which enabled scientists to construct human genome draft sequences. The RIKEN team led by Yoshiyuki Sakaki (the author) played remarkable roles in the draft sequencing and completion of chromosomes 21, 18, and 11. Additionally, the Keio University team led by Nobuyoshi Shimizu made noteworthy contributions to the completion of chromosomes 22, 21, and 8. In April 2003, the Japanese team joined the international consortium in declaring the completion of the human genome sequence. Consistent with the HGP mandate, Japan has successfully developed a wide range of ambitious genomic sciences.

Keywords: Human Genome Project, genome sciences, DNA sequencing, chromosome 21, Human Genome Research Center, RIKEN Genomic Sciences Center

Introduction

The international Human Genome Project (HGP) achieved a monumental breakthrough — the decoding of the entire human genome, our own genetic blueprint, and opened a new chapter in medicine and human life sciences.

This historic project was initially contemplated by leading scientists such as Renato Dulbecco1) and proposed in the mid-1980s by the U.S. Department of Energy2) and later by the National Institute of Health in the U.S. Because the proposed project was so huge, it was subjected to extensive deliberation concerning not only scientific and technological aspects but also its social and economic viability.3) Finally, on the recommendation of the National Research Council,4) the HGP was established as a national project in the U.S. (http://www.esp.org/misc/genome/firstfiveyears.

pdf#searchFHumanDGenomeDProject%2CDUSA%2CD1991).

In 1981, a Japanese researcher Akiyoshi Wada proposed a pioneering project towards the automation of DNA sequencing technology, an essential step for the HGP. Despite Wada’s pioneering work, Japan failed to launch a unified national project. Instead, the Ministry of Education, Science and Culture (Monbusho) and the Science and Technology Agency (STA) independently launched new programs on human genome research. Other countries also launched human genome research programs, and the HGP started in 1991 as an international cooperative project between the U.S., U.K., Japan, France, and Germany (China joined later).

The HGP was designed as a project in three stages, specifically, a genome mapping stage, a whole genome sequencing stage, and a whole genome sequence annotation stage. Towards these goals, the international consortium introduced various new technologies, concepts, and schemes, which brought about revolutionary changes in life sciences. For example, the project actively incorporated the power of engineering, particularly the automation of DNA sequencing technology, which enabled scientists to...
generate large amounts of data and brought about a new style of the life sciences driven by huge sets of data. Such data-driven research elevated bioinformatics as a scientific discipline in life sciences. In addition, the consortium introduced a new open data-sharing concept called the Bermuda Principles for successful international cooperation (https://web.ornl.gov/sci/techresources/Human.Genome/research/bermuda.shtml). During the course of the project, the consortium faced various challenges such as that presented by a private company Celera Genomics. Nevertheless the consortium overcome such impediments, and finally declared the completion of the human genome in April, 2003.

The history of the international HGP has already been well described in various forms, mainly from the viewpoints of the U.S. and U.K., the major players in the Project. However, little has been expressed from the viewpoints of other founding members. In this regards, this review is unique and provides insight into achievements of key Japanese researchers during the HGP. Japan was one of the founding members of the HGP, and it made significant contribution from its earliest stage to the final stage. These contributions include the pioneering challenge of Akiyoshi Wada for the automation of DNA sequencing technology, the leading roles of the Japanese team in the completion of chromosomes 21 and 11, and also the successful development of a key device for high-speed multi-capillary DNA sequencers by Hideki Kambara. Moreover, in concert with the international project, the Japanese genome community successfully organized the domestic alliances necessary to promote a wide range of genome sciences, including unique functional cDNA programs, a variety of medical genomics, and the development of unique functional genome databases.

**Prologue to the Human Genome Project in Japan**

“Wada Project”. Japanese history in the HGP began with the so-called “Wada Project” which proposed, for the first time, the automation of DNA sequencing technology.

DNA sequencing technology developed in the late 1970s brought about a revolution in molecular biology. Nevertheless it required exceptional skills, and the production rate of one skilled researcher was only about 1000 bases per day. To sequence even a small gene, tedious and monotonous work was required.

In 1981, Wada proposed the automation of DNA sequencing technology by introducing high-quality robot technologies from Japanese industry (Fig. 1). Wada argued that this would not only release...
researchers from laborious tasks but also strongly promote life sciences and biotechnology. Wada’s proposal was strongly supported by the Special Coordination Funds of the Science and Technology Agency for Promoting Science and Technology, and carried out with cooperation between universities and companies such as Hitachi, Fuji Photo Film, and Seiko. The budget for the first three years was around 900 million yen. In 1987, Wada published a report titled “Automated high-speed DNA sequencing” in Nature, which stated that the technology developed by the project has reached a sequencing rate of one million bases a day at a cost of 9.2 yen per base.

In the same year, Wada organized an international workshop on automated DNA sequencing at Okayama under the auspices of Hayashibara Foundation. The impact of his report in Nature was enough to draw many leading scientists and engineers including Walter Gilbert (Nobel Prize Laureate), Leroy Hood (the leader in DNA sequencing automation in the U.S.), and Lennart Phillipson (the Director of the European Molecular Biology Organization) (Fig. 2). The participants actively discussed the advancement of automated DNA sequencing technologies as well as the impact of this technology on the HGP.\(^6\)

Wada also proposed an ambitious vision for an international DNA sequencing center equipped with automated high-speed DNA sequencing lines. In his report in Nature,\(^6\) he described “…..we foresee that DNA-sequencing super center will be set up in several countries and will become symbols of the effort of nations to broaden and build on human knowledge…….”

Based on his proposal, the construction of an automated high-speed DNA sequencing line started in 1988 at RIKEN. Around the same time, the U.S. company Applied Biosystems, and slightly later Hitachi, released compact, laboratory-size automated DNA sequencers to the market. These compact DNA sequencers have been widely utilized by the international genome research community. The RIKEN system called HUGA-1 was completed in 1991, but very regrettably it seemed too late. The compact DNA sequencers had already been utilized in most genome research laboratories/centers. Thus there remained less opportunity for the RIKEN system to contribute to genome sequencing.

Wada’s ambitious proposal of an international DNA sequencing center was realized in a different style by the big genome centers in the U.S. and U.K. in the mid-1990s and the RIKEN Genomic

![Fig. 2. Group photo of the participants of the Hayashibara Forum: International Workshop on Automatic and High Speed DNA-Base Sequencing in July 1987. Wada, Hood, Tom Caskey, Charles Cantor, Kambara, and Shimizu are seen at the front row.](image-url)
Sciences Center (GSC) in Japan in 1998, which is described below.

Towards the Japanese Human Genome Project. In spite of Wada’s pioneering proposal, the concept of the HGP was established in the U.S. and shared with various countries including Japan. In Japan, several Ministries and Agencies such as the Science and Technology Agency (STA), Ministry of Education, Science and Culture (Monbusho), Ministry of Health and Welfare (MHW), and Ministry of International Trade and Industry (MITI) took actions for the HGP. In 1989, the recommendation for the promotion of the Human Genome Project was published by the Science Council of Japan (http://www.scj.go.jp/ja/info/kohyo/12/14-02-k.pdf). However, Japan failed to establish a unified national program, partly due to bureaucratic sectionalism.

The slow progress by Japan in the international HGP seemed to concern James Watson, the leader of the U.S. and international HGPs. In July 1989, he wrote a letter to Ken-ichi Matsubara, a Japanese leader of the genome project, threatening to deny Japanese scientists access to DNA sequence databases unless Japan pays its “fair share” for the genome project.11)

Under such severe criticism, Monbusho established a comprehensive human genome research program under the exceptional leadership of Matsubara. Through the 2-year pilot study, Matsubara and colleagues, including Mitsuoki Yoshida, Hiroshi Yoshikawa, Minoru Kanehisa, Mitsuru Takanami, and the author (Fig. 3) shared a common conviction that the stream of the HGP should bring about revolutionary changes in the framework of life sciences in Japan in the near future, and proposed a comprehensive (human) genome research program, which included not only structural analysis of the human genome but also a wide range of functional analyses (Fig. 4). In addition, the proposal emphasized the importance of bioinformatics for the future development of genome sciences. Furthermore, the establishment of the Human Genome Research Center was recommended as a core facility to promote these programs.

The proposal was approved by the policy making committee of Monbusho, and two new programs were established under the Grant-in-Aid Scientific Research (KAKENHI) system: the “Human-genome analysis” program (principal investigator [PI]: K. Matsubara) and the “Genome information” program (PI: M. Kanehisa). These two programs started in 1991, with the participation by more than 100 researchers. The “Human-genome analysis” program was financially well-supported by a new 5-year special program of KAKENHI (budget was around 500 million yen per year). The budget for the “Genome information” program was around 100 million yen per year.
In contrast, STA launched a RIKEN-based human genome analysis program that sharply focused on several chromosomes of the human genome by a limited number of researchers, including Misao Ohki, Yusuke Nakamura, Nobuyoshi Shimizu, and the author. The contribution of the STA to the HGP was limited in the early stage, but in the late 1990s, the STA provided the powerful support to the HGP, as discussed later.

Beyond the human genome, the Ministry of Agriculture, Forestry and Fisheries (MAFF) planned the Rice Genome Project, MITI proposed the development of biotechnology instrumentation for the genome study, and MHW planed a program for the analysis of genes related to aging and diseases, particularly oncogenes.

In line with international efforts to decode the human genome, the necessity of an organization in charge of international cooperation was recognized, and an international organization called HUGO (Human Genome Organization) (http://www.hugointernational.org/) was founded in 1988. Under the presidency of Victor McKusick, HUGO started its activity with the participation of 220 scientists from 23 countries in 1989.12) From Japan, Ken-ichi Matsubara was nominated as the Vice President of HUGO and also nine more scientists participated in HUGO activities, including Nobuyoshi Shimizu and Tasuku Honjo as founding members and also Mitsuki Yoshida, Minoru Kanehisa, Akiyoshi Wada, Susumu Nishimura, Yoji Ikawa, Michio Oishi, and the author.

First stage of the HGP (1991–1995)

Outline. The first stage of the HGP focused on the construction of genome maps. In Japan, map construction was conducted as part of the Monbusho “Human-genome analysis” program and the RIKEN-based STA program. In addition to map construction, the Monbusho program promoted a wide range of genome sciences, including the functional analysis of human and experimental organisms and technology development for genome researches. Furthermore, bioinformatics, an emerging key player in genome sciences, was promoted by the other Monbusho program “Genome information” led by Kanehisa. In order to support and promote these new activities in genome sciences, the Human Genome Research Center was established in 1992 at the Institute of Medical Sciences, the University of Tokyo. Overall, Japan successfully constructed and developed a well-balanced platform for the genome sciences during the first stage of the HGP.

Construction of the human genome map. In the first stage of the HGP, Japanese researchers focused on map construction for specific chromosomes, including chromosomes 6, 8, 11, 21, and 22. Among them, progress in chromosome 21 mapping was notable. The NotI restriction map of chromosome 21 by Misao Ohki’s group13) was noted as a landmark achievement in the international community. Sequence-tagged NotI and SfiI linking clones by Ohki’s and the author’s teams provided solid anchor-points for construction of a chromosome 21 yeast artificial chromosome (YAC) contig map by an international team led by Daniel Cohen.14) Furthermore, the group of Nobuyoshi Shimizu constructed unique, chromosome-specific libraries by sorted chromosomes 21 and 22 for map construction.15) Yusuke Nakamura and his team was the most active group in clone-based map construction. They focused on chromosomes 6 and 8, where various genetic diseases and tumor suppressor genes were mapped.16)
Promotion of functional genomics. The Monbusho program also provided significant achievements in functional genomics, particularly in cDNA analysis. For example, the oligo-cap method by Sumio Sugano and Kazuo Maruyama was a profound breakthrough for full-length cDNA cloning. Traditional cDNA synthesis started from the oligo A-tail at the 3′ end of mRNAs. However, this new method enabled cDNA synthesis to start from the cap-structure at the 5′ end of the mRNA. This method was widely employed for full-length cDNA library constructions, particularly in late 1990s by the MITI-supported Helix Research Institute as described later. The “Body Map” project by Kousaku Okubo and colleagues provided a new and unique concept in functional genomics. They collected the mRNA-based gene expression profiles (map) of more than 100 tissues and cells of human and mouse.

The Monbusho program also promoted the genome analysis of experimental model organisms, including Caenorhabditis elegans, Saccharomyces cerevisiae, and Bacillus subtilis. Among these, the accomplishments of Naoki Ogasawara and colleagues was noteworthy. They completed the sequencing and functional annotation of B. subtilis genome in collaboration with a French group.

The Human Genome Research Center and promotion of bioinformatics. The Human Genome Research Center, equipped with automated DNA sequencers (Fig. 5) and ultra-parallelled computers, was founded at the Institute of Medical Sciences, the University of Tokyo in 1992 as the core facility to promote the Monbusho genome programs. The center consisted of three laboratories, specifically the Genome Database Laboratory led by Toshihisa Takagi, the Human Genome Analysis Laboratory led by the author, and the Genome Informatics Laboratory led by Minoru Kanelhisa. The center played leading roles in establishment of platforms for genome sciences in Japan during the first stage of the project. In particular, the Genome Database and the Genome Informatics Laboratories, together with researchers from the “Genome information” program, constructed a solid platform of genome informatics, including many tools for the genome analysis (http://hgc.jp/japanese/software.html) and a variety of databases (http://hgc.jp/japanese/database.html). In addition, they successfully recruited bioinformatics researchers by organizing the training courses and tutorials.

International cooperation and collaboration. A number of international collaborative researches, international meetings, workshops, and conferences were held by the HGP. Among them, the international meeting HGM93 and the chromosome 21 sequencing consortium should be mentioned as activities in which the Japanese research community took leading roles.

The HGM (initially Human Gene Mapping, later Human Genome Meeting) was one of the largest and the highest quality meetings in genomics. The HGM was held annually in the U.S. since the 1980s. However, the 1993 meeting (HGM93) was held in Kobe on November 14–17, 1993. The meeting chaired by Ken-ichi Matsubara had more than 700 participants, including nearly 400 foreign participants. The guest speakers included Victor McKusick, Leroy Hood, Francis Collins, Mark Lathrop, and Jean Weissenbach. The meeting was a great opportunity to demonstrate the progress of the Japanese genome activity to the international community and as well as for Japanese researchers to learn the cutting-edge genomics in the U.S. and Europe.

At the next meeting of HGM at Washington D.C. in October 1994, however, the atmosphere was quite different from the previous meeting in Kobe. Most of the leaders of the big genome centers in the U.S. and U.K. presented their plan and progress towards the coming large-scale DNA sequencing stage of the HGP. The international project was clearly evolving towards this next stage. It seemed obvious that Japan may fall behind without suitable action. In response, Yoshiyuki Sakai (the author), the director of the Human Genome Research Center at the University of Tokyo and the head of the Human

Fig. 5. The DNA sequencing laboratory in the Human Genome Research Center at the Institute of Medical Sciences, the University of Tokyo, in 1992.
Genome Research Group of the Monbusho program, immediately proposed, along with David Patterson, to initiate an international consortium for sequencing chromosome 21. Chromosome 21 was the most well-mapped, symbolic chromosome for the Japanese genome community. The proposal was favorably approved by the international chromosome 21 research community, and the consortium started under the leadership of Japanese and German researchers in November 1994 (Fig. 6). This consortium became an important strategic base for Japan to expand its activity in the HGP, as described later.

**ELSI and other domestic activities.** The HGP gave considerable weight to the ethical, legal, and social issues (ELSI) surrounding decoding human genetic information, the ultimate personal information. The “Human-genome analysis” program allotted nearly 2% of its budget for ELSI endeavors. A unique ELSI-related activity was a series of tutorials for journalists and science writers, which provided a precious opportunity for genome researchers to communicate with public via mainstream media on the significance, impact, and progress of the HGP.

Similarly, the Japan Society for the Promotion of Sciences (JSPS) backed a series of special forums for leading genome scientists to discuss the promotion and future directions of genome sciences in Japan.

**Second stage of the HGP (1996–2000)**

**Outline.** In the second stage of the HGP, research efforts centered upon the sequencing of the human genome. The international consortium had a meeting in Bermuda in February 1996, in which the important principles called the Bermuda Principles for the rapid release of DNA sequence data to the public were proposed and agreed.

By 1998, a dramatic evolution in sequencing occurred with the invention of the high-performance multi-capillary sequencer using technology developed by Hideki Kambara, which brought one-order higher sequencing capability than the traditional approach. This innovation led a company, Celera Genomics, to declare the whole genome sequencing at a commercial level. This caused serious conflict between the public and the private sectors. Under unusual competition between the two sectors, two draft sequences of the human genome were constructed.

In order to keep up with the phase shift to large-scale sequencing, the RIKEN Genomic Sciences
Center was established in Japan in 1998. The center made a significant contribution to the human genome draft sequencing. Furthermore, the RIKEN Center, in cooperation with the Keio University and three German teams, took leading roles in the completion of the chromosome 21.7

In the late 1990s, genomic sciences evolved into a big stream of life sciences in Japan, and various genome-based new programs were initiated. Monbusho started a new 5-year program “Genome Science: New Frontiers in Biosciences”, and MITI supported the establishment of the Helix Research Institute, a joint venture with ten pharmaceutical companies. MAFF stepped up the Rice Genome Project to the whole genome sequencing stage.22) Furthermore, many new programs including the SNP Research Center at RIKEN started under the Millennium Project of Japan.

International Human Genome Sequencing Consortium and the Bermuda Principles. The HGP shifted from the initial mapping stage to the second sequencing stage around 1995, and in February 1996, the leaders of the HGP from each country and big centers gathered in Bermuda towards better cooperation for the whole human genome sequencing. From Japan, Matsubara, the author, and Nakamura were initially invited to attend, but were unable to attend due to other commitments. Instead Naoki Ogasawara and Masahira Hattori attended the meeting as the Japanese representatives. At the meeting a series of important principles, the Bermuda Principles20), were proposed. Most importantly, sequence assemblies of 2 kb or more should be released within 24 hours to the international public DNA sequence databases (GenBank, EMBL, and DDBJ) without any restriction. The centers and groups that agreed with the Bermuda Principles formed the International Human Genome Sequencing Consortium (Fig. 7). However, the Bermuda Principles were not smoothly accepted in Japan, because the principles were inconsistent with the traditional policy of the funding agency and policymakers in Japan, namely, competition-based promotion of science and technology. However, the ambitious goals of the project demanded compliance with international norms of cooperation, and the Bermuda Principles were finally accepted, and the author and Shimizu’s groups joined the consortium. At the meeting, the participants also decided the sequencing share of each country and center, in which Japan, together with Germany, claimed chromosome

Fig. 7. The International Human Genome Sequencing Consortium members in Bermuda in 1988. James Watson at the center of the front row.
21 as its first target based on the international chromosome 21 sequencing consortium established in late 1994.

**Multi-capillary sequencer and Hideki Kambara.** The slab-gel type automated DNA sequencer developed in the late 1980s significantly contributed to the early stage of the HGP. However, the data production speed was limited because a slab-gel electrophoresis system must operate with a limited high-voltage electric field to avoid Joule overheating. Such a limitation can be overcome using a capillary electrophoresis system, which allows electrophoresis with a much higher voltage electric field. The capillary electrophoresis system, however, had some technical difficulties in the simultaneous signal detection in multi-capillary array systems. Hideki Kambara, a Fellow of Hitachi, successfully developed and patented a unique and powerful method, called the “multiple sheath-flow technique”, and demonstrated simultaneous signal detection in a multi-capillary array system in 1993.9) The multi-capillary DNA sequencer equipped with Kambara’s device was commercialized in 1998 in alliance with Hitachi and Applied Biosystems (Fig. 8). The new sequencer showed more than one-order higher data production rate and brought about the dramatic advance in the HGP. However, it also triggered severe competition between the public and private sectors, as described below.

**The draft sequencing of the human genome.** In May 1998, Craig Venter announced the establishment of a venture company called Celera Genomics, with its aspiration to sequence the whole human genome at a commercial level.21) Equipped with more than a hundred multi-capillary sequencers, Celera took a strategy called whole genome shotgun sequencing approach, namely, all of the fragments of the human genome are sequenced randomly and the data are then assembled by computer to determine the whole genome sequence. Venter declared that the whole genome sequencing would be completed within years and that Celera would obtain patents for the most commercially interesting genes.

Their strategy was precipitous in contrast to the clone-by-clone, step-by-step strategy of the international consortium, which produced and released the sequence data including chromosomal positions with the highest quality for the public without any restriction for use. Celera’s action was clearly a serious rebuttal to the public welfare efforts of the international consortium. The consortium discussed the situation and adopted a clone-based whole genome shotgun sequencing approach as a temporary strategy to counter Celera’s aggression.23) In the public sector, 20 centers and laboratories participated, including RIKEN Genomic Sciences Center (GSC; Fig. 9) and Keio University from Japan.

After the severe competition for around 2 years, both sectors declared the end of the draft sequencing. Although the RIKEN GSC was a relatively late comer in comparison with the big centers of the U.S. and U.K., the RIKEN GSC made a significant contribution (the sixth largest amount of sequence data) to the draft sequence (Table 1).

A ceremony to celebrate the finishing of the human draft sequence took place in Washington D.C. on June 26, 2000 in the presence of President Bill Clinton with Francis Collins and Craig Venter in attendance. The Prime Minister of the U.K., Tony Blair also attended remotely via television. The Prime Minister of Japan, Yoshiro Mori, was unable to participate due to the general election of the House of Representative in Japan.

Summary papers on the draft sequence of the human genome were published separately from both sectors in February 2001 in *Nature*24) and *Science*,25) respectively. The draft sequence covered around 92% of the genome, around 2.7 Gb, with roughly 145,000 gaps. The accuracy of the data was estimated as 99.9% or more. The paper successfully drew the rough first landscape of the human genome, and the total protein-coding gene number was estimate around 32,000 (Table 2).
Japanese genome sequencing efforts and RIKEN Genomic Sciences Center. When the HGP entered its second stage, the Japanese contribution to the large-scale sequencing was limited due to insufficient funding in comparison to that of the U.S. and U.K. Monbusho had little funding for the large-scale sequencing. STA raised the fund for the large-scale human genome sequencing (total 400

Table 1. Major centers involved with the draft and finished sequences

| Center                                                                 | Data production (kb: kilo base) |
|-----------------------------------------------------------------------|---------------------------------|
| Whitehead Institute for Biomedical Research, Center for Genome Research (U.S.) | 1,196,888 562,096                |
| The Sanger Institute (U.K.)                                           | 970,789 919,388                 |
| Washington University Genome Sequencing Center (U.S.)                 | 765,898 645,062                 |
| U.S. DOE Joint Genome Institute (U.S.)                                | 377,998 485,085                 |
| Baylor College of Medicine, Human Genome Sequencing Center (U.S.)      | 345,125 320,733                 |
| RIKEN Genomic Sciences Center (Japan)                                 | 203,166 155,769                 |
| Genoscope (France)                                                    | 85,995 99,970                   |
| GTC Sequencing Center (U.S.)                                          | 71,367 45,710                   |
| University of Washington Genome Center (U.S.)                          | 29,115 145,745                  |

Table 2. Summary of the draft and the complete sequences

| Items                      | Draft* | Completed** |
|----------------------------|--------|-------------|
| Total sequence             | 2.69 Gb| 2.85 Gb     |
| Gaps                       | ~145,000| 341        |
| Accuracy                   | >99.9% | >99.99%     |
| Estimated gene number      | ~32,000| 20,000–25,000|

*Data from ref. 24. **Data from ref. 30. Gb: giga base.
In line with the draft sequencing efforts, the international teams made continuous efforts to complete a high-quality human genome sequence based on the clone-by-clone strategy. The first chromosome completed was chromosome 22,29) which was carried out by the U.K. Sanger Center in collaboration with the U.S. and Japanese (Keio) teams in December 1999. Six months later, in May 2000, the second completed chromosome was announced, chromosome 21, which was completed by the chromosome 21 sequencing consortium led by the Japanese and German teams.7) In particular, the RIKEN GSC team played leading roles in the completion. Although it was the second completed chromosome, it allowed us to draw a unique landscape of the human genome by incorporating data from the first finished chromosome 22. First, it showed extremely uneven distribution of genes, particularly the presence of an extremely gene poor region called “gene desert” (Fig. 10). Furthermore, the total number of protein-coding genes in the entire human genome was estimated to be close to 40,000. This estimated number was extremely low in comparison with the previous estimate of 70,000–140,000 genes and stimulated active discussion on the total gene number in the human genome. The estimated number of 40,000 genes was close to the gene number of 20,000–32,000 in the later published draft and finished human genome sequences.24),30) The completion of chromosome 21 was a landmark for Japan to show its presence in the international community.

Progress and expansion of genome sciences in Japan. Consistent with the human genome sequencing, the Japanese genome community successfully expanded and strengthened the research platform for genome sciences.

Monbusho launched a new 5-year program “Genome Sciences” (PI: the author) in 1996. The new program consisted of three research groups: the Human Genome research group; the Functional Genomics research group; and the Genome Informatics research group. The Human Genome research group focused on the analysis of functionally important genes and regions, such as genes responsible for incurable diseases,31) cancer-related regions,32) and the genomic imprinting genes.33) The Functional Genomics research group expanded the full-length cDNA analysis34) and also functional genomics of experimental model organisms such as Drosophila melanogaster, C. elegans,35) and yeast.36) The huge amount of sequence data generated by the high-performance DNA sequencers raised bioinformatics to a major discipline in genome sciences. To cope with this drastic change of circumstances, the Genome

million yen, or roughly $3.6 million per year), which was distributed to four leading teams: the Nakamura group for sequencing of cancer-related regions of chromosomes 3 and 13, the Shimizu group for chromosomes 22 and 21, the author’s group for chromosome 21 and the Hidetoshi Inoko group for the HLA region of chromosome 6. However, the scale of the funding was extremely small in comparison with the U.S. human genome budget, $243 million in fiscal year 1996 and $266 million the following year.

It was obvious that the budget of Japan was insufficient to play a significant role in the international project, so Japanese genome researchers, including the author, made vigorous efforts in fund-raising for large-scale DNA sequencing. After rigorous discussion and investigation, STA started fund-raising efforts and successfully launched the Genomic Sciences Center (GSC) at RIKEN in October 1998 (Fig. 9, https://www.gsc.riken.jp/eng/ayumi/index.html). This was an extremely important step in the Japanese history of the HGP. Without the RIKEN GSC, the contribution of Japan would have been extremely small in the international community. The Center became the flagship for Japanese genome sciences.

The RIKEN GSC was directed by Wada and comprised three research groups: the Human Genome Research Group led by the author; the transcriptome research group led by Yoshihide Hayashizaki; and the structural biology research group led by Shigeki Yokoyama. The Human Genome Research Group was equipped with more than ten high-performance multi-capillary DNA sequencers and played remarkable roles in the international HGP. As mentioned previously, the Center made the sixth largest contribution to the human genome draft sequence and also significant contribution to the final complete sequences (Table 1). Furthermore, the group played leading roles in the completion of human chromosome 21,7) as described below. The transcriptome research group developed a comprehensive and systematic mouse cDNA analysis26) and played a leading role in the transcriptome world through a project called the FANTOM.27) The structural biology research group, equipped with a large number of high-performance NMR machines, developed a systematic analysis of protein structures as part of the Protein 3000 project.28)
Informatics research group extensively developed and constructed various new tools, algorithms, and databases. Among them, KEGG (Kyoto Encyclopedia of Genes and Genomes) constructed by Kanehisa and colleagues should be noted.\(^{37}\) KEGG was the first database that integrated genomic, chemical, and systematic functional information, which is now globally well recognized and utilized (https://www.genome.jp/kegg/). To strengthen data-driven genome sciences, Monbusho introduced a supercomputer systems to the Human Genome Research Center at the University of Tokyo (http://www.ims.u-tokyo.ac.jp/imsut/en/lab/hgclink/) and the DNA Data Base Japan (DDBJ, https://www.ddbj.nig.ac.jp/), the international public database at the National Institute of Genetics. These database activities are now integrated with the NBDC (National Bioscience Database Center, https://biosciencedbc.jp/en/).

In 2000, Monbusho and STA were reorganized into the Ministry of Education, Culture, Sports, Science, and Technology of Japan (MEXT), and MEXT started a new 5-year program for genome research promotion (PI: Yuji Kohara) under a Grant-in-Aid for Scientific Research on Priority Areas.

In addition to the promotion of the basic academic research, the government shifted the focus towards the application of basic research in biotechnologies.\(^ {38}\) MITI supported pharmaceutical companies to launch a joint company, Helix Research Institute, which aimed to produce and patent large numbers of full-length cDNA. MAFF promoted the Rice Genome Project to step up to the whole genome sequencing stage as an international project.\(^ {39}\) Furthermore, as part of the Japanese millennium project (http://www.kantei.go.jp/jp/mille/genomu/index.html), major investments were made in healthcare and biotechnology. For example, MEXT launched the SNP Research Center (Director: Yusuke Nakamura) at RIKEN towards the establishment of personalized medicine (http://www.riken.jp/en/about/reports/evaluation/src/). MITI launched several new programs, including the genome analysis of industrial micro-organisms, functional analysis of full-length human cDNAs, functional proteomics, the standard SNP project, and technology development for bioinformatics. MAFF also launched a new program for the genome-based breeding of crops.

**International cooperation and HUGO Pacific.** The HGP widely stimulated the genome research across various countries and regions. In such circumstances, HUGO took various actions to stimulate international cooperation, such as the organization of the annual HGM conference. Japanese researchers actively participated in HUGO activities to promote genome sciences in the Asia region. HUGO Pacific meetings have been held at various cities in Asia including Busan (Korea), Bali (Indonesia), and Pataya (Thailand) with support from the HUGO Pacific office in Tokyo. Among these, notable HUGO Pacific activities included the HUGO Pan-Asia SNP Initiative, which aimed to characterize Asian populations based on their genomic variations. This initiative was launched at HGM 2002 in Shanghai, under the author’s proposal as the President of HUGO and implemented under the leadership of Edison Liu, with participation by nearly one hundred genome scientists from Asian countries and regions. Progress was presented in HGM 2005 in Kyoto, and the final report was published in Science\(^ {40}\) in 2009.

---

Fig. 10. Extremely uneven distribution of G+C contents and genes in human chromosome 21.
Final stage of the HGP (2000–2003): Completion of the human genome sequence

Outline. After the end of the draft sequencing phase, the international consortium continued efforts to complete the human genome sequence at the highest quality in a chromosome-by-chromosome manner. In addition to chromosomes 21 and 22, Japanese researchers played leading roles in the completion of chromosome 11 and also made a significant contribution to the completion of chromosomes 18 and 8. Finally, on April 14, 2003, the 50th Anniversary of the discovery of the double helix structure of DNA by Watson and Crick, the international consortium declared the completion of the human genome sequence. This historic achievement was celebrated globally. In Japan, representatives of the Japanese teams visited Prime Minister Jun-ichiro Koizumi and offered their gratitude for the long-term support of the Japanese Government during the project. Furthermore, a Joint Proclamation was made by the Heads of Governments of six countries to celebrate this historic achievement.

Completion of the human genome sequence. After the end of the draft sequencing phase, the international team turned back to the original strategy to complete the human genome sequence at the highest quality in a chromosome-by-chromosome manner. The consortium set the final goal of its completion in April 2003, the 50th Anniversary of the discovery of the double helix structure of DNA by Watson and Crick.

The “draft sequence” required mainly technological speed and power of sequencing. However, the complete sequencing was quite different. It required substantial gap filling efforts with professional knowledge and eyes. In the consortium, progress was checked and reviewed almost every 6 months at strategy meetings. The last strategy meeting was held at RIKEN GSC in August 2002 (Fig. 11).

On April 14, 2003, the international team finally declared the completion of the human genome sequence (https://www.genome.gov/11006929/2003-release-international-consortium-completes-hgp/ and https://www.ddbj.nig.ac.jp/news/ja/wn030423. html).

The “completed data” covered 2.85 Gb, with the more than 99.99% accuracy and an estimated 341 small gaps (Table 2). About 1% of the genome was not covered because the regions were unable to be cloned by the technology of the time. The data were disclosed to the public through public databases, and the summary paper on the finished whole genome sequence was published in 2004. Detailed analysis of each chromosome was published separately one-by-one.

The contribution of the Japanese researchers to the whole genome sequencing project was remarkable. The RIKEN and Keio teams completed chromosome 21 in May 2000 in collaboration with three German teams. The Keio team contributed to the completion of chromosome 22 in collaboration with U.K. and U.S. teams in December 1999. In addition to these previously mentioned achievements, the RIKEN team completed chromosome 11 as the principal investigator and also chromosome 18 as the sub-principal investigator in collaboration with the Broad Institute of MIT. The Keio team also contributed to the sequencing of chromosome 8 in collaboration with the Broad Institute of MIT. In addition, this author would like to note that the international team led by RIKEN GSC successfully completed the sequencing of chimpanzee chromosome 22, the counterpart of the
human chromosome 21, which enabled unique insights into the human evolution.\textsuperscript{43}

The Japanese contribution is summarized in Fig. 12. The regions sequenced by the Japanese teams are shown with the national flag of Japan. The national flag at the bottom of each chromosome shows the nationality of the leading center for each chromosome. For chromosomes 11 and 21, RIKEN GSC played the leading roles.

Celebrations and joint proclamation. The completion of the human genome was celebrated globally. In Japan, the four leaders of the Japanese teams, Yoshiyuki Sakaki (the author), Nobuyoshi Shimizu, Hidetoshi Inoko, and Hideaki Sugawara (DDBJ), visited the Prime Minister Jun-ichiro Koizumi with Atsuko Tohyama, the MEXT, to report the completion of the human genome and to offer special gratitude for the long-term support of the Japanese Government. At the ceremony, the leaders gave a gift of a set of CD-ROMs containing the entire sequence data to the Prime Minister (Fig. 13). In addition, to celebrate the completion of this historical achievement, the Heads of Government of six countries that participated in the HGP released the “Joint Proclamation” (http://japan.kantei.go.jp/koizumispeech/2003/04/14sengen_e.html).

The Joint Proclamation stated:

“We, the Heads of Government of the United States of America, the United Kingdom, Japan, France, Germany and China, are proud to announce that scientists from our six countries have completed the essential sequence of three billion base pairs of DNA of the human genome. .........................

This genetic sequence provides us with the fundamental platform for understanding ourselves, from which revolutionary progress will be made in biomedical sciences and in the health and welfare of humankind. .........................

We congratulate all the people who participated in this project on their creativity and dedication. Their outstanding work will be noted in the history of science and technology, and as well in the history of humankind, as a landmark achievement. ....

..........”

A side story. The idea of the Joint Proclamation initially arose at an informal meeting of Toh-ichi Sakata (Director-General, Research Promotion Bureau, MEXT), Kazuo Todani (Director, Life Sciences Division, MEXT) and the author for planning the cerebration of the completion of the human genome in Japan. The idea was proposed and favorably approved by the international consortium at Cold Spring Harbor Laboratory in May, 2002. The first draft was prepared by the author, and later revised by Michael Morgan and Francis Collins to obtain the approval of the Heads of Government of six countries.

Epilogue

The HGP is a historic achievement by global researchers, who successfully completed the human genome sequence, the genetic blueprint of mankind,
and provided a solid basis for understanding ourselves. The Japanese genome community is proud of its significant and visible contribution to this historic achievement. It should also be noted that the Japanese genome community, in concert with this international project, successfully organized and developed the domestic alliances necessary to promote wide range of ambitious genome sciences as described above.

The completion of the human genome was not the final goal, but it was a start of new era of life sciences. The completion of the human genome brought about revolutionary changes in genomic sciences, particularly in medicine at a global level. For example, we are now able to determine the whole genome sequence of ourselves in a very short period of time using so-called next-next generation DNA sequencers to choose the best medical care based on our own genetic background. We are now in the age of personalized medicine. In Japan, genome-based medicine has been actively promoted through various programs, which are now well coordinated by the Japan Agency for Medical Research and Development (AMED), the headquarter of the national medical research programs.

Furthermore, the current advances in sequencing technologies, or the next generation of DNA sequencers, may allow us to sequence and analyze the genomes of all biological organisms on the planet. Such a huge amount of the genome information of all organisms is a treasury of the wisdom that living organisms acquired through evolution of life over 3.6 billion years. We are now advancing towards a new stage to utilize the wisdom of organisms for the welfare of the mankind, coupled with new emerging technologies such as genome editing and synthetic biology.

It is entirely fair to say that the HGP triggered a revolution in life sciences and biotechnology in the 21st century.

**Human Genome Archive.** The HGP is a historic achievement by mankind and should be inherited as a historical legacy of science and technology, which will be subjected to study by historians and researchers in the future. For example, it may be interesting to investigate more why Japan failed to establish a unified national program in spite of Wada’s ambitious proposal. In this context, it is important to preserve various documents such as meeting reports/memorandums, personal and official letters/e-mails, photographs, interviews, unpublished scientific reports/drafts, and other important documents related to the project. However, most of those materials are stored in personal collections, institu-
tional stocks, and so on, and are publicly unavailable. The international community participated in the HGP launched the Human Genome Archive activity to collect and edit these documents and materials in publicly available forms. In Japan, Human Genome Archive activity started in 2016 with support from the Genome Technology Committee of the JSPS and is now in progress. The archives will be available in a few years, which will help us more deeply understand Japanese history of involvement in the human genome programs described in this review.

Acknowledgement

The author thanks all the members who participated in the Japanese human genome programs for their long-term collaboration and cooperation. The author would like to deliver special thanks to Akiyoshi Wada and Ken-ichi Matsubara for their pioneering, leading roles and supports of the Japanese human genome programs, and Masahira Sumio Sugano, Masamitsu Saito, and Mayumi Kobayashi for their devoted efforts to the draft and the complete sequencing of the human genome. Toshihisa Takagi kindly provided the records of Japanese efforts in genome informatics. Sumio Sugano, Masamitsu Saito, and Mayumi Yokota provided important documents on the Japanese human genome programs under the Japanese Human Genome Archive Project.

The author also thanks colleagues of the international Human Genome Project for their cooperation and collaboration, particularly Francis Collins, John Sulston, and Michel Morgan for their kind understanding and support of Japanese efforts in the Human Genome Project. In addition, the author would like to deliver special thanks to the German colleagues, Marie-Loo Yaspo, Hans Lehrach, Andre Rosenthal, and Helmut Blocker for their fruitful collaborations for the completion of the chromosome 21. Todd Taylor and Desire Smith kindly checked the manuscript.

References

1) Dulbecco, R. (1986) Turning point in cancer research: Sequencing the human genome. Science 231, 1055–1056.
2) Palca, J. (1986) Human genome: Department of Energy on the map. Nature 321, 371.
3) Lewin, R. (1986) Proposal to sequence the human genome stirs debate. Science 232, 1598–1600.
4) National Research Council (1988) Mapping and Sequencing the Human Genome. The National Academy Press, Washington D.C.
5) Smaglik, P. (2000) Critics challenge Celera’s claim over human genome sequence. Nature 404, 691–692.
6) Wada, A. (1986) Automated high-speed DNA sequencing. Nature 325, 771–772.
7) The chromosome 21 mapping and sequencing consortium (2000) The DNA sequence of human chromosome 21. Nature 405, 311–319.
8) Taylor, T., Noguchi, H., Totoki, Y., Toyota, A., Kuroki, Y., Dewar, K. et al. (2006) Human chromosome 11 DNA sequence and analysis including novel gene identification. Nature 440, 497–500.
9) Kambara, H. and Takahashi, H. (1993) Multiplesheathflow capillary array DNA analyser. Nature 361, 565–567.
10) Swinbanks, D. (1987) Japanese eye human genome. Nature 328, 195.
11) Roberts, L. (1989) Watson versus Japan. Science 246, 576–578.
12) McKusick, V. (1989) The human genome organization. History, purpose and membership. Genomics 5, 385–387.
13) Ichikawa, H., Hosada, F., Arai, Y., Shimizu, K., Ohhira, M. and Ohki, M. (1993) A NotI restriction map of the entire long arm of human chromosome 21. Nat. Genet. 4, 361–366.
14) Chumakov, I.M., Rigault, P., Guillou, S., Ougen, P., Billault, A., Guasoni, G. et al. (1992) Continuum of overlapping clones spanning the entire human chromosome 21. Nature 359, 380–387.
15) Kawasaki, K., Minoshima, S., Kudoh, J. and Shimizu, N. (1992) Shotgun polymerase chain reaction: Construction of clone libraries specific to a NotI fragment of flow-sorted human chromosome 22. Genomics 13, 109–114.
16) Nakamura, Y., Okui, K., Takahashi, E. and 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. Cyto genet. Cell Genet. 65, 111–118.
17) Marnyama, K. and Sugano, S. (1994) Oligo-capping: A simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. Gene 138, 171–174.
18) Hishiki, T., Kawamoto, S., Morishita, S. and Okubo, K. (2000) BodyMap: A human and mouse gene expression database. Nucleic Acids Res. 28, 136–138.
19) Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V. et al. (1997) The complete genome sequence of the gram-positive bacterium Bacillus subtilis. Nature 390, 249–256.
20) Guyer, M. (1998) Statement of the rapid release of genomic DNA sequence. Genome Res. 8, 413.
21) Wadman, M. (1998) Company aims to beat NIH human genome efforts. Nature 393, 101–102.
22) Sasaki, T. (1998) The rice genome project in Japan. Proc. Natl. Acad. Sci. U.S.A. 95, 2027–2028.
23) Wadman, M. (1998) ‘Rough draft’ of human genome
International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. Nature **409**, 860–921.

Venter, C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J. and Sutton, G.G. (2001) The sequence of the human genome. Science **291**, 1304–1351.

Kawai, J., Shinagawa, A., Shibata, A., Yoshino, M., Itoh, M., Ishii, Y. et al. (2001) Functional annotation of a full-length mouse cDNA collection. Nature **409**, 685–690.

Del Hoon, M., Shin, J.W. and Carninci, P. (2015) Paradigm shifts in genomics through the FANTOM project. Mamm. Genome **26**, 391–402.

Handa, N., Nureki, O., Kurimoto, K., Kim, I., Sakamoto, H., Shimura, Y. et al. (1999) Structural basis for recognition of the transcript RNA precursor by the Sex-lethal protein. Nature **398**, 579–585.

Dunham, L., Shimizu, N., Roe, B.A., Chissoe, S., Hunt, A.R., Collins, J.E. et al. (1999) The DNA sequence of human chromosome 22. Nature **401**, 489–495.

International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. Nature **431**, 931–945.

Kobayashi, K., Nakahori, Y., Miyake, M., Matsumura, K., Kondo-Iida, E., Nomura, Y. et al. (1998) An ancient retroprenpositional insertion causes Fukuyama-type congenital muscular dystrophy. Nature **394**, 323–327.

Arai, Y., Hosada, F., Nakayama, K. and Ohki, M. (1996) An yeast artificial chromosome contig and Not1 restriction map that spans the tumor suppressor gene(s) locus, 11q22.2-q23.3. Genomics **35**, 196–206.

Ishihara, K., Hatano, N., Furuumi, H., Kato, R., Iwaki, T., Miura, K. et al. (2000) Comparative genomic sequencing identified novel tissue-specific enhancers and sequence elements for methylation-sensitive factors implicated in Igf2/H19 imprinting. Genome Res. **10**, 664–671.

Suzuki, Y., Yoshimoto, K., Maruyama, K., Suyama, A. and Sugano, S. (1997) Construction and characterization of a full length-enriched and a 5′-end-rich cDNA library. Gene **200**, 149–156.

Tabara, H., Hill, R.J., Mello, C., Priess, J. and Kohara, Y. (1999) Pos-1 encodes a cytoplasmic zinc finger protein essential for germine specification in *C. elegans*. Development **126**, 1–11.

Ito, T., Tashiro, K., Muta, S., Ozawa, R., Chiba, T., Nishizawa, M. et al. (2000) Towards protein-protein interaction map of the budding yeast. A comprehensive system to examine two-hybrid interaction in all possible combination between the yeast proteins. Proc. Natl. Acad. Sci. U.S.A. **97**, 1143–1147.

Kanehisa, M. and Goto, S. (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. **28**, 27–30.

Saegusa, A. (1999) Japan acts to speed technology transfer from universities. Nature **401**, 3.

International Rice Genome Sequencing Consortium (2005) The map-based sequence of the rice genome. Nature **36**, 793–800.

The HUGO Pan-Asian SNP Consortium (2009) Mapping human genetic diversity in Asia. Science **326**, 1541–1545.

Nusbaum, C., Zody, M.C., Borowsky, M.L., Kamal, M., Kodira, C.D., Taylor, D.T. et al. (2005) DNA sequence and analysis of human chromosome 18. Nature **437**, 551–555.

Nusbaum, C., Mikkelsen, T.S., Zody, M.C., Asamawa, S., Taudien, S., Garber, M.M. et al. (2006) DNA sequence and analysis of human chromosome 8. Nature **439**, 331–335.

The International Chimpanzee Chromosome 22 Consortium (2004) DNA sequence and comparative analysis of chimpanzee chromosome 22. Nature **29**, 382–388.

(Received Feb. 5, 2019; accepted June 4, 2019)
Profile

Yoshiyuki Sakaki was born in 1942 in Nagoya. He graduated from the Faculty of Science, the University of Tokyo and received his Ph.D. degree in 1971. He worked at University of California Berkeley as a post-doctoral fellow (1971–73) and then at the Mitsubishi-Kasei Institute of Life Sciences as a researcher (1973–77). He moved to Kyushu University as a lecturer and an associate professor at the Medical School (1977–85), and then served as the professor of the Research Laboratory of Genetic Information (1985–91). In 1991 he moved to the Institute of Medical Sciences, the University of Tokyo and served as the professor of the Human Genome Research Center (1991–2004). In 1998, he received a joint appointment at the RIKEN Genomic Sciences Center, and then served as the Director of the RIKEN Genomic Sciences Center (2004–2008). In 2008, he was nominated as the President of Toyohashi University of Technology (2008–2014). He is an Emeritus Professor of the University of Tokyo, Kyushu University, and Toyohashi University of Technology. He was also an Emeritus Researcher of RIKEN. He is a world-leader in human genomics, and represented Japan in the International Human Genome Project. He also served as the President of HUGO (Human Genome Organization) from 2002 to 2005.

He received “Chevalier” from France Government (2001), the Award of the Japanese Society of Human Genetics (2001), the Chunichi Culture Award from the Chunichi Culture Foundation (2003), the Medal with Purple Ribbon (2003), a Person of Cultural Merit (2013), and the Order of the Sacred Treasure, Gold and Silver Star (2016) from the Japanese Government.