The pig genome consists of 18 pairs of autosomes and one pair of sex chromosomes. The size of the entire genomic sequence is about 2.7 Gb, comparable with or slightly smaller than that of human. Genome information is essential for physiological analyses in pigs as well as other organisms. The international community for pig researchers has made efforts to develop components needed for sequencing of the entire pig genome, such as construction of a physical map of the pig genome by using radiation hybrid panels. Based on such analyses, sequencing of the entire pig genome was planned by the International Swine Genome Sequencing Consortium (SGSC), which was established in 2003. SGSC consists of research institutes and universities in 12 countries and regions, including the National Institute of Agrobiological Sciences (NIAS) and the Institute of Society for Techno-innovation of Agriculture, Forestry and Fisheries (STAFF) from Japan.

SGSC used a female Duroc pig, named T. J. Tabasco (Fig. 1), for the genome sequencing. Sequencing was conducted by the “hybrid approach” composed of hierarchical and whole-genome shotgun sequencing. For hierarchical shotgun sequencing, a bacterial artificial chromosome (BAC) library was constructed with genomic DNA of T. J. Tabasco. The library was used for construction of a minimum-tiling path of the BAC clones covering the entire pig genome. The clones in the minimum-tiling path were subjected to shotgun sequencing at a 4 to 8′ depth. The shotgun sequencing was mainly performed by the Wellcome Trust Sanger Institute, and the Japanese research team, Animal Genome Research Program (AGP) which consists of NIAS and STAFF, took charge of 254 BAC clones located on chromosome 6 and 7 (Fig. 2). The 254 BAC clones corresponded to 1.6% of the entire genome. AGP conducted the shotgun sequencing of each BAC clone at a 7.92 × depth on average. Finally, SGSC conducted sequencing of 17,000 BAC clones and announced completion of draft sequencing of the pig genome in November 2009. In the announcement SGSC declared completion in sequencing of 98% of the physical map of the pig genome (Fig. 3). The assembled pig genome sequence can be available in the Ensembl Database of the European Molecular Biology Laboratory (http://www.ensembl.org/Sus_scrofa/Info/Index). On the other hand, whole genome shotgun sequencing of...
The genome sequence in the Ensembl database is now being subjected to genome annotation by several groups that are interested in particular fields such as immunity or lipid metabolism. In the genome annotation, information of gene transcripts, particularly full-length cDNA sequences, is invaluable for correct mapping of exons of genes. In addition to the genome sequencing, we have also conducted full-length-enriched cDNA library construction for various tissues and cell populations of pigs, analyzed expressed sequence tags (ESTs), and sequenced the entire inserts of the cDNA clones. We have accumulated more than 280,000 ESTs mainly derived from full-length-enriched cDNA libraries, and completed sequencing of about 25,000 cDNA clones extracted from the clones subjected to the EST analysis. Information of the cDNA and ESTs may be accessed through the Pig Expression Data Explorer (http://pede.dna.affrc.go.jp/) with results of BLAST similarity search, as well as through public nucleotide databases (DDBJ/EMBL/GenBank). This is a significant contribution to the annotation process of the pig genome.

The draft sequence of the pig genome will generate much benefit in research of livestock and biomedical sciences. It enables us to localize many markers on the whole genome and to conduct analyses for detection of relationships between useful traits and genomic regions, which will contribute to pig breeding based on molecular information. Furthermore, information about the pig genome accelerates the use of pigs as biomedical model animals, because the body size of pig is comparable to human. Moreover, physiological systems such as cardiovascular system in pigs show high similarity to the human system. Pigs are regarded as large experimental animals for analyses that cannot be conducted with mouse or other small animals. To make the genome sequence more valuable, SGSC will continue the effort to improve the accuracy of the pig genome sequence.

Fig. 2 Contribution of the Animal Genome Research Program (AGP) in the International Swine Genome Sequencing Consortium (SGSC)
AGP conducted sequencing of the genomic regions, which correspond to 42.3 Mb, on chromosome 6 and 7.

Fig. 3 Sequencing status of the pig genome at the announcement of completion of draft genome sequencing by SGSC
A bar indicates the ratio of the region that is already sequenced in each chromosome.