

FUGU FACTS

Q&A

Fugu Genome Sequencing Consortium FACT SHEET

What is the Fugu Genome Project?

The Fugu Genome Project is an international program aimed at determining the complete DNA sequence of the genome of the Japanese pufferfish, *Fugu rubripes*. Despite the obvious differences between fish and humans, it is expected that comparisons of the human genome with that of Fugu will shed light on the common genetic systems shared by these two animals, and help us understand the information encoded in the human genome.

What is Fugu?

Fugu is a teleost fish belonging to the order Tetraodontiformes (four toothed puffers) and a member of the gnathostomes (jawed vertebrates). There are over 100 species of pufferfish with diverse salt water and fresh water habitats. Fugu species are farmed in Japan and the flesh is consumed as a delicacy—although certain organs of the fish must be avoided because they contain a potent neurotoxin.

What was actually accomplished?

Today's announcement marks the completion of the draft sequence of the Fugu genome. Over the past year, nearly four million pieces of Fugu genome sequence were determined by the Fugu Genome Consortium. These genomic fragments, averaging around 600 DNA bases in length, overlap each other,

which allows them to be reassembled computationally to reconstruct long stretches of the Fugu genome, spanning tens of thousands of DNA bases in length. Fugu is the first animal genome to be sequenced and assembled in the public sector using this “whole genome shotgun” sequencing approach.

Why is this sequencing achievement important?

The Fugu genome is the first vertebrate genome to be draft sequenced after human. Its compact form and similarity to the human genome will make it an important tool for getting at the information encoded in the human sequence. We now have in hand the basic gene-level description of two vertebrates. Comparing and contrasting them will allow us to discover new human genes and, importantly, elements which control or regulate the activity of genes. Using genomes in this way has been compared to the way in which ancient languages were decoded using the Rosetta Stone—one common text translated side-by-side into different languages.

But I thought all of the human genes were identified by now?

No. The computational methods for predicting genes are imperfect, and we know that a significant number of human genes remain to be discovered in the human genome sequence. Comparison



between genomes is a powerful way to find such genes. More significantly, there are no good computational methods for reliably finding the elements which surround genes and control their expression, that is, determine when and where a gene will be turned on or off, and how much protein should be made. For example, genes that are used in the kidney may not be used in the brain. We are still learning to detect these genomic signals. This is an important missing piece of the human genome puzzle.

So what makes Fugu such a good choice for comparison with human?

First, Fugu is a vertebrate—despite their apparent differences, fish have nearly all of the same organ systems and physiology as humans, in contrast to the more distantly related invertebrate animals already sequenced, like flies and worms. Just as important, however, is that the Fugu genome is unusually small for a vertebrate. The pioneering work on Fugu, published in the journal *Nature* in 1993, showed that despite a similar gene content, the entire Fugu genome is only 1/8th the size of the human. Even among fish, Fugu is special: most fish genomes are several times longer than Fugu's.

How big is the Fugu genome?

Pufferfish have the smallest known vertebrate genomes, around 350-400 million bases long, or 350-400 megabases. (These bases are denoted by letters — A, C, T, or G — which represent the chemical units that are strung together to make genes and chromosomes.) Fugu has 22 pairs of chromosomes, though these have no direct correspondence with the 23 pairs of human chromosomes. For comparison, the human genome is about three billion bases long. Despite this size difference, however, both Fugu and human are expected to have a similar repertoire of genes

How can fish and humans have the same set of genes? Aren't they very different?

It depends on what you mean by “the same.” It has been amply demonstrated that many human genes including, for example, “disease genes” like dystrophin, whose mutation causes muscular dystrophy, have close relatives in Fugu. These related genomic features can be detected computationally by comparing the two genomes and looking for simi-

lar sequences. The Fugu and human genomes are similar by virtue of their shared vertebrate heritage. Of course, humans (and Fugu) will have their own unique genes that are special for human-ness and fish-ness that the genome comparisons will also bring to light. But even these fish- and human-specific genes are likely to share a common genetic heritage.

If they have similar gene content, why is the Fugu genome so much smaller than the human genome?

Genomes contain more than just genes. In fact, only a few percent of the human genome actually represents “coding sequence,” the functional parts of genes. The rest of the human sequence is dominated by highly repetitive non-gene DNA—for example, regions that read “ACACACAC ...” for hundreds of bases, or have longer sequences that are scattered throughout the human genome hundreds of thousands of times. While these repeats make up 40% of the human sequence, the Fugu genome has much less repetitive content—for mysterious reasons that should be illuminated by the genomic sequence now in hand. But it's not only the relative lack of repeats that makes Fugu special—Fugu genes themselves are more compact than human genes, and packed more tightly on the genome. This is the main reason Fugu was chosen for sequencing—as a cost-effective, more-genes-for-the-buck shortcut to a vertebrate gene set, the gene-rich Fugu can't be beat.

Why sequence Fugu rather than another mammal like mouse or rat?

Genome sequencing shouldn't be thought of as an either-or proposition. It is essential that a broad range of animal genomes be sequenced, to shed light on the underlying similarities and essential differences between species. Fugu is just the beginning. The ongoing mouse and rat genome projects are critical for biomedicine, and will be particularly powerful tools because these animals are mammals (more closely related to human) and can be bred and studied more easily than Fugu. The Fugu genome provides a more distant evolutionary comparison (400 million years, versus 100 million years for mouse and rat) that permits a more accurate triangulation of genome function than mouse or rat alone. Genomic features that are common to Fugu, rodents, and human will focus our attention on the essential core genes that define being a vertebrate.



How are Fugu and humans related by evolution?

About 450 to 500 million years ago, the first vertebrates (animals with segmented backbones made of cartilage or bone) appeared in the early oceans. Their descendents split into two main groups: the ray-finned fishes—which include Fugu and most fish familiar to us from the dinner table—and the lobe-finned fishes, a more obscure group with fleshy paddle-like appendages in place of the paper-thin fins of the ray-finned fish. Over millions of years, these lobe-fins evolved into the limbs possessed by all four-limbed creatures (the tetrapods, including reptiles, amphibians, birds, and mammals). So Fugu are our very distant cousins, sharing a common ancestor with us nearly half a billion years ago. Remarkably, this common ancestry is still recorded in our genes.

Where can I learn more about the Fugu and its genome?

Below are a few of the key scientific publications which describe landmarks in the work on Fugu:

Brenner et al *Nature* 366:265-8(1993)

Aparicio et al *PNAS* 92:1684-8 (1995)

Trower et al *PNAS* 93:1366-9 (1996)

Elgar et al *Genome Res.* 9:960-71 (1999)

Also, each of the consortium members has a web site which describes facts and features of the Fugu. These may be found at:

www.fugu-sg.org

www.jgi-psf/fugu

fugu.hgmp.mrc.ac.uk



KEY PUBLICATIONS

Who are the Consortium Members?

The U.S. Department of Energy Joint Genome Institute

The Department of Energy's Joint Genome Institute, established in 1997, is one of the largest publicly funded human genome sequencing institutes in the world. The JGI was founded by the three University of California-managed national laboratories: Lawrence Berkeley National Laboratory and Lawrence Livermore National Laboratory in California and the Los Alamos National Laboratory in New Mexico. The JGI is led by Dr. Trevor Hawkins and has its main headquarters and Production Genomics Facility in Walnut Creek, California. The JGI employs about 240 full-time people and has programs in genomic sequencing, computation, functional genomics, genomic diversity and new technology development. In its role as a partner in the Human Genome Project, JGI was responsible for sequencing human chromosomes 5, 16, and 19, which make up 11% of the human genome. Funding for the JGI is predominantly from the DOE Office of Science with additional funding from NIH, NSF, USDA and NASA. Information about the JGI can be found on the genome portal at www.jgi.doe.gov.

Institute of Molecular and Cell Biology (IMCB)

IMCB was established in 1987 at the National University of Singapore (NUS). Its mission is to develop and foster a vibrant research culture for biological and biomedical sciences which will support the development of biotechnology for the human health care industry in Singapore. IMCB is one of the 5 biomedical sciences research institutes funded by the Biomedical Research Council. From a modest start with 38 scientists it today has a research staff of 250. These comprise an internationally diverse group representing Asia, North America, and Europe. Their research focuses on cell regulation and signal transduction, development, functional genomics, immunology, virology, infectious diseases and drug discovery. The IMCB has established collaborations with industry, universities, and research institutions worldwide. <http://imcb.nus.edu.sg/>

The Singapore Biomedical Research Council (BMRC)

is part of the National Science and Technology Board (NSTB), Singapore's national agency for science, technology and research. BMRC oversees and provides support to public sector biomedical research and development activities in Singapore. The Council also aims to strengthen collaborative public research in the biomedical sciences in Singapore. The Council's objectives are to support, sustain and stimulate excellent research for maintaining and improving human health, train people in high quality research skills to meet Singapore's needs of health, quality of life and global economic competitiveness, and promote societal awareness of biomedical research. www.biomed-singapore.com

The Medical Research Council (MRC)

is a national organization funded by the United Kingdom taxpayers. Its business is medical research aimed at improving human health. The research it supports and the scientists it trains meet the needs of the health services, the pharmaceutical and other health-related industries and the academic world. MRC has funded work which has led to some of the most significant discoveries and achievements in medicine in the UK. About half of the MRC's expenditure of £345 million is invested in over 50 of its Institutes and Units, where it employs its own research staff. The remaining half goes in the form of grant support and training awards to individuals and teams in universities and medical schools.

www.mrc.ac.uk

The Cambridge University Department of Oncology

United Kingdom, hosts the Cancer Genomics Program, a multidisciplinary group of investigators in areas ranging from comparative genomics, genetics, cell biology and translational clinical research, to address basic and translational research questions applicable to the understanding of cancer. The University of Cambridge is one of the oldest universities in the world, and one of the largest in the United Kingdom. It has a world-wide reputation for outstanding academic achievement and the high quality of research undertaken in a wide range of science and arts subjects. The University's achievements in the sciences can be measured by the sixty or more Nobel Prizes awarded to its members over the years. www.hutchison-mrc.cam.ac.uk

The Institute for Systems Biology

located in Seattle, Washington, was co-founded in 2000 by Lee Hood, Ruedi Aebersold, and Alan Aderem as a private, non-profit research institute devoted to systems biology, an emerging field made possible by rapid advancements in genomic, proteomic and computer technologies. Unlike traditional scientific approaches that examine single genes or proteins, systems biology focuses on studying the complex interaction of vast numbers of biological elements. The Institute is also pioneering new science education and increasing public awareness of biotechnology issues.

www.systemsbiology.org/

The Celera Genomics Group

headquartered in Rockville, MD, is engaged principally in integrating advanced technologies to create therapeutic discovery and development capabilities for internal use and for its customers and collaborators. Celera's businesses are its online information business and its therapeutics discovery business. The online information business is a leading provider of information based on the human genome and other biological and medical information. Through the therapeutic discovery business, Celera intends to leverage its genomic and proteomic capabilities to identify drug targets and diagnostic marker candidates, and to discover novel therapeutic candidates. www.celera.com/

Myriad Genetics, Inc.

is a leading biopharmaceutical company focused on the development of novel healthcare products. Based in Salt Lake City, Utah, the company has established two wholly owned subsidiaries. Myriad Pharmaceuticals, Inc. develops and intends to market therapeutic products, and Myriad Genetic Laboratories, Inc. develops and markets proprietary predictive medicine and personalized medicine products. The company has established strategic alliances with Bayer, Eli Lilly, Hitachi, Novartis, Oracle, Pharmacia, Roche, Schering AG, Schering-Plough and Syngenta. www.myriad.com/

