

**Institute of Genomics and Integrative Biology
(Council of Scientific and Industrial Research)
Mall Road,
Delhi 110 007**

Telephone: 27667602, 27667439, 27666156/157
Website <http://igib.res.in>

PRESS RELEASE

WHOLE GENOME SEQUENCING OF THE ZEBRAFISH WILDTYPE STRAIN

Institute of Genomics and Integrative Biology (IGIB), which is a constituent laboratory of the Council of Scientific and Industrial Research (CSIR) has completed the whole genome sequencing of a Wildtype Strain of Zebrafish (*Danio rerio*). This work marks India's entry into the arena of whole genome sequencing of animals.

The Zebrafish genome is about half the size of the human genome, containing about 1700 million DNA base pairs. The research team at IGIB generated over 89 Gigabases of DNA sequences in two months time resulting in over 20X coverage of the Zebrafish genome. The Solexa / Illumina sequencing technology was employed for sequencing the Zebrafish Wildtype strain. This next-generation sequencing technology enables massively parallel sequencing of millions of genomic fragments ranging from 36 to 76 base pairs, which are then mapped back to the reference genome. This humongous exercise was made possible with the CSIR Supercomputing facility at IGIB.

Zebrafish, a four cm-long fish native to Indian rivers, has attracted considerable scientific interest worldwide primarily as a non-mammalian vertebrate model organism. Being a vertebrate, Zebrafish has blood, eye, heart, kidney and other biological processes that share many features of the human system. Because of the optical transparency of the embryos, small size at maturity and ease of culture, Zebrafish has become a popular organism to study embryonic development for biologist worldwide. Genetic screens in Zebrafish have led to the discovery of large number of mutations that disrupt embryonic development. Many of these mutations serve as models for several human syndromes. Over the years, studies in zebrafish have enabled discovery of novel genes, pathways and models of human diseases in addition to understanding mechanisms underlying genome evolution. Zebrafish models are also amenable to high-throughput chemical compound screening, thereby enabling discovery of novel drug compounds.

Why sequence the Wildtype Strain of Zebrafish?

Till date, the main genomics work has been focused on lab grown Zebrafish strains that functionally represent genetic clones; little or no true genetic diversity was captured in these initial genome studies. Therefore to address questions related to genomic variations, we have sequenced a Wildtype Strain of Zebrafish caught directly from the water bodies in India.

Understanding the maximum possible natural variation in genomes of organisms is important as it will provide clues to discovery of regions of the genome that are not/less likely to tolerate variations versus regions that are constantly subjected to genetic alterations in addition to understanding the overall architecture of the genome. This is best addressed by comparing the genome of a parent organism with that of its progeny or by comparing the genomes of multiple related siblings. Such a study will require analysis of genomes of multiple siblings (ranging from few tens to hundreds of siblings) from a single male and female parent and is near impossible in humans because of limiting family size (limited number of siblings). Therefore we have used Zebrafish, females of which produce about 200 embryos per week, for

addressing questions related to vertebrate genomic variations. Using the new advancements in sequencing technologies alongside cutting-edge bioinformatics capabilities, we want to explore genetic variation through comparing the genomes of this single Wildtype Strain Zebrafish parent and approximately 100 of its offsprings, which are currently bred and phenotyped at IGIB. We have named this endeavor as “**Project Kaurava**” after the legendary Kaurava brothers from the Mahabharata.

The pipeline established for the Zebrafish in terms of natural genetic variation represents a proof-of-concept for long-term goals of integrating human genetic variation into personalized medicine. In the near future (approx 5 years), sequencing of individual human genomes will become economical both in terms of cost and time. To make best use of this new technology, we will require meaningful interface tools to assess and visualize any potential genetic variation. The zebrafish represents a test case whose pipeline will be used downstream for genetic variation analysis with the goal of improving human health.

Timelines

The whole genome sequencing of a single Wildtype Strain of Zebrafish was completed in about two months. The whole genome sequencing of approximately 100 of the offspring is expected to complete by 2012.

Funding

The whole genome sequencing of the Wildtype Strain of Zebrafish was made possible through the CSIR sponsored Eleventh Five Year Plan project entitled “National Facility for Functional Genomics Research” (FAC0002).

Partners

The Zebrafish Wildtype Strain Genome Project was led by the CSIR, India. Mayo Clinic, Rochester, USA will join this CSIR led project for genome annotation.

Additional information on the Zebrafish Wildtype Strain Genome Project

Institute of Genomics and Integrative Biology,

Website: www.igib.res.in

Project Website: <http://genome.igib.res.in>

Contact information

Dr. Rajesh Gokhale,

Director, IGIB,

rsg@igib.res.in

Dr. Vinod Scaria,

09212371663

vinods@igib.res.in

Dr. Sridhar Sivasubbu

09871780662

s.sivasubbu@igib.res.in