BGI presents a high-quality gene catalog of human gut microbiome

July 7 2014

Researchers from BGI, working within the Metagenomics of the Human Intestinal Tract (MetaHIT) project, and in collaboration with other institutions around the world , have established the highest quality integrated gene set for the human gut microbiome to date- a close-to-complete catalogue of the microbes that reside inside us and massively outnumber our own cells. While the roughly 20,000 genes in the human genome have been available for over a decade, the gene catalog of the microbiome, our much larger "other genome", has to date been much more poorly understood and characterized.

The data released from this study should facilitate further research on the interactions between human and microbial genomes, and brings us closer to an understanding of how to maintain the microbial balance that keeps us healthy. The latest study was published online today in the journal *Nature Biotechnology*.

Each of our guts is colonized by more than 3 pounds of microorganisms that can break down toxins, manufacture vitamins and essential amino acids, and form a barrier against invaders. However, until now there has been a lack of comprehensive and uniformly processed database resources cataloging the human gut microbiota around the world, which has hindered our knowledge of the genetic and functional mechanism of human gut microbes.

In this study, researchers established a catalog of the human gut microbial genes by processing 249 newly sequenced samples and 1,018

published samples from MetaHIT, Human Microbiome Project (HMP) and a large diabetes study from China, as well as 511 sequenced genomes of gut-related bacteria and archaea. This expanded research is at least three times larger than the cohorts used for previous gene catalogs.

Based upon the catalog, researchers investigated the gut microbiota of healthy Chinese and Danish adults, and found the two cohorts greatly differed in nutrient metabolism as well as xenobiotic detoxification, which might be influenced by the differences in diet and environment. In addition, they observed enrichment in possible antibiotic resistance genes both at the population level (penicillin resistance in Danes and multidrug resistance in Chinese) and in the individual-specific genes, which highlighted the need for close monitoring of direct and indirect exposure to antibiotics.

Individual-specific genes contributed overwhelmingly to the increased total gene number in the integrated gene catalog and were overrepresented in genes responsible for the synthesis of cell wall components, DNA-related functions such as transposases, endonucleases and DNA methylases and encoding phage-related proteins. Such individual-specific genes likely reflect adaptation and might reflect the distinct combination of genetic, nutritional and medical factors in a host.

This nonredundant reference catalog of over 9.8 million genes is freely accessible through the website and the data have also been deposited in BGI's GigaScience Database, GigaDB and the SRA. It provides a much expanded and invaluable resource for global researchers to more deeply explore the geographical, genetic, temporal and physiological characteristics of gut microbes.

Junhua Li, Research Scientist from BGI, said, "Catalogs of reference genes in the human gut microbiome should facilitate quantitative

characterization of multi-omic data from the gut microbiome to understand its variation across populations in human health and disease."

Provided by BGI Shenzhen