Low frequency genetic variation may determine complex diseases

In an international project jointly funded by NSFC, researchers in BGI-Shenzhen, China, together with international collaborators from UC Berkeley, University of Copenhagen and some other European institutions conducted research on the re-sequencing and analysis of 200 human exomes, established the largest data set for human exomes published so far and revealed an excess of low frequency deleterious non-synonymous genetic mutations. Their findings were reported in a paper in *Nature Genetics* on October 4, 2010.

According to a report by EurekAlert, the team used Nimble Gen 2.1M exon capture array to targeted capture 18 654 coding genes of human genome and sequenced 200 individuals from Denmark. A large number of unknown SNPs were found and most of them appeared with low frequency. This study has developed the largest scale and the highest resolution genetic map of human exomes so far. Moreover, the massive data and in-depth analysis demonstrated that the excess of low frequency genetic mutations may cause the variations of protein amino acid sequences which would influence human health and under natural selection regulation.

Recently, a number of scientific researches indicated that association studies of complex diseases, following theoretical strategy, has identified numerous candidate genes in various experiments, but the results can only explain a limited fraction of the heritability of complex diseases. This missing heritability is also a major problem in complex disease genetics research. This study for the first time confirmed that genetic mutations associated with human health and disease susceptibility are generally at low frequency and are associated with multiple loci. Previous association studies using genotyping microarray can only detect common genetic variation and overlooks low frequency genetic mutations which may related with complex diseases, thus causing missing heritability.

This study not only shows the drawback of current disease research approach but also raise the revolutionary approach to apply genome sequencing technology instead of genotyping in association studies of disease. As a milestone, it will change the research methods of complex diseases and lead to the developments of human health and medical research.

This study is a part of the collaborative Sino-Danish Diabetes-associated Genes and Variations Study (LuCAMP). LuCAMP aims to detect novel rare and common genetic variations related with metabolic disorders through the exome sequencing of 1000 patients and 1000 controls.

Previously, a paper in *Science* (Science. 2010 July; 329(5987): 75-78) reported sequencing the exomes of 50 Tibetan individuals and found evidence for high altitude adaption of Tibetan populations. It shows that next generation sequencing is getting more applications and will have great potential in genomics research, drug discovery and personalized medical treatment.