The Taiwan Biobank Project: For the Health of Future Generation

Chia-Hao Ou, Chen-Yang Shen
Institute of Biomedical Sciences

Due to advancements in biomedical research and the completion of the sequencing of the human genome, geneticists can now provide biomedical researchers with a wider range of information on disease than before, aiding them to carry out more in-depth research on disease-causing mechanisms. In recent years, research on hereditary syndromes has shown there is high correlation between disease development and genetic variation.

From studies on the gene, environment, and disease, scientists have learned that there is an intricate relationship between genes and environmental factors in common diseases. In cancer formation, a series of studies have shown the importance of an interaction between metabolic genes, such as CYP1A1, NAT2, CYP2E1, and GSTM1, and smoking in disease development. Epidemiology research has shown that both genes and environmental factors play a role in disease-causing mechanisms. Many studies have shown that a single genetic factor is often inadequate to explain chronic disease and that we must also consider external environmental and behavioral factors. Even when we try to consider everything, we can obtain inconsistent results because of the effect of the interaction between genes and the environment. This could be due to several factors, as, in previous research (1) the samples were small, which results in a low statistical power producing unreliable results; (2) there has been incomplete or inappropriate measurement of exposure to dangerous factors and a lack of appropriate control groups; (3) there has not been formal statistical examination of various kinds of disease effects, especially of interactions; (4) there has been no complete picture of the combined effects of genes and the environment; and (5) a retrospective case-control design was used, which can only include the people who survive long enough to participate in the study and might be influenced by recall bias.

Small size studies are only suitable for rare genetic diseases caused by single gene mutation. In contrast, in common diseases, such as arthritis or stroke, multiple genes are thought to be affected, and life factors, such as diet or smoking, are also involved. If we could screen the gene profile of more people and combine phenotype and health information, we could obtain more information on common diseases than in the past. We know that, in common diseases, the effect of the specific gene and the environment can be very small, so large sample size studies will be helpful.

Because of the research limitations due to studies by individual laboratories or small sample size, we face the difficult situation of inference with study results, and researchers in various countries have begun to think about the need for large-scale resources to help biomedical research. If Taiwan can establish a large-scale resource for research, i.e. a biological database (Biobank), which includes gene and other medical information, it will provide a great help to future biomedical research. The Biobank will (1) help to determine the prevalence rate for a specific gene or variant in the general population; (2) help simplify the procedure for searching for a biological marker molecule; (3) be used to improve research into, and the development of, new medicines to cure disease; (4) help simplify tactics for preventing disease; and (5) offer the basis for decisions for hygiene policy. Most supporters of biological databases think that databases can improve the

quality of medical treatment. If we can define the genes involved in common diseases, we can quantify the danger associated with a particular gene, which will be helpful in the development of new treatments. Pharmaceutical companies will be able to use the database to design new drugs to match a specific gene profile and doctors will be able to offer individualized treatments and prevention programs.

The Taiwan Biobank plans to use prospective cohort studies based on ethnicity (population-based) that will help determine the effects of the environment or gene alone and of gene-gene interactions and gene- environmental factor interactions in common diseases. It will use nested case-control studies to track the incidence of different common diseases. This design structure is similar to that of the UK Biobank. Because the onset of disease, especially chronic disease, is usually the result of long-term environment-gene effects, we will only be able to establish causality by the use of long-term extendable cohort studies.

The feasibility research study for the Taiwan Biobank was launched in July 2006. Before July 2007, we expect to invite 1000 people to participate in this study in the Chia-yi area. We will build up the fundamental health database using questionnaires and simple and easy physical examinations using non-invasive methods. By carrying out this study, we will be able to assess the difficulties that the formal study may meet during launching and how to design an appropriate procedure. The assessment includes how interviewers inform the subjects of the purpose of the study, the procedure for obtaining consent to participate, the procedure for, and the duration of, the interview, document delivery, handling, and storage, and the participant's impression of the interview. It will be the reference for the execution and adjustment of the formal study in the future.

Nearly 700 people will have participated in the feasibility study by the end of March 2007. The early phase for Taiwan Biobank plans to invite 15,000 people between November 2007 and November 2009.

During the construction phase of the Taiwan Biobank, many issues will surface and need to be addressed, such as how to deal with myriads of biological samples and associated information, ethical and social considerations, sample storage/management, data processing, and analysis. Confidentiality and protection of personal information will be of particular concern. The most important thing is to communicate with the general public and to make them understand the importance of the Biobank and educate them in the development of science and genetic technology.

We expect that the Taiwan Biobank will not only help to identify disease-causing factors and mechanisms, but will also provide a huge resource for biomedical research, thus avoiding the difficulties encountered in the past in studying the relationship between genes and environmental factors. The ethnicity of Taiwan is unique and the Taiwan Biobank can be expected to develop into the supply centre for a Chinese ethnicity database. We could utilize the biomedical resources in Taiwan to carry out extensive cohort research into common diseases in Taiwan, which will not only help in identifying the disease-causing factors and mechanisms of common disease, but also help improve the tactics for treatment and prevention and so reduce medical costs and make it possible to achieve the goal of improving our nation's health.