

## Using genomic databases to study complex diseases

May 04, 2007 - For centuries humans have debated the relative roles of nature and nurture in shaping a person, but only recently has science begun to provide the tools needed to disentangle the two. What has become clear is that while some conditions — such as Huntington disease and Tay-Sachs — have solely genetic roots, most result from a complex interplay of multiple genes and an individual's environment. A person's environment includes everything from the physical environment in which they live and work, what they eat, how much they exercise, what medications they take, and whether they smoke, to what goes on inside their cells and how their genes interact with each other.

One example of this complex interplay is Type 2 diabetes. Diabetes runs in some families and appears to be more common in some ethnic groups. Researchers have identified genetic variants that increase a person's risk of developing the disease during their lifetime. However, not everyone with these variants will develop diabetes, and many people with diabetes do not carry the variants. There may be other, as yet unidentified, genetic variants that also are important. Diet and exercise also play important roles in diabetes risk. Most of the common disorders in the general population – heart disease, cancer, etc. – are similarly complex.

Genomic databases, or biobanks, increasingly are used to study complex conditions like diabetes. These large databases contain genetic, environmental, and health information collected from many people over long periods of time, and can be used to look for genetic and environmental contributors to disease. However, there are many significant scientific and practical challenges to building these kinds of research databases. For example, the issue of informed consent is complicated because the eventual uses for the data collected, and the types of research results generated, cannot be anticipated at the beginning of the study. Database administrators must also grapple with whether, and how, to share individual research results with participants.

Data security is another major challenge. Health data can be stripped of identifying information before researchers analyze it, but for a longitudinal study, in which participants are tracked over time, data must be identifiable in order to be updated. Therefore, plans to protect the privacy of longitudinal research subjects must be in place. Given the amount of information collected on each person, whether complete privacy can ever be assured is debatable.

The scientific challenges of studying complex diseases include understanding and confirming genetic associations identified through research, and accurately assessing environmental influences. Current methods of quantifying environmental influences are relatively primitive. For example, studies have shown that peoples' recollections of what – and how much – they've eaten often are inaccurate even if the meal was recent,

making food diaries an unreliable source for researchers seeking to link diet with health outcomes. Exposure to small quantities of chemicals over long periods is another influence that can be very difficult to measure. Better tools -- such as small, wearable sensors -- could greatly improve the accuracy of environmental measurements.

Studies that take both genes and environment into account hold tremendous promise for understanding the causes of many common health conditions. Many small studies are underway using existing cohort data. Large, prospective studies that enroll tens of thousands of individuals and follow them over time have started in several countries. Many citizens of Iceland have contributed samples and health information to the company deCODE, which is building an extensive database. Great Britain recently started a national project, UK Biobank, that will enroll half a million adults aged 40-69. A large-scale, prospective study of comparable scale and scope has not yet been funded in this the United States, but there is interest at the National Institutes of Health and other federal health agencies in conducting such a study.

Compiled by Shawna Williams Updated 03/28/08 by Joan Scott and Juli Murphy