News Focus

European genomics: think big or small?

With the first draft of the human genome completed, Michael Gross looks at some of the plans to study a variety of European populations to exploit the new information.

Now that the human genome is sequenced (more or less), what shall we do with it? If we want to benefit from it to gain an improved understanding of the genetic basis of disease and the different responses to drugs, knowing the three billion bases that all of us have in common is only the beginning. The second leg that future medical genetics will need to stand on is the knowledge of the several million bases that we don't have in common, from which the most widespread types of genetic variation originate. The most important group of variabilities is described as 'single nucleotide polymorphisms' or SNPs. They are defined as base exchanges occurring in at least 1% of a population, as opposed to a random mutation that is normally found with a frequency of much less than 1%.

SNPs essentially define the genetic differences between individuals, including — most importantly for medical research — their different susceptibilities to diseases and different responses to drugs. So far, the links between genes and diseases have only been established in relatively few, mostly rare and monocausal diseases, like cystic fibrosis. How can the more common multifactorial diseases (like coronary heart disease or type II diabetes) be linked to patterns in genetic variability? This is one of the major challenges awaiting science in the postgenomic era, and there are different approaches to it, which can be broadly divided into those which use small, inbred populations and those based on large populations.

Iceland cometh: Local people celebrating National Day in Reykjavik may be among the first to provide insight into the links between human genes and common diseases through the establishment of a national genetic database.

One large-scale study is being launched in the UK with funding from the Medical Research Council and the Wellcome Trust. The plan is for a major long-term study of around 500,000 middle-aged volunteers as a representative sample of the UK population. Alongside genetic analysis, basic measurements such as height, weight and blood pressure will be made and followed up with data on health and lifestyle. The hope is that within a few years meaningful research on a significant number of ill people should be possible.

Researchers who believe in the benefit of studying small populations are the luckier ones in practical terms, as they get to work in sunnier climates. Hotspots of genomics research have emerged in remote Italian villages such as Perdasdefogu and Talana, both situated in the Ogliastra region in Northwestern Sardinia, near the town of Alghero. Between them, these villages have less than 4000 inhabitants, who mostly descend from a small group of families who founded the villages around 500 years ago. In the case of Talana, it is precisely known that all 1400 villagers descend from eight males and eight females. Less than 5% of marriages have been to outsiders.

The man who pushed these villages to the forefront of genomics research is Mario Pirastu, a native Sardinian who returned to his island after research work in San Francisco and is now the director of the Istituto de Genetica Molecolare in Alghero, funded by Italy's national research council, CNR. He is a firm believer in the approach that uses small inbred populations which should be less than 200 generations old. Outside Sardinia, the best-known cases are the Amish, some communities in Newfoundland, along with some villages in Switzerland. Pirastu and his colleagues argue that the genetic homogeneity of such small inbred populations makes it easier to track down genetic factors contributing to multifactorial diseases, as differences irrelevant to the questions under study will be producing less noise than they would in a large and diverse population. The main criterion used in such studies is linkage disequilibrium which essentially relies on the fact that genes are more likely to be passed on together, the closer they are on the chromosome.

Last summer, Pirastu obtained large scale support from various sources to set up an international research centre at Perdasdefogu and make the most of the rare and valuable gene pool found in this village. For the native population, this comes mainly as a welcome boost to the local job market, so the local politicians are quite supportive of the research. In other remote parts of Italy, similar projects are on the way. Paolo Gasparini and his coworkers from the Medical Genetics Service at Rotondo have picked the Southern Italian village of Carlantino, where the rate of marriage within the community was 99.5% over the last century. And in the surroundings of Naples, Graziella Persico is studying a whole cluster of similarly isolated villages.

However, this cottage industry also has its limitations. Although some disease-related variants will be abnormally frequent in such inbred populations, others will be totally absent and thus cannot be studied until one finds another suitable group that has them. This is where large-scale, nationwide genome projects are hoping to make their mark. The oldest study of this kind is the one in Iceland. Like the Sardinian villages, Iceland has remained isolated for centuries, but it started from a much bigger gene pool, gradually expanding to today's population of 275,000.

In December 1998, the Icelandic parliament passed a bill enabling the creation of a centralized medical database, after a year-long public debate on ethical and privacy problems arising from this proposal. Much of the debate was based on the issue that a private biotech company, namely deCode Genetics, chaired by the Icelandic geneticist Kari Stefansson, is running and controlling the database. To defuse this issue, the parliament created a Governing Committee including members of the public, that acts as a kind of consumer watchdog to oversee this project.

Public debate is remarkably absent in a country that tries to beat Iceland's genome project with larger numbers and a more diverse gene pool: Estonia. Although the core Estonian population has been sedentary on the Baltic shore for more than 5000 years, some mingling with neighbours must have taken place. Among today's 1.45 million citizens of the Estonian Republic, Russians form the strongest ethnic minority with 28%. Thus, the proponents of the Estonian Genome Project argue, Estonians are less exceptional than the isolated populations, while still being sufficiently homogenous for meaningful genetic studies.

The Estonian Genome Foundation (EGF) is currently seeking corporate sponsors and parliamentary approval for what would become the world's largest database of medical, genetic and genealogical data. More than 90% of the Estonian population are expected to consent to part with 50 ml blood and their medical records.If successful, this project would result in a database holding sensitive information on 1.3 million individuals, including people's health status, diseases, important health and behavioural risks and genealogical information, as the EGF cheerfully proclaims on its website. Even though the names of the subjects will be replaced by anonymous codes, it is hard to imagine how a database of this size and with such sensitive personal information could be handled without at least a small risk of leakage and misuse of data. Using only DNA samples from deceased hospital patients (who have consented before dying) may be a way around the privacy problems arising from such large projects.

While the village-sized projects certainly have the advantage that the smaller databases are more easily kept safe from misuse, ethical issues are far from absent in these cases. Like in anthropological studies of native tribes in Africa or South America, the very remoteness of the villages under study throws up ethical concerns. In this case the question has to be asked whether agreement to give a blood sample constitutes informed consent to a full scale characterization of genomic markers. As even the geneticists don't know yet what these markers will tell them, the subjects of the research cannot possibly know — and probably lack the ability to imagine — what kind of information about themselves they are handing over. If geneticists don't want to limit themselves to studying the DNA of deceased patients, they may have to redefine what informed consent means in the postgenomic era.

Michael Gross is a science writer based in Oxford, UK. URL: www.michaelgross.co.uk