NEWS & VIEWS

GENOMICS

Guilt by association

Anne M. Bowcock

In a tour-de-force demonstration of feasibility, a consortium of 50 research teams uses 500,000 genetic markers from each of 17,000 individuals to identify 24 genetic risk factors for 7 common human diseases.

Mr Woodhouse, the comical hypochondriac of Jane Austen's *Emma*, takes great comfort in blaming his various ailments on the rain, the cold and an unfortunate piece of wedding cake. He would, no doubt, have been greatly surprised to learn that even his most rudimentary ailments resulted, at least in part, from genetic factors. Reporting on page 661 of this issue[1], a consortium of more than 50 British groups, known collectively as the Wellcome Trust Case Control Consortium (WTCCC), asserts just that. In the largest study of its type so far, the WTCCC has examined the genetic underpinnings of seven common human diseases: rheumatoid arthritis, hypertension, Crohn's disease (the most common form of inflammatory bowel disease), coronary artery disease, bipolar disorder — also known as manic depression — and type 1 and type 2 diabetes.

The WTCCC study is groundbreaking in various respects. It not only confirms the involvement of some genes for which disease association has previously been reported, but it also identifies several novel genes that affect susceptibility to common diseases. Moreover, it models a successful and instructive approach to large-scale genomic scans of this type, showing that a set of common controls can be used for a variety of diseases with relatively little loss of analytical power. Its success also provides strong grounds for performing such studies on an even larger scale.

The WTCCC investigators examined genetic variation at 500,000 different positions within the genomes of 17,000 individuals living in Britain using a genome-wide association scan (Fig. 1). This statistical approach compares the frequencies of genetic variation in disease cases and in healthy controls from the same population. Using the signal from each position as an indicator for the DNA sequence that surrounds it, genome-wide association scans examine the relationship between each DNA position and a particular trait (such as diabetes). Strong 'association' between a DNA position and a trait marks the general locale of the offending alteration, even if it is not itself the cause.

The concept of drawing an association between biological traits and disease is hardly new[2], but the scope and scale that the WTCCC
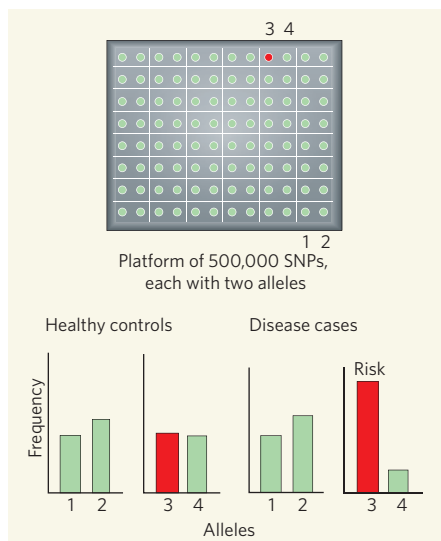

Figure 1 | Genome-wide association scan. To identify genetic risk factors for common diseases, the WTCCC researchers[1] scanned DNA from patients (2,000 per disease) and controls (3,000 shared for all seven diseases studied) for the frequency with which they contained each of the 500,000 genetic markers, or single nucleotide polymorphisms (SNPs), from the human genome. After statistical evaluation of the data, they found that most markers showed very little difference in the frequency of their two constituent forms — or alleles — between controls and cases. However, some SNPs occurred at a greater frequency in patients. Such alleles (one is shown in red) can be considered a genetic risk factor for a particular disease.

attained in their application of this concept is unprecedented. Crucial to both the success of this study and keeping its cost reasonable were DNA from large numbers of unrelated patients; the availability of the complete DNA sequence of the human genome; the subsequent cataloguing of a large component of variation in the genome in the form of single nucleotide polymorphisms (SNPs)[3]; the completion of the HapMap project[4], which provided information on the statistical relatedness of SNPs; and the availability of high-throughput technologies that allowed for parallel typing of 500,000 markers representing most of the common variation in the genome.

For the seven diseases studied by the WTCCC, strong statistical evidence for association was obtained for 12 previously identified genomic regions and a similar number of new regions. Although this WTCCC report is based on initial studies, independent groups[5–9] have confirmed the involvement of all but one of these most significant regions through replication studies. Some of the other identified regions with less statistically significant disease association are also likely to be true indicators of genetic risk; so these will need to be further evaluated in additional large sets of patients and controls. Indeed, because the WTCCC data will be publicly available, they will be a useful resource to other groups and consortia embarking on similar efforts to investigate genetic-association markers in these and other diseases. These researchers include members of the Genetic Association Information Network[10] (GAIN), the Framingham Genetic Research Study and the Women's Health Study.

With many of the genomic regions identified by the WTCCC, the next step will be to study the exact nature of the disease-causing variants, rather than the marker SNP with which each is associated. From this and previous studies, it seems that variations leading to common disease are diverse; some alter the coding sequences of genes, others lie within their noncoding sequences, and some are even located within gene deserts — regions of a chromosome that contain no genes. So understanding the biological function of disease-risk-associated genomic regions will be challenging.

Two replication studies relating to the WTCCC findings are also published today[5,6], revealing connections between the genomic regions associated with the risk of type 1 diabetes and Crohn's disease and their underlying biology. Some of the known and newly identified genetic risk factors for type 1 diabetes alter the development or function of immune cells, leading to aberrant recognition of pancreatic islet cells as foreign particles. But additional susceptibility genes identified recently[5] do not fit easily into this simple model.

For Crohn's disease, one of the newly identified[6] susceptibility genes is of particular interest because it is proposed to control the spread

645

of intracellular pathogens by autophagy — the process of cellular self-digestion. This is the second gene to be implicated in Crohn's disease through involvement in autophagy; the first was identified earlier this year[11,12]. Moreover, an increasing body of evidence, including the latest replication study[6], points to defects in the early immune response and the handling of intracellular gut bacteria in the pathogenesis of Crohn's disease.

The overall increase in risk (1.2–1.5 times) conferred by the genetic factors identified in the WTCCC study[1] is in agreement with those reported by others. However, these factors are unlikely to explain completely the clustering of any of these diseases in families, and there are other genes (possibly many of very small effect) — or rare variants of genes — that are still to be identified for these and other diseases.

One unexpected result of the WTCCC study was the identification of 13 regions with pronounced geographical variation within Britain. Among these regions is a large cluster of genes that encodes the major histocompatibility complex, which is well known for its function in the immune response and autoimmune disease[13], and a gene that is involved in lactase persistence, or the ability to digest milk[14,15]. Some of the other regions are thought to function in preventing diseases such as pellagra, tuberculosis and leprosy. Although the infectious agents responsible for tuberculosis and leprosy are now rare in Britain, they have left behind genetic footprints in the existing population that probably led to some degree of protection in the past. Several of these are also candidate genes for autoimmune disease[5].

Despite the magnitude and wealth of information that this study[1] provides, other questions about the genetic basis of common disease remain. The answers will become increasingly important as we enter an era of personalized medicine, in which therapy is tailored to an individual's genetic constitution. It will become crucial to discover which genes predispose individuals to these diseases; how genes interact with each other to increase the risk of a particular disease; and what proportion of disease is due to rare variants that would be hard to detect with current approaches.

We will also want to know whether different patients can be stratified into subpopulations on the basis of genetic risk factors, and what role the environment has in triggering disease. The Genes, Environment and Health Initiative (GEI) of the US National Institutes of Health already aims to develop tools to assess environmental contribution and to answer some of the other questions. Ultimately, comprehensive answers that would allow the translation of genetic susceptibility into scientifically sound medical practice will require much larger patient populations, well-annotated clinical databases and sophisticated environmental assessment. One wonders what Mr Woodhouse would have to say to that. ■

Anne M. Bowcock is in the Departments of Genetics, Pediatrics and Medicine, Division of Human Genetics, Washington University School of Medicine, 4559 Scott Avenue, Saint Louis, Missouri 63110, USA.
e-mail:bowcock@genetics.wustl.edu

1. Wellcome Trust Case Control Consortium *Nature* **447,** 661–678 (2007).
2. Buckwalter, J. A., Wohlwend, C. B., Colter, D. C., Tidrick, R. T. & Knowler, L. A. *Surg. Gynecol. Obstet.* **104,** 176–179 (1957).
3. Carlson, C. S. *et al. Am. J. Hum. Genet.* **74,** 106–120 (2004).
4. International HapMap Consortium *Nature* **437,** 1299–1320 (2005).
5. Todd, J. A. *et al. Nature Genet.* doi:10.1038/ng2068 (2007).
6. Parkes, M. *et al. Nature Genet.* doi:10.1038/ng2061 (2007).
7. Zeggini, E. *et al. Science* doi:10.1126/science.1142364 (2007).
8. Saxena, R. *et al. Science* doi:10.1126/science.1142358 (2007).
9. Frayling, T. M. *et al. Science* doi:10.1126/science.1141634 (2007).
10. www.fnih.org/GAIN2/home_new.shtml
11. Hampe, J. *et al. Nature Genet.* **39,** 207–211 (2007).
12. Rioux, J. D. *et al. Nature Genet.* **39,** 596–604 (2007).
13. Tomlinson, I. P. & Bodmer, W. F. *Trends Genet.* **11,** 493–498 (1995).
14. Cavalli-Sforza, L. *Am. J. Hum. Genet.* **25,** 82–104 (1973).
15. Enattah, N. S. *et al. Nature Genet.* **30,** 233–237 (2002).

SPECTROSCOPY

The magic of solenoids

Arthur S. Edison and Joanna R. Long

A technique known as magic-angle spinning has helped make nuclear magnetic resonance spectroscopy as sensitive for solids as it is for solutions. Inductive thinking leads to even better signal detection.

The great strength of nuclear magnetic resonance (NMR) spectroscopy is that it can determine, non-invasively and at atomic resolution, the chemistry, structure, dynamics and overall architecture of samples in solid, liquid or even gaseous forms. The liquid version of the technique, solution NMR, is used routinely to identify small molecules, study protein structures and dynamics, and probe intermolecular interactions. Solid-state NMR teases out the structure and properties of materials, surfaces and biological solids such as human tissue. But compared with many other analytical techniques, NMR has extremely poor sensitivity. A great deal of research has sought to improve this situation: on page 694 of this issue[1], Sakellariou *et al.* describe a potential leap forward for solid-state NMR.

When atomic nuclei with non-zero spin are placed in an external magnetic field, they become polarized, precessing rather as a gyroscope does in Earth's gravitational field. When electromagnetic radiation of a frequency (energy) that corresponds exactly to that of the energy gap between two states of different polarization is applied to the sample, the nuclei resonate, jumping between those states. The accompanying gyroscopic precession of the spins induces a current in a conducting coil placed around the sample. This basic principle is both NMR's blessing and its bane as a spectroscopic technique: the small energies make the approach non-destructive, but they also make it difficult to distinguish the characteristic polarization (or signal) from thermal noise.

The signal-to-noise ratio in NMR measurements can be improved by either one of two general routes. The first of these is enhancing
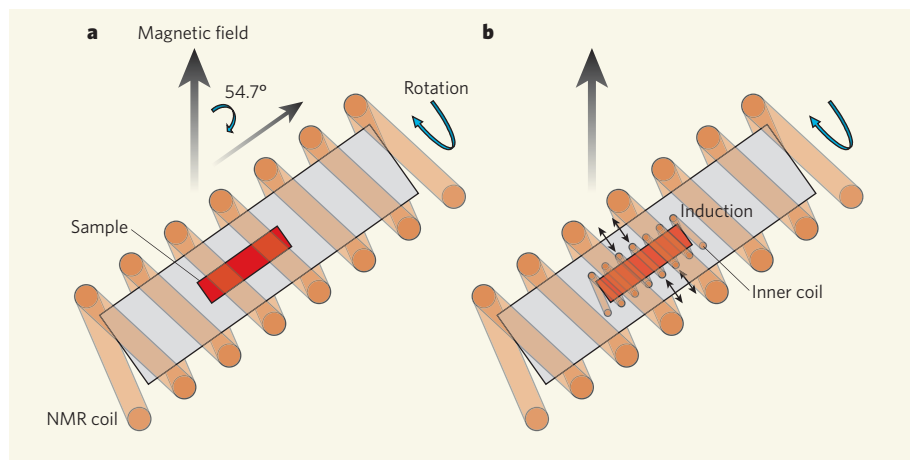

Figure 1 | Inductive logic. a, In the traditional 'magic-angle spinning' approach to solid-state NMR, a spectrum of better resolution is achieved by rapidly rotating the sample, at an angle of 54.7° relative to the main magnetic field, within a static coil assembly. **b,** Sakellariou and colleagues' alternative approach[1] uses the inductive coupling of a smaller coil rotating with the sample to the larger static coil to produce a similar effect. The result is a higher sensitivity and the capability to investigate smaller samples.