Cracking the code of autoimmunity

Advances in genetic analysis are providing a wealth of insights into the causes of autoimmune disease, but translating this understanding into treatments is a considerable challenge. **By Simon Makin**

In 2005, immunologist David Hafler, then at Harvard Medical School in Boston, Massachusetts, visited the laboratory of geneticist John Todd at the University of Cambridge, UK. Hafler was studying multiple sclerosis (MS), whereas Todd was mainly concerned with type 1 diabetes. Both were using a then-new approach known as genome-wide association studies (GWAS) to look for the same thing: regions of genetic code that are associated with disease.

Hafler and Todd both had some leads, but nothing conclusive. After all, stringent statistical tests are required when making so many comparisons, to ensure that associations are not down to chance. "We had no idea if association studies would work," says Hafler, who is now at Yale School of Medicine in New Haven, Connecticut. Chatting over beer in a local pub, the pair agreed to peek at each

other's data. Despite the fact that they were studying different autoimmune diseases, they noticed a lot of overlap in the genetic regions they had flagged — and that gave them an idea. "We said, let's pick the ones that are in both diseases," says Hafler. In 2007, their cooperation led them each to identify several genetic regions that are strongly associated with autoimmune disease — some of the first such regions uncovered by genome-wide analysis.

GWAS have now implicated hundreds of regions of the genome in various autoimmune diseases. These variants typically have small individual effects, but collectively they contribute significantly to risk and reveal the basic biology of autoimmunity. Most lie in non-coding regions, but epigenetic sequencing technologies are helping researchers to unravel their complex function. And advances in sequencing are accelerating efforts to find

rare variants involved in autoimmune diseases, which could further aid the development of treatments.

Faltering first steps

Autoimmune diseases arise from a combination of genetic predisposition and environmental triggers that disrupt the immune system's ability to ignore a person's own tissue and cells. In rare cases, an autoimmune disease is monogenic, caused by mutations in a single gene. But most of the time, autoimmunity is polygenic, with genetic risk determined by the interplay of numerous genes.

In the 1970s, the human leukocyte antigen (HLA) region became the first area of the genome to be linked to autoimmune disease. The region contains genes involved in presenting antigens to immune cells. It harbours risk factors for numerous autoimmune diseases,

and typically contributes the largest risk of any genetic factor. Around half of genetic susceptibility to type 1 diabetes, for instance, is accounted for by HLA variation.

This large effect made HLA fairly straightforward to identify, and other genes with relatively large impacts were soon similarly spotted. *PTNP22*, for example, which encodes an enzyme involved in multiple immune-system signalling pathways, has been linked to numerous conditions. Certain rare syndromes caused by a single gene also provided some early insight. Study of IPEX (immunodysregulation polyendocrinopathy enteropathy X-linked) syndrome, a monogenic condition that counts multiple autoimmune problems among its symptoms, showed that the protein FOXP3 is crucial to the development and function of regulatory T cells[1], which typically act as brakes on the immune response (see page S60).

However, early attempts to detect variants that make smaller contributions to risk were largely unsuccessful. Together these variants represent a considerable proportion of genetic predisposition to autoimmune disease, but the primary method of genetic analysis at the time — painstaking investigations of inheritance patterns, known as linkage studies — lacked the power to detect their modest individual effects.

The rise of GWAS

The sequencing of the human genome in 2003 enabled GWAS to transform the study of human genetics. Previous efforts were either linkage studies or hypothesis-driven studies of individual genes, but GWAS enable researchers to scan the entire genome for clues.

The technique exploits knowledge of common single-letter changes in the genetic code, known as single nucleotide polymorphisms (SNPs), to quickly compare the frequency with which they occur in people with and without disease. The lists of genetic regions generated are called loci. As Hafler and Todd found with their early efforts, establishing which loci are linked to disease by more than chance is difficult and often requires large sample sizes. "You need tens of thousands of patients to get the proper signal," says Alexander Marson, a geneticist at the University of California, San Francisco. As a result, large consortia have been formed to study almost every common autoimmune condition.

GWAS have been especially fruitful in the study of autoimmune disease. Although individual variants in the hundreds of associated loci have only a small impact on risk, they bring about significant changes at a molecular level. "The phenotypes are quite

remarkable," says Hafler.

Many of the genes implicated by these variants appear repeatedly across conditions, in a similar way to HLA and *PTNP22*. The overlap suggests that autoimmune diseases share some underlying mechanisms, which seem to involve either increased activity of pro-inflammatory T cells or impaired function of regulatory T cells or B cells.

One region that is linked to many autoimmune diseases, including Crohn's disease, systemic lupus erythematosus and psoriasis, contains the *TYK2* gene. This encodes an enzyme crucial to the function of some cytokines — signalling molecules that are central to coordinating immune responses. Two of the variants associated with disease by GWAS involve changes in *TYK2* itself, which some researchers had suggested reduced its function. But the variants' effects on risk were odd: one was thought to confer protection against every associated disease, whereas the other conferred protection against some, but increased risk for others. "Something didn't add up," says immunologist Calliope Dendrou at the University of Oxford, UK.

Dendrou and her colleagues conducted experiments in human cells that showed that only one of the two suspect SNPs directly alters *TYK2* function[2]. Their findings suggested that links between the second suspect and autoimmune disease are instead due to a different, co-occurring SNP that affects the gene *ICAM1*.

"If you have a lot of the common variants, you really are at increased risk."

The SNP that they did pin down in *TYK2* had an unusually strong protective effect. In a mouse model of MS, the team showed that animals with two copies of the variant experienced a near total loss of *TYK2* function. This impaired signalling involving the cytokines interferon-I, interleukin (IL)-12 and IL-23, and conferred complete protection. "The animals didn't get disease at all," says Dendrou. "That really made a case that if you could strongly block *TYK2* function, that should have a protective effect across multiple conditions."

A phase II clinical trial of a drug that blocks *TYK2* function in psoriasis reported that up to three-quarters of patients experienced a 75% reduction in skin lesion severity[3]. One-quarter saw their lesions clear completely. A phase III trial is under way, as are early-stage trials in Crohn's disease, lupus and psoriatic arthritis — conditions that are known to involve IL-12 and IL-23. "The genetics alone tells us one could

trial this across a whole range of conditions," says Dendrou. "But it makes sense to start with conditions where we know that some of the key cytokine pathways are involved."

Refining output

The avalanche of clues produced by GWAS has brought fresh problems. Because DNA is inherited in chunks called haplotypes, the borders of which are not random, disease-associated loci contain multiple variants that frequently co-occur. This makes it tricky to work out which variants are truly driving disease.

A first step towards overcoming this difficulty is to generate higher-quality genetic data. Sampling the genome more densely in disease-relevant areas, combining multiple sources of genetic information and applying sophisticated statistical techniques can all help to map the genome more finely.

But that still leaves another problem. Unlike the *TYK2* variant that Dendrou characterized, most variants involved in autoimmune diseases lie in non-coding regions — parts of the genome that do not contain instructions for making a protein. Many of these act as switches, turning coding genes on or off, and regulating their activity levels. Unfortunately, the biological consequences of non-coding variants are difficult to interpret. "We're much less sophisticated at understanding variants that affect those switches," says Marson.

In 2015, a team co-led by Marson combined fine-mapping, gene expression data and analysis of epigenetics (mechanisms that influence gene activity without changing the underlying code) to estimate that around 90% of causal variants in autoimmune disease are non-coding[4]. Using new sequencing technologies, the researchers constructed maps of epigenetic features, such as where proteins that control gene transcription can attach. This led them to estimate that around 60% of causal variants lie inside immune-cell enhancers — regulatory regions involved in determining which genes are activated, when, and in which cells.

Regulatory machinery involved in immunity is usually context-specific: if researchers are not looking in the right type of cell at the right time, the impact of a variant might be missed. Marson's team therefore generated maps in a range of immune cells in various activation states. They found that SNPs associated with autoimmune disease mapped most closely to enhancers that are switched on in activated immune cells — especially helper T cells, which activate other immune cells.

Working out which cells are being influenced by a non-coding variant is one thing; knowing what effect the variant is having

Immunologist Calliope Dendrou and colleagues discuss *TYK2*-mediated signalling.

on the cell is quite another. To fully understand the biology underlying autoimmunity, researchers must also work out which genes variants regulate, and the effect that this has on disease. In the past, testing the functions of precise genetic sequences was extremely challenging, but the advent of gene editing has simplified things considerably. "Now we have the ability not only to pinpoint the important sequences, but actually perturb them and see what happens," says Marson.

In a 2017 study[5], Marson and colleagues used CRISPR to screen for enhancers that regulate the *IL2Ra* gene. "This has been highlighted by many GWAS as a critical risk locus, but it's really complex," he says. "The variants are all in non-coding regions, and there's been controversy about which are actually causing disease." Marson's team found an *IL2Ra* enhancer that corresponded to a variant that was flagged as a risk factor in multiple autoimmune diseases. The team then used CRISPR to engineer mice with the variant, and saw that it altered the timing and levels of *IL2Ra* expression in response to stimulation with certain antibodies. "We could actually watch how that single variant changes the way *IL2Ra* gets turned on," says Marson. "It doesn't determine whether it's on or off permanently — it just tunes the rate it's turned on in T cells, and that seems sufficient already to start changing risk."

Monoclonal antibodies against *IL2Ra* already exist, and new therapies targeting *IL2Ra* are under development for multiple diseases. Marson hopes that studies such as his will help to inform the use of such therapies.

Translation challenge

Common variation of the kind identified by GWAS represents one end of a spectrum, with rare monogenic conditions at the other extreme. Monogenic disorders can provide considerable insight into a disease because of their profound impact on a person's biology. Conversely, the effects of common variants are typically much smaller. Although these variants can help us to understand the fundamental biology of autoimmunity, some researchers think it might be difficult to target them with treatments. "If you have a lot of the common variants, you really are at increased risk, but they're all contributing a little bit," says Peter Gregersen, a geneticist at the Feinstein Institutes for Medical Research in Manhasset, New York. "The problem is teasing out what's really useful, versus what's giving insight into pathogenesis but may not be addressable."

Rather than decoding the function of common variants, Gregersen is returning to the rare end of the spectrum. "I'm increasingly interested in genes with variants that have big effects, even if they're monogenic," he says. "Because they might point you in the direction of a therapeutic."

For example, a rare form of lupus is caused by loss of function of a single gene called *DNASE1L3*. This gene encodes an enzyme called DNase-γ that helps to clear away the DNA of dead cells. Without it, DNA accumulates and stimulates an autoimmune response. "You have bits of DNA floating around that your immune system recognizes and starts making antibodies to," explains Karen Cerosaletti, a geneticist at the Benaroya Research Institute in Seattle, Washington.

Common variants in *DNASE1L3* also modestly increase risk for lupus, rheumatoid arthritis and scleroderma. This year, Gregersen's group showed that a SNP in *DNASE1L3* that had been linked to these diseases by GWAS does not eliminate the function of DNase-γ entirely, as had been thought, but rather reduces levels of the enzyme secreted by cells. The fact that this same pathway is disrupted in both rare and common forms of disease, and across multiple diseases, makes it a promising target for treatment. "That really suggests this is an important protective pathway," Gregersen says. Neutrolis, a biotechnology company in Cambridge, Massachusetts, is investigating a DNase-γ replacement therapy.

There is also increasing interest in variants that are rare but still cause only polygenic disease. They are likely to account for a small amount of risk in the population, but the effects of these rare variants on individuals can be large. "Rare variants are a tool to really unearth what mechanism a gene is using to contribute to autoimmunity," says Cerosaletti. "Then you can target that." The search for such variants is at an early stage in autoimmune disease, but efforts are being accelerated by the arrival of sequencing technologies that can decode the genome at single-nucleotide resolution — avoiding the problem of tagged regions containing multiple variants.

Some researchers think that common and rare variants might often point in the same direction. For example, whereas the common variants in *IL2Ra* are non-coding, Marson and his colleagues are studying a family with *IL2Ra* mutations that cause monogenic disease. "We're looking at the full spectrum," says Marson. They have shown they can use CRISPR to correct the mutation in immune cells derived from the family, suggesting that gene editing could treat rare autoimmune conditions. Marson contends that such examples of rare and common variants converging on the same pathways are probably the rule rather than the exception. "Together, they point to powerful levers in the immune system," he says.

Simon Makin is a freelance writer in London.

1. Bennett, C. L. *et al. Nature Genet.* **27**, 20–21 (2001).
2. Dendrou, C. A. *et al. Sci. Transl. Med.* **8**, 363ra149 (2016).
3. Papp, K. *et al. N. Engl. J. Med.* **379**, 1313–1321 (2018).
4. Farh, K. K.-H. *et al. Nature* **518**, 337–343 (2015).
5. Simeonov, D. R. *et al. Nature* **549**, 111–115 (2017).