

Translating GWAS in rheumatic disease: approaches to establishing mechanism and function for genetic associations with ankylosing spondylitis

Julie A. Osgood and Julian C. Knight

Corresponding author: Julian C. Knight, Wellcome Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK. Tel.: +44 1865 287651; E-mail: julian@well.ox.ac.uk

Abstract

Ankylosing spondylitis (AS) is a highly heritable chronic inflammatory arthritis characterized by osteoproliferation, fusion of affected joints and systemic manifestations. Many disease associations for AS have been reported through genome-wide association studies; however, identifying modulated genes and functional mechanism remains challenging. This review summarizes current genetic associations involving AS and describes strategic approaches for functional follow-up of disease-associated variants. Fine mapping using methods leveraging Bayesian approaches are outlined. Evidence highlighting the importance of context specificity for regulatory variants is reviewed, noting current evidence in AS for the relevant cell and tissue type to conduct such analyses. Technological advances for understanding the regulatory landscape within which functional variants may act are discussed using exemplars. Approaches include defining regulatory elements based on chromatin accessibility, effects of variants on genes at a distance through evidence of physical interactions (chromatin conformation capture), expression quantitative trait loci mapping and single-cell methodologies. Opportunities for mechanistic studies to investigate the function of specific variants, regulatory elements and genes enabled by genome editing using clustered regularly interspaced short palindromic repeats/Cas9 are also described. Further progress in our understanding of the genetics of AS through functional genomic and epigenomic approaches offers new opportunities to understand mechanism and develop innovative treatments.

Key words: ankylosing spondylitis; genome-wide association study; regulatory genetic variants; expression quantitative trait loci; chromatin conformation capture

Introduction

Ankylosing spondylitis (AS) is a severe chronic inflammatory arthritis and part of a group of diseases collectively known as spondyloarthropathies [1]. AS is characterized by inflammation and osteoproliferation, followed by bone fusion of affected areas, which are typically the spine and sacroiliac joints (Figure 1). The disease affects men two to three times more often than women and shares many features seen in other immune-mediated traits, including peripheral arthritis, anterior uveitis, psoriasis, inflammatory bowel disease (IBD), osteoporosis and cardiovascular disease (aortitis, aortic valve disease,

conduction disturbances, cardiomyopathy and ischaemic heart disease) [2]. It is associated with significant morbidity and mortality, including chronic pain, disability and accompanying comorbidities such as cardiovascular and gastrointestinal complications. This has a significant socio-economic impact, as AS particularly affects young adults [3]. The introduction of biologic therapies [4] highlights significant opportunities for improved care but currently do not affect disease progression, and the pathophysiology of AS remains relatively poorly understood [5, 6], limiting early effective intervention. Genetics offers significant opportunities to address this challenge given the high

Julie Osgood is a doctoral student at the University of Oxford conducting research into the functional genomics of ankylosing spondylitis.

Julian Knight is Professor of Genomic Medicine at the University of Oxford. His laboratory studies the impact of genetic variation on the innate immune response.

© The Author(s) 2018. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

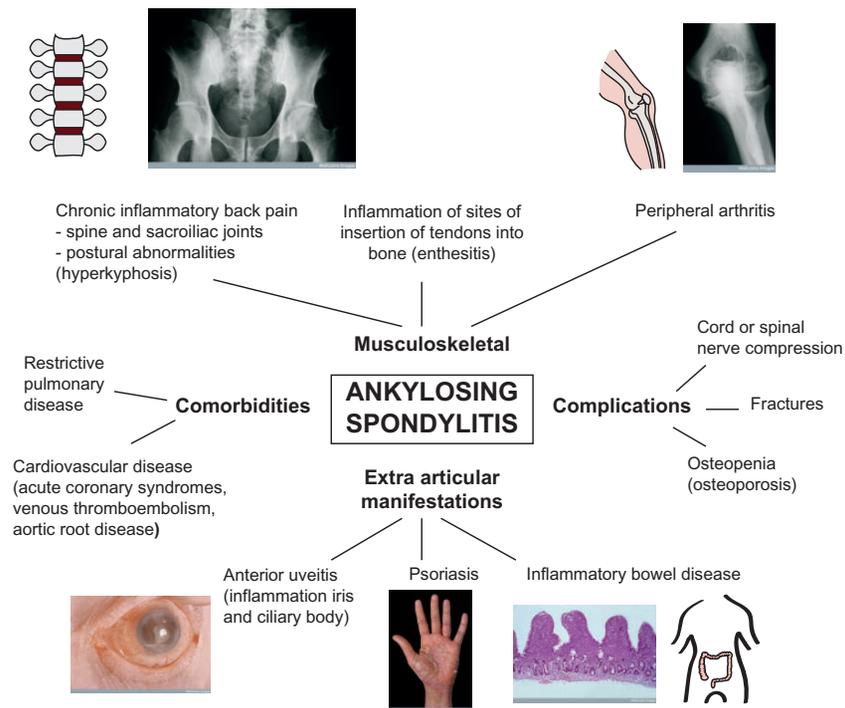


Figure 1: Overview of the musculoskeletal and extra-articular manifestations of AS together with comorbidities and major complications. Images from Wellcome Images under Creative Commons.

heritability of AS [7, 8]. Following the early discoveries in the 1970s regarding the role of human leukocyte antigen (HLA), substantial recent progress has been made in establishing genetic susceptibility to AS through genome-wide association studies (GWASs) [9–12]. This has the potential to provide new insights into disease pathogenesis and opportunities for therapeutic intervention. However, our current ability to establish function and mechanism for reported genetic associations is limited, requiring new approaches and innovative strategies to maximize the translational impact of GWAS in AS. In this article, we outline current progress and future strategies for addressing this important area of unmet need (Figure 2).

Genetics of AS

AS is a highly heritable disease with first-degree relatives being >50 times more likely to develop the disease than unrelated individuals, a parent-child recurrence risk of 7.9% and overall heritability of 90% [7, 8, 13]. An early breakthrough in genetic research into AS came in 1973 with the landmark discovery of association between disease occurrence and the major histocompatibility complex (MHC) Class 1 allele, HLA-B27 [14, 15]. Among AS patients, 96% are HLA-B27 positive however, only a minority of HLA-B27 positive individuals develop AS, and current understanding suggests that while HLA-B27 is the strongest genetic risk factor it is not sufficient for the disease to develop [7, 13, 16–18]. Monozygotic twins have a disease concordance rate of 63%, compared with 12.5% for dizygotic twins (27% in HLA-B27 positive dizygotic twins) [7].

The first genome-wide association study for AS was performed by the Wellcome Trust Case Consortium and the Australo-Anglo-American Spondylitis Consortium in 2007 [17]. This focused on a genome-wide set of 14 436 non-synonymous SNPs together with 897 MHC SNPs, identifying ERAP1 and IL23R as

new susceptibility loci. Subsequent GWAS and Immunochip studies identified 48 genomic loci reaching genome-wide significance and have further highlighted the role of non-MHC genes (Figure 3) [9, 11, 12, 17, 18]. Most disease-associated GWAS SNPs are not found in exons but rather in intronic and intergenic regions with putative regulatory effects [19, 20] and those associated with AS are no exception, including, for example, associations in at least two gene deserts. Many such regulatory variants can be key modulators of gene expression through generation or disruption of splice sites, change in transcription factor affinity, altered physical chromatin interactions, change in microRNA action or modification of DNA methylation, resulting in increased or decreased chromatin accessibility [21–24]. All of these can cause local effects, or can act at a distance, meaning the modulated gene(s) responsible for the observed disease association with the genetic marker can be hard to establish. Some of the most notable genetic findings from recent studies involving AS are the discoveries implicating involvement of the aminopeptidases, ERAP1 and ERAP2, and genes in both the tumor necrosis factor (TNF) and Interleukin-23 (IL-23) pathways [9, 11, 17].

ERAP1 is a promising candidate for disease, and is an important example of epistasis, as it has been shown to be associated with AS only in HLA-B27 positive individuals [11]. ERAP1 encodes an aminopeptidase involved in the processing of peptides before their presentation on MHC Class I molecules [25, 26]. There is evidence that an AS-protective single nucleotide polymorphism (SNP), rs31087, has a decreased ability to cleave peptides before MHC presentation and therefore slows down the rate of cleavage [11, 27], but the associations are complex with several different alleles involved [28]. A decrease in ERAP1 activity has also been associated with a decrease in the stability of HLA-B27 [29]. Chen and colleagues [30] examined the effect of silencing ERAP and found that there was an extension of both the N and C terminals of HLA-B27 epitopes in the absence of

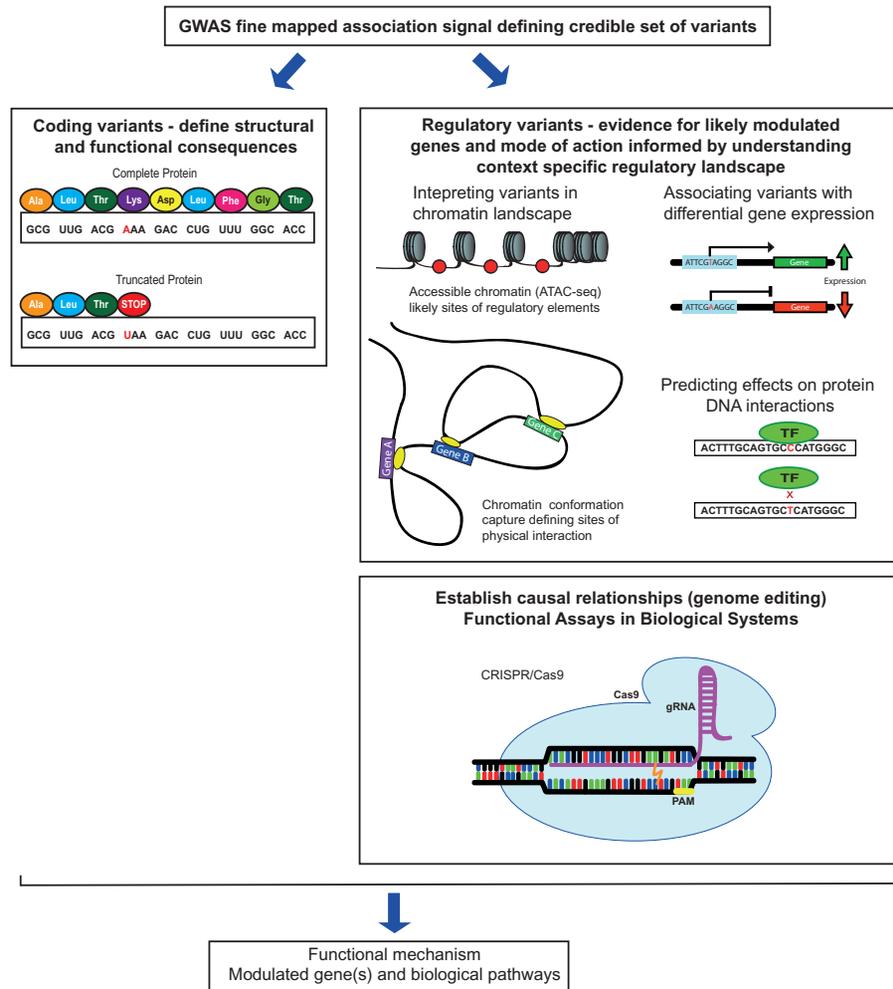


Figure 2: Illustration of some approaches enabling understanding of the functional basis of GWAS.

ERAP. It has also been found that ERAP1 activity governs HLA-B27 free heavy chain expression on the cell surface and may even be involved in Th17 promotion [31]. However, the exact mechanisms on how these findings may relate to disease remain unclear and more experimental evidence is required. ERAP2 has also been implicated in disease [9, 17]. It is a similar aminopeptidase responsible for N terminal cleavage of peptides before antigen presentation, and can form a heterodimer with ERAP1 [32, 33]. Other aminopeptidase genes associated with AS are LNPEP and NPEPPS [9].

IL-23 is a major proinflammatory cytokine, and several genes involved in IL-23 signalling have been implicated in AS through genetic studies, including tyrosine kinase 2 (TYK2), *IL12*, nuclear factor kappa B subunit 1 (*NFKB1*) and *IL23R*, encoding the receptor for IL-23 [9, 17, 18, 34]. SNPs found in *IL23R* have also been shown to be associated with other autoimmune diseases such as psoriasis [35], Crohn's disease [36], IBD [37], ulcerative colitis [38], psoriatic arthritis [39] and Behcet's disease [40]. On ligation with IL-23, *IL23R* initiates a signalling cascade through JAK2 and STAT3, leading to the production of IL-17 and IL-22, all of which are involved in inflammation. IL-22 is also thought to be implicated with the expression of genes involved in new bone growth [41].

GWASs have also uncovered SNPs near/in three major genes in the TNF pathway: *TNFRSF1A* [11, 17, 42–44], *TRADD* [45] and

TNFSF15 [46]. TNF is an important mediator of inflammation and remains the main target for biologic treatments [47]. It is proposed that TNF is involved in the initial inflammation of joints, which then leads to bone degradation followed by new bone growth and fusion of joints as a reparative attempt [48]. However, evidence from animal models suggests that anti-TNF therapies help treat inflammation but not the bone growth [49–51].

Fine mapping and resolving likely causal variants from GWAS

A major challenge to the translation of GWAS into mechanistic understanding is determining the causal variant(s), as this may not have been directly genotyped, and typically, there are many SNPs in linkage disequilibrium with the lead GWAS marker(s). Fine-mapping approaches often use a combination of statistics and functional annotations for variants. These include the use of genotyping arrays developed to study a specific set of SNPs, such as the ImmunoChip for immune-relevant variants, and statistical approaches that can define a small subset of statically likely causal SNPs, known as a credible set, that can then undergo functional annotations through the use of data sets and online predictors. Fine mapping typically involves

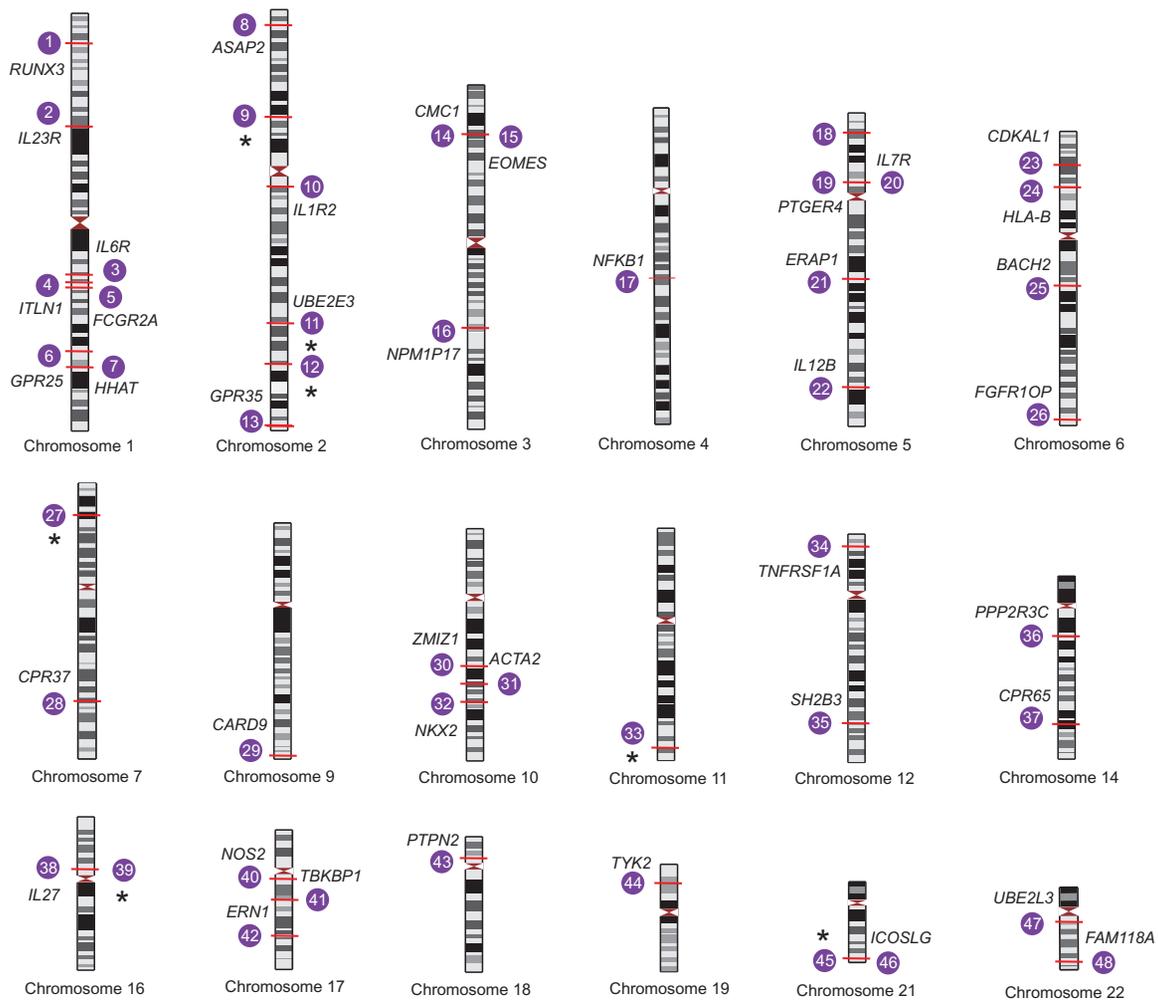


Figure 3: Overview of AS GWAS loci. Loci associated with AS at genome-wide significance reported by Ellinghaus et al. [20] are plotted (numbered 1–48). The implicated gene and chromosomal position (marked in red) are shown for each locus; where there is no reported gene, this is indicated by an asterisk. In the majority of cases, the genes listed are candidates based on proximity and biological plausibility, and causality has not been established except in a small number of instances.

imputation [52] and targeted re-sequencing [53] with stepwise conditional analysis to define independent association signals within a locus [54]. To narrow down the associations to particular variants, calculating a posterior probability in a Bayesian approach can define a credible set [54]. Credible sets can then be subjected to functional annotations such as arising from ENCODE [55] and NIH Roadmap [56] studies. A Bayesian fine-mapping approach has been used to examine 50 susceptibility loci for type 1 diabetes [57]. The authors identified 29 SNPs in 99% credible sets for enhancer regions in the thymus, T cells, B cells and CD34+ stem cells, allowing for future functional assays to be carried out using a manageable set of potential causative variants. Application in AS is enabled by international collaborative efforts to maximize samples sizes and power for GWAS such as through the International Genetics of Ankylosing Spondylitis Consortium (IGAS).

Characterizing role of causal genetic variants in landscape of gene regulation

Given the majority of GWAS SNPs for complex traits are found within introns and intergenic regions, considering the

diversity of effects on gene regulation involving allelic variation is essential [19]. Moreover, knowing how the genotype of a non-coding variant relates to gene expression is complex. For example, a single non-coding SNP may modulate the regulation of multiple genes through a given enhancer and multiple SNPs may exert a combined (haplotypic) effect on the regulation of a given gene with evidence that SNPs in linkage disequilibrium (LD) can act on multiple enhancers, all near each other, and collectively contribute to observed expression of a gene [58]. Complex processes govern gene regulation including post-translational modifications, such as methylation and acetylation, the physical conformation and looping of chromatin to bring, for example, enhancer elements proximal to regulated genes, microRNA activity, relative DNA-binding affinity for transcription factors and alternative splicing [21, 22, 24, 59]. Some of these processes are context specific in that an enhancer may only be active in a certain cell type or under a particular stimulus, or that regions of open chromatin will vary between cell types [60–62]. Variants that impact transcription factor-binding motifs, chromatin accessibility or physical chromatin interactions could all have significant effects on gene expression and relate to disease. Therefore, to better understand the role of non-coding variants in disease pathogenesis, there is value in

understanding the chromatin landscape across cell types and its interactions with both nearby and distant elements.

Defining chromatin accessibility

A highly informative predictor of the location of regulatory elements is identification of regions of open chromatin accessible to transcription factors, co-regulatory factors and chromatin interactions. DNase I mapping has been the golden standard for defining accessible chromatin and was used by the ENCODE Consortium for genome wide mapping of open chromatin regions [55, 63]. It relies on the use of the DNase I enzyme to create double-stranded breaks in regions of accessible open chromatin to generate fragments [64]. These fragments are then amplified and analysed by Southern blotting, real-time quantitative polymerase chain reaction (PCR) or next-generation sequencing to provide a visual map of regions of open chromatin [65, 66]. The assay for transposase-accessible chromatin using high-throughput sequencing (ATAC-Seq) is a more recently developed method to assay chromatin accessibility and was first introduced in 2013 by Buenrostro *et al.* [67] as an improvement on other methods used to perform epigenetic profiling. Like DNase I mapping, it is able to evaluate transcription factor-binding sites, nucleosome positions and regions of open chromatin in a single assay. However, unlike DNaseI, it works effectively with low cell numbers, maps transcription factor and nucleosome binding, and is credited with being more robust [63, 67]. Fragmentation is achieved using a Tn5 transposase that has the ability to not only fragment the DNA but will also simultaneously tag it with sequencing adaptors. In regions where the DNA is tightly wound around histones or bound by proteins, the transposase will be unable to cut the DNA, and therefore data are generated about where chromatin is easily accessible for DNA-binding proteins. This type of data can be overlapped with GWAS hits to annotate which variants are likely to be involved in gene regulation, and which are non-functional SNPs in regions of dense chromatin. For example, ATAC-seq was used recently to help prioritize schizophrenia GWAS risk variants. Forrest *et al.* [68] used excitatory neuronal differentiation from hiPSCs to examine genome-wide areas of open chromatin using ATAC-seq. From this, they were able to prioritize potential risk variants affecting neural development and at a specific risk locus, *MIR137*, narrowed down the list of potential causal SNPs to just one found in open chromatin. Furthermore, they were able to confirm using CRISPR that this SNP decreased expression of *MIR137*.

Genetic variants can influence gene expression by disrupting transcription factor binding, creating or removing methylation sites, changing splice sites, altering nucleosome structure or create new promoter-like elements [69–72]. These changes in gene expression can be of small magnitude or substantial, as seen with loss of GATA-1 binding associated with a single-nucleotide change in the *DARC* promoter in Duffy blood group antigen-negative individuals [73]. Mutations that change methylation sites can cause genes to become hypo- or hyper-methylated and can result in significant changes in expression by creating or removing areas in which proteins can bind to initiate transcription [70]. The acetylation of histones can also have an impact on gene expression and disease. Rubinstein-Taybi syndrome is an autosomal dominant disorder characterized by developmental delay and congenital abnormalities. It has been associated with mutations in the cAMP response element-binding protein, a histone acetyl transferase, which inhibit its function and decreases transcription [74]. Modulation

of splicing is a further important mechanism by which genetic variants may act, reported for example with *CTLA4* and associated with autoimmunity [75], while effects on mRNA stability may also be important as illustrated by *CDSN* and susceptibility to psoriasis [76]. Uncovering functional variants modulating such mechanisms would be a major step in understanding the underlying pathogenesis of AS.

Defining chromatin interactions

Understanding the 3D landscape of the genome is another current challenge to characterizing non-coding disease variants. Distant, physical interactions between *cis*-acting regulatory elements and promoters can occur to regulate transcription through the bending and folding of chromatin. Topologically associated domains (TADs) describe regions of frequent physical chromatin interactions [77]. Cohesion and the transcriptional repressor CCCTC-binding factor help regulate TAD borders and facilitate interactions within, including bringing enhancer elements close to different promoters [78]. Chromatin is highly likely to interact with another region of chromatin within the same TAD to influence the expression of a gene, but unlikely to interact with chromatin from another TAD [79, 80]. Therefore, understanding these 3D interactions provides insight into the more complex regulation of gene expression and helps further annotate GWAS SNPs. Chromatin conformation capture (3C) is one approach to analyse these physical interactions and provide evidence connecting regions of the genome with the promoters they interact with, and can therefore associate SNPs with the genes they may regulate [81]. In 3C, DNA is crosslinked within the cell and then digested and re-ligated. Chromatin that physically interacts will ligate together and produce products that can be quantified by real-time PCR. A significant advance was the development of Hi-C using next-generation sequencing as a high-throughput way of examining the whole genome [82, 83]. Complementing this, in 2014, Hughes and colleagues [84] developed Capture C allowing analysis of hundreds of target regions at high resolution. In Capture C, a 3C library is generated by sonicating DNA into small fragments and the ligation adaptors. An oligonucleotide capture is then used by binding a region with a known sequence, usually an enhancer or promoter of interest. This is followed by library amplification, and finally, paired-end sequencing to show which enhancers and promoters physically interact with certain loci, indicating a *cis*-interaction. For diseases like AS where large numbers of variants need to be analysed, chromatin conformation capture is potentially useful in providing evidence for the potential regulatory mechanism of SNPs identified through GWAS.

Chromatin conformation capture technology has already proven instrumental in understanding regulatory effects in other diseases. Smemo *et al.* [85] used this approach in studying obesity, to undercover an interaction occurring between associated variants in an intron and a distant enhancer. Previous GWAS data had indicated that variants within introns of the *FTO* gene were associated with obesity [86–88]. *FTO* was then found to encode an enzyme involved in metabolism and regulation of body weight [89, 90]. Initial efforts focused on whether these variants modulated *FTO* gene expression. Smemo *et al.* [85] then performed a 3C analysis on mice and found that the intronic region of *FTO* containing these variants was interacting with the *Irx3* gene, encoding a homeodomain transcription factor, located over 500 kb away. In human brain samples tested, these obesity-related variants showed an association with expression levels of *IRX3*, but surprisingly not with *FTO* as

previously predicted. Continued investigation into *Irx3* showed that a 30% reduction in body weight occurred in mice with an *Irx3* knockout, further validating this interaction. The use of 3C allowed for the discovery of a new role for *IRX3*, as it had never before been associated with obesity. This finding demonstrated the power of chromatin conformation capture to uncover unique regulatory interactions occurring at great distances across the genome, and this technology is likely to be instrumental in defining regulatory variants in other complex disorders, such as AS.

Mapping gene expression as a quantitative trait

Expression quantitative trait loci (eQTL) mapping has been another valuable tool in understanding the function of non-coding variants by establishing genetic association for a given variant with differences in gene expression. eQTL have been found to be often highly context specific, dependent on cell and tissue type [91–93] as well as activation state [94, 95] and disease context [96]. The Genotype-Tissue Expression (GTEx) study provides an extremely powerful resource for exploring tissue-specific effects, including eQTL for 53 different tissues [97, 98]. AS is an example of a disease in which the most relevant cell type to disease pathogenesis remains unresolved, as well as the specific disease-relevant conditions to evaluate how variants may be affecting gene expression, although there are now diverse strands of evidence prioritizing specific cell types (discussed in the following section). Further work evaluating eQTL using cells and tissue from patients with active disease could provide the most accurate representation of the correlation between genotype and gene expression.

Applying -omic approaches in context: biology of AS

The tools described above can be useful in helping us understand GWAS variants, but the value of the data produced relies heavily on how relevant the cell and tissue types from which the data were collected are for disease. Major cell types involved in autoimmune disease, and potentially important in AS, are T cells, natural killer (NK) cells, neutrophils, dendritic cells, monocytes and B cells [99]. Evidence for the involvement for specific circulating immune cell types in AS, which are accessible to investigators for functional genomic studies, is outlined below recognizing that future work on synovial tissue as well as cartilage and connective tissue around the facet joints of the spine and sacroiliac joints will be essential and aided by development of minimally invasive methods for sampling.

One theory of AS pathogenesis, known as the arthritogenic peptide hypothesis, relies on the involvement of CD8+ T cells [99, 100]. This hypothesis suggests that the self-peptides displayed by HLA-B27 bear a resemblance to peptides produced by foreign microbes and become the target of autoreactive CD8+ T cells. These cells then initiate an immune response and inflammation ensues. Moreover, it was found that CD4+ T cells express KIR3DL2, which is a receptor that recognizes homodimers of HLA-B27 but does not recognize heterodimers of HLA-B27 [101]. Interestingly, it was found that AS patients who are HLA-B27 positive have an increased number of cells expressing KIR3DL2. The binding of KIR3DL2 to HLA-B27 homodimers could trigger an immune response and lead to inflammation, and it is also suggested that binding triggers Th17 responses [31]. These theories suggest a connection between T cells and HLA-B27, but there are other factors involved in inflammation relevant to AS.

There is much evidence on the role of IL-23 in AS pathogenesis [9, 17]. To produce an inflammatory response, IL-23 binds to its receptor, IL-23R, which then induces IL-17. IL-17 activation leads to the production of IL-6 and IL-8, which are major contributors to inflammation. T helper 17 (Th17) cells are the main contributors of IL-17, but there is evidence that it is also secreted by NK cells, mast cells, CD4+ T cells and neutrophils [102, 103]. Th17 cells are promising as a key cell type involved in AS pathogenesis, as they are involved in recruiting neutrophils and monocytes to a site of inflammation and encourage the development of osteoclasts, which absorb bone tissue, potentially one of the triggers for bone growth to occur at a later stage [104]. More evidence supporting the role of Th17 cells is that high levels of Th17 cells have been found in the peripheral blood of AS patients and that anti-TNF therapy, which is the most effective current treatment available for AS, reduces the number of Th17 cells in the blood [105].

Functional genomics studies have also provided evidence for the involvement of T cells in AS. Farh *et al.* [106] conducted a study to investigate fine mapped GWAS SNPs across immune diseases. To understand which cell types were contributing to disease, they performed a comparative analysis between SNP location and chromatin maps showing *cis*-regulatory elements across different cell types and found enrichment for those in Th17, Th0, Th1 cells and monocytes in AS. Monocytes in spondyloarthritis patients exhibit an upregulation of genes involved in essential inflammation pathways. One study used label-free quantitative expression profiling to examine protein expression [107]. The authors noted upregulation of proteins involved in leukocyte recruitment, and important signalling pathways, such as vascular endothelial growth factor, Janus kinases/Signal Transducer and Activator of Transcription proteins and toll-like receptor in monocytes from AS patients. Moreover, there was upregulation of genes in the ubiquitin proteasome pathway in AS monocytes [107]. This pathway is involved in the formation of peptides to be presented by HLA Class I proteins, like HLA-B27.

In terms of other circulating immune cells, one study looked at the involvement of regulatory B cells in AS and found that there was a defect in these B-cell populations in AS [108]. B cells can suppress T cells by secreting IL-10 [109]. The authors found that IL-10 secreting B cells were reduced in AS and found that controls were better able to suppress memory CD8+ T cells and naïve T cells [108]. The same study also found that B-cell population numbers were increased in AS, suggesting that they may play a significant role in disease. As mentioned above, IL-17 is a key regulator of inflammation, and there is evidence that B cells are receptive to it and are known to accumulate at the site of inflammation [102]. This could be further evidence implicating B cells in AS pathogenesis. It has been suggested that there are specific receptors found on NK cells, involved in the innate immune response and inflammation, that are able to interact with homodimers of HLA-B27 and contribute to AS [101]. Dendritic cells are important in antigen presentation and also secrete IL-23, which is thought to be a major signalling pathway in AS [110]. Further knowledge is needed relating to which cells are the most involved in the inflammation and bone growth occurring in AS.

Given this current heterogeneity in terms of the key immune cell types in AS, single-cell genomics is likely to significantly impact on our understanding of the functional genomics of disease [111]. While single-cell transcriptomics has been widely adopted, opportunities for epigenomic profiling are increasing. For example, Buenrostro *et al.* [112] used a microfluidic approach to isolate single cells from eight different cell types and then

performed ATAC-seq. They believed that *trans*-factors would cause variability in chromatin accessibility between individual cells and found that this was related to specific transcription factor binding, for example, *GATA1* and *GATA2* in K562 cells. They also found that this cell-to-cell variation is related to cell type. *NFKB*, one of the major inflammatory mediators, seemed to increase chromatin state variability between cells in one cell type but not in others. Epigenetics was further examined by investigators looking at protein binding and control of regulatory elements at a single-cell level [113]. They studied H3K4me2 in embryonic stem (ES) cells and found that enhancer and promoter occupancy varied between cells, and they were able to categorize cells from the ES population based on epigenetic profile into three groups with different protein binding. These studies suggest that looking at genomics from a single-cell point of view can provide us with a new depth of information. In the context of AS, this technology could be used to examine expression and epigenetic profiles from different cell types or tissues to help resolve relevant cell types and context-specific interactions.

Assigning causal relationships for disease-associated variants

Our current tools used for evaluating GWAS hits can help to identify the genes and genomic regions that variants are interacting with and how they may influence gene expression, but they do not necessarily provide causal evidence for mechanism. Advances in genome editing based on clustered regularly interspaced short palindromic repeats (CRISPR) are a powerful approach to address this challenge. CRISPR are short segments of prokaryotic DNA that, when combined with CRISPR associated protein 9 (Cas9) and specific guide RNAs (gRNAs), have the ability to selectively bind to regions of DNA and create a double-stranded break [114–116]. Once this break has been created, the cell will repair the damage via non-homologous end joining (hence disrupting the normal sequence and successfully used to knockout genes), or homology-directed repair to specifically introduce a new sequence of interest [114–116]. This method has been rapidly gaining momentum because of its ease of use and efficiency, and is replacing zinc finger nucleases and transcription activator-like effector nucleases as a means of genome editing. Using data generated through chromosome conformation capture, ATAC-Seq and eQTL mapping can help researchers generate an informed hypothesis that can then be tested by producing these mutations using CRISPR/Cas9, making it a highly valuable tool. Fine-mapped GWAS SNPs of interest and regulatory elements could be tested by editing within an appropriate cell type or animal model to experimentally validate its mechanistic effect.

This approach has already been used with success by Soldner *et al.* [117] in 2016 in Parkinson's disease. The authors were able to uncover a risk variant in a non-coding enhancer element that regulates *SNCA* expression by creating multiple SNPs of interest using CRISPR/Cas9, and they then measured expression. It was found that a SNP, rs356168, lies within an enhancer outside of the *SNCA* gene, which is associated with Parkinson's disease. The A allele of this SNP is protective and enables binding of a transcription factor to control the expression of *SNCA*. The G allele, however, prevents this transcription factor from binding and causes upregulation of *SNCA*, leading to disease. This causative SNP was in LD with others, and by using

CRISPR, the investigators were able to determine which SNP was disease relevant. Another study using CRISPR to examine putative causal SNPs for autoimmune diseases also proved successful when investigators related the risk variant, rs6927172, to the expression of *TNFAIP3* and *IL-20RA*, which had not been previously associated with immune disease [118]. Unfortunately, while highly informative, generating and testing variants in this way is currently low throughput.

Genome-wide CRISPR/Cas9 screens are a higher-throughput approach. CRISPR screens involve knocking out many genes or creating mutations in one experiment. They involve designing a library of gRNAs and then transducing cells at a low multiplicity of infection to ensure that one cell will only uptake one gRNA [119]. The cells are then selected using a selection marker or treatment, and sequenced to determine which gRNAs led to the desired phenotype. Such screens can also be applied to characterize the non-coding genome [120]. One example involved characterising non-coding regions that may be influencing gene expression leading to melanoma drug resistance [121]. The investigators designed gRNAs in the regions surrounding three genes already associated with vemurafenib drug resistance in melanoma through a CRISPR/Cas9 knockout screen [122]. Melanoma cells containing a *BRAF* mutation, which is inhibited by vemurafenib, were transduced with these gRNAs and then treated with vemurafenib for selection. The cells resistant to this treatment were mainly enriched for gRNAs surrounding the *CUL3* gene and decreased expression of this gene, leading to the assumption that these are key regulatory areas in gene expression and drug resistance.

Researchers have also started moving beyond conventional CRISPR and modifying Cas9 to do more than just cleave DNA. Simeonov *et al.* [123] implemented a CRISPR activation (CRISPRa) method to uncover novel enhancers that may only be active under certain conditions. Instead of generating a double-stranded break, the Cas9 is fused to VP64, which is used for transcriptional activation. CRISPRa in this experiment identified novel elements within the *IL2RA* super-enhancer responsible for inducible expression of *IL2RA*. The ability to study inducible enhancers without needing to know the context in which they are active can be used in the study of immune disease in which enhancers are often only active when stimulated with molecules such as lipopolysaccharide. In addition to activation, genes can also be repressed through the use of modified Cas9 by methylation or deacetylation of histones at a specific locus of interest [124, 125].

Conclusions

GWAS has been a powerful tool in uncovering genetic risk variants for complex disorders, however refining GWAS variants remains a significant challenge. In this review, we have outlined the techniques being implemented in post-GWAS studies to close the gap between finding a disease-associated variant and identifying the causal variant(s), with particular reference to AS. Fine mapping is an essential step and provides a statistical approximation of those variants most likely to be disease causing. With most of the GWAS variants being located in non-coding regions, understanding the chromatin landscape is important. Techniques such as chromatin conformation capture can help us understand physical interactions occurring in the chromatin, while other approaches such as ATAC-seq establish where regions of open chromatin are located. eQTL data can further suggest regulatory elements by providing a link between the expression of a specific gene and a genetic variant. This

information can then be combined to generate hypotheses about the mechanism of action of variants, which can be experimentally validated in a biological system using CRISPR/Cas9 technology. However, challenges remain in understanding context specificity of variant interactions and knowing what the relevant cell types and conditions are for biological assays to best mimic disease. Using this approach to functionally validate GWAS variants is key to understanding the underlying genetic aetiology of AS and to using genetic data to identify and prioritize drug targets.

Key Points

- AS is a highly heritable multifactorial trait with multiple genomic loci implicated by GWAS.
- Functional genomic approaches in disease-relevant cell and tissue types are needed to generate hypotheses regarding the likely modulated gene(s) and mechanism of action.
- Genome editing approaches such as using CRISPR/Cas9 provide new ways of testing hypotheses by examining the biological effects of GWAS variants in model systems to establish causal relationships.

Funding

This work was supported by Arthritis Research UK (grant number 20773 to J.C.K.); Wellcome Trust Investigator Award (grant number 204969/Z/16/Z to J.C.K.), Wellcome Trust Grant (grant number 090532/Z/09/Z to core facilities Wellcome Centre for Human Genetics) and the NIHR Oxford Biomedical Research Centre.

References

1. Raychaudhuri SP, Deodhar A. The classification and diagnostic criteria of ankylosing spondylitis. *J Autoimmun* 2014; **48–49**:128–33.
2. Stolwijk C, van Tubergen A, Castillo-Ortiz JD, et al. Prevalence of extra-articular manifestations in patients with ankylosing spondylitis: a systematic review and meta-analysis. *Ann Rheum Dis* 2015; **74**(1):65–73.
3. Palla I, Trieste L, Tani C, et al. A systematic literature review of the economic impact of ankylosing spondylitis. *Clin Exp Rheumatol* 2012; **30**(4 Suppl 73):S136–41.
4. Goh L, Samanta A. Update on biologic therapies in ankylosing spondylitis: a literature review. *Int J Rheum Dis* 2012; **15**(5):445–54.
5. Hreggvidsdottir HS, Noordenbos T, Baeten DL. Inflammatory pathways in spondyloarthritis. *Mol Immunol* 2014; **57**(1):28–37.
6. Sherlock JP, Buckley CD, Cua DJ. The critical role of interleukin-23 in spondyloarthropathy. *Mol Immunol* 2014; **57**(1):38–43.
7. Brown MA, Kennedy LG, Macgregor AJ, et al. Susceptibility to ankylosing spondylitis in twins the role of genes, HLA, and the environment. *Arthritis Rheum* 1997; **40**(10):1823–8.
8. Pedersen OB, Svendsen AJ, Ejstrup L, et al. Ankylosing spondylitis in Danish and Norwegian twins: occurrence and the relative importance of genetic vs. environmental effectors in disease causation. *Scand J Rheumatol* 2008; **37**:120–6.
9. Cortes A, Hadler J, Pointon JP, et al. Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci International Genetics of Ankylosing Spondylitis Consortium (IGAS). *Nat Genet* 2013; **45**(7):730–8.
10. Robinson PC, Brown MA. Genetics of ankylosing spondylitis. *Mol Immunol* 2014; **57**(1):2–11.
11. Evans DM, Spencer CCA, Pointon JJ, et al. Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nat Genet* 2011; **43**(8):761–7.
12. Reveille JD, Sims AM, Danoy P, et al. Genome-wide association study of ankylosing spondylitis identifies non-MHC susceptibility loci. *Nat Genet* 2010; **42**(2):123–7.
13. Brown MA, Laval SH, Brophy S, et al. Recurrence risk modelling of the genetic susceptibility to ankylosing spondylitis. *Ann Rheum Dis* 2000; **59**(11):883–6.
14. Schlosstein L, Terasaki PI, Bluestone R, et al. High association of an HL-A antigen, W27, with ankylosing spondylitis. *N Engl J Med* 1973; **288**(14):704–6.
15. Caffrey MFP, James DCO. Human lymphocyte antigen association in ankylosing spondylitis. *Nature* 1973; **242**(5393):121.
16. van der Linden SM, Valkenburg HA, de Jongh BM, et al. The risk of developing ankylosing spondylitis in HLA-B27 positive individuals. A comparison of relatives of spondylitis patients with the general population. *Arthritis Rheum* 1984; **27**:241–9.
17. Burton PR, Clayton DG, Cardon LR; Wellcome Trust Case Control Consortium, Australo-Anglo-American Spondylitis Consortium (TASC), et al. Association scan of 14, 500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat Genet* 2007; **39**(11):1329–37.
18. Ellinghaus D, Jostins L, Spain SL, et al. Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. *Nat Genet* 2016; **48**(5):510–18.
19. Maurano MT, Humbert R, Rynes E, et al. Systematic localization of common disease-associated variation in regulatory DNA. *Science* 2012; **337**(6099):1190–5.
20. Schaub MA, Boyle AP, Kundaje A, et al. Linking disease associations with regulatory information in the human genome. *Genome Res* 2012; **22**(9):1748–59.
21. Ward LD, Kellis M. Interpreting noncoding genetic variation in complex traits and human disease. *Nat Biotechnol* 2012; **30**(11):1095–106.
22. Knight JC. Resolving the variable genome and epigenome in human disease. *J Intern Med* 2012; **271**(4):379–91.
23. Kilpinen H, Waszak SM, Gschwind AR, et al. Coordinated effects of sequence variation on DNA binding, chromatin structure, and transcription. *Science* 2013; **342**(6159):744–7.
24. McVicker G, van de Geijn B, Degner JF, et al. Identification of genetic variants that affect histone modifications in human cells. *Science* 2013; **342**(6159):747–9.
25. Chang SC, Momburg F, Bhutani N, et al. The ER aminopeptidase, ERAP1, trims precursors to lengths of MHC class I peptides by a "molecular ruler" mechanism. *Proc Natl Acad Sci USA* 2005; **102**(47):17107–12.
26. York IA, Chang SC, Saric T, et al. The ER aminopeptidase ERAP1 enhances or limits antigen presentation by trimming epitopes to 8–9 residues. *Nat Immunol* 2002; **3**(12):1177–84.
27. Kochan G, Krojer T, Harvey D, et al. Crystal structures of the endoplasmic reticulum aminopeptidase-1 (ERAP1) reveal the molecular basis for N-terminal peptide trimming. *Proc Natl Acad Sci USA* 2011; **108**(19):7745–50.
28. Roberts AR, Appleton LH, Cortes A, et al. ERAP1 association with ankylosing spondylitis is attributable to common

- genotypes rather than rare haplotype combinations. *Proc Natl Acad Sci USA* 2017;**114**(3):558–61.
29. García-Medel N, Sanz-Bravo A, Van Nguyen D, et al. Functional interaction of the ankylosing spondylitis-associated endoplasmic reticulum aminopeptidase 1 polymorphism and HLA-B27 in vivo. *Mol Cell Proteomics* 2012;**11**(11):1416–29.
 30. Chen L, Fischer R, Peng Y, et al. Critical role of endoplasmic reticulum aminopeptidase 1 in determining the length and sequence of peptides bound and presented by HLA-B27. *Arthritis Rheumatol* 2014;**66**(2):284–94.
 31. Chen L, Ridley A, Hammitzsch A, et al. Silencing or inhibition of endoplasmic reticulum aminopeptidase 1 (ERAP1) suppresses free heavy chain expression and Th17 responses in ankylosing spondylitis. *Ann Rheum Dis* 2016;**75**(5):916–23.
 32. Evnouchidou I, Weimershaus M, Saveanu L, et al. ERAP1-ERAP2 dimerization increases peptide-trimming efficiency. *J Immunol* 2014;**193**(2):901–8.
 33. Martín-Esteban A, Guasp P, Barnea E, et al. Functional interaction of the ankylosing spondylitis-associated endoplasmic reticulum aminopeptidase 2 with the HLA-B*27 peptidome in human cells. *Arthritis Rheumatol* 2016;**68**(10):2466–75.
 34. Evans DM, Spencer CC, Pointon JJ, et al. Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nat Genet* 2011;**43**(8):761–67.
 35. Nair RP, Duffin KC, Helms C, et al. Genome-wide scan reveals association of psoriasis with IL-23 and NF- κ B pathways. *Nat Genet* 2009;**41**(2):199–204.
 36. Barrett JC, Hansoul S, Nicolae DL, et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet* 2008;**40**(8):955–62.
 37. Liu JZ, van Sommeren S, Huang H, et al. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet* 2015;**47**(9):979–86.
 38. McGovern DPB, Gardet A, Törkvist L, et al. Genome-wide association identifies multiple ulcerative colitis susceptibility loci. *Nat Genet* 2010;**42**(4):332–7.
 39. Stuart PE, Nair RP, Tsoi LC, et al. Genome-wide association analysis of psoriatic arthritis and cutaneous psoriasis reveals differences in their genetic architecture. *Am J Hum Genet* 2015;**97**(6):816–36.
 40. Mizuki N, Meguro A, Ota M, et al. Genome-wide association studies identify IL23R-IL12RB2 and IL10 as Behçet's disease susceptibility loci. *Nat Genet* 2010;**42**(8):703–6.
 41. Sherlock JP, Joyce-Shaikh B, Turner SP, et al. IL-23 induces spondyloarthropathy by acting on ROR- γ t+ CD3+CD4–CD8– enthesal resident T cells. *Nat Med* 2012;**18**(7):1069–76.
 42. Corona-Sanchez EG, Muñoz-Valle JF, Gonzalez-Lopez L, et al. –383 A/C tumor necrosis factor receptor 1 polymorphism and ankylosing spondylitis in Mexicans: a preliminary study. *Rheumatol Int* 2012;**32**(8):2565–8.
 43. Davidson SI, Liu Y, Danoy PA, et al. Association of STAT3 and TNFRSF1A with ankylosing spondylitis in Han Chinese. *Ann Rheum Dis* 2011;**70**(2):289–92.
 44. Karaderi T, Pointon JJ, Wordsworth TW, et al. Evidence of genetic association between TNFRSF1A encoding the p55 tumour necrosis factor receptor, and ankylosing spondylitis in UK Caucasians. *Clin Exp Rheumatol* 2011;**30**:110–13.
 45. Pointon JJ, Harvey D, Karaderi T, et al. The chromosome 16q region associated with ankylosing spondylitis includes the candidate gene tumour necrosis factor receptor type 1-associated death domain (TRADD). *Ann Rheum Dis* 2010;**69**(6):1243–6.
 46. Zinovieva E, Bourgain C, Kadi A, et al. Comprehensive linkage and association analyses identify haplotype, near to the TNFSF15 gene, significantly associated with spondyloarthritis, Kerr K, editor. *PLoS Genet* 2009;**5**(6):e1000528.
 47. Braun J, Brandt J, Listing J, et al. Treatment of active ankylosing spondylitis with infliximab: a randomised controlled multicentre trial. *Lancet* 2002;**359**(9313):1187–93.
 48. Sieper J, Appel H, Braun J, et al. Critical appraisal of assessment of structural damage in ankylosing spondylitis: implications for treatment outcomes. *Arthritis Rheum* 2008;**58**(3):649–56.
 49. Lories RJU, Derese I, Luyten FP. Inhibition of osteoclasts does not prevent joint ankylosis in a mouse model of spondyloarthritis. *Rheumatology* 2008;**47**(5):605–8.
 50. Schett G, Stolina M, Dwyer D, et al. Tumor necrosis factor α and RANKL blockade cannot halt bony spur formation in experimental inflammatory arthritis. *Arthritis Rheum* 2009;**60**(9):2644–54.
 51. van Duivenvoorde LM, Dorris ML, Satumtira N, et al. Relationship between inflammation, bone destruction, and osteoproliferation in the HLA-B27/human β_2 -microglobulin-transgenic rat model of spondylarthritis. *Arthritis Rheum* 2012;**64**(10):3210–19.
 52. Howie B, Fuchsberger C, Stephens M, et al. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet* 2012;**44**(8):955–9.
 53. Salomon MP, Li WLS, Edlund CK, et al. GWASseq: targeted resequencing follow up to GWAS. *BMC Genomics* 2016;**17**:176.
 54. Spain SL, Barrett JC. Strategies for fine-mapping complex traits. *Hum Mol Genet* 2015;**24**(R1):R111–19.
 55. Dunham I, Kundaje A, Aldred SF, et al. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012;**489**(7414):57–74.
 56. Romanoski CE, Glass CK, Stunnenberg HG, et al. Epigenomics: roadmap for regulation. *Nature* 2015;**518**(7539):314–16.
 57. Onengut-Gumuscu S, Chen WM, Burren O, et al. Fine mapping of type 1 diabetes susceptibility loci and evidence for colocalization of causal variants with lymphoid gene enhancers. *Nat Genet* 2015;**47**(4):381–6.
 58. Corradin O, Saiakhova A, Akhtar-Zaidi B, et al. Combinatorial effects of multiple enhancer variants in linkage disequilibrium dictate levels of gene expression to confer susceptibility to common traits. *Genome Res* 2014;**24**(1):1–13.
 59. Rivera CM, Ren B. Mapping human epigenomes. *Cell* 2013;**155**(1):39–55.
 60. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* 2003;**33**(3s):245–54.
 61. Agelopoulos M, McKay DJ, Mann RS. Developmental regulation of chromatin conformation by hox proteins in *Drosophila*. *Cell Rep* 2012;**1**(4):350–9.
 62. Lelli KM, Slattery M, Mann RS. Disentangling the many layers of eukaryotic transcriptional regulation. *Annu Rev Genet* 2012;**46**(1):43–68.
 63. Tsompana M, Buck MJ. Chromatin accessibility: a window into the genome. *Epigenetics Chromatin* 2014;**7**(1):33.
 64. Weintraub H, Groudine M. Chromosomal subunits in active genes have an altered conformation. *Science* 1976;**193**(4256):848–56.

65. Keene MA, Corces V, Lowenhaupt K, et al. DNase I hypersensitive sites in Drosophila chromatin occur at the 5' ends of regions of transcription. *Proc Natl Acad Sci USA* 1981;**78**(1):143–6.
66. Crawford GE, Holt IE, Whittle J, et al. Genome-wide mapping of DNase hypersensitive sites using massively parallel signature sequencing (MPSS). *Genome Res* 2005;**16**(1):123–31.
67. Buenrostro JD, Giresi PG, Zaba LC, et al. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* 2013;**10**(12):1213–18.
68. Forrest MP, Zhang H, Moy W, et al. Open chromatin profiling in hiPSC-derived neurons prioritizes functional noncoding psychiatric risk variants and highlights neurodevelopmental loci. *Cell Stem Cell* 2017;**21**(3):305–18.e8.
69. Laurila K, Lähdesmäki H. Systematic analysis of disease-related regulatory mutation classes reveals distinct effects on transcription factor binding. *In Silico Biol* 2009;**9**:209–24.
70. El-Maarri O. DNA methylation and human disease. *Nat Rev Genet* 2003;**5**:44–54.
71. Hendrich B, Bickmore W. Human diseases with underlying defects in chromatin structure and modification. *Hum Mol Genet* 2001;**10**(20):2233–42.
72. De Gobbi M, Viprakasit V, Hughes JR, et al. A regulatory SNP causes a human genetic disease by creating a new transcriptional promoter. *Science* 2006;**312**(5777):1215–17.
73. Tournamille C, Colin Y, Cartron JP, et al. Disruption of a GATA motif in the Duffy gene promoter abolishes erythroid gene expression in Duffy-negative individuals. *Nat Genet* 1995;**10**(2):224–8.
74. Petrij F, Giles RH, Dauwerse HG, et al. Rubinstein-Taybi syndrome caused by mutations in the transcriptional coactivator CBP. *Nature* 1995;**376**(6538):348–51.
75. Ueda H, Howson JMM, Esposito L, et al. Association of the T-cell regulatory gene CTLA4 with susceptibility to autoimmune disease. *Nature* 2003;**423**:506–11.
76. Capon F, Allen MH, Ameen M, et al. A synonymous SNP of the corneodesmosin gene leads to increased mRNA stability and demonstrates association with psoriasis across diverse ethnic groups. *Hum Mol Genet* 2004;**13**(20):2361–8.
77. Dixon JR, Selvaraj S, Yue F, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 2012;**485**(7398):376–80.
78. Cubeñas-Potts C, Corces VG. Architectural proteins, transcription, and the three-dimensional organization of the genome. *FEBS Lett* 2015;**589**(20 Pt A):2923–30.
79. Phillips-Cremins JE, Sauria MEG, Sanyal A, et al. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* 2013;**153**(6):1281–95.
80. Nora EP, Lajoie BR, Schulz EG, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 2012;**485**(7398):381–5.
81. Dekker J, Rippe K, Dekker M, et al. Capturing chromosome conformation. *Science* 2002;**295**(5558):1306–11.
82. Belton JM, McCord RP, Gibcus JH, et al. Hi-C: a comprehensive technique to capture the conformation of genomes. *Methods* 2012;**58**(3):268–76.
83. Lieberman-Aiden E, van Berkum NL, Williams L, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 2009;**326**(5950):289–93.
84. Hughes JR, Roberts N, McGowan S, et al. Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. *Nat Genet* 2014;**46**:205–12.
85. Smemo S, Tena JJ, Kim KH, et al. Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature* 2014;**507**(7492):371–5.
86. Dina C, Meyre D, Gallina S, et al. Variation in FTO contributes to childhood obesity and severe adult obesity. *Nat Genet* 2007;**39**(6):724–6.
87. Scuteri A, Sanna S, Chen WM, et al. Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genet* 2007;**3**(7):e115.
88. Frayling TM, Timpson NJ, Weedon MN, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 2007;**316**(5826):889–94.
89. Fischer J, Koch L, Emmerling C, et al. Inactivation of the Fto gene protects from obesity. *Nature* 2009;**458**(7240):894–8.
90. Church C, Moir L, McMurray F, et al. Overexpression of Fto leads to increased food intake and results in obesity. *Nat Genet* 2010;**42**(12):1086–92.
91. Dimas AS, Deutsch S, Stranger BE, et al. Common regulatory variation impacts gene expression in a cell type-dependent manner. *Science* 2009;**325**(5945):1246–50.
92. Nica AC, Parts L, Glass D, et al. The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. *PLoS Genet* 2011;**7**(2):e1002003.
93. Fairfax BP, Makino S, Radhakrishnan J, et al. Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nat Genet* 2012;**44**(5):502–10.
94. Fairfax BP, Humburg P, Makino S, et al. Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. *Science* 2014;**343**(6175):1246949.
95. Lee MN, Ye C, Villani AC, et al. Common genetic variants modulate pathogen-sensing responses in human dendritic cells. *Science* 2014;**343**(6175):1246980.
96. Peters JE, Lyons PA, Lee JC, et al. Insight into genotype-phenotype associations through eQTL mapping in multiple cell types in health and immune-mediated disease. Plagnol V, editor. *PLoS Genet* 2016;**12**(3):e1005908.
97. Lonsdale J, Thomas J, Salvatore M, et al. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013;**45**(6):580–85.
98. GTEx Consortium. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 2015;**348**:648–60.
99. Zambrano-Zaragoza JF, Agraz-Cibrian JM, González-Reyes C, et al. Ankylosing spondylitis: from cells to genes. *Int J Inflamm* 2013;**2013**:501653.
100. Smith JA, Märker-Hermann E, Colbert RA. Pathogenesis of ankylosing spondylitis: current concepts. *Best Pract Res Clin Rheumatol* 2006;**20**(3):571–91.
101. Chan AT, Kollnberger SD, Wedderburn LR, et al. Expansion and enhanced survival of natural killer cells expressing the killer immunoglobulin-like receptor KIR3DL2 in spondylarthritis. *Arthritis Rheum* 2005;**52**(11):3586–95.
102. Miossec P, Kolls JK. Targeting IL-17 and TH17 cells in chronic inflammation. *Nat Rev Drug Discov* 2012;**11**(10):763–76.
103. Noordenbos T, Yeremenko N, Gofita I, et al. Interleukin-17-positive mast cells contribute to synovial inflammation in spondylarthritis. *Arthritis Rheum* 2012;**64**(1):99–109.
104. Limon-Camacho L, Vargas-Rojas MI, Vazquez-Mellado J, et al. In vivo peripheral blood proinflammatory T cells in

- patients with ankylosing spondylitis. *J Rheumatol* 2012;**39**(4): 830–5.
105. Xueyi L, Lina C, Zhenbiao W, et al. Levels of circulating Th17 cells and regulatory T cells in ankylosing spondylitis patients with an inadequate response to anti-TNF- α therapy. *J Clin Immunol* 2013;**33**(1):151–61.
106. Farh KK-H, Marson A, Zhu J, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* 2015;**518**(7539):337–43.
107. Wright C, Edelmann M, DiGleria K, et al. Ankylosing spondylitis monocytes show upregulation of proteins involved in inflammation and the ubiquitin proteasome pathway. *Ann Rheum Dis* 2009;**68**(10):1626–32.
108. Chen M, Zhang L, Ren Y, et al. Defective function of CD24(+)CD38(+) regulatory B cells in ankylosing spondylitis. *DNA Cell Biol* 2016;**35**(2):88–95.
109. Mauri C, Gray D, Mushtaq N, et al. Prevention of arthritis by interleukin 10-producing B cells. *J Exp Med* 2003;**197**(4): 489–501.
110. Miossec P, Kolls JK. Targeting IL-17 and TH17 cells in chronic inflammation. *Nat Rev Drug Discov* 2012;**11**(10):763–76.
111. Gawad C, Koh W, Quake SR. Single-cell genome sequencing: current state of the science. *Nat Rev Genet* 2016;**17**(3):175–88.
112. Buenrostro JD, Wu B, Litzenburger UM, et al. Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* 2015;**523**(7561):486–90.
113. Rotem A, Ram O, Shores N, et al. Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nat Biotechnol* 2015;**33**(11):1165–72.
114. Gasiunas G, Barrangou R, Horvath P, et al. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci USA* 2012;**109**(39):E2579–86.
115. Mali P, Yang L, Esvelt KM, et al. RNA-guided human genome engineering via Cas9. *Science* 2013;**339**(6121):823–6.
116. Cong L, Ran FA, Cox D, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* 2013;**339**(6121): 819–23.
117. Soldner F, Stelzer Y, Shivalila CS, et al. Parkinson-associated risk variant in distal enhancer of α -synuclein modulates target gene expression. *Nature* 2016;**533**(7601):95–9.
118. Wu J, Yang S, Yu D, et al. CRISPR/cas9 mediated knockout of an intergenic variant rs6927172 identified IL-20RA as a new risk gene for multiple autoimmune diseases. *Genes Immun* 2018. doi: 10.1038/s41435-018-0011-6.
119. Wang T, Wei JJ, Sabatini DM, et al. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* 2014;**343**(6166): 80–4.
120. Montalbano A, Canver MC, Sanjana NE. High-throughput approaches to pinpoint function within the noncoding genome. *Mol Cell* 2017;**68**(1):44–59.
121. Sanjana NE, Wright J, Zheng K, et al. High-resolution interrogation of functional elements in the noncoding genome. *Science* 2016;**353**(6307):1545–9.
122. Shalem O, Sanjana NE, Hartenian E, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 2014;**343**(6166):84–7.
123. Simeonov DR, Gowen BG, Boontanart M, et al. Discovery of stimulation-responsive immune enhancers with CRISPR activation. *Nature* 2017;**549**(7670):111–15.
124. Lei Y, Zhang X, Su J, et al. Targeted DNA methylation in vivo using an engineered dCas9-MQ1 fusion protein. *Nat Commun* 2017;**8**:16026.
125. Kwon DY, Zhao YT, Lamonica JM, et al. Locus-specific histone deacetylation using a synthetic CRISPR-Cas9-based HDAC. *Nat Commun* 2017;**8**:15315.