REVIEW

Fine mapping with epigenetic information and 3D structure

Gisela Orozco[1,2] 🔘

Abstract

Since 2005, thousands of genome-wide association studies (GWAS) have been published, identifying hundreds of thousands of genetic variants that increase risk of complex traits such as autoimmune diseases. This wealth of data has the potential to improve patient care, through personalized medicine and the identification of novel drug targets. However, the potential of GWAS for clinical translation has not been fully achieved yet, due to the fact that the functional interpretation of risk variants and the identification of causal variants and genes are challenging. The past decade has seen the development of great advances that are facilitating the overcoming of these limitations, by utilizing a plethora of genomics and epigenomics tools to map and characterize regulatory elements and chromatin interactions, which can be used to fine map GWAS loci, and advance our understanding of the biological mechanisms that cause disease.

Keywords Autoimmune disease · Genome-wide association studies · Fine mapping · Functional genomics · Epigenetics · Chromatin conformation

Introduction

Complex diseases arise from a combination of genetic and environmental factors. Genome-wide association studies (GWAS) have identified thousands of genetic regions, or loci, that contain single nucleotide polymorphisms (SNPs) associated with complex diseases [1]. For autoimmune diseases, multiple, large-scale genetic association studies have been carried out, including GWAS, fine mapping studies, follow-up studies and meta-analyses, contributing to the identification of a large proportion of the heritability of these diseases.

✉ Gisela Orozco
gisela.orozco@manchester.ac.uk

1 Centre for Genetics and Genomics Versus Arthritis, Division of Musculoskeletal and Dermatological Sciences, School of Biological Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, AV Hill Building, Oxford Road, Manchester M13 9LJ, UK

2 NIHR Manchester Biomedical Research Centre, Manchester University NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK

The ultimate goal of genetic studies is to improve human health, by enabling a better understanding of the biological mechanisms that lead to disease. This, in turn, can help identify novel therapeutic targets for autoimmune diseases for which there is a lack of specific treatments, or where a large proportion of patients do not respond appropriately to available treatments, such as psoriatic arthritis (PsA), Crohn's disease, ankylosing spondylitis (AS) or rheumatoid arthritis (RA), among many others. Indeed, several studies have shown that drug development programs with incidental genomic support had a higher rate of developmental success [2–5]. In this regard, a coding variant in *TYK2*, which has a protective effect in multiple sclerosis (MS), AS, ulcerative colitis (UC) and Crohn's disease, leads to a weaker response of CD4 + T helper 1 and 17 type cells to proinflammatory signals [6], which suggested that reducing the activity of TYK2 could be a potential therapeutic option for these diseases. There are now multiple clinical trials testing TYK2 inhibitors in several autoimmune diseases [7].

Furthermore, biological mechanisms revealed by GWAS can suggest compounds suitable for repurposing, which involves the use in a new disease of a drug approved for treatment of a different condition and allows cheaper and quicker translation implementation of treatments into the clinic compared to de novo drug discovery [8]. For example, the monoclonal antibodies ustekinumab and risankizumab,

first used to treat psoriasis, target components of the interleukin-23 (IL-23) signalling pathway. One of the most robust GWAS signals in Crohn's disease is a missense mutation on the IL-23 receptor (*IL23R*) gene, which inspired the exploration of the above-mentioned treatments in Crohn's disease [9]. These biologics have now been proven to be successful in the treatment of Crohn's and in other diseases that also present genetic associations with genes in the IL-23 pathway such as PsA and AS [10–14].

Genetics also have the potential to aid precision medicine approaches to target the best treatments available to specific groups of patients who are more likely to respond appropriately or avoid particular treatments in patients who are predicted to present adverse reactions [15]. For example, certain HLA alleles have been shown to be correlated with treatment response in psoriasis [16] and rheumatoid arthritis (RA) [17]. In addition, genetic risk scores (GRS) can help identifying those at highest risk of developing disease. Although they are less powerful in predicting traits with low prevalence, disclosure of high genetic risk to individuals in the clinic can lead to behavioural changes to minimize exposure to environmental risk factors [15]. The ongoing investment in long-term prospective cohort studies that include genetic data combined with detailed electronic health records, such as the UK Biobank [18], will help advance the field of precision medicine [19].

However, despite these exciting and promising advances, the full potential of GWAS for clinical translation has not been achieved yet, due to several caveats that make the interpretation of GWAS variants difficult [20]. First, each individual GWAS-associated variant acts as a signpost or "tag" for a haplotype containing multiple neighbouring SNPs that are inherited together in a block, i.e. they are in high linkage disequilibrium (LD). Therefore, GWAS alone cannot distinguish the causal variants underpinning the association from that of the other variants in LD with them. Second, over 90% of GWAS variants map to non-coding regions of the genome, and therefore, they do not directly affect the coding sequence of a gene, making it challenging to identify the biological mechanism by which they cause disease [21, 22]. Third, in many cases, it is unclear which are the causal genes, since GWAS loci often contain multiple genes in the vicinity or map at very large distances from coding genes. And finally, we lack a complete understanding of the context, i.e. the cell types and stimuli, under which disease associated variants have an effect.

Fine mapping using statistical methods, for example, Bayesian approaches, credible SNP sets and trans-ethnic meta-analyses, have been successful in prioritizing genetic variants for further study [23, 24], but to overcome the limitations listed above and fully realize the potential of GWAS to understand disease biology, follow-up functional studies are needed. In this review, different functional genomics approaches to fine map GWAS loci and prioritize causal genes in autoimmune diseases will be discussed, with examples illustrating how these methods have been used to translate genetic findings into functional understanding of etiopathological mechanisms of autoimmune disease.

Functional genomics

The non-coding portion of the human genome is thought to play a pivotal role in the highly complex process that is the regulation of gene expression, which is mediated by regulatory elements. Regulatory elements are short, non-coding functional DNA sequences that can regulate transcription of their target genes and contain binding sites for regulatory proteins. There are several classes of regulatory elements, such as enhancers, promoters, insulators or silencers, that are characterized by slightly different features, such as specific histone modifications or binding or particular transcription factors [25, 26]. As above mentioned, the vast majority of GWAS-associated variants maps to non-coding regions of the genome, and they seem to be predominantly enriched in cell type-specific enhancers [27].

The mapping and characterization of regulatory elements such as enhancers using functional genomics has therefore been used extensively to functionally interpret GWAS signals across all complex traits [28, 29] (Fig. 1). Functional genomics involves the use of genomic data to investigate gene expression and function at genome-wide scale, usually by applying high-throughput methods (Table 1). Regulatory elements generally map to regions of open chromatin, are flanked by nucleosomes with histone modifications that are associated with active transcription, are bound by specific transcription factors and are likely to physically come into contact with their target genes through chromatin looping [28]. Major large-scale international projects, like Encode [30], Roadmap [31] or the IHEC consortium [32], have used techniques to detect such features at the genome-wide scale (like ATAC-Seq, ChIP-Seq and Hi-C, Table 1) in a wide range of cell and tissue types, which has become an invaluable resource in the post-GWAS era.

Fine mapping using epigenetic information and colocalization

Understanding the mechanism by which SNPs lead to changes in molecular phenotypes in disease relevant tissues is required for the identification of causal variants in risk loci identified by GWAS. The majority of GWAS SNPs are thought to disrupt regulatory elements, and therefore, these mechanisms most commonly involve alteration of gene expression levels, disruption of transcription factor binding
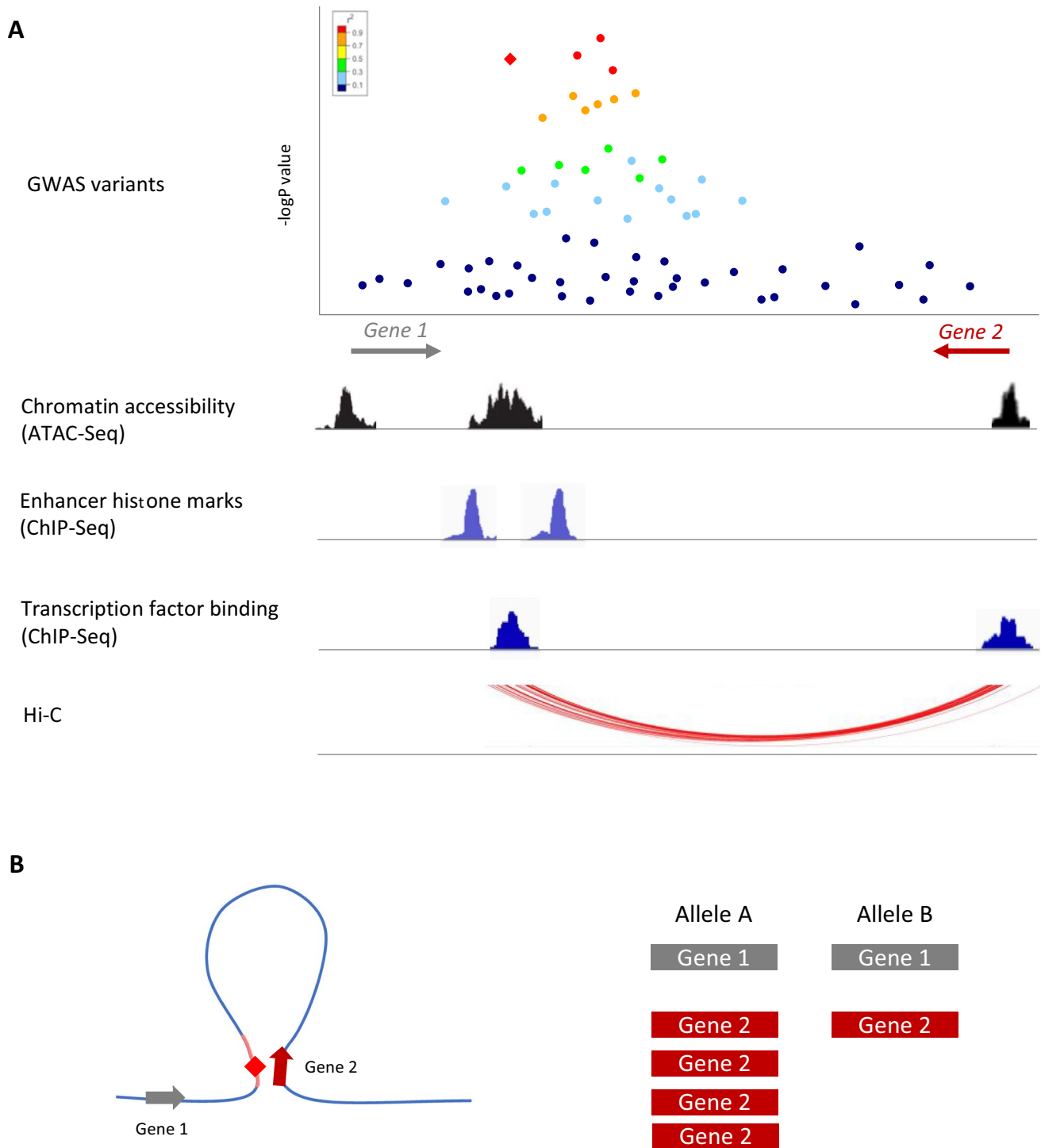
Fig. 1 Schematic representation of how functional genomics data can be used to functionally interpret GWAS loci. **A** Fine mapping using epigenetic information and colocalization: amongst the most strongly associated SNP within a disease GWAS locus (red dots), only one (red diamond) overlaps an active enhancer, suggesting that this variant may be the causal SNP. **B** The disease enhancer containing the risk variant, although closer in the linear conformation to *Gene 1*, interacts with *Gene 2* through chromatin looping. In addition, the disease SNP is an eQTL for *Gene 2* but has no effect on the expression of *Gene 1*, suggesting that *Gene 2* may be the causal gene at this locus

Table 1 Summary of the genomics, epigenomics and transcriptomics techniques most commonly used to interpret GWAS signals

Type	Technique	Description
Transcription	RNA-seq, GRO-cap, CAGE [33, 34]	Whole-transcriptome RNA sequencing can identify transcription of active enhancers (enhancer RNA or eRNAs)
Transcription	eQTLs [35]	Correlation between genetic variation and levels of gene expression
Chromatin accessibility	MNase-seq, DNase-seq, ATAC-seq [36–38]	High-throughput detection of open chromatin by micrococcal nuclease digestion, cleavage by DNase I or library construction using the hyperactive transposase Tn5, respectively, followed by next-generation sequencing (NGS)
Histone marks	ChIP-seq, Cut&Run [39, 40]	Detection of post-translational histone modifications by immuno-precipitation with specific antibodies, e.g. enhancer or promoter-associated histone modifications such as H3K4me1 or H3K27ac, followed by NGS
Protein binding	ChIP-seq, Cut&Run [39, 40]	Detection of DNA bound regulatory proteins and transcription factors by immunoprecipitation with specific antibodies, e.g. RNA polymerase II or NFκB, followed by NGS
3D proximity	Chromatin conformation capture (3C) methods, i.e. Hi-C, Capture Hi-C (CHi-C), HiChIP [41–43]	Family of methods to detect looping and spatial organization of DNA. The chromatin is digested with enzymes, and then interacting regions are re-ligated together. The resulting products are sequenced and analysed to quantify the frequency of interactions

and perturbations of enhancer activity, in the form of alterations in chromatin accessibility and histone modifications [44]. Therefore, numerous functional fine-mapping methods are based on the colocalization of GWAS variants in high LD with functional annotations such as promoters and enhancers using data from large-scale projects like Encode or Roadmap, described above. If a variant, or variants, within a locus overlaps active functional elements in tissues of interest, they will be more likely to be causal than variants in LD that do not [45] (Fig. 1). It is important to note that although many fine-mapping methods assume that only one SNP per locus contributes to disease, it has been shown that multiple variants within a single locus can play a role through additive or epistatic effects, which can context cell type specific [46].

Applying this colocalization approach, Trynka et al. examined 15 epigenetic chromatin marks using ChIP-seq data from 14 immune cell types from the ENCODE project [47]. After overlapping with all associated SNPs for 31 complex traits, they found that H3K4me3 was the most cell type-specific chromatin mark. Indeed, there was an enrichment of risk SNPs in peaks for this histone mark in cell types that are keys for disease biology. For example, in RA, GWAS SNPs were enriched in H3K4me3 peaks in CD4 + Treg cells. Interestingly, when the analysis was repeated using newly defined index SNPs from the dense genotyping fine-mapping study Immunochip [48], instead of index SNPs from the largest RA GWAS at that time [49], the significance of the enrichment for CD4 + Treg cells increased. In addition, an RA-associated locus tagged by rs13119723, an SNP mapping to an intron of the gene of unknown function *KIAA1109* and containing multiple variants in LD spanning over 500 Kb,

was fine mapped to a single SNP mapping to a H3K4me3 peak specific to CD4 + Treg cells thought to be involved in the regulation of *IL2* expression. These results demonstrate how colocalization approaches of GWAS and functional annotations can inform both the most relevant cell types implicated in disease and the fine mapping of associated SNPs to identify causal variation.

In another seminal study, Farh et al. aimed at fine-mapping GWAS loci from 21 autoimmune diseases by developing an algorithm, Probabilistic Identification of Causal SNPs (PICS), that estimates the probability that an individual SNP is a causal variant given the haplotype structure and observed pattern of association at the locus, followed by integration of transcription and cis-regulatory element annotations from the NIH Roadmap and similar data [22]. They estimated that 60% of causal variants map to immune enhancers, but only 10–20% directly alter recognizable transcription factor binding motifs. Multiple methods have been developed recently to enhance statistical fine mapping with functional annotations, significantly improving the identification of causal variants. For example, Weissbrod et al. developed a functionally informed method, PolyFun, that specifies prior causal probabilities for subsequent fine-mapping methods, providing > 20% power increase over nonfunctionally informed fine-mapping methods [50].

A number of other studies have also showed that overlap of disease variants with functional annotations can identify tissue-specific enrichments [51, 52]. In addition to cell type specificity, disease variants can have an effect on phenotype in specific cell states or under specific stimulatory conditions. Soskic et al. stimulated T cells, and macrophages with an array of 13 different cytokines found that immune

disease-associated variants are enriched in chromatin regions that are open and active in early rather than late activation of memory CD4[+] T cells [53].

The most complete human epigenome reference to date, EpiMap (epigenome integration across multiple annotation projects), was recently assembled, by compiling 10,000 epigenomic maps from the main large-scale consortia studies such as ENCODE, Roadmap and IHEC, including ChIP-Seq data for numerous histone marks and transcription factors, open chromatin data from DNAse-seq and ATAC-seq experiments, across 800 samples from multiple cell types and tissues [54]. These datasets were used to define chromatin states, enhancers and target genes, which were then used to annotate 30,000 genetic loci associated with 540 traits from the GWAS catalogue, predicting key disease tissues, causal SNPs enriched in enhancers and candidate target genes for each locus.

Transcription factor binding profiling has also been used to functionally interpret non-coding variation. IMPACT is a recently developed genome annotation method that identifies regulatory elements defined by cell-state-specific TF binding profiles, learned from 515 chromatin and sequence annotations [55]. When integrated with RA GWAS data, it was found that the top 5% of CD4 + Treg regulatory elements identified by IMPACT capture 85.7% of RA genetic heritability, outperforming methods that ignore differential functionality of effector cell states. This further strengthens the importance of selecting the right tissues and cell types in functional studies. This annotation method has also been proven to be useful in the identification of causal variants that are common for different populations, despite the presence of different LD patterns at associated loci, which improves the trans-ancestry portability of polygenic risk scores [56].

On the other hand, high-throughput protein-DNA binding assays, rather than predictions, have been used to elucidate the biological function of non-coding variation. Resources such as the systematic assessment of the binding of 270 human transcription factors to 95,886 non-coding variants in the human genome using one of such methods, SNP-SELEX, further facilitates the understanding of the pathways involved in disease [57]. When applied to T2D, this study found that SNPs that demonstrate differential binding of TFs were highly enriched in the set of SNPs that had been previously reported as likely causal. The T2D-associated SNP rs7578326, which overlaps a candidate enhancer and had been linked to the *IRS1* gene by long-range chromatin interactions in HepG2 cells, was found to affect binding of the liver-specific transcription factor CEBPB. Using CRISPR interference, the enhancer was silenced in HepG2 cells, which resulted in significantly reduced expression of *IRS1*.

In addition to enabling the identification of regulatory elements across the genome, the outputs of high-throughput

functional genomics methods described above can be utilised to assess whether genetic variants present allele-specific effects. The allele-specific abundance of sequencing reads can help us generate hypothesis about the biological function of disease variants (quantitative trait loci mapping, QTL). This type of study can profile traits such as chromatin accessibility (caQTL), histone marks (hQTL), exon splicing (sQTL) gene expression (eQTL) or protein expression (pQTL).

Several studies have shown that many caQTLs overlap TF binding sites and motifs, and a subset of them colocalize with eQTLs and GWAS variants, which suggests that SNPs that map to these loci influence GWAS traits by altering chromatin accessibility [58–62]. Likewise, the presence of hQTLs that affect enhancer-associated histone ChIP-Seq peaks like H3K27ac or H3K4me1 further suggests that these variants have an effect on cell type-specific enhancer activity [52]. Pelikan et al. identified numerous hQTLs in lymphoblastoid cell lines derived from systemic lupus erythematosus (SLE) patients, which were enriched in autoimmune disease risk haplotypes and influenced gene expression variability compared with non-hQTL variants in strong LD, suggesting that this type of data can be used in fine-mapping efforts [63].

Gene prioritization using expression quantitative trait loci (eQTLs)

The examples described in the previous section illustrate how functional genomics annotations have identified thousands of enhancers that are disrupted by GWAS variants. However, only a handful of these have been confidently linked to their target genes, limiting the biological knowledge of the pathways that are altered in disease that we can extrapolate from these studies. Due to the robustness of RNA-sequencing technologies, eQTL studies represent a common approach to prioritize potential causal genes at risk loci.

Several methods have been used to integrate eQTL maps with GWAS data, from just assessing whether GWAS variants were also significant eQTLs, to more sophisticated colocalization statistical analyses [29]. Early studies that overlapped the most strongly associated GWAS variants at a given locus with the top eQTL within the same variants described how risk SNPs are more likely to be eQTLs than random non-associated SNPs [64, 65]. Examples of how this type of study has been used to pinpoint disease genes include the analysis conducted by Westra et al. where an SLE-associated SNP was also found to be an eQTL for the TF *IKF1*. This eQTL also affected expression of a number of other genes that are in turn regulated by IKF1 [66].

More accurate statistical colocalization methods were later developed, to account for the fact that the large amount of eQTLs present in the genome can result in a high degree of false positives due to chance, multiple genes can be affected by the same variant (pleiotropy), and that the top eQTL SNP in a locus may not always be the same or in tight LD with the top GWAS SNP [67–70]. It is also important to know that eQTLs can be context specific and can change with stimulation with pathogens or IFN-γ, for example [71–73].

Large-scale projects that characterize the genetic factors underlying gene expression across tissues and cell types provide a highly valuable resource. The Genotype-Tissue Expression (GTEx) consortium represents one of the most comprehensive eQTL resources [74]. GTEx's approach is to perform both DNA sequencing and multi-tissue RNA-seq across many post-mortem donor samples, to identify genetic variants that are correlated with the expression levels of genes across the whole genome. The third and final phase of the project was recently completed, with a dataset of 838 donors and 15,201 samples across 49 tissues, identifying more than 4.2 million eQTLs which were enriched in disease-associated variants from the GWAS catalogue [75].

Another noteworthy large-scale eQTL mapping effort relevant for autoimmune diseases is eQTLGen, which includes whole blood-derived expression from 31,684 individuals [76]. One strength of this study is the capability to identify trans-effects, which are harder to detect than cis-eQTLs. Forty-seven GWAS traits for which at least four independent variants affected the same gene in trans were identified; in SLE, 13 genes in the SLE interferon signature were affected by at least three SLE-associated genetic variants each.

GTEx and eQTLGen data are generated from bulk tissue samples, and given the cell type specificity of the regulation of gene expression, their sample heterogeneity may confound mechanistic follow-up of GWAS loci. Kim-Hellmuth et al. have developed a computational deconvolution method to estimate the proportion of different cell types in GTEx tissues, so then cell type information can be accounted for in eQTL analyses [77]. But, ideally, disease relevant cell types, such as individual immune cell populations isolated from blood, can be more informative for the colocalization of eQTLs with GWAS variants in autoimmunity. The DICE study studied 13 different immune cell types from 91 donors and discovered that a large fraction (41%) of the 12,254 genes for which eQTLs were identified showed a strong *cis*-association with genotype only in a single cell type [78]. Of note, the ImmuNexU study identified immune cell type- and disease-specific eQTLs, by studying 28 distinct immune cell subsets from patients with 10 different immune-mediated diseases and healthy individuals in Japanese population. Most large-scale eQTL studies such as GTEx, eQTLGen or DICE have been generated in European donors; the ImmuNexU study revealed eQTL variants in East Asian populations that had not been detected in European populations, highlighting the need for ancestry-matched eQTL databases for GWAS fine mapping and improvement of functional annotation [79].

There are many other eQTL studies that have made their datasets publicly available, which represents an incredibly valuable resource. Recently, the eQTL catalogue provided uniformly processed gene expression and splicing QTLs from all available public studies, facilitating the widespread use of these resources [80].

Gene prioritization using chromatin 3D structure

Enhancers and their target genes, which might be distant from one another in the linear DNA primary structure, are brought close together by chromatin loops. Therefore, the genes affected by GWAS variants that disrupt regulatory elements can be identified by characterizing 3D chromatin structure in relevant cell types using methods like Hi-C [81, 82]. Hi-C is based on chromosome conformation capture (3C), a method to detect looping and spatial organization of DNA. The chromatin is digested with enzymes, and interacting regions are re-ligated together. The resulting products are sequenced and analysed to quantify the frequency of interactions (Table 1). In addition to chromatin loops, Hi-C data can be used to map higher order chromatin structures, such as active/inactive A/B compartments and topologically associating domains (TADs) representing large domains that display a marked regulatory potential, that can help us understand the chromatin context of disease associated loci.

It has been shown that abnormal 3D chromatin organization can lead to disease by rewiring interactions between genes and regulatory elements [83] and that there is a direct relationship between the strength of chromatin interaction, DNA activity, and gene expression, where small changes in interaction intensity mediated by disease-associated regulatory genetic variation lead to large functional effects in gene expression [84, 85]. In addition, recent evidence has demonstrated that SNPs can influence chromatin interaction strength, along with accessibility and gene expression [86] and, interestingly, that T1D risk haplotypes demonstrate a greater level of interactions and gene activity in mice [87].

In a seminar paper, Javierre et al. applied a technique based on Hi-C that incorporates a sequence capture step to enrich for interactions that involve specific regions of interest, capture Hi-C (CHi-C), to identify interacting regions of all known promoters in the human genome, in 17 human primary hematopoietic cell types [88]. They found that chromatin interactions between promoters and promoter-interacting regions (PIRs) are highly cell type specific.

Interestingly, PIRs were found to be significantly enriched in active chromatin regions and eQTLs. Similarly, PIRs were also enriched in GWAS variants associated with autoimmune diseases in lymphoid cells, and the data was subsequently used to identify putative causal genes in these traits. Of note, multiple genes such as *RP4-753F5.1*, *CD101*, *TTF2* and *TRIM45* were prioritized as RA candidate genes at the 1p13.1 GWAS locus, which had been initially assigned to *CD2* due to proximity.

Alternatively, Hi-C libraries can be enriched by immunoprecipitation using an antibody targeting a protein or histone modification of interest, in a method called HiChIP. Recently, Chandra et al.used H3K27ac HiChIP to map interactions between eQTLs overlapping active *cis*-regulatory elements and their target genes in five immune cell types, increasing our understanding the mechanism by which disease risk variants exert their effects on gene expression, in a cell-specific manner [89].

Other studies have successfully used a similar approach to link associated variants to their target genes in autoimmune diseases. A study targeting all GWAS loci for RA, T1D, PsA and JIA (as opposed to capturing gene promoters) in T and B cell lines also showed that chromatin interactions can be cell type specific [90]. Multiple putative causal genes were identified, and it was found that some disease-associated SNPs do not interact with the nearest gene but with other candidate genes that can be situated several megabases away and that regions associated with different autoimmune diseases can interact with each other and the same promoter. In addition, this data proved to be useful in the identification of drugs that could be repositioned to treat rheumatic diseases [91].

Other studies have highlighted the importance of non-immune cell types in autoimmune diseases. For example, stromal cells, such as synovial fibroblasts (FLS), play an important role in RA, but functional genomics datasets in this cell type were lacking. A recent study aimed at characterizing DNA architecture, 3D chromatin interactions, DNA accessibility and gene expression in FLS samples from RA patients; it was found that FLS account for up to 24% of RA heritability and 10% of the RA-associated regions contain SNPs located within enhancers that are exclusive to FLS, which suggests an additional, independent role of FLS in driving the genetic risk of developing RA [92]. The same study found that TNF stimulation of FLS alters the organization of topologically associating domains, chromatin state and the expression of putative RA causal genes such as *TNFAIP3* and *IFNAR1*. Therefore, future functional studies in non-immune cell types that contribute to the pathogenesis of autoimmune diseases will likely advance fine mapping of GWAS loci.

These studies and many others [93, 94] exemplify how 3D chromatin structure can aid the functional interrogation of GWAS loci. However, Hi-C methods often lack resolution to characterize individual enhancer-promoter interactions at sufficient depth since interaction maps are constrained by the restriction enzyme used, typically *HindIII*. This can be only partially solved with the use of multiple restriction enzymes that digest the genome into smaller fragments, like the strategy used in commercially available kits (Arima Genomics). A new analysis method, based on a Bayesian sparse variable selection approach, was developed to fine map chromatin contacts [95]. On the other hand, exciting novel technological advances, such as the development of a new 3C technique called Micro-Capture-C, are starting to pave the way to increase resolution of chromatin interactions at the base-pair level [96]. Although this method cannot be applied at the genome wide level yet, it provides an unprecedented level of detail about how gene expression is regulated by enhancers containing disease-associated variants and will facilitate fine mapping of GWAS loci.

Conclusions

In the past decade, advances in large-scale methods for the functional characterization of autoimmunity risk loci have enabled a deeper understanding of the genetics underlying disease. However, the complexity of the regulatory networks that are affected by disease variants and the ever increasing number of genetic variants to dissect [97] means that further work is needed to accelerate clinical translation of genetics findings.

In this regard, the high degree of context specificity of some disease-associated regulatory elements means that bulk chromatin profiling methods may be limited for the detection of regulatory elements that are active in very specific cell types or developmental/stimulatory conditions. This challenge will likely be overcome by recently developed single cell (sc) methods like scATAC-seq [98], scChIP-seq [99] and scHi-C [100], among others. These methods are likely to reveal even more complex layers to the dysregulation of gene expression that characterize complex autoimmune diseases and will probably replace bulk assays in the future. For example, the single-cell eQTLGen Consortium has recently been established to investigate the consequences of disease-associated genetic variants in specific immune cell types at the gene expression level [101]. Furthermore, multi-sector initiatives such as The Accelerating Medicines Partnership (AMP) program will advance our knowledge of the key cell types that need to be targeted to understand the genetic basis of autoimmune disease. Indeed, using single cell transcriptomics, distinct cell states that are expanded in RA have been identified as part of phase I of the AMP program [102].

Functional investigation of GWAS variants using epigenomics and 3D chromatin structure is extremely valuable

in the generation of hypotheses as to which SNPs within a disease locus might be causal, the regulatory mechanism they might be disrupting and in which cell or tissue type, and their likely causal or target gene/s. These hypotheses then need to be experimentally validated. Although still somewhat limited by effect size, efficiency and implementation in disease relevant primary cells, clustered regularly interspaced short palindromic repeats (CRISPR)-based methods have become the gold standard for the functional dissection of GWAS loci in recent years [103]. The use of CRISPR-based technologies and massively parallel reporter assays (MPRAs) to functionally interpret GWAS signals is reviewed extensively elsewhere in this issue of *Seminars in Immunopathology* [104].

Finally, meaningful integration of the vast wealth of functional data and knowledge that has been generated in the past few years is essential for elucidating biological networks that are altered in disease and for fulfilling the potential of GWAS for drug development and repositioning. Initiatives such as Open Targets [105, 106] and bioinformatics pipelines like EpiMap [54], IMPACT [55] and others [45, 107, 108] provide frameworks to prioritize potential targets by integrating GWAS data with genomic features, disease ontologies and network connectivity. However, data integration still currently remains a challenge, and further research is needed in this area.

These exciting developments, together with future advances and the use of patient-derived biological material and electronic health records, will further our understanding of complex autoimmune diseases, assist drug development and enable precision medicine, delivering on the promise of GWAS for patient benefit.

Data availability Not applicable.

Code availability Not applicable.

Declarations

Conflict of interest The author declares no competing interests.

References

1. Buniello A, Macarthur JAL, Cerezo M et al (2018) The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res 47:1005–1012. https://doi.org/10.1093/nar/gky1120
2. Nelson MR, Tipney H, Painter JL et al (2015) The support of human genetic evidence for approved drug indications. Nat Genet 47:856–860. https://doi.org/10.1038/ng.3314
3. Cook D, Brown D, Alexander R et al (2014) Lessons learned from the fate of AstraZeneca's drug pipeline: a five-dimensional framework. Nat Rev Drug Discov 13:419–431. https://doi.org/10.1038/nrd4309
4. Plenge RM, Scolnick EM, Altshuler D (2013) Validating therapeutic targets through human genetics. Nat Rev Drug Discov 12:581–594. https://doi.org/10.1038/nrd4051
5. Pritchard J-LE, O'Mara TA, Glubb DM (2017) Enhancing the promise of drug repositioning through genetics. Front Pharmacol 8:896. https://doi.org/10.3389/fphar.2017.00896
6. CA Dendrou A Cortes L Shipman et al 2016 Resolving TYK2 locus genotype-to-phenotype differences in autoimmunity SciTransl Med 8 https://doi.org/10.1126/SCITRANSLMED.AAG1974
7. D Hromadová D Elewaut RD Inman et al 2021 From Science to success? Targeting tyrosine kinase 2 in spondyloarthritis and related chronic inflammatory diseases Front Genet 1108 https://doi.org/10.3389/FGENE.2021.685280
8. Reay WR (2021) Cairns MJ (2021) Advancing the use of genome-wide association studies for drug repurposing. Nat Rev Genet 2210(22):658–671. https://doi.org/10.1038/s41576-021-00387-z
9. Duerr RH, Taylor KD, Brant SR et al (2006) A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science 314:1461–1463. https://doi.org/10.1126/science.1135245
10. Sandborn WJ, Feagan BG, Fedorak RN et al (2008) A randomized trial of ustekinumab, a human interleukin-12/23 monoclonal antibody, in patients with moderate-to-severe Crohn's disease. Gastroenterology 135:1130–1141. https://doi.org/10.1053/J.GASTRO.2008.07.014
11. Sandborn WJ, Gasink C, Gao L-L, et al (2012) Ustekinumab induction and maintenance therapy in refractory Crohn's disease. 367 1519 1528 https://doi.org/10.1056/NEJMOA1203572
12. Feagan BG, Sandborn WJ, Gasink C, et al (2016) Ustekinumab as induction and maintenance therapy for Crohn's disease. 375 1946 1960 https://doi.org/10.1056/NEJMOA1602773
13. McInnes IB, Sieper J, Braun J et al (2014) Efficacy and safety of secukinumab, a fully human anti-interleukin-17A monoclonal antibody, in patients with moderate-to-severe psoriatic arthritis: A 24-week, randomised, double-blind, placebo-controlled, phase ii proof-of-concept trial. Ann Rheum Dis 73:349–356. https://doi.org/10.1136/annrheumdis-2012-202646
14. Sieper J, Deodhar A, Marzo-Ortega H et al (2017) Secukinumab efficacy in anti-TNF-naive and anti-TNF-experienced subjects with active ankylosing spondylitis: results from the MEASURE 2 Study. Ann Rheum Dis 76:571–575. https://doi.org/10.1136/annrheumdis-2016-210023

15. Zeggini E, Gloyn AL, Barton AC, Wain LV (2019) Translational genomics and precision medicine: moving from the lab to the clinic. Science 365:1409–1413. https://doi.org/10.1126/science.aax4588

16. Dand N, Duckworth M, Baudry D et al (2019) HLA-C*06:02 genotype is a predictive biomarker of biologic treatment response in psoriasis. J Allergy Clin Immunol 143:2120–2130. https://doi.org/10.1016/J.JACI.2018.11.038

17. Viatte S, Plant D, Han B et al (2015) Association of HLA-DRB1 haplotypes with rheumatoid arthritis severity, mortality, and treatment response. JAMA 313:1645. https://doi.org/10.1001/jama.2015.3435

18. Sudlow C, Gallacher J, Allen N et al (2015) UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLOS Med 12:e1001779. https://doi.org/10.1371/JOURNAL.PMED.1001779

19. Glicksberg BS, Johnson KW, Dudley JT (2018) The next generation of precision medicine: observational studies, electronic health records, biobanks and continuous monitoring. Hum Mol Genet 27:R56–R62. https://doi.org/10.1093/HMG/DDY114

20. Tam V, Patel N, Turcotte M et al (2019) Benefits and limitations of genome-wide association studies. Nat Rev Genet 20:467–484. https://doi.org/10.1038/s41576-019-0127-1

21. French JD, Edwards SL (2020) The role of noncoding variants in heritable disease. Trends Genet 36:880–891. https://doi.org/10.1016/j.tig.2020.07.004

22. Farh KKH, Marson A, Zhu J et al (2015) Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature 518:337–343. https://doi.org/10.1038/nature13835

23. Caliskan M, Brown CD, Maranville JC (2021) A catalog of GWAS fine-mapping efforts in autoimmune disease. Am J Hum Genet 108:549–563. https://doi.org/10.1016/j.ajhg.2021.03.009

24. DJ Schaid W Chen NB Larson 2018 From genome-wide associations to candidate causal variants by statistical fine-mapping Nat Rev Genet 1 https://doi.org/10.1038/s41576-018-0016-z

25. Chatterjee S, Ahituv N (2017) Gene regulatory elements, major drivers of human disease. Annu Rev Genomics Hum Genet 18:45–63. https://doi.org/10.1146/ANNUREV-GENOM-091416-035537

26. Andersson R, Sandelin A (2020) Determinants of enhancer and promoter activities of regulatory elements. Nat Rev Genet 21:71–87. https://doi.org/10.1038/s41576-019-0173-8

27. Maurano MT, Humbert R, Rynes E et al (2012) Systematic localization of common disease-associated variation in regulatory DNA. Science 337:1190–1195. https://doi.org/10.1126/SCIENCE.1222794

28. Gasperini M, Tome JM, Shendure J (2020) Towards a comprehensive catalogue of validated and target-linked human enhancers. Nat Rev Genet 21:292–310. https://doi.org/10.1038/s41576-019-0209-0

29. Cano-Gamez E, Trynka G (2020) From GWAS to function: using functional genomics to identify the mechanisms underlying complex diseases. Front Genet 11:424. https://doi.org/10.3389/fgene.2020.00424

30. Dunham I, Kundaje A, Aldred SF et al (2012) An integrated encyclopedia of DNA elements in the human genome. Nature 489:57–74. https://doi.org/10.1038/nature11247

31. Bernstein BE, Stamatoyannopoulos JA, Costello JF et al (2010) The NIH Roadmap Epigenomics Mapping Consortium. Nat Biotechnol 2810(28):1045–1048. https://doi.org/10.1038/nbt1010-1045

32. Stunnenberg HG, Abrignani S, Adams D et al (2016) The International Human Epigenome Consortium: a blueprint for scientific collaboration and discovery. Cell 167:1145–1149. https://doi.org/10.1016/J.CELL.2016.11.007

33. Core LJ, Martins AL, Danko CG et al (2014) Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. Nat Genet 4612(46):1311–1320. https://doi.org/10.1038/ng.3142

34. Andersson R, Gebhard C, Miguel-Escalada I et al (2014) An atlas of active enhancers across human cell types and tissues. Nature 507:455–461. https://doi.org/10.1038/nature12787

35. Aguet F, Brown AA, Castel SE et al (2017) Genetic effects on gene expression across human tissues. Nature 550:204–213. https://doi.org/10.1038/nature24277

36. Boyle AP, Davis S, Shulha HP et al (2008) High-resolution mapping and characterization of open chromatin across the genome. Cell 132:311–322. https://doi.org/10.1016/J.CELL.2007.12.014

37. Schones DE, Cui K, Cuddapah S et al (2008) Dynamic regulation of nucleosome positioning in the human genome. Cell 132:887–898. https://doi.org/10.1016/J.CELL.2008.02.022

38. Buenrostro JD, Giresi PG, Zaba LC et al (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods 10:1213–1218. https://doi.org/10.1038/nmeth.2688

39. Ren B, Robert F, Wyrick JJ et al (2000) Genome-wide location and function of DNA binding proteins. Science 290:2306–2309. https://doi.org/10.1126/SCIENCE.290.5500.2306

40. Skene PJ, Henikoff S (2017) An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. Elife 6:e21856. https://doi.org/10.7554/ELIFE.21856

41. Lieberman-Aiden E, van Berkum NL, Williams L et al (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326:289–293. https://doi.org/10.1126/SCIENCE.1181369

42. Dryden NH, Broome LR, Dudbridge F et al (2014) Unbiased analysis of potential targets of breast cancer susceptibility loci by Capture Hi-C. Genome Res 24:1854–1868. https://doi.org/10.1101/gr.175034.114

43. Mumbach MR, Rubin AJ, Flynn RA et al (2016) HiChIP: efficient and sensitive analysis of protein-directed genome architecture. Nat Methods 13:919–922. https://doi.org/10.1038/nmeth.3999

44. Klemm SL, Shipony Z, Greenleaf WJ (2019) Chromatin accessibility and the regulatory epigenome. Nat Rev Genet 20:207–220. https://doi.org/10.1038/s41576-018-0089-8

45. Pickrell JK (2014) ARTICLE joint analysis of functional genomic data and genome-wide association studies of 18 human traits. Am J Hum Genet 94:559–573. https://doi.org/10.1016/j.ajhg.2014.03.004

46. Maurano MT, Haugen E, Sandstrom R et al (2015) Large-scale identification of sequence variants influencing human transcription factor occupancy in vivo. Nat Genet 47:1393–1401. https://doi.org/10.1038/ng.3432

47. Trynka G, Sandor C, Han B et al (2013) Chromatin marks identify critical cell types for fine mapping complex trait variants. Nat Genet 45:124–130. https://doi.org/10.1038/ng.2504

48. Eyre S, Bowes J, Diogo D et al (2012) High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. Nat Genet 44:1336–1340. https://doi.org/10.1038/ng.2462

49. Stahl EA, Raychaudhuri S, Remmers EF et al (2010) (2010) Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat Genet 426(42):508–514. https://doi.org/10.1038/ng.582

50. Weissbrod O, Hormozdiari F, Benner C et al (2020) Functionally informed fine-mapping and polygenic localization of complex

trait heritability. Nat Genet 52:355–1363. https://doi.org/10.1038/s41588-020-00735-5

51. Onengut-Gumuscu S, Chen WM, Burren O et al (2015) Fine mapping of type 1 diabetes susceptibility loci and evidence for colocalization of causal variants with lymphoid gene enhancers. Nat Genet 47:381–386. https://doi.org/10.1038/ng.3245

52. Chen L, Ge B, Casale FP et al (2016) Genetic drivers of epigenetic and transcriptional variation in human immune cells. Cell 167:1398-1414.e24. https://doi.org/10.1016/j.cell.2016.10.026

53. Soskic B, Cano-Gamez E, Smyth DJ et al (2019) Chromatin activity at GWAS loci identifies T cell states driving complex immune diseases. Nat Genet 51:1486–1493. https://doi.org/10.1038/s41588-019-0493-9

54. Boix CA, James BT, Park YP et al (2021) Regulatory genomic circuitry of human disease loci by integrative epigenomics. Nature 590:300–307. https://doi.org/10.1038/s41586-020-03145-z

55. Amariuta T, Luo Y, Gazal S et al (2019) IMPACT: genomic annotation of cell-state-specific regulatory elements inferred from the epigenome of bound transcription factors. An J Hum Genet 104:879–895. https://doi.org/10.1016/j.ajhg.2019.03.012

56. Amariuta T, Ishigaki K, Sugishita H et al (2020) Improving the trans-ancestry portability of polygenic risk scores by prioritizing variants in predicted cell-type-specific regulatory elements. Nat Genet 52:1346–1354. https://doi.org/10.1038/s41588-020-00740-8

57. Yan J, Qiu Y, Ribeiro dos Santos AM et al (2021) Systematic analysis of binding of transcription factors to noncoding variants. Nature 591:147–151. https://doi.org/10.1038/s41586-021-03211-0

58. Qu K, Zaba LC, Giresi PG et al (2015) Individuality and variation of personal regulomes in primary human T cells. Cell Syst 1:51–61. https://doi.org/10.1016/J.CELS.2015.06.003

59. Alasoo K, Rodrigues J, Mukhopadhyay S et al (2018) Shared genetic effects on chromatin and gene expression indicate a role for enhancer priming in immune response. Nat Genet 50:424–431. https://doi.org/10.1038/s41588-018-0046-7

60. Degner JF, Pai AA, Pique-Regi R et al (2012) DNase I sensitivity QTLs are a major determinant of human expression variation. Nature 482:390–394. https://doi.org/10.1038/nature10808

61. Gate RE, Cheng CS, Aiden AP et al (2018) Genetic determinants of co-accessible chromatin regions in activated T cells across humans. Nat Genet 50:1140–1150. https://doi.org/10.1038/s41588-018-0156-2

62. Khetan S, Kursawe R, Youn A et al (2018) Type 2 diabetes–associated genetic variants regulate chromatin accessibility in human islets. Diabetes 67:2466–2477. https://doi.org/10.2337/DB18-0393

63. Pelikan RC, Kelly JA, Fu Y et al (2018) Enhancer histone-QTLs are enriched on autoimmune risk haplotypes and influence gene expression within chromatin networks. Nat Commun 9:2905. https://doi.org/10.1038/s41467-018-05328-9

64. PC Dubois Trynka G Franke L et al 2010 Multiple common variants for celiac disease influencing immune gene expression Nat Genet 42 295 302 https://doi.org/10.1038/ng.543

65. Nicolae DL, Gamazon E, Zhang W et al (2010) Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. PLoS Genet 6:e1000888. https://doi.org/10.1371/journal.pgen.1000888

66. Westra H-J, Peters MJ, Esko T et al (2013) Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat Genet 45:1238–1243. https://doi.org/10.1038/ng.2756

67. Giambartolomei C, Vukcevic D, Schadt EE et al (2014) Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLOS Genet 10:e1004383. https://doi.org/10.1371/JOURNAL.PGEN.1004383

68. Watanabe K, Taskesen E, van Bochoven A, Posthuma D (2017) Functional mapping and annotation of genetic associations with FUMA. Nat Commun 8:1826. https://doi.org/10.1038/s41467-017-01261-5

69. Nica AC, Montgomery SB, Dimas AS et al (2010) Candidate causal regulatory effects by integration of expression QTLs with complex trait genetic associations. PLoS Genet 6:e1000895. https://doi.org/10.1371/journal.pgen.1000895

70. He X, Fuller CK, Song Y et al (2013) Sherlock: detecting gene-disease associations by matching patterns of expression QTL and GWAS. Am J Hum Genet 92:667–680. https://doi.org/10.1016/j.ajhg.2013.03.022

71. Piasecka B, Duffy D, Urrutia A et al (2018) Distinctive roles of age, sex, and genetics in shaping transcriptional variation of human immune responses to microbial challenges. Proc Natl Acad Sci 115:E488–E497. https://doi.org/10.1073/PNAS.1714765115

72. Zhernakova DV, Deelen P, Vermaat M et al (2016) Identification of context-dependent expression quantitative trait loci in whole blood. Nat Genet 49:139–145. https://doi.org/10.1038/ng.3737

73. Fairfax BP, Humburg P, Makino S et al (2014) Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. Science 343:1246949. https://doi.org/10.1126/SCIENCE.1246949

74. Consortium TG (2013) The Genotype-Tissue Expression (GTEx) project. Nat Genet 45:580–585. https://doi.org/10.1038/ng.2653

75. Aguet F, Barbeira AN, Bonazzola R et al (2020) The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science 369:1318–1330. https://doi.org/10.1126/SCIENCE.AAZ1776

76. Võsa U, Claringbould A, Westra H-J et al (2021) Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. Nat Genet 53:1300–1310. https://doi.org/10.1038/s41588-021-00913-z

77. Kim-Hellmuth S, Aguet F, Oliva M, et al (2020) Cell type specific genetic regulation of gene expression across human tissues. Science 369: eaaz8528. https://doi.org/10.1126/science.aaz8528

78. BJ Schmiedel D Singh A Madrigal et al 2018 Impact of genetic polymorphisms on human immune cell gene expression resource impact of genetic polymorphisms on human immune cell gene expression Cell 175 https://doi.org/10.1016/j.cell.2018.10.022

79. Ota M, Nagafuchi Y, Hatano H et al (2021) Dynamic landscape of immune cell-specific gene regulation in immune-mediated diseases. Cell 184:3006–3021. https://doi.org/10.1016/j.cell.2021.03.056

80. Kerimov N, Hayhurst JD, Peikova K et al (2021) A compendium of uniformly processed human gene expression and splicing quantitative trait loci. Nat Genet 53:1290–1299. https://doi.org/10.1038/s41588-021-00924-w

81. Krijger PHL, de Laat W (2016) Regulation of disease-associated gene expression in the 3D genome. Nature Publishing Group

82. AM Oudelaar DR Higgs 2020 The relationship between genome structure and function Nat Rev Genet 22 https://doi.org/10.1038/s41576-020-00303-x

83. Anania C, Lupiáñez DG Order and disorder: abnormal 3D chromatin organization in human disease. Brief Funct Genomics 19:128–138. https://doi.org/10.1093/bfgp/elz028

84. Greenwald WW, Li H, Benaglio P et al (2019) Subtle changes in chromatin loop contact propensity are associated with differential gene regulation and expression. Nat Commun 10:1–17. https://doi.org/10.1038/s41467-019-08940-5

85. J Yang A McGovern P Martin et al 2020 Analysis of chromatin organization and gene expression in T cells identifies functional genes for rheumatoid arthritis 11 1 13 https://doi.org/10.1038/s41467-020-18180-7

86. Gorkin DU, Qiu Y, Hu M et al (2019) Common DNA sequence variation influences 3-dimensional conformation of the human genome. Genome Biol 20:255. https://doi.org/10.1186/s13059-019-1855-4

87. Fasolino M, Goldman N, Wang W et al (2020) Genetic variation in type 1 diabetes reconfigures the 3D chromatin organization of T cells and alters gene expression. Immunity 52:257-274.e11. https://doi.org/10.1016/j.immuni.2020.01.003

88. Javierre BM, Sewitz S, Cairns J et al (2016) Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters. Cell 167:1369-1384.e19. https://doi.org/10.1016/j.cell.2016.09.037

89. Chandra V, Bhattacharyya S, Schmiedel BJ et al (2021) Promoter-interacting expression quantitative trait loci are enriched for functional genetic variants. Nat Genet 53:110–119. https://doi.org/10.1038/s41588-020-00745-3

90. Martin P, McGovern A, Orozco G et al (2015) Capture Hi-C reveals novel candidate genes and complex long-range interactions with related autoimmune risk loci. Nat Commun 6:10069. https://doi.org/10.1038/ncomms10069

91. Martin P, Ding J, Duffus K et al (2019) Chromatin interactions reveal novel gene targets for drug repositioning in rheumatic diseases. Ann Rheum Dis 78:1127–1134. https://doi.org/10.1136/annrheumdis-2018-214649

92. Ge X, Frank-Bertoncelj M, Klein K et al (2021) Functional genomics atlas of synovial fibroblasts defining rheumatoid arthritis heritability. Genome Biol 22:247. https://doi.org/10.1186/S13059-021-02460-6

93. Fu Y, Tessneer KL, Li C, Gaffney PM (2018) From association to mechanism in complex disease genetics: the role of the 3D genome. Genome Biol 20:216. https://doi.org/10.1186/s13075-018-1721-x

94. Ding J, Frantzeskos A, Orozco G (2020) Functional genomics in autoimmune diseases. Hum Mol Genet 29:R59–R65. https://doi.org/10.1093/hmg/ddaa097

95. Eijsbouts CQ, Burren OS, Newcombe PJ (2018) Wallace C (2019) Fine mapping chromatin contacts in capture Hi-C data. BMC Genomics 201(20):77. https://doi.org/10.1186/S12864-018-5314-5

96. Hua P, Badat M, Hanssen LLP et al (2021) Defining genome architecture at base-pair resolution. Nature 595:125–129. https://doi.org/10.1038/s41586-021-03639-4

97. Mathieson I (2021) The omnigenic model and polygenic prediction of complex traits. Am J Hum Genet 108:1558–1563. https://doi.org/10.1016/J.AJHG.2021.07.003

98. Buenrostro JD, Wu B, Litzenburger UM et al (2015) Single-cell chromatin accessibility reveals principles of regulatory variation. Nature 523:486–490. https://doi.org/10.1038/nature14590

99. Rotem A, Ram O, Shoresh N et al (2015) Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. Nat Biotechnol 33:1165–1172. https://doi.org/10.1038/nbt.3383

100. Nagano T, Lubling Y, Stevens TJ et al (2013) Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. Nature 502:59–64. https://doi.org/10.1038/nature12593

101. MGP Wijst van der DH Vries de HE Groot et al 2020 The single-cell eQTLGen consortium Elife 9 e52155 https://doi.org/10.7554/ELIFE.52155

102. Zhang F, Wei K, Slowikowski K et al (2019) Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. Nat Immunol 20:928–942. https://doi.org/10.1038/s41590-019-0378-1

103. L Przybyla LA Gilbert 2021 A new era in functional genomics screens Nat Rev Genet https://doi.org/10.1038/s41576-021-00409-w

104. J Ding A Frantzeskos G Orozco 2021 (2021) Functional interrogation of autoimmune disease genetics using CRISPR/Cas9 technologies and massively parallel reporter assays SeminImmunopathol https://doi.org/10.1007/s00281-021-00887-4

105. Ghoussaini M, Mountjoy E, Carmona M et al (2020) Open Targets Genetics: systematic identification of trait-associated genes using large-scale genetics and functional genomics. Nucleic Acids Res 49:1311–1320. https://doi.org/10.1093/nar/gkaa840

106. Mountjoy E, Schmidt EM, Carmona M et al (2021) An open approach to systematically prioritize causal variants and genes at all published human GWAS trait-associated loci. Nat Genet 253:1527–1533. https://doi.org/10.1038/s41588-021-00945-5

107. Iotchkova V, Ritchie GRS, Geihs M et al (2019) GARFIELD classifies disease-relevant genomic features through integration of functional annotations with association signals. Nat Genet 51:343–353. https://doi.org/10.1038/s41588-018-0322-6

108. Fang H, Beckmann G, Bountra C et al (2019) A genetics-led approach defines the drug target landscape of 30 immune-related traits. Nat Genet 51:1082–1091. https://doi.org/10.1038/s41588-019-0456-1