Exploring the world of non-coding genes in stem cells and autoimmunity.

Messemaker, T.C.

Chapter 3

Immunogenetics of rheumatoid arthritis: Understanding functional implications

J Autoimmun. 2015 Nov;64:74-81. doi: 10.1016/j.jaut.2015.07.007.

Messemaker TC, Huizinga TW, Kurreeman F

Abstract

The last decade has seen a dramatic technological revolution. The characterisation of the majority of the common variations in our genetic code in 2003 precipitated the discovery of the genetic risk factors predisposing to Rheumatoid Arthritis development and progression. Prior to 2007, only a handful of genetic risk factors had been identified, HLA, PTPN22 and CTLA4. Since then, over 100 genetic risk loci have been described, with the prediction that an ever-increasing number of risk alleles with consistently decreasing effect sizes will be discovered in the years to come. Each risk locus harbours multiple candidate genes and the proof of causality of each of these candidates is as yet unknown. An enrichment of these RA-associated genes is found in three pathways: T-cell receptor signalling, JAK-STAT signalling and the NF-κB signalling cascade, and currently drugs targeting these pathways are available for the treatment of RA. However, the role that RA-associated genes have in these pathways and how they contribute to disease is not always clear. Major efforts in understanding the contribution of genetic risk factors are currently under way with studies querying the role of genetic variation in gene expression of coding and non-coding genes, epigenetic marks and other regulatory mechanisms yielding ever more valuable insights into mechanisms of disease. Recent work has suggested a possible enrichment of non-coding RNAs as well as super-enhancers in RA genetic loci indicating possible new insights into disease mechanism. This review brings together these emerging genetic data with an emphasis on the immunogenetic links these findings have provided and what we expect the future will bring.

1. Introduction

Rheumatoid arthritis (RA) is a heterogeneous chronic (auto)immune disease associated with significant morbidity and reduced life expectancy. Global prevalence of RA has been estimated to be around 0.2–0.5% on average, with a large variation across regions [1,2]. The highest prevalence has been detected in Europe and North America with lower prevalence in Africa and Southeast Asia. In general, there is a two-fold higher occurrence in females than in males. Given the common prevalence and the lack of a cure for RA, the socio-economic burden remains large and is predicted to rise with an increasingly ageing population [3]. Rheumatoid Arthritis is characterized by chronic inflammation and destruction of the synovial joints leading to progressive joint damage and disability.

Autoimmunity, identified by the production of auto-antibodies such as rheumatoid factor (RF) or anti-citrullinated protein antibodies (ACPA) precedes the clinically detectable onset of inflammatory arthritis and can last for years (these aspects have been reviewed elsewhere in detail) [4]. Individuals who harbour autoantibodies tend to have a more severe disease course and respond differently to treatment as compared to those who do not [5]. Interestingly though, at the time of diagnosis, no difference has so far been detected in clinical presentation of autoantibody positive patients versus autoantibody negative patients.

Both genetic and environmental factors are thought to play a role in disease development and disease progression. The heritable component of RA is evident from the 15% concordance rate observed in monozygotic twin pairs and increased familial clustering [6,7]. Heritability estimates of autoantibody positive individuals are similar to autoantibody negative individuals (~40–50%) indicating a significant contribution of genetic factors to both subgroups [8].

Identifying genetic factors has largely been hampered by the existence of genetic heterogeneity, low penetrance of individual disease alleles and the potential for gene–gene/gene–environment interactions. Nevertheless, candidate gene studies but to a larger extent genome-wide association studies querying ~10 million variants in the human genome in ~100,000 individuals have led to the identification of >100 loci that are associated with RA [9]. These loci individually confer only modest effects decreasing their potential utility in the clinic as a prediction tool but do provide important insights into relevant pathways involved in the disease process. It is important to note that the majority of association studies have been performed in individuals of European ancestry and patients who harbour autoantibodies and our review mainly discusses the immunogenetic pathways from this relatively homogenous group of patients.

2. Immunogenetics of the HLA association with RA

The main genetic region linked to RA over thirty years ago, before the advent of genome wide association studies (GWAS), is the HLA region which is encoded by the major histocompatibility complex (MHC). The MHC locus spans approximately 4 Mb and contains approximately 250 genes, of which ~60% have immune-

related functions. The region is characterised by extended and complex linkage disequilibrium patterns that have made it notoriously difficult to pinpoint the causal gene(s) in the region. The initial association between HLA and RA was made in 1976 with the observation of an overrepresentation of HLA-DR4 in mixed lymphocyte cultures of RA patients [10]. Other HLA-DR molecules associate with RA defined by a common amino acid sequence in the HLA-DRB1 chain, termed the HLA shared epitope (HLA-SE) [11].

Over the last few years with the advent of GWAS to measure millions of variants along with the possibility to deeply sequence our genome, significant progress has been made to assess the association of the HLA region to autoantibody positive RA. More precisely, amino acid positions 11, 13, 71 and 74 at the HLA-DRB1 chain as well as position 9 of HLA-B and position 9 of HLA-DPB1 have now been identified as being the most statistically significant associations [12]. These positions are located within the antigen-binding groove to the HLA molecule further supporting the role of T cells in RA. Similar associations at the HLA region do exist in African Americans and East Asians, indicating possible shared mechanisms in different ethnic groups although more well powered studies need to be performed to dissect the overlap and differences at the HLA alleles [13]. In contrast, clearly distinct association signals (e.g. HLA-DR3) have been observed at the HLA locus in ACPA-negative individuals of European descent, shedding light on different genetic predispositions in the two disease subgroups [14–16]. Other HLA haplotypes such as HLA-DRB1*13 carrying the five amino acid sequence DERAA at positions 70–74 protect against development of RA [17,18]. These protective effects are confined to ACPA + patients indicating a possible overlap in pathways mediating risk and protection. While methods have been developed to allow the simultaneous query of hundreds of thousands of samples (which represents significant progress), very few new insights have been generated in elucidating the functional mechanisms underlying the HLA association with RA.

Importantly, a recent study has shown that HLA-DQ molecules, which are in full linkage disequilibrium with HLA-SE alleles, are able to efficiently present DERAA epitopes derived from microorganisms as well as from a self-protein known as vinculin [19]. DERAA-directed T cells can provide help to B cells ultimately leading to ACPA production. Individuals who carry HLA-DR13 tend have an decreased number of DERAA-directed T cells likely due to negative selection in the thymus

providing some additional clue of the role of the HLA locus in disease development [19]. Such studies provide an exciting avenue for future research on how HLA-peptide interactions shape the T-cell repertoire. Interestingly it has also been described that non-inherited maternal antigens expressed by the mother but not by the child are also able to provide protection [20]. This observation holds the promise that exposure to external antigens such as DERAA derived from micro-organisms in individuals with distinct genetic background may lead to protection from developing RA. The future will learn whether this pathway can be exploited to prevent RA in high risk individuals.

3. Non-HLA genetic risk factors

Prior to 2007 only a handful of genes outside of the HLA region had been identified including PTPN22 and CTLA4 in Europeans and PADI4 in Asians. In 2007, the TRAF1-C5 locus was concurrently discovered by a candidate gene approach as well the first genome-wide association study in RA. Research in this area, propelled by unparallelled efforts to (i) sequence human genomes (since 2001) [21], (ii) to characterise the most common genetic variations in human populations (HapMap www.HapMap.org [22,23], 1000 genomes project [24], since 2003), (iii) to reliably impute unmeasured genetic variation through robust statistical methods [25] (iv) to define more homogenous groups of patients (e.g through ACPA positivity), has seen a tremendous increase in the number of genetic regions associated with the susceptibility to RA. 101 loci have now been identified either at genome-wide significant thresholds and/or with evidence from replication studies [9,26–42].

The latest major study encompassed a combined analysis of 100,000 individuals of European and Asian descent with the query of ~10 million single nucleotide variants across the human genome. HLA remains the strongest association to disease with an odds ratio of 2–3 with the second strongest genetic risk being conferred by PTPN22 (OR 1.8) (Fig. 1). The remainder of genetic risk factors have modest effect sizes (<1.2) with a prediction of ever-decreasing odds ratios paired with an ever increasing number of risk alleles which will be discovered as sample sizes increase [43]. HLA explains the majority of the genetic risk ~13% with an additional 5% of the genetic risk being explained by an additional 100 loci discovered to date [9,12].

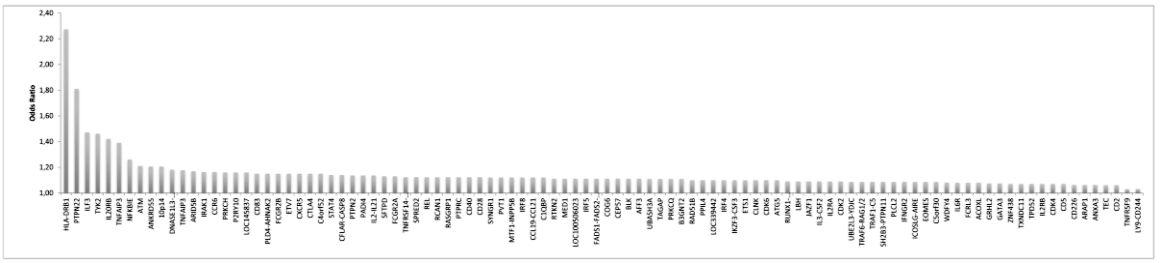

Fig 1. 101 Genetic risk loci predisposing to the development of Rheumatoid Arthritis. Susceptibility loci are ranked by effect size (Odds ratio, y axis) as observed in a meta-analysis performed in 100,000 individuals of European and Asian Ancestry. Each locus is identified by the most likely biological candidate in the associated region as provided by Okada *et al*. [9].

4. Functional implications of genetic risk loci identified to date

There are major challenges to understanding how genetic variation is involved in disease development. An association with a genetic variant does not directly lead to either a causal variant or a causal gene, making the task of translating the functional consequences of genetic variation in diseases where ORs are very low rather challenging. Importantly, the fact that parts of our human genome are inherited in blocks (linkage disequilibrium, LD) [22,23] makes the identification of causal genes and causal variants complicated. The approach currently employed in the identification of causal variants is (i) identify all variants that are (highly) linked to the best signal of association (ii) determine what functional consequences these variants may have (ie are they located in an exon, intron or intergenic region and do they result in a change in protein structure, function or expression). In the end, empirical experimental evidence is required to determine the effects of causal variants and genes and their contribution to the pathogenesis of disease.

In order to understand the functional consequences of genetic findings, there are a few crucial questions. (i) Which SNP will be chosen (ii) what is the endpoint to be measured (for example which gene expression should be measured) and finally (iii) in which cell-type should it be measured and should the cell be activated in order to detect putative differences? Despite the simplicity of the questions, there is complexity at all levels all of which are being addressed by the research community and which will in the end hopefully help us in understanding disease processes better.

Among the 100 non-HLA genetic loci identified to date (Supplementary Table 1), associated regions contain on average ~4 putative candidate genes (in total 377 genes across 100 loci). Within these regions, genetic variants in coding genes that lead to changes in amino acids resulting in dysfunctional proteins are very few (19 out of the 377 genes observed from the latest GWAS analyses, Table 1). Among the genes that have been reported to harbour at least one genetic variant with missense SNP(s) in high LD (R2 > 0.8), only a few have been linked to functional changes in candidate genes (Table 1).

One example of a functional missense variant was identified in PTPN22. The variant (rs2476601) is located in exon 14 and induces an amino acid substitution from an arginine to a trypthophan. This R620W conversion is located in protein motif thought to be required for protein–protein interaction [44]. Elegant studies have shown that the mutant (risk) allele results in decreased TCR and BCR signalling in lymphocytes [45] (reviewed recently by Rawlings *et al*. and by Burn *et al*. [46,47]). Interestingly, in mice, the homologous R620W variant in PTPN22 known as R619W located in Ptpn22 reduced protein levels of Ptpn22 and was shown to manifest in thymic and splenic enlargements [44,48]. This reduced Ptpn22 protein expression has been shown to diminish its inhibitory effect on T- and B-cell activation leading to an increased number of T cells, and an enhanced T-cell, dendritic-cell and B-cell activation. Furthermore the R-to-W conversion seems to increase the resistance of B cells to apoptosis and expands the pool of transitional and auto-reactive B cells. More recently, PTPN22 has been reported to interact with PADI4 (Peptidyl Arginine Deaminase 4), also a risk factor for RA in Asians and European [9]. PADI4 is involved in regulating the citrullination process through which ACPA may be generated. Polymorphisms in the PADI4 haplotype have been shown to affect the mRNA stability of the gene [49]. This is an example of new insight gained into pathogenesis generated by genetic findings. It is known that tolerance to citrullinated antigens is broken in RA and genetic studies have involved HLA in this process [50]. It is not known that the amount of antigen is involved in breaking of tolerance and if the stability of mRNA of PADI4 translates in different levels of the enzyme and subsequent amount of antigen, this is a new hypothesis put forward by genetic data. Interestingly, PTPN22 was recently shown to physically interact with PADI4. Deficiency of PTPN22 led to enhanced protein citrullination and formation of neutrophil extracellular traps (a mechanism in place to combat pathogens [51]). The data suggests that the

PTPN22 risk allele disrupts the interaction of PTPN22 with PADI4 leading to hypercitrullination in peripheral blood mononuclear cells [52]. Although these experiments require independent replication, they do provide novel insights into previously unknown mechanisms that could be at play in disease process and highlight the need to look beyond known interactions between proteins.

Table 1. GWAS candidate genes harbouring missense SNPs (Adapted from Okada *et al*.). Missense SNPs in high linkage disequilibrium with the lead SNP (Supplementary Table 1) with R2 > 0.8 from genetic loci associated with disease are provided below.*,**,*** indicates candidate genes derived from the same genetic locus.

Gene	Missense Variants	R^2 with RA	Functional Effects
PADI4	Gly55Ser, Val82Ala, Gly112Ala	0.95	Affects PADI4 mRNA stability
PTPN22	Arg620Trp	1	Affects BCR and TCR signalling
IL6R	Asp358Ala	1	Impairs classical IL6R signalling
NCK1	Ala116Val	0.92	-
NFKBIE*	Val194Ala, Pro175Leu	1	Decreased NF-κB activity
TCTE1*	Arg95His	0.94	-
AARS2*	Val730Met	0.88	-
TNFAIP3	Phe127Cys	1	Affects TNFAIP3 mRNA and NFKB activity
WDFY4	Arg1816Gln	0.84	-
RTKN2	Ala288Thr	0.88	Increased mRNA levels and increased NF-κB activity
CD5	Ala471Val	0,9	Increased T-cell proliferation and cytokine release
SH2B3	Trp262Arg	0.86	-
PRKCH**	Val374Ile	1	-
AHNAK2*	Gly1901Ser	1	-
ZPBP2*	Ser151Ile	0.99	-
GSDMB**	Pro298Ser, Gly291Arg	0.99	-
TYK2	Pro110Ala	0.87	-
ICOSLG	Trp353Arg	0.94	-
IRAK1	Phe196Ser, Ser453Leu	0.96	-

Other examples of functional studies to elucidate the mechanism of action of the putative causal variant include PADI4 [49], TNFAIP3 [53], IL6R [54], NFKBIE [55],

CD5 [56] (Table 1) as well as TYK2, CCR6, IL2RA and CD40. The CD40 variant leads to increased cell surface expression of CD40 protein on B cells, leading to enhanced NF-κB pathway activation [57]. Chemokine receptor 6 (CCR6) is considered as a surface marker for Th17 cells [58]. Expression of CCR6 correlates with the polymorphism rs3093024 that was found associated with RA [59]. The CCR6 genotype was also correlated with induced IL-17 levels in the sera of RA patients [59]. Interestingly, an increased number of CCR6+T-cells were identified in peripheral blood, synovial fluid and inflamed synovial tissues of RA patients, highlighting an important role in RA pathogenesis [60]. TNFAIP3 encodes a ubiquitin-modifying enzyme (known as A20) that was identified as a component of the NF-κB signalling pathway [61] (Fig. 2). Three independent polymorphisms were identified in this locus to associate with RA susceptibility [37,62]. One of these variants reduced avidity for transcription factor binding by NF-κB resulting in decreased mRNA expression and reduced A20 protein levels [53]. Moreover, mice lacking A20 in myeloid cells developed spontaneous polyarthritis sharing many features with rheumatoid arthritis [63]. These data indicate the importance of the NF-κB signalling cascade in rheumatoid arthritis and highlight how genes like TNFAIP3 can disturb such immune homeostasis. Another functional variant was identified near IL2RA which encodes the IL2 receptor subunit α. This variant (rs12722495) was shown to correlate with both IL2RA mRNA and protein levels in stimulated monocytes, CD4+ Naïve T cells and memory T-cells [64]. The few attempts at ad-hoc functional characterization of causal variants and causal candidate genes in loci identified in genetic studies of complex diseases has proven to be laborious and challenging, yielding limited advance on our understanding of disease pathogenesis but have highlighted important challenges (i) How do we identify the causal variants (ii) How do we identify the causal genes (iii) which relevant cell types are affected.

5. Pathways involved in RA identified by genetic studies

In anticipation of elucidating this three-pronged puzzle, attempts to identify whether candidate genes are enriched in certain molecular pathways have provided some relevant insight into the mechanisms that are at play. Among the 100 loci associated with RA to date, 377 candidate genes have been identified. Candidate genes are often annotated based on their immune function and their closeness to the lead SNP. More sophisticated ways have been developed and

continue to be implemented [65,66]. Pathway analysis of the 377 candidate genes or 100 prioritised genes using StringDB [67] reveals largely similar pathways as previously described with the top three pathways enriched in RA-associated genes being the JAK-STAT signalling, NF-κB signalling and T-cell receptor signalling pathways (Table 2, Fig. 2).

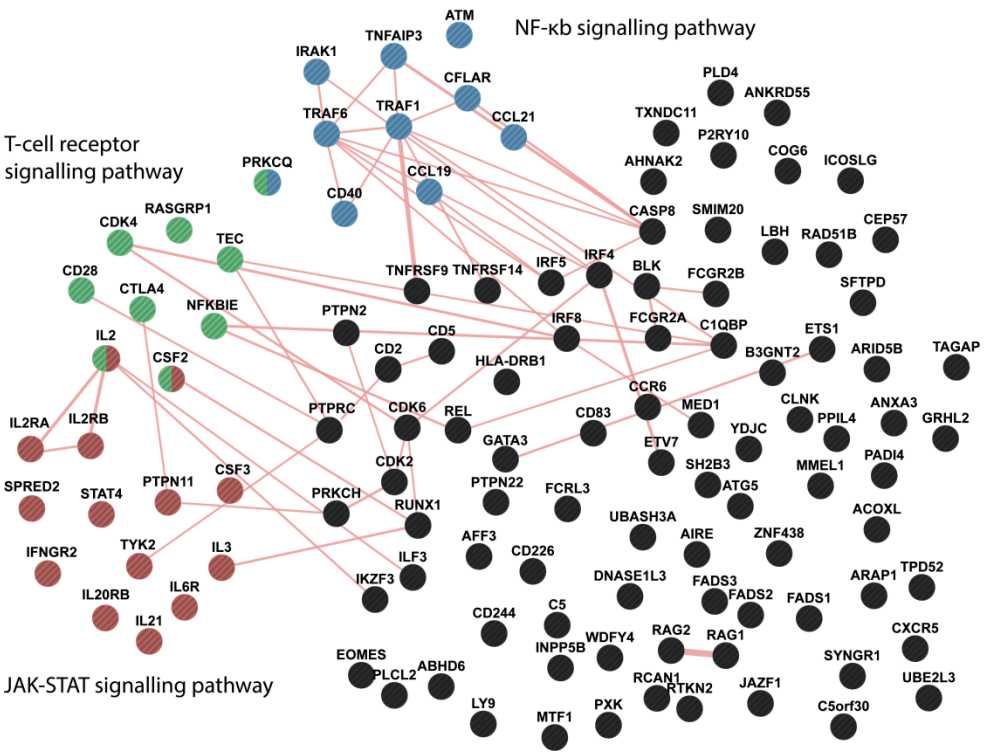

Fig. 2. Interactive overview of top three pathways enriched for rheumatoid arthritis susceptibility genes. Each node represents a susceptibility gene and coloured nodes represent genes in the enriched pathways: JAK-STAT signalling pathway (red), T-cell receptor signalling pathway (green), NF-κb signalling pathway (blue). Protein–protein interaction between susceptibility genes is indicated by red lines. Pathway enrichment analysis was performed in stringDB using KEGG-pathways [117]. GeneMENIA was used for visualisation and protein–protein interactions [118].

Interestingly, several JAK inhibiting drugs are currently under development, with one drug, tofacitinib being approved for treatment of RA [68–71]. In addition, enrichment for B cell and cytokine signalling have also been reported [9]. It is important to map whether “causal” genes or pathways are down or up-regulated

by risk alleles to gain insight into how disease mechanisms operate. However, pathway analyses are limited by what is already known and are often biased towards mechanisms that are mostly studied in the context of common diseases and may therefore limit the possibility of novel hypothesis generating exercises. Bearing this in mind, in order to further our understanding of genes identified and how they play a role in disease, there is a need to move beyond what is already known.

Table 2. Top three pathways identified from candidate genes across 100 loci associated with rheumatoid arthritis (StringDB, Kegg pathway enrichment).

GO ID	Kegg Pathway	# of genes	FDR P	GENES
4630	Jak-STAT signaling pathway	14	2.74E-11	CSF2, CSF3, IL21, IL3, IL20RB, IL2RB, IL6R, IL2, SPRED2, IFNGR2, PTPN11, IL2RA, TYK2, STAT4
4064	NF-kappa B signaling pathway	10	4.75E-9	CCL19, CCL21, PRKCQ, CD40, TNFAIP3, IRAK1, TRAF6, CFLAR, TRAF1, ATM
4660	T cell receptor signaling pathway	9	2.43E-7	IL2, CD28, TEC, CTLA4, CSF2, RASGRP1, CDK4, PRKCQ, CSF2, NFKBIE

6. Non-coding variation, super-enhancers and non-coding RNAs

While coding variants span ~1% of the human genome and explain <10% of the heritability across immune diseases [72], the majority of likely "causal" variants lie in non-coding regions of the genome outside known protein-coding genes and are likely to affect expression of candidate genes [73,74]. In fact, RA-risk SNPs have been found in 44 cis-acting expression quantitative trait loci (cis-eQTL) identified in peripheral blood mononuclear cells [75,76]. Similar observations have been reported for other autoimmune diseases and other cell types, indicating that the quantitative differences of RNA expression with respect to risk alleles may provide clues to disease pathogenesis. Recent studies in CD4+T cells under stimulation conditions revealed cis-eQTLs at 46 genes, 11 of which were previously undetected in peripheral blood mononuclear cells, highlighting the value of well powered cell-type specific analyses to gain novel insights into disease mechanism [77]. There is a growing body of eQTL studies being performed in individual (primary) cell types under basal as well as stimulation conditions [77–89]. Growing evidence suggests that these GWAS signals are enriched in cell-type-specific [90,91], large active regulatory regions of the genome [92–94], known as super-enhancers [95,96]. Based on the analysis of 21

autoimmune GWAS, a recent paper by Farh and colleagues, describes the development of a unique resource for assigning a probability of single nucleotide polymorphisms (SNPs) being causal in disease [97]. 60% of these likely causal variants are located within stimulus and cell-type specific enhancers, identified through both histone modifications and the transcription of noncoding RNAs.

Histone modifications are markers of different chromatin states with methylation or acetylation of specific histones strongly correlating with promoter or enhancer positions and activity [98]. Using this method, Farh *et al*. calculate that the lead SNP is less likely to be the causal variant and is a median 14 kb distance away from the predicted causal SNP. Interestingly, despite the close colocalisation of causal variants to transcription binding motifs, the authors suggest that altering the motif itself to affect binding is unlikely, implying that other as yet unknown mechanisms of mediating their effects remain to be identified.

A more recent study led by Vahedi *et al*. confirms the enrichment of super-enhancers in CD4+T cells in addition to CD56 + NK cells and CD14 + monocytes. CD4+T cells have been repeatedly identified as critical cell types in RA [78,79]. CD4+T cells after being activated and differentiated into distinct effector subtypes play a major role in mediating immune response through the secretion of specific cytokines.

The CD4+T cells carry out multiple functions, ranging from activation of the cells of the innate immune system, B-lymphocytes, cytotoxic T cells, as well as non-immune cells, and also play critical role in the suppression of immune reaction. Importantly, the levels of super-enhancers at GWAS loci detected in CD4+T cells seem to be preferentially affected by the JAK-STAT inhibitor tofacitinib as compared to super-enhancers in non-CD4+T cells, providing indications that genes involved in the disease pathway are likely being targeted [95] and provides hope for such approaches to at least yield valuable drug targets. In addition, these studies help to highlight that genome regulation is dynamic, cell specific and much more complex than previously envisaged [99]. Similar endeavours need to be undertaken for other relevant immune cell (sub)types. Taking into account this complexity, no matter how challenging, will undoubtedly lead us to novel insight as an ever increasing amount of data begins to emerge.

7. Noncoding RNAs as novel candidate genes for rheumatoid arthritis

In recent years, non-coding RNAs have gained much interest as their prevalence in the human genome is much larger than previously anticipated. The ~20,000 protein coding genes occupy less than 2% of total human genome sequence [100]. Not surprisingly, at least 90% of the genome is actively transcribed into noncoding RNAs (ncRNAs), which have no protein coding potentiality [101]. A heterogeneous, novel class of long noncoding RNAs (lncRNAs) with length longer than 200 nucleotides is generally characterized as non-protein transcript [102]. Over the past decade more than 18,000 transcripts have been discovered/annotated as lncRNAs in mammalian transcriptomes [103–105]. Numerous studies have revealed that lncRNAs are believed to form a major proportion of novel transcripts and are known to be involved in number of functionally distinct biological and physiological processes including chromatin remodelling, gene transcription, RNA splicing [106,107] and directly linked to human diseases including various cancers and autoimmune diseases [108–110]. Furthermore, lncRNAs act as a key regulator of inflammatory gene expression by a collaboration involving signal-dependent activation of transcription factors, transcriptional coregulators, and chromatin-modifying factors [111].

Ding and colleagues identified 12 lncRNAs in GWAS regions (LD region defined as R2 > 0.8 with lead SNP) [112]. These lncRNAs were expressed in RNAseq data from the Illumina Human Body Map which consist of 16 human tissue types, including adrenal, adipose, brain, breast, colon, heart, kidney, liver, lung, lymph, ovary, prostate, skeletal muscle, testes, thyroid, and white blood cells. 9 of these were sufficiently far away from protein-coding genes to suggest that they are the putative causal genes in the associated regions. In addition, there are a large number of lncRNAs across the majority of loci identified so far implicating in part that they may be as yet uninvestigated candidate genes. Noncoding RNAs do not easily come up in the list of candidate genes as mostly their functions are not as yet well characterized. Elucidating the roles of lncRNAs and the impact of RA-associated variants have on their function will be an important area of research aimed at elucidating mechanisms of disease susceptibility. Recently, observations of differential regulation of lncRNA pathways relevant to RA have been observed. Various studies show either up or down regulation of lncRNAs after specific immune stimuli [111,113,114]. Specifically, a study evaluated lncRNA expression in CD14 + monocytes from RA patients before and after anti-TNFa or anti-IL6

treatment [115]. 55 lncRNAs were differentially expressed upon TNFa inhibition, while 25 distinct non-overlapping lncRNAs were differentially regulated upon anti-IL6 treatment. Another study performed by Kumar and colleagues has also provided evidence of lncRNA eQTLs from GWAS SNPs (lncRNA eQTLs) further emphasizing the potential role of lncRNAs in the aetiology of disease [116]. Much like enhancers, these lncRNA eQTLs were tissue specific. The average expression of lncRNAs under basal conditions is lower than protein-coding genes. Many more lncRNAs are likely to exist but are as yet undiscovered due to limitations in detection thresholds using quantitative PCR and current depth of RNA sequencing. Further efforts need to be made to further characterize the expression and the function of these lncRNAs in specific cell subsets. There is therefore supporting evidence for a role of lncRNAs in RA and in autoimmunity in general and future studies focussing on this field in autoimmunity is likely going to reveal much about pathogenic mechanisms in disease. In particular, once the role of lncRNAs are better mapped, their consideration as "causal" candidate genes can be included to discover and understand the contribution of genetic findings to disease pathways.

8. Summary and perspectives

In summary, we are currently observing enormous changes in the landscape of moving from genetics to understanding immune function. We are generating new insights in the role of HLA in disease onset through complex laborious functional experiments. In addition, specific cell types including CD4+T cells, B cells as well as novel pathways like the JAK-STAT pathway has been definitively established. Genetic studies have also revealed a striking diversity of molecular pathways to disease, including unexpectedly important contributions of non-coding genetic variation in modulating regulatory elements and immune genes. In particular, advances in technology is increasingly making it possible to (i) prioritise variants more accurately, (ii) prioritise genes more accurately and (iii) prioritise cell-types more accurately. The challenge ahead is to carve out suitable strategies to gain insight into cell-type specific molecular processes and pathways underlying the discovered GWAS signals.

Technologies like mass cytometry, (single cell) gene expression profiling by RNA sequencing and multiplexed functional assays can be leveraged and will enable

the analysis of immune cell function with unprecedented detail and promise not only a deeper understanding of pathogenesis, but also the discovery of novel biomarkers. The large and complex data sets generated by these technologies require specialized approaches for analysis and visualization of results which is a rapidly moving field.

Despite the obvious challenges we have faced and those remaining ahead, it is imperative that we remember that it was barely a decade ago that the first genomes were published and that we are now starting to catalogue a comprehensive list of genetic variation that associates with disease. In contrast to what we mostly expected, that large and obvious changes would be detected in the coding region of the human genome, we are gaining more insight into how our genome is non-linear and dynamic and that taking snapshots of functionality at a given time and under one given condition may have restricted our discovery efforts until now.

Supplementary information

Supplementary Table 1 is available online at the journal of autoimmunity (http://dx.doi.org/10.1016/j.jaut.2015.07.007).

Acknowledgements

This work was supported by the Dutch Arthritis Foundation, a Leiden University Medical Center (LUMC) fellowship (11-1-406), the IMI JU (115142) funded project BeTheCure and European Union (Health-F2-2008-223404) (Seventh Framework Programme integrated project Masterswitch).

References

[1] D.G. Hoy, E. Smith, M. Cross, L. Sanchez-Riera, F.M. Blyth, R. Buchbinder, *et al*. Reflecting on the global burden of musculoskeletal conditions: lessons learnt from the global burden of disease 2010 study and the next steps forward. Ann. Rheum. Dis., 74 (1) (2015 Jan), pp. 4-7

[2] I. Rudan, S. Sidhu, A. Papana, S.J. Meng, Y. Xin-Wei, W. Wang, *et al*. Prevalence of rheumatoid arthritis in low- and middle-income countries: A systematic review and analysis. J. Glob. Health, 5 (1) (2015 Jun), p. 010409

[3] I.B. McInnes, G. Schett The pathogenesis of rheumatoid arthritis. N. Engl. J. Med., 365 (23) (2011 Dec 8), pp. 2205-2219

[4] A. Willemze, L.A. Trouw, R.E. Toes, T.W. Huizinga The influence of ACPA status and characteristics on the course of RA. Nat. Rev. Rheumatol., 8 (3) (2012 Mar), pp. 144-152

[5] K.P. Machold, T.A. Stamm, V.P. Nell, S. Pflugbeil, D. Aletaha, G. Steiner, *et al*. Very recent onset rheumatoid arthritis: clinical and serological patient characteristics associated with radiographic progression over the first years of disease. Rheumatol. Oxf., 46 (2) (2007 Feb), pp. 342-349

[6] A.J. MacGregor, H. Snieder, A.S. Rigby, M. Koskenvuo, J. Kaprio, K. Aho, *et al*. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. Arthritis Rheum., 43 (1) (2000 Jan), pp. 30-37

[7] A.J. Silman, A.J. MacGregor, W. Thomson, S. Holligan, D. Carthy, A. Farhan, *et al*. Twin concordance rates for rheumatoid arthritis: results from a nationwide study. Br. J. Rheumatol., 32 (10) (1993 Oct), pp. 903-907

[8] D. van der Woude, J.J. Houwing-Duistermaat, R.E. Toes, T.W. Huizinga, W. Thomson, J. Worthington, *et al*. Quantitative heritability of anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-negative rheumatoid arthritis. Arthritis Rheum., 60 (4) (2009 Apr), pp. 916-923

[9] Y. Okada, D. Wu, G. Trynka, T. Raj, C. Terao, K. Ikari, *et al*. Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature, 506 (7488) (2014 Feb 20), pp. 376-381

[10] P. Stastny Association of the B-cell alloantigen DRw4 with rheumatoid arthritis. N. Engl. J. Med., 298 (16) (1978 Apr 20), pp. 869-871

[11] P.K. Gregersen, J. Silver, R.J. Winchester The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. Arthritis Rheum., 30 (11) (1987 Nov), pp. 1205-1213

[12] S. Raychaudhuri, C. Sandor, E.A. Stahl, J. Freudenberg, H.S. Lee, X. Jia, *et al*. Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. Nat. Genet., 44 (3) (2012 Mar), pp. 291-296

[13] R.J. Reynolds, A.F. Ahmed, M.I. Danila, L.B. Hughes, P.K. Gregersen, S. Raychaudhuri, *et al*. HLA-DRB1-associated rheumatoid arthritis risk at multiple levels in African Americans: hierarchical lassification systems, amino acid positions, and residues. Arthritis Rheumatol., 66 (12) (2014 Dec), pp. 3274-3282

[14] B. Han, D. Diogo, S. Eyre, H. Kallberg, A. Zhernakova, J. Bowes, *et al*. Fine mapping seronegative and seropositive rheumatoid arthritis to shared and distinct HLA alleles by adjusting for the effects of heterogeneity. Am. J. Hum. Genet., 94 (4) (2014 Apr 3), pp. 522-532

[15] P. Irigoyen, A.T. Lee, M.H. Wener, W. Li, M. Kern, F. Batliwalla, *et al*. Regulation of anti-cyclic citrullinated peptide antibodies in rheumatoid arthritis: contrasting effects of HLA-DR3 and the shared epitope alleles. Arthritis Rheum., 52 (12) (2005 Dec), pp. 3813-3818

[16] K.N. Verpoort, F.A. van Gaalen, A.H. van der Helm-van Mil, G.M. Schreuder, F.C. Breedveld, T.W. Huizinga, *et al*. Association of HLA-DR3 with anti-cyclic citrullinated peptide antibody-negative rheumatoid arthritis. Arthritis Rheum., 52 (10) (2005 Oct), pp. 3058-3062

[17] S. Oka, H. Furukawa, A. Kawasaki, K. Shimada, S. Sugii, A. Hashimoto, *et al*. Protective effect of the HLA-DRB1*13:02 allele in Japanese rheumatoid arthritis patients. PLoS One, 9 (6) (2014), p. e99453

[18] D. van der Woude, B.A. Lie, E. Lundstrom, A. Balsa, A.L. Feitsma, J.J. Houwing-Duistermaat, *et al*. Protection against anti-citrullinated protein antibody-positive rheumatoid arthritis is predominantly associated with HLA-DRB1*1301: a meta-analysis of HLA-DRB1 associations with anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-negative rheumatoid arthritis in four European populations. Arthritis Rheum., 62 (5) (2010 May), pp. 1236-1245

[19] H.J. van, D.T. Jansen, S. Polydorides, A.K. Moustakas, M. Bax, A.L. Feitsma, *et al*. Crossreactivity to vinculin and microbes provides a molecular basis for HLA-based protection against rheumatoid arthritis. Nat. Commun., 6 (2015), p. 6681

[20] A.L. Feitsma, J. Worthington, A.H. van der Helm-van Mil, D. Plant, W. Thomson, J. Ursum, *et al*. Protective effect of noninherited maternal HLA-DR antigens on rheumatoid arthritis development. Proc. Natl. Acad. Sci. U. S. A., 104 (50) (2007 Dec 11), pp. 19966-19970

[21] Finishing the euchromatic sequence of the human genome. Nature, 431 (7011) (2004 Oct 21), pp. 931-945

[22] A haplotype map of the human genome. Nature, 437 (7063) (2005 Oct 27), pp. 1299-1320

[23] The International HapMap Project. Nature, 426 (6968) (2003 Dec 18), pp. 789-796

[24] G.R. Abecasis, A. Auton, L.D. Brooks, M.A. DePristo, R.M. Durbin, R.E. Handsaker, *et al*.

An integrated map of genetic variation from 1,092 human genomes. Nature, 491 (7422) (2012 Nov 1), pp. 56-65

[25] B. Howie, J. Marchini, M. Stephens. Genotype imputation with thousands of genomes. G3 (Bethesda), 1 (6) (2011 Nov), pp. 457-470

[26] S. Eyre, J. Bowes, D. Diogo, A. Lee, A. Barton, P. Martin, *et al*. High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. Nat. Genet., 44 (12) (2012 Dec), pp. 1336-1340

[27] K.P. Liao, F. Kurreeman, G. Li, G. Duclos, S. Murphy, R. Guzman, *et al*. Associations of autoantibodies, autoimmune risk alleles, and clinical diagnoses from the electronic medical records in rheumatoid arthritis cases and non-rheumatoid arthritis controls. Arthritis Rheum., 65 (3) (2013 Mar), pp. 571-581

[28] Y. Okada, C. Terao, K. Ikari, Y. Kochi, K. Ohmura, A. Suzuki, *et al*. Meta-analysis identifies nine new loci associated with rheumatoid arthritis in the Japanese population. Nat. Genet., 44 (5) (2012 May), pp. 511-516

[29] R. Chen, E.A. Stahl, F.A. Kurreeman, P.K. Gregersen, K.A. Siminovitch, J. Worthington, *et al*. Fine mapping the TAGAP risk locus in rheumatoid arthritis. Genes Immun., 12 (4) (2011 Jun), pp. 314-318

[30] A. Zhernakova, E.A. Stahl, G. Trynka, S. Raychaudhuri, E.A. Festen, L. Franke, *et al*. Meta-analysis of genome-wide association studies in celiac disease and rheumatoid arthritis identifies fourteen non-HLA shared loci. PLoS Genet., 7 (2) (2011 Feb), p. e1002004

[31] F. Kurreeman, K. Liao, L. Chibnik, B. Hickey, E. Stahl, V. Gainer, *et al*. Genetic basis of autoantibody positive and negative rheumatoid arthritis risk in a multi-ethnic cohort derived from electronic health records. Am. J. Hum. Genet., 88 (1) (2011 Jan 7), pp. 57-69

[32] E.A. Stahl, S. Raychaudhuri, E.F. Remmers, G. Xie, S. Eyre, B.P. Thomson, *et al*. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat. Genet., 42 (6) (2010 Jun), pp. 508-514

[33] S. Raychaudhuri, B.P. Thomson, E.F. Remmers, S. Eyre, A. Hinks, C. Guiducci, *et al*. Genetic variants at CD28, PRDM1 and CD2/CD58 are associated with rheumatoid arthritis risk. Nat. Genet., 41 (12) (2009 Dec), pp. 1313-1318

[34] F.A. Kurreeman, N.A. Daha, M. Chang, J.J. Catanese, A.B. Begovich, T.W. Huizinga, *et al*. Association of IL2RA and IL2RB with rheumatoid arthritis: a replication study in a Dutch population. Ann. Rheum. Dis., 68 (11) (2009 Nov), pp. 1789-1790

[35] P.K. Gregersen, C.I. Amos, A.T. Lee, Y. Lu, E.F. Remmers, D.L. Kastner, *et al*. REL, encoding a member of the NF-kappaB family of transcription factors, is a newly defined risk locus for rheumatoid arthritis. Nat. Genet., 41 (7) (2009 Jul), pp. 820-823

[36] S. Raychaudhuri, E.F. Remmers, A.T. Lee, R. Hackett, C. Guiducci, N.P. Burtt, *et al*. Common variants at CD40 and other loci confer risk of rheumatoid arthritis. Nat. Genet., 40 (10) (2008 Oct), pp. 1216-1223

[37] R.M. Plenge, C. Cotsapas, L. Davies, A.L. Price, P.I. de Bakker, J. Maller, *et al*. Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. Nat. Genet., 39 (12) (2007 Dec), pp. 1477-1482

[38] F.A. Kurreeman, L. Padyukov, R.B. Marques, S.J. Schrodi, M. Seddighzadeh, G. Stoeken-Rijsbergen, *et al*. A candidate gene approach identifies the TRAF1/C5 region as a risk factor for rheumatoid arthritis. PLoS Med., 4 (9) (2007 Sep), p. e278

[39] E.F. Remmers, R.M. Plenge, A.T. Lee, R.R. Graham, G. Hom, T.W. Behrens, *et al*. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N. Engl. J. Med., 357 (10) (2007 Sep 6), pp. 977-986

[40] R.M. Plenge, M. Seielstad, L. Padyukov, A.T. Lee, E.F. Remmers, B. Ding, *et al*. TRAF1-C5 as a risk locus for rheumatoid arthritis–a genomewide study. N. Engl. J. Med., 357 (12) (2007 Sep 20), pp. 1199-1209

[41] E.W. Karlson, L.B. Chibnik, J. Cui, R.M. Plenge, R.J. Glass, N.E. Maher, *et al*. Associations between human leukocyte antigen, PTPN22, CTLA4 genotypes and rheumatoid arthritis phenotypes of

autoantibody status, age at diagnosis and erosions in a large cohort study. Ann. Rheum. Dis., 67 (3) (2008 Mar), pp. 358-363

[42] R.M. Plenge, L. Padyukov, E.F. Remmers, S. Purcell, A.T. Lee, E.W. Karlson, *et al*. Replication of putative candidate-gene associations with rheumatoid arthritis in >4,000 samples from North America and Sweden: association of susceptibility with PTPN22, CTLA4, and PADI4. Am. J. Hum. Genet., 77 (6) (2005 Dec), pp. 1044-1060

[43] E.A. Stahl, D. Wegmann, G. Trynka, J. Gutierrez-Achury, R. Do, B.F. Voight, *et al*. Bayesian inference analyses of the polygenic architecture of rheumatoid arthritis. Nat. Genet., 44 (5) (2012 May), pp. 483-489

[44] J.F. Cloutier, A. Veillette. Cooperative inhibition of T-cell antigen receptor signaling by a complex between a kinase and a phosphatase. J. Exp. Med., 189 (1) (1999 Jan 4), pp. 111-121

[45] M. Rieck, A. Arechiga, S. Onengut-Gumuscu, C. Greenbaum, P. Concannon, J.H. Buckner. Genetic variation in PTPN22 corresponds to altered function of T and B lymphocytes. J. Immunol., 179 (7) (2007 Oct 1), pp. 4704-4710

[46] D.J. Rawlings, X. Dai, J.H. Buckner. The role of PTPN22 risk variant in the development of autoimmunity: finding common ground between mouse and human. J. Immunol., 194 (7) (2015 Apr 1), pp. 2977-2984

[47] G.L. Burn, L. Svensson, C. Sanchez-Blanco, M. Saini, A.P. Cope. Why is PTPN22 a good candidate susceptibility gene for autoimmune disease? FEBS Lett., 585 (23) (2011 Dec 1), pp. 3689-3698

[48] J. Zhang, N. Zahir, Q. Jiang, H. Miliotis, S. Heyraud, X. Meng, *et al*. The autoimmune disease-associated PTPN22 variant promotes calpain-mediated Lyp/Pep degradation associated with lymphocyte and dendritic cell hyperresponsiveness. Nat. Genet., 43 (9) (2011 Sep), pp. 902-907

[49] A. Suzuki, R. Yamada, X. Chang, S. Tokuhiro, T. Sawada, M. Suzuki, *et al*. Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. Nat. Genet., 34 (4) (2003 Aug), pp. 395-402

[50] T.W. Huizinga, C.I. Amos, A.H. van der Helm-van Mil, W. Chen, F.A. van Gaalen, D. Jawaheer, *et al*. Refining the complex rheumatoid arthritis phenotype based on specificity of the HLA-DRB1 shared epitope for antibodies to citrullinated proteins. Arthritis Rheum., 52 (11) (2005 Nov), pp. 3433-3438

[51] V. Brinkmann, U. Reichard, C. Goosmann, B. Fauler, Y. Uhlemann, D.S. Weiss, *et al*. Neutrophil extracellular traps kill bacteria. Science, 303 (5663) (2004 Mar 5), pp. 1532-1535

[52] H.H. Chang, N. Dwivedi, A.P. Nicholas, I.C. Ho. The W620 polymorphism of PTPN22 disrupts its interaction with PAD4 and enhances citrullination and NETosis. Arthritis Rheumatol. (2015 May 27), 10.1002/art.39215

[53] I. Adrianto, F. Wen, A. Templeton, G. Wiley, J.B. King, C.J. Lessard, *et al*. Association of a functional variant downstream of TNFAIP3 with systemic lupus erythematosus. Nat. Genet., 43 (3) (2011 Mar), pp. 253-258

[54] R.C. Ferreira, D.F. Freitag, A.J. Cutler, J.M. Howson, D.B. Rainbow, D.J. Smyth, *et al*. Functional IL6R 358Ala allele impairs classical IL-6 receptor signaling and influences risk of diverse inflammatory diseases. PLoS Genet., 9 (4) (2013 Apr), p. e1003444

[55] K. Myouzen, Y. Kochi, Y. Okada, C. Terao, A. Suzuki, K. Ikari, *et al*. Functional variants in NFKBIE and RTKN2 involved in activation of the NF-kappaB pathway are associated with rheumatoid arthritis in Japanese. PLoS Genet., 8 (9) (2012 Sep), p. e1002949

[56] M.C. Cenit, M. Martinez-Florensa, M. Consuegra, L. Bonet, E. Carnero-Montoro, N. Armiger, *et al*. Analysis of ancestral and functionally relevant CD5 variants in systemic lupus erythematosus patients. PLoS One, 9 (11) (2014), p. e113090

[57] G. Li, D. Diogo, D. Wu, J. Spoonamore, V. Dancik, L. Franke, *et al*. Human genetics in rheumatoid arthritis guides a high-throughput drug screen of the CD40 signaling pathway. PLoS Genet., 9 (5) (2013 May), p. e1003487

[58] J.D. Haas, F.H. Gonzalez, S. Schmitz, V. Chennupati, L. Fohse, E. Kremmer, *et al*. CCR6 and NK1.1 distinguish between IL-17A and IFN-gamma-producing gammadelta effector T cells. Eur. J. Immunol., 39 (12) (2009 Dec), pp. 3488-3497

[59] Y. Kochi, Y. Okada, A. Suzuki, K. Ikari, C. Terao, A. Takahashi, *et al*. A regulatory variant in CCR6 is associated with rheumatoid arthritis susceptibility. Nat. Genet., 42 (6) (2010 Jun), pp. 515-519

[60] J.P. van Hamburg, P.S. Asmawidjaja, N. Davelaar, A.M. Mus, E.M. Colin, J.M. Hazes, *et al*. Th17 cells, but not Th1 cells, from patients with early rheumatoid arthritis are potent inducers of matrix metalloproteinases and proinflammatory cytokines upon synovial fibroblast interaction, including autocrine interleukin-17A production. Arthritis Rheum., 63 (1) (2011 Jan), pp. 73-83

[61] H.Y. Song, M. Rothe, D.V. Goeddel. The tumor necrosis factor-inducible zinc finger protein A20 interacts with TRAF1/TRAF2 and inhibits NF-kappaB activation. Proc. Natl. Acad. Sci. U. S. A., 93 (13) (1996 Jun 25), pp. 6721-6725

[62] G. Orozco, A. Hinks, S. Eyre, X. Ke, L.J. Gibbons, J. Bowes, *et al*. Combined effects of three independent SNPs greatly increase the risk estimate for RA at 6q23. Hum. Mol. Genet., 18 (14) (2009 Jul 15), pp. 2693-2699

[63] M. Matmati, P. Jacques, J. Maelfait, E. Verheugen, M. Kool, M. Sze, *et al*. A20 (TNFAIP3) deficiency in myeloid cells triggers erosive polyarthritis resembling rheumatoid arthritis. Nat. Genet., 43 (9) (2011 Sep), pp. 908-912

[64] C.A. Dendrou, V. Plagnol, E. Fung, J.H. Yang, K. Downes, J.D. Cooper, *et al*. Cell-specific protein phenotypes for the autoimmune locus IL2RA using a genotype-selectable human bioresource. Nat. Genet., 41 (9) (2009 Sep), pp. 1011-1015

[65] S. Raychaudhuri, R.M. Plenge, E.J. Rossin, A.C. Ng, S.M. Purcell, P. Sklar, *et al*. Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. PLoS Genet., 5 (6) (2009 Jun), p. e1000534

[66] T.H. Pers, J.M. Karjalainen, Y. Chan, H.J. Westra, A.R. Wood, J. Yang, *et al*. Biological interpretation of genome-wide association studies using predicted gene functions. Nat. Commun., 6 (2015), p. 5890

[67] D. Szklarczyk, A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, *et al*. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res., 43 (Database issue) (2015 Jan), pp. D447-D452

[68] K. Garber. Pfizer's JAK inhibitor sails through phase 3 in rheumatoid arthritis. Nat. Biotechnol., 29 (6) (2011 Jun), pp. 467-468

[69] K. Nishimura, J. Saegusa, F. Matsuki, K. Akashi, G. Kageyama, A. Morinobu. Tofacitinib facilitates the expansion of myeloid-derived suppressor cells and ameliorates arthritis in SKG mice. Arthritis Rheumatol., 67 (4) (2015 Apr), pp. 893-902

[70] D.L. Boyle, K. Soma, J. Hodge, A. Kavanaugh, D. Mandel, P. Mease, *et al*. The JAK inhibitor tofacitinib suppresses synovial JAK1-STAT signalling in rheumatoid arthritis. Ann. Rheum. Dis., 74 (6) (2015 Jun), pp. 1311-1316

[71] K. Garber. Pfizer's first-in-class JAK inhibitor pricey for rheumatoid arthritis market. Nat. Biotechnol., 31 (1) (2013 Jan), pp. 3-4

[72] A. Gusev, S.H. Lee, G. Trynka, H. Finucane, B.J. Vilhjalmsson, H. Xu, *et al*. Partitioning heritability of regulatory and cell-type-specific variants across 11 common diseases. Am. J. Hum. Genet., 95 (5) (2014 Nov 6), pp. 535-552

[73] D.L. Nicolae, E. Gamazon, W. Zhang, S. Duan, M.E. Dolan, N.J. Cox. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. PLoS Genet., 6 (4) (2010 Apr), p. e1000888

[74] A.C. Nica, S.B. Montgomery, A.S. Dimas, B.E. Stranger, C. Beazley, I. Barroso, *et al*. Candidate causal regulatory effects by integration of expression QTLs with complex trait genetic associations. PLoS Genet., 6 (4) (2010 Apr), p. e1000895

[75] H.J. Westra, M.J. Peters, T. Esko, H. Yaghootkar, C. Schurmann, J. Kettunen, *et al*. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat. Genet., 45 (10) (2013 Oct), pp. 1238-1243

[76] P.C. Dubois, G. Trynka, L. Franke, K.A. Hunt, J. Romanos, A. Curtotti, *et al*. Multiple common variants for celiac disease influencing immune gene expression. Nat. Genet., 42 (4) (2010 Apr), pp. 295-302

[77] X. Hu, H. Kim, T. Raj, P.J. Brennan, G. Trynka, N. Teslovich, *et al*. Regulation of gene expression in autoimmune disease loci and the genetic basis of proliferation in CD4+ effector memory T cells. PLoS Genet., 10 (6) (2014 Jun), p. e1004404

[78] X. Hu, H. Kim, E. Stahl, R. Plenge, M. Daly, S. Raychaudhuri. Integrating autoimmune risk loci with gene-expression data identifies specific pathogenic immune cell subsets. Am. J. Hum. Genet., 89 (4) (2011 Oct 7), pp. 496-506

[79] T. Raj, K. Rothamel, S. Mostafavi, C. Ye, M.N. Lee, J.M. Replogle, *et al*. Polarization of the effects of autoimmune and neurodegenerative risk alleles in leukocytes. Science, 344 (6183) (2014 May 2), pp. 519-523

[80] S. Kim, J. Becker, M. Bechheim, V. Kaiser, M. Noursadeghi, N. Fricker, *et al*. Characterizing the genetic basis of innate immune response in TLR4-activated human monocytes. Nat. Commun., 5 (2014), p. 5236

[81] M. Rotival, T. Zeller, P.S. Wild, S. Maouche, S. Szymczak, A. Schillert, *et al*. Integrating genome-wide genetic variations and monocyte expression data reveals trans-regulated gene modules in humans. PLoS Genet., 7 (12) (2011 Dec), p. e1002367

[82] L. Franke, R.C. Jansen. eQTL analysis in humans. Methods Mol. Biol., 573 (2009), pp. 311-328

[83] H.J. Westra, D. Arends, T. Esko, M.J. Peters, C. Schurmann, K. Schramm, *et al*. Cell Specific eQTL Analysis without Sorting Cells. PLoS Genet., 11 (5) (2015 May), p. e1005223

[84] B.P. Fairfax, J.C. Knight. Genetics of gene expression in immunity to infection. Curr. Opin. Immunol., 30 (2014 Oct), pp. 63-71

[85] B.P. Fairfax, P. Humburg, S. Makino, V. Naranbhai, D. Wong, E. Lau, *et al*. Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. Science, 343 (6175) (2014 Mar 7), p. 1246949

[86] P.L. De Jager, N. Hacohen, D. Mathis, A. Regev, B.E. Stranger, C. Benoist. ImmVar project: Insights and design considerations for future studies of “healthy” immune variation. Semin. Immunol., 27 (1) (2015 Feb), pp. 51-57

[87] M.N. Lee, C. Ye, A.C. Villani, T. Raj, W. Li, T.M. Eisenhaure, *et al*. Common genetic variants modulate pathogen-sensing responses in human dendritic cells. Science, 343 (6175) (2014 Mar 7), p. 1246980

[88] J.K. Pickrell, J.C. Marioni, A.A. Pai, J.F. Degner, B.E. Engelhardt, E. Nkadori, *et al*. Understanding mechanisms underlying human gene expression variation with RNA sequencing. Nature, 464 (7289) (2010 Apr 1), pp. 768-772

[89] S.B. Montgomery, M. Sammeth, M. Gutierrez-Arcelus, R.P. Lach, C. Ingle, J. Nisbett, *et al*. Transcriptome genetics using second generation sequencing in a Caucasian population. Nature, 464 (7289) (2010 Apr 1), pp. 773-777

[90] G. Trynka, S. Raychaudhuri. Using chromatin marks to interpret and localize genetic associations to complex human traits and diseases. Curr. Opin. Genet. Dev., 23 (6) (2013 Dec), pp. 635-641

[91] G. Trynka, C. Sandor, B. Han, H. Xu, B.E. Stranger, X.S. Liu, *et al*. Chromatin marks identify critical cell types for fine mapping complex trait variants. Nat. Genet., 45 (2) (2013 Feb), pp. 124-130

[92] An integrated encyclopedia of DNA elements in the human genome. Nature, 489 (7414) (2012 Sep 6), pp. 57-74

[93] M.T. Maurano, R. Humbert, E. Rynes, R.E. Thurman, E. Haugen, H. Wang, *et al*. Systematic localization of common disease-associated variation in regulatory DNA. Science, 337 (6099) (2012 Sep 7), pp. 1190-1195

[94] R.E. Thurman, E. Rynes, R. Humbert, J. Vierstra, M.T. Maurano, E. Haugen, *et al*. The accessible chromatin landscape of the human genome. Nature, 489 (7414) (2012 Sep 6), pp. 75-82

[95] G. Vahedi, Y. Kanno, Y. Furumoto, K. Jiang, S.C. Parker, M.R. Erdos, *et al*. Super-enhancers delineate disease-associated regulatory nodes in T cells. Nature, 520 (7548) (2015 Apr 23), pp. 558-562

[96] D. Hnisz, B.J. Abraham, T.I. Lee, A. Lau, V. Saint-Andre, A.A. Sigova, *et al*. Super-enhancers in the control of cell identity and disease. Cell, 155 (4) (2013 Nov 7), pp. 934-947

[97] K.K. Farh, A. Marson, J. Zhu, M. Kleinewietfeld, W.J. Housley, S. Beik, *et al*. Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature, 518 (7539) (2015 Feb 19), pp. 337-343

[98] J. Ernst, P. Kheradpour, T.S. Mikkelsen, N. Shoresh, L.D. Ward, C.B. Epstein, *et al*. Mapping and analysis of chromatin state dynamics in nine human cell types. Nature, 473 (7345) (2011 May 5), pp. 43-49

[99] A. Kundaje, W. Meuleman, J. Ernst, M. Bilenky, A. Yen, A. Heravi-Moussavi, *et al*. Integrative analysis of 111 reference human epigenomes. Nature, 518 (7539) (2015 Feb 19), pp. 317-330

[100] P. Carninci, T. Kasukawa, S. Katayama, J. Gough, M.C. Frith, N. Maeda, *et al*. The transcriptional landscape of the mammalian genome. Science, 309 (5740) (2005 Sep 2), pp. 1559-1563

[101] E. Birney, J.A. Stamatoyannopoulos, A. Dutta, R. Guigo, T.R. Gingeras, E.H. Margulies, *et al*. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature, 447 (7146) (2007 Jun 14), pp. 799-816

[102] D. Bu, K. Yu, S. Sun, C. Xie, G. Skogerbo, R. Miao, *et al*. NONCODE v3.0: integrative annotation of long noncoding RNAs. Nucleic Acids Res., 40 (Database issue) (2012 Jan), pp. D210-D215

[103] D. Bhartiya, K. Pal, S. Ghosh, S. Kapoor, S. Jalali, B. Panwar, *et al*. lncRNome: a comprehensive knowledgebase of human long noncoding RNAs. Database (Oxford), 2013 (2013), p. bat034

[104] J.E. Wilusz, H. Sunwoo, D.L. Spector. Long noncoding RNAs: functional surprises from the RNA world. Genes Dev., 23 (13) (2009 Jul 1), pp. 1494-1504

[105] J. Zhao, B.K. Sun, J.A. Erwin, J.J. Song, J.T. Lee. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science, 322 (5902) (2008 Oct 31), pp. 750-756

[106] T.R. Mercer, M.E. Dinger, J.S. Mattick. Long non-coding RNAs: insights into functions. Nat. Rev. Genet., 10 (3) (2009 Mar), pp. 155-159

[107] K.C. Wang, H.Y. Chang. Molecular mechanisms of long noncoding RNAs. Mol. Cell, 43 (6) (2011 Sep 16), pp. 904-914

[108] J. Li, Z. Xuan, C. Liu. Long non-coding RNAs and complex human diseases. Int. J. Mol. Sci., 14 (9) (2013), pp. 18790-18808

[109] B. Gupta, R.D. Hawkins. Epigenomics of autoimmune diseases. Immunol. Cell Biol., 93 (3) (2015 Mar), pp. 271-276

[110] T. Gutschner, S. Diederichs. The hallmarks of cancer: a long non-coding RNA point of view. RNA Biol., 9 (6) (2012 Jun), pp. 703-719

[111] S. Carpenter, D. Aiello, M.K. Atianand, E.P. Ricci, P. Gandhi, L.L. Hall, *et al*. A long noncoding RNA mediates both activation and repression of immune response genes. Science, 341 (6147) (2013 Aug 16), pp. 789-792

[112] J. Ding, S. Eyre, J. Worthington. Genetics of RA susceptibility, what comes next? RMD Open. 2015; 1(1): e000028.

[113] H. Cui, N. Xie, Z. Tan, S. Banerjee, V.J. Thannickal, E. Abraham, *et al*. The human long noncoding RNA lnc-IL7R regulates the inflammatory response. Eur. J. Immunol., 44 (7) (2014 Jul), pp. 2085-2095

[114] A.P. Mao, J. Shen, Z. Zuo. Expression and regulation of long noncoding RNAs in TLR4 signaling in mouse macrophages. BMC Genomics, 16 (2015), p. 45

[115] N. Muller, F. Doring, M. Klapper, K. Neumann, D.M. Schulte, K. Turk, *et al*. Interleukin-6 and tumour necrosis factor-alpha differentially regulate lincRNA transcripts in cells of the innate immune system in vivo in human subjects with rheumatoid arthritis. Cytokine, 68 (1) (2014 Jul), pp. 65-68

[116] V. Kumar, H.J. Westra, J. Karjalainen, D.V. Zhernakova, T. Esko, B. Hrdlickova, *et al*. Human disease-associated genetic variation impacts large intergenic non-coding RNA expression. PLoS Genet., 9 (1) (2013), p. e1003201

[117] L.J. Jensen, M. Kuhn, M. Stark, S. Chaffron, C. Creevey, J. Muller, *et al*. STRING 8–a global view on proteins and their functional interactions in 630 organisms. Nucleic Acids Res., 37 (Database issue) (2009 Jan), pp. D412-D416

[118] D. Warde-Farley, S.L. Donaldson, O. Comes, K. Zuberi, R. Badrawi, P. Chao, *et al*. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res., 38 (Web Server issue) (2010 Jul), pp. W214-W220