Journal of Autoimmunity xxx (2015) 1–8

Contents lists available at ScienceDirect

Journal of Autoimmunity

journal homepage: www.elsevier.com/locate/jautimm

Review article

Immunogenetics of rheumatoid arthritis: Understanding functional implications

Tobias C. Messemaker [a, b], Tom W. Huizinga [a, **], Fina Kurreeman [a, *]

[a] Department of Rheumatology, Leiden University Medical Center, Leiden, Netherlands
[b] Department of Molecular Cell Biology, Leiden University Medical Center, Leiden, Netherlands

ARTICLE INFO

Article history:
Received 13 July 2015
Accepted 14 July 2015
Available online xxx

Keywords:
Rheumatoid arthritis
Autoimmunity
Genetics
GWAS
Non-coding RNA

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

* Corresponding author.
** Corresponding author.
E-mail addresses: t.w.j.huizinga@lumc.nl (T.W. Huizinga), b.a.s.kurreeman@lumc.nl (F. Kurreeman).

http://dx.doi.org/10.1016/j.jaut.2015.07.007
0896-8411/

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].

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

Please cite this article in press as: T.C. Messemaker, et al., Immunogenetics of rheumatoid arthritis: Understanding functional implications, Journal of Autoimmunity (2015), http://dx.doi.org/10.1016/j.jaut.2015.07.007