Review article

Genetics of rheumatoid arthritis: a new boost is needed in Latin American populations

Patricia Castro-Santos\textsuperscript{a}, Roberto Díaz-Peña\textsuperscript{a,b,*}

\textsuperscript{a} Facultad de Ciencias de la Salud, Universidad Autónoma de Chile, Talca, Chile
\textsuperscript{b} Department of Immunology, Hospital Universitario Central de Asturias, Oviedo, Spain

\begin{abstract}
Rheumatoid arthritis (RA) is an autoimmune inflammatory rheumatic disease which affects several organs and tissue, predominantly the synovial joints. Like many other autoimmune diseases, RA is a complex disease, where genetic variants, environmental factors and random events interact to trigger pathological pathways. Genetic implication in RA is evident, and recent advances have expanded our knowledge about the genetic factors that contribute to RA. An exponential increment in the number of genes associated with the disease has been described, mainly through gene wide screen studies (GWAS) involving international consortia with large patient cohorts. However, there are a few studies on Latin American populations. This article describes what is known about the RA genetics, the future that is emerging, and how this will develop a more personalized approach for the treatment of the disease. Latin American RA patients cannot be excluded from this final aim, and a higher collaboration with the international consortia may be needed for a better knowledge of the genetic profile of patients from this origin.

© 2015 Elsevier Editora Ltda. All rights reserved.
\end{abstract}

\begin{resumo}
A artrite reumatoide (AR) é uma doença reumática inflamatória autoimune que atinge vários órgãos e tecidos, predominantemente as articulações sinoviais. Como muitas outras doenças autoimunes, a AR é uma doença complexa, em que variáveis genéticas, fatores ambientais e eventos aleatórios interagem e desencadeiam vias patológicas. A implicação genética na AR é evidente e avanços recentes têm expandido nosso conhecimento sobre os fatores genéticos que contribuem para a doença. Houve um incremento exponencial na quantidade de genes associados à doença descritos, principalmente por estudos de associação genômica ampla (GWAS) que envolveram consórcios internacionais com grandes

© 2015 Elsevier Editora Ltda. All rights reserved.

\* Corresponding author.
\textsuperscript{E-mail: rdiazp@uautonoma.cl (R. Díaz-Peña).}
\textsuperscript{http://dx.doi.org/10.1016/j.rbre.2015.10.004}
\textsuperscript{2255-5021/© 2015 Elsevier Editora Ltda. All rights reserved.}
Rheumatoid arthritis

Rheumatoid arthritis (RA) is an autoimmune inflammatory rheumatic disease that affects many tissues and organs, mainly synovial joints. This disease leads to progressive destruction of articular cartilage and ankylosis of the joints.\(^1\) Subsequent, pannus formation may lead to destruction of underlying cartilage and bony erosions. RA diagnosis is based on clinical criteria and laboratory tests.\(^2\) Anti-citrullinated protein autoantibodies (ACPA) show a high specificity for RA, even ACPA testing has become a substantial component of the current American College of Rheumatology (ACR)-European League Against Rheumatism (EULAR) classification criteria for RA.\(^3\) Additionally, it has been described that ACPA may play a role in disease pathogenesis.\(^4\)

RA affects approximately 1% of the population worldwide.\(^5\) In the last years, several epidemiological studies of RA have been published, showing variations in the incidence and prevalence of RA across populations. Most of the studies have been developed in countries from the North Europe and North America, estimating prevalences of 0.5–1.1%.\(^5\) Another studies made mainly in countries from South Europe reported a lower prevalence around 0.3–0.7%.\(^6\)-\(^8\) The lowest prevalence data have been reported in areas from Africa and Asia, and the highest in Native American populations.\(^7\) In fact, the prevalence of RA is 10 times higher among Canadian or Native Americans than Europeans (3% and 0.3%, respectively).\(^9\)-\(^10\) Although the disease can develop at any age, RA affects females more frequently than males and it is diagnosed mainly in age 40–60 years, although the mechanism by which gender influences the susceptibility to RA remains unclear. Other characteristic of RA is heterogeneity: patients do not form a homogenous population and some clinical RA subgroups, such as ACPA seropositive versus seronegative, erosive versus non-erosive, progressive versus mild-course, have been identified.\(^11\)-\(^13\)

RA genetics and pathogenesis

Like many autoimmune diseases, the etiology of RA is multifactorial. Genetic susceptibility is evident in familial clustering and monozygotic twin studies, with a 50% of RA risk attributable to genetic factors, and heritability of RA has been estimated to be about 60%.\(^4\) Moreover, disease progression, outcome and RA phenotype have been associated with genetic factors.\(^11\)-\(^14\) Thus, understanding the genetics basis of RA is required in order to develop a more personalized approach for the disease treatment. RA genetic risk factors can be classified into two groups: (1) major histocompatibility complex (MHC) genes and (2) non-MHC regions. Interestingly, HLA and some non-HLA genes have been linked to the development of antibodies against citrullinated proteins, differentiating between two entities with distinctive characteristics, ACPA seropositive and seronegative RA.\(^15\) Interestingly, several genetic polymorphisms have been described associated to environmental factors in RA patients, primarily smoking.\(^16\) Smoking and possibly other environmental factors may trigger ACPA production and the development of ACPA seropositive RA (Fig. 1).\(^17\) Although the etiology of RA has not been elucidated yet, their symptoms develop gradually in different phases.\(^18\) In this development of the disease has been described a “preclinical phase”, in which several immunological markers, as ACPA or rheumatoid factor (RF), become positive sometimes years before of the onset of clinical symptoms. To sum up, RA develops in genetically pre-disposed individuals subjected to an unclear set of life events, specially smoking (Fig. 1).

HLA region

The genomic map of the human MHC (HLA) spans about 7.6 Mb and contains approximately 421 gene loci on a contiguous region on chromosome 6.\(^19\) The classical HLA loci, which play a central role in the immune system, are called -A, -B, and -C (class I) and -DRB1, -DQB1, and -DPB1 (class II). Particularly, the HLA class I and class II genes encode for proteins that bind to small antigen peptides and carry them into the cell surface thus presenting them to the immune system. Therefore, this genomic region is crucial for the organism resistance and susceptibility to pathogenic factors.

It has been 35 years since it was published that the HLA region contributes to RA susceptibility, specifically HLA-DR4 allele,\(^20\) but the exact mechanism that determines the predisposition is unknown. Among the HLA genes, the HLA-DRB1 shared epitope (SE) alleles that encode for a common amino acid sequence, is the most important risk factor described for RA susceptibility and progression.\(^21\) The presence of SE suggest that the HLA alleles containing it bind the same antigen, postulating the presentation of arthritogenic self-peptides or molecular mimicry with foreign antigens,\(^22\)-\(^23\) and/or shaping the T-cell-antigen repertoire.\(^24\) HLA-DRB1 SE alleles are strongly associated with ACPA-positive RA. Indeed, HLA-DRB1 SE alleles contribute in 18% to the heritability of ACPA-positive RA, whereas they only contribute in 2.4% to the heritability of ACPA-negative RA.\(^25\) The relationship between HLA-DRB1 SE
Fig. 1 – Hypothetical model for initiation of RA in ACPA-positive individuals. In an induction phase, environmental factors could contribute to stimulate the innate immunity. Apoptosis, necrosis, or both of some cells could cause citrullination in certain proteins in the lungs (due to the increase in the activity of peptidylarginine deiminases enzymes, PAD). Some of these modified proteins bind specifically to HLA-DR molecules on dendritic cells or macrophages resulting in high titers of ACPA. Citrullination proteins in the joints due to infection, trauma, exercise, etc., could lead to immune complex formation between modified proteins and ACPA, which further bind to Fc receptors on the surface of synovial macrophages, contributing to the perpetuation of inflammation.

and ACPA in the pathogenesis of RA has been explained by citrullinated peptide binding into the pocket of DRB1 molecules containing the shared epitope, and the consequent activation of CD4+ T cells and polarization to Th17 cells, a Th subpopulation involved primarily in autoimmune processes. HLA-DRB1 SE alleles are present in 64–70% of RA patients and in 55% of their first-degree relatives; this frequency is significantly higher to the one observed in control

Fig. 2 – Timeline of discovery of several genes associated with RA.
populations (35.8%).\textsuperscript{26,27} In ACPA-positive RA patients, 80% have at least one SE, while 49% of ACPA-negative RA patients have SE. This interaction among genetic risk factors and the presence of autoantibodies increase the risk of developing RA in first-degree relatives of RA patients.\textsuperscript{27,28}

The SE hypothesis remains controversial, because suggest the existence of an autoantigenic peptide that has not identified yet. Several other diseases, like Type 1 diabetes, psoriatic arthritis, lupus, early-onset chronic lymphoid leukemia, and other conditions,\textsuperscript{29} and this promiscuity are incongruous with tenets of MHC-restricted antigen presentation theory. Although HLA-DRB1 alleles containing the epitope are established genetic risk factors in RA, the precise immunological implications of their expression are not clear. Furthermore, it has been reported that shared epitope alleles at the HLA-DRB1 locus do not completely explain the association of the MHC region with the disease.\textsuperscript{30-32}

**Non-HLA genetic associations**

The pathogenesis of RA has a polygenic basis. About 50% of RA risk is thought to be genetic and one-third of this risk belongs to the HLA locus.\textsuperscript{3} Thus, genetic variation can be explained by RA risk alleles in non-HLA locus. There has been an exponential increase in the number of genes associated with RA in the last several years, as shown in Fig. 2. Specifically, in addition to the HLA-DRB locus, over 46 non-HLA RA risk loci have emerged from genome-wide association studies (GWAS) and subsequent GWAS meta-analysis of GWAS datasets,\textsuperscript{33,34} all of them in individuals of European ancestry. Another meta-analysis of GWAS in Japanese population was reported, which identified nine novel loci associated with RA.\textsuperscript{35} The cited study provided evidence of significant overlap in the RA genetic risks between Japanese and European population, contributing to further understanding of the RA etiology. GWAS are considered to be one of the primary tools for determining genetic links to diseases. These analyses have been abundant in recent scientific researches. In each of these studies at least 100,000 single nucleotide polymorphisms (SNPs) are genotyped, taking an unbiased view of the whole genome and therefore have a higher probability of detecting an association with a genetic marker, providing the studies with sufficient power. Fig. 2 captures the top regions of RA associations that are statistically significant. A recent study discovered 42 novel RA risk loci at a genome-wide level of significance, bringing the total to 101.\textsuperscript{36} These study genotyped around 10 million SNPs on a total of over 100,000 subjects of European and Asian ancestry, 29,880 RA patients and 73,758 controls. In summary, the researchers were able to establish 98 genes that could potentially contribute to the onset of RA. Many of these genes also play a role in other diseases, including human primary immunodeficiency disorders and blood cancers. They discovered many genes that overlap to contribute to the condition that are already being targeted by existing drugs, but was not known when the drugs were developed. This study provides evidences that genetic of diseases could contribute to biological insight and drug discovery.

**Situation in Latin American populations**

The study of genetics in Latin American populations is not a trivial topic. The expression of genetic variants is modified by many environmental factors, and the significance of ethnicity in genetics is controversial.\textsuperscript{37} Latin Americans have been wrongly designated as “Hispanics” and considered homogenous. Actually, the origins and destinations of non-Amerindian populations have depended on the time and reasons for the migration, and the degree of admixture varies between Latin American countries according to the major ancestry population component.\textsuperscript{38}

There are important challenges in finding susceptibility genes for RA in these populations. The Hispanic community is an admixed population, and the allelic frequency differences across ethnic groups can interfere with association studies and lead to false-positive results. Thus, in GWAS, candidate genes and replication studies of GWAS, differences found in the allele frequencies may be originated more by differences in the populations structure than by the phenotype of the disease. However, there are approaches to overcoming the problem of population structure, like use ancestry informative markers (AIMs),\textsuperscript{39} or include structured association test (structure) and principal component analysis for adjusting population stratifications in the studies.\textsuperscript{40,41}

Although a high progress has been made in detecting the genes implicated in RA susceptibility, little is known about genetic susceptibility in the “Hispanic” populations of the Americas. This is largely due to the difficulty of performing association studies in admixed populations and the fact that the power required to identify genetic associations in these populations is greater than in more homogeneous populations. Some AR association studies have been previously reported in Amerindian and mixed Hispanic populations, and the strongest association observed was in the HLA class II region. Specifically, genetic associations of RA with HLA-DRB1 alleles have been reported in Native Americans, Mexican American ancestry, Colombian population, Chilean population, Peruvian population, Brazilian population and Mexican Mestizo population with a larger proportion of European ancestry.\textsuperscript{42-48} A meta-analysis carried out across Latin American populations estimated the relevance of HLA-DRB1 alleles on RA susceptibility, confirming a significant association between RA and HLA-DRB1 gene and revalidating the shared epitope hypothesis in Latin American populations.\textsuperscript{49} A recent study examined susceptibility loci for RA in Latin American individuals with admixed European and Amerindian genetic ancestry.\textsuperscript{50} These study genotyped 196,524 markers, covering the previously associated loci with various autoimmune diseases, in 1,475 RA patients and 1,213 controls. A strong genetic association of RA with the MHC region was observed, with three independent effects, probably due to the diverse origin of the samples. In the same study,\textsuperscript{50} RA associations previously reported in GWAS (European and Asian populations) were found, but with moderate significant values (including STAT4, IRF5, IL2RA, SPRED2, CCL21 and PTPN22 genes). Additionally, two novel putative associations in ENOX1 gene on chromosome 13 and NNA25 gene on chromosome 12 were identified. The results of this large-scale association study provided new
perspectives into the RA genetic basis in Latin-Americans individuals. Several of these findings require replication and supply an impetus for future studies. Moreover, they provide interesting conclusions of the observed complexity of RA associations with HLA region, probably as a consequence of the origin diversity.

Genetic analyses undertaken in the recent years have revealed a new picture for RA pathogenesis and made us aware of heterogeneity among individuals and populations. Genomics research is advancing rapidly, through SNP genotyping and the next genome sequencing, two techniques that are improving our understanding of the RA etiopathogenesis. The final goal in the coming years is to identify genetic variants involved in the different clinical manifestations and RA-associated features, and thereby predict the evolution of the disease, and finally, to establish new treatments for RA based on the prognosis of individuals, enabling the development of personalized therapies for RA. Other aspects such as epigenetics and pharmacogenetics, require further investigation in order to establish any role they may have in RA. Regarding this last point, the final objective of pharmaco- genetics in rheumatology is to define genetically distinct patient subsets, which have differential responses to the various therapies used to treat rheumatic diseases. A vast growing body of literature describes the pharmacogenetics of drugs used in RA treatment. However, there are no data about RA pharmacogenetic in Latin-American populations.

Variations in the frequency of certain genotypes across ethnic groups may occur, and due to this, genetic association studies conducted in Latin American must have a powerful “control population”. The large and diverse population of Latin America is a powerful resource for elucidating the genetic basis of complex traits as RA.

Conclusions

Multicentric studies have shown a high relevance in the understanding of the risk genetic factors in complex diseases. Therefore, an additional effort in the search for unknown genetic predispositions and clarify differences in roles among ethnic groups, including Latin American populations, is needed. Investigation in genomics area has advanced very quickly through SNPs genotyping and GWAS, and will advance even more with the new massive sequencing techniques. In this way, a better knowledge about genetic basis of RA in Latin American populations undoubtedly would contribute to a better understanding of this disease pathology.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgment

This work was supported by Fondecyt, grant no. 11130198.

REFERENCES


42. Anaya JM, Correa PA, Mantilla RD, Arcos-Burgos M. Rheumatoid arthritis association in Colombian population is restricted to HLA-DRB1*04 QRAA alleles. Genes Immun. 2002;3:56–8.


