PTPN22: a confirmed rheumatoid arthritis susceptibility gene?

Anne Hinks
University of Manchester, Arthritis Research Campaign Epidemiology Unit, Manchester, M13 9PT, UK Tel.: +44 161 275 1669; Fax: +44 161 275 5043; Anne.Hinks@manchester.ac.uk

‘It will be important...to verify whether the function of the (PTPN22) gene in autoimmune disease susceptibility is the same for all diseases.’

2005 was a landmark year for research into the genetics of rheumatoid arthritis (RA). It was the year that a new confirmed susceptibility gene was identified, along with a new pathological pathway important in susceptibility to RA. The protein tyrosine phosphatase, nonreceptor type 22 (PTPN22) gene is a member of the protein tyrosine phosphatase (PTP) family of proteins, which are key regulatory proteins of the immune system. The PTPs work in conjunction with protein tyrosine kinases (PTKs), regulate the reversible phosphorylation of tyrosine residues and are fundamental in controlling many physiological processes [1,2].

RA, similar to other autoimmune diseases, is a complex genetic disease and dissecting the contributing genetic variants is challenging. The association of the shared epitope alleles of major histocompatibility complex, class II, DRβ1 (HLA-DRB1) with RA is well established and, until this year, has been the only consistently replicated RA susceptibility gene and the only region to be identified in all four whole genome screens. All of the studies found no consistent gender interaction for association with RA, or were located in linkage regions identified in whole genome screens. A missense SNP within the PTPN22 gene was found to be associated with RA in their initial discovery data set (p = 0.0007, OR: 1.65, 95% CI: 1.23–2.20) and also in a second replication data set (p = 2.1 x 10−8, OR: 1.97, 95% CI: 1.55–2.5) [7]. The polymorphism is a C→T substitution (rs2476601) at nucleotide position 1858 that results in a tryptophan (W) for arginine (R) transition at codon 620. Interestingly, an earlier independent association study on a different autoimmune disease, Type 1 diabetes (T1D), demonstrated an association with the same SNP [8].

The original association in RA has subsequently been replicated by many groups in all populations studied, including British [9,10], Spanish [11], Norwegian [12], Dutch [13], Canadian [14], New Zealand [15] and Finnish [16] populations. In all studies, the direction of the association is the same; there is an increase in T allele frequency in cases compared with controls, with the ORs for RA conferred by carriage of the T allele ranging from 1.38 to 2.04. Together, these data provide compelling evidence that PTPN22 is a genuine susceptibility gene for RA.

All of the PTPN22 and RA association studies have examined whether there are any potential interactions between PTPN22 and HLA-DRB1. In the first study, Begovich and colleagues performed conditional logistic regression to adjust for the HLA-DRB1 genotype and found that it had little impact on risk estimates [7]. Many other studies have also stratified the data according to shared epitope (SE) status and in all cases, no significant difference in allele or genotype frequencies between SE-positive and -negative cases has been observed [9,10]. All of these analyses suggest that PTPN22 is acting independently of HLA-DRB1.

However, in 2005, Begovich and colleagues performed a case-control association study of potential functional single nucleotide polymorphisms (SNPs) located in candidate genes for association with RA, or were located in linkage regions identified in whole genome screens. A missense SNP within the PTPN22 gene was found to be associated with RA in their initial discovery data set (p = 0.0007, OR: 1.65, 95% CI: 1.23–2.20) and also in a second replication data set (p = 2.1 x 10−8, OR: 1.97, 95% CI: 1.55–2.5) [7]. The polymorphism is a C→T substitution (rs2476601) at nucleotide position 1858 that results in a tryptophan (W) for arginine (R) transition at codon 620. Interestingly, an earlier independent association study on a different autoimmune disease, Type 1 diabetes (T1D), demonstrated an association with the same SNP [8].

The original association in RA has subsequently been replicated by many groups in all populations studied, including British [9,10], Spanish [11], Norwegian [12], Dutch [13], Canadian [14], New Zealand [15] and Finnish [16] populations. In all studies, the direction of the association is the same; there is an increase in T allele frequency in cases compared with controls, with the ORs for RA conferred by carriage of the T allele ranging from 1.38 to 2.04. Together, these data provide compelling evidence that PTPN22 is a genuine susceptibility gene for RA.

All of the PTPN22 and RA association studies have examined whether there are any potential interactions between PTPN22 and HLA-DRB1. In the first study, Begovich and colleagues performed conditional logistic regression to adjust for the HLA-DRB1 genotype and found that it had little impact on risk estimates [7]. Many other studies have also stratified the data according to shared epitope (SE) status and in all cases, no significant difference in allele or genotype frequencies between SE-positive and -negative cases has been observed [9,10]. All of these analyses suggest that PTPN22 is acting independently of HLA-DRB1.

Many studies have also examined the effect of gender, age at disease-onset, rheumatoid factor (RF) status, family history of disease and severity of disease course upon the RA–PTPN22 association. All of the studies found no consistent gender difference in the association of PTPN22 with RA. Some studies have shown that the association is
stronger in RA cases with a young age at disease-onset [9,10], whilst others found this had no effect on the association [11,14].

The issue that appears to be most contradictory across the different studies is from the results of the stratification according to RF status. The initial association study by Begovich and colleagues suggested that the association with PTPN22 was only found in the subgroup of patients that were RF positive [77]. This was also borne out when the researchers further increased their sample size [17] and in another RA cohort [10]. Conversely, a number of other studies have demonstrated an association in both RF-positive and -negative subgroups [9,14,15].

The issue here is that, as RF is present in 75–80% of RA cases, the RF-negative subgroup sample sizes in the different studies are modest. However, the RF-negative subgroup sample size was also used by Lee and coworkers and should have had sufficient power to detect an association with PTPN22. It is possible that there has been a misclassification of RF-positive RA cases into the RF-negative RA case subgroup in the studies where there is an association with PTPN22, possibly due to different RF detection methodologies, or to the time points, in terms of disease duration, where these tests were performed. Alternatively, these differences may reflect clinical heterogeneity across different populations. Additional studies of larger cohorts of RA cases and controls, as well as combined data or meta-analyses, will be required to give a better picture of which diseases are and are not associated with PTPN22 and its clinical presentation.

Since the original associations in both T1D and RA, there has been a flurry of reports of the association of PTPN22 with other autoimmune diseases. Diseases that show evidence for association with PTPN22 include RA [7,9–12,14–17], juvenile idiopathic arthritis (JIA) [9,12,16], T1D [8,18–23], Graves’ disease [21,24,25], systemic lupus erythematosus (SLE) [11,26,27], Hashimoto’s thyroiditis [28], autoimmune Addison’s disease [29], Sjögren’s syndrome [39], celiac disease [40,41], Crohn’s disease [14] and systemic sclerosis [S Parameshwar & J Worthington, Unpublished Data].

The function of the protein encoded by PTPN22 may provide clues to the role of PTPN22 in autoimmune diseases and the pathways that might be important. PTPN22 encodes the protein lymphoid-specific phosphatase (Lyp), which is composed of an N-terminal phosphatase domain and a noncatalytic C-terminal end containing several proline-rich motifs. It is expressed in hemopoietic tissues, thymus, spleen and bone marrow, as well as in all subtypes of peripheral blood mononuclear cells, including T and B cells, monocytes, neutrophils and natural killer cells [7]. Initial studies of the function of the protein have largely been determined through experiments on the mouse homolog, PEST domain-enriched tyrosine phosphatase (PEP), which is encoded by the gene PTPN8. Cloutier and Veilette showed that the PTPN8 missense SNP is not universally associated with autoimmune diseases and a better picture of which diseases are and are not associated with PTPN22 will provide a better understanding of different autoimmune disease pathways. Diseases that, to date, do not appear to be associated with PTPN22 include multiple sclerosis (MS) [9,28,36,37], psoriasis [9,38], psoriatic arthritis [9], Sjögren’s syndrome [39], celiac disease [40,41], Crohn’s disease [14] and systemic sclerosis [S Parameshwar & J Worthington, Unpublished Data].

The fact that PTPN22 is associated with numerous autoimmune diseases lends support to the hypothesis that there are common pathophysiological pathways underlying certain autoimmune diseases. The hypothesis of overlapping susceptibility to autoimmune diseases is not a new one, as it was originally proposed following observations of clustering of different autoimmune diseases within families. This has been confirmed through epidemiological studies that have observed an increased statistical risk that an individual with an autoimmune disease will have a blood relative with either the same or another autoimmune disease [30–33]. The hypothesis has been strengthened further in recent years, following whole genome screens of different autoimmune diseases. Analysis of the results of autoimmune disease whole genome screens identified 18 distinct clusters, with evidence of linkage to two or more autoimmune or immune-related diseases [34]. This overlap of loci has also been observed in mouse models of autoimmunity [35].

‘The association of the PTPN22 gene with RA is the first widely replicated genetic association with this disease since HLA.’

However, it is also now becoming clear that the PTPN22 missense SNP is not universally associated with autoimmune diseases and a better picture of which diseases are and are not associated with PTPN22 will provide a better understanding of different autoimmune disease pathways. Diseases that, to date, do not appear to be associated with PTPN22 include multiple sclerosis (MS) [9,28,36,37], psoriasis [9,38], psoriatic arthritis [9], Sjögren’s syndrome [39], celiac disease [40,41], Crohn’s disease [14] and systemic sclerosis [S Parameshwar & J Worthington, Unpublished Data].

The function of the protein encoded by PTPN22 may provide clues to the role of PTPN22 in autoimmune diseases and the pathways that might be important. PTPN22 encodes the protein lymphoid-specific phosphatase (Lyp), which is composed of an N-terminal phosphatase domain and a noncatalytic C-terminal end containing several proline-rich motifs. It is expressed in hemopoietic tissues, thymus, spleen and bone marrow, as well as in all subtypes of peripheral blood mononuclear cells, including T and B cells, monocytes, neutrophils and natural killer cells [7]. Initial studies of the function of the protein have largely been determined through experiments on the mouse homolog, PEST domain-enriched tyrosine phosphatase (PEP), which is encoded by the gene PTPN8. Cloutier and Veilette showed that there is a synergistic relationship between a PTP (PEP) and a PTK (C-terminal Src tyrosine kinase [Csk]), which together inhibit signaling of Src family kinases, such as Lyk, Fyn and ZAP-20, and thus mediate T-cell inactivation [42]. They demonstrated that PEP inhibited T-cell receptor signaling by dephosphorylation
The R620W SNP is situated in the proline-rich P1 domain of PTPN22, which, from the work of Cloutier and Villette, is known to bind to the SH3 binding domain of Csk [42]. Functional studies have attempted to establish whether the substitution affects binding to Csk. Immunoprecipitation analysis suggested that the R620 variant binds Csk, whereas the W620 has reduced binding with Csk than individuals who are heterozygous. Therefore, it is likely that PTPN22 is important in setting thresholds for T-cell receptor signaling. Defective binding of Lyp to Csk would, in effect, lead to lower thresholds for T-cell activation, overall increased activity of the immune system and thus a greater potential of mounting an autoimmune response.

However, recent data provided evidence to the contrary, suggesting that carriage of the R620W variant is associated with more efficient inhibition of T-cell activation. This contradicts the more simplistic model of autoimmunity, which would predict that T cells with defects in T-cell activation would be likely to cause disease. It has been speculated that the increased efficiency of PTPN22 Trp620 to inhibit TCR signaling may lead to weaker signaling and a failure to delete autoreactive T cells during thymic selection or insufficient activity of T regulatory cells [44].

Recent data proposing the gain-of-function mutation suggests a role of T regulatory cells in susceptibility. Experiments on collagen-induced arthritis (CIA) in mice suggests that intrinsic CD25+ regulatory T cells modulate the severity of the disease, which is an accepted model of RA, and that adoptive transfer of these cells can be used for the treatment of CIA [45].

All of the functional data so far has focused on T cells; however, it is also evident that PTPN22 is expressed in a wide range of hematopoietic cell types, including neutrophils, macrophages and natural killer cells, and the role of PTPN22 in these cell types has yet to be established, leaving an array of potential disease mechanisms still to be explored [7].

Another hypothesis is related to the role of PTPN22 in the humoral immune response. This hypothesis has originated from the initial nonassociation reports, where it appeared that the diseases that were not associated with PTPN22 were diseases that are not classically associated with circulating autoantibodies (MS, psoriasis, psoriatic arthritis and celiac disease). This hypothesis has originated from the initial nonassociation reports, where it appeared that the diseases that were not associated with PTPN22 were diseases that are not classically associated with circulating autoantibodies (MS, psoriasis, psoriatic arthritis and celiac disease).

The development of autoantibodies in diseases such as RA, T1D and autoimmune thyroid disease (AITD) often predates the development of overt clinical disease by a number of years and autoantibodies are also observed in normal healthy individuals. It will be interesting to examine whether PTPN22 is also associated with autoantibody production in normal individuals. However, we should be cautious, as there is a number of observations that challenge this data. First, a number of studies, including our study, found that association with PTPN22 is also significant in RF-negative RA cases and in antinuclear antibody-negative JIA cases. Second, two autoimmune diseases that do
not appear to be associated with PTPN22 are systemic sclerosis and Sjögren's syndrome, two diseases that are typically associated with the presence of autoantibodies.

Despite the convincing functional data supporting the PTPN22 R620W variant as the causal SNP in these associations with autoimmune diseases, only one study so far has investigated the possibility that the observed association is due to linkage disequilibrium (LD) with another SNP, or indeed that there are additional PTPN22 variants associated with disease but independent of the R620W variant. Carlton and colleagues reported the most thorough investigation of PTPN22 to date. The coding regions of the PTPN22 gene were screened for previously unidentified SNPs and these were then genotyped along with all other known SNPs from public databases. Strong LD across the gene and ten common haplotypes (frequency >1%) were identified. Only one of these haplotypes contained the W620 risk allele and this haplotype is strongly associated with RA. Another haplotype that was identical at all other SNPs, apart from carrying the R620 allele, showed no association, providing evidence that the W620 allele is the disease-predisposing allele on this haplotype. However, there was evidence that the R620W SNP does not explain all the association with RA, as there were three SNPs, rs1310182, rs3811021 and rs3789604, which were associated with RA independently of R620W. The latter two SNPs are in complete LD, suggesting they represent a single association, whilst the other cannot be excluded as a disease-predisposing locus, its association with RA is not independent of both R620W and rs3789604 [46]. It will be a priority to analyze these other variants in other RA populations, but it would also be interesting to genotype these variants in the autoimmune diseases that showed no association with the R620W variant. Studies of the wider region around PTPN22 found that the PTPN22 associated SNP lies within a 293 kb LD block of 41 haplotype map SNPs. At least six other known genes map to that LD block and on the National Center for Biotechnology Information SNP database there are 625 known SNPs that fall within that region [101]. One of these SNPs was found to be in complete LD (R² = 1) with the R620W variant, meaning that it is impossible to distinguish the effects of these SNPs using genetic data alone. The SNP was situated between two genes and was predicted to disrupt potential transcription factor binding sites [21]. Therefore, whilst it cannot be assumed with complete certainty that the PTPN22 R620W variant is the one and only autoimmune-predisposing SNP in the region, the functional data and the PTPN22 haplotype analysis provide compelling evidence that it plays a significant role.

The association of the PTPN22 gene with RA is the first widely replicated genetic association with this disease since HLA. It has been a success story in the field of rheumatological genetics and in the wider spectrum of autoimmune disease. It will be important to understand the finer details of its role in immune regulation and to verify whether the function of the gene in autoimmune disease susceptibility is the same for all diseases. There are a number of potential mechanisms through which it functions and these will require exploration in more detail. The association with PTPN22 is only one component of the complex genetic and environmental pathogenesis of diseases such as RA; however, a greater understanding of the role of PTPN22 and its pathway in these diseases may lead to the discovery of other interacting susceptibility genes and potential therapies.

Bibliography

8. Bottini N, Musumeci L, Alonso A et al.: A functional variant of lymphoid tyrosine phosphatase is associated with Type I
PTPN22: a confirmed rheumatoid arthritis susceptibility gene? – EDITORIAL

20. Sedin M F, Shigeta R, Lahko K et al.: Finnish case-control and family studies support PTPN22 R620W polymorphism as a risk factor in rheumatoid arthritis, but suggest only minimal or no effect in juvenile idiopathic arthritis. Genes Immun. 6(8), 702–722 (2005).

www.futuremedicine.com


