The association of \textit{PTPN22} with rheumatoid arthritis and juvenile idiopathic arthritis

Despite the wealth of evidence to support the involvement of multiple genetic factors in rheumatoid arthritis (RA), since the identification of the link between RA and HLA class II genes over 30 yr ago no single convincing non-HLA gene has emerged. Finally last year came a breakthrough with the association of a single-nucleotide polymorphism (SNP) in a candidate gene with RA, a finding replicated in every one of the eight RA cohorts subsequently examined \cite{1–9}. The gene is \textit{PTPN22}, a negative regulator of T-cell activation and the polymorphism is a C→T substitution (rs2476601) at nucleotide position 1858 that leads to a tryptophan (W) for arginine (R) transition at codon 620. This has a demonstrable functional effect, reducing the interaction between the protein and C-terminal Src tyrosine kinase (Csk) and resulting in an increased risk of inappropriate T-cell activation \cite{1, 10}.

What makes the finding particularly interesting is that it is not unique to RA but common to a number of autoimmune diseases, including type 1 diabetes (T1D) \cite{8, 10–15}, in which the association was first observed, Graves’ disease \cite{14, 16, 17} and Hashimoto’s thyroiditis \cite{18}. From a rheumatological point of view the association is also found in systemic lupus erythematosus (SLE) \cite{4, 19, 20} and juvenile idiopathic arthritis (JIA) \cite{2, 5}, suggesting some commonality of pathobiology for these very distinct conditions.

Like so many important findings in biology, the \textit{PTPN22} story emerged almost simultaneously from two groups using completely different approaches. A candidate gene strategy was adopted by Bottini \textit{et al.} \cite{10}, and led them to identify the protein tyrosine phosphatases (PTPs) as potential candidate genes for T1D and resulted in their investigation of the missense SNP in \textit{PTPN22} for association with this disease. This initial study found that the T allele was increased in T1D cases compared with healthy controls in two populations. In contrast, Begovich \textit{et al.} \cite{1} used a screening strategy examining 16 000 potential functional SNPs in genes that were candidates for RA or were located in linkage regions identified in RA whole-genome screens. Remarkably, the same missense SNP within the \textit{PTPN22} gene was the most significantly associated marker with RA in their initial ‘discovery’ data set of 475 cases ($P = 0.0005$; odds ratio for carriage of the T allele $= 1.73$, 95% confidence interval $= 1.27–2.38$). This association with RA has subsequently been replicated by many groups studying different populations, including UK \cite{2, 3}, Spanish \cite{4}, Norwegian \cite{5}, Canadian \cite{6}, New Zealand \cite{7}, Dutch \cite{8} and Finnish \cite{9} 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, the odds ratios for RA conferred by carriage of the T allele ranging from 1.38 to 2.04 (Fig. 1). Together, these data provide compelling evidence that \textit{PTPN22} is a true susceptibility gene for RA. However, many questions remain to be answered before we will fully understand how carriage of the \textit{PTPN22} risk allele influences the disease process in RA and other autoimmune diseases.

It is generally accepted that the aetiology of complex disease will only be fully explained once we understand the way in which multiple genetic and non-genetic factors interact. It is therefore not surprising that all the RA studies to date have examined the relationship between \textit{PTPN22} and \textit{HLA-DRB1}.  

![Fig. 1. Odds ratio plots for carriage of the PTPN22 T allele with RA.](image)
Begovich et al. [1] performed conditional logistic regression to adjust for the HLA-DRB1 genotype and found that it had little impact on risk estimates. Many of the other studies have also stratified their data according to shared epitope (SE) status, and in all cases no significant difference in allele or genotype frequencies between SE-positive and SE-negative cases has been observed [2, 3]. All these analyses suggest that PTPN22 is acting independently of HLA-DRB1, although none of the studies has specifically addressed the question of interaction.

All the studies have also investigated the effect of phenotypic heterogeneity upon the RA PTPN22 association, the most interesting but controversial observation relating to RF status. Begovich et al. [1] suggested that the association with PTPN22 was only in the subgroup of patients that were RF-positive [1]; this was borne out when they increased their sample size further [21]. Another study found that the frequency of the T allele was higher in RF-positive cases compared with RF-negative cases; however, the number of cases in the RF-negative data set was small (n = 58) [3]. Conversely, some studies have found association in RF-negative subgroups [2, 6, 7] and one study found no significant difference in the percentage of cases with RF in the TT and CT cases compared with the CC cases [4]. The inconsistencies may reflect the fact that in most studies the sample size of the RF-negative subgroup is modest (n = 145–207), although the study by Lee et al. [21] of 445 RF-negative cases also failed to find an association. Alternatively, the discrepancies may simply reflect clinical heterogeneity across different populations. Another concern may be possible misclassification of RF-positive RA cases as RF-negative, perhaps because of different RF detection methods or to the time points, in terms of disease duration, when these tests are performed.

The association of PTPN22 with RA led to the investigation of the gene in other rheumatic diseases, particularly JIA. Studies of Norwegian [5], UK [2] and Finnish [9] JIA cohorts all provide some evidence of association with PTPN22. The largest study of 661 UK cases and 595 controls revealed highly significant association of the PTPN22 missense SNP and JIA (P = 0.0005; odds ratio for carriage of the T allele = 1.6, 95% confidence interval = 1.23–2.09) [2]. JIA is an even more heterogeneous condition than RA, and consequently analysis of the clinically defined International League of Associations for Rheumatology (ILAR) subgroups was particularly revealing. Two observations stand out: firstly, systemic-onset JIA, often considered the most clinically distinct of the ILAR subgroups, showed no association with PTPN22 (P = 0.83; odds ratio for carriage of the T allele = 0.91, 95% confidence interval = 0.5–1.5), suggesting that the molecular pathway involving the PTPN22 variant is not important in disease pathogenesis for this subgroup. Secondly, in contrast with some of the studies of RA, the strongest association was observed with the RF-negative polyarthritis subgroup (P = 0.00009; odds ratio for carriage of the T allele = 2.4, 95% confidence interval = 1.6–3.6). The persistent and extended oligoarthritis subgroups, which are also RF-negative, also showed an increase in T-allele frequency in cases compared with controls. A similar finding was also observed by Viken et al. [5].

Since the original associations reported by both Bottini et al. [10] and Begovich et al. [1], there have been a flurry of reports on the association of PTPN22 variants with other autoimmune diseases, but it is becoming clear that the gene is not important for all autoimmune diseases. Evidence of association with PTPN22 has been reported for TID [8, 11–15], Graves’ disease [14, 16, 17], SLE [4, 19, 20], Hashimoto’s thyroiditis [18] plus, in two small studies, Addison’s disease [16] and generalized vitiligo [22]. Diseases that, to date, appear not to be associated with PTPN22 include multiple sclerosis [2, 18, 23, 24], psoriasis [2, 25], psoriatic arthritis [2], Sjögren’s syndrome [26], coeliac disease [8, 27], Crohn’s disease [6] and systemic sclerosis (Parameshwar and Worthington, unpublished data). These findings will have great importance in exploring the pathogenic mechanisms that underlie the different autoimmune diseases.

The function of the protein encoded by PTPN22 may give us clues to the role of PTPN22 in autoimmune disease and the pathways that might be important. PTPN22 encodes the protein lymphoid-specific phosphatase or Lyp, which is composed of an N-terminal phosphatase domain and a non-catalytic C-terminal end containing several proline-rich motifs. It is expressed in haemopoietic 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 [1]. The function of the protein has largely been investigated through experiments on the mouse homologue, PEST domain-enriched tyrosine phosphate (PEP), which is encoded by the gene Ptnph. Cloutier and Veillette [28] showed that there is a synergistic relationship between a protein tyrosine phosphate (PTP) (PEP) and a protein tyrosine kinase (PTK) (Csk), which together inhibit signalling of Src family kinases such as Lyk, Fyn and ZAP-70 and thus mediate T-cell inactivation. They showed that PEP inhibited T-cell receptor signalling by dephosphorylation of the Src family kinases but that this was dependent on the interaction of PEP with the SH3 domain of Csk. The importance of this pathway in RA disease aetiology is supported by the observation that a mutation in the Zap-70 gene causes a spontaneous autoimmune arthritis in mice [29]. A PEP knockout mouse on a non-autoimmune background has been created and observations of the phenotype have confirmed the role of PEP in T-cell function. There was evidence of enlargement of spleen and lymph nodes, and increased numbers of effector/memory T cells, particularly in older mice. There is also evidence for its role in the humoral immune response, with the development of germinal centres and increased concentration of certain antibody isotypes. There was no evidence of increased levels of autoantibodies, however, and no signs of overt autoimmune disease. Therefore, other initiators, genetic or environmental, are necessary for the development of autoimmune disease in these mice [30].

The associated SNP is situated in the proline-rich P1 domain of PTPN22, which binds to the SH3-binding domain of Csk [28]. Immunoprecipitation analysis suggested that the R620W variant binds Csk whereas the W620 has reduced binding [1, 10]. It is proposed that this in turn leads to reduced ability to down-regulate T-cell activation. It has been observed in some of the PTPN22 association studies that there is a dosage effect and that individuals homozygous for the PTPN22*T allele would have a more severely reduced binding with Csk than would individuals who are heterozygous. Therefore, it is likely that PTPN22 is important in setting thresholds for T-cell receptor signalling. Defective binding of Lyk 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 for mounting an autoimmune response.

Despite the very convincing functional data supporting the PTPN22 R620W variant as the causal SNP in autoimmune disease, the association could arise as a result of linkage disequilibrium with another SNP, or indeed there may be additional PTPN22 variants associated with disease but independent of the R620W variant. In the most recent study of PTPN22 and RA, 36 SNPs have been identified and 10 haplotypes (frequency > 1%) characterized. The 185ST allele was present on only one haplotype and completely accounted for the association of that haplotype with RA. There was evidence, however, that three additional SNPs were associated with RA independently of R620W [31]. It will be a priority to analyse these variants in other RA populations and also in JIA, and perhaps particularly in the autoimmune diseases that showed no association with the R620W variant. Work in diabetes has established that PTPN22 lies within a 293 kb linkage disequilibrium block to which at least six other known genes map, emphasizing the potential complexity of SNP associations to this region [15]. Existence of multiple genetic factors would be consistent with studies in murine collagen-induced arthritis, in which analysis of the linked region containing
the PTPN22 mouse homologue, Ptpn8, also implicates a role for other genes [32].

How the PTPN22 gene specifically contributes to RA and JIA susceptibility is still unclear, but there are a number of hypotheses. Firstly, its role in T-cell inactivation is attractive, knowing the probable role of T cells in RA and JIA pathogenesis. The characteristic pathology of both RA and JIA is of infiltration of the joint space by immune cells, largely CD4+ T cells, B cells and macrophages. Carriage of the associated PTPN22*6 allele, the consequent defect in interaction of Lyp with Csk and less effective down-regulation of T-cell activation may translate to hyperactive T cells within the synovial joint with a greater potential for mounting an autoimmune response. The observation that PTPN22 association is stronger in autoimmune cases with a younger age at onset may reflect this T-cell hyperactivity, with lower thresholds for an autoimmune response and hence earlier disease onset.

Another hypothesis revolves around the role of PTPN22 in the humoral immune response. This has stemmed from the initial non-association reports, where it appeared that the diseases that were not associated with PTPN22 were diseases that are not classically associated with circulating autoantibodies (multiple sclerosis, psoriasis, psoriatic arthritis and coeliac disease). This interpretation was also concurrent with the observations that the association in RA appeared in some studies to be restricted to RA cases that were autoantibody RF-positive. Further evidence from animal models suggests that PTPN22 is also important in B-cell function, with increased antibody levels and also increased numbers of germinal centres observed in the PEP knockout mouse. This has led to speculation that PTPN22 may be associated with the generation of disease-associated autoantibodies. The development of autoantibodies in diseases such as RA, TID and autoimmune thyroid disease 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. We should be cautious, however, as there are a number of observations that challenge these data. Firstly, a number of studies, including our own, found that association with PTPN22 is significant in RF-negative RA cases and also in ANA-negative JIA cases. This is in contrast with the original PTPN22 and RA study of US cases, which did not detect an association with the RF-negative subgroup. Secondly, 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.

It is also evident that PTPN22 is expressed in a wide range of haemopoietic cell types, including neutrophils, macrophages and natural killer cells [1], and the role of PTPN22 in these cell types has yet to be established. This leaves an array of potential disease mechanisms yet to be explored. The association of the PTPN22 gene with RA and JIA is the best-replicated genetic association with these diseases since HLA. It is certainly intriguing that a single polymorphism in this gene appears to be of importance not only in RA and JIA but also in a number of other diseases. Perhaps this fact will ensure rapid progress in the dissection of the pathways involving this gene and a greater understanding of autoimmune diseases.

The authors have declared no conflicts of interest.

References


