*Am. J. Hum. Genet.* 75:504–507, 2004

# Report

## Genetic Association of the R620W Polymorphism of Protein Tyrosine Phosphatase PTPN22 with Human SLE

Chieko Kyogoku,[1*] Carl D. Langefeld,[2*] Ward A. Ortmann,[1] Annette Lee,[3] Scott Selby,[1] Victoria E. H. Carlton,[4] Monica Chang,[4] Paula Ramos,[1] Emily C. Baechler,[1] Franak M. Batliwalla,[3] Jill Novitzke,[1] Adrienne H. Williams,[2] Clarence Gillett,[1] Peter Rodine,[1] Robert R. Graham,[5] Kristin G. Ardlie,[6] Patrick M. Gaffney,[1] Kathy L. Moser,[1] Michelle Petri,[7] Ann B. Begovich,[4] Peter K. Gregersen,[3] and Timothy W. Behrens[1]

[1]Department of Medicine, University of Minnesota School of Medicine, Minneapolis; [2]Section on Biostatistics, Wake Forest University School of Medicine, Winston-Salem, NC; [3]Robert S. Boas Center for Genomics and Genetics, North Shore Long Island Jewish Research Institute, Manhasset, NY; [4]Celera Diagnostics, Alameda, CA; [5]Department of Medicine, Massachusetts General Hospital, Boston; [6]Genomics Collaborative, Inc. (GCI), Cambridge, MA; and [7]Department of Medicine, Johns Hopkins University School of Medicine, Baltimore

**We genotyped 525 independent North American white individuals with systemic lupus erythematosus (SLE) for the PTPN22 R620W polymorphism and compared the results with data generated from 1,961 white control individuals. The R620W SNP was associated with SLE (genotypic $P = .00009$), with estimated minor (T) allele frequencies of 12.67% in SLE cases and 8.64% in controls. A single copy of the T allele (W620) increases risk of SLE (odds ratio [OR] = 1.37; 95% confidence interval [CI] 1.07–1.75), and two copies of the allele more than double this risk (OR = 4.37; 95% CI 1.98–9.65). Together with recent evidence showing association of this SNP with type 1 diabetes and rheumatoid arthritis, these data provide compelling evidence that PTPN22 plays a fundamental role in regulating the immune system and the development of autoimmunity.**

Systemic lupus erythematosus (SLE [MIM 152700]) is a chronic and severe systemic autoimmune disease associated with high titers of antinuclear antibodies and clinical involvement of many different organs and tissues, including skin, kidney, lungs, heart, and brain. SLE affects ~0.1% of the North American population, and women are nearly 10 times more frequently affected with disease than men. Epidemiologic evidence, together with recent linkage and association studies, suggest that SLE susceptibility in humans is strongly influenced by genetic factors (Wakeland et al. 2001). Similarly, studies of lupus-prone mice demonstrate the importance of genes in driving the onset, progression, and end-organ targeting of SLE (Wakeland et al. 2001). Abnormalities in B and T lymphocytes are found frequently in patients with SLE, and many studies have documented defects in proximal signal transduction downstream of the T- (TCR) and B-cell receptors (Kammer et al. 2002; Khan et al. 2003).

Begovich and colleagues (2004) recently discovered that a missense polymorphism (rs2476601; 1858C→T) in protein tyrosine phosphatase N22 (PTPN22, a key molecule regulating TCR signaling in memory/effector T lymphocytes [Hasegawa et al. 2004]), was strongly associated with human rheumatoid arthritis (RA). The polymorphism occurs in the proximal proline-rich SH3-binding domain of PTPN22, which results in substitution of a highly conserved arginine with tryptophan (R620W). This proline-rich region is an important docking site for C-terminal Src tyrosine kinase (CSK) (Cloutier and Veillette 1996), a molecule that downmodulates TCR signaling by phosphorylating regulatory tyrosines on the Src family kinase Lck (Cloutier and Veillette 1999). In vitro experiments show that the R620W polymorphism affects the ability of PTPN22 to bind CSK (Begovich et al. 2004; Bottini et al. 2004). Since this SNP has also independently been associated with type 1 diabetes (T1D) (Bottini et al. 2004) and previous data have linked defi-

Received May 5, 2004; accepted for publication June 23, 2004; electronically published July 23, 2004.

Address for correspondence and reprints: Dr. Timothy W. Behrens, Department of Medicine, University of Minnesota School of Medicine, 6-126 Basic Sciences Biomedical Engineering Building, 312 Church Street SE, Minneapolis, MN 55455. E-mail: behre001@umn.edu

* These two authors contributed equally to this work.


0002-9297/2004/7503-0016$15.00

ciency in phosphatase activity with lupus in mice (Schultz et al. 1993; O'Keefe et al. 1996; Di Cristofano et al. 1999), we investigated association of the R620W SNP with SLE.

Initially, we compared R620W genotype frequencies of cohort A, in which a single subject with SLE was randomly selected from each of 185 white families with SLE-affected sib pairs from the University of Minnesota collection (Gaffney et al. 2000), with results generated from 926 white controls (control 1) (Begovich et al. 2004). The C/T genotype was observed in 38 of 185 (20.5%) SLE cases and the T/T genotype in 6 (3.2%) cases (table 1). Compared with controls, the C/T and T/T genotypes were significantly overrepresented in the SLE cases ($P = .0109$, Fisher's exact test). Similar results were obtained when genotypes from a second affected subject from the families were examined and compared with the same controls (cohort B; $n = 180$; C/T 16.1%, T/T 4.4%; $P = .0060$). In these two sample sets, the presence of the T allele increased the odds of SLE by 1.46 (cohort A) and 1.10 (cohort B). Although we emphasize the need for caution in drawing conclusions from these results because of the limited number of genotypes, the odds ratios (ORs) for T/T homozygotes were 3.66 in cohort A and 4.81 in cohort B and were significant in a recessive model (table 1).

We then confirmed this finding using two independent replication cohorts. The first was a collection of 201 white individuals with SLE (cohort C), recruited at the University of Minnesota, from across North America as part of a trio family collection. Overall, this group showed very comparable PTPN22 R620W genotype frequencies (22.3% C/T and 1.0% T/T) to those observed in cohorts A and B. We also tested allele frequencies in a second independent control group (control 2 [$n = 1,035$]), which showed very similar allele frequencies compared with the control 1 group. When cohort C was compared with the control 2 group, association of the SNP with SLE was significant ($P = .0356$). A second replication group was an independent collection of 139 whites with SLE (cohort D), derived from the Hopkins Lupus Cohort (Petri 2000). Again, the R620W genotypic frequencies observed ($n = 139$; 17.3% C/T, 3.6% T/T) were very comparable to those from the Minnesota cohorts; when compared with the control 2 frequencies, the R620W SNP was associated with SLE ($P = .0015$).

Combination of data from the various independent case cohorts and comparison with the controls provided us with increased precision in our estimates of effect size and statistical power for testing the hypothesized relationship between R620W and SLE. Joint analysis of cohorts A, C, and D (525 individuals with SLE) yielded the following estimates for SLE genotype frequencies: 20.4% C/T, 2.5% T/T ($P = .00009$; compared with the combined control groups: 16.1% C/T and 0.6% T/T). Examination of the ORs in the combined data set again suggested a dose effect, with heterozygotes at increased risk relative to C/C homozygotes (OR = 1.37; 95% CI 1.07–1.75) and T/T homozygotes, with more than twice the risk of heterozygotes (OR = 4.37; 95% CI 1.98–9.65). Very similar results were obtained when cohorts B, C, and D were combined and analyzed together (B+C+D [table 1]).

The overall risk-allele frequency of R620W in 525 individuals with SLE (1,050 chromosomes, 133 T alleles) was 12.67%, compared with an allele frequency of 8.64% in 1,961 white control individuals (3,922 chromosomes, 339 T alleles; $P < .0001$) The risk allele was present in 22.8% of individuals with SLE, compared with 16.7% of control individuals.

After adjustment for age and sex, we found no significant differences in frequencies of lupus subphenotypes in subjects carrying one or more copies of the risk allele, compared with individuals lacking the risk allele (as defined by the criteria used to diagnose SLE: malar rash, discoid rash, arthritis, oral ulcers, serositis, renal disorder, CNS disorder, hematologic disorder, anti-dsDNA, Sm or anticardiolipin antibodies, and antinuclear antibodies). The risk allele is less common in African American and Hispanic/Latino populations than in North American whites (Begovich et al. 2004), and our current collection is not sufficiently powered to assess the possible influence of R620W on SLE in these populations. Finally, we note that there is currently no significant evidence for linkage at the 1p13 locus in the Minnesota family collection (Gaffney et al. 2000). The transmission/disequilibrium test (TDT), when applied to combined cohorts A and C, showed 70 transmissions and 57 nontransmissions of the T allele from heterozygous founders, a result that did not reach significance ($P = .22$). The transmission: nontransmission ratio of 1.23 is consistent with the OR for this allele (1.37 for C/T heterozygotes). Given an allele frequency in the subject population of 12.67% and ORs of 1.37 (1.07–1.75) for C/T heterozygotes and 4.37 (1.98–9.65) for T/T homozygotes, 229 trios (66–2,578) are required for 80% power at $P < .05$. Thus, the current trio collection is likely underpowered to detect the effect by TDT.

These data, together with the recent evidence for association of R620W with T1D (Bottini et al. 2004) and RA (Begovich et al. 2004), suggest that the minor allele is a potent genetic risk factor for both organ-specific (T1D) and systemic (RA and SLE) autoimmune syndromes. Knockout mice deficient for PTPN22 show selective dysregulation in the effector/memory T-cell compartment, with hyperproliferation and exaggerated early signaling responses in restimulated T cells, compared with essentially normal responses in naive T cells (Hasegawa et al. 2004). PTPN22 knockouts also demonstrated high-level spontaneous germinal center formation and ele-

## Table 1

### Elevated Frequency of PTPN22 R620W C/T and T/T Genotypes in North American White Patients with SLE

| Sample | No. of Individuals | No. of Individuals with Genotype (Genotype Frequency) | | | | OR for Genotype[a] (95% CI) | | Genotypic Association[b] | OR for Model (95% CI) [$P^c$] | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | C/C | C/T | T/T | C/T or T/T | C/T | T/T | *P* Value | Dominant | Additive | Recessive |
| Control 1[d] | 926 | 774 (83.6%) | 143 (15.4%) | 9 (1.0%) | 152 (16.4%) | | | | | | |
| Cohort A[e] | 185 | 141 (76.2%) | 38 (20.5%) | 6 (3.2%) | 44 (23.8%) | 1.46 (.98–2.18) | 3.66 (1.28–10.44) | .0109 | 1.59 (1.09–2.32) [.0199] | 1.60 (1.15–2.23) [.0054] | 3.42 (1.20–9.71) [.0264] |
| Cohort B[f] | 180 | 143 (79.4%) | 29 (16.1%) | 8 (4.4%) | 37 (20.5%) | 1.10 (.71–1.70) | 4.81 (1.83–12.68) | .0060 | 1.32 (.88–1.97) [.1938] | 1.45 (1.04–2.04) [.0298] | 4.74 (1.80–12.45) [.0027] |
| Control 2 | 1,035 | 860 (83.1%) | 172 (16.6%) | 3 (.3%) | 175 (16.9%) | | | | | | |
| Cohort C | 201 | 154 (76.6%) | 45 (22.3%) | 2 (1.0%) | 47 (23.3%) | 1.46 (1.01–2.12) | 3.72 (.62–22.46) | .0356 | 1.50 (1.04–2.16) [.0346] | 1.51 (1.07–2.14) [.0199] | 3.46 (.57–20.82) [.1882] |
| Cohort D | 139 | 110 (79.1%) | 24 (17.3%) | 5 (3.6%) | 29 (20.9%) | 1.09 (.68–1.75) | 13.03 (3.07–55.28) | .0015 | 1.30 (.83–2.01) [.2826] | 1.49 (1.01–2.22) [.0470] | 12.84 (3.03–54.32) [.0009] |
| Control 1+2 | 1,961 | 1,634 (83.3%) | 315 (16.1%) | 12 (.6%) | 327 (16.7%) | | | | | | |
| Cohorts A+C+D | 525 | 405 (77.1%) | 107 (20.4%) | 13 (2.5%) | 120 (22.8%) | 1.37 (1.07–1.75) | 4.37 (1.98–9.65) | .00009 | 1.48 (1.17–1.87) [.0014] | 1.53 (1.23–1.89) [.00010] | 4.12 (1.87–9.09) [.0006] |
| Cohorts B+C+D | 520 | 407 (78.3%) | 98 (18.8%) | 15 (2.9%) | 113 (21.7%) | 1.25 (.97–1.61) | 5.02 (2.33–10.80) | .00006 | 1.39 (1.09–1.76) [.0081] | 1.47 (1.19–1.82) [.0004] | 4.82 (2.24–10.37) [.00008] |

[a] Odds ratio (OR) for T/T and C/T genotype frequencies between control individuals and each subgroup of subjects; reference group is the CC genotype.

[b] 2 df.

[c] *P* value for dominant and recessive models calculated using Fisher's exact test. For the additive model, *P* values were obtained using the Wald test.

[d] The North American white control populations were recruited and genotyped as described in detail elsewhere (Begovich et al. 2004). They comprise three independently collected cohorts: healthy individuals enrolled in the New York Cancer Project (NYCP) ($n = 926$), control individuals from the GCI Discovery cohort ($n = 475$), and a third cohort of randomly selected white individuals from across the United States ($n = 560$). There were no significant differences in genotype frequency between men and women or within or between groups, and there were no age-related frequency differences. The NYCP cohort was used as control group 1. The GCI cohort and the random collection of control individuals were combined to form control group 2 ($n = 1,035$). Cohorts A and B were compared individually with control group 1, whereas cohorts C and D were compared individually with control group 2. Summary analysis compares SLE case data (A+C+D or B+C+D) with the genotype frequency data from the combined control sample (control groups 1 and 2; $n = 1,961$).

[e] First affected subject within each Minnesota sib-pair family. There were no significant differences in allele frequencies between male and female subjects across all the cohorts.

[f] Second affected subject from each Minnesota sib-pair family.

vated titers of T-dependent antibodies IgG1 and IgG2a (Hasegawa et al. 2004). All three of the human autoimmune diseases shown to date to be associated with PTPN22 R620W are characterized by the production of autoantibodies (anti-GAD Abs in T1D, anti-citrulline Abs and rheumatoid factor in RA, and a vast array of autoantibodies in SLE), and the appearance of these antibodies often predates clinical disease (Arbuckle et al. 2003; Rantapaa-Dahlqvist et al. 2003). We speculate that the 620W variant of PTPN22 may predispose individuals to autoimmunity by facilitating the generation of certain disease-associated autoantibodies, thereby contributing to disease onset and progression.

## Acknowledgments

We thank all the individuals with SLE, family members, and referring physicians for their ongoing participation in this work. We also thank Catherine Jensen, Neil Wenberg, and Thearith Koueth for technical assistance. This study was supported by grants (R01 AR32274, R01 AR43727) and contracts from the National Institute of Arthritis and Musculoskeletal and Skin Diseases and the National Institute of Allergy and Infectious Diseases, Johns Hopkins General Clinical Research Center grant MO1-RR00052, and grants from the Minnesota Lupus Foundation, the Mary Kirkland Center for Lupus Research, and the Alliance for Lupus Research.

## Electronic-Database Information

The URL for data presented herein is as follows:

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for SLE)

## References

Arbuckle MR, McClain MT, Rubertone MV, Scofield RH, Dennis GJ, James JA, Harley JB (2003) Development of autoantibodies before the clinical onset of systemic lupus erythematosus. N Engl J Med 349:1526–1533

Begovich AB, Carlton VEH, Honigberg LA, Schrodi SJ, Chokkalingam AP, Alexander HC, Ardlie KG, et al (2004) A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase *(PTPN22)* is associated with rheumatoid arthritis. Am J Hum Genet 75:330–337

Bottini N, Musumeci L, Alonso A, Rahmouni S, Nika K, Rostamkhani M, MacMurray J, Meloni GF, Lucarelli P, Pellecchia M, Eisenbarth GS, Comings D, Mustelin T (2004) A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. Nat Genet 36:337–338

Cloutier JF, Veillette A (1996) Association of inhibitory tyrosine protein kinase p50csk with protein tyrosine phosphatase PEP in T cells and other hemopoietic cells. EMBO J 15: 4909–4918

——— (1999) Cooperative inhibition of T-cell antigen receptor signaling by a complex between a kinase and a phosphatase. J Exp Med 189:111–121

Di Cristofano A, Kotsi P, Peng YF, Cordon-Cardo C, Elkon KB, Pandolfi PP (1999) Impaired Fas response and autoimmunity in $Pten^{+/-}$ mice. Science 285:2122–2125

Gaffney PM, Ortmann WA, Selby SA, Shark KB, Ockenden TC, Rohlf KE, Walgrave NL, Boyum WP, Malmgren ML, Miller ME, Kearns GM, Messner RP, King RA, Rich SS, Behrens TW (2000) Genome screening in human systemic lupus erythematosus: results from a second Minnesota cohort and combined analyses of 187 sib-pair families. Am J Hum Genet 66:547–556

Hasegawa K, Martin F, Huang G, Tumas D, Diehl L, Chan AC (2004) PEST domain-enriched tyrosine phosphatase (PEP) regulation of effector/memory T cells. Science 303:685–689

Kammer GM, Perl A, Richardson BC, Tsokos GC (2002) Abnormal T cell signal transduction in systemic lupus erythematosus. Arthritis Rheum 46:1139–1154

Khan IU, Tsokos GC, Kammer GM (2003) Abnormal B cell signal transduction in systemic lupus erythematosus. Curr Dir Autoimmun 6:89–104

O'Keefe TL, Williams GT, Davies SL, Neuberger MS (1996) Hyperresponsive B cells in CD22-deficient mice. Science 274: 798–801

Petri M (2000) Hopkins Lupus Cohort: 1999 update. Rheum Dis Clin North Am 26:199–213

Rantapaa-Dahlqvist S, de Jong BA, Berglin E, Hallmans G, Wadell G, Stenlund H, Sundin U, van Venrooij WJ (2003) Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. Arthritis Rheum 48:2741–2749

Schultz LD, Schweitzer PA, Rajan TV, Yi T, Ihle JN, Mathews J, Thomas ML, Beier DR (1993) Mutation at murine moth-eaten locus are within the hematopoetic cell protein-tyrosine phosphatase *(Hcph)* gene. Cell 73:1445–1454

Wakeland EK, Liu K, Graham RR, Behrens TW (2001) Delineating the genetic basis of systemic lupus erythematosus. Immunity 15:397–408