

A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes

Nunzio Bottini¹, Lucia Musumeci¹, Andres Alonso¹, Souad Rahmouni¹, Konstantina Nika¹, Masoud Rostamkhani², James MacMurray², Gian Franco Meloni³, Paola Lucarelli⁴, Maurizio Pellecchia¹, George S Eisenbarth⁵, David Comings² & Tomas Mustelin¹

We report that a single-nucleotide polymorphism (SNP) in the gene (*PTPN22*) encoding the lymphoid protein tyrosine phosphatase (LYP), a suppressor of T-cell activation, is associated with type 1 diabetes mellitus (T1D). The variants encoded by the two alleles, 1858C and 1858T, differ in a crucial amino acid residue involved in association of LYP with the negative regulatory kinase Csk. Unlike the variant encoded by the more common allele 1858C, the variant associated with T1D does not bind Csk.

SNPs are common in the human genome and often provide correlative evidence for the involvement of specific genes in human disease¹. SNPs that affect the function of crucial components of the T-cell antigen receptor (TCR) signaling pathways could have profound effects on the function of the immune system and thus the development of autoimmune diseases². Protein tyrosine phosphatases³ (PTPs) are particularly good candidates for carrying disease-related SNPs because they are involved in preventing spontaneous T-cell activation and they restrict the response to antigen by dephosphorylating and inactivating TCR-associated kinases and their substrates^{4,5}. PTPs are also needed to revert activated T blasts to a resting phenotype⁶.

The lymphoid-specific phosphatase (LYP) encoded by *PTPN22* is a 110-kDa PTP consisting of an N-terminal phosphatase domain and a long noncatalytic C terminus with several proline-rich motifs^{4,7}. LYP is expressed in lymphocytes⁷, where it physically associates through its most N-terminal proline-rich motif (called P1) with the SH3 domain of the Csk kinase⁸, an important suppressor of the Src family kinases Lck⁹ and Fyn, which mediate TCR signaling. LYP is among the most powerful inhibitors of T-cell activation^{4,10}, a task accomplished by dephosphorylation of TCR-associated kinases, like Lck⁴, Fyn¹⁰ and ZAP-70 (ref. 10). Dephosphorylation of ZAP-70 may be connected to the negative regulatory function of the c-Cbl proto-oncogene, as LYP forms a complex with c-Cbl⁷.

We found that *PTPN22* contains a SNP at nucleotide 1858 in codon 620 (Supplementary Fig. 1 online), which encodes an arginine (CGG) in most donors and in all published human and mouse LYP sequences, but encodes a tryptophan (TGG) in at least one allele of *PTPN22* in ~16.2% of genotyped healthy, non-Hispanic white individuals from the North American population (Table 1). Only 4 of 395 healthy individuals were homozygous with respect to the *PTPN22* 1858T allele. Because of the proposed role of LYP as a gatekeeper in T-cell activation, we determined whether the SNP was correlated with autoimmune disease.

Insulin-dependent type 1 diabetes mellitus (T1D) affects 0.5% of the human population and ~1.4 million people in the United States.

The disease is believed to arise as the consequence of an autoimmune destruction of insulin-producing β -cells by cytotoxic CD8⁺ T cells with help from CD4⁺ T cells. First, we carried out a genetic association study using a sample of 294 North American individuals with T1D and 395 healthy subjects from the same ethnic background. We genotyped these individuals and found that the *PTPN22* allele 1858T was more frequent in individuals with T1D than in healthy individuals: 30.6% of individuals with T1D were heterozygous with respect to 1858T, compared with only 21.3% of the healthy controls, and 3.7% of individuals with T1D were homozygous with respect to 1858T, compared with only 1.0% of the healthy controls ($\chi^2 = 14.84$ with 2 degrees of freedom, $P < 0.001$; odds ratio for the 1858T carriers versus noncarriers = 1.83, 95% confidence interval = 1.284–2.596). As the cohorts were matched for age and race, these results suggest that the *PTPN22* allele 1858T predisposes an individual to developing T1D.

This result was supported by genotyping a second, completely independent and ethnically separate cohort of 388 individuals from Sardinia, Italy. All affected individuals and controls came from northern Sardinia, a region characterized by an ethnically homogenous population. In this study, more than twice as many individuals with T1D as controls were heterozygous with respect to 1858T (8.6% versus 4.2%, respectively) and the only individual who was homozygous with respect to 1858T in the cohort had the disease ($\chi^2 = 3.93$ with 1 degree of freedom, $P < 0.05$; odds ratio for the 1858T carriers versus noncarriers = 2.31, 95% confidence interval = 0.932–5.819). This confirms that the *PTPN22* allele 1858T confers predisposition to T1D, even in a different background in which the allele frequency is lower. As LYP is expressed in lymphocytes and the disease is caused by lymphocytes, this correlation probably relates to altered lymphocyte function. Chromosomal region 1p13, where the

Table 1 Frequency of *PTPN22* genotypes in individuals with T1D in two independent samples

	Percentage and number of individuals with each genotype ^a		
	C/C	C/T	T/T
North American samples^b			
T1D ^c ($n = 294$)	65.6 (193)	30.6 (90)	3.7 (11)
Controls ^d ($n = 395$)	77.7 (307)	21.3 (84)	1.0 (4)
$\chi^2 = 14.84$, d.f. = 2, $P < 0.001$; OR 1858T carriers versus noncarriers = 1.83 (95% c.i. = 1.284–2.596)			
Italian samples^b			
T1D ^e ($n = 174$)	90.8 (158)	8.6 (15)	0.6 (1)
Controls ^f ($n = 214$)	95.8 (205)	4.2 (9)	0 (0)
$\chi^2 = 3.93$, d.f. = 1, $P < 0.05$; OR 1858T carriers versus noncarriers = 2.31 (95% c.i. = 0.932–5.819)			

^aGenotypes were determined as described in the legend to Supplementary Fig. 1 online. ^bThe study was approved by the Institutional Review Boards of the Institutions at which the subjects were recruited. The nature of the study was explained to all study participants and informed consent was obtained from each participant. ^cThis T1D sample consisted of unrelated, non-Hispanic, white subjects diagnosed with T1D at the Barbara Davis Center for Autoimmune Diabetes, Denver, Colorado, USA. ^dThis control sample consisted of 189 unrelated, non-Hispanic, white college students from California; 160 healthy unmarried, unrelated, non-Hispanic, white parents of the Minnesota Twin Study; and 46 healthy, unrelated, non-Hispanic, white controls from the Barbara Davis Center for Autoimmune Diabetes. These samples were not statistically significantly different from each other. ^eThis T1D sample consisted of unrelated subjects diagnosed with T1D at the Department of Pediatrics of the University of Sassari, Italy. Mean age of the subjects was 7.4 \pm 0.3 (s.e.) years, and 60.2% were males. ^fThis control sample consisted of healthy individuals from the same Italian population.

d.f., degrees of freedom; c.i., confidence interval; OR, odds ratio.

¹Program of Signal Transduction, Cancer Research Center, The Burnham Institute, 10901 North Torrey Pines Road, La Jolla, California 92037, USA. ²Department of Medical Genetics, City of Hope Medical Center, Duarte, California 91010, USA. ³Department of Pediatrics, University of Sassari, Sassari, Italy. ⁴Italian National Research Council, Rome, Italy. ⁵Barbara Davis Center for Childhood Diabetes, University of Colorado Health Science Center, Denver, Colorado 80262, USA. Correspondence should be addressed to T.M. (tmustelin@burnham-inst.org).

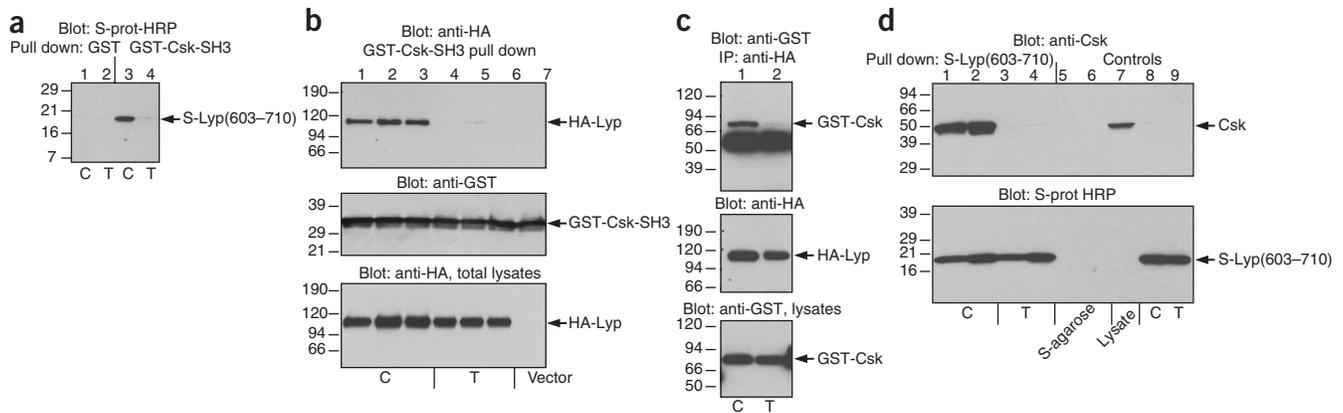


Figure 1 LYP with Trp620 (allele 1858T) does not bind Csk. **(a)** Binding of an S-tagged LYP polypeptide corresponding to amino acid residues 603–710 with Arg620 (allele 1858C) or Trp620 (allele 1858T) to GST (lanes 1 and 2) or GST-Csk-SH3 (lanes 3 and 4). The bound material was detected with S-protein-horseradish peroxidase (HRP). Both forms of the peptide were equally expressed in *E. coli* BL21. **(b)** Immunoblotting with antibody to hemagglutinin (HA; top panel) of material precipitated with GST-Csk-SH3 from COS cells transfected with three different amounts of plasmid expressing LYP with either Arg620 or Trp620. Blotting with antibody to GST of the same filter (middle panel) and blotting with antibody to hemagglutinin (HA) of lysates from the same transfectants (bottom panel) show equal expression of the two LYP variants. **(c)** Immunoblotting with antibody to GST (top panel) of material immunoprecipitated with LYP from COS cells transfected with GST-Csk and the two LYP proteins. Blotting with antibody to hemagglutinin (HA; middle panel) and with antibody to GST (bottom panel) of lysates from the same transfectants show equal expression. **(d)** Immunoblotting with antibody to Csk (upper panel) of endogenous Csk precipitated with the S-tagged LYP(603–710) protein with either Arg620 or Trp620 from Jurkat T cells. Lane 6 is material bound to S-agarose alone, lane 7 is total lysates of Jurkat cells and lanes 8 and 9 are the S-tagged protein alone. S-protein-HRP blot of the same samples (lower panel). Expression of recombinant proteins, protein binding, transfections, immunoprecipitation and immunoblotting were done as described^{14,15}.

PTPN22 gene resides, is also weakly linked to systemic lupus erythematosus¹¹ and rheumatoid arthritis¹², suggesting that it might be worth addressing a possible role for LYP in systemic autoimmune diseases.

To gain some insight into the possible molecular basis of the disease correlation, we examined the location and role of the amino acid encoded by codon 620. This residue resides in the P1 proline-rich motif involved in binding to the SH3 domain of Csk (**Supplementary Fig. 1** online). The three-dimensional structure of the complex¹³ shows that the side chain of Arg620 fits into a somewhat acidic depression within the ligand binding cleft of the SH3 domain, where it interacts with a tryptophan (Trp47) of Csk (**Supplementary Fig. 1** online). Replacing Arg620 with a tryptophan residue is expected to disrupt this interaction. The side chain of tryptophan is also bulkier and would not fit into the pocket. Indeed, when we expressed amino acids 603–710 of LYP as an S-tagged protein in *Escherichia coli*, only the Arg620 construct was precipitated by the SH3 domain of Csk (**Fig. 1a**). Similarly, full-length LYP with Arg620 (allele 1858C) expressed in COS cells was readily precipitated by the Csk SH3 domain, whereas LYP with Trp620 (allele 1858T) was not (**Fig. 1b**). Full-length LYP with Arg620 and Csk also coimmunoprecipitated (**Fig. 1c**), and endogenous Csk was readily precipitated from T cells by S-tagged LYP peptide 603–710 with Arg620 but not Trp620 (**Fig. 1d**). Together, these results indicate that only LYP with Arg620 (allele 1858C) forms a complex with the Csk kinase, whereas LYP with Trp620 (allele 1858T) does not.

Our findings suggest that a SNP in lymphoid tyrosine phosphatase is part of the complex genetic background of a common autoimmune disease, T1D. The SNP directly disrupts the formation of the LYP-Csk complex, the physiological relevance of which is well documented^{4,10}. The mouse ortholog of LYP (called PEP) did not inhibit T-cell activation in an antigen-specific T-cell hybridoma when Csk association was disrupted by mutation of the P1 motif¹⁰. We found that mouse⁴ as well as human (data not shown) LYP inhibits activation of human Jurkat T cells in a synergistic manner with Csk⁴. Thus, individuals heterozygous with respect to the *PTPN22* allele 1858T may have reduced

amounts of LYP-Csk complexes, and homozygous individuals no such complexes. Therefore, hyper-reactive T cells are more prone to mount a destructive immune response against autoantigens after viral infection or other immune stresses.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

This work was supported by fellowships from American-Italian Cancer Foundation, Rotary International, Van Beirs Foundation, Centre Anticancéreux près l'Université de Liège, the Spanish Ministry of Education and Culture, and by grants from the US National Institutes of Health.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 15 December 2003; accepted 29 January 2004

Published online at <http://www.nature.com/naturegenetics/>

- Bottini, N., Bottini, E., Gloria-Bottini, F. & Mustelin, T. *Arch. Immunol. Ther. Exp.* **50**, 95–104 (2002).
- Mustelin, T. & Altman, A. in *The Molecular Pathology of Autoimmune Diseases* 2nd edn. (A. Theofilopoulos & C. A. Bona, eds.) 127–141 (Taylor & Francis, New York, New York, 2002).
- Mustelin, T. *et al. Front. Biosci.* **7**, 85–142 (2002).
- Gjörloff-Wingren, A., Saxena, M., Williams, S., Hammi, D. & Mustelin, T. *Eur. J. Immunol.* **29**, 3845–3854 (1999).
- Mustelin, T., Abraham, R.T., Rudd, C.E., Alonso, A. & Merlo, J.J. *Front. Biosci.* **7**, 918–969 (2002).
- Iivanainen, A.V., Lindqvist, C., Mustelin, T. & Andersson, L.C. *Eur. J. Immunol.* **20**, 2509–2512 (1990).
- Cohen, S., Dadi, H., Shaoul, E., Sharfe, N. & Roifman, C.M. *Blood* **93**, 2013–2024 (1999).
- Cloutier, J.-F. & Veillette, A. *EMBO J.* **15**, 4909–4918 (1996).
- Bergman, M. *et al. EMBO J.* **11**, 2919–2924 (1992).
- Cloutier, J. & Veillette, A. *J. Exp. Med.* **189**, 111–121 (1999).
- Gaffney, P.M. *et al. Proc. Natl. Acad. Sci. USA* **95**, 14875–14879 (1998).
- Jawaheer, D. *et al. Arth. Rheum.* **48**, 906–916 (2003).
- Ghose, R., Shekhtman, A., Goger, M.J., Ji, H. & Cowburn, D. *Nat. Struct. Biol.* **8**, 998 (2001).
- Saxena, M., Williams, S., Taskén, K. & Mustelin, T. *Nat. Cell Biol.* **1**, 305–311 (1999).
- Alonso, A. *et al. Nat. Immunol.* **4**, 44–48 (2002).