Published in final edited form as:
Genes Immun. 2009 December ; 10(Suppl 1): S16–S20. doi:10.1038/gene.2009.86.

Analyses of multiple single-nucleotide polymorphisms in the SUMO4/IDDM5 region in affected sib-pair families with type I diabetes

R Podolsky, **MV Prasad Linga-Reddy**, **J-X She**, and **the Type I Diabetes Genetics Consortium**
Center for Biotechnology and Genomic Medicine, Medical College of Georgia, Augusta, GA, USA

Abstract

Previous studies suggested that the SUMO4 gene, located in the IDDM5 interval on chromosome 6q25, was associated with type I diabetes (T1D) and several other autoimmune diseases. Subsequent analyses of the SUMO4 variants with T1D suggested that the association was stronger and more consistent in the Asian populations. In addition, considerable heterogeneity has been observed in the Caucasian populations. In this report, a 40-kb genomic interval including the SUMO4 gene was tagged with 15 single-nucleotide polymorphisms. A total of 2317 affected sib-pair families from the Type I Diabetes Genetic Consortium were genotyped using both the Illumina and Sequenom genotyping platforms. In these Caucasian families, we found little evidence supporting an association between SUMO4 and T1D.

Keywords

type I diabetes; SUMO4; genetic susceptibility; linkage disequilibrium; single-nucleotide polymorphism

Introduction

SUMO4 is a member of a highly conserved family of genes that encode small ubiquitin-related modifiers (SUMO).[1] The process to attach SUMO to proteins is known as sumoylation. The sumoylation system, initially characterized in 1996,2 is a critical protein modification system found in all eukaryotic kingdoms.3›4 Sumoylation is a reversible modification involved in a variety of important processes of eukaryotic cells. Sumoylation seems to be a highly selective process both in terms of the choice of substrates and the timing of the modification, which is probably related to the maintenance of cellular homeostasis and the defensive response to stress or inflammatory insults.[5,6] *SUMO4* has a major function in NFκB and JAK/STAT signaling pathways and can regulate the activity of other critical immune molecules such as AP-1.[1,5,6]

SUMO4 is located in the *IDDM5* interval on chromosome 6q25, a type I diabetes (T1D) candidate region supported by numerous linkage studies. Multiple single-nucleotide polymorphisms (SNPs) surrounding the *SUMO4* gene have been shown to be associated with T1D.[1] The associated *SUMO4* SNP rs237025 has a single base-pair change (A→G) that causes an amino-acid change from methionine (M) to valine (V) at position 55 (M55V) in exon 1 of

Correspondence: Dr J-X She, Center for Biotechnology and Genomic Medicine, Medical College of Georgia, 1120 15th Street, Augusta, GA 30912, USA. jshe@mail.mcg.edu.

Conflict of interest

The authors declare no conflict of interest.

the gene. This SNP was strongly associated with T1D; further, the SNP is a potential ‘causal variant’ as the M55V substitution seemed to influence NFκB activity.[1,6] Subsequently, this initial observation was a subject of intensive debate because of the failure to replicate the association in several Caucasian populations.[6–9] Since the initial report, a number of studies have also revealed significant associations between SNPs in *SUMO4* with other autoimmune diseases,[10–14] type II diabetes,[15] and diabetic complications.[16,17] These studies have provided unambiguous evidence supporting a critical function of *SUMO4* in many inflammatory diseases.

Several important questions concerning T1D and other disease associations with SNPs in *SUMO4* remain to be addressed. One of the most important questions is why the *SUMO4*–T1D association is relatively strong in Asian populations, whereas the association evidence is inconsistent or undetectable in many Caucasian populations. The second question is whether the candidate *SUMO4* (M55V) SNP is the only functional variant in the *IDDM5* region. In the Type I Diabetes Genetic Consortium (T1DGC) Rapid Response project, 15 SNPs in the *SUMO4* region were genotyped using the Illumina and the Sequenom platforms. Here, we report the data analysis on a worldwide collection of 2317 T1D affected sib-pair families from multiple populations of Caucasian ethnicity.

Results

A total of 2317 families were genotyped for 15 SNPs in the *SUMO4* region (Supplementary Table 1) using both the Illumina and Sequenom platforms. The average call rate was 99.5% for the Illumina platform and 98.6% for the Sequenom platform. The average agreement of the genotype calls between the two platforms is 98.6% when the SNP rs237032 (agreement = 33%) is excluded from analysis. The agreement for 4 of the 14 SNPs (Supplementary Table 1) is below 97% (average = 95.8%). Although the agreement rates between platforms seem to be good, they may not be sufficient for association studies when the effect size of the genes is very small. The poor inter-platform agreement for the rs237032 and rs652921 SNPs (Supplementary Table 1) seems to be due to the poor call rates, resulting in significant deviation from Hardy–Weinberg equilibrium in the Sequenom data. Combining the genotype call rates and Hardy–Weinberg results, it seems that the Illumina platform may perform slightly better for the *SUMO4* SNPs.

Pedigree disequilibrium test analysis was first performed over all families in the T1DGC dataset (Table 1). Three SNPs (rs7742990, rs237032, and rs236999) had P-values <0.05 in the Illumina data, but only one of these three SNPs (rs236999) was marginally significant (P =0.07) in the Sequenom dataset (Table 1). Analysis of the data by source of the family collection also revealed several SNPs with marginally significant P-values in the Illumina panel for the British Diabetes Association and North America datasets (Supplementary Table 2). The earlier studied *SUMO4* rs237025 (M55V) SNP showed significant associations with T1D in the British Diabetes Association dataset for both Illumina (P =0.02) and Sequenom (P =0.04).

We next examined association between T1D and *SUMO4* SNP genotypes with respect to the interaction with HLA genotypes using a logistic regression model. The majority of the T1DGC affected sib-pair family members had been earlier typed for HLA-DR alleles. Family members were grouped into three HLA categories (HLA-DR4/4, -DR4/X, and -DRX/X, in which ‘DRX’ DR4). As shown in Table 2, the rs366905 SNP showed a significant association with T1D (P<0.05) in both the Illumina and Sequenom datasets. The rs366905 SNP also exhibited a marginally significant association with T1D that depended on HLA-DR4 genotype (DR4/4, DR4/X, DRX/X). The rs480034 SNP also showed evidence of significant association with T1D in both panels overall as well as with inclusion of HLA-DR4 genotype. The greatest contributor to the association seems to be in the DR4/4 subgroup (data not shown). The linkage disequilibrium (LD), measured by R^2 value, was calculated for all marker pairs and shown in Figure 1. Although there are strong LD between some markers pairs, there is no strong LD for this 45 kb region. This type of LD pattern is very unusual.

Discussion

This large T1DGC dataset failed to provide strong evidence for association between T1D and SNPs in the *SUMO4* region in Caucasian affected sib-pair families. These results further support that there is no evidence for association between *SUMO4* SNPs and T1D in Caucasian populations. In contrast, *SUMO4* is consistently associated with T1D in the Asian populations.[1,6,13,18–20] Furthermore, *SUMO4* is associated with several other inflammatory diseases including rheumatoid arthritis,[10–13] autoimmune thyroid disease,[13] autoimmune Behcet's disease,[14] diabetic retinopathy in T1D,[17] T2D,[15] and nephropathy in T2D.[16] These studies suggested that *SUMO4* is involved in many inflammatory diseases in both Asians and Caucasians.

A key question remains why the association between *SUMO4* SNPs and T1D is only observed in some Caucasian datasets, but not in others? The simplest interpretation of the data is that *SUMO4* is not involved in T1D susceptibility in Caucasians and the observed association evidence is purely because of chance. One potential hypothesis is that a gene (*SUMO4*) is involved in the risk of disease (T1D) only in certain ethnic groups, a hypothesis that we do not believe. An alternative explanation that we favor more is that the differences in association of SNPs in *SUMO4* and T1D between populations could be explained by differences in the patterns of LD between populations. This is a distinct possibility given the unusual LD pattern in this region. A third explanation is that *SUMO4* may only be an important risk factor in a subset of T1D patients and that the frequency and/or sampling of the patient subsets may determine whether a significant association can be observed. This interpretation is consistent with the weak and marginally significant associations observed in the majority of the Caucasian datasets, including this T1DGC dataset. These results are also consistent with the observations that association may be improved in Caucasian and Asian datasets after stratification with other genetic factors (such as HLA)[13,]21 or other autoimmune diseases.13 However, the main genetic or environmental factors that may interact with *SUMO4* remain to be identified. Ultimately, the true function of *SUMO4* in T1D susceptibility can only be understood when all the genetic factors and their interactions with other genes and environmental risks are fully characterized.

Materials and methods

Patients and methods

SNPs were genotyped in a set of 11 250 individuals comprising 5047 T1D affected individuals collected throughout the world by the T1DGC. In total, there are 2317 T1D nuclear families, of which 2126 families were of European origin and 191 were Asian-Pacific. Genotyping was carried out using the Illumina Infinium II Human-Hap550 BeadChip technology (Illumina, Inc., San Diego, USA) as well as Sequenom high throughput SNP genotyping platform. Details of the patient samples and the quality control procedures can be found in this volume (Brown *et al*.22).

Statistical analysis

Tests for Hardy–Weinberg frequencies were conducted by randomly sampling one normal subject from each pedigree. These tests were conducted using SAS v9.1. UNPHASED v2.404 was used with the PDTPHASE option to conduct the pedigree disequilibrium test analyses.[23] Single marker, and two- and three-marker haplotype association tests were used. Generalized estimating equations were used to analyze logistic regression models in examining interactions with HLA genotypes. The general model included SNP genotype, the number of HLA-DR4 alleles, and contributing cohort, as well as interactions among these three variables. Pedigree within cohort was used as the variable that accounted for repeated measures. These analyses were conducted using the Proc GENMOD procedure in SAS v9.1. Haploview was used to analyze LD.[24]

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The dataset analyzed in this study was provided by the Type I Diabetes Genetics Consortium (T1DGC), a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Human Genome Research Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418. Genotyping was performed at the Broad Institute Center for Genotyping and Analysis is supported by grant U54 RR020278 from the National Center for Research Resources.

References

1. Guo D, Li M, Zhang Y, Yang P, Eckenrode S, Hopkins D, et al. A functional variant of SUMO4, a new I kappa B alpha modifier, is associated with type 1 diabetes. Nat Genet 2004;36:837–841. [PubMed: 15247916]
2. Matunis MJ, Coutavas E, Blobel G. A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex 1. J Cell Biol 1996;135:1457–1470. [PubMed: 8978815]
3. Muller S, Hoege C, Pyrowolakis G, Jentsch S. SUMO, ubiquitin's mysterious cousin. Nat Rev Mol Cell Biol 2001;2:202–210. [PubMed: 11265250]
4. Li M, Guo D, Isales CM, Eizirik DL, Atkinson M, She J-X, et al. SUMO wrestling with type 1 diabetes. J Mol Med 2005;83:504–513. [PubMed: 15806321]
5. Guo D, Han J, Adam BL, Colburn NH, Wang MH, Dong Z, et al. Proteomic analysis of SUMO4 substrates in HEK293 cells under serum starvation-induced stress. Biochem Biophys Res Commun 2005;337:1308–1318. [PubMed: 16236267]
6. Wang CY, She J-X. SUMO4 and its role in type 1 diabetes pathogenesis. Diabetes Metab Res Rev 2008;24:93–102. [PubMed: 17990297]
7. Smyth DJ, Howson JM, Lowe CE, Walker NM, Lam AC, Nutland S, et al. Assessing the validity of the association between the SUMO4 M55V variant and risk of type 1 diabetes. Nat Genet 2005;37:110–111. [PubMed: 15678134]
8. Qu H, Bharaj B, Liu XQ, Curtis JA, Newhook LA, Paterson AD, et al. Assessing the validity of the association between the SUMO4 M55V variant and risk of type 1 diabetes. Nat Genet 2005;37:111–112. [PubMed: 15678135]
9. Kosoy R, Concannon P. Functional variants in SUMO4, TAB2, and NF[kappa]B and the risk of type 1 diabetes. Genes Immun 2005;6:231–235. [PubMed: 15729364]
10. Glaser B, Nikolov I, Chubb D, Hamshere M, Segurado R, Moskvina V, et al. Analyses of single marker and pairwise effects of candidate loci for rheumatoid arthritis using logistic regression and random forests. BMC Proc 2007;1(Suppl 1):S54. [PubMed: 18466554]
11. Yoo Y, Gao G, Zhang K. Case-control association analysis of rheumatoid arthritis with candidate genes using related cases. BMC Proc 2007;1(Suppl 1):S33. [PubMed: 18466531]
12. Ding Y, Cong L, Ionita-Laza I, Lo SH, Zheng T. Constructing gene association networks for rheumatoid arthritis using the backward genotype-trait association (BGTA) algorithm. BMC Proc 2007;1(Suppl 1):S13. [PubMed: 18466472]

13. Tsurumaru M, Kawasaki E, Ida H, Migita K, Moriuchi A, Fukushima K. Evidence for the role of small ubiquitin-like modifier 4 as a general autoimmunity locus in the Japanese population. J Clin Endocrinol Metab 2006;91:3138–3143. [PubMed: 16735488]
14. Hou S, Yang P, Du L, Zhou H, Lin X, Liu X, et al. SUMO4 gene polymorphisms in Chinese Han patients with Behcet's disease. Clin Immunol 2008;129:170–175. [PubMed: 18657476]
15. Noso S, Fujisawa T, Kawabata Y, Asano K, Hiromine Y, Fukai A, et al. Association of small ubiquitin-like modifier 4 (SUMO4) variant, located in IDDM5 locus, with type 2 diabetes in the Japanese population. J Clin Endocrinol Metab 2007;92:2358–2362. [PubMed: 17374705]
16. Lin HY, Wang CL, Hsiao PJ, Lu YC, Chen SY, Lin KD, et al. SUMO4 M55V variant is associated with diabetic nephropathy in type 2 diabetes. Diabetes 2007;56:1177–1180. [PubMed: 17229939]
17. Rudofsky G, Schlotterer A, Humpert PM, Tafel J, Morcos M, Nawroth PP, et al. A M55V polymorphism in the SUMO4 gene is associated with a reduced prevalence of diabetic retinopathy in patients with type 1 diabetes. Exp Clin Endocrinol Diabetes 2008;116:211–214. [PubMed: 18072015]
18. Park Y, Park S, Kang J, Yang S, Kim D. Assessing the validity of the association between the SUMO4 M55V variant and risk of type 1 diabetes. Nat Genet 2005;37:112–113. [PubMed: 15678137]
19. Noso S, Ikegami H, Fujisawa T, Kawabata Y, Asano K, Hiromine Y, et al. Genetic heterogeneity in association of the SUMO4 M55V variant with susceptibility to type 1 diabetes. Diabetes 2005;54:3582–3586. [PubMed: 16306380]
20. Ikegami H, Kawabata Y, Noso S, Fujisawa T, Ogihara T. Genetics of type 1 diabetes in Asian and Caucasian populations. Diab Res Clin Prac 2007;77:S116–S121.
21. Sedimbi SK, Luo XR, Sanjeevi CB, Lernmark A, et al. Swedish Childhood Diabetes Study Group, Diabetes Incidence in Sweden Study Group. SUMO4 M55V polymorphism affects susceptibility to type 1 diabetes in HLA DR3- and HLA DR4-positive Swedish patients. Genes Immun 2007;8:518–521. [PubMed: 17554341]
22. Brown WM, Pierce JJ, Hilner JE, Perdue LH, Lohman K, Lu L, et al. the Type I Diabetes Genetics Consortium. Overview of the Rapid Response data. Genes Immun 2009;10(Suppl 1):S5–S15. [PubMed: 19956101]
23. Dudbridge F. Pedigree disequilibrium tests for multilocus haplotypes. Genet Epidemiol 2003;25:115–121. [PubMed: 12916020]
24. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;15:263–265. [PubMed: 15297300]

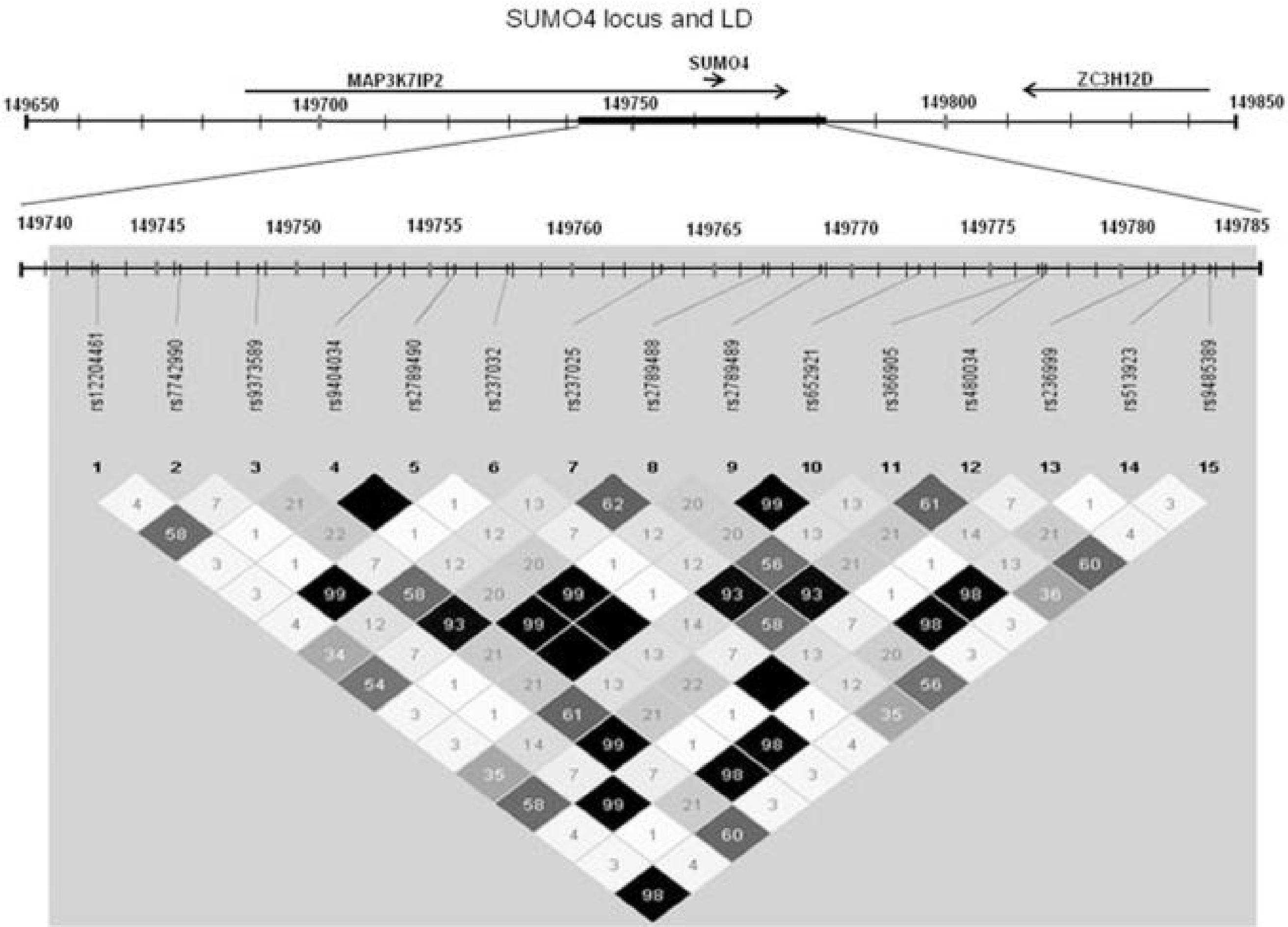

Figure 1.

Linkage disequilibrium map of the 15 SNPs on the basis of the Illumina genotyping data analyzed by the Haploview v4.0 software; r^2 values (%) are shown in the boxes. The black boxes have r^2 =1 and the white boxes have r^2 =0.

Table 1

PTD test for 15 SNPs in the *SUMO4* interval in all families

SNP	Illumina				Sequenom		
	A1:A2*	T:U_PDT	Aff:UnafSib	P_PDT	T:U_PDT	Aff:UnafSib	P_PDT
rs12204461	A:T	4387:4315	2418:2422	ns	4649:4601	2442:2447	ns
rs7742990	G:A	757:697	410:390	0.0208	682:667	394:387	ns
rs9373589	A:G	4057:4025	2091:2129	ns	3954:3901	2009:2034	ns
rs9404034	C:G	650:645	403:368	ns	674:661	399:372	ns
rs2789490	C:G	644:633	400:365	ns	633:618	395:369	ns
rs237032	G:A	757:699	410:389	0.0145	107:143	112:116	ns
rs237025	A:G	3048:3022	1695:1635	ns	3035:3026	1661:1605	ns
rs2789488	G:A	3941:3916	2022:2064	ns	2303:2320	1205:1194	ns
rs2789489	A:T	648:645	398:364	ns	659:640	398:373	ns
rs652921	T:C	646:642	400:365	ns	641:638	372:347	ns
rs366905	A:T	2976:2955	1658:1601	ns	2866:2873	1543:1502	ns
rs480034	T:C	4050:4020	2091:2129	ns	3777:3745	2011:2041	ns
rs236999	A:G	759:702	410:390	0.0202	728:703	392:385	0.0729
rs513923	G:A	636:626	407:368	ns	669:660	399:374	ns
rs9485389	C:A	4697:4649	2496:2499	ns	4595:4554	2409:2426	ns

* A1 allele is the transmitted allele.

Table 2

Logistic regression analysis of association between T1D and SNP genotype (T1D column) or dependent on DR4 genotypes (T1D*DR4)

SNP	T1D		T1D*DR4	
	Illumina	Sequenom	Illumina	Sequenom
rs12204461	0.0868	ns	ns	ns
rs7742990	ns	ns	ns	ns
rs9373589	0.0672	0.0138	0.0643	ns
rs9404034	ns	ns	ns	ns
rs2789490	ns	ns	ns	ns
rs237032	ns	**	ns	**
rs237025	ns	ns	0.0885	ns
rs2789488	ns	0.0904	0.0683	0.0786
rs2789489	ns	ns	ns	ns
rs652921	ns	ns	ns	ns
rs366905	0.0465	0.0418	0.0688	0.0642
rs480034	0.0660	0.0554	0.0646	0.0373
rs236999	ns	ns	ns	ns
rs513923	ns	ns	ns	ns
rs9485389	ns	ns	ns	ns

** Insufficient data.