Brief Genetics Report

Association of Human Endogenous Retrovirus K-18 Polymorphisms With Type 1 Diabetes

Samuel Marguerat,[1] William Y.S. Wang,[2] John A. Todd,[2] and Bernard Conrad[1]

Several lines of evidence suggest the involvement of the human endogenous retrovirus (*HERV)-K18* in the etiology of type 1 diabetes. *HERV-K18* encodes for a T-cell superantigen (SAg). T-cells with T-cell receptor Vβ7 chains reactive to the SAg and *HERV-K18* mRNA were enriched in the tissues at the onset of the disease. *HERV-K18* transcription and SAg function in cells capable of efficient presentation are induced by proinflammatory stimuli such as viruses and interferon-α and may trigger progression of disease to insulitis or from insulitis to overt diabetes. Allelic variation of *HERV-K18* or the DNA flanking it, the CD48 gene, could modulate genetic susceptibility. Analysis of 14 polymorphisms in the locus using 754 diabetic families provided positive evidence of association of three variants belonging to a single haplotype ($P = 0.0026$), present at 21.8% frequency in the population. Genotype analysis suggested a dominantly protective effect of this haplotype ($P = 0.0061$). Further genetic and functional analyses are required to confirm these findings. *Diabetes* 53: 852–854, 2004

Type 1 diabetes is a multifactorial disease caused by the T-cell–mediated destruction of the insulin-producing β-cells owing to a complex, and largely unknown, interaction with the environment (1). Although the major locus has been discovered, the HLA complex on chromosome 6p21, and other loci have been associated with disease susceptibility including the insulin gene and the T-cell inhibition gene *CTLA4* (2), many other genes probably contribute to the familial clustering of the disease. The candidate gene approach remains a powerful method for susceptibility gene identification, especially if the candidacy is specific. Such a strong candidacy exists for the human endogenous retrovirus (*HERV)-K18* locus on human chromosome 1.

Several lines of evidence suggest the involvement of *HERV-K18* in the etiology of type 1 diabetes. *HERV-K18* encodes T-cell superantigen (SAg), and T-cells with *HERV-K18* SAg reactive T-cell receptor Vβ7 chains were found to be enriched in the pancreas, in the spleen (3,4), and in circulation (5) at disease onset. *HERV-K18* mRNA expression was also enhanced in inflammatory lesions of patients with recent-onset type 1 diabetes (6). *HERV-K18* transcription and SAg function in cells capable of efficient presentation are induced by proinflammatory stimuli (7,8) with established immunopathological potential, namely viruses (9) and interferon-α (10). *HERV-K18* SAg may thus trigger progression of disease to insulitis, or from insulitis to overt diabetes, and allelic variation of the HERV or the DNA flanking it, the CD48 gene, could modulate genetic susceptibility (7).

HERV-K18 is 9,235 bp in length and located within intron 1 of *CD48* on human chromosome 1q. We previously characterized the locus and determined its haplotype diversity in the European population (7). Three main haplotypes were identified that differed in amino acid sequence at five positions within the SAg coding region (7). Two of them have or could have biochemical consequences for SAg structure and function, namely a Y/C substitution in haplotypes 1 and 3 at position 97 and a premature stop codon in haplotype 1 at position 154. The latter variant produces a soluble COOH-terminally truncated SAg protein for haplotype 1 and full-length envelope proteins requiring intracellular cleavage for haplotypes 2 and 3. The Y/C substitution in haplotypes 1 and 3 could interfere with intra- and interchain disulfide bonding that is critical for the maturation of secreted and membrane proteins. Despite these differences, the three *HERV-K18* alleles all encode SAgs with an indistinguishable capacity to stimulate mature T-cells and T-cell hybrids.

To evaluate the association of *HERV-K18* polymorphisms with type 1 diabetes, we undertook a large family-based association study involving 14 single nucleotide polymorphisms (SNPs) and 754 families (see RESEARCH DESIGN AND METHODS), focusing on the main haplotypes of HERV-K18 and flanking them with *CD48* SNPs to delimit linkage disequilibrium.

Table 1 shows results of allelic and genotypic association analyses. Hardy-Weinberg equilibriums were observed for all SNPs. Figure 1 shows the pattern of linkage disequi-

From the [1]Department of Genetics and Microbiology, University of Geneva Medical School, Geneva, Switzerland; and the [2]Juvenile Diabetes Research Foundation/Wellcome Trust Diabetes and Inflammation Laboratory, Cambridge Institute for Medical Research, University of Cambridge, Cambridge, U.K.

Address correspondence and reprint requests to B. Conrad or John A. Todd, Department of Genetics & Microbiology, CMU, 1 rue Michel Servet, University of Geneva Medical School, 1211 Geneva 4, Switzerland. E-mail: b.conrad@medecine.unige.ch.

Received for publication 13 October 2003 and accepted in revised form 17 November 2003.

S.M. and W.Y.S.W. contributed equally to this work

B.C. holds stock in NovImmune, a company that develops diabetes drugs.

HERV, human endogenous retrovirus; SAg, T-cell superantigen; SNP, single nucleotide polymorphism.

TABLE 1
SNPs used in this study, and the results of allelic and genotypic association analysis

	Allele		Minor allele				Genotype relative risks	
	Major	Minor	Frequency	Transmitted	Nontransmitted	*P*	Minor homozygotes	Major homozygotes
rs3795324	C	A	0.183	379	365	0.533	1.22 (0.81–1.84)	0.99 (0.84–1.17)
rs3766366	G	A	0.098	175	208	0.113	1.40 (0.57–3.45)	1.24 (0.99–1.55)
rs3766367	G	A	0.019	40	60	0.0817	too rare	1.67 (1.06–2.63)*
rs3796502	T	A	0.263	475	531	0.0890	1.07 (0.81–1.40)	1.19 (1.02–1.39)*
rs2295615	C	G	0.137	303	337	0.195	1.08 (0.65–1.78)	1.15 (0.96–1.37)
HERV-6836	A	G	0.383	678	652	0.492	1.12 (0.93–1.36)	1.01 (0.87–1.17)
HERV-7007	G	A	0.446	655	612	0.232	1.03 (0.86–1.23)	0.90 (0.77–1.07)
HERV-8148	G	A	0.214	370	450	0.00821	0.98 (0.68–1.41)	1.27 (1.07–1.49)†
HERV-8594	C	T	0.216	374	459	0.00432	0.89 (0.62–1.27)	1.25 (1.07–1.47)†
HERV-8460	T	C	0.218	385	476	0.00260	0.84 (0.59–1.19)	1.25 (1.07–1.46)†
rs3766369	A	G	0.347	621	588	0.364	1.17 (0.94–1.45)	1.00 (0.86–1.16)
rs352683	G	A	0.214	391	427	0.232	1.17 (0.83–1.64)	1.17 (0.99–1.38)
rs352684	G	A	0.438	623	671	0.197	1.10 (0.91–1.32)	1.21 (1.04–1.42)*
rs352685	C	G	0.322	605	559	0.170	1.13 (0.91–1.40)	0.94 (0.82–1.09)

Genotype relative risks were based on genotype distortion analyses with 95% CIs shown. Heterozygote risks = 1. *P* values were determined by transmission disequilibrium testing. *$P < 0.05$; †$P < 0.01$.

librium of the CD48 locus with three regions of high linkage disequilibrium and breaks at rs3766367 and rs352683.

We have found some evidence of association ($P = 0.0026$) between three *HERV-K18* SNPs at nt8146, nt8594, and nt8460 and type 1 diabetes. These three SNPs were in perfect linkage disequilibrium ($r^2 > 0.98$) with each other and belonged to the same previously identified haplotype, *HERV-K18.3*, which is unique in having the combination of a tyrosine at position 97 and no stop codon at position 154 (7). The slight differences in their results were due to different genotyping failure rates. Genotype-based analyses suggest that haplotype 3 may be dominantly protective for the disease ($P = 0.0061$), but the final determination of the mode of inheritance will require much larger datasets due to the small number of individuals who are homozygous for the minor alleles. U.S. and U.K. samples showed similar results with 56 and 54% transmission, respectively.

In a small pilot study involving 74 Japanese case subjects and 54 control subjects (11), no significant association was shown between *HERV-K18* SNPs at nucleotide positions 6836 and 7007, and borderline association ($P = 0.03$) was demonstrated only in subgroup analyses. That study, however, was underpowered. It nevertheless demonstrated significant differences in haplotype frequencies between the Japanese and Caucasian populations (7).

A number of explanations can be given for our observed associations between *HERV-K18* SNPs and diabetes. Given that the three associated SNPs were highly correlated with each other, Hardy-Weinberg equilibrium was observed, and the misinheritance rate was very low,

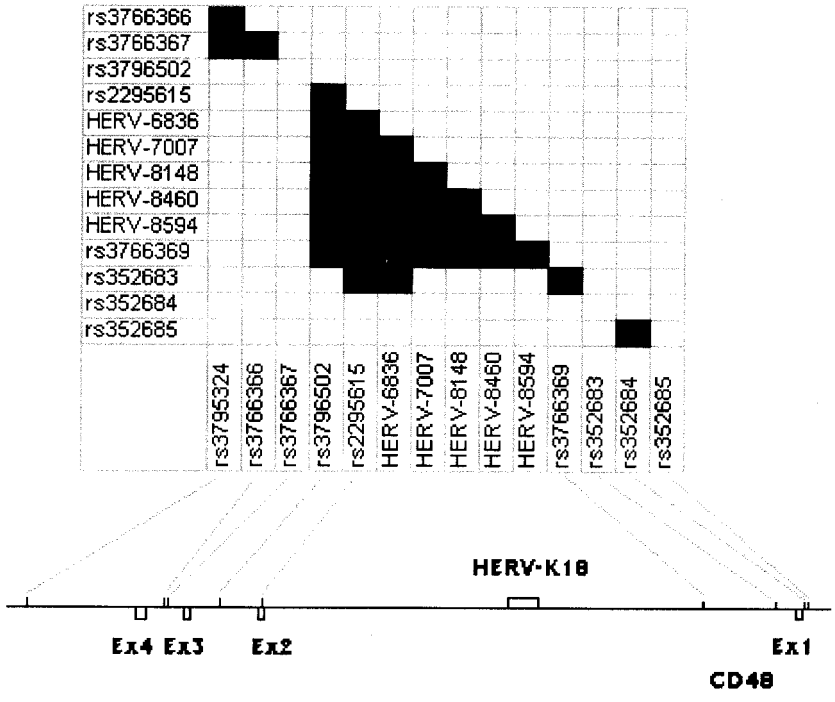

FIG. 1. The pattern of linkage disequilibrium of the CD48 locus with three regions of high linkage disequilibrium and breaks at rs3766367 and rs352683.

associations due to technical errors (12) seem unlikely. However, it is possible that our result is a statistical false-positive, even given the functional candidacy of *HERV-K18*. Replication studies with other large datasets are thus warranted. If the associations are true positives, the causative variant may lie anywhere between rs3766367 and rs352683, a 30-kb region from intron 1 to exon 3 of *CD48*. The high degree of correlation between SNPs would make isolating the disease variant difficult (2).

The association of HERV-K18 in the context of other genes involved in type 1 diabetes will warrant further examination. HERV-K18 SAgs are exquisitely major histocompatibility complex class II dependent (7,8), and therefore, genetic epistasis between the two loci is possible. However, given the moderate genetic effect we have reported, interactions analyses between HERV-K18 and HLA and with other loci firmly established in type 1 diabetes will require much larger sample sizes to ensure statistical power. This is provided that our current primary association can be confirmed. The involvement of HERV-K18 in other autoimmune diseases also remains to be tested (2). Our PCR assay and genotyping methods, and our determination of linkage disequilibrium structure in this genomic region, will facilitate those future investigations.

The HERV-K18 locus appears to be unique among the endogenous retrovirus HERV-K family with respect to gene regulatory features that could predispose it for autoimmunity, namely its transcriptional induction by proinflammatory stimuli with immunopathological potential and its constitutive expression in the thymus (B.C. and F. Meylan, unpublished data). Both of these characteristics are not shared by other known HERV-K proviruses, and our observation that *cis* regulatory elements residing outside the HERV-K18 provirus in the *CD48* gene are required for these responses could constitute at least in part a basis for why disease susceptibility has yet to be detected with other HERV loci.

In conclusion, our current study supports the *HERV-K18/CD48* locus in the genetic etiology of type 1 diabetes. This complex region, however, will require further genetic and functional analyses to firmly establish its role in the disease.

RESEARCH DESIGN AND METHODS

The study population consisted of 754 families each composed of both parents and at least two affected offspring. This included 472 Diabetes U.K. Warren 1 multiplex families (13) and 282 multiplex families obtained in the U.S. from the Human Biological Data Interchange (14).

SNP selection and genotyping. Five SNPs within *HERV-K18* and nine SNPs in the flanking *CD48* locus were genotyped. Using HERV-K18 nucleotide positions as reference, genotyped *HERV-K18* SNPs were at nt6836, nt7007, nt8146, nt8594, and nt8460 (7). Their correspondence with *HERV-K18* haplotypes has been described previously (7). CD48 SNPs genotyped were dbSNP rs3795324, rs3766366, rs3766367, rs3796502, rs2295615, rs3766369, rs352683, rs352684, and rs352685. The first five were telomeric to *HERV-K18*, and the rest were centromeric.

Genotyping was performed using Taqman MGB chemistry (Applied Biosystems, Foster City, CA) (15) on either PCR products (7) for the five *HERV-K18* SNPs or genomic DNA.

Statistical analysis. Statistical analyses were performed in Stata using Genassoc routines (available from http://www.gene.cimr.cam.ac.uk/clayton/software/stata/). For the transmission disequilibrium test and tests of genotype distortion, the "robust cluster (pedigree)" option was used to provide valid tests of association when there were more than one affected member per family. Between-marker linkage disequilibrium analyses involved the use of only parents.

ACKNOWLEDGMENTS

This work was funded by the Wellcome Trust, the Juvenile Diabetes Research Foundation International, the Swiss National Science Foundation, and the Helmut Horten Foundation. W.Y.S.W. receives scholarships from the University of Cambridge Clinical School, Gonville and Caius College, and the University of Sydney.

We thank the Human Biological Data Interchange and Diabetes U.K. for the family collections.

REFERENCES

1. Todd JA: From genome to aetiology in a multifactorial disease, type 1 diabetes. *Bioessays* 21:164–174, 1999
2. Ueda H, Howson JM, Esposito L, Heward J, Snook H, Chamberlain G, Rainbow DB, Hunter KM, Smith AN, Di Genova G, Herr MH, Dahlman I, Payne F, Smyth D, Lowe C, Twells RC, Howlett S, Healy B, Nutland S, Rance HE, Everett V, Smink LJ, Lam AC, Cordell HJ, Walker NM, Bordin C, Hulme J, Motzo C, Cucca F, Hess JF, Metzker ML, Rogers J, Gregory S, Allahabadia A, Nithiyananthan R, Tuomilehto-Wolf E, Tuomilehto J, Bingley P, Gillespie KM, Undlien DE, Ronningen KS, Guja C, Ionescu-Tirgoviste C, Savage DA, Maxwell AP, Carson DJ, Patterson CC, Franklyn JA, Clayton DG, Peterson LB, Wicker LS, Todd JA, Gough SC: Association of the T-cell regulatory gene *CTLA4* with susceptibility to autoimmune disease. *Nature* 423:506–511, 2003
3. Somoza N, Vargas F, Roura-Mir C, Vives-Pi M, Fernandez-Figueras MT, Ariza A, Gomis R, Bragado R, Marti M, Jaraquemada D: Pancreas in recent onset insulin-dependent diabetes mellitus: changes in HLA, adhesion molecules and autoantigens, restricted T cell receptor V beta usage, and cytokine profile. *J Immunol* 153:1360–1377, 1994
4. Conrad B, Weidmann E, Trucco G, Rudert WA, Behboo R, Ricordi C, Rodriquez-Rilo H, Finegold D, Trucco M: Evidence for superantigen involvement in insulin-dependent diabetes mellitus aetiology. *Nature* 371:351–355, 1994
5. Luppi P, Zanone MM, Hyoty H, Rudert WA, Haluszczak C, Alexander AM, Bertera S, Becker D, Trucco M: Restricted TCR V beta gene expression and enterovirus infection in type I diabetes: a pilot study. *Diabetologia* 43:1484–1497, 2000
6. Conrad B, Weissmahr RN, Boni J, Arcari R, Schupbach J, Mach B: A human endogenous retroviral superantigen as candidate autoimmune gene in type I diabetes. *Cell* 90:303–313, 1997
7. Stauffer Y, Marguerat S, Meylan F, Ucla C, Sutkowski N, Huber B, Pelet T, Conrad B: Interferon-alpha-induced endogenous superantigen: a model linking environment and autoimmunity. *Immunity* 15:591–601, 2001
8. Sutkowski N, Conrad B, Thorley-Lawson DA, Huber BT: Epstein-Barr virus transactivates the human endogenous retrovirus HERV-K18 that encodes a superantigen. *Immunity* 15:579–589, 2001
9. Jaeckel E, Manns M, Von Herrath M: Viruses and diabetes. *Ann N Y Acad Sci* 958:7–25, 2002
10. Ioannou Y, Isenberg DA: Current evidence for the induction of autoimmune rheumatic manifestations by cytokine therapy. *Arthritis Rheum* 43:1431–1442, 2000
11. Kinjo Y, Matsuura N, Yokota Y, Ohtsu S, Nomoto K, Komiya I, Sugimoto J, Jinno Y, Takasu N: Identification of nonsynonymous polymorphisms in the superantigen-coding region of IDDMK1,2 22 and a pilot study on the association between IDDMK1,2 22 and type 1 diabetes. *J Hum Genet* 46:712–716, 2001
12. Gordon D, Heath SC, Liu X, Ott J: A transmission/disequilibrium test that allows for genotyping errors in the analysis of single-nucleotide polymorphism data. *Am J Hum Genet* 69:371–380, 2001
13. Bain SC, Todd JA, Barnett AH: The British Diabetic Association–Warren repository. *Autoimmunity* 7:83–85, 1990
14. Lernmark A, Ducat L, Eisenbarth G, Ott J, Permutt MA, Rubenstein P, Spielman R: Family cell lines available for research. *Am J Hum Genet* 47:1028–1030, 1990
15. Ranade K, Chang MS, Ting CT, Pei D, Hsiao CF, Olivier M, Pesich R, Hebert J, Chen YD, Dzau VJ, Curb D, Olshen R, Risch N, Cox DR, Botstein D: High-throughput genotyping with single nucleotide polymorphisms. *Genome Res* 11:1262–1268, 2001