

Genetic discrimination between LADA and childhood-onset type 1 diabetes within the MHC

Short running title: MHC discriminators for Type 1 Diabetes vs. LADA

Rajashree Mishra, PhD^{1,2*}, Mikael Åkerlund, PhD^{3*}, Diana L. Cousminer, PhD^{1,4*}, Emma Ahlqvist, PhD³, Jonathan P. Bradfield, BS⁵, Alessandra Chesi, PhD¹, Kenyaita M. Hodge, MPH¹, Vanessa C. Guy, MS^{1^}, David J. Brillon, MD⁶, Richard E. Pratley, MD⁷, Michael R. Rickels, MD, MS⁸, Adrian Vella, MD⁹, Fernando Ovalle, MD¹⁰, Ronald I. Harris, MD¹¹, Olle Melander, MD, PhD³, Stephen Varvel, PhD¹², Hakon Hakonarson, MD, PhD^{5,13}, Phillippe Froguel, MD, PhD^{14,15}, John T. Lonsdale, PhD¹⁶, Didac Mauricio, MD, PhD¹⁷, Nanette C. Schloot, MD^{18‡}, Kamlesh Khunti, MD, PhD¹⁹, Carla J. Greenbaum, MD²⁰, Knud B. Yderstræde, MD, PhD²¹, Tiinamaija Tuomi, MD, PhD^{22,23, 29}, Benjamin F. Voight, PhD^{4,8,24,25}, Stanley Schwartz, MD²⁶, Bernhard O. Boehm, MD^{27,28}, Leif Groop, MD, PhD^{3,29}, Richard David Leslie, MD³⁰ and Struan F.A. Grant, PhD^{1,4,5,8,13}

*equal contribution

¹Division of Human Genetics, The Children's Hospital of Philadelphia, Philadelphia, PA, USA; ²Graduate Group in Genomics and Computational Biology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; ³Lund University Diabetes Centre, Department of Clinical Sciences Malmö, Lund University, Skåne University Hospital, SE-21428 Malmö, Sweden; ⁴Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; ⁵Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia, PA, USA; ⁶Weill Cornell Medical College, NY, USA; ⁷Florida Hospital Translational Research Institute for Metabolism and Diabetes, Orlando FL, USA; ⁸Institute for Diabetes, Obesity and Metabolism, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; ⁹Mayo Clinic, Rochester, MN, USA; ¹⁰University of Alabama, Birmingham, AL, USA; ¹¹Geisinger Health System, Danville, PA, USA; ¹²Health Diagnostic Lab Inc., Richmond, VA; ¹³Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; ¹⁴Centre National de la Recherche Scientifique (CNRS) 8199, Université Lille Nord de France, Pasteur Institute, Lille, France; ¹⁵Department of Genomics of Common Disease, Imperial College London, London, UK; ¹⁶National Disease Research Interchange, Philadelphia, PA, USA; ¹⁷Hospital de la Santa Creu i Sant Pau, CIBERDEM, Barcelona, Spain; ¹⁸German Diabetes Center, Düsseldorf, Germany; ¹⁹Diabetes Research Centre, University of Leicester, Leicester, UK; ²⁰Benaroya Research Institute, Seattle, WA, USA; ²¹Odense University Hospital, Odense, Denmark; ²²Helsinki University Hospital, Department of Endocrinology, Helsinki, Finland; ²³Folkhälsan Research Centre, Helsinki Finland; Research Programs Unit, Clinical and Molecular Metabolism, University of Helsinki; Helsinki, Finland; ²⁴Department of Systems, Pharmacology and Translational Therapeutics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; ²⁵Institute for Translational Medicine and Therapeutics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; ²⁶Main Line Health System, Wynnewood, PA, USA; ²⁷Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore and Imperial College, London, UK; ²⁸Department of Internal Medicine I, Ulm University Medical Centre, Ulm, Germany; ²⁹Institute for Molecular Medicine Finland FIMM, Helsinki, Finland; ³⁰Department of Immunobiology, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK;

‡ Currently employed by Lilly Germany

^ Currently employed by Science 37

Corresponding author:

Struan F. A. Grant

Divisions of Human Genetics and Endocrinology

Children's Hospital of Philadelphia

3615 Civic Center Boulevard, Room 1102D

Philadelphia, PA 19104

Phone: 267-426-2795; Fax: 267-426-0363

Email: grants@email.chop.edu

Word count: 3797

Tables: 1

Figures: 2

Abstract

Objective: The major histocompatibility complex (MHC) region harbors the strongest loci for latent autoimmune diabetes in adults (LADA); however, the strength of association is likely attenuated compared to childhood-onset type 1 diabetes. In this study, we recapitulate independent effects in the MHC Class I region in a type 1 diabetes population, then determine whether such conditioning in LADA yields potential genetic discriminators between the two subtypes within this region.

Research Design and Methods: Chromosome 6 was imputed using SNP2HLA, with conditional analysis performed in type 1 diabetes cases ($n = 1,985$) and controls ($n=2,219$). The same approach was applied to a LADA cohort ($n=1,428$) using population-based controls ($n=2,850$), and in a separate replication cohort (656 type 1 diabetes cases, 823 LADA cases, and 3,218 controls).

Results: The strongest associations in the MHC Class II region ($rs3957146$, Beta (SE) = 1.44 (0.05)), as well as the independent effect of MHC Class I genes, on type 1 diabetes risk, particularly *HLA-B*39* (Beta (SE) = 1.36 (0.17)) were confirmed. The conditional analysis in LADA versus controls showed significant association in the MHC Class II region ($rs3957146$, Beta (SE) = 1.14 (0.06)); however, we did not observe significant independent effects of MHC class I alleles in LADA.

Conclusion: In LADA, the independent effects of MHC class I observed in type 1 diabetes were not observed after conditioning on the leading MHC class II associations, suggesting that the MHC class I association may be a genetic discriminator between LADA and childhood-onset type 1 diabetes.

‘Latent autoimmune diabetes in adults’ (LADA) is typically defined as initial insulin independency for at least six months after diagnosis and the presence of diabetes associated autoantibodies (1). Despite such features, autoantibody screening is not routinely carried out in routine clinical practice, resulting in frequent misdiagnosis. For instance, in a cohort of apparent type 2 diabetes cases, as many as 8-10% can actually represent misdiagnosed autoimmune diabetes cases (2,3). Hence, there is a need to identify biomarkers to aid in accurately diagnosing LADA as well as other diabetes subtypes(4).

A comprehensive analysis of the genetic etiology of LADA has, until recently, not been performed (5). Previous genetic studies have suggested the condition comprised both type 1 diabetes and type 2 diabetes components either because it is an intermediate form of diabetes or because it is a mixture of type 2 diabetes in a predominantly type 1 diabetes cohort owing to a high false positive detection rate using autoantibodies when screening. There is some debate as to whether LADA is in fact a distinct clinical entity or simply a category imposed on continuous features such as age of onset and time to insulin. However, since LADA is currently defined as a slowly progressive form of type 1 diabetes (6), it is crucial to define genetic differences between childhood-onset type 1 diabetes and LADA if we are to clarify the clinical utility of identifying adult-onset autoimmune diabetes.

Previous genetic studies in LADA have shown a strong association signal in the major histocompatibility complex (MHC), although with diminished effect sizes compared to observations in childhood onset (5,7). The MHC region is located on chromosome 6 and harbors over 400 genes, with two main classes, MHC Class I and MHC Class II, which together harbor

classic human leukocyte antigen (HLA) genes (*HLA-A*, *HLA-B*, *HLA-C* and *HLA-DRB*, *HLA-DQA*, *HLA-DQB*, *HLA-DPA*, and *HLA-DPB*, respectively). The HLA encodes cell surface proteins for antigen presentation and accounts for approximately 50% of the genetic heritability of type 1 diabetes, with susceptibility principally harbored within the MHC Class II genes *HLA-DQB1* and *HLA-DRB1*. However, in addition to Class II genes, previous studies have also suggested MHC Class I genes in susceptibility to type 1 diabetes (8–10); in particular, variation within the MHC class I genes *HLA-A* and *HLA-B* variation has been shown through conditional analysis to further increase type 1 diabetes risk (11). MHC Class I markers have also been shown to be associated with younger age-at-diagnosis in type 1 diabetes, and given the adult-onset phenotype of LADA, we hypothesized that this genetic variation will be less enriched in LADA.

In this effort, we first attempted to recapitulate the independent effects of MHC Class I variants using the SNP2HLA imputation tool followed by stepwise forward logistic regression in the same type 1 diabetes cohort as the previous study (11). In addition, we set out to identify distinguishing features within the MHC between childhood-onset type 1 diabetes from adult-onset LADA, by performing the same conditional analysis followed by a replication attempt in a second case/control set. Finally, we compared beta regression coefficients for each disease to determine whether or not effect sizes differ between LADA cases and controls versus type 1 cases and controls.

Materials & Methods

Study populations

(I) LADA cases: 1,492 LADA cases were derived from multiple cohorts across the United Kingdom, Germany and the United States. Details on the participants can be found in **Supplementary Table 1**. All participants were diagnosed with LADA if they fulfilled the following criteria: age at diagnosis 30-70 years old, tested positive for at least one diabetes-associated autoantibody (most cases were positive for glutamic Acid Decarboxylase autoantibodies (GADA)) and were not on insulin treatment for at least 6 months after diagnosis.

(II) Controls: The LADA population-based controls comprised of two cohorts (n=2,979). The first cohort consisted of 1,296 non-diabetic children and adolescents of European ancestry, aged 5-20 years, enrolled in the Bone Mineral Density in Childhood Study (BMDCS (12)). The second control cohort consisted of 1,683 adults of European ancestry from a Non-Hodgkin lymphoma GWAS available in dbGaP (www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000818.v2.p1) (13). Details on the control cohorts can be found in **Supplementary Table 1**.

(III) Recapitulating a previous study: We also leveraged 3,000 healthy adult British Birth Cohort controls, 2,000 individuals with childhood-onset type 1 diabetes and 1,999 individuals with type 2 diabetes from the Wellcome Trust Case Control Consortium (WTCCC)(14) to recapitulate observations found in a previous study (11).

Individual data from the WTCCC is available through the Consortium's Data Access Committee

(<http://www.wtccc.org.uk>). More details on cohort information can be found in **Supplementary**

Table 1. (IV) Replication: A cohort of individuals from the All New Diabetics In Scania (ANDIS) and Scania Diabetes Registry (SDR) studies were used for further recapitulation and replication, including case subjects with type 1 diabetes (N = 656), LADA cases (n=823) and

population-based controls (N=3,218). Details on the participants can be found in

Supplementary Table 1. See flow chart for overview of datasets and workflow

(Supplementary Figure 1).

Genotyping

All samples, except the WTCCC data, were genotyped using the Illumina OmniExpress genotyping chip. WTCCC type 1 diabetes and type 2 diabetes cases were genotyped using Affymetrix 500K and WTCCC controls were genotyped on the Illumina 1.2M BeadChip. Quality control was performed using PLINK(15). Individuals with ambiguous sex, genotype missingness >5%, genome-wide heterozygosity (3 standard deviations from the mean), duplicates and related-individuals were excluded (See **Supplementary Table 1** for details). Principal component (PC) analysis was performed using PLINK, and outliers were removed to exclude individuals with non-European ancestry. Single nucleotide polymorphisms (SNPs) with missing rate <5%, minor allele frequency (MAF) <1% and Hardy-Weinberg equilibrium exact test P-value below 1×10^{-5} were removed before HLA imputation.

HLA imputation

Starting from the genotyped SNPs, we imputed chromosome 6 using the HLA imputation software SNP2HLA along with the Type 1 Diabetes Genetics Consortium (T1DGC) reference panel (16). A marker window size of 1,000 bp and a posterior probability (gprob) threshold of 0.5 were used. The HLA alleles of LADA cases (n = 1,428) and WTCCC type 1 diabetes cases (n = 1,985) were imputed to both 2-digit resolution and 4-digit resolution for increased coverage and resolution of HLA alleles. In total, there were 5,698 SNPs, 424 HLA alleles and 1,276 HLA

amino acids. In this study, we focused on a subset of SNPs and HLA alleles which had a MAF greater than 1% in all three control cohorts (159 HLA alleles and 5,506 SNPs remained).

Power calculations

Power calculations were performed using the Genetic Association Study (GAS) Power Calculator (<http://csg.sph.umich.edu/abecasis/cats/>). Assumptions included a multiplicative model, a disease incidence of 0.0036, 1,428 cases and 2,979 controls and a significance level of 8.83×10^{-6} , based on a Bonferroni correction for the 5,665 variants tested (**Supplementary Table 2**).

Recapitulation of a previously published conditional analysis for type 1 diabetes

Logistic regression using SNPTEST (17) was used to test all HLA alleles and SNPs with MAF >1% in all three control cohorts. Sex and the 12 broad geographical regions, provided by the WTCCC, were included as covariates in the analysis. The analyses were performed in the WTCCC type 1 diabetes vs. control datasets using forward stepwise conditional logistic regression until there were no significant signals remaining after correction for multiple testing.

Conditional analysis in LADA vs. population-based controls

Conditional logistic regression was performed using SNPTEST in the LADA versus population-based controls, including sex and the first 4 principal components as covariates.

Replication

To further validate MHC Class I independent effects in type 1 diabetes and lack of MHC Class I independent effects in LADA, we implemented approximate conditional analyses (COJO) in GCTA (18) on summary statistics from the Swedish replication cohort. Association analysis was

performed using SNPTEST, and sex and the first four PCs were used as covariates. There were 656 cases with type 1 diabetes vs 3,218 population-based controls, and 823 cases with LADA vs 3,211 population-based controls.

Sensitivity analysis

We performed sensitivity analysis to determine whether the lack of independent type 1 diabetes-associated signals in MHC Class I genes in LADA cases could be due to a lack of power. We randomly sampled 1,428 type 1 diabetes cases and 714 type 1 diabetes cases (subsets equating to the same size as the LADA cohort and half the size of the LADA cohort, respectively) and 2,219 controls to determine whether the type 1 diabetes-associated signals could be still be detected. Stepwise conditional logistic regression using SNPTEST was performed as above. To test the hypothesis that LADA is simply a mixture of type 1 diabetes and type 2 diabetes cases, we performed a further constrained conditional analysis in 714 randomly sampled type 1 diabetes cases and 714 randomly sampled type 2 diabetes cases (total n = 1,428 cases) and 2,219 WTCCC controls.

Further validating independent signals

PLINK was used to calculate pair-wise linkage disequilibrium (LD) between variants to further validate that the associated variants were truly independent of each other. To confirm independent association of HLA-B*39, the specific HLA-B*39 subtype HLA-B*3906 was tested in the WTCCC type 1 diabetes cases (n = 1985) vs controls (n = 2219) dataset using the presence of *DQB1**0402 and *DQB1**0501 as covariates.

Results

Confirming independent effects of MHC Class I signals in WTCCC type 1 diabetes vs Controls

Before conditioning, we observed rs3957146 as the strongest association signal in the type 1 diabetes vs WTCCC controls analysis ($P = 8.94 \times 10^{-165}$; **Figure 1A**). rs3957146 is in strong LD with a classical HLA subtype allele, *HLA-DQB1*0302* ($r^2 = 0.99$). After conditioning on the top signal, rs3957146, and subsequent independent MHC Class II signals (*HLA-DQB1*0201* and rs9268633), we observed the reported independent significant association of MHC Class I variants rs1610649 (*HLA-G*, $P = 6.89 \times 10^{-23}$) and *HLA-B*39* ($P = 6.89 \times 10^{-23}$) (**Figure 1B; Table 1; Supplementary Table 3; Supplementary Figure 2**). Conditioning on these variants in addition to the MHC class II variants, also demonstrated significant association with the *HLA-A* locus(rs9259852, $P = 2.04 \times 10^{-8}$).

Conditional analysis in LADA vs population-based controls

We then went on to perform stepwise conditional analysis in 1,428 LADA cases and 2,979 controls. Similar to observations in the type 1 diabetes vs WTCCC controls dataset, before conditioning on any variants, the strongest association signal in LADA vs population-based controls was also rs3957146 ($P = 1.80 \times 10^{-68}$; **Figure 2A**). Although we had 98% power to detect *HLA-B*39* with an allele frequency of 2% and an odds ratio of 2.5 (**Supplementary Table 2**), when conditioning on the most highly significant MHC Class II alleles (rs3957146, *HLA-DRB1*03*, rs9269081, *DRB1*0404* and *DQB1*0602*), there were no remaining independent signals in the MHC Class I region reaching significance after correction for multiple comparisons ($P < 8.83 \times 10^{-6}$; **Figure 2B**). Furthermore, we also noted independent effects in the MHC Class III region (rs2143462, $P = 8.24 \times 10^{-8}$) and the MHC Class II region (*HLA-DPA1*02*, $P = 1.62 \times 10^{-6}$ and *HLA-DPBI* variant rs3130192, $P = 5.32 \times 10^{-6}$), which are known to be

associated with type 1 diabetes(19). Here, *HLA-DPB1* variant is in strong LD with rs2301225 ($r^2=0.85$) and is independently associated with type 1 diabetes. MHC Class I variants were not observed to be independently associated with LADA after correcting for multiple comparison (**Table 1; Supplementary Figure 2; Supplementary Table 4**).

Sensitivity analysis in reduced sample of type 1 diabetes vs controls

To ensure that the lack of significant associations with MHC Class I genes in the LADA cohort was not explained by reduced power, we conducted a sensitivity analysis by systematically decreasing the sample size of the type 1 diabetes vs WTCCC control cohort to match the size of the discovery LADA cohort (n= 1428 type 1 diabetes cases and 2219 controls) and performing conditional analysis. Independent significant association signals at *HLA-G* ($P = 1.37 \times 10^{-17}$), *HLA-B* ($P = 5.58 \times 10^{-14}$) and *MUC22* (rs9262545, $P = 3.36 \times 10^{-9}$; rs9262547, $P = 9.69 \times 10^{-14}$) were still observed in this reduced type 1 diabetes sample size (**Supplementary Table 5**), although these signals were missing in the comparatively-sized LADA vs controls dataset. Similarly, independent significant association signals at *HLA-B* ($P = 1.26 \times 10^{-10}$), *HLA-G* ($P = 0.002$), and *MUC22* (rs9262545, $P = 1.57 \times 10^{-5}$) remained after further reducing the type 1 diabetes cohort size to equate with half the LADA cohort size (**Supplementary Table 6**).

Sensitivity analysis in an randomly mixed cohort of type 1 diabetes and type 2 diabetes cases vs controls

Another explanation for the lack of independent, significant associations across MHC Class I genes in LADA could be due to the possibility that the LADA cohort simply represents an approximately 50/50 mixture of misdiagnosed type 1 diabetes and type 2 diabetes cases. Therefore, we randomly sampled 714 cases with type 1 diabetes, 714 cases with type 2 diabetes and 2,219 controls, creating a “mixture” cohort. We performed the same conditional analysis

described above and observed that the *HLA-B*, *HLA-G* and two *MUCC2* signals in the MHC Class I regions remained independently significant in this mixed cohort, driven by the type 1 diabetes case subset (**Supplementary Table 7**).

Replication

We leveraged summary statistics data from Swedish cohorts to attempt replication of our findings. In type 1 diabetes vs controls, the strongest association was rs9275206 ($P = 6.35 \times 10^{-89}$), which is in strong LD with *HLA-DQB1*0302* ($r^2 = 0.99$). After conditioning on rs9275206 and subsequent top signals (**Supplementary Table 8**), we again observed significant association signals at the *HLA-G* ($P = 1.74 \times 10^{-10}$) and *HLA-B* (1.10×10^{-9}) loci. However, when conditional analysis was performed in LADA vs controls, there were no such signals across MHC Class I genes, and very sparse signals in the MHC Class II region (**Supplementary Table 9**). Furthermore, we observed a significant association signal at the *NOTCH4* (rs397081, $P = 1.11 \times 10^{-10}$) locus, the *MUC22* locus (rs9262545, $P = 7.83 \times 10^{-11}$ and rs9262547, $P = 7.17 \times 10^{-17}$) and the *HLA-A* locus (rs9259852, $P = 5.84 \times 10^{-14}$) (**Figure 2C**). Notably, rs9259852 is in strong LD with the classic HLA subtype allele, HLA-A*32 ($r^2 = 0.96$).

Further validating *HLA*B*39*

It has been shown that the *HLA-B*3906* allele is associated with a high risk of diabetes only for specific HLA-DR/DQ haplotypes, *DRB1*0801-DQB1*0402* and *DRB1*0101-DQB1*0501*(20). When specifically conditioning on these HLA-DR/DQ haplotypes in the WTCCC type 1 diabetes and control cohort, the independent significant association of the more specific *HLA-B*3906* subtype still remained (OR (95% CI) = 4.57 (3.08-6.80); $P = 5.84 \times 10^{-14}$).

Conclusion

The main objective of this study was to perform conditional analysis of the HLA region in LADA, which has been under explored to date in this disease context. The few genetic studies in LADA (5,21,22) only focused on the HLA Class II *DRB1* and *DQB1* haplotypes. Such studies, in populations of both European and Chinese ancestry, show that type 1 diabetes risk haplotypes are less frequent in LADA compared to childhood-onset type 1 diabetes cases, whereas type 1 diabetes protective haplotypes are more frequent in LADA, suggesting that LADA is a genetically attenuated form of type 1 diabetes. By extending the analysis of HLA in LADA beyond the MHC Class II region, we were able to observe further genetic differences between LADA and childhood-onset type 1 diabetes.

Although previous studies have reported MHC Class I independent effects in type 1 diabetes, those studies used directly HLA-typed cases and controls. Given the cost and challenges of direct HLA typing, we utilized the imputation tool SNP2HLA on genotyping data. SNP2HLA has been commonly used in the field to assess the genetics of autoimmune diseases (16,25,29,30). Furthermore, given this approach differs from that of Nejentsev *et al.*, it was crucial to first ensure that we could recapitulate the previously reported type 1 diabetes observations in the same cohort. First, we leveraged the WTCCC type 1 diabetes and control dataset, as a positive control, with previous studies identifying MHC Class I independent type 1 diabetes associations in the MHC Class I region (8,10,11,23). Since these studies were reported, imputation tools have allowed the analysis of the HLA region more cheaply and, in general, more practically. Before investigating MHC Class II independent LADA associations in the MHC Class I region, given the difference in our analytical approach, we recapitulated the observations in previous studies (10,11), by leveraging the same WTCCC type 1 diabetes and

control datasets. We confirmed that MHC Class I variants are significantly associated with type 1 diabetes, independent of the MHC Class II region using this imputation-based approach followed by stepwise conditional logistic regression. The conditional analysis was repeated in the LADA cohort, which consisted of cases and population-based controls. Crucially, there were no significant independent effects in the MHC Class I region remaining after correction for multiple comparisons; furthermore this observation was replicated in a separate Swedish cohort of type 1 diabetes cases, LADA cases and population-based controls.

The MHC Class I variant *HLA-B*39* is an established locus associated with type 1 diabetes risk (10,11,24). More specifically, studies suggested a strong association with type 1 diabetes for the subtype *HLA-B*3906*, which is now used in type 1 diabetes genetic risk scores to predict type 1 diabetes diagnosis (25). It has also been shown that the *B*3906* allele significantly enhances the risk of type 1 diabetes when present on specific *HLA-DR/DQ* haplotypes (e.g. *DRB1 0801-DQB1 0402* and *DRB1 0101-DQB1 0501*). The frequency of *HLA-B*3906* is different among different populations, and here did not survive our filter of having a MAF > 1% in the replication control cohort of Swedes. Thus, it was excluded in the analyses across the three cohorts. However, we confirmed that the *HLA-B *3906* allele remained significantly associated with type 1 diabetes after conditioning on the presence of the *DRB1 0801-DQB1 0402* and *DRB1 0101-DQB1 0501* haplotypes. Additionally, *HLA-B*3906* is associated with younger age-at-diagnosis in type 1 diabetes (9,11). A recent study using a NOD mouse model showed that *HLA-B*3906* mediates the development of CD8⁺ T cells required for type 1 diabetes onset; moreover, in the context of reduced immunological tolerance to insulin, *HLA-B*3906*-transgenic NOD mice develop type 1 diabetes at an accelerated rate (26). The lack of an independent *HLA-B*39*

association observed in the adult-onset phenotype of LADA further confirms the link between *HLA-B*39* with autoimmune progression with earlier onset of clinical disease.

HLA-B associations have been confirmed in a previous study (23), as well as associations around *HLA-G*, which is expressed in human pancreas (27) and may play a role in autoimmune progression (28). However, the MHC Class I variant rs1619379, located in *HLA-G* and ~100kb telomeric of *HLA-A*, may be less informative compared to *HLA-A* variants in predicting type 1 diabetes risk (10). This particular MHC Class I variant was independently significant in the downsampled type 1 diabetes cohort, but is in strong linkage disequilibrium with *HLA-G* variants, rs1610649 and rs2735028, which were significantly associated in the full type 1 diabetes set, the mixture cohort consisting of type 1 diabetes and type 2 diabetes cases, and the type 1 diabetes Swedish replication cohort. Additionally, the MHC Class I variants located in the *MUC22* locus have not been replicated in separate cohorts, and likely form haplotypes with HLA Class I alleles.

One limitation of this study was that we only tested variants with a MAF > 1% in all three control cohorts, which resulted in filtering out many informative alleles such as *HLA-B*3906*. By filtering to include only common alleles we limited potential discrepancies between populations, and were able to replicate our observations across cohorts with different frequencies of known risk variants. Furthermore, our study is limited in power to assess the underlying continuous traits of age at onset, time to insulin, and autoantibody titer; future well-designed, large studies are needed to enable those analyses. With larger and more complete data in individuals, we can then formally test the many competing hypotheses regarding the state of LADA in the field. The hypothesis that LADA exists as a different disease to type 1 diabetes with both overlapping genes and distinct genes is unlikely, as we did not clearly observe distinct

susceptible loci that were unique to LADA in this study or our previous GWAS (5). Future studies leveraging cases diagnosed with type 1 diabetes and LADA across the age of onset range will be crucial to test the remaining hypotheses, which are: i) LADA is type 1 diabetes with misdiagnosed type 2 cases that are false positive for autoantibodies ii) LADA cases are essentially type 1 diabetes at later onset with lower rates of progression iii) LADA is a form of diabetes where cases have both type 1 and type 2 risk alleles present at individual level. This first hypothesis (LADA is type 1 diabetes with misdiagnosed type 2 cases that are false positive for autoantibodies) motivated the sensitivity analysis, in which we randomly sampled cases from the WTCCC type 1 diabetes and type 2 diabetes cohort to create an random LADA cohort under the assumption that LADA would be a “mixture” of actual type 1 diabetes and type 2 diabetes cases. In this analysis, we still observed the same independent effects of MHC Class I variants, showing that the type 1 diabetes signature remained in the “mixture” cohort despite not being observed in LADA. However, we recognize that type 1 diabetes cases were sampled from the cohort of childhood-onset type 1 diabetes and to properly test this hypotheses LADA cases would have to be ascertained from another cohort, but also further stratified by autoantibody titre.

Although the underlying populations from which the type 1 diabetes and LADA sets were derived are the same, we addressed if the LD structure in the HLA region could be different between the two sample sets, which in turn could have resulted in inaccurate imputation. However, when we calculated LD for the MHC region in these datasets, we found it to be highly correlated (Pearson correlation coefficient = 0.97). Future studies are needed to address this question in depth as well as validate these findings in a cohort directly typed for MHC Class II and MHC Class I HLA alleles. Additionally, to further delineate this putative distinguishing

genetic feature between LADA and childhood-onset type 1 diabetes, it will be crucial to investigate how the HLA profile compares across the diabetes age continuum stratified for different autoantibody positivity status. A previous study observed different independent effects of MHC Class I variants to GAD autoantibodies and insulinoma-associated antigen-2 autoantibodies (31). Additionally, studies have shown that children with type 1 diabetes who are positive for a single autoantibody are more like to show type 2 diabetes features (32,33), for instance, a significant association with type 2 diabetes GWAS-implicated variants. Overall, our results point to key differences in the genetic signature in the MHC region, especially Class I markers, between LADA and childhood-onset type 1 diabetes. This study highlights the clinical utility of genetic screening in adult-onset diabetes that may be autoimmune in origin. The potential of defining these subjects who are at risk of rapid loss of insulin secretion using genetic characteristics could enable targeted immune-based, disease-modifying therapy.

ACKNOWLEDGEMENT

We gratefully acknowledge the participation of all the patients and control subjects. We would like to thank the JDRF, German Research Council (DFG: SFB 518, A1), German Diabetes Foundation, EUFP5 (Action LADA), All New Diabetes in Scania (ANDIS), Scania Diabetes Registry (SDR), and Malmö Diet and Cancer Study–Cardiovascular Cohort (MDC-CC) for providing samples for this research. We also would like to thank the Bone Mineral Density in Childhood Study, a multicenter, longitudinal study of bone accrual in healthy children. Authors are listed in Supplementary Material. This study makes use of data generated by the Wellcome Trust Case Control Consortium. A full list of the investigators who contributed to the generation of the data is available from www.wtccc.org.uk. Funding for the project was provided by the Wellcome Trust under award 076113. We acknowledge Stephen S. Rich, Marcella Devoto, and Andrew Wells for all the fruitful discussions.

FUNDING

D.L.C. is supported by the National Institutes of Health (1K99HD099330-1). D.M. is supported by CIBERDEM, Instituto de Salud Carlos III (Spain). B.O.B. is supported by the German Research Council (SFB 518, A1), LKC School of Medicine's Ong Tiong Tat Chair of Diabetes Research, the state of Baden-Württemberg, and a Ministry of Education, Singapore startup grant. S.F.A.G. is supported by the National Institutes of Health (R01-DK-085212) and the Children's Hospital of Philadelphia Daniel B. Burke Endowed Chair for Diabetes Research.

CONFLICTS OF INTEREST

The authors declare that there are no potential conflicts of interest relevant to this article.

Rajashree Mishra is currently employed by GlaxoSmithKline, Nanette C. Schloot is currently employed by Lilly Germany, and Vanessa C. Guy is currently employed by Science 37.

ETHICAL APPROVAL

This study was approved by local institutional ethical review boards.

AUTHOR CONTRIBUTIONS

Study concept and design: R.M., D.L.C., J.P.B, E.A., M.A., R.D.L., S.F.A.G.; Analysis and interpretation of data: R.M., M.A., D.L.C., R.D.L., S.F.A.G.; Resources: K.M.H., V.C.G., M.A., D.J.B., O.M., R.E.P., M.R.R., A.V., F.O., R.I.H., S.V., H.H., P.F., J.T.L., D.M., N.C.S., K.K., C.J.G., K.B.Y.,T.T., B.O.B., L.G., R.D.L., S.F.A.G.; Drafting and critical revision of the manuscript: R.M., D.L.C., E.A., M.A., A.C., N.C.S., B.O.A., B.F.V., B.O.B., R.D.L., S.F.A.G.; Obtained funding: S.S., B.O.B., R.D.L., S.F.A.G. All authors contributed to the final version of the manuscript. R.M., S.F.A.G. are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

References

1. Tuomi T, Carlsson A, Li H, Isomaa B, Miettinen A, Nilsson A, et al. Clinical and genetic characteristics of type 2 diabetes with and without GAD antibodies. *Diabetes* [Internet]. 1999;48(1):150–7. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/9892237>
2. Hawa MI, Kolb H, Schloot N, Beyan H, Paschou SA, Buzzetti R, et al. Adult-onset autoimmune diabetes in Europe is prevalent with a broad clinical phenotype: Action LADA 7. *Diabetes Care* [Internet]. 2013;36(4):908–13. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/23248199>
3. Mishra R, Hodge KM, Cousminer DL, Leslie RD, Grant SFA. A Global Perspective of Latent Autoimmune Diabetes in Adults. *Trends Endocrinol Metab* [Internet]. 2018;29(9):638–50. Available from: <https://doi.org/10.1016/j.tem.2018.07.001>
4. Ahlqvist E, Storm P, Karajamaki A, Martinell M, Dorkhan M, Carlsson A, et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *Lancet Diabetes Endocrinol* [Internet]. 2018; Available from: <https://www.ncbi.nlm.nih.gov/pubmed/29503172>
5. Cousminer DL, Ahlqvist E, Mishra R, Andersen MK, Chesi A, Hawa MI, et al. First Genome-Wide Association Study of Latent Autoimmune Diabetes in Adults Reveals Novel Insights Linking Immune and Metabolic Diabetes. *Diabetes Care* [Internet]. 2018;41(November):dc181032. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/30254083>
6. Alberti KG, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet Med* [Internet]. 1998;15(7):539–53. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/9686693>
7. Mishra R, Chesi A, Cousminer DL, Hawa MI, Bradfield JP, Hodge KM, et al. Relative contribution of type 1 and type 2 diabetes loci to the genetic etiology of adult-onset, non-insulin-requiring autoimmune diabetes. *BMC Med*. 2017;
8. Noble JA, Valdes AM, Bugawan TL, Apple RJ, Thomson G, Erlich HA. The HLA class I A locus affects susceptibility to type 1 diabetes. *Hum Immunol*. 2002;
9. Valdes AM, Erlich HA, Noble JA. Human leukocyte antigen class I B and C loci contribute to Type 1 Diabetes (T1D) susceptibility and age at T1D onset. *Hum Immunol*. 2005;
10. Howson JMM, Walker NM, Clayton D, Todd JA. Confirmation of HLA class II independent type 1 diabetes associations in the major histocompatibility complex including HLA-B and HLA-A. *Diabetes, Obes Metab*. 2009;
11. Nejentsev S, Howson JM, Walker NM, Szeszeko J, Field SF, Stevens HE, et al. Localization of type 1 diabetes susceptibility to the MHC class I genes HLA-B and HLA-A. *Nature* [Internet]. 2007;450(7171):887–92. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/18004301>
12. Kalkwarf HJ, Zemel BS, Gilsanz V, Lappe JM, Horlick M, Oberfield S, et al. The bone mineral density in childhood study: bone mineral content and density according to age, sex, and race. *J Clin Endocrinol Metab* [Internet]. 2007;92(6):2087–99. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/17311856>
13. Berndt SI, Skibola CF, Joseph V, Camp NJ, Nieters A, Wang Z, et al. Genome-wide association study identifies multiple risk loci for chronic lymphocytic leukemia. *Nat Genet*. 2013;

14. Wellcome Trust Case Control C. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* [Internet]. 2007;447(7145):661–78. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/17554300>
15. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* [Internet]. 2007;81(3):559–75. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/17701901>
16. Jia X, Han B, Onengut-Gumuscu S, Chen WM, Concannon PJ, Rich SS, et al. Imputing amino acid polymorphisms in human leukocyte antigens. *PLoS One* [Internet]. 2013;8(6):e64683. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/23762245>
17. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* [Internet]. 2007;39(7):906–13. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/17572673>
18. Yang J, Ferreira T, Morris AP, Medland SE, Madden PAF, Heath AC, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat Genet*. 2012;
19. Varney MD, Valdes AM, Carlson JA, Noble JA, Tait BD, Bonella P, et al. HLA DPA1, DPB1 alleles and haplotypes contribute to the risk associated with type 1 diabetes: Analysis of the type 1 diabetes genetics consortium families. *Diabetes*. 2010;
20. Baschal EE, Baker PR, Eyring KR, Siebert JC, Jasinski JM, Eisenbarth GS. The HLA-B*3906 allele imparts a high risk of diabetes only on specific HLA-DR/DQ haplotypes. *Diabetologia*. 2011;
21. Luo S, Lin J, Xie Z, Xiang Y, Zheng P, Huang G, et al. HLA genetic discrepancy between latent autoimmune diabetes in adults and type 1 diabetes: LADA China study No. 6. *J Clin Endocrinol Metab*. 2016;
22. Desai M, Zeggini E, Horton VA, Owen KR, Hattersley AT, Levy JC, et al. An association analysis of the HLA gene region in latent autoimmune diabetes in adults. *Diabetologia* [Internet]. 2007;50(1):68–73. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/17143607>
23. Eike MC, Becker T, Humphreys K, Olsson M, Lie BA. Conditional analyses on the T1DGC MHC dataset: Novel associations with type 1 diabetes around HLA-G and confirmation of HLA-B. *Genes Immun*. 2009;
24. Noble JA, Valdes AM, Varney MD, Carlson JA, Moonsamy P, Fear AL, et al. HLA class I and genetic susceptibility to type 1 diabetes: results from the Type 1 Diabetes Genetics Consortium. *Diabetes*. 2010;
25. Sharp SA, Rich SS, Wood AR, Jones SE, Beaumont RN, Harrison JW, et al. Development and standardization of an improved type 1 diabetes genetic risk score for use in newborn screening and incident diagnosis. *Diabetes Care* - Press. 2019;42(February):200–7.
26. Ali R, DiLorenzo TP, Racine JJ, Chapman HD, Serreze D V., Schloss J. HLA-B*39:06 Efficiently Mediates Type 1 Diabetes in a Mouse Model Incorporating Reduced Thymic Insulin Expression. *J Immunol*. 2018;
27. Cirulli V, Zalatan J, McMaster M, Prinsen R, Salomon DR, Ricordi C, et al. The class I HLA repertoire of pancreatic islets comprises the nonclassical class Ib antigen HLA-G. *Diabetes*. 2006;
28. Shiroishi M, Tsumoto K, Amano K, Shirakihara Y, Colonna M, Braud VM, et al. Human inhibitory receptors Ig-like transcript 2 (ILT2) and ILT4 compete with CD8 for MHC

- class I binding and bind preferentially to HLA-G. *Proc Natl Acad Sci*. 2003;
29. Karnes JH, Steiner HE, Phillips EJ, Roden DM, Bastarache L, Mosley JD, et al. Comparison of HLA allelic imputation programs. *PLoS One*. 2017;
 30. Hu X, Deutsch AJ, Lenz TL, Onengut-Gumuscu S, Han B, Chen WM, et al. Additive and interaction effects at three amino acid positions in HLA-DQ and HLA-DR molecules drive type 1 diabetes risk. *Nat Genet*. 2015;
 31. Howson JMM, Stevens H, Smyth DJ, Walker NM, Chandler KA, Bingley PJ, et al. Evidence that HLA class I and II associations with type 1 diabetes, autoantibodies to GAD and autoantibodies to IA-2, are distinct. *Diabetes*. 2011;
 32. Redondo MJ, Muniz J, Rodriguez LM, Iyer D, Vaziri-Sani F, Haymond MW, et al. Association of TCF7L2 variation with single islet autoantibody expression in children with type 1 diabetes. *BMJ Open Diabetes Res Care* [Internet]. 2014;2(1):e000008. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/25452857>
 33. Redondo MJ, Geyer S, Steck AK, Sosenko J, Anderson M, Antinozzi P, et al. TCF7L2 Genetic Variants Contribute to Phenotypic Heterogeneity of Type 1 Diabetes. *Diabetes Care* [Internet]. 2017; Available from: <https://www.ncbi.nlm.nih.gov/pubmed/29025879>

Table 1. Comparison of beta coefficients between conditional analyses in type 1 diabetes and LADA cohorts. Beta, and standard error(SE) are calculated for the risk allele frequency (RAF) from stepwise regression conditional on all SNP/HLA-Alleles in rows above (first column) in 1985 type 1 diabetes cases vs. 2219 controls. Beta coefficients and standard error for LADA versus controls correspond to maximum effect size in conditional analysis. P-value is derived from two sample Z test to formally test whether beta coefficients are significantly different or not. Forest plot is shown in Supplementary Figure 2. Three independent signals appeared in both type 1 diabetes and LADA conditional analyses (rs3957146, HLA-DRB1*0404, and rs9269081), however of these three signals only rs3957146 had a significant difference in effect size between type 1 diabetes and LADA (interaction p-value = 4.15×10^{-10}).

SNP/HLA-Allele	Locus	Position	Alleles (Risk/ Other)	WTCCC type 1 diabetes vs WTCCC Controls				LADA vs Controls				P-value
				RAF Cases	RAF Controls	Beta	SE	RAF Cases	RAF Controls	Beta	SE	
rs3957146	<i>HLA-DQA2</i>	32789508	T/C	0.385	0.113	1.44	0.05	0.251	0.101	1.14	0.06	1.22×10^{-4}
DQB1*0201	<i>HLA-DQB1</i>	32739039	P/A	0.338	0.140	1.56	0.06	0.209	0.119	1.05	0.07	3.17×10^{-8}
rs9268633	<i>HLA-DRA</i>	32514451	G/A	0.983	0.803	1.46	0.10	0.919	0.812	0.83	0.06	6.58×10^{-8}
rs1610649	<i>HLA-G</i>	29876896	G/A	0.616	0.582	0.61	0.06	0.592	0.586	0.16	0.05	8.33×10^{-9}
B*39	<i>HLA-B</i>	31431272	P/A	0.043	0.016	1.36	0.17	0.023	0.019	0.58	0.19	2.22×10^{-3}
DRB1*0404	<i>HLA-DRB1</i>	32660042	P/A	0.082	0.048	1.04	0.13	0.037	0.035	1.01	0.14	0.88
rs17427599	<i>HLA-DQB1</i>	32775342	T/C	0.849	0.755	0.59	0.09	0.818	0.775	0.31	0.07	1.41×10^{-2}
rs2301225	<i>HLA-DPA1</i>	33143838	T/C	0.941	0.891	0.72	0.11	0.924	0.886	0.42	0.08	2.74×10^{-2}
rs397081	<i>NOTCH4</i>	32300595	T/C	0.095	0.045	0.79	0.12	0.075	0.054	0.64	0.10	0.34
rs9262545	<i>MUC22</i>	31101041	A/G	0.913	0.881	0.67	0.11	0.862	0.858	0.09	0.07	8.65×10^{-6}
rs9262547	<i>MUC22</i>	31101206	T/A	0.135	0.119	1.59	0.19	0.137	0.142	-0.09	0.07	1.07×10^{-16}
rs9259852	<i>HLA-A</i>	30004400	C/T	0.977	0.959	1.00	0.18	0.968	0.963	0.19	0.13	2.64×10^{-4}
rs9269081	<i>HLA-DRA</i>	32549078	A/C	0.890	0.735	0.57	0.11	0.821	0.685	0.71	0.05	0.25
rs1978029	<i>HLA-DQB2</i>	32839688	C/T	0.651	0.537	0.36	0.08	0.565	0.475	0.39	0.05	0.75

Figure 1. Conditional analysis in 1985 type 1 diabetes cases and 2219 WTCCC controls. A) Logistic regression analysis without conditioning on MHC Class II alleles. B) Logistic regression analysis conditioning on MHC Class II alleles. C) Logistic regression analysis conditioning on MHC Class II and MHC Class I signals.

Figure 2. Conditional analysis in 1,428 LADA cases and 2,979 WTCCC controls. A) Logistic regression analysis without conditioning on MHC Class II alleles. B) Logistic regression analysis conditioning on MHC Class II alleles.