Published in final edited form as:

Diabetes Metab Res Rev. 2016 March ; 32(3): 297–307. doi:10.1002/dmrr.2720.

T-cell receptor repertoire variation may be associated with type 2 diabetes mellitus in humans

Joseph A. Frankl¹, Marie S. Thearle¹, Cindy Desmarais², Clifton Bogardus¹, and Jonathan Krakoff^{1,*}

¹Obesity and Diabetes Clinical Research Section, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Phoenix, AZ, USA

²Adaptive Biotechnologies, Seattle, WA, USA

Abstract

Background—Recent work in Pima Indians, a population with high rates of obesity and type 2 diabetes mellitus (T2DM), demonstrated that human leukocyte antigen haplotype DRB1*02 carriers have an increased acute insulin response and decreased risk for the development of T2DM, implicating loss of self-tolerance in the pathogenesis of T2DM. Advances in genomic sequencing have made T-cell receptor repertoire analysis a practical mode of investigation.

Methods—High-throughput sequencing of T-cell receptor complementarity-determining region 3 was carried out in male Pima Indians with normal glucose regulation ($n = 11$; age = 31 ± 8 years; %fat = $30.2 \pm 8.7\%$) and the protective DRB1*02 haplotype *versus* those with T2DM without DRB1*02 ($n = 7$; age = 34 ± 8 years; %fat = $31.2 \pm 4.7\%$). Findings were partially replicated in another cohort by assessing the predictive ability of T-cell receptor variation on risk of T2DM in Pima Indian men ($n = 27$; age = 28.9 ± 7.1 years; %fat = $28.8 \pm 7.1\%$) and women ($n = 20$; age = 29 ± 7.0 years; %fat = $37.1 \pm 6.8\%$) with baseline normal glucose regulation but without the protective haplotype who were invited to follow-up examinations as frequently as every 2 years where diabetes status was assessed by a 75-g oral glucose tolerance test. Of these subjects, 13 developed diabetes.

Results—T-cell receptor complementarity-determining region 3 length was shorter in those with T2DM, and a one-nucleotide decrease in complementarity-determining region 3 length was associated with a nearly threefold increase in risk for future diabetes. The frequency of one variable gene, *TRBV7-8*, was higher in those with T2DM. A 1% increase in *TRBV7-8* frequency was associated with a greater than threefold increase in diabetes risk.

Conclusions—These results indicate that T-cell autoimmunity may be an important component in progression to T2DM in Pima Indians.

*Correspondence to: Jonathan Krakoff, 5th Floor, Phoenix Indian Medical Center, 4212 N. 16th St, Phoenix, AZ 85016, USA. jkrakoff@mail.nih.gov.

This article has been contributed to by US Government employees and their work is in the public domain in the USA.

Conflicts of interest
None declared.

Keywords

autoimmunity; human; Pima Indians; survival analysis; T-cell receptor; type 2 diabetes mellitus

Introduction

Diabetes has traditionally been classified into two main types: an autoimmune disorder (type 1 diabetes mellitus) characterized by T-cell-mediated destruction of pancreatic beta cells [1] and a non-autoimmune metabolic disease (type 2 diabetes mellitus) preceded by insulin resistance that leads to pancreatic beta-cell failure [2]. There are individuals who present as phenotypically similar to type 2 diabetes mellitus but who have measureable antibody titres and islet cell reactive T-cells, attributes indicative of autoimmunity [3–6]. These individuals with apparently latent autoimmune diabetes do not fit into the traditional classification scheme and have intermediate autoimmune responses to beta-cell antigens [7]. A subgroup of autoantibody negative but autoreactive T-cell positive individuals has also been identified [8] indicating that islet cell antibodies are not absolutely necessary for adaptive immune system participation in development of diabetes. T-cells may mediate the chronic low-grade inflammation that accompanies obesity and promotes decreased insulin action [9,10].

Pima Indians have high rates of type 2 diabetes mellitus and obesity, and development of diabetes is usually considered to be primarily a metabolic disorder in this population [11]. GAD antibodies (GAD65Ab) [12] and other known islet cell antibodies [13] are not found at elevated levels in this population. However, our group recently identified a human leukocyte antigen (HLA) haplotype in Pima Indians, *HLA-DRB1*02*, that is protective against type 2 diabetes mellitus and is associated with increased insulin secretion [14] indicating a potential role for loss of self-tolerance in the development of type 2 diabetes mellitus.

T-cell receptor antigen recognition may play a defining role in pancreatic beta-cell destruction [3,4]. Advances in genomic sequencing have made T-cell receptor repertoire analysis practical and affordable. Sequencing can be restricted to the complementarity-defining region 3 because peptide contact and recognition are primarily made within this hypervariable region [15]. Complementarity-determining region 3 (CDR3) variability arises from somatic recombination of one each of 54 Variable, 2 Diversity and 13 Joining genes plus addition and deletion of nucleotides at the V–D and D–J junctions [16]. However, these recombinations are not completely random, as healthy adult humans have elevated frequencies of a similar subset of V and J genes, and it is estimated that a healthy adult human has $\sim 3 \times 10^6$ T-cell receptor CDR3 sequences, a small fraction of the theoretically possible sequences [17,18].

We hypothesized that there would be differences in the T-cell receptor repertoire between full-heritage Pima Indians with normal glucose regulation and the protective *HLA-DRB1*02* haplotype [14] and individuals with type 2 diabetes mellitus without the haplotype. We also hypothesized that at least some of the T-cell receptor differences evident in the first cohort would be predictive of type 2 diabetes mellitus onset in a larger cohort of Pima Indians with normal glucose regulation at baseline but without the protective haplotype.

Materials and methods

Subject selection

Volunteers participated in a longitudinal inpatient study examining risk factors for type 2 diabetes mellitus at the Obesity and Diabetes Clinical Research Section of the National Institutes of Diabetes and Digestive and Kidney Diseases in Phoenix, AZ. Upon admission, volunteers were fed a weight-maintaining diet [19]. On day 2 of admission, participants underwent measurement of body composition using dual energy X-ray absorptiometry (DPX-L; Lunar Corp), and per cent body fat (%fat), fat free mass and fat mass were calculated as previously described [20]. On day 4, a 75-g oral glucose tolerance test was administered. Diabetes and glucose tolerance were determined as per 2003 American Diabetes Association guidelines [21]. Glucose concentrations were measured using the glucose oxidase method (Beckman Instruments, Fullerton, CA). Blood samples for DNA extraction were collected with the fasting blood draw during the oral glucose tolerance test. Participants subsequently had a measurement of acute insulin secretion using an IV glucose tolerance test and insulin action using the euglycemic-hyperinsulinemic clamp, as described previously [22–24].

Volunteers also participated in an outpatient longitudinal study of health where participants were invited to examinations as frequently as every 2 years and asked to fast overnight. Follow-up diabetes status was determined from these examinations, which all included a 75-g oral glucose tolerance test. Glucose concentrations for the out-patient study were measured by the hexokinase method (Ciba-Corning, Palo Alto, CA). The purposes of both studies were discussed with the volunteers, and written informed consent was obtained for each study before participation. The Institutional Review Board of the National Institutes of Diabetes and Digestive and Kidney Diseases approved both protocols.

Two separate groups were selected for T-cell repertoire analysis from among participants in the inpatient phenotype study (participants were studied between 1983 and 2007). No effect of storage time on results of the T-cell receptor high-throughput sequencing was observed. The first group was selected to maximize differences between comparators as an exploratory analysis to determine whether any differences existed in the T-cell repertoire. This exploratory group included 11 male full-heritage Pima Indian volunteers with both the HLA haplotype DRB1*02 and normal glucose regulation and seven full-heritage Pima Indian male volunteers without the *HLA-DRB1*02* haplotype and with type 2 diabetes mellitus. The second group was selected as a validation analysis to verify that results found with the first group were related to diabetes development rather than HLA type and to explore associations with phenotypic measures. The validation cohort included 49 full-heritage Pima men ($n = 29$) and women ($n = 20$) with normal glucose regulation but without the protective *HLA-DRB1*02* haplotype. Within this validation group, 88% ($n = 43$) had at least one follow-up visit with data from a 75-g oral glucose tolerance test with an average follow-up time of 7.2 ± 4.2 years. Two subjects, both male, were subsequently excluded from analyses because of T-cell repertoire reports that failed to meet quality control standards. Of the remaining 41, 13 developed type 2 diabetes mellitus.

High-throughput sequencing

TCR β CDR3 regions, the region between the conserved 3' cysteine in the V gene and phenylalanine in the J gene [25], were amplified and sequenced from 600 ng of genomic DNA using the immunoSEQ assay (Adaptive Biotechnologies, Seattle, WA) with primers specific for all 54 known expressed V β and all 13 known expressed J β genes [16]. A standard algorithm was used to identify V, D and J genes as well as inserted and deleted nucleotides within each sequence [25], and the copy number was determined after merging erroneous sequences derived from polymerase chain reaction (PCR) and/or sequencing errors with the counts of true biological sequences and adjusting for PCR bias. Specifically, PCR bias was accounted for by sequencing a library of constructs consisting of all V/J pairs, which are represented at equal amounts, to look for and account for differential amplification. Productive read counts ranged from 182 786 to 5 383 392, and unique clonotype counts, defined by unique amino acid sequences, ranged from 4099 to 109 886 per subject. Reads that were out of frame or had a stop codon totaled 42 178 to 891 559 and 805 to 3121 unique clonotypes per subject, respectively.

Statistics

Given the hypothesis-generating nature of the exploratory cohort and the confirmation of these results with the validation cohort, a p-value less than 0.05 was considered statistically significant, and we did not adjust for multiple comparisons. Baseline characteristics of individuals with type 2 diabetes mellitus *versus* normal glucose regulation in the exploratory cohort and diabetes progressors *versus* non-progressors in the validation cohort were assessed using Student's *t*-test for continuous variables and Fisher's exact test for categorical variables. Baseline demographics were described using means and standard deviations for data that were normally distributed and medians with interquartile ranges for skewed data. To meet the assumptions of linear regression, skewed data were log transformed for analyses.

In both cohorts, the clonotype frequency distribution was heavily skewed towards low frequency clones. Before assessing individual TR β V, -D and -J gene frequencies, we filtered the data so as to eliminate low frequency clonotypes by visual inspection of the clonotype frequency distribution. Gene frequencies were determined as the percentage of productive sequences in an individual's analysed repertoire for that particular gene. Low frequency was defined as clonotypes representing less than 0.005% of the productive repertoire in the exploratory analysis and less than 0.0015% in the validation analysis, values similar to those that have been previously used [26], because of assay differences. Student's *t*-tests weighted by total normalized number of read copies were used to find differences in TR β V, -D and -J gene frequency among clonotypes between individuals with type 2 diabetes mellitus *versus* normal glucose regulation in the exploratory analysis, and results that met the $p < 0.05$ threshold were analysed in the validation cohort between diabetes progressors and non-progressors. Linear regression analyses were used to analyse the associations between gene frequencies and continuous variables.

Differences in CDR3 length, defined as the number of nucleotides between the conserved 3' cysteine in the V gene and the phenylalanine in the J gene [25], and net CDR3 insertions and

deletions between groups in both the exploratory and validation analyses were assessed using mixed models weighted by the normalized number of read copies with current or future diabetes status as a classification variable and accounting for repeated measure. Linear regression was used to analyse associations between mean CDR3 length and CDR3 insertions minus deletions with continuous variables. In the validation cohort, Cox proportional hazards models weighted by the total normalized number of read copies were used to examine the predictive power of mean CDR3 length and net insertions and deletions on the risk of developing type 2 diabetes mellitus. Cox proportional hazards models were also used to assess the predictive value of differences in gene frequencies identified in the exploratory analysis on developing type 2 diabetes mellitus. Time was calculated from the date of the blood draw used for T-cell receptor repertoire analysis to the onset of diabetes or to the last follow-up visit. Age, sex and %fat were serially added to proportional hazards models to assess whether they attenuated the results. Cumulative incidence for development of type 2 diabetes mellitus was analysed for those above and below the *TRBV7-8* frequency and CDR3 length median, separately and then combined (e.g. four groups encompassing above/below medians for each variable).

Diversity of the T-cell receptor repertoires was calculated using Simpson's Diversity Index [27]. The sample sizes were normalized by performing a random draw procedure to create a dataset equal to smallest sample in the cohort (182 786 reads in the exploratory group and 184 103 in the validation group), and this was repeated 999 times. Simpson's Diversity Index was then generated for all 1000 representative datasets, and then the median diversity score was generated as the final diversity index for each subject. Correlations between diversity scores and continuous variables were calculated using Spearman's rank correlation coefficient, and diversity score differences between groups were assessed with the Wilcoxon rank-sum test. Analyses were performed using SAS 9.2 and SAS Enterprise Guide 4.2 (Cary, NC).

Results

Exploratory analysis

Baseline characteristics of subjects included in the exploratory analysis are shown in Table 1. The clonotype frequency distribution was heavily skewed towards low frequency clones in those with both type 2 diabetes mellitus and normal glucose regulation. Even when low frequency clonotypes (<0.005%) were excluded, mean V gene frequencies varied more than 1000-fold (Figure 1A). The 13 J genes' frequencies had a considerably smaller order of magnitude variation with extremes of 13.7926% and 1.6852% (Figure 1B). The TRBD1-1 : TRBD1-2 ratio was approximately 2 : 1 in both the type 2 diabetes mellitus and normal glucose regulation groups. Four TR β V and one TR β J genes had higher frequencies in subjects with type 2 diabetes mellitus and without the protective *HLA-DRB1*02* haplotype, whereas three TR β V and two TR β J genes were represented at lower frequencies in those with type 2 diabetes mellitus (Figure 1).

Complementarity-determining region 3 length ranged from 15 to 60 nucleotides and was normally distributed with a mode of 36 nucleotides. CDR3 length was significantly shorter in those with type 2 diabetes mellitus than those with normal glucose regulation (mean $\pm$

standard error, 36.92 ± 0.02 vs 37.34 ± 0.02 nucleotides; $p < 0.0001$; Figure 2). The net number of insertions and deletions in this region was more negative in those with type 2 diabetes mellitus consistent with a greater number of deletions ($-7.50 \pm .02$ vs -7.14 ± 0.02 nucleotides; $p < 0.0001$).

T-cell repertoire diversity was quantified using the Simpson Index, which gives the probability that any two T-cell receptors chosen at random from a sample will have different CDR3 amino acid sequences [28]. The diversity score did not differ between groups and was not correlated with %fat in either group. There was a negative correlation between the diversity score and age in both groups that did not reach statistical significance (Figure 5A).

Validation analysis

All significant gene frequency differences from the initial exploratory analysis were also assessed in a larger cohort ($n = 47$) followed as part of a longitudinal study seeking to determine environmental and genetic determinants of diabetes (Figure 3). Follow-up data were available for 41 of the 47 subjects over an average of 7.2 ± 4.2 years, among whom 13 developed type 2 diabetes mellitus. Baseline characteristics of the subjects included in the validation analysis are shown in Table 2.

The V, D and J gene frequencies were similar in this group when compared with the exploratory cohort (Figure 3A and B). Of the ten genes that significantly differed by frequency in the exploratory analysis, only *TRBV7-8* was more frequent in those who developed diabetes. *TRBJ1-4* was also significantly different but in the validation analysis was more frequent in those who did not develop diabetes in opposition to the results of the exploratory analysis. In a proportional hazards model, we found that increased frequency of *TRBV7-8* predicted development of type 2 diabetes mellitus (hazard rate ratio (HRR) = 3.13 per 1% increase in *TRBV7-8*; $p < 0.0001$). Serial addition of age, sex and %fat did not attenuate the HRR of *TRBV7-8* percentage in the model.

Similar to the findings in the exploratory analysis, CDR3 length was also shorter in diabetes progressors versus non-progressors (37.101 ± 0.006 vs 37.304 ± 0.004 , $p < 0.0001$). In a proportional hazards model, shorter mean CDR3 length predicted development of type 2 diabetes mellitus (HRR = 2.99 per 1 nucleotide decrease in CDR3 length; $p < 0.0001$). This effect was also seen when assessing the net number of insertions and deletions (HRR = 3.32 per 1 nucleotide decrease in the direction of more deletions; $p < 0.0001$). Serial addition of age, sex and %fat did not attenuate the HRR of either CDR3 length or the net number of insertions and deletions in these models. The mean number of net insertions and deletions strongly correlated with the mean CDR3 length ($\rho = 0.90$, $p < 0.0001$). Eleven-year cumulative incidence for type 2 diabetes mellitus was higher in those above the *TRBV7-8* frequency and below the CDR3 length medians (Figure 4A and B). Furthermore, as shown in Figure 4C, an additive effect for those with higher *TRBV7-8* frequency and shorter CDR3 length was observed. However, there was no significant difference in diabetes incidence in those above compared with below the *TRBV7-8* frequency median within the CDR3 long group or in those above compared with below the CDR3 length median within the *TRBV7-8* low group. In a proportional hazards model with *TRBV7-8* frequency and CDR3 length, both parameters predicted development of type 2 diabetes mellitus (HRR =

3.20 per 1% increase in *TRBV7-8*, $p < 0.0001$ and HRR = 1.99 per 1 nucleotide decrease in the direction of more deletions; $p < 0.0001$). CDR3 length did not correlate with *TRBV7-8* frequency.

Simpson's Diversity Index did not differ between type 2 diabetes mellitus progressors and non-progressors or by gender. There was no correlation between diversity and %fat. The diversity index was, however, negatively correlated with age ($\rho = -0.39$, $p = 0.0063$) in the validation cohort (Figure 5B).

In a regression model adjusted for age, sex and %fat, neither *TRBV7-8* frequency, mean CDR3 length nor the diversity score was associated with baseline insulin action (M). In a regression model adjusted for age, %fat, M and sex, there were no associations with acute insulin release. *TRBV7-8* and CDR3 length were also not associated with %fat after adjusting for age and sex.

Discussion

Using high-throughput sequencing of T-cell receptor CDR3 regions, we examined differences in T-cell repertoires that may be associated with type 2 diabetes mellitus in an exploratory analysis and partially replicated our findings in a validation analysis. T-cell receptor repertoire parameters, including TR β V and -J gene frequencies and CDR3 length, were identified that differed between groups in the exploratory cohort. Differences in TR β V7-8 frequency and CDR3 length were replicated in the validation cohort between type 2 diabetes mellitus and nonprogressors, and both increased risk for type 2 diabetes mellitus onset.

Prior studies have identified individuals with phenotypic type 2 diabetes mellitus who display signs of islet cell autoimmunity and specifically T-cell reactivity. Subjects with type 2 diabetes mellitus and the presence of autoantibodies were found to have peripheral mononuclear cell activity to islet antigens [7]. Furthermore, in subjects with type 2 diabetes mellitus and autoantibodies, T-cell reactivity to islet antigens specifically correlated with impairment in beta-cell function [3]. T-cell reactivity to islet antigens has also been demonstrated in antibody negative subjects with type 2 diabetes mellitus [8]. Our results extend and build on these previous studies by using an alternative method that involved direct sequencing of the variable region of the T-cell receptor and allowed us to identify specific T-cell receptor differences that provide potential mechanistic information on how autoimmunity might develop. We were also able to confirm two of our cross-sectional observations in a prospective study, strengthening our findings. Pima Indians have been characterized as a population with 'classic' type 2 diabetes mellitus, in whom metabolic risk factors, including increased adiposity, reduced insulin action and insulin secretion, are known to predict diabetes. Thus, finding an association between T-cell receptor repertoire parameters and diabetes in this population provides a unique intersection between metabolic and possible autoimmune mediators in the development of type 2 diabetes mellitus.

It should be noted that beta-cell destruction is not the only mechanism through which T-cell activity could contribute to diabetes development. Although we did not find an association

between *TRBV7–8* frequency or CDR3 length with insulin action, other studies in both rodents and humans have shown that adipose tissue T-cells appear to play a role in the chronic low-grade inflammation and worsening insulin action associated with increased adiposity. CD4⁺Foxp3⁺, the most potent regulatory cells and a defence against inappropriate immune responses, are depleted in obesity [9]. Other studies indicate a role for both CD4⁺ T helper type 2 and CD4⁺Foxp3⁺ cells in mediating insulin action [10]. CD8⁺ cells appear to precede macrophage infiltration of adipose tissue, and depletion of CD8⁺ cells prevents macrophage infiltration and improves insulin action [29].

Consistent with previous studies [17,18,30], V, D and J gene frequencies were non-random. We found that *TRBV20–1*, *TRBV5–1* and *TRBV19*, and *TRBJ2–1*, *TRBJ2–3* and *TRBJ2–7* were among the most frequently used V and J genes in this group of Pima Indians. *TRBD1–1* was the more common D gene with a frequency greater than 50% in all subjects, which has not been observed in all other populations [17] and may be specific to Pima Indians. Four TRβV and one TRβJ genes occurred with significantly greater frequency among the repertoires of individuals with type 2 diabetes mellitus and without the protective *HLA-DRB1*02* haplotype, and three TRβV and two TRβJ genes were more frequent among the repertoires of individuals with normal glucose regulation. We anticipated that the majority of the over-represented V and J genes would be found among subjects with type 2 diabetes mellitus compared with subjects with normal glucose regulation due to theorized oligoclonal expansions of autoreactive T-cells [31,32]. It is possible that we did not see this because of the small sample size in the exploratory cohort. Of the ten genes that differed in frequency in the exploratory analysis, only *TRBV7–8* had a higher frequency in type 2 diabetes mellitus progressors in the validation cohort. The proportional hazards analysis indicated that a 1% increase in frequency of *TRBV7–8* was associated with a threefold increase in diabetes risk.

Prior to the advent of next-generation sequencing, CDR3 length spectratyping was used to identify antigen-driven clonotype expansions (thus producing loss of diversity) within T-cells expressing particular TRβV genes [33]. Now, with a larger number of sequences available for analysis, Simpson's Diversity Index provides a better assessment of that characteristic. CDR3 length may also serve as a global measurement of repertoire differences. CDR3 length was shorter in those with type 2 diabetes mellitus than normal glucose regulation in the exploratory cohort, and shorter CDR3 length predicted development of type 2 diabetes mellitus in the validation cohort. Consistent with these findings, the net number of inserted and deleted nucleotides in the CDR3 sequence was more negative (overall more deletions) among those with type 2 diabetes mellitus in the exploratory cohort and type 2 diabetes mellitus progressors in the validation cohort. Shorter CDR3 length may indicate relative proliferation of autoreactive T-cells via selective activation and survival of particular T-cell subsets, consistent and subtle differences in TRβV and J gene use or alterations in recombination machinery [34].

Although peripheral T-cell diversity was not directly associated with diabetes status, it was negatively associated with age in the larger validation cohort. This cohort was relatively young with a mean age under 30. Loss of self-tolerance appears to occur with ageing. The CD8⁺ fraction declines with age, while the CD4⁺ fraction does not [35–37]. Higher avidity clonotypes may increase in frequency [37] as the body tries to maintain normal immune

function despite loss of T-cell numbers [38]. relationship with age was not seen in the smaller exploratory cohort, likely because of the smaller sample size. It is also possible that those with the *HLA-DRB1*02* haplotype (present only in those with normal glucose regulation in the exploratory group) are protected from an age-related decline in diversity. Although Pima Indians develop type 2 diabetes mellitus at younger age than other populations, age is still an important risk factor for development of type 2 diabetes mellitus in this group [39]. Loss of T-cell diversity with age, especially if more rapid in those without the protective HLA allele, might contribute to age-related increase in diabetes risk.

Several limitations in our study need to be acknowledged. First, the T-cell receptor repertoires were analysed in two different batches. However, despite this, we were able to confirm our initial findings in the validation cohort, which only strengthens the confidence in our observations. Our validation cohort only allowed for partial replication of findings in our exploratory cohort because this group included only those without the *HLA-DRB1*02* allele. However, this allowed us to evaluate if the findings on the exploratory cohort were specifically related to development of diabetes. Because we used genomic DNA, we could not distinguish between the CD4⁺ and CD8⁺ fractions. However, analysing genomic DNA allowed for precise quantification of T-cell count [40] and analysis of T-cell repertoires from subjects with up to 17 years of follow-up. We also did not have antibody measurements from our subjects. However, direct beta-cell destruction is more likely to be a T-cell-mediated event [4], so direct examination of the T-cell receptor repertoire may lead to discovery of novel pathways to beta-cell damage. In addition, all analyses were performed in a limited number of full-heritage Pima Indians, which may limit the generalizability of our findings. However, the pathophysiology of type 2 diabetes mellitus in Pima Indians mirrors that of other populations. Nevertheless, replication is needed in a larger more diverse cohort or in other populations at risk for type 2 diabetes.

We used a novel approach to assess the peripheral blood T-cell repertoire. In both an exploratory and a validation analysis, we found that frequency of a specific V gene, *TRBV7-8*, and CDR3 length were associated with and predictive of the development of type 2 diabetes mellitus. These findings provide additional evidence indicating that, even in a population considered to have classic type 2 diabetes mellitus, there may be an important autoimmune component.

Acknowledgements

The authors thank the volunteers who participated in the studies. They also thank the clinical staff of the Phoenix Epidemiology and Clinical Research Branch for conducting the examinations.

This work was supported by the Intramural Research Program of the National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases.

C. D. owns stock in and is an employee of Adaptive Biotechnologies. No other potential conflicts of interest relevant to this article were reported.

References

1. Bluestone JA, Herold K, Eisenbarth G. Genetics, pathogenesis and clinical interventions in type 1 diabetes. *Nature* 2010; 464(7293): 1293–1300. [PubMed: 20432533]

2. DeFronzo RA. Pathogenesis of type 2 diabetes mellitus. *Med Clin North Am* 2004; 88(4): 787–835. [PubMed: 15308380]
3. Goel A, Chiu H, Felton J, Palmer JP, Brooks-Worrell B. T-cell responses to islet antigens improves detection of auto-immune diabetes and identifies patients with more severe beta-cell lesions in phenotypic type 2 diabetes. *Diabetes* 2007; 56(8): 2110–2115. [PubMed: 17473222]
4. Brooks-Worrell B, Narla R, Palmer JP. Islet autoimmunity in phenotypic type 2 diabetes patients. *Diabetes Obes Metab* 2013; 15(Suppl 3): 137–140. [PubMed: 24003930]
5. Itariu BK, Stulnig TM. Autoimmune aspects of type 2 diabetes mellitus – a mini-review. *Gerontology* 2014; 60: 189–196. [PubMed: 24457898]
6. Wentworth JM, Furlanos S, Harrison LC. Reappraising the stereotypes of diabetes in the modern diabetogenic environment. *Nat Rev Endocrinol* 2009; 5(9): 483–489. [PubMed: 19636326]
7. Brooks-Worrell BM, Juneja R, Minokadeh A, Greenbaum CJ, Palmer JP. Cellular immune responses to human islet proteins in antibody-positive type 2 diabetic patients. *Diabetes* 1999; 48(5): 983–988. [PubMed: 10331401]
8. Brooks-Worrell BM, Reichow JL, Goel A, Ismail H, Palmer JP. Identification of autoantibody-negative autoimmune type 2 diabetic patients. *Diabetes Care* 2011; 34(1): 168–173. [PubMed: 20855551]
9. Feuerer M, Herrero L, Cipolletta D, et al. Fat Treg cells: a liaison between the immune and metabolic systems. *Nat Med* 2009; 15(8): 930–939. [PubMed: 19633656]
10. Winer S, Chan Y, Paltser G, et al. Normalization of obesity-associated insulin resistance through immunotherapy. *Nat Med* 2009; 15(8): 921–929. [PubMed: 19633657]
11. Knowler WC, Pettitt DJ, Savage PJ, Bennett PH. Diabetes incidence in Pima Indians: contributions of obesity and parental diabetes. *Am J Epidemiol* 1981; 113(2): 144–156. [PubMed: 7468572]
12. Dabelea D, Palmer JP, Bennett PH, Pettitt DJ, Knowler WC. Absence of glutamic acid decarboxylase antibodies in Pima Indian children with diabetes mellitus. *Diabetologia* 1999; 42(10): 1265–1266. [PubMed: 10525671]
13. Knowler WC, Bennett PH, Bottazzo GF, Doniach D. Islet cell antibodies and diabetes mellitus in Pima Indians. *Diabetologia* 1979; 17(3): 161–164. [PubMed: 389722]
14. Williams RC, Muller YL, Hanson RL, et al. HLA-DRB1 reduces the risk of type 2 diabetes mellitus by increased insulin secretion. *Diabetologia* 2011; 54(7): 1684–1692. [PubMed: 21484216]
15. Rudolph MG, Stanfield RL, Wilson IA. How TCRs bind MHCs, peptides, and coreceptors. *Annu Rev Immunol* 2006; 24(1): 419–466. [PubMed: 16551255]
16. Robins HS, Campregher PV, Srivastava SK, et al. Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. *Blood* 2009; 114(19): 4099–4107. [PubMed: 19706884]
17. Robins HS, Srivastava SK, Campregher PV, et al. Overlap and effective size of the human CD8+ T cell receptor repertoire. *Sci Transl Med* 2010; 2(47): 47ra64.
18. Sherwood AM, Desmarais C, Livingston RJ, et al. Deep sequencing of the human TCR γ and TCR β repertoires suggests that TCR β rearranges after $\alpha\beta$ and $\gamma\delta$ T cell commitment. *Sci Transl Med* 2011; 3(90): 90ra61.
19. Ferraro R, Boyce VL, Swinburn B, De Gregorio M, Ravussin E. Energy cost of physical activity on a metabolic ward in relationship to obesity. *Am J Clin Nutr* 1991; 53(6): 1368–1371. [PubMed: 2035463]
20. Tataranni PA, Ravussin E. Use of dual-energy X-ray absorptiometry in obese individuals. *Am J Clin Nutr* 1995; 62(4): 730–734. [PubMed: 7572700]
21. Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Report of the expert committee on the diagnosis and classification of diabetes mellitus. *Diabetes Care* 2003; 26(suppl 1): s5–s20. [PubMed: 12502614]
22. Bunt JC, Krakoff J, Ortega E, Knowler WC, Bogardus C. Acute insulin response is an independent predictor of type 2 diabetes mellitus in individuals with both normal fasting and 2-h plasma glucose concentrations. *Diabetes Metab Res Rev* 2007; 23(4): 304–310. [PubMed: 17001619]
23. Bogardus C, Lillioja S, Howard BV, Reaven G, Mott D. Relationships between insulin secretion, insulin action, and fasting plasma glucose concentration in nondiabetic and noninsulin-dependent diabetic subjects. *J Clin Invest* 1984; 74(4): 1238–1246. [PubMed: 6384267]

24. Lillioja S, Bogardus C. Obesity and insulin resistance: lessons learned from the Pima Indians. *Diabetes Metab Rev* 1988; 4(5): 517–540. [PubMed: 3061759]
25. Yousfi Monod M, Giudicelli V, Chaume D, Lefranc M-P. IMGT/JunctionAnalysis: the first tool for the analysis of the immunoglobulin and T cell receptor complex V-J and V-D-J JUNCTIONS. *Bioinforma Oxf Engl* 2004; 20(Suppl 1): i379–i385.
26. Marrero I, Hamm DE, Davies JD. High-throughput sequencing of islet-infiltrating memory CD4(+) T cells reveals a similar pattern of TCR V β usage in prediabetic and diabetic NOD mice. *PLoS One* 2013; 8(10): e76546. [PubMed: 24146886]
27. Simpson EH. Measurement of diversity. *Nature* 1949; 163: 688.
28. Venturi V, Kedzierska K, Turner SJ, Doherty PC, Davenport MP. Methods for comparing the diversity of samples of the T cell receptor repertoire. *J Immunol Methods* 2007; 321(1–2): 182–195. [PubMed: 17337271]
29. Nishimura S, Manabe I, Nagasaki M, et al. CD8+ effector T cells contribute to macrophage recruitment and adipose tissue inflammation in obesity. *Nat Med* 2009; 15(8): 914–920. [PubMed: 19633658]
30. Serana F, Sottini A, Caimi L, et al. Identification of a public CDR3 motif and a biased utilization of T-cell receptor V beta and J beta chains in HLA-A2/Melan-A-specific T-cell clonotypes of melanoma patients. *J Transl Med* 2009; 7: 21. [PubMed: 19317896]
31. Reijonen H, Mallone R, Heninger A-K, et al. GAD65-specific CD4+ T-cells with high antigen avidity are prevalent in peripheral blood of patients with type 1 diabetes. *Diabetes* 2004; 53(8): 1987–1994. [PubMed: 15277377]
32. Diz R, Garland A, Vincent BG, et al. Autoreactive effector/memory CD4+ and CD8+ T cells infiltrating grafted and endogenous islets in diabetic NOD mice exhibit similar T cell receptor usage. *PLoS One* 2012; 7(12): e52054. [PubMed: 23251685]
33. Hall MA, Reid JL, Lanchbury JS. The distribution of human TCR junctional region lengths shifts with age in both CD4 and CD8 T cells. *Int Immunol* 1998; 10(10): 1407–1419. [PubMed: 9796907]
34. Pannetier C, Cochet M, Darche S, Casrouge A, Zöller M, Kourilsky P. The sizes of the CDR3 hypervariable regions of the murine T-cell receptor beta chains vary as a function of the recombined germ-line segments. *Proc Natl Acad Sci U S A* 1993; 90(9): 4319–4323. [PubMed: 8483950]
35. Goronzy JJ, Lee W-W, Weyand CM. Aging and T-cell diversity. *Exp Gerontol* 2007; 42(5): 400–406. [PubMed: 17218073]
36. Rudd BD, Venturi V, Li G, et al. Nonrandom attrition of the naive CD8+ T-cell pool with aging governed by T-cell receptor:pMHC interactions. *Proc Natl Acad Sci* 2011; 108: 13694–13699. [PubMed: 21813761]
37. Boyd SD, Liu Y, Wang C, Martin V, Dunn-Walters DK. Human lymphocyte repertoires in ageing. *Curr Opin Immunol* 2013; 25(4): 511–515. [PubMed: 23992996]
38. Le Saux S, Weyand CM, Goronzy JJ. Mechanisms of immunosenescence: lessons from models of accelerated immune aging. *Ann N Y Acad Sci* 2012; 1247: 69–82. [PubMed: 22224726]
39. Pavkov ME, Hanson RL, Knowler WC, Bennett PH, Krakoff J, Nelson RG. Changing patterns of type 2 diabetes incidence among Pima Indians. *Diabetes Care* 2007; 30(7): 1758–1763. [PubMed: 17468358]
40. Robins H Immunosequencing: applications of immune repertoire deep sequencing. *Curr Opin Immunol* 2013; 25(5): 646–652. [PubMed: 24140071]

Figure 1.

(A) V and (B) J gene frequencies in the exploratory cohort. Asterisk denotes $p < 0.05$. Student's t -test used to assess differences between the normal glucose regulation and type 2 diabetes groups for frequencies of TR β V and TR β J genes among productive sequences making up at least 0.005% of an individual's T-cell receptor repertoire

Figure 2.

(A) Frequency distribution of complementarity-determining region 3 (CDR3) length in the exploratory cohort and (B) mean CDR3 length in those with type 2 diabetes mellitus and without the *HLA-DRB1*02* haplotype and those with normal glucose regulation and the *HLA-DRB1*02* haplotype

Figure 3.

(A) V and (B) J gene frequencies in the validation analysis. Differences identified in the exploratory cohort were confirmed in this group with Student's *t*-tests. Asterisk denotes $p < 0.05$ in this analysis. TRBV4-2 and 4-3 were not included in the second assay report. Insets highlight gene frequencies in validation cohort that were significantly different in exploratory cohort

Figure 4.

Eleven-year cumulative incidence for development of type 2 diabetes in those above or below the median for (A) *TRBV7-8* frequency (high = above median; low = below median); (B) CDR3 length (long = above median; short = below median); and (C) combined (e.g. four groups encompassing above/below medians for each variable). $p < 0.01$ between *TRBV7-8* high/CDR3 short and all other groups, $p < 0.05$ between *TRBV7-8* low/CDR3 short and *TRBV7-8* high/CDR3 long and $p > 0.05$ between all other groups

Figure 5.

Association of age and Simpson's Diversity Index in both cohorts. (A) Age was not associated with the diversity index in either subgroup in the exploratory cohort. The correlation was $\rho = -0.21$ ($p = 0.54$) among those with normal glucose regulation and the *HLA-DRB1*02* haplotype (black squares) and $\rho = -0.50$ ($p = 0.25$) among those with type 2 diabetes and without the *HLA-DRB1*02* haplotype (white squares). Spearman's rank correlation coefficient was used to test associations. (B) Age was negatively associated with the diversity index of T-cell repertoire in the *HLA-DRB1*02* negative normal glucose regulation validation cohort. $\rho = -0.39$, $p = 0.0063$. Association remained significant after removing the outlier

Table 1.

Demographics of the exploratory cohort

	Type 2 diabetes/ <i>HLA-DRB1*02</i> negative (<i>n</i> = 7)	Normal glucose regulation/ <i>HLA-DRB1*02</i> positive (<i>n</i> = 11)	<i>p</i>
Per cent body fat (%)	31.2±4.7	30.2± 8.7	0.78
Age (years)	34.0±8.2	31.3± 8.2	0.50
Fasting glucose (mg/dL)	117.7±19.6	84.5± 10.1	<0.01
2-h glucose (mg/dL)	261.4±28.3	105.9± 14.1	<0.01
AIR (pmol/L)	12.5 (0, 30.5)	208.4 (136.9, 311.3)	<0.01
M ^a [mg (kg EMBS/min)]	1.74 (1.7, 2.3)	2.4 (1.7, 3.2)	0.10

AIR, acute insulin response; EMBS, estimated metabolic body size.

Student's *t*-test used to assess differences.^aInsulin-stimulated glucose disposal rate.

Table 2.
Demographics of the validation cohort by future development of type 2 diabetes and gender

	Progressors (<i>n</i> = 13)	Non-progressors (<i>n</i> = 28)	<i>p</i>	Male (<i>n</i> = 23)	Female (<i>n</i> = 18)	<i>p</i>
Gender (male: female)	6:7	17:11	0.50			
Future diabetes status				6:17	7:11	0.50
(progressors : non-progressors)						
Per cent body fat (%)	35.2±5.6	31.5±9.1	0.18	28.3±7.3	38.2±5.8	<0.01
Age (years)	29.0±6.2	29.2±7.9	0.91	28.9±7.6	29.5 ±7.2	0.78
Fasting glucose (mg/dL)	89.4±8.5	84.6±7.0	0.06	83.9±8.1	88.9±6.4	0.04
2-h glucose (mg/dL)	105.5±17.1	103.6±20.7	0.78	99.8±22.3	109.8±13.5	0.08
AIR (pmol/L)	202.5 (163.5, 261.3)	238.5 (164.9, 343.7)	0.79	224.4 (145.4, 317.1)	224.3 (183.2, 341.9)	0.61
M ^a [mg (kg EMBS/min)]	2.2 (2.0, 2.6)	2.5 (2.2, 3.4)	0.01	2.6 (2.1, 3.3)	2.4 (2.2, 2.7)	0.38
Follow-up time (years)	5.9±3.2	7.8±4.5	0.18	7.9±4.2	6.3±4.2	0.25

AIR, acute insulin response; EMBS, estimated metabolic body size.

Student's *t*-test used to assess differences between continuous variables and chi-squared test for categorical variables.

^aInsulin-stimulated glucose disposal rate.