Pediatric Diabetes

ISPAD International Society for Pediatric and Adolescent Diabetes

Pediatric Diabetes 2011
doi: 10.1111/j.1399-5448.2011.00774.x

Original Article

The Environmental Determinants of Diabetes in the Young (TEDDY): genetic criteria and international diabetes risk screening of 421 000 infants

Hagopian WA, Erlich H, Lernmark Å, Rewers M, Ziegler AG, Simell O, Akolkar B, Vogt Jr R, Blair A, Ilonen J, Krischer J, She J, and the TEDDY Study Group. The Environmental Determinants of Diabetes in the Young (TEDDY): genetic criteria and international diabetes risk screening of 421 000 infants.
Pediatric Diabetes 2011.

Aims: The Environmental Determinants of Diabetes in the Young (TEDDY) study seeks to identify environmental factors influencing the development of type 1 diabetes (T1D) using intensive follow-up of children at elevated genetic risk. This study requires a cost-effective yet accurate screening strategy to identify the high-risk cohort.
Methods: The TEDDY cohort was identified through newborn screening using human leukocyte antigen (HLA) class II genes based on criteria established with pre-TEDDY data. HLA typing was completed at six international centers using different genotyping methods that can achieve >98% accuracy.
Results: TEDDY developed separate inclusion criteria for the general population (GP) and first-degree relatives (FDRs) of T1D patients. The FDR eligibility includes nine haplogenotypes (*DR3/4, DR4/4, DR4/8, DR3/3, DR4/4b, DR4/1, DR4/13, DR4/9,* and *DR3/9*) for broad HLA diversity, whereas the GP eligibility includes only the first four haplogenotypes with *DRB1*0403* as an exclusion allele. TEDDY has screened 414 714 GP infants, of which 19 906 (4.8%) were eligible, whereas 1415 of the 6333 screened FDR infants (22.2%) were eligible. High-resolution confirmation testing of the eligible subjects indicated that the low-cost and low-resolution genotyping techniques employed at the screening centers yielded an accuracy of 99%. There were considerable variations in eligibility rates among the centers for GP (3.5–7.4%) and FDR (19–32%) subjects. The eligibility rates among US ethnic groups were 0.9, 1.3, 5.0, and 6.9% for Asians, Black, Caucasians, and Hispanics, respectively.
Conclusions: Different low-cost and low-resolution genotyping methods are useful for the efficient and accurate identification of a high-risk cohort for follow-up based on the TEDDY HLA inclusion criteria (ClinicalTrials.gov NCT00279318).

William A Hagopian[a], Henry Erlich[b], Åke Lernmark[c], Marian Rewers[d], Anette G Ziegler[e], Olli Simell[f], Beena Akolkar[g], Robert Vogt Jr[h], Alan Blair[b], Jorma Ilonen[i,j], Jeffrey Krischer[k], JinXiong She[l] and the TEDDY Study Group[†]

[a]Pacific Northwest Diabetes Research Institute, Seattle, WA, USA; [b]Roche Molecular Systems, Pleasanton, CA, USA; [c]Department of Clinical Sciences, Lund University, University Hospital UMAS, Malmö, Skåne, Sweden; [d]Barbara Davis Center for Childhood Diabetes, University of Colorado, Aurora, CO, USA; [e]Diabetes Research Institute, Technische Universität München, Munich, Germany; [f]Department of Pediatrics, University of Turku, Turku, Finland; [g]DDEMD, NIDDK, Bethesda, MD, USA; [h]Newborn Screening Branch Division of Laboratory Sciences, Centers for Disease Control and Prevention, Atlanta, GA, USA; [i]Department of Clinical Microbiology, University of Kuopio, Kuopio, Finland; [j]Immunogenetics Laboratory, Department of Pediatrics, University of Turku, Turku Finland; [k]Department of Informatics and Biostatistics, Pediatrics Epidemiology Center, University of South Florida, Tampa, FL, USA; and [l]Center for Biotechnology and Genomic Medicine, Medical College of Georgia, Augusta, GA, USA

[†]Members of the TEDDY Study Group are listed in the Appendix.

Key words: autoimmunity – genetic screening – population based – prediction – T1D

Corresponding author:
Dr William Hagopian,
Pacific Northwest
Diabetes Research Institute,
720 Broadway, Seattle, WA 98122,
USA.
Tel: +1.206.860.6759;
fax: +1.206.320.1448;
e-mail: wah@u.washington.edu

Submitted 21 October 2010.
Accepted for publication 2 March 2011

Type 1 diabetes (T1D) results from poorly defined interactions between susceptibility genes and environmental determinants. T1D susceptibility is primarily defined by genetic factors within the human leukocyte antigen (HLA) complex on chromosome 6. The main disease factors are the HLA-DQ molecule encoded by *DQA1* and *DQB1* genes and the HLA-DR molecule defined by *DRB1* alleles (1). In addition, recent genome-wide association studies have identified 40 other association intervals that may harbor T1D susceptibility/protection genes (2–5). In contrast to the rapid progress in finding T1D genes, identification and confirmation of environmental determinants remain a formidable challenge. The reasons underlying the lack of progress are multi-faceted. First, different categories and large numbers of environmental determinants could contribute to the triggering or protection of T1D. Although many candidates have been suggested by previous studies (6, 7), few have been definitively proven beyond reasonable doubt. Second, exposures may occur any time before the onset of disease, from *in utero* to disease onset. Third, environmental determinants may differ in different populations, partly depending on the genetic architecture. Fourth, the individual risk of developing T1D in the general population (GP) is not very high and quite variable in different populations. Therefore, large study populations with elevated T1D risk must be identified. Although first-degree relatives (FDRs) of T1D patients certainly have elevated risk, subjects from the GP must be included as well because 85–90% of diagnosed patients do not have an FDR with the disease.

Identification of environmental determinants requires frequent follow-up studies of large number of subjects from early in life until disease onset for a variety of exposures using both epidemiological and laboratory methodologies. To accomplish such ambitious goals, long-term multicenter prospective studies on a cohort at high risk of developing the disease are necessary. The Environmental Determinants of Diabetes in the Young (TEDDY), an NIH-funded prospective observational study, was designed to accomplish this goal. The TEDDY design addresses the main concerns related to the studies of environmental exposures (8). TEDDY has identified a large cohort of infants that have increased genetic risk for developing islet autoantibodies and T1D by screening several hundred thousand newborns. The high-risk cohort is closely monitored beginning at approximately 3 months of age for the development of islet autoantibodies and T1D for 15 yr, during which environmental exposures are extensively and intensively measured. These exposures include diet, infectious agents, psychosocial stress, and other lifestyle and location-based factors. Exposures are captured via frequent biological samples from the participating children as well as extensive questionnaire-acquired data (8).

For the TEDDY study to be cost-effective, the intention was to apply the long and intense follow-up protocol only to children at elevated risk of T1D. Development of a high-risk study cohort of sufficient size required multiple strategies including an international consortium of large clinical centers, screening of both FDR and GP infants, and study inclusion criteria based on genetic risk screening applicable in this diverse setting. Despite the available information on multiple T1D susceptibility genes, the only genes useful for screening purpose were, and still are, the HLA class II genes (*DRB1, DQA1,* and *DQB1*), which account for some 50% of the total genetic contribution to T1D. Therefore, these genes were chosen for TEDDY screening. Here, we describe the development of the TEDDY HLA strategy, its successful implementation in the screening centers, the overall results as the screening nears completion, and the associated quality control programs and outcomes.

Methods

Pre-TEDDY data collection

To develop an HLA screening strategy for TEDDY, HLA data on the healthy background population, T1D patients and their FDRs were assembled from the six TEDDY clinical centers based in Colorado (COL), Washington State (WAS), Georgia/Florida (GEO), Finland (FIN), Germany (GER), and Sweden

(SWE). All centers provided historical data on their background populations that were used to develop the TEDDY HLA strategy for GP subjects. Two centers (GER and SWE) also provided FDR data. Some of the pre-TEDDY data from this study have been published previously (9–14).

TEDDY genotyping methods

Samples were in all cases obtained from subjects under informed consent of parents and with IRB/Ethics Board approval. Each center was allowed to develop its own genotyping methods as long as a minimum accuracy of 98% was achieved. Five HLA screening laboratories were chosen for TEDDY screening and they employed four different genotyping strategies. Screening genotyping results were expected to be available by the time the infant was 2 months of age. Low-cost genotyping was achieved by adopting a two-stage screening strategy in four laboratories. In the first stage, approximately 90% of the ineligible subjects are excluded by the presence of specific alleles that can be detected inexpensively. In the second stage, detailed genotyping of *DQB1* and *DQA1* or *DQB1* and *DRB1* alleles is determined. For the GP, the *DRB1*0403* allele is usually determined by a restriction digest of the exon 2 amplicon. The first-stage strategy used by the Finnish and Swedish screening laboratories was to exclude certain resistant alleles while requiring certain susceptible alleles, and was previously described (15). The WAS laboratory used a first-stage strategy of exclusion of *DQB1*05, DQB1*06, DQB1*0301,* and *DQA1*02* followed by direct exon 2 sequencing of specific *DQB1* and *DQA1* alleles in the second-stage genotyping. The GEO screening laboratory excluded subjects with *DQB1*05, DQB1*06,* and *DQB1*0301* using allele-specific amplifications in the first stage. The potentially eligible subjects were further genotyped for *DQB1* by denaturing gradient gel electrophoresis using a previously published protocol (9) and *DRB1* by Luminex beads. Samples from the COL center were genotyped in the laboratory of Dr Erlich using a reverse line blot technique with a panel of immobilized oligonucleotide probes for *DRB1* and *DQB1* alleles (11). The same laboratory also served as the TEDDY HLA Reference Laboratory to carry out confirmatory tests of enrolled subjects from all six clinical centers using separate *DRB1, DQA1,* and *DQB1* reverse line blots, each with a much higher resolution panel of immobilized probes (16).

Results

Development of the TEDDY HLA strategy

To design a study-wide TEDDY HLA strategy, the TEDDY investigators assembled HLA genotyping data from all six TEDDY clinical centers. These data represent the populations near the three US centers and three European centers. All six data sets consisted primarily of Caucasian subjects from the study areas. Odds ratios (ORs) for association with T1D were calculated for each haplogenotype in each population. Genotypes were then ranked by the OR in each of the study populations. Interestingly, the rank order for the top five high-risk haplogenotypes was identical for all six data sets. From the combined data set, we were able to identify nine high-risk haplogenotypes which had an estimated relative risk of >3 in all six data sets (Table 1). Although several other haplogenotypes had increased OR in one or several data sets, their ORs were not consistent in all study populations and thus were excluded from further consideration.

During the TEDDY design stage, consensus favored the adoption of inclusion of specific HLA haplogenotypes eligible for TEDDY follow-up with specific exclusion of dominantly protective alleles. The data in Table S1 (Supporting information) summarize the cumulative frequencies of the top two, four, or nine

Table 1. Human leukocyte antigen eligibility for FDR and GP newborns*

Code	Haplotype genotypes	Abbreviation	FDR	GP†
A	*DR4-DQA1*030X-DQB1*0302/DR3-DQA1*0501-DQB1*0201*	*DR3/4*	Y	Y
B	*DR4-DQA1*030X-DQB1*0302/DR4-DQA1*030X-DQB1*0302*	*DR4/4*	Y	Y
C	*DR4-DQA1*030X-DQB1*0302/DR8-DQA1*0401-DQB1*0402*	*DR4/8*	Y	Y
D	*DR3-DQA1*0501-DQB1*0201/DR3-DQA1*0501-DQB1*0201*	*DR3/3*	Y	Y
E	*DR4-DQA1*030X-DQB1*0302/DR4- DQA1*030X-DQB1*020X*	*DR4/4b*	Y	N
F	*DR4-DQA1*030X-DQB1*0302/DR1- DQA1*0101-DQB1*0501*	*DR4/1*	Y	N
G	*DR4-DQA1*030X-DQB1*0302/DR13-DQA1*0102-DQB1*0604*	*DR4/13*	Y	N
H	*DR4-DQA1*030X-DQB1*0302/DR9- DQA1*030X-DQB1*0303*	*DR4/9*	Y	N
I	*DR3-DQA1*0501-DQB1*0201/DR9- DQA1*030X-DQB1*0303*	*DR3/9*	Y	N

FDR, first-degree relative; GP, general population; TEDDY, the Environmental Determinants of Diabetes in the Young.
*Although *DQB1*0302* is shown above, *DQB1*0304* is acceptable in its place for TEDDY inclusion. Subtyping was not required for further characterization of *DQB1*020X* and *DQA1*030X* genotypes. Y = eligible and N = Not eligible for TEDDY inclusion.
†*DR4* subtyping was required to exclude GP newborns with *DRB1*0403*, but no other *DRB1* subtyping was required.

haplogenotypes in T1D and control populations, the estimated odds ratio and absolute risk in each of the six clinical centers, and the combined data for all populations. As expected, inclusion of the top two haplogenotypes (*DR3/4* and *DR4/4*, denoted A and B, respectively) is a strategy that yields the highest specificity (96.7%) and good AR (5.5%), but only 39.3% of the future T1D cases can be identified by these two haplogenotypes. In contrast, inclusion of nine genotypes (A–I) would increase the average sensitivity to 63% while decreasing the specificity to 90% and the AR to 2.4%. By consensus, the TEDDY adopted the compromise strategy that included four high-risk haplogenotypes (A–D) for the GP infants (Table 1). The GP inclusion criteria were expected to yield a sensitivity of 50%, a specificity of 94%, an average OR of 10, and an average AR of 3.4%, assuming equal screening numbers of Caucasians in all six clinical centers. Using these inclusion criteria, 5.7% of the GP infants were expected to be eligible for follow-up studies. It should be noted that the pre-TEDDY estimates included all *DR4* subtypes in the calculation, while haplotypes with a *DRB1*0403* subtype are excluded from the actual TEDDY follow-up, which should decrease the observed eligibility rate below the 5.7% estimate.

Because FDR subjects had higher risk compared to GP subjects, it was agreed to expand the inclusion criteria to include all nine haplogenotypes in Table 1 for FDR infants. It should be noted that *DRB1*0403* was not used as an exclusion criteria for FDR subjects. An estimated 31% of the FDR population would be eligible for follow-up and an estimated 69% of future T1D cases from the FDR population would be included in the eligible population with an estimated AR of 13%.

Screening results

From September 2004 to February 2010, TEDDY screened a total of 414 714 GP newborns. Of these newborns, 19 906 were found to be eligible for follow-up, representing 4.8% of the screened GP subjects (Table 2). The overall eligibility rate was lower than the eligibility rate estimated using pre-TEDDY data. More than one third (39.5%) of the eligible GP infants were *DR3/4* (haplogenotype A), while each of the other three eligible genotypes accounted for approximately 20% of the entire cohort of eligible infants (Fig. 1). There was considerable variability in the total eligibility rate as well as the frequencies of the eligible genotypes across the six clinical centers (Table 2). Most notably, the SWE center had the highest eligibility rate (7.4%; $p < 0.0001$) compared to all other centers, which ranged from 3.5 to 5.6%. This was primarily because of the high frequency of the *DR3/4* haplogenotype at the SWE center ($p < 0.0001$ vs. the other centers). The overall eligibility rates for the FIN and COL clinical centers (5.6 and 5.5%, respectively) were also higher than the GER (4.0%), WAS (4.0%), and GEO (3.5%) clinical centers (Table 2).

TEDDY also screened 6333 FDR subjects, of which 1415 were eligible for the follow-up studies based on the nine eligible haplogenotypes (Table 2). The mean eligibility rate for all six major clinical centers and two small centers was 22.2%. The eligibility rates were quite similar in five of the six large clinical centers (19.1–23.2%), whereas the FIN center had a higher eligibility rate for FDRs (31.2%) compared to the other centers ($p < 0.0001$). As expected, the *DR3/4* genotype was the most common haplogenotype in five of the six major clinical centers; however, *DR4/1* (haplogenotype F) was the most common eligible haplogenotype in the FIN center (29.2% of the Finnish eligible genotypes). Interestingly, the greater overall eligibility rate for FDRs in the FIN center is primarily because of this greater *DR4/1* frequency among eligible Finnish FDRs, compared to the other centers ($p < 0.0001$). The *DR4/9* (haplogenotype H) is also significantly more common among eligible Finnish FDRs vs. the other centers ($p < 0.0001$). *DR4/1* and *DR4/4* (haplogenotype B) are the second and third most common haplogenotypes in the overall study population (20.1 and 15.6%, respectively). *DR3/3* (haplogenotype D) and *DR4/8* (haplogenotype C) represent 12.7 and 8.5% of the overall FDR eligible population, respectively. The other three genotypes are less common, together representing only 9.2% of the overall eligible population (Table 2 and Fig. 1).

Ethnic differences in eligibility rate

Although the newborns screened in the three European centers are primarily Caucasians, the screened newborns in the USA included all minority populations reflecting the increasingly diverse characteristics of these screening centers. Overall, the screened US cohort includes Asian-Americans (6%), Hispanics (10%), African-Americans (14%), Caucasians (58%), and other ethnic groups (13%). Although Caucasians represent 56–60% of the screened cohort in all three US TEDDY centers, each center has a different predominant minority group, Hispanics in the COL center (27%), African-Americans in the GEO center (26%), and Asian-Americans in the WAS center (10%). The entire GP cohort screened in the US centers was analyzed for the distribution of eligible genotypes according to race/ethnicity (Table 3). The *DR3/3* genotype is the most common (~50%) eligible genotype in both Asian-American and African-American groups. In contrast, the *DR3/4* genotype is common in both Hispanics and Caucasians. Surprisingly, the *DR4/4* and *DR4/8* genotypes are very common in the Hispanic group and these two genotypes are primarily

Table 2. TEDDY human leukocyte antigen screening and eligibility results for GP (top) and FDR (bottom) newborns

Clinical center	Screened (*n*)	Eligible (*n*)	Eligible (%)	Percent of screened GP				Percent of eligible GP			
				A	B	C	D	A	B	C	D
COL	75 213	4170	5.5	2.0	1.4	1.2	1.0	35.2	24.6	21.3	18.9
GEO	85 358	2995	3.5	1.5	0.6	0.4	0.9	44.2	18.1	11.0	26.7
WAS	113 056	4510	4.0	1.6	0.7	0.6	1.0	41.0	18.0	14.7	26.3
FIN	59 754	3370	5.6	1.9	1.0	1.8	0.9	33.6	17.7	32.7	16.0
GER	34 218	1353	4.0	1.7	0.7	0.4	1.1	43.8	18.8	10.6	26.9
SWE	47 115	3508	7.4	3.2	1.6	1.0	1.7	42.5	21.7	13.6	22.2
Total	414 714	19 906	4.8	1.9	1.0	0.9	1.1	39.5	20.0	18.1	22.4

Clinical center	Screened (*n*)	Eligible (*n*)	Eligible (%)	Percent of eligible FDR								
				A	B	C	D	E	F	G	H	I
COL	945	210	22.2	29.5	17.1	9.5	15.7	0.5	17.1	6.2	0.5	3.8
GEO	973	186	19.1	39.2	14.5	4.8	16.7	0.5	15.1	4.3	1.1	3.8
WAS	898	208	23.2	37.0	16.3	5.8	15.9	1.4	14.9	5.3	1.9	1.4
FIN	924	288	31.2	20.8	14.2	13.9	5.6	0.0	29.2	5.9	6.9	3.5
GER	1518	297	19.6	31.6	15.2	6.7	12.8	0.3	22.6	7.1	1.3	2.4
SWE	1019	215	21.1	33.5	20.5	7.9	11.6	0.5	16.7	7.0	1.4	0.9
NBD	31	3	9.7	33.3	33.3	0.0	33.3	0.0	0.0	0.0	0.0	0.0
CHP	25	8	32.0	50.0	0.0	12.5	25.0	0.0	12.5	0.0	0.0	0.0
Total	6333	1415	22.2	31.5	15.6	8.5	12.7	0.5	20.1	6.0	2.4	2.6

CHP, Children's Hospital of Philadelphia; FDR, first-degree relative; GP, general population; NBD, Naomi Berrie Diabetes Center; TEDDY, the Environmental Determinants of Diabetes in the Young.
These two small centers joined TEDDY and contributed a small number of FDRs to this study.

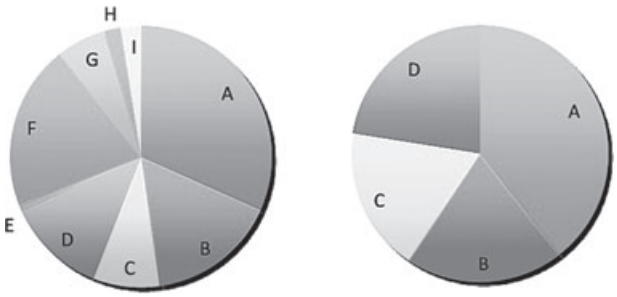

Fig. 1. Distribution of human leukocyte antigen (HLA) haplogenotypes in the eligible first-degree relatives (left) and general populations (right). HLA letter abbreviations are as follows: A, *DR3/4*; B, *DR4/4*; C, *DR4/8*; D, *DR3/3*; E, *DR4/4b*; F, *DR4/1*; G, *DR4/13*; H, *DR4/9*; and I, *DR3/9*. Full haplogenotypes are specified in Table 1.

responsible for the high eligibility rate for the Hispanic group in Colorado. Overall, the eligibility rates are significantly lower for Asian-American (0.9%) and African-American (1.3%) infants than for Hispanic (6.9%) and Caucasian (5.0%) infants (Table 3).

Because the annual incidence of T1D varies greatly among these ethnic groups, it is important to view the HLA eligibility rates in the context of the annual incidences to determine whether the eligibility is proportional to the incidence in each ethnic group. For this purpose, we used the annual incidence of T1D in each of the ethnic groups from the SEARCH study, which includes regions identical or highly similar to each of the three US TEDDY centers (17). For comparative purposes, we calculated the relative eligibility rates observed in TEDDY and the relative incidence rates based on published SEARCH data, both normalized relative to Caucasians. Finally, we determined the ratio of these two relative rates, which is denoted as the weighted eligibility rate. As shown in Table 3, relative eligibility based on the TEDDY inclusion criteria differed significantly among ethnic groups, ranging from 18% in Asian-Americans to 138% in Hispanics. The relative incidence rates also differed significantly among ethnic groups, being 22% in Asian-Americans and in the 50% range for Hispanics and African-Americans, relative to Caucasians (Table 3). Importantly, the derived weighted eligibility rates clearly show that: (i) African-Americans are underrepresented by the eligibility criteria (47%); (ii) Hispanics are overrepresented (266%); and (iii) Asian-Americans are represented nearly proportionally to their incidence (81%). A similar analysis was not made for the FDR eligibility criteria at this time because of the much smaller number of subjects.

Quality control programs for HLA genotyping

TEDDY developed two quality control programs to ensure the quality and accuracy of the HLA screening data. The first component is an annual HLA proficiency test administered by the Newborn Screening Branch of the National Center for Environmental Health at the Centers for Disease Control (CDC). For each test, a set of 50 coded blood samples (40 designated as GP and 10 as FDRs) are genotyped by the participating laboratories using their genotyping methods. The typing results (eligibility status and eligible genotype code) are returned to the CDC within 30 d of receiving the test samples. A minimum accuracy of 98% is deemed acceptable. Failure to meet this requirement calls for an immediate repeat test with a different set of samples. The genotyping laboratory is suspended if it fails both consecutive tests. Four separate tests were carried out in 2004, 2005, 2006, and 2008, respectively. Each laboratory passed all tests with 100% accuracy

Table 3. Screening results in different ethnic groups in the three US centers

Race/ethnicity	Asian-Americans	Hispanic Americans	African-Americans	Caucasian	Other*
Screened (*n*)	3715	6611	9231	38 240	8409
Percentage of screened: COL	2.0	27.1	6.2	55.8	9.0
Percentage of screened: GEO	4.7	0.8	26.0	60.1	8.4
Percentage of screened: WAS	10.0	7.1	4.9	56.4	21.6
Percentage of screened: average	6	10	14	58	13
Eligible (%)					
Genotype A (*DR3/4*)	0.2	1.9	0.5	2.3†	0.9
Genotype B (*DR/4/4*)	0.1	1.9‡	0.2	0.9	0.5
Genotype C (*DR4/8*)	0.1	2.2‡	0.1	0.5	0.6
Genotype D (*DR3/3*)	0.5‡	0.8	0.6‡	1.3	0.6
All genotypes	0.9‡	6.9	1.3‡	5.0	2.5
Incidence of T1D§	6	14	15	27	NA
Relative incidence (vs. Caucasians) (%)	22	52	56	100	NA
Relative eligibility (vs. Caucasians) (%)	18	138	26	100	NA
Weighted eligibility rate (%)	81	266	47	100	NA

*Native Americans, Pacific Islanders, multiracial, or unknown.
†Uncorrected $p < 0.01$ (chi-square test) vs. total for all ethnicities.
‡Uncorrected $p < 0.001$ (chi-square test) vs. total for all ethnicities.
§Incidence per 100 000/yr.

with the exception that one laboratory scored 98% once and one laboratory scored 96% once. On the basis of the established TEDDY quality control procedures, the latter necessitated a repeat test, which was passed with 100% accuracy.

The second quality control program consists of confirmatory repeat genotyping of all eligible subjects by the central HLA reference laboratory. The confirmatory genotyping serves three primary purposes: (i) identify genotyping errors or inaccuracies that occurred in the screening laboratories; (ii) identify potential sample mislabeling that occurred anywhere from hospitals to clinical centers to genotyping laboratories; and (iii) perform high-resolution genotyping for three HLA class II loci, *DRB1, DQA1,* and *DQB1*. To achieve these goals, a blood sample is collected at the 9-month or 12-month follow-up visit for each enrolled infant. Genotyping results on the new sample from the HLA reference laboratory are considered the gold standard, and are compared with the initial screening results from the clinical centers. A minimum agreement of 98% must be achieved by each clinical center. This requirement is more stringent than the proficiency test because all errors including genotyping mistakes, sample contaminations, errors in inferred haplotypes, and sample labeling mistakes can contribute to the overall discordance rate. Despite the multiple sources of potential errors, the screening laboratories using low-cost and low-resolution genotyping methods yielded remarkably accurate data, as shown by the 98–100% accuracies in all screening laboratories and the 99% accuracy for the overall cohort (Table S2). The confirmatory test ensures that genotyping results are 100% correct for all infants who continue enrollment in the long-term follow-up phase of the TEDDY study.

Discussion

HLA class II genes are the most important susceptibility genes for T1D, accounting for approximately 50% of the genetic contribution to the disease. The HLA criteria used to select a high-risk population are not trivial, because of the extremely high degree of polymorphism in these genes, ethnic variability (18, 19), and the hierarchical nature of the risks conferred by the large number of distinct haplogenotypes. These selection criteria are also compromises between considerations of sensitivity, specificity, and typing costs (11, 20–25). Investigators may exclude specific alleles or haplotypes, include specific haplotypes, or use a combination of inclusion and exclusion criteria. For example, the US DAISY study required the susceptibility haplotypes *DRB1*03* and/or *DRB1*04-DQB1*0302* for inclusion, whereas *DRB1*15/16* was used as an exclusion criterion (25). The Finnish DIPP study used *DQB1*0302/X*, where *X* was either *DQB1*02* or any allele other than *DQB1*0602 or DQB1*0301* (13). The Trial to Reduce IDDM in the Genetically at Risk study of FDRs required susceptibility haplotypes *DQB1*0302, DQA1*05-DQB1*02,* and/or *DQA1*03-DQB1*02*, excluded all subjects with *DQB1*0602* or *DQB1*0301*, and conditionally excluded *DQB1*0603* or *DQA1*0201-DQB1*02* depending on the susceptibility haplotype (26).

Others, like TEDDY, have used strategies with detailed inclusion haplogenotypes. For example, the Belgian Diabetes Registry defined a list of four susceptibility genotypes including *DQA1*0301-DQB1*0302/DQA1*0501-DQB1*0201, DQA1*0301-DQB1*0302/DQA1*0301-DQB1*0302, DQA1*0501-DQB1*0201/DQA1*0501-DQB1*0201,* and *DQA1*0301-DQB1*0302/X*, where *X* is any one of the 11 generally disease-neutral *DQA1-DQB1* haplotypes (27). The TEDDY strategy includes the first three of these susceptibility genotypes, but limits the fourth for GP subjects to $X = DQA1$*0401-DQB1*0402*. The latter choice increased the overall risk level of the TEDDY cohort by limiting the size of the moderate risk portion of the included subjects. For GP screening, TEDDY also excluded *DRB1*0403* from eligible *DR4* haplotypes because these haplotypes are generally disease resistant (28). Neither limitation was necessary for FDRs because of their higher absolute disease risk.

For TEDDY, the HLA screening strategy had to meet many requirements: (i) identify an eligible cohort with high risk for developing islet autoantibodies and T1D, the primary and secondary end-points of the TEDDY study; (ii) minimize the number of subjects requiring screening to accumulate the cohort; (iii) select a relatively genetically homogenous cohort to achieve sufficient power to identify environmental determinants; (iv) include a sufficiently diverse set of HLA genotypes to determine whether there are different environmental determinants for different HLA genotypes; (v) employ laboratory methods that are both accurate and highly cost-effective; (vi) allow efficient risk stratification for both GP and FDR populations; (vii) be applicable to an international multi-site study that studies multiple ethnic groups; and (viii) provide screening results quickly enough to recruit subjects to follow-up by the deadline of 4.5 months of age per the TEDDY protocol (8). The TEDDY HLA strategy was a successful compromise to fulfill all these requirements under many constraints.

The TEDDY screening laboratories utilized a variety of genotyping strategies to accomplish the goal. These methods were usually simple, low-cost, and could efficiently handle tens of thousands of samples a year. These strategies worked exceptionally well as shown by the near-perfect score on proficiency tests, as well as the 99% confirmation rate for retyping of eligible subjects. This level of performance over 421 000

screened subjects is remarkable given the demand for low genotyping cost, the large numbers of samples to be collected, and the rapid turnaround time required. In fact, the median time to completion of screening typing was 41 d of age, and 95% of infants had complete TEDDY genotyping by 74 d of age.

Of the 414 714 GP newborns screened by TEDDY, 4.8% were genetically eligible for the follow-up study. This observed eligibility rate is significantly less than the 5.7% eligibility rate estimated using the pre-TEDDY data from all six clinical centers. The lesser overall eligibility rate is primarily because of overestimates in the COL center (8.5% estimated vs. 5.6% observed) and the WAS center (6.0% estimated vs. 4.0% observed). For the four other major TEDDY centers (GER, SWE, FIN, and GEO centers), observed rates were similar to estimated rates. The lower observed eligibility rate in Washington was partly explained by the 44% non-Caucasian infants in their screened cohort, which is much greater than that in pre-TEDDY sample because of increasing ethnic diversity in the region. The overestimated eligibility rate in the Colorado population may be because of a combination of factors such as inclusion of *DRB1*0403* subjects in the pre-TEDDY data and other differences in genotyping methodologies between pre-TEDDY and TEDDY. In fact, for all centers, the eligibility estimates using pre-TEDDY data did not exclude *DRB1*0403*, which is excluded for the TEDDY GP cohort.

The TEDDY strategy may not appear to be easy to implement for genotyping purpose because it includes very specific haplogenotypes and excludes the protective *DRB1*0403* allele for GP infants. However, the TEDDY strategy actually does promote economic and accurate genotyping because the four GP genotypes consist of only three haplotypes: *DR4 (DRB1*04-DQA1*0301-DQB1*0302)*, *DR3 (DRB1*03-DQA1*0501-DQB1*0201)*, and *DR8 (DRB1*08-DQA1*0401-DQB1*0402)*.

The TEDDY data on different ethnic groups in the USA provided valuable information for future population screening for T1D. For various reasons discussed earlier, TEDDY elected to adopt a uniform HLA strategy for all ethnic groups. It was not surprising that different ethnic groups had highly different eligibility rates. We indeed observed very low eligibility rates for two populations (0.9 and 1.3% for Asian-American and African-American, respectively) and high eligibility rate for the Hispanic group (6.9%). As these ethnic groups also have lower T1D incidence, the lower eligibility rates may be appropriate if the eligibility rates are proportional to the annual disease incidence in the corresponding populations. This is indeed true for the Asian-American population. However, African-Americans are underrepresented even after correction for the disease incidence (Table 3). As for Africans in general, many African-American T1D patients have a greater diversity of HLA haplotypes. Additional T1D risk haplogenotypes would therefore be required to increase the sensitivity of screening for this group. On the other hand, the Hispanic group is overrepresented by the TEDDY inclusion criteria (Table 3). These results suggest that ethnic-specific criteria, while more difficult to implement, should be considered for population-wide screening to maximize sensitivity and specificity. Nevertheless, efficient and accurate TEDDY HLA screening of more than 421 000 infants from multiple international sites, diverse ethnic groups, and different risk strata (FDR vs. GP) was successfully completed. This experience supports the notion that population-wide genetic screening for T1D risk may ultimately be a practical goal for public health infrastructures as a part of population-wide T1D prediction and prevention in the future.

Acknowledgements

This study was funded by DK 63829, 63861, 63821, 63865, 63863, 63836, and 63790 and contract no. HHSN267200700014C from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Institute of Child Health and Human Development (NICHD), National Institute of Environmental Health Sciences (NIEHS), Juvenile Diabetes Research Foundation (JDRF), and Centers for Disease Control and Prevention (CDC). We wish to thank all TEDDY children and their families for their participation in this study and thank all office staff and laboratory personnel who made significant contribution to this study. Dr Ingrid Kockum provided unpublished FDR genotype data from the Swedish Childhood Diabetes Registry Study.

Appendix

The Teddy Study Group

Committees

[1]Ancillary Studies; [2]Diet; [3]Genetics; [4]Human Subjects/Publicity/Publications; [5]Immune Markers; [6]Infectious Agents; [7]Laboratory Implementation; [8]Maternal Studies; [9]Psychosocial; [10]Quality Assurance; [11]Steering; [12]Study Coordinators; [13]Celiac Disease; [14]Clinical Implementation; and [15]Quality Assurance Subcommittee on Data Quality

Colorado Clinical Center

Marian Rewers, MD, PhD, PI[1,4,6,10,11]; Katherine Barriga[12]; Judith Baxter[9,12,15]; George Eisenbarth, MD, PhD; Nicole Frank[2]; Patricia Gesualdo[2,12,14,15]; Michelle Hoffman[12,13,14]; Lisa Ide; Jill Norris, PhD[2,12]; Jessie Robinson[12]; and Kathleen Waugh[7,12,15]

University of Colorado at Denver and Health Sciences Center

Barbara Davis Center for Childhood Diabetes
Aurora, CO, USA

Georgia/Florida Clinical Center

Jin-Xiong She, PhD, PI[1,3,4,11]; Desmond Schatz, MD[4,5,7,8]; Diane Hopkins[12]; Leigh Steed[6,12,13,14,15]; Angela Choate[12]; Katherine Silvis[2]; Meena Shankar[2]; Yi-Hua Huang, PhD; Ping Yang; Hong-Jie Wang; Jessica Leggett; Kim English; Richard McIndoe, PhD; Angela Dequesada[12]; Michael Haller, MD[14]; and Stephen W Anderson, MD

Medical College of Georgia
Augusta, GA, USA
University of Florida
Gainesville, FL, USA
Pediatric Endocrine Associates
Atlanta, GA, USA

Germany Clinical Center

Anette G Ziegler, MD, PI[1,3,4,11]; Heike Boerschmann[14]; Ezio Bonifacio, PhD[5]; Melanie Bunk; Johannes Försch; Lydia Henneberger[2,12]; Michael Hummel, MD[13]; Sandra Hummel, PhD[2]; Gesa Joslowski[2]; Mathilde Kersting, PhD[2]; Annette Knopff[7]; Nadja Kocher; Sibylle Koletzko, MD[13]; Stephanie Krause; Claudia Lauber; Ulrike Mollenhauer; Claudia Peplow; Maren Pflüger[6]; Daniela Pöhlmann; Claudia Ramminger; Sargol Rash-Sur; Roswith Roth, PhD[9]; Julia Schenkel; Leonore Thümer; Katja Voit; Christiane Winkler, PhD[2,12,15]; and Marina Zwilling

Diabetes Research Institute
Center for Regenerative Therapies
TU Dresden, Germany
Institute of Psychology
University of Graz
Graz, Austria
Department of Gastroenterology
Dr von Hauner Children's Hospital
Ludwig Maximillians University Munich
Munich, Germany
Research Institute for Child Nutrition
Dortmund, Germany

Finland Clinical Center

Olli G Simell, MD, PhD, PI[1,4,11,13]; Kirsti Nanto-Salonen, MD, PhD[12]; Jorma Ilonen, MD, PhD[3]; Mikael Knip, MD, PhD; Riitta Veijola, MD, PhD; Tuula Simell, PhD[9,12]; Heikki Hyöty, MD, PhD[6]; Suvi M Virtanen, MD, PhD[2]; Carina Kronberg-Kippilä[2]; Maija Torma[12,14]; Barbara Simell[12,15]; Eeva Ruohonen; Minna Romo; Elina Mantymaki; Heidi Schroderus; Mia Nyblom; and Aino Stenius

University of Turku
Turku, Finland
University of Tampere
Tampere, Finland
University of Oulu
Oulu, Finland
Turku University Hospital
Turku, Finland
Tampere University Hospital
Tampere, Finland
Oulu University Hospital
Oulu, Finland
National Public Health Institute
Helsinki, Finland
University of Kuopio
Kuopio, Finland

Sweden Clinical Center

Åke Lernmark, PhD, PI[1,3,4,8,10,11,15]; Daniel Agardh, MD, PhD[13]; Peter Almgren; Eva Andersson; Carin Andrén-Aronsson[2,13]; Maria Ask; Ulla-Marie Karlsson; Corrado Cilio, MD, PhD[5]; Jenny Bremer; Emilie Ericson-Hallström; Thomas Gard; Joanna Gerardsson; Ulrika Gustavsson; Gertie Hansson[12,14]; Monica Hansen; Susanne Hyberg; Rasmus Håkansson; Sten Ivarsson, MD, PhD[6]; Fredrik Johansen; Helena Larsson, MD, PhD[14]; Barbro Lernmark, PhD[9,12]; Maria Markan; Theodosia Massadakis; Jessica Melin; Maria Månsson-Martinez; Anita Nilsson; Emma Nilsson; Kobra Rahmati; Sara Rang; Monica Sedig Järvirova; Sara Sibthorpe; Birgitta Sjöberg; Carina Törn, PhD[3,15]; Anne Wallin; and Åsa Wimar

Lund University
Lund, Sweden

Washington Clinical Center

William A Hagopian, MD, PhD, PI[1,3,4,5,6,7,11,13,14]; Xiang Yan, MD; Michael Killian[6,7,12,13]; Claire Cowen Crouch[12,14,15]; Kristen M Hay[2]; Stephen Ayres; Carissa Adams; Brandi Bratrude; Greer Fowler; Czarina Franco; Carla Hammar; Diana Heaney; Patrick Marcus; Arlene Meyer; Denise Mulenga; Elizabeth Scott; Jennifer Skidmore; Erin Small; Joshua Stabbert; and Viktoria Stepitova

Pacific Northwest Diabetes Research Institute
Seattle, WA, USA

Pittsburgh Satellite Center

Dorothy Becker, MD; Margaret Franciscus[12]; MaryEllen Dalmagro-Elias[2]; and Ashi Daftary, MD
Children's Hospital of Pittsburgh of UPMC
Pittsburgh, PA, USA

Data Coordinating Center

Jeffrey P Krischer, PhD, PI[1,4,5,10,11]; Michael Abbondondolo; Lori Ballard[3,9,14,15]; Rasheedah Brown[12,15]; David Cuthbertson; Christopher Eberhard; Veena Gowda; Hye-Seung Lee, PhD[3,6,13,15]; Shu Liu; Kristian Lynch, PhD[9]; Jamie Malloy; Cristina McCarthy[12,15]; Wendy McLeod[2,5,6,13,15]; Laura Smith, PhD[9]; Stephen Smith; Susan Smith[12,15]; Ulla Uusitalo, PhD[2,15]; Kendra Vehik, PhD[4,5,9,14,15]; and Jimin Yang, PhD[2,15]

University of South Florida
Tampa, FL, USA

Project Officer

Beena Akolkar, PhD

National Institutes of Diabetes and Digestive and Kidney Diseases
Bethesda, MD, USA

Other Contributors

Thomas Briese, PhD; Henry Erlich, PhD; Suzanne Bennett Johnson, PhD; and Steve Oberste, PhD

Columbia University
New York, NY, USA
Children's Hospital Oakland Research Institute
Oakland, CA, USA
Florida State University
Tallahassee, FL, USA
Centers for Disease Control and Prevention
Atlanta, GA, USA

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Comparison of three different human leukocyte antigen inclusion strategies for general population and first-degree relative newborns.

Table S2. Confirmation of human leukocyte antigen typing results by central laboratory.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

References

1. She JX. Susceptibility to type I diabetes: HLA-DQ and DR revisited. Immunol Today 1996: 17: 323–329.
2. Smyth DJ, Cooper JD, Bailey R et al. A genome-wide association study of nonsynonymous SNPs identifies a type 1 diabetes locus in the interferon-induced helicase (IFIH1) region. Nat Genet 2006: 38: 617–619.
3. WTCCC. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007: 447: 661–678.
4. Barrett JC, Clayton DG, Concannon P et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet 2009: 41: 703–707.
5. Hakonarson H, Grant SFA, Bradfield JP et al. A genome-wide association study identifies KIAA0 350 as a type 1 diabetes gene. Nature 2007: 448: 591–594.
6. Knip M, Veijola R, Virtanen SM et al. Environmental triggers and determinants of type 1 diabetes. Diabetes 2005: 54: S125–S136.
7. Peng H, Hagopian W. Environmental factors in the development of type 1 diabetes. Rev Endocr Metab Disord 2006: 7: 149–162.
8. TEDDY Study Group. The Environmental Determinants of Diabetes in the Young (TEDDY) study: study design. Pediatr Diabetes 2007: 8: 286–298.
9. She JX, Bui MM, Tian XH et al. Additive susceptibility to insulin-dependent diabetes conferred by HLA-DQB1 and insulin genes. Autoimmunity 1994: 18: 195–203.
10. Bonifacio E, Hummel M, Walter M et al. IDDM1 and multiple family history of type 1 diabetes combine to identify neonates at high risk for type 1 diabetes. Diabetes Care 2005: 27: 2695–2700.
11. Emery LM, Babu S, Bugawan TL et al. Newborn HLA-DR, DQ genotype screening: age- and ethnicity-specific type 1 diabetes risk estimates. Pediatr Diabetes 2005: 6: 136–144.
12. Kockum I, Sanjeevi CB, Eastman S et al. Complex interaction between HLA DR and DQ in conferring risk for childhood type 1 diabetes. Eur J Immunogenet 1999: 26: 361–372.
13. Hermann R, Turpeinen H, Laine AP et al. HLA DR-DQ-encoded genetic determinants of childhood-onset type 1 diabetes in Finland: an analysis of 622 nuclear families. Tissue Antigens 2003: 62: 162–169.
14. Wion E, Brantley M, Stevens J et al. Population-wide infant screening for HLA-based type 1 diabetes risk via dried bloodspots from the public health infrastructure. Annals NY Acad Sci 2003: 1005: 400–404.
15. Kiviniemi M, Hermann R, Nurmi J et al. A high-throughput population screening system for the estimation of genetic risk for type 1 diabetes: an application for the TEDDY (the Environmental Determinants of Diabetes in the Young) study. Diabetes Technol Ther 2007: 9: 460–472.
16. Erlich H, Valdes AM, Noble J et al. HLA DR-DQ haplotypes and genotypes and type 1 diabetes risk: analysis of the type 1 diabetes genetics consortium families. Diabetes 2008: 57: 1084–1092.
17. Liese AD, D'Agostino RB Jr, Hamman RF et al. The burden of diabetes mellitus among US youth: prevalence estimates from the SEARCH for Diabetes in Youth Study. Pediatrics 2006: 118: 1510–1518.
18. Agrawal S, Khan F, Bharadwaj U. Human genetic variation studies and HLA class II loci. Int J Immunogenet 2007: 34: 247–252.
19. Thomson G, Valdes AM, Noble JA et al. Relative predispositional effects of HLA class II DRB1-DQB1

haplotypes and genotypes on type 1 diabetes: a meta-analysis. Tissue Antigens 2007: 70: 110–127.
20. Carmichael SK, Johnson SB, Baughcum A et al. Prospective assessment in newborns of diabetes autoimmunity (PANDA): maternal understanding of infant diabetes risk. Genet Med 2003: 5: 77–83.
21. Nejentsev S, Sjoroos M, Soukka T et al. Population-based genetic screening for the estimation of type 1 diabetes mellitus risk in Finland: selective genotyping of markers in the HLA-DQB1, HLA-DQA1 and HLA-DRB1 loci. Diabet Med 1999: 16: 985–992.
22. Kupila A, Muona P, Simell T et al. Feasibility of genetic and immunological prediction of type I diabetes in a population-based birth cohort. Diabetologia 2001: 44: 290–297.
23. Ilonen J, Sjoroos M, Knip M et al. Estimation of genetic risk for type 1 diabetes. Am J Med Genet 2002: 115: 30–36.
24. Ilonen J, Reijonen H, Knip M et al. Population-based genetic screening for IDDM susceptibility as a source of HLA-genotyped control subjects. Diabetologia 1996: 39: 123.
25. Rewers M, Bugawan TL, Norris JM et al. Newborn screening for HLA markers associated with IDDM: Diabetes Autoimmunity Study in the Young (DAISY). Diabetologia 1996: 39: 807–812.
26. Akerblom HK, Knip M, Becker D et al. The TRIGR trial: testing the potential link between weaning diet and type 1 diabetes. Immunol Endocr Metab Agents Med Chem 2007: 7: 251–263.
27. Van der Auwera BJ, Schuit FC, Weets I et al. Relative and absolute HLA-DQA1-DQB1 linked risk for developing type I diabetes before 40 years of age in the Belgian population: implications for future prevention studies. Hum Immunol 2002: 63: 40–50.
28. Van der Auwera B, Van Waeyenberge C, Schuit F et al. DRB1*0403 protects against IDDM in Caucasians with the high-risk heterozygous *DQA1*0301-DQB1*0302/DQA1*0501-DQB1*0201* genotype. Diabetes 1995: 44: 527–530.