

# Glycotoxin and Autoantibodies Are Additive Environmentally Determined Predictors of Type 1 Diabetes

## A Twin and Population Study

Huriya Beyan,<sup>1</sup> Harriette Riese,<sup>2,3</sup> Mohammed I. Hawa,<sup>1</sup> Guisi Beretta,<sup>1</sup> Howard W. Davidson,<sup>4</sup> John C. Hutton,<sup>4</sup> Huibert Burger,<sup>3,5</sup> Michael Schlosser,<sup>6,7</sup> Harold Snieder,<sup>2</sup> Bernhard O. Boehm,<sup>7</sup> and R. David Leslie<sup>1</sup>

In type 1 diabetes, diabetes-associated autoantibodies, including islet cell antibodies (ICAs), reflect adaptive immunity, while increased serum *N*<sup>ε</sup>-carboxymethyl-lysine (CML), an advanced glycation end product, is associated with proinflammation. We assessed whether serum CML and autoantibodies predicted type 1 diabetes and to what extent they were determined by genetic or environmental factors. Of 7,287 unselected schoolchildren screened, 115 were ICA<sup>+</sup> and were tested for baseline CML and diabetes autoantibodies and followed (for median 7 years), whereas a random selection (*n* = 2,102) had CML tested. CML and diabetes autoantibodies were determined in a classic twin study of twin pairs discordant for type 1 diabetes (32 monozygotic, 32 dizygotic pairs). CML was determined by enzyme-linked immunosorbent assay, autoantibodies were determined by radioimmunoprecipitation, ICA was determined by indirect immunofluorescence, and HLA class II genotyping was determined by sequence-specific oligonucleotides. CML was increased in ICA<sup>+</sup> and prediabetic schoolchildren and in diabetic and nondiabetic twins (all *P* < 0.001). Elevated levels of CML in ICA<sup>+</sup> children were a persistent, independent predictor of diabetes progression, in addition to autoantibodies and HLA risk. In twins model fitting, familial environment explained 75% of CML variance, and nonshared environment explained all autoantibody variance. Serum CML, a glycotoxin, emerged as an environmentally determined diabetes risk factor, in addition to autoimmunity and HLA genetic risk, and a potential therapeutic target. *Diabetes* 61:1192–1198, 2012

From the <sup>1</sup>Centre for Diabetes and Metabolic Medicine, Blizard Institute, Queen Mary, University of London, London, U.K.; the <sup>2</sup>Unit of Genetic Epidemiology & Bioinformatics, Department of Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands; the <sup>3</sup>Interdisciplinary Center for Psychiatric Epidemiology, Department of Psychiatry, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands; the <sup>4</sup>Barbara Davis Center, University of Colorado Denver, Aurora, Colorado; the <sup>5</sup>Department of Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands; the <sup>6</sup>Institute of Pathophysiology, Ernst Moritz Arndt University of Greifswald, Greifswald, Germany; and the <sup>7</sup>Division of Endocrinology and Diabetes, University Medical Center Ulm, Ulm, Germany.

Corresponding author: R. David Leslie, r.d.g.leslie@qmul.ac.uk.

Received 12 July 2011 and accepted 20 January 2012.

DOI: 10.2337/db11-0971

H.Be., H.R., M.I.H., B.O.B., and R.D.L. contributed equally to this study.

© 2012 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. See <http://creativecommons.org/licenses/by-nc-nd/3.0/> for details.

**A**utoimmune diseases affect ~10% of the population and result from the interaction of genetic and nongenetic (probably environmental) factors (1). Type 1 diabetes is a prototypic autoimmune disease due to immune-mediated destruction of insulin-secreting islet cells involving cells of both the innate and adaptive immune systems (2,3). However, not all genetically susceptible individuals develop type 1 diabetes; consequently, even identical twins often remain discordant for the disease (2–5). A limited number of predictive biomarkers are associated with risk of diabetes, including genetic features (i.e., HLA polymorphisms) and endophenotypes (i.e., reduced insulin secretory capacity and diabetes-associated autoantibodies) (2,3). The latter include islet cell antibodies (ICAs), serum autoantibodies to GAD antibody (GADA), insulinoma-associated protein 2 antigen (IA-2A), and zinc transporter 8 (ZnT8A) (6–10). In addition, there are innate immune changes involving altered monocyte/macrophage and proinflammatory responses (11,12), with the latter including increased advanced glycation end products (AGEs) (13). In normal physiology, AGEs, such as *N*<sup>ε</sup>-carboxymethyl-lysine (CML), are generated endogenously and can be strongly inherited, (14) but tissue and serum AGE levels also relate to exogenous intestinal sources, especially heat-treated dietary factors (13).

Since both increased serum levels of CML and serum autoantibodies are associated with type 1 diabetes, we sought to discover whether each is genetically or environmentally determined and antedates the disease. We, therefore, performed a population-based cohort study to ascertain whether this marker panel predicted diabetes. Having established that CML and diabetes-associated autoantibodies were predictive of type 1 diabetes, we then performed a twin study on a cohort of monozygotic (MZ) and dizygotic (DZ) pairs initially discordant for type 1 diabetes to investigate and control for potential confounding effects on these predictors of the disease itself (cotwin case-control design) and to determine the impact of genetic and environmental factors on them (classic twin design).

### RESEARCH DESIGN AND METHODS

#### Participants

**Population study.** We screened 7,287 unselected school children of European origin between 1989 and 2008 in Germany. Of these, 115 participants (51.3% women) who were ICA<sup>+</sup> at ascertainment were followed for diabetes

TABLE 1  
Characteristics of subjects in the population study

Variable	T1D (n = 33)	No T1D (n = 82)	P value
Age at inclusion (years), mean (SD) and [range]	13.48 (4.07) [6–19]	14.24 (4.01) [6–19]	NS
Follow-up (months)	77 [58–115]	94 [67–130]	NS
CML <sub>inclusion</sub> (ng/mL)	669 [524–761]	512 [424–619]	<0.001
CML <sub>last visit</sub> (ng/mL) <sup>a</sup>	773 [611–831]	424 [381.5–452]	<0.001
Male	18 (54.5)	38 (46.3)	NS
Antibody positivity			
GADA	29 (87.9)	57 (69.5)	0.04
IA-2A	17 (51.5)	14 (17.1)	<0.001
ZnT8A <sup>b</sup>	21 (80.8)	14 (28.6)	<0.001
HLA risk status			
Low <sup>c</sup>	3 (11.1)	20 (40.8)	
High <sup>c</sup>	24 (88.9)	29 (59.2)	0.007

Of 7,287 normal subjects, 115 were ICA<sup>+</sup> and 33 developed type 1 diabetes (T1D). Data are median [interquartile range] or *n* (%) unless otherwise indicated. <sup>a</sup>*n* = 25 T1D, *n* = 48 no T1D. <sup>b</sup>*n* = 26 T1D, *n* = 49 no T1D. <sup>c</sup>*n* = 27 T1D, *n* = 49 no T1D: high HLA risk status means that subjects had any of the DRB1\*03, DRB1\*04, and/or DQB1\*302 alleles.

development (15–17). Baseline serum CML was tested in these 115 ICA<sup>+</sup> subjects and a random selection (*n* = 2,102) of the remainder of the sample. Moreover, in the 115 ICA<sup>+</sup> subjects, GADA and IA-2A were tested, and depending on serum availability, ZnT8A was tested (*n* = 75) (Table 1) (16). Of 115 ICA<sup>+</sup> subjects, 73 had >1 CML measurement prediabetes, the final sample (CML<sub>last visit</sub>) being 5–10 months prediagnosis. All subjects gave informed consent, and the study was approved by the ethical committee of the University Medical Center Ulm.

**Twin study.** A cohort of MZ and DZ twins were tested for serum CML and diabetes-associated autoantibodies (GADA, IA-2A, and ZnT8A) to establish if they were genetically determined. Twin pairs were selected from the British Diabetic Twin Study and ascertained by referral through their physicians from 1971 to present (4,5). From our collection of 546 twin pairs, we identified all 32 initially disease-discordant DZ twin pairs and ascertained 32 MZ pairs discordant for type 1 diabetes of similar age at diagnosis and disease duration at sampling (Table 2). These subjects fulfilled the following criteria: 1) European origin, 2) twin pairs initially disease discordant, 3) both twins available for

study, 4) neither twin receiving drugs other than human insulin, 5) all had normal plasma creatinine, 6) diabetes initially excluded in the cotwin by oral glucose tolerance test and random whole blood glucose <7.0 mmol/L. Zygosity was determined as described previously (4), and type 1 diabetes was defined by standard criteria (15). As controls for serum CML, we tested 168 nondiabetic female twins (39 MZ, 45 DZ pairs, mean age 51.3 [SD = 14.1], range 21–73) (14). All subjects gave informed consent, and the East London Health Authority Research Ethics Committee approved the study (Ref 07/Q0604/10).

**CML assay.** Serum samples were tested for the protein- and lipid-derived stable glycoxidative product, CML. CML is a dominant circulating AGE, the best characterized of all the AGEs. CML was measured using a competitive ELISA (AGE-CML ELISA; MicroCoat, Penzberg, Germany), which is based on a CML-specific monoclonal antibody and a glycosylated and biotinylated BSA coated to a streptavidin microplate (18). This assay has been validated, is specific, and shows no cross-reactivity with other compounds. CML levels are reported as monomeric epitopes in nanogram per milliliter serum, irrespective of the

TABLE 2  
Characteristics of 32 MZ and 32 DZ twin pairs in the twin study

Twin study	Diabetic twin	Nondiabetic cotwin
MZ ( <i>n</i> )	32	32
Male, <i>n</i> (%)	20 (62.5)	20 (62.5)
Age at sampling (years)	28.09 (16.41)	28.09 (16.41)
Duration of follow-up (years)	26.9 (11)	26.9 (11)
Age at diagnosis (years)	17.06 (12.32)	NA
Diabetes duration (years)	11.03 (9.2)	NA
Antibody positivity		
GADA	12	2
IA-2A	7	2
ZnT8A	2	2
Serum CML (ng/mL), median [interquartile range]	2,638.5 [1,911.5–3,550.5]	2,248.5 [1,708.5–2,991.0]
DZ ( <i>n</i> )	32	32
Male, <i>n</i> (%)	13 (40.6)	12 (37.5)
Age at sampling (years)	28.00 (18.16)	28.00 (18.16)
Duration of follow-up (years)	19.9 (8.3)	19.8 (8.3)
Age at diagnosis (years)	17.53 (14.45)	NA
Diabetes duration (years)	10.47 (8.65)	NA
Antibody positivity		
GADA	8	2
IA-2A	7	0
ZnT8A	1	2
CML (ng/mL), median [interquartile range]	1,776.5 [1,185.5–2,077.5]	1,616.5 [1,252.0–1,854.0]

Features of MZ and DZ twin pairs discordant for type 1 diabetes including disease duration and results at sampling. Data are means (SD) unless otherwise indicated. NA, not applicable.

protein to which the CML is attached. Assay sensitivity was 5 ng CML/mL with intra- and interassay variability <4 and <5%, respectively (18).

**Diabetes-associated autoantibody levels or assay**

**Population study.** All subjects were tested for ICA in a single laboratory (Ulm) by indirect immunofluorescence on unfixed sections of human pancreas, with a detection limit of 5 JDFU and >20 JDFU considered positive (17)

ICA<sup>+</sup> subjects were tested for GADA and IA-2A, and 65% of these subjects were tested for ZnT8A (London, U.K.), using radioimmunoprecipitation assays (8,10). Ulm assay characteristics were GADA 86% sensitivity and 95% specificity and IA-2A 73% sensitivity and 99% specificity (16). Characteristics of the ZnT8A assay are described above.

**Twin study.** All twins were tested for serum autoantibodies to GADA, IA-2A, and ZnT8A using established radioimmunoprecipitation assays (8,10). All twin samples were tested at a single laboratory (London) in batched assays with values expressed as categorical (positive/negative) and continuous traits. Positive results were duplicated, reducing false positives to <0.2%. In the latest Diabetes Antibody Standardization Program (DASP 2008), London assay characteristics were as follows: GADA sensitivity 90%, specificity 93%; IA-2A sensitivity 68%, specificity 95% (19); and ZnT8A sensitivity 60%, specificity 88% (M. Hawa, unpublished data).

**HLA genotype.** HLA class II genotyping was performed in twin (data not shown) and population studies with sequence-specific oligonucleotides after DNA amplification with HLA-DRB1- and HLA-DQB1-specific primers (16). HLA risk status was defined as a categorical trait (risk/high risk). Carriers of any of the DRB1\*03, DRB1\*04, and/or DQB1\*302 alleles were considered high risk.

**Statistical analysis**

**Diabetes prediction.** To test whether serum autoantibodies, CML, and HLA risk status predict type 1 diabetes in 115 ICA<sup>+</sup> subjects, we estimated positive and negative predictive values as well as sensitivity and specificity of the markers using the Kaplan-Meier method. Multivariable Cox proportional hazards models were used to estimate hazard ratios (HR) and 95% CIs adjusted for age and sex. Survival time was defined as time from inclusion (months) to either diagnosis or last follow-up. Variables, irrespective of statistical significance in the univariable analyses, were CML, GADA, IA-2A, ZnT8A at inclusion, and HLA risk status. Model 1a included the classical risk factors HLA, GADA, IA-2A, and ZnT8A; model 1b included only CML; model 1c included all factors in models 1a and b; model 1d included only the significant predictors of models 1a and b. For models 1a–d, analyses were performed on case subjects with complete data (*n* = 73) on all predictors. In addition, two post hoc models were tested: model 2 included CML at onset categorized as high (>600 ng/mL) or low (<600 ng/mL) for all 115 ICA<sup>+</sup> subjects; model 3 included ΔCML (CML<sub>last visit</sub> – CML<sub>inclusion</sub>) calculated for the 73 subjects from whom serum was available at two different time points.

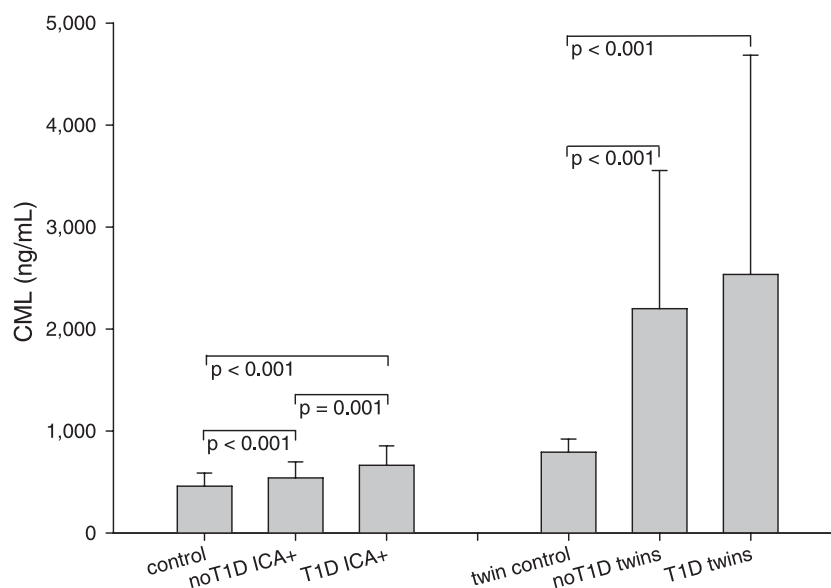
The proportional hazards assumption was examined using the log-minus-log plot of the survival function. To assess the overall discriminatory power for each model, we calculated the concordance (c)-statistic, an index comparable to the area under the receiver operating curve. Data handling and (preliminary) analyses were done with STATA 11.1 (StataCorp, College Station, Texas). Quantitative genetic modeling was performed using Mx software (20,21). Variables were natural logarithm transformed if distribution deviated from normal before statistical analyses. Tests were two-tailed, and *P* < 0.05 was considered significant.

**Twin study.** Because twin pairs were initially discordant for diabetes (cotwin case-control design) matched for age, genes (for MZs completely, for DZs in part), and shared childhood environmental exposures, paired Student *t* test examined differences within twin pairs (i.e., disease effects) after adjustment for sex. Determinants of CML levels and GADA, IA-2A, and ZnT8A were tested within a regression framework using generalized estimating equations, which takes the nonindependency of twin data into account to generate unbiased *P* values (20). To estimate the influence of genetic and environmental factors, we conducted quantitative genetic model fitting (21). In brief, we compared covariances (or correlations) in MZ and DZ twin pairs and quantified sources of individual differences by separation of observed phenotypic variance into additive (A) genetic, common (shared) (C), and unique (or nonshared) (E) environmental components. The significance of components A and C was assessed by testing deterioration in model fit after each component was dropped from the full model (ACE). Standard hierarchic  $\chi^2$  tests were used to select the best fitting model in combination with Akaike Information Criterion (AIC =  $\chi^2 - 2$  df). Mean levels of CML and GADA, IA-2A, and ZnT8A were adjusted for age and sex before calculating twin correlations with residuals used in model fitting.

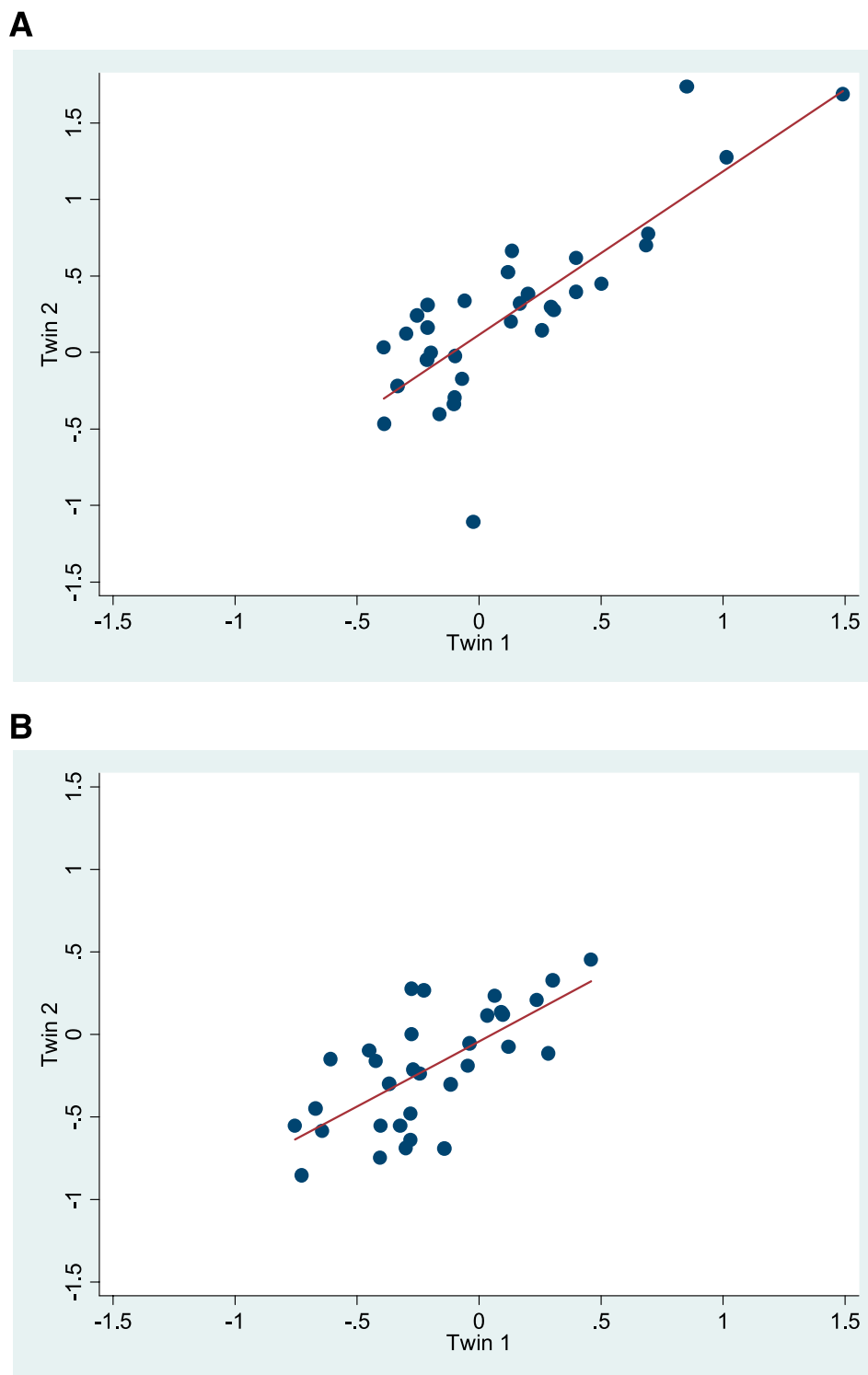
**RESULTS**

**Population study.** Of 115 ICA<sup>+</sup> subjects (Table 1), 33 developed type 1 diabetes after follow-up. The follow-up duration was similar in those who did and did not develop diabetes (*P* = 0.14). Compared with subjects who did not develop diabetes, prediabetic subjects had higher CML levels both at inclusion (*P* < 0.001) and at last follow-up (*P* < 0.001) even after correction for age and sex (*P* < 0.001 for both).

**Twin study.** In MZ and DZ twin cohorts, serum CML was not influenced by age, sex, diabetes status, or disease duration. CML levels were higher in MZ than in DZ twins (*P* <



**FIG. 1.** Mean (SD) baseline serum CML in the population and twin study. Graded CML increase in childhood population and adult twins according to ICA and/or type 1 diabetes (T1D). Control children (*n* = 2,102) CML mean (SD) [interquartile range] = 458.1 ng/mL (128.4) [372.0–534.0]; adult twins (*n* = 168) CML mean (SD) [interquartile range] = 792.2 ng/mL (127.4) [708.5–867.5]. No T1D ICA<sup>+</sup> (*n* = 82); T1D ICA<sup>+</sup> (*n* = 33); no T1D twins = nondiabetic cotwins (*n* = 64); T1D twins (*n* = 64). Untransformed CML values are plotted in the figure; however, all statistical analyses were performed on natural logarithm-transformed CML values.



**FIG. 2.** Scatterplots of diabetic (twin 1) vs. nondiabetic (twin 2) for natural logarithm-transformed serum CML. Serum CML corrected for age and sex for MZ (A) and DZ (B) twins shows strong correlations in both MZ ( $r = 0.81$ ) and DZ ( $r = 0.69$ ) twins irrespective of disease. (A high-quality color representation of this figure is available in the online issue.)

0.001) (Table 2), which we considered in quantitative genetic model fitting analyses. Diabetic twins had similar serum CML levels compared with their nondiabetic cotwins, but both were raised compared with normal control singletons and normal control twins (Fig. 1). Twin correlations for serum CML levels were strong and similar in MZ and DZ twin pairs (Fig. 2A and B). Model fitting showed

that additive genetic influences (A) could be dropped from the full ACE model without deterioration in fit (ACE vs. CE:  $\Delta\chi^2$  [df = 1] = 92.57–92.57 = 0,  $P = \text{NS}$ ). Shared environment (C) could not be dropped from the model because the fit deteriorated (ACE vs. AE:  $\Delta\chi^2$  [df = 1] = 103.64–92.57 = 11.07,  $P = 0.001$ ). Thus, the CE model showed the best fit, confirmed by the lowest AIC. In the best fitting

model, shared environmental factors explained 75% (95% CI 62–84) of individual differences, with the remainder due to nonshared environment (25% [16–38]).

#### Diabetes-associated autoantibodies

**Population study.** Of the 115 ICA<sup>+</sup> subjects, 33 developed diabetes (Table 1) and compared with those who did not, were more often positive for GADA ( $P = 0.04$ ), IA-2A ( $P < 0.001$ ), and ZnT8A ( $P < 0.001$ ). Subjects who developed diabetes were more often at risk based on their HLA alleles ( $P = 0.007$ ).

**Twin study.** More diabetic twins, compared with their nondiabetic twins, were GADA, IA-2A, and ZnT8A positive, irrespective of zygosity. Of 64 twin pairs, 31 diabetic twins had autoantibodies compared with 9 nondiabetic twins ( $P < 0.0001$ ) (Table 2). When analyzed as continuous traits, diabetic twins, compared with their nondiabetic cotwins, had higher values for GADA ( $P = 0.02$ ) and IA-2A ( $P = 0.001$ ) but not ZnT8A. Neither age nor disease duration affected GADA levels, but older subjects had less IA-2A ( $P < 0.001$ ) and ZnT8A ( $P = 0.002$ ). Twin correlations ( $r$ ) were weak for GADA,  $r(\text{MZ}) = -0.03$ ,  $r(\text{DZ}) = 0.40$ ; IA-2A,  $r(\text{MZ}) = 0.06$ ,  $r(\text{DZ}) = -0.03$ ; and ZnT8A,  $r(\text{MZ}) = 0.15$ ,  $r(\text{DZ}) = 0.15$ . Model fitting showed that for GADA, IA-2A, and ZnT8A, the so-called E model (including only the unique environmental [E] variance component) was most parsimonious with best fit.

TABLE 3  
Prediction models for developing type 1 diabetes in the population study

	<i>n</i>	HR [95% CI]	<i>P</i> value	Harrell C
<b>Model 1a</b>				
HLA risk	73	2.12 [0.58–7.72]	0.25	0.73
GADA		1.24 [0.33–4.65]	0.75	
IA-2A		1.44 [0.60–3.47]	0.41	
ZnT8A		3.10 [1.01–9.53]	0.048	
<b>Model 1b</b>				
CML <sub>inclusion</sub>	73	1.005 [1.002–1.007]	0.001	0.67
<b>Model 1c</b>				
CML <sub>inclusion</sub>	73	1.003 [1.0003–1.006]	0.03	0.74
HLA risk		1.82 [0.50–0.62]	0.36	
GADA		1.20 [0.33–4.35]	0.78	
IA-2A		1.38 [0.56–3.41]	0.48	
ZnT8A		2.45 [0.77–7.76]	0.13	
<b>Model 1d</b>				
CML <sub>inclusion</sub>	73	1.003 [1.0006–1.006]		0.73
ZnT8A		3.04 [1.05–8.80]	0.04	
<b>Model 2</b>				
CML <sub>inclusion</sub> (high/low)	115	4.98 [2.25–10.99]	<0.001	0.66
<b>Model 3</b>				
ΔCML	73	1.005 [1.003–1.008]	<0.001	0.76

Results of different risk factors using multivariable Cox proportional hazards models. Serum CML when high at inclusion in study or change in CML during study were best predictors, and only CML and ZnT8A independently predicted type 1 diabetes. All models included age and sex. Model 1a included the classical risk factors HLA, GADA, IA-2A, and ZnT8A; model 1b included only CML; model 1c included all factors in models 1a and b; model 1d included only the significant predictors of models 1a and b. For models 1a–d, analyses were performed on cases with complete data ( $n = 73$ ) on all predictors. In addition, two post hoc models were tested: model 2 included CML at onset categorized as high (>600 ng/mL) or low (<600 ng/mL) for all 115 ICA<sup>+</sup> subjects; model 3 included ΔCML (CML<sub>last visit</sub> – CML<sub>inclusion</sub>) calculated for the 73 subjects from whom serum was available at two different time points.

#### Diabetes prediction

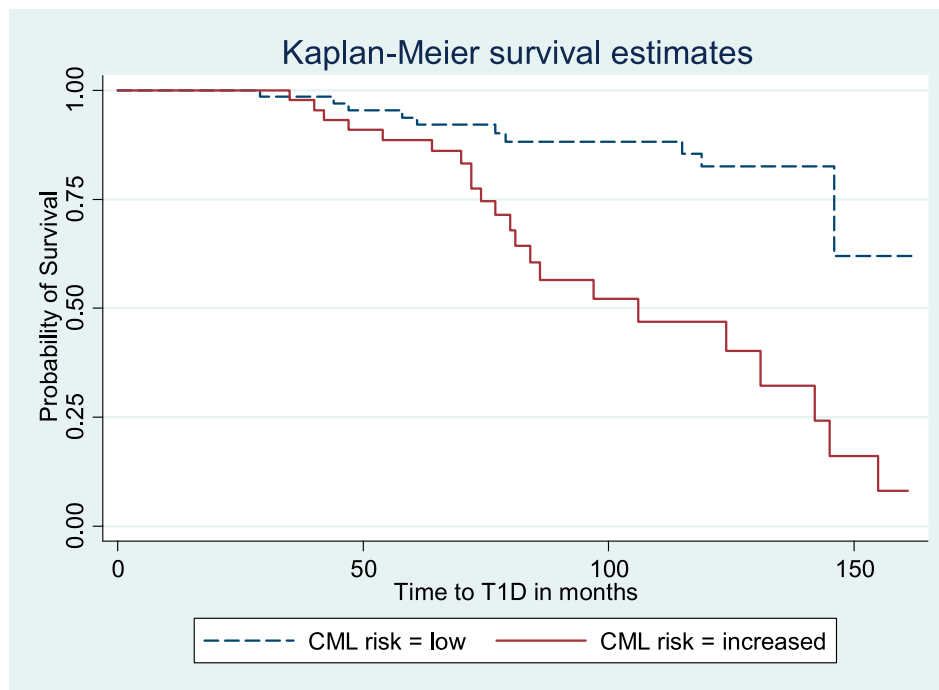
**Population study.** Kaplan-Meier survival tables were used to calculate absolute 10-year risk according to CML risk status based on the 115 ICA<sup>+</sup> subjects. Positive and negative predictive values, respectively, were as follows: raised CML<sub>inclusion</sub> 52.2 and 82.6% (sensitivity 67.7%, specificity 72.2%); GADA 36.0 and 86.2% (sensitivity 88.6%, specificity 31.3%); IA-2A 54.8 and 81.0% (sensitivity 51.5%, specificity 82.9%); and ZnT8A 45.7 and 87.5% (sensitivity 76.2%, specificity 64.8%). When combining the CML risk status with each autoantibody status, the accuracy increased since the positive predictive values increased and the negative predictive values remained stable: GADA 51.4 and 79.5% (positive and negative, respectively) (sensitivity 54.3%, specificity 77.5%); IA-2A 70.6 and 78.6% (sensitivity 36.4%, specificity 93.9%); and ZnT8A 62.5 and 79.7% (sensitivity 45.5%, specificity 88.7%). In Cox proportional hazards models, age or sex did not predict diabetes (Table 3). Model 1a shows that of the classic risk markers, only ZnT8A significantly contributed to prediction of diabetes. Model 1b shows that the new marker CML<sub>inclusion</sub> significantly predicted diabetes. Robustness analyses using all ICA<sup>+</sup> subjects ( $n = 115$ ) rather than those with complete marker data ( $n = 73$ ) gave comparable results to model 1b (HR 1.003 [95% CI 1.001–1.006];  $P = 0.04$ , Harrell C = 0.73). Testing significant predictors from models 1a and b showed CML<sub>inclusion</sub> and ZnT8A independently predicted diabetes, with a high c-statistic (model 1d).

**Post hoc tests of predictive thresholds.** To determine potential threshold levels for predictive capacity, CML<sub>inclusion</sub> was divided into deciles to use as a categorical predictor. Relative risk of diabetes increased with each decile, but only the 4 upper deciles contributed substantially to diabetes prediction (HR 7.05–11.1;  $P = 0.08–0.03$ ); when categorized into 4 upper deciles (CML<sub>inclusion</sub> >600 ng/mL) and a lower decile (CML<sub>inclusion</sub> <600 ng/mL), the former had fivefold diabetes risk (Fig. 3 and model 2 in Table 3). Persistence of CML levels (i.e., ΔCML = CML<sub>last visit</sub> – CML<sub>inclusion</sub>) predicted diabetes (model 3 in Table 3), reflecting the persistent higher CML levels in subjects who developed type 1 diabetes and that CML levels further increased in subjects who developed type 1 diabetes but decreased in subjects who did not develop type 1 diabetes.

**Twin study.** Of 64 nondiabetic twins, 62 (96.9%) had a raised serum CML (>1,097 ng/mL as 99th centile of normal twin values), as did 57 (89.1%) of the diabetic twins (Fig. 1); 10 of 64 had diabetes-associated autoantibodies, and 2 of these latter 10 subsequently developed diabetes (Table 2).

#### DISCUSSION

These observations identify distinct biomarkers, one novel (serum CML), the other an established disease predictor (diabetes-associated autoantibodies), which are now shown to have strong additive and quantitative predictive value in a large population study. Using the markers at inclusion only, an optimally predicting model consisted of CML and ZnT8A with reasonable discriminatory power. Studying a selected twin cohort excluded genetic factors as sole determinants of both serum CML levels and serum diabetes-associated autoantibodies. It follows that non-genetic, most probably environmental, factors contribute to these diabetes-associated features, where the term *environmental* represents all nongenetic contributions to the variation in a trait, whether internal or external to the



**FIG. 3.** Kaplan-Meier survival curves for type 1 diabetes (T1D) risk based on baseline CML levels. Subjects in the upper 4 deciles of the CML distribution were found to have high risk and subjects in the bottom 6 deciles low risk of developing T1D. The curves plotted in the figure are a graphic representation of the analysis presented in Table 3, model 2 ( $n = 115$ ). (A high-quality color representation of this figure is available in the online issue.)

individual. Our study indicated that familial (that is, shared) environmental factors are important determinants of serum CML levels, while nonshared environmental factors are important determinants of autoantibody frequency. It follows that these predictive biomarkers are largely determined by distinct environmental effects.

Despite the limited heritability of serum CML and diabetes-associated autoantibodies, in this highly selected cohort of discordant twins, these two factors are likely to be influenced by genetic factors. For example, additive genetic effects (heritability) explained 74% of normal population variance in serum CML (14). Previous twin studies of diabetes-associated autoantibodies also support a genetic influence, though one such study examines only nondiabetic twins without matching them (8,9). In family studies, those with high-risk HLA genes show diabetes-associated autoantibodies earlier, while ZnT8A is associated with specific genetic susceptibility (7,22). Nevertheless, the risk of autoantibodies appearing is influenced by dietary modification in infancy (23,24). HbA<sub>1c</sub> (available for 108 of 115 ICA<sup>+</sup> subjects), however, was not associated with CML (data not shown), replicating results reported in healthy twins (14). Raised serum CML emerges as a novel and potent biomarker of diabetes risk—a predictive effect most marked at the highest serum CML levels. In this quantitative predictive effect, CML resembles autoantibodies that are most predictive at high titers and in combination with other autoantibodies (25).

This study has limitations. Because our patient cohort was composed exclusively of Caucasian participants, generalization to other ethnic groups is not possible. Both serum CML and diabetes-associated autoantibodies are likely to be only surrogate markers of putative destructive innate and adaptive immune changes, respectively. A twin study should ideally be performed prospectively in a population-based cohort from birth to determine the rate of

induction of autoantibodies and diabetes, and because our twins were initially disease discordant, disease concordance rate is underestimated (26), but our analysis limits bias from the strong disease association with these autoantibodies (2). If autoantibodies were genetically determined, even nondiabetic MZ twins would eventually show them, which, in general, they did not. While we noted higher CML levels in twins, it would not affect our interpretation of the results, though it could be due to differences between the Ulm population study and the U.K. twin study in subjects' age and diabetes duration and differences in sample storage ( $-80$  and  $-30^{\circ}\text{C}$ , respectively). Furthermore, the cause of the raised serum CML, and its relationship to the origin of diabetes, is unclear, though a dietary source is most likely, consistent with a shared familial origin irrespective of zygosity (13). Specifically, thermally sensitive nutrients, including infant formula cow's milk and heat-treated animal fat, are major sources of AGEs, and AGEs, including CML, can reach adult serum levels by 1 year of age and have been implicated in sustaining an altered inflammatory response (13). Indeed, while we found raised AGE levels, in the form of serum CML, in young children before the onset of type 1 diabetes, AGE levels also increase with age and disease duration and are associated with diabetes complications (18). Different AGEs and different AGE receptors are responsible for their inflammatory action, and further studies will be required to define which are involved in the prediabetic process (e.g., increase in the AGE serum carboxyethyl-lysine is associated with relapses in multiple sclerosis) (27,28).

The present observations, which imply that at least two distinct environmental events cause type 1 diabetes, could consolidate different theories regarding the pathogenesis of autoimmune diabetes (e.g., hygiene hypothesis and

accelerator hypothesis), as well as different schools of opinion regarding candidate environmental agents (e.g., enteroviruses and diet) (1–3,13). That only two biomarkers (serum levels of a proinflammatory AGEs and diabetes-associated autoantibodies) can predict most children who develop diabetes suggests that there may be a finite number of critical environmental factors causing the disease. Weaning to a highly hydrolyzed casein formula in place of conventional cow's milk-based formula recently has been shown to reduce the frequency of appearance of diabetes-associated autoantibodies (24). The potent predictive value of the AGE biomarker in this present study, supported by another recent study, raises the possibility that dietary modifications in early life to reduce AGE intake could prevent progression to type 1 diabetes (29). Furthermore, reduction of AGEs by diet or drug therapy in animal models of autoimmune diabetes reduces not only AGE levels but also progression to diabetes (30).

#### ACKNOWLEDGMENTS

J.C.H. was supported by the Children's Diabetes Foundation in Denver, the University of Colorado Denver Diabetes and Endocrinology Research Center (National Institutes of Health [NIH] Grant P30-DK-57516), NIH Grant R01-DK-052068, and the Juvenile Diabetes Research Foundation International Autoimmunity Center Consortium; B.O.B. was supported by Deutsche Forschungsgemeinschaft (DFG SFB 518/ GRK 1041) and State Baden-Wuerttemberg Centre of Excellence "Metabolic Disorders"; and R.D.L. was supported by grants from the British Diabetic Twin Research Trust and the Juvenile Diabetes Research Foundation International.

H.Be. was in receipt of an Eli Lilly award. No other potential conflicts of interest relevant to this article were reported.

H.Be., M.S., and M.I.H. obtained and analyzed samples. H.R., H.S., and H.Bu. researched data. G.B. obtained samples. H.W.D. and J.C.H. supplied assay material. B.O.B. and R.D.L. conceived the project and wrote the manuscript. All authors contributed to the final manuscript. R.D.L. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Parts of this study were presented in abstract form at the 70th Scientific Sessions of the American Diabetes Association, Orlando, Florida, 4–8 June 2011, and at the 46th Annual Meeting of the European Association for the Study of Diabetes, Lisbon, Portugal, 12–16 September 2011.

#### REFERENCES

- Bach JF. The effect of infections on susceptibility to autoimmune and allergic diseases. *N Engl J Med* 2002;347:911–920
- Bluestone JA, Herold K, Eisenbarth G. Genetics, pathogenesis and clinical interventions in type 1 diabetes. *Nature* 2010;464:1293–1300
- Todd JA. Etiology of type 1 diabetes. *Immunity* 2010;32:457–467
- Redondo MJ, Yu L, Hawa M, et al. Heterogeneity of type 1 diabetes: analysis of monozygotic twins in Great Britain and the United States. *Diabetologia* 2001;44:354–362
- Barnett AH, Eff C, Leslie RD, Pyke DA. Diabetes in identical twins. A study of 200 pairs. *Diabetologia* 1981;20:87–93
- Leslie RD, Delli Castelli M. Age-dependent influences on the origins of autoimmune diabetes: evidence and implications. *Diabetes* 2004;53:3033–3040
- Kupila A, Keskinen P, Simell T, et al. Genetic risk determines the emergence of diabetes-associated autoantibodies in young children. *Diabetes* 2002;51:646–651
- Hawa M, Rowe R, Lan MS, et al. Value of antibodies to islet protein tyrosine phosphatase-like molecule in predicting type 1 diabetes. *Diabetes* 1997;46:1270–1275
- Redondo MJ, Rewers M, Yu L, et al. Genetic determination of islet cell autoimmunity in monozygotic twin, dizygotic twin, and non-twin siblings of patients with type 1 diabetes: prospective twin study. *BMJ* 1999;318:698–702
- Wenzlau JM, Juhl K, Yu L, et al. The cation efflux transporter ZnT8 (Slc30A8) is a major autoantigen in human type 1 diabetes. *Proc Natl Acad Sci U S A* 2007;104:17040–17045
- Beyan H, Drexhage RC, van der Heul Nieuwenhuijsen L, et al. Monocyte gene-expression profiles associated with childhood-onset type 1 diabetes and disease risk: a study of identical twins. *Diabetes* 2010;59:1751–1755
- Beyan H, Goodier MR, Nawroly NS, et al. Altered monocyte cyclooxygenase response to lipopolysaccharide in type 1 diabetes. *Diabetes* 2006;55:3439–3445
- Mericq V, Piccardo C, Cai W, et al. Maternally transmitted and food-derived glycotoxins: a factor preconditioning the young to diabetes? *Diabetes Care* 2010;33:2232–2237
- Leslie RD, Beyan H, Sawtell P, Boehm BO, Spector TD, Snieder H. Level of an advanced glycated end product is genetically determined: a study of normal twins. *Diabetes* 2003;52:2441–2444
- Genuth S, Alberti KG, Bennett P, et al.; Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Follow-up report on the diagnosis of diabetes mellitus. *Diabetes Care* 2003;26:3160–3167
- Boehm BO, Manfras B, Seissler J, et al. Epidemiology and immunogenetic background of islet cell antibody-positive nondiabetic schoolchildren. Ulm-Frankfurt population study. *Diabetes* 1991;40:1435–1439
- Endl J, Rosinger S, Schwarz B, et al. Coexpression of CD25 and OX40 (CD134) receptors delineates autoreactive T-cells in type 1 diabetes. *Diabetes* 2006;55:50–60
- Boehm BO, Schilling S, Rosinger S, et al. Elevated serum levels of N(epsilon)-carboxymethyl-lysine, an advanced glycation end product, are associated with proliferative diabetic retinopathy and macular oedema. *Diabetologia* 2004;47:1376–1379
- Törn C, Mueller PW, Schlosser M, Bonifacio E, Bingley PJ; Participating Laboratories. Diabetes Antibody Standardization Program: evaluation of assays for autoantibodies to glutamic acid decarboxylase and islet antigen-2. *Diabetologia* 2008;51:846–852
- Neale M, Cardon LR. *Methodology for Genetic Studies in Twins and Families*. Dordrecht, the Netherlands, Kluwer Academic Publishers, 1992
- Neale M, Boker SM, Xie G, Maes HH. *Mx: Statistical Modeling*. Richmond, VA, Department of Psychiatry, Virginia Commonwealth University, 1999
- Wenzlau JM, Liu Y, Yu L, et al. A common nonsynonymous single nucleotide polymorphism in the SLC30A8 gene determines ZnT8 autoantibody specificity in type 1 diabetes. *Diabetes* 2008;57:2693–2697
- Ziegler AG, Schmid S, Huber D, Hummel M, Bonifacio E. Early infant feeding and risk of developing type 1 diabetes-associated autoantibodies. *JAMA* 2003;290:1721–1728
- Knip M, Virtanen SM, Seppä K, et al.; Finnish TRIGR Study Group. Dietary intervention in infancy and later signs of beta-cell autoimmunity. *N Engl J Med* 2010;363:1900–1908
- Ziegler AG, Nepom GT. Prediction and pathogenesis in type 1 diabetes. *Immunity* 2010;32:468–478
- Redondo MJ, Jeffrey J, Fain PR, Eisenbarth GS, Orban T. Concordance for islet autoimmunity among monozygotic twins. *N Engl J Med* 2008;359:2849–2850
- Sternberg Z, Hennies C, Sternberg D, et al. Diagnostic potential of plasma carboxymethyllysine and carboxyethyllysine in multiple sclerosis. *J Neuroinflammation* 2010;7:72
- Yan SF, Ramasamy R, Schmidt AM. Mechanisms of disease: advanced glycation end-products and their receptor in inflammation and diabetes complications. *Nat Clin Pract Endocrinol Metab* 2008;4:285–293
- Coughlan MT, Yap FY, Tong DC, et al. Advanced glycation end products are direct modulators of  $\beta$ -cell function. *Diabetes* 2011;60:2523–2532
- Peppas M, He C, Hattori M, McEvoy R, Zheng F, Vlassara H. Fetal or neonatal low-glycotxin environment prevents autoimmune diabetes in NOD mice. *Diabetes* 2003;52:1441–1448