The American Journal of Human Genetics, Volume *86*

Supplemental Data

Evolutionary and Functional Analysis of Celiac Risk Loci Reveals SH2B3 as a Protective Factor against Bacterial Infection

Alexandra Zhernakova, Clara C. Elbers, Bart Ferwerda, Jihane Romanos, Gosia Trynka, Patrick C. Dubois, Carolien G.F. de Kovel, Lude Franke, Marije Oosting, Donatella Barisani, Maria Teresa Bardella, Finnish Celiac Disease Study Group, Katri Kaukinen, Kalle Kurppa, Markku Mäki, Leo A.B. Joosten, Paivi Saavalainen, David A. van Heel, Carlo Catassi, Mihai G. Netea, and Cisca Wijmenga

Figure S1

EHH plot of ancestral and derived haplotypes per population in *IL12A*, *IL18RAP* and *SH2B3* gene loci.

Table S1

Distribution of iHS values in European controls and in the Saharawi population.

Table S2

Estimation of the age of the selective sweep in European populations, with different EHH thresholds.

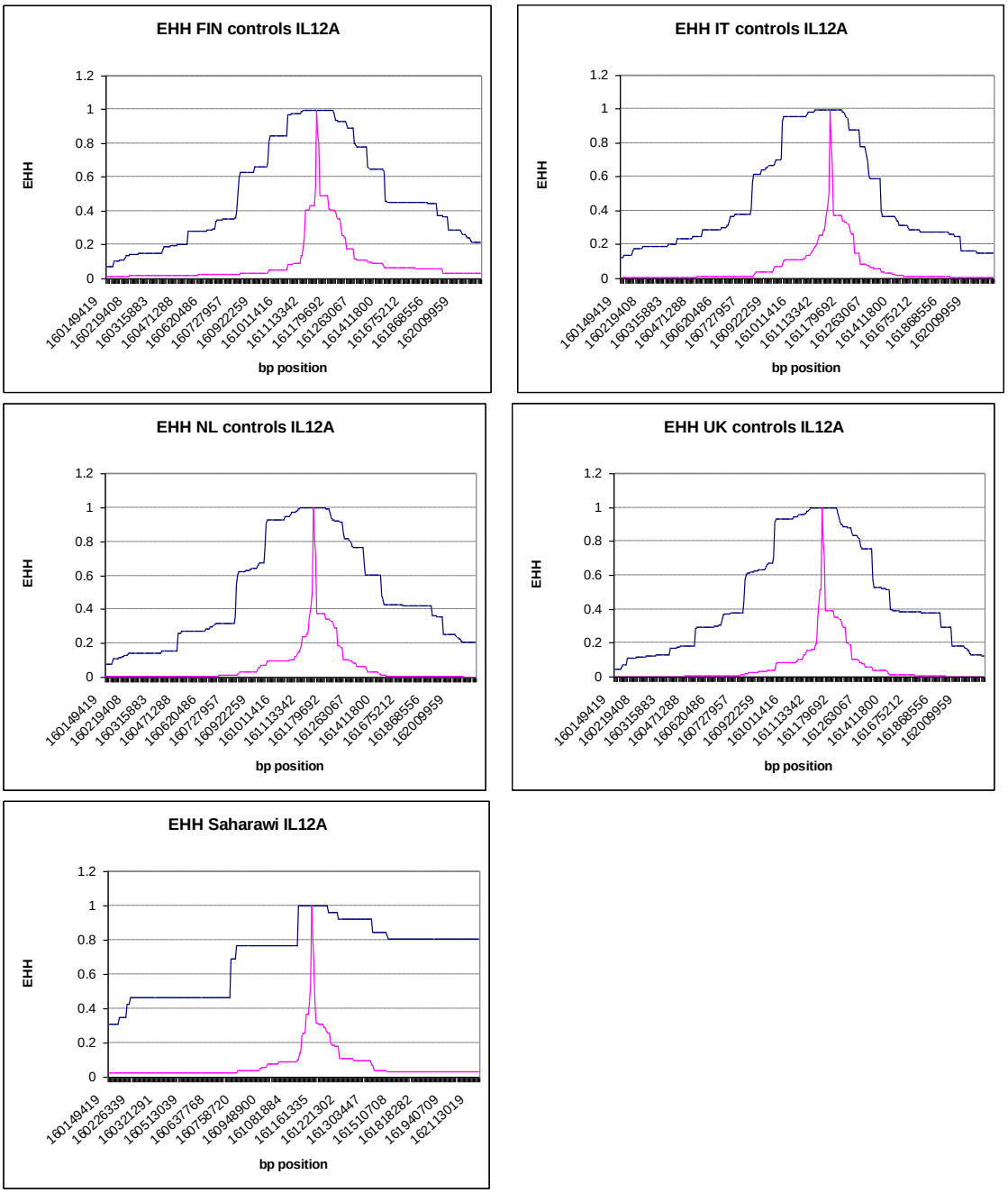
Full acknowledgments

Figure S1

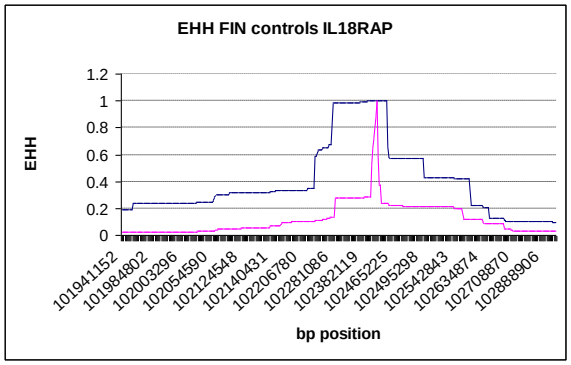
EHH plot of derived (blue) and ancestral (red) haplotypes of

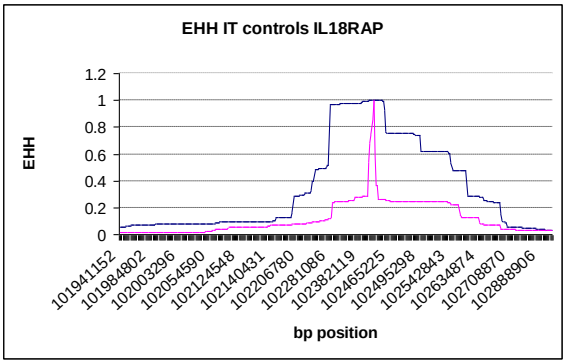
a. rs17810546 SNP in *IL12A* locus (1 Mb left and right from rs17810546)
b. rs917997 SNP in *IL18RAP* locus (500 kb left and right from rs917997)
c. rs3184504 SNP in *SH2B3* locus (1.4 Mb left and right from rs3184504 are included).

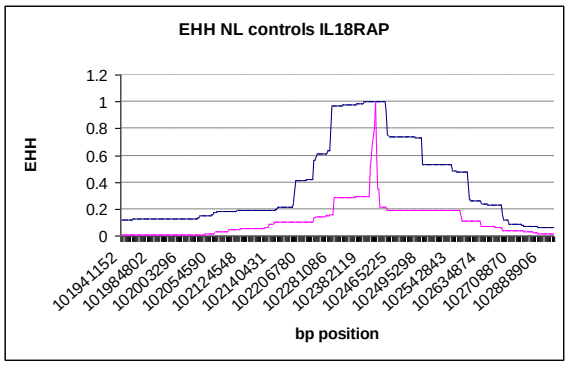
A.

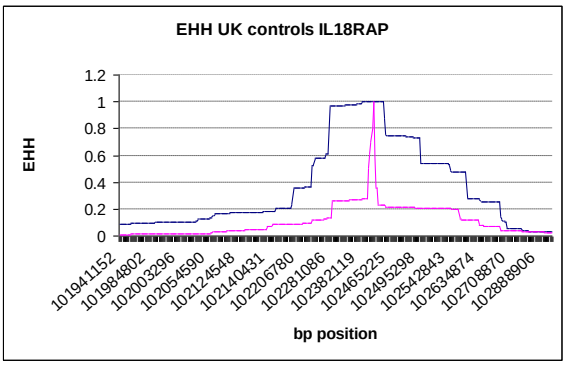


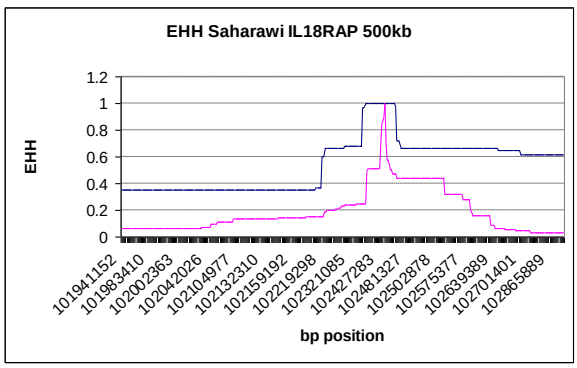
B.


EHH FIN controls IL18RAP
EHH
1.2
1
0.8
0.6
0.4
0.2
0
101941152
101984802
102003296
102054590
102124548
102140431
102206780
102281086
102382119
102465225
102495298
102542843
102634874
102708870
102888906
bp position

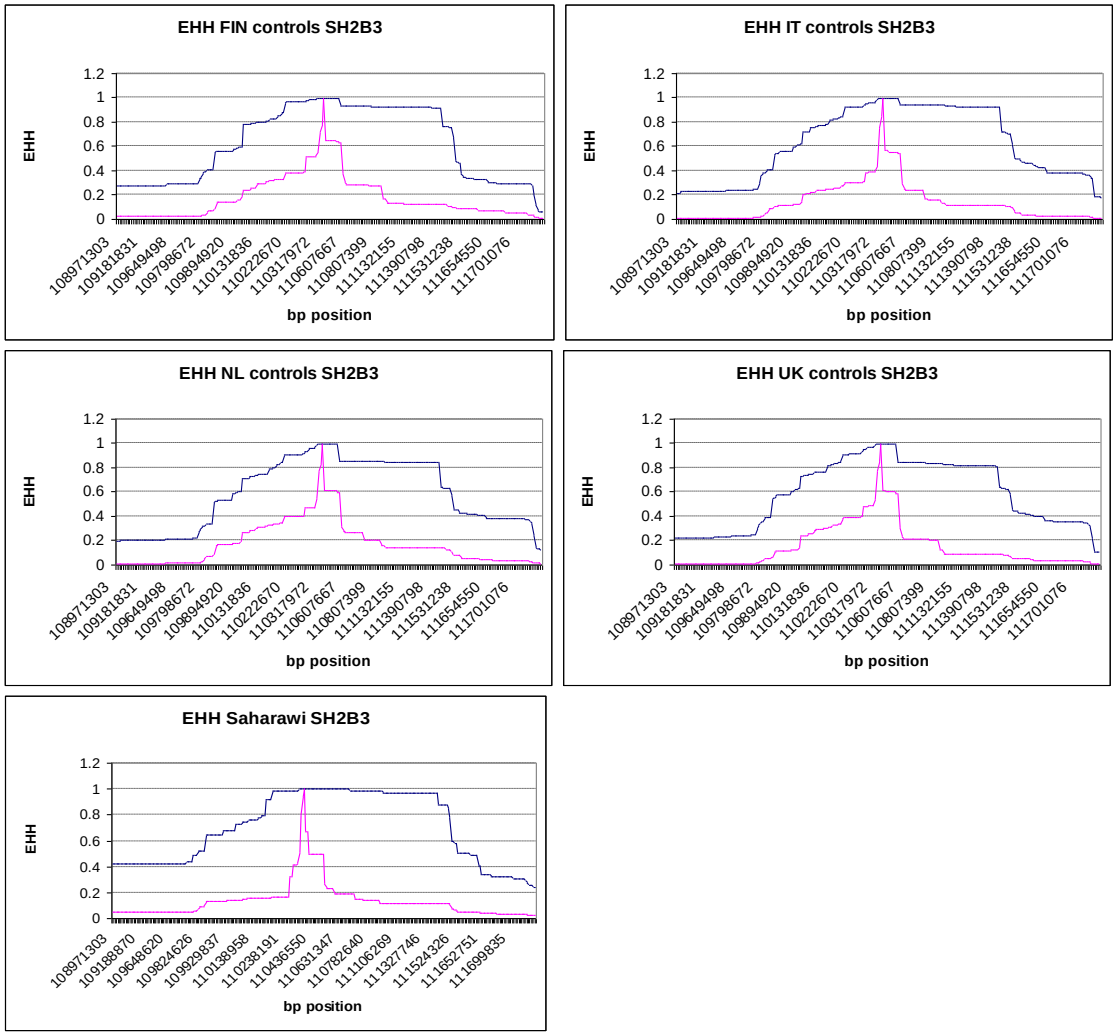

EHH IT controls IL18RAP
EHH
1.2
1
0.8
0.6
0.4
0.2
0
101941152
101984802
102003296
102054590
102124548
102140431
102206780
102281086
102382119
102465225
102495298
102542843
102634874
102708870
102888906
bp position


EHH NL controls IL18RAP
EHH
1.2
1
0.8
0.6
0.4
0.2
0
101941152
101984802
102003296
102054590
102124548
102140431
102206780
102281086
102382119
102465225
102495298
102542843
102634874
102708870
102888906
bp position


EHH UK controls IL18RAP
EHH
1.2
1
0.8
0.6
0.4
0.2
0
101941152
101984802
102003296
102054590
102124548
102140431
102206780
102281086
102382119
102465225
102495298
102542843
102634874
102708870
102888906
bp position


EHH Saharawi IL18RAP 500kb
EHH
1.2
1
0.8
0.6
0.4
0.2
0
101941152
101983410
102002363
102042026
102104977
102132310
102159192
102219298
102321085
102427283
102481327
102502878
102575377
102639389
102701401
102865889
bp position

C.



Populations: FIN – Finnish, IT – Italian, NL – Dutch, UK – United Kingdom

Table S1

Distribution of iHS values in European controls and in the Saharawi population

gene	population	RS ID	N_ind	Allele Frq	IHSnoncorrected	IHScorrected	p value
CCR3	FIN controls	rs6441961	1829	0.294	-0.688	0.056	0.955
CCR3	IT controls	rs6441961	543	0.324	0.234	1.562	0.118
CCR3	NL controls	rs6441961	846	0.31	-0.789	-0.534	0.594
CCR3	Saharawi	rs6441961	195	0.249	-1.221	-0.422	0.673
CCR3	UK3 controls	rs6441961	4936	0.297	-0.939	-0.728	0.467
CCR3	HapMapCEU	rs6441961	60	0.317		-0.025	
CCR3	HapMapYRI	rs6441961	60	0.075		1.431	
CCR3	HapMapAZN	rs6441961	89	0.339		-0.757	
IL12A	FIN controls	rs9811792	1829	0.475	-1.239	-1.595	0.111
IL12A	IT controls	rs9811792	543	0.42	-1.159	-1.342	0.180
IL12A	NL controls	rs9811792	846	0.444	-1.218	-1.424	0.154
IL12A	Saharawi	rs9811792	195	0.344	-1.373	-1.174	0.240
IL12A	UK3 controls	rs9811792	4936	0.46	-1.063	-1.157	0.247
IL12A	HapMapCEU	rs9811792	60	0.525		-0.937	
IL12A	HapMapYRI	rs9811792	60	0.3		-0.556	
IL12A	HapMapAZN	rs9811792	89	0.61		-2.187	
IL12A	FIN controls	rs17810546	1829	0.101	-2.792	-2.936	0.0033
IL12A	IT controls	rs17810546	543	0.067	-2.729	-2.923	0.0035
IL12A	NL controls	rs17810546	846	0.118	-2.874	-3.434	0.0006
IL12A	Saharawi		195	0.06	not calculated		
IL12A	UK3 controls	rs17810546	4936	0.128	-2.811	-3.321	0.0009
IL12A	HapMapCEU	rs17810546	60	0.1		-2.492	
IL12A	HapMapYRI	rs17810546	60	not present		n/a	
IL12A	HapMapAZN	rs17810546	89	not present		n/a	
IL18RAP	FIN controls	rs917997	1829	0.192	-1.822	-1.383	0.167
IL18RAP	IT controls	rs917997	543	0.23	-1.847	-2.036	0.042
IL18RAP	NL controls	rs917997	846	0.216	-1.859	-1.911	0.056
IL18RAP	Saharawi	rs917997	195	0.154	-2.562	-1.703	0.089
IL18RAP	UK3 controls	rs917997	4936	0.238	-1.877	-2.036	0.042
IL18RAP	HapMapCEU	rs917997	60	0.233		-1.739	
IL18RAP	HapMapYRI	rs917997	60	0.067		-0.450	
IL18RAP	HapMapAZN	rs917997	89	0.406		-2.115	
IL2_IL21	FIN controls	rs13151961	1829	0.115	-1.167	-0.334	0.738
IL2_IL21	IT controls	rs13151961	543	0.12	-1.046	-0.483	0.629
IL2_IL21	NL controls	rs13151961	846	0.194	-1.164	-0.516	0.606
IL2_IL21	Saharawi		195	0.02	not calculated		
IL2_IL21	UK3 controls	rs13151961	4936	0.167	-1.182	-0.521	0.603
IL2_IL21	HapMapCEU	rs13151961	60	0.192		-1.976	
IL2_IL21	HapMapYRI	rs13151961	60	0.017		not calculated	
IL2_IL21	HapMapAZN	rs13151961	89	not present		n/a	
LPP	FIN controls	rs1464510	1829	0.547	0.265	0.416	0.677
LPP	IT controls	rs1464510	543	0.471	0.186	0.722	0.470
LPP	NL controls	rs1464510	846	0.493	0.112	0.560	0.576

LPP	Saharawi	rs1464510	195	0.313	0.079	1.104	0.270
LPP	UK3 controls	rs1464510	4936	0.435	0.099	0.668	0.504
LPP	HapMapCEU	rs1464510	60	0.425		-0.673	
LPP	HapMapYRI	rs1464510	60	0.183		-0.629	
LPP	HapMapAZN	rs1464510	89	0.511		0.794	
REL	FIN controls	rs842647	1829	0.323	-0.989	-0.852	0.394
REL	IT controls	rs842647	543	0.263	-1.065	-0.764	0.445
REL	NL controls	rs842647	846	0.346	-1.152	-1.100	0.271
REL	Saharawi	rs842647	195	0.136	-1.685	-0.512	0.608
REL	UK3 controls	rs842647	4936	0.348	-1.122	-1.022	0.307
REL	HapMapCEU	rs842647	60	0.3		-0.516	
REL	HapMapYRI	rs842647	60	0.475		-1.793	
REL	HapMapAZN	rs842647	89	0.828		-0.279	
RGS1	FIN controls	rs2816316	1829	0.857	0.701	-0.110	0.912
RGS1	IT controls	rs2816316	543	0.831	0.752	0.184	0.854
RGS1	NL controls	rs2816316	846	0.711	-0.081	-0.756	0.450
RGS1	Saharawi	rs2816316	195	0.756	0.657	0.106	0.915
RGS1	UK3 controls	rs2816316	4936	0.822	0.754	0.042	0.966
RGS1	HapMapCEU	rs2816316	60	0.783		0.695	
RGS1	HapMapYRI	rs2816316	60	0.733		-0.334	
RGS1	HapMapAZN	rs2816316	89	0.767		1.039	
SH2B3	FIN controls	rs3184504	1829	0.398	-1.766	-2.129	0.033
SH2B3	IT controls	rs3184504	543	0.494	-1.977	-2.597	0.009
SH2B3	NL controls	rs3184504	846	0.465	-1.680	-2.114	0.035
SH2B3	Saharawi	rs3184504	195	0.151	-2.209	-1.224	0.221
SH2B3	UK3 controls	rs3184504	4936	0.468	-1.736	-2.214	0.027
SH2B3	HapMapCEU	rs3184504	60	0.408		-2.756	
SH2B3	HapMapYRI	rs3184504	60	not present		n/a	
SH2B3	HapMapAZN	rs3184504	89	not present		n/a	
TAGAP	FIN controls	rs1738074	1829	0.58	-1.150	-1.673	0.094
TAGAP	IT controls	rs1738074	543	0.587	-1.029	-1.450	0.147
TAGAP	NL controls	rs1738074	846	0.605	-1.042	-2.059	0.039
TAGAP	Saharawi	rs1738074	195	0.531	-1.141	-1.425	0.154
TAGAP	UK3 controls	rs1738074	4936	0.577	-1.056	-1.499	0.134
TAGAP	HapMapCEU	rs1738074	60	0.508		-1.287	
TAGAP	HapMapYRI	rs1738074	60	0.283		-1.275	
TAGAP	HapMapAZN	rs1738074	89	0.339		0.637	
TNFAIP3	FIN controls	rs2327832	1829	0.185	-1.420	-0.740	0.460
TNFAIP3	IT controls	rs2327832	543	0.157	-1.475	-1.220	0.222
TNFAIP3	NL controls	rs2327832	846	0.205	-1.566	-1.430	0.153
TNFAIP3	Saharawi	rs2327832	195	0.249	-1.354	-0.616	0.538
TNFAIP3	UK3 controls	rs2327832	4936	0.23	-1.584	-1.531	0.126
TNFAIP3	HapMapCEU	rs2327832	60	0.175		-1.491	
TNFAIP3	HapMapYRI	rs2327832	60	0.108		-0.544	
TNFAIP3	HapMapAZN	rs2327832	89	0.006		not calculated	

Table S2

Estimation of the age of the selective sweep in European populations, with different EHH thresholds

A. IL12A

threshold EHH	**0.25**	**0.3**	**0.35**
FIN controls	2248	2303	2054
IT controls	2479	3381	3446
NL controls	2231	2297	2182
UK controls	2423	2410	2107

B. IL18RAP

threshold EHH	**0.25**	**0.3**	**0.35**
FIN controls	6282	5560	6822
IT controls	7481	7700	6822
NL controls	7606	7111	6200
UK controls	7136	7111	6200

C. SH2B3

threshold EHH	**0.25**	**0.3**	**0.35**
FIN controls	1248	1608	1502
IT controls	1727	1525	1339
NL controls	1756	1541	1374
UK controls	1756	1525	1370

Full acknowledgments

This study was supported by the Celiac Disease Consortium, an Innovative Cluster approved by: the Netherlands Genomics Initiative and partially funded by the Dutch Government (BSIK03009), the Netherlands Organization for Scientific Research (NWO, VICI grant 918.66.620 to CW), EU STREP KP6 (grant 036383), SenterNovem (IOP genomics IGE05012), and the Wellcome Trust (grant WT084743MA to DAvH). We acknowledge use of DNA from the British 1958 Birth Cohort collection, funded by the UK Medical Research Council (grant G0000934) and the Wellcome Trust (grant 068545/Z/02). GT was awarded a Ter Meulen Fund travel grant by the Royal Netherlands Academy of Arts and Sciences (KNAW). MGN was supported by a VIDI grant from the Netherlands Organization for Scientific Research (NWO). PCD is an MRC Clinical Training Fellow (G0700545). We thank all the clinicians and Coeliac UK for their help in recruiting individuals for this study. The Finnish Celiac Disease Study Group is represented by Katri Kaukinen, Kalle Kurppa, and Markku Mäki. We thank C.J.J. Mulder, M.L. Mearin, J. Sweizer, W.H.M. Verbeek, G.J. Tack, V.M. Wolters, and R.H.J. Houwen for collecting Dutch samples and F. Cucca, S. Teresi, G. Iacono, R. Francavilla, S. Castellaneta, M.D. Cantarero and A. Ravelli for collecting the Saharawi families. We thank P. Deloukas, R. Gwilliam and the Wellcome Trust Sanger Institute for genotyping the UK controls. We thank K. Hunt, V. Mistry and G. Heap for sample preparation and A. Curtotti and C. Mein for assistance with genotyping. This study used data generated by the Wellcome Trust Case-Control Consortium 2. A full list of the investigators who contributed to generating the data is available from www.wtccc.org.uk. Funding for the project was provided by the Wellcome Trust under award 085475. This research utilized resources provided by the Type 1 Diabetes Genetics Consortium, a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Human Genome Research Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418. We thank all the individuals who participated in the study, J. Senior for critically reading the manuscript, and F. Mulder, M. Charlaganov and M. Golov for help with the data analysis.