

Henna Kallionpää,^{1,2,3} Laura L. Elo,^{1,4} Essi Laajala,^{1,3,5,6} Juha Mykkänen,^{1,3,7,8} Isis Ricaño-Ponce,⁹ Matti Vaarma,¹⁰ Teemu D. Laajala,¹ Heikki Hyöty,^{11,12} Jorma Ilonen,^{13,14} Riitta Veijola,¹⁵ Tuula Simell,^{3,7} Cisca Wijmenga,⁹ Mikael Knip,^{3,16,17,18} Harri Lähdesmäki,^{1,3,5} Olli Simell,^{3,7,8} and Riitta Lahesmaa^{1,3}



Innate Immune Activity Is Detected Prior to Seroconversion in Children With HLA-Conferred Type 1 Diabetes Susceptibility

Diabetes 2014;63:2402–2414 | DOI: 10.2337/db13-1775

The insult leading to autoantibody development in children who will progress to develop type 1 diabetes (T1D) has remained elusive. To investigate the genes and molecular pathways in the pathogenesis of this disease, we performed genome-wide transcriptomics analysis on a unique series of prospective whole-blood RNA samples from at-risk children collected in the Finnish Type 1 Diabetes Prediction and Prevention study. We studied 28 autoantibody-positive children, out of which 22 progressed to clinical disease. Collectively, the samples covered the time span from before the development of autoantibodies (seroconversion) through the diagnosis of diabetes. Healthy control subjects matched for date and place of birth, sex, and *HLA-DQB1* susceptibility were selected for each case. Additionally, we genotyped the study subjects with Immunochip to identify potential genetic variants associated with the observed transcriptional

signatures. Genes and pathways related to innate immunity functions, such as the type 1 interferon (IFN) response, were active, and IFN response factors were identified as central mediators of the IFN-related transcriptional changes. Importantly, this signature was detected already before the T1D-associated autoantibodies were detected. Together, these data provide a unique resource for new hypotheses explaining T1D biology.

By the time of clinical diagnosis of type 1 diabetes (T1D), an ~80–90% of the insulin-producing pancreatic β -cells have been destroyed by autoimmune mechanisms with immune cells infiltrating the islets. Autoantibodies against β -cell-specific antigens, such as proinsulin and insulin, GAD65, islet antigen-2, and zinc transporter 8,

¹Turku Centre for Biotechnology, University of Turku and Åbo Akademi University, Turku, Finland

²Turku Doctoral Programme of Biomedical Sciences, Turku, Finland

³The Finnish Centre of Excellence in Molecular Systems Immunology and Physiology Research, Helsinki, Finland

⁴Biomathematics Research Group, Department of Mathematics, University of Turku, Turku, Finland

⁵Department of Information and Computer Science, Aalto University School of Science, Aalto, Finland

⁶The National Graduate School in Informational and Structural Biology, Turku, Finland

⁷Department of Pediatrics, University of Turku, Turku, Finland

⁸Department of Pediatrics, Turku University Hospital, Turku, Finland

⁹Department of Genetics, University of Groningen, University Medical Centre, Groningen, the Netherlands

¹⁰Department of Signal Processing, Tampere University of Technology, Tampere, Finland

¹¹Department of Virology, University of Tampere, Tampere, Finland

¹²Fimlab Laboratories, Pirkanmaa Hospital District, Tampere, Finland

¹³Immunogenetics Laboratory, University of Turku, Turku, Finland

¹⁴Department of Clinical Microbiology, University of Eastern Finland, Kuopio, Finland

¹⁵Department of Pediatrics, University of Oulu and Oulu University Hospital, Oulu, Finland

¹⁶Children's Hospital, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland

¹⁷Folkhälsan Research Center, Helsinki, Finland

¹⁸Department of Pediatrics, Tampere University Hospital, Tampere, Finland

Corresponding author: Riitta Lahesmaa, riitta.lahesmaa@btk.fi.

Received 22 November 2013 and accepted 13 February 2014.

This article contains Supplementary Data online at <http://diabetes.diabetesjournals.org/lookup/suppl/doi:10.2337/db13-1775/-/DC1>.

H.K., L.L.E., and E.L. contributed equally to this study.

© 2014 by the American Diabetes Association. See <http://creativecommons.org/licenses/by-nc-nd/3.0/> for details.

See accompanying articles, pp. 2203 and 2538.

can usually be detected in serum before the clinical onset of the disease, but although the number of circulating autoantibodies is predictive for the development of T1D (1), predicting the onset of clinical disease remains a challenge.

Susceptibility to T1D is inherited, with the human leukocyte antigen (HLA) class II haplotypes encoding DR3-DQ2 (*DRB1*03-DQA1*0501-DQB1*0201*) and DR4-DQ8 (*DRB1*0401-DQA1*0301-DQB1*0302*) molecules conferring an ~50% of the total genetic risk. In addition, >40 non-HLA genes with more modest effect on disease risk, such as *INS*, cytotoxic T-lymphocyte-associated protein 4, protein-tyrosine phosphatase, nonreceptor-type 22, and interleukin-2 receptor, have been discovered. When combined with a susceptible genetic background, largely unknown random environmental factors probably contribute to the breakdown of self-tolerance. For instance, dietary factors, such as early exposure to foreign complex proteins (2) and vitamin D deficiency, have been associated with development of T1D, and viral infections are believed to trigger autoimmunity (3,4). Prevalence of T1D is also associated with Westernized lifestyle and level of hygiene (5).

To improve our understanding of the autoimmune processes driving the pathogenesis in the prediabetic phase in humans, we performed genome-wide transcriptomics with two different array formats to determine gene expression changes that correlate with the development of diabetes-associated autoantibodies and progression toward clinical diabetes. We analyzed a unique series

of prospective whole-blood RNA samples collected in the Finnish Type 1 Diabetes Prediction and Prevention (DIPP) study in which HLA-susceptible individuals are observed at frequent intervals from birth (6). For this study, we analyzed blood samples taken from 28 at-risk children that collectively covered the time span from before the appearance of autoantibodies (seroconversion) through the diagnosis of T1D (Fig. 1). Healthy, persistently autoantibody-negative children matched for date and place of birth, sex, and the *HLA-DQB1*-conferred genetic risk were chosen as control subjects (Supplementary Table 1). Additionally, we genotyped the study individuals using the ImmunoChip SNP array (7) to identify potential genetic variants associated with the observed transcriptional signatures.

RESEARCH DESIGN AND METHODS

Subjects and Sample Collection

All children studied were participants in the DIPP study (6), in which venous blood was drawn into PAXgene Blood RNA tubes (PreAnalytix) one to four times a year at the study clinic in Turku, Finland. The samples were incubated for 2 h at room temperature and then stored at -70°C until analyzed. Islet cell autoantibodies, insulin autoantibodies, and autoantibodies to GAD, islet antigen-2, and zinc transporter 8 were measured from all individuals. Of 28 subjects in this study, 9 were sampled starting before or at the time of the appearance of autoantibodies (seroconversion) (S1–5 and S7–10, Table 1); 6 case

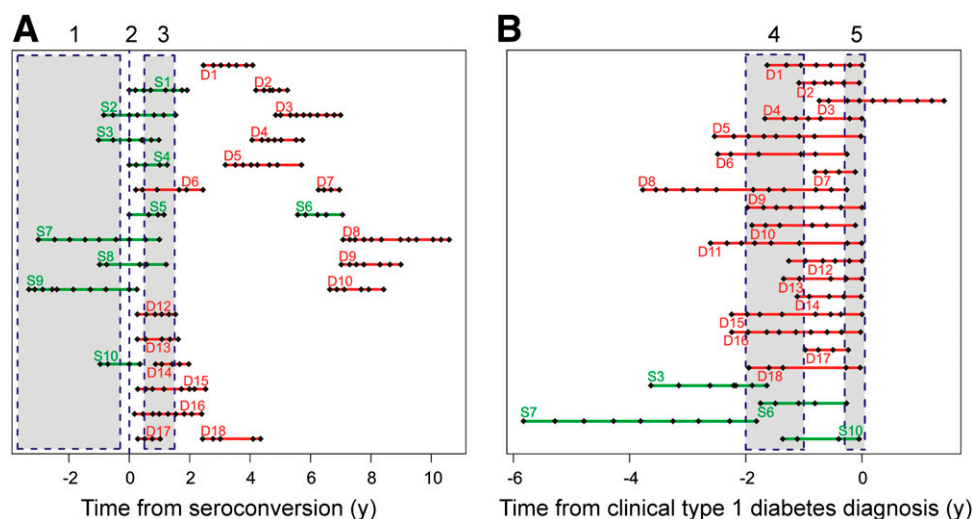


Figure 1—Timing of blood sample collection showing time windows used for temporal analysis of the data. A: Sample series of seroconverted children (S, marked in green) and T1D progressors (D, in red) relative to the time from seroconversion (in years). Each black diamond represents a blood sample that was subsequently analyzed by transcriptional profiling. For the analysis of temporal changes in gene expression, samples representing three time windows relative to seroconversion were selected as follows: 1, before seroconversion (22 samples from 6 case-control pairs); 2, at the time of the seroconversion (8 samples from 8 pairs); and 3, 6–18 months after seroconversion (40 samples from 14 pairs). B: Sample series of seroconverted progressors (S, in green) and T1D progressors (D, in red) relative to the time from clinical diagnosis (in years). Samples representing two time windows relative to diagnosis were selected as follows: 4, 1 to 2 years before clinical diagnosis (48 samples from 19 pairs); and 5, at the time of clinical diagnosis (18 samples from 18 pairs). A and B: The time windows were based on the maximal representation of the samples. See Table 1 and Supplementary Table 1 for more details on study subjects.

subjects (S2, S3, S7–10) were sampled before seroconversion, and 4 case subjects (S3, S6, S7, and S10) progressed to T1D. The remaining 18 children (D1–18, Table 1) were all sampled starting after seroconversion, and all progressed to diabetes (D1–18). A persistently autoantibody-negative control child was matched with each case, based on the date and place of birth, sex, and *HLA-DQB1* genotype (Supplementary Table 1). In all, 356 blood samples (191 from seroconverters and T1D children and 165 from healthy control subjects) were processed for genome-wide transcriptional analysis.

Microarray Hybridizations

Total whole-blood RNA was extracted from the samples using PAXgene Blood RNA kit (Qiagen), and each sample was hybridized in duplicate both on Affymetrix and Illumina arrays according to the manufacturer's instructions. For Affymetrix, RNA was processed with the Ovation RNA amplification system v2 including the Ovation whole-blood reagent and hybridized to GeneChip Human Genome U219 array plate (Affymetrix). For Illumina arrays, RNA was processed with RiboAmp OA 1 round RNA amplification kit (case-control pairs S1–S6; Applied Biosystems/

Table 1—Summary of the two study groups: seroconverted children ($n = 10$) and T1D progressors ($n = 18$) studied with genome-wide transcriptomics

Case number	Number of longitudinal RNA samples	First sample from seroconversion (months)	Last sample from seroconversion (months)	Time of T1D diagnosis	Age at T1D diagnosis (years)
S1	8	0	23	NA	
S2	7	–10	19	NA	
S3	7	–12	12	20 months after last sample	7.7
S4	5	0	15	NA	
S5	4	0	14	NA	
S6	6	68	89	0 months after last sample	8.6
S7	9	–37	12	22 months after last sample	9.3
S8	6	–12	15	NA	
S9	10	–40	3	NA	
S10	4	–12	4	1 month after last sample	2.4
	Number of longitudinal RNA samples	First sample from seroconversion (months)	First sample from diagnosis (months)	Time of T1D diagnosis	Age at T1D diagnosis (years)
D1	7	17	–33	At last sample	6.6
D2	6	51	–13	0.6 months after last sample	6.6
D3	12	59	–9	20.5 months prior to last sample	6.9
D4	7	49	–20	At last sample	7.8
D5	8	39	–31	At last sample	8.3
D6	6	3	–30	3.2 months after last sample	8.8
D7	4	76	–10	1.4 months after last sample	9.2
D8	12	86	–46	3.2 months after last sample	12.1
D9	7	85	–24	At last sample	10
D10	6	81	–23	1.4 months after last sample	12.5
D11	8	NA	–32	At last sample	13
D12	6	3	–15	At last sample	2.3
D13	5	3	–16	At last sample	3
D14	5	11	–14	At last sample	3.2
D15	8	3	–27	At last sample	3.6
D16	9	2	–27	0.3 months after last sample	3.8
D17	4	3	–12	2.9 months after last sample	4.2
D18	5	30	–24	0.4 months after last sample	4.9

Seroconverted children (S1–S5 and S7–S10) were sampled starting before or at the time of the appearance of autoantibodies (seroconversion), except for S6, who was seropositive since the beginning of the follow-up. Case subjects S3, S6, S7, and S10 progressed to T1D, and the other six did not. The remaining 18 children (D1–18) were all sampled starting after seroconversion, and all progressed to diabetes (D1–18). See Supplementary Table 1 for more information on the case subjects and their matched control subjects. D, T1D progressors; NA, not applicable; S, seroconverted children.

Arcturus) or Ovation RNA amplification system v2 including the Ovation whole-blood reagent (case-control pairs S7–S10 and all T1D case-control pairs; NuGEN Technologies) and hybridized to Illumina Sentrix human WG6 v2 expression bead chips (seroconverted case-control pairs) or Illumina Human HT-12 Expression BeadChips, version 3 (T1D case-control pairs).

Microarray Data Processing and Statistical Analysis

The gene-expression microarray data were preprocessed using the RMA procedure for the Affymetrix data and the quantile normalization procedure for the Illumina data using R/Bioconductor (8). Differentially expressed genes were identified by comparing subjects with their matched healthy control subjects. Since the follow-up series is not fully synchronized in time between the different individuals, we first applied a similar approach as in Elo et al. (9) and Huang et al. (10) that avoids the need of a direct alignment. More specifically, for each individual, the expression intensity value of a particular probe/probe set x at each time point was given a z -score, defined as $z = (x - m)s^{-1}$, where m is the mean, and s is the SD calculated using all the time points in the matched control time series. Such z -scores penalize those probes/probe sets that have high variability in the control series. To identify significant up- or down-regulation across the individuals, the rank product algorithm was applied to casewise maximum/minimum z -scores (11). Genes with a false discovery rate (FDR) <0.05 , estimated as the percentage of false-positive predictions, were considered as differentially regulated. To focus on T1D- or seroconversion-specific changes, the same procedure was applied to the data after swapping the case and control pairs. Only those probes were retained that did not pass the criterion FDR <0.05 in the swapped analysis. The aim of the swapping procedure was to exclude from further analyses such non-diabetes-related changes that were similarly detected also in the control subjects. Finally, the findings from Affymetrix and Illumina arrays were compared, and genes showing concordant changes were regarded as final, validated findings (Supplementary Tables 1 and 3). Additionally, we required that the genes were determined as present in $>50\%$ of all individuals (at least in one sample per individual) using a two-component Gaussian mixture model for the Affymetrix data and detection P values for the Illumina data. Affymetrix and Illumina data were combined using the gene symbols provided by Ingenuity Pathway Analysis (IPA). To combine the data from the Illumina WG-6 v2 and HT-12 arrays, the probes between the array types were matched based on their sequence similarity. A total of 70% of the probes in Human WG-6 v2 array had remained completely identical to the corresponding probes in the Human HT-12 arrays. In addition, we mapped 6% of the probes by requiring a minimum sequence overlap of 25 of 50 consecutive bases (one mismatch allowed). The probes also had to target the same gene according to annotation. The

remaining 24% of the probes were excluded from the analyses, for which the WG-6 v2 and HT-12 data needed to be combined.

Pathway Analysis

The functional annotation tool Database for Annotation, Visualization and Integrated Discovery (DAVID) was used to identify enriched biological pathways (Biological Biochemical Image Database, BioCarta, Kyoto Encyclopedia of Genes and Genomes, Panther, and Reactome) among the regulated genes (12,13). Pathways with FDR $<20\%$ were reported (modified Fisher exact test).

Transcription Binding Motif Analysis

Overrepresented transcription factor binding sites in the promoter regions (± 2 kb around the transcription start site) of the regulated genes were identified using the transcription factor target sets in the Molecular Signatures Database (14). Binding sites with a P value <0.05 were reported (hypergeometric test).

Detecting Differential Expression in Time Windows

In order to identify transcripts active in different phases of the autoimmune process, the rank product method was applied to find differentially expressed genes between the case subjects and their control subjects in five time windows (Fig. 1). The division was based on the maximal sample representation in each time window. Inside each time window, the fold change between the case and the matched control was calculated using linear inter-/extrapolation. More specifically, the control sample series was matched to the time points of the case sample series. For the time points that were inside the range of real control time points, this was performed by linear interpolation. For the time points needed outside this range, the expression values were approximated by constant extrapolation (set equal to the closest real measurement). Genes with FDR <0.05 were considered differentially regulated. Additionally, we required that the genes were determined as present, as described above. The findings validated with all array types (as described above) were regarded as changed.

In addition, we did a comparison between those seroconverted children (case subjects S3, S7, and S10) who had later progressed to T1D and those who have not been diagnosed with T1D to date. An unpaired two-group rank product analysis was performed for the case-control ratios to compare the affected and unaffected case subjects. The analysis was performed for three time windows: before seroconversion, at seroconversion, and 6–18 months after seroconversion.

Differentially Expressed Genes With Functions Related to the Innate Immunity Responses

The human genes with literature-annotated function in the innate immunity were downloaded from <http://www.innatedb.org> (15). The enrichment of these genes among

our differentially expressed genes was calculated using Fisher exact test.

Transcription Module Analysis

A module-based method modified from Chaussabel et al. (16) was used to survey coordinately expressed sets of genes (modules) functionally annotated with literature search. The gene symbols belonging to each module were downloaded from <http://www.biir.net/modules>, and differentially expressed genes not belonging to any of the modules presented in Chaussabel et al. (16) were excluded from the analysis. The significance of overlap of each of the 28 modules with each list of differentially expressed genes was calculated using Fisher exact test. Modules up- or downregulated with the Benjamini-Hochberg corrected P value <0.05 were visualized in the module maps with red or blue (up- and downregulation, respectively), with the intensity of the color corresponding to the percentage of the regulated probe sets that belong to the module.

ImmunoChip Data Processing and Statistical Analysis

DNA from whole blood was genotyped with the ImmunoChip array (7) in accordance with Illumina protocols at the Department of Genetics, University Medical Centre Groningen (Groningen, the Netherlands). Single nucleotide polymorphisms (SNPs) were mapped to National Center for Biotechnology Information build 36 (hg18). Quality control was applied using PLINK (17). Samples with a call rate $<95\%$, as well as SNPs with a call rate $<98\%$ were excluded. Markers were excluded when they deviated from Hardy-Weinberg equilibrium in the control subjects ($P < 0.0001$). SNPs with minor allele frequency $<10\%$ were removed ($N = 80,259$). Finally, we pruned our dataset based on linkage disequilibrium between markers ($r^2 > 0.8$). Our final dataset comprised 30,463 SNPs. *Cis*-acting expressed quantitative trait loci (*cis* eQTL) effects were calculated for each SNP residing within ± 250 kb around the gene coordinates of each differentially regulated gene. A linear model was fit for genotype-gene expression, and a P value was calculated for the null hypothesis that genotype has no effect on gene expression (slope = 0). Effects with Benjamini-Hochberg-corrected P value <0.05 are reported (Supplementary Table 1). The SNPs and their proxies ($r^2 > 0.8$) that had an eQTL effect on our differentially expressed genes were searched for associations with autoimmune diseases in genome-wide association studies, as listed in Ricaño-Ponce and Wijmenga (18) and/or T1Dbase (19). The proxies for these SNPs were found based on HapMap3 (release 2) and 1000 Genomes in the CEU population panel by using SNP annotation and proxy search (<http://www.broadinstitute.org/mpg/snap/ldsearch.php>).

Accession Codes

The gene expression data discussed in this publication are accessible through Gene Expression Omnibus SuperSeries accession number GSE30211 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE30211>).

RESULTS

Gene Expression Signatures Detected During Seroconversion to Autoantibody Positivity

When the transcriptional profiles of S1–10 (Table 1) were compared with their matched healthy control subjects, 109 unique genes were differentially expressed (FDR <0.05 ; Supplementary Table 1), with 66% of genes being upregulated and 34% being downregulated. Genes residing in the genomic T1D susceptibility loci [T1Dbase (19)] included the upregulated oncostatin M and the downregulated killer cell lectin-like receptor (KLR) subfamily B member 1 (Supplementary Fig. 1). The upregulated genes were enriched in the retinoic acid-inducible gene 1-like receptor signaling pathway involved in recognizing viral double-stranded RNA (Supplementary Table 1).

Altogether, 12 unique transcription factor genes were differentially regulated in the seroconverted children (molecular type = transcription regulator, Supplementary Table 1). Analysis of known gene and protein interactions among these transcription factors and other genes in the dataset using the IPA network neighborhood tool showed that interconnected interferon (IFN) regulatory factors 5 and 7 (*IRF5* and *IRF7*) had several coregulated network partners (Fig. 2B and Supplementary Fig. 1). Additionally, *IRF7* binding sites were enriched among the promoters of the upregulated genes ($P = 0.000309$; Fig. 2A and Supplementary Table 1). Both *IRF5* and *IRF7* are highly expressed in lymphoid tissues and drive the activation of a wide range of antiviral genes. Target genes in the identified network (Fig. 2B) included, for example, the DEAD box protein 58 (*DDX58*, better known as retinoic acid-inducible gene 1), which codes a helicase involved in viral double-stranded RNA recognition, as well as T1D-associated 2'-5'-oligoadenylate synthetase 1 (*OAS1*), which codes for an enzyme involved in activating RNase L, resulting in viral RNA degradation.

Gene Expression During the Autoantibody-Positive Period Preceding Clinical T1D

A comparison of the samples obtained from the 18 progressors (D1–18, Table 1) to their matched control subjects revealed 472 differentially expressed genes (FDR <0.05 ; Supplementary Table 1). Notably, gene expression was suppressed, as 73% of the differentially regulated genes in the progressors were downregulated. The upregulated genes residing in the genomic T1D susceptibility loci [T1Dbase (19)] included the tumor necrosis factor receptor-type zinc finger domain containing 1 and signal transducer and activator of transcription 2. In contrast, the diabetes-associated grancalcin and the von Hippel-Lindau binding protein 1 (20–22) were downregulated (Supplementary Fig. 1). The top downregulated pathways were related to immune system signaling, blood coagulation, and complement (Supplementary Table 1).

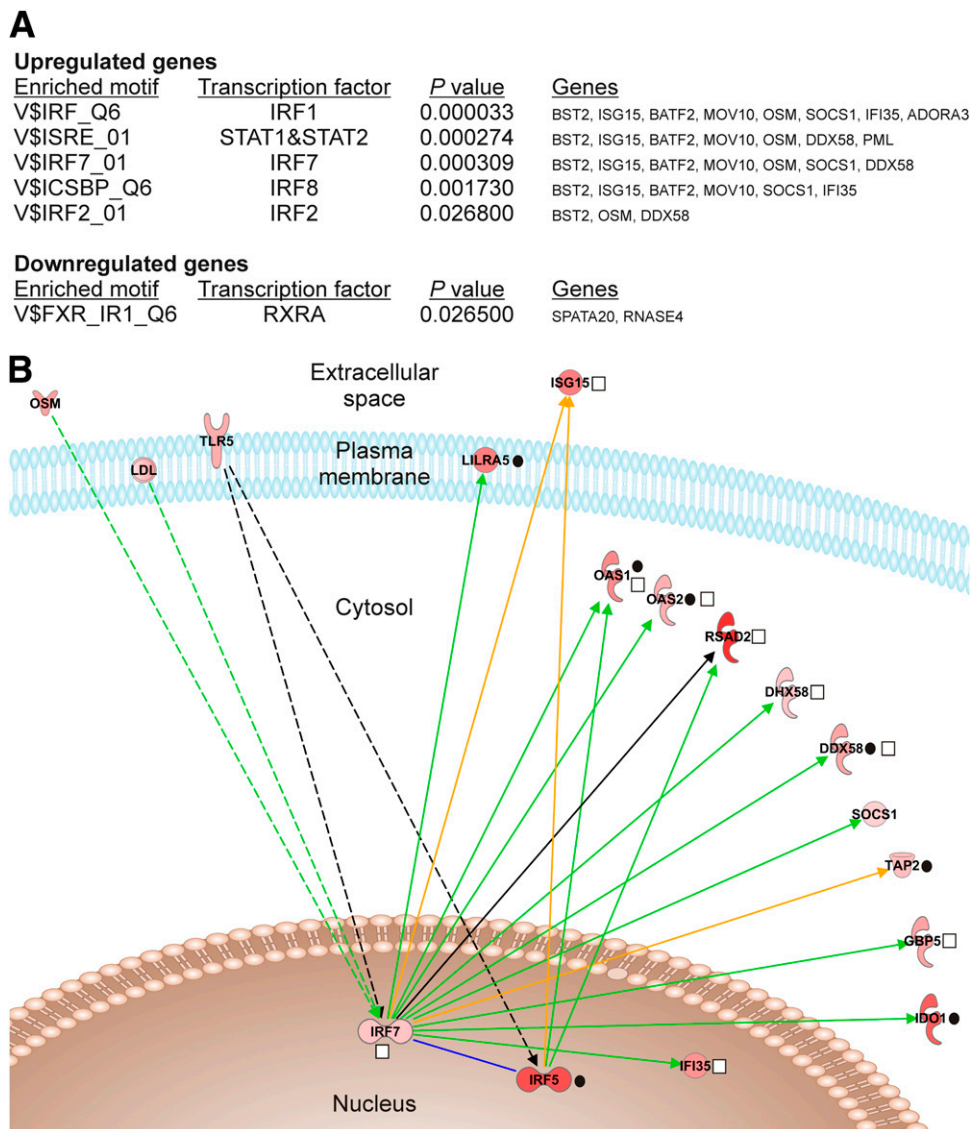


Figure 2—Transcriptional regulation during the pathogenesis of T1D. **A:** Transcription factor binding sites enriched among the promoters of the differentially up- and downregulated genes in the seroconverted individuals. Overrepresented sites in the ± 2 -kb region around the transcription start site of the regulated genes were identified using the transcription factor target sets in the Molecular Signatures Database (14). Binding sites with $P < 0.05$ and a known transcription factor are presented (hypergeometric test). See Supplementary Table 1 for more details. **B:** IRF5 and IRF7, the most highly regulated transcription factors in the seroconverted children, and their central role in the seroconversion-specific transcription. The network was generated with the IPA network neighborhood tool, and it shows all of the IRF5- and -7-connected genes (nodes) that were differentially regulated in the seroconverters ($FDR < 0.05$). The genes are organized according to their cellular localization, and red color indicates upregulation of the gene in the seroconverted case subjects compared with the matched control subjects. The genes regulating IRF5 or -7 are presented on the *left*, and the downstream target genes of IRF5 or -7 on the *right*. Solid edge indicates direct interaction and dashed edge indirect interaction. Transcriptional induction/reduction is marked with green color, promoter binding with orange, and mixed interactions with black. Black circles represent genes having a *cis* eQTL effect with SNPs residing in a window of ± 250 kb around the gene coordinates (Supplementary Table 1); white squares indicate genes belonging to the T1D-associated human IRF7-mediated immune response network genetically identified by Heinig et al. (44). RXRA, retinoid X receptor, alpha; STAT, signal transducer and activator of transcription.

Identification of Transcriptional Signatures Characteristic for Distinct Phases of the Diabetic Disease Process

To identify transcripts that are active in different phases of the diabetic disease, we divided and analyzed the data separately within five time windows (Fig. 1). The number of samples available in each time window is summarized

in Fig. 3. In the time windows before, at, or 6–18 months after seroconversion, 124, 33, and 551 genes were identified as differentially expressed ($FDR < 0.05$), respectively; only 14 genes were identified that were common to all three time windows (Fig. 3A and Supplementary Table 1). During the interval of 1 to 2 years before clinical diagnosis and at clinical diagnosis, 388 and 211 genes,

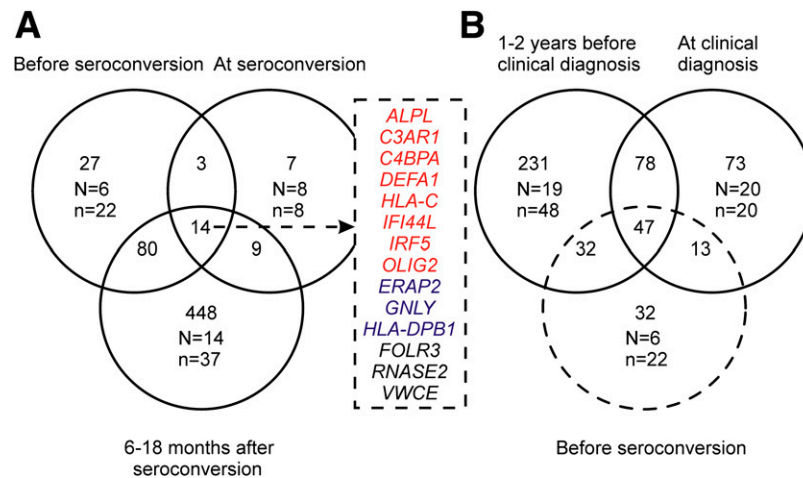


Figure 3—Transcriptional changes in distinct phases of the diabetic autoimmune process. Numbers of regulated genes (FDR <0.05; Supplementary Table 1) are listed for each time window in the analysis. **A:** Time windows before seroconversion, at seroconversion, and 6–18 months after seroconversion. The overlapping 14 genes are listed on the right. Red, upregulated; blue, downregulated; and black, mixed regulation between the time windows. **B:** Time windows 1 to 2 years before clinical T1D and at clinical T1D diagnosis and their overlap. The time window before seroconversion is presented with a dotted circle. *N*, number of case-control pairs in the analysis; *n*, number of samples in the analysis.

respectively, were differentially expressed between the progressors and control subjects, with 125 common genes (Fig. 3B and Supplementary Table 1). The two latest time windows overlapped with 47 genes that were differentially expressed before seroconversion (Fig. 3B). Different time windows showed enrichment of genes bearing different transcription factor binding sites; for example, IRF and/or signal transducer and activator of transcription binding sites were enriched in genes upregulated before seroconversion and in the interval between seroconversion and diagnosis (Supplementary Table 1). The overall suppression of the ribosomal proteins was presented as downregulated pathways named ribosome, metabolism of proteins, and 3' untranslated region (UTR)-mediated translational regulation (Table 2 and Supplementary Table 1). In addition, by using a whole-blood sample data analysis tool (16) that searches for presence of pre-defined groups of coordinately expressed transcripts (functionally related, literature-annotated modules), we observed that the changes in transcriptional signatures were not stable but dynamic relative to different stages of T1D pathogenesis (Fig. 4 and Supplementary Table 1). For example, the cytotoxic cell module (M2.1) was first downregulated at the earliest time window, but then induced at the time of clinical diagnosis. This module included, for example, *KLRD1* and *KLRF1* that are expressed on natural killer cells, as well as granulysin that is expressed by both cytotoxic T cells and natural killer cells. However, a module consisting of IFN-induced transcripts (M3.1), including, for example, the above-mentioned antiviral *OAS1* and *DDX58*, was constantly showing changes at almost every stage of the development of T1D.

Transcriptional Signatures Distinct for Progressors Among the Seroconverters

Finally, we performed a separate time-window analysis for the seroconverted children who later progressed to clinical T1D (case subjects S3, S6, S7, and S10; Table 1) compared with seroconverted children who have remained nondiabetic for at least 76 months. For case subjects S3, S7, and S10, data before, at, and 6–18 months after seroconversion were available.

Before seroconversion, nine genes were identified as differentially regulated (FDR <0.05; Supplementary Table 1). At seroconversion and 6–18 months after seroconversion, the difference had increased to 13 and 54 genes, respectively. One of the genes showing constant and high upregulation in the progressors was ribonuclease, RNase A family, 2 (*RNASE2*; also known as eosinophil-derived neurotoxin; Fig. 5). It encodes a secreted protein with several immunomodulatory functions. The ribonucleolytic activity of *RNASE2* plays an important role in eosinophil-mediated antiviral activity against single-stranded RNA viruses. *RNASE2* is an endogenous Toll-like receptor 2 (TLR2) ligand (23) that could play a role in the induction of IFNs. Most interestingly, *RNASE2* expression has been reported to be upregulated in the peripheral blood mononuclear cells of patients with autoimmune diseases, such as rheumatoid arthritis (24) and systemic lupus erythematosus (SLE) (25).

Identification of the Genomic Variants Affecting the Expression Levels

In order to study the effect of possible SNPs as the cause for the detected expression differences, we used the ImmunoChip SNP array (Illumina) (7) for genotyping. Using representative (linkage disequilibrium-pruned) SNP

Table 2—Summary of the DAVID pathway analysis results in the five time windows

Time window	Pathway term (DAVID)	FDR (%)	P value	Genes
Before seroconversion	<i>hsa03010:Ribosome</i>	8.84	0.011911	<i>RPL26, RPS27L, RSL24D1</i>
At seroconversion	No pathways			
6–18 months after seroconversion	hsa04610:Complement and coagulation cascades	0.49	0.000477	C3AR1, F5, SERPING1, C4BPA, C2, CFD
	<i>hsa03010:Ribosome</i>	0.03	0.000024	<i>RPL17, MRPL13, RPL7, RPL31, RPS3A, RPL9, RPL26, RPS15A, RPS27L, RPL39, RPS23</i>
	<i>REACT_17015:Metabolism of proteins</i>	0.19	0.000210	<i>RPL17, RPL26, VBP1, RPS15A, RPL39, PIGK, PIGF, RPL7, RPS3A, RPL31, RPL9, PFDN5, RPS14, DPM1, RPS23, PIGA</i>
	<i>REACT_1762:3'-UTR-mediated translational regulation</i>	1.06	0.001189	<i>RPL17, RPL7, RPL31, RPS3A, RPL9, RPS14, RPL26, RPS15A, RPL39, RPS23</i>
	<i>h_DNAfragmentPathway: Apoptotic DNA fragmentation and tissue homeostasis</i>	1.48	0.001414	<i>HMGB1, HMGB2, CASP3, TOP2A</i>
	<i>REACT_6167:Influenza infection</i>	4.21	0.004784	<i>RPL17, RPL7, RPL31, RPS3A, RPL9, RPS14, RPL26, RPS15A, RPL39, RPS23, SNRPG</i>
	<i>hsa05012:Parkinson's disease</i>	9.00	0.008457	<i>NDUFA4, CASP3, COX7A2, NDUFS4, UBE2J1, COX7C, NDUFA1, UQCRQ, COX6C</i>
1 to 2 years before T1D diagnosis	hsa05332:Graft-versus-host disease	0.73	0.000710	PRF1, HLA-A, HLA-C, KIR2DL3, HLA-DQA1
	hsa05330:Allograft rejection	6.36	0.006311	PRF1, HLA-A, HLA-C, HLA-DQA1
	REACT_6900:Signaling in immune system	5.48	0.007180	SELP, LY96, HLA-A, HLA-C, TLR5, C2, KIR2DL3, PROS1, HLA-DQA1
	hsa04940:Type I diabetes mellitus	9.63	0.009703	PRF1, HLA-A, HLA-C, HLA-DQA1
	<i>hsa03010:Ribosome</i>	0.00	0.000002	<i>RPL36A, MRPL13, RPL31, RPS3A, RPL9, RPL26, RPS15A, RPS27L, RPL39, RPS23</i>
	<i>REACT_1762:3'-UTR-mediated translational regulation</i>	0.44	0.000519	<i>RPL36A, RPL31, RPS3A, RPL9, RPL26, RPS15A, RPL39, RPS23</i>
	<i>REACT_6167:Influenza infection</i>	0.85	0.001011	<i>RPL36A, RPL31, RPS3A, RPL9, RPL26, RPS15A, RPL39, EIF2AK2, RPS23</i>
	<i>REACT_17015:Metabolism of proteins</i>	1.73	0.002063	<i>RPL36A, RPL31, RPS3A, RPL9, RPL26, RPS15A, RPL39, PROS1, RPS23, PIGA</i>
	<i>REACT_6900:Signaling in immune system</i>	3.36	0.004040	<i>THBD, LY96, CD160, C1R, TREM1, KIR2DL3, CFD, KLRD1, PROS1, HLA-DOB, TLR8</i>
At T1D diagnosis	hsa05332:Graft-versus-host disease	0.16	0.000178	HLA-A, GZMB, HLA-C, KLRD1
	hsa05330:Allograft rejection	4.25	0.004829	HLA-A, GZMB, HLA-C
	hsa04650:Natural killer cell mediated cytotoxicity	5.53	0.006321	HLA-A, GZMB, HLA-C, KLRD1
	hsa04940:Type I diabetes mellitus	5.72	0.006533	HLA-A, GZMB, HLA-C
	hsa05320:Autoimmune thyroid disease	8.24	0.009527	HLA-A, GZMB, HLA-C

Continued on p. 2410

Table 2—Continued

Time window	Pathway term (DAVID)	FDR (%)	P value	Genes
	<i>hsa03010:Ribosome</i>	0.03	0.000027	<i>RPS26, RPL31, RPL9, RPL26, RPS15A, RSL24D1, RPL39, RPS23</i>
	<i>hsa05310:Asthma</i>	0.21	0.000206	<i>FCER1A, HLA-DRB4, MS4A2, HLA-DOB, HLA-DQA1</i>
	<i>hsa05330:Allograft rejection</i>	0.48	0.000484	<i>HLA-A, HLA-DRB4, HLA-C, HLA-DOB, HLA-DQA1</i>
	<i>REACT_1762:3'-UTR-mediated translational regulation</i>	0.37	0.000493	<i>RPS26, RPL31, RPL9, RPL26, RPS15A, RPL39, RPS23</i>
	<i>hsa05332:Graft-versus-host disease</i>	0.66	0.000661	<i>HLA-A, HLA-DRB4, HLA-C, HLA-DOB, HLA-DQA1</i>
	<i>hsa04940:Type I diabetes mellitus</i>	0.87	0.000878	<i>HLA-A, HLA-DRB4, HLA-C, HLA-DOB, HLA-DQA1</i>
	<i>hsa05320:Autoimmune thyroid disease</i>	1.81	0.001830	<i>HLA-A, HLA-DRB4, HLA-C, HLA-DOB, HLA-DQA1</i>
	<i>REACT_6167:Influenza infection</i>	2.61	0.003467	<i>RPS26, RPL31, RPL9, RPL26, RPS15A, RPL39, RPS23</i>
	<i>hsa05416:Viral myocarditis</i>	5.93	0.006104	<i>HLA-A, HLA-DRB4, HLA-C, HLA-DOB, HLA-DQA1</i>

Pathways with FDR <10% are presented. See Supplementary Table 4 for further details. Boldface text indicates pathways enriched among the upregulated genes; italicized text indicates pathways enriched among the downregulated genes. COX, cytochrome c oxidase; GZMB, granzyme B; RP, ribosomal protein; VBP, von Hippel-Lindau binding protein.

markers residing inside a window of ± 250 kb around the gene coordinates, 118 differentially expressed genes had a *cis* eQTL effect with 1–54 SNPs per gene (FDR <0.05; Supplementary Table 1). Comparison with eQTLs recently identified in whole blood from 5,311 individuals (26) validated 27% of the detected *cis* effects (Supplementary Table 1). Fourteen of the identified eQTL genes (the identified SNPs or their proxies; $r^2 > 0.8$), such as histone cluster 1 H2bd, have also been associated with autoimmune diseases in genome-wide association studies, as listed in Ricaño-Ponce and Wijmenga (18) and/or T1Dbase (19) (Supplementary Table 1). Interestingly, eQTL effects were found with SNPs adjacent to *IRF5*, *OAS1*, *OAS2*, *DDX58*, transporter 2 ATP-binding cassette subfamily B, indoleamine 2,3-dioxygenase 1, and leukocyte immunoglobulin-like receptor subfamily A (with TM domain) member 5, which were identified as the core of the central IFN stimulatory network in the seroconverted individuals (Fig. 2B). Previously, SNPs overlapping the transporter 2 ATP-binding cassette subfamily B gene residing in the human major histocompatibility complex locus have been associated with T1D in several studies (27,28), and the *IRF5* polymorphisms have been connected to SLE, rheumatoid arthritis, and inflammatory bowel disease (29–31).

DISCUSSION

This study provides a comprehensive analysis of transcriptional changes that occur over time in children at risk for or developing T1D and provide a unique resource for new hypotheses explaining the disease biology. Our results demonstrate that diabetes-specific changes are

evident in the whole-blood transcriptome, affecting hundreds of genes. As the gene expression analysis was performed on whole-blood samples, we exploited the module-based tool to suggest which cell types are important in different phases of the pathogenesis of T1D (Fig. 4 and Supplementary Table 1). For example, the neutrophil module, including granule proteins such as lipocalin 2 and matrix metalloprotease 9, was first activated, after which the neutrophil-associated transcripts were suppressed before diagnosis, consistent with a recent report on reduced numbers of circulating neutrophils during the autoantibody-positive phase as well as at the diagnosis of T1D in humans (32). Also, the erythrocyte-related module was activated at early time windows and before diagnosis. The platelet module, including coagulation factor XIII A1 polypeptide, was activated after seroconversion until clinical disease, which is in line with the reports on platelet hyperreactivity in T1D (33). Interestingly, the cytotoxic cell module was suppressed before seroconversion but highly activated at the time of diagnosis, indicating the differential cytotoxic cell activity in the blood early versus late in the disease process. However, as detailed recording of the proportions of different cell populations at the time of the blood sample draw is not available, these conclusions based on heterogeneous whole-blood samples should be further explored in future studies.

Taken together, our results indicate for the first time that type 1 IFN-mediated innate immune system is activated even prior to seroconversion and throughout the development of T1D. The induction is temporal, as expected for this systemic response, but some individuals show several signature peaks during the follow-up period

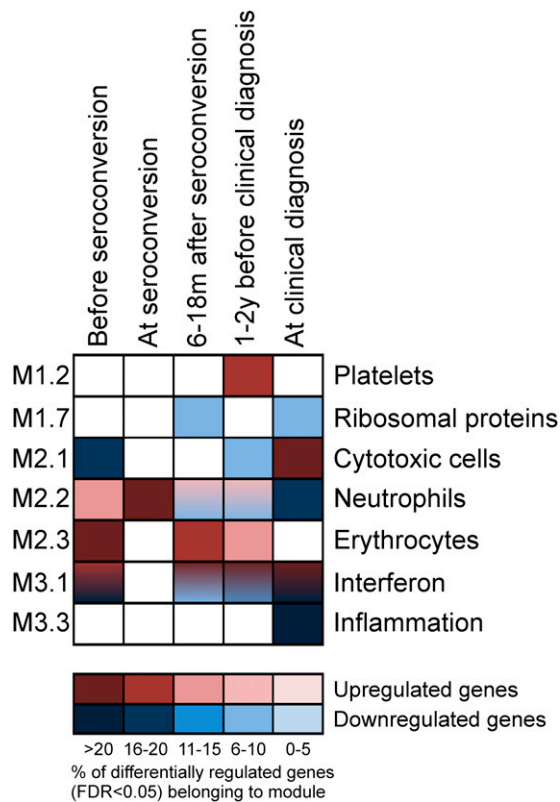


Figure 4—Summary of the functional modules regulated in the distinct phases of development of T1D. Upregulation or downregulation of the module (annotated groups of coordinately expressed transcripts from Chaussabel et al. [16]) is marked with red or blue, respectively, and the color intensity indicates the percentage of the regulated genes belonging to that module. Module maps and genes belonging to each module in each time window are presented in Supplementary Table 1.

(Supplementary Fig. 1). Importantly, our findings are consistent with the results of the accompanying study by Ferreira et al. (34), conducted in an independent birth cohort of German children genetically susceptible for T1D. They also report that type 1 IFN signature is temporally increased in susceptible children prior to the development of autoantibodies. These two independent longitudinal studies provide a unique insight into the preclinical stage of T1D.

In agreement with our results, Reynier et al. (35) recently reported that 12 IFN response genes are upregulated in the whole blood of 30% of autoantibody-positive prediabetic children, but not in healthy control subjects or recently diagnosed patients. Our data indicate that the IFN response can be detected before the appearance of autoantibodies (Fig. 4) and that it consists of a much larger number of genes than previously reported. In fact, comparison of the differentially regulated genes to the genes listed to be involved in functions related to human innate immune responses in the InnateDB (15) provides us with 17.4% overlap for the seroconverted children (19 genes; $P = 0.0000003$) and 9.5% overlap

with the progressors (45 genes; $P = 0.0000013$), respectively (Supplementary Tables 1 and 3, InnateDB columns). IFN- α and - β were not differentially regulated in our data, indicating that the blood cells are showing an indirect response to cytokines produced elsewhere. For instance, enteroviral infections have been detected in the β -cells of newly diagnosed patients (36), and signs of infections have been observed a few months prior to seroconversion to autoantibody positivity (37) and during the 6-month period preceding seroconversion (3,4). Moreover, the accompanying study (34) demonstrates an association between upper respiratory tract infections and upregulation of IFN-responsive genes. The correlation of gene expression profiles to markers of virus infections in susceptible children could make it possible to explore the mechanisms leading to the activation of the IFN system.

In addition, bacterial DNA, lipopolysaccharide, and flagellin are efficient triggers of TLR signaling and IFN response, providing an intriguing link between the gut immune system and autoimmunity. Both preclinical and established T1D patients exhibit increased gut permeability (38,39), leading to enhanced immune responses in the intestinal tissues (40). Also, a preliminary study has shown that gut microbiome diversity is reduced in children progressing toward T1D compared with healthy control subjects (41). Moreover, the IFN response can be activated without infection by bacteria or viruses; for example, in SLE, IFN- α production is driven by TLR7-mediated signaling and is induced by autoantibody-protein-RNA complexes derived from apoptotic cells. Also, endogenous TLR ligands such as RNASE2, which was identified to be upregulated in the seroconverters who later presented with T1D, could play a role as it has been shown that spontaneous IFN production drives autoimmune diabetes in NOD mice (42).

In this study, the matching of the case and control subjects was based on the *HLA-DQB1* genotype. To take into account further susceptibility alleles and alterations affecting the transcriptional changes, Immunochip SNP detection was performed for combinatorial analysis of the genome and transcriptome. This revealed that ~10% of the genes differentially expressed between the case and control subjects were affected in *cis* by the genetic variation. The affected variants and genes have previously been linked to several autoimmune diseases, highlighting the shared genetic basis for these disorders. For example, in the case of SLE, in which a strong type 1 IFN signature is also observed, the associated variations in *IRF5* have been shown to control cytokine responses upon TLR ligation (43). This pathway could be also genetically linked to T1D, since only a small part of the genetic factors conferring susceptibility to this disease have been identified. In support of our findings, Heinig et al. (44) performed an eQTL analysis across seven rat tissues combined with transcription factor binding site enrichment analysis, resulting in the identification of 23 *Irf7* target

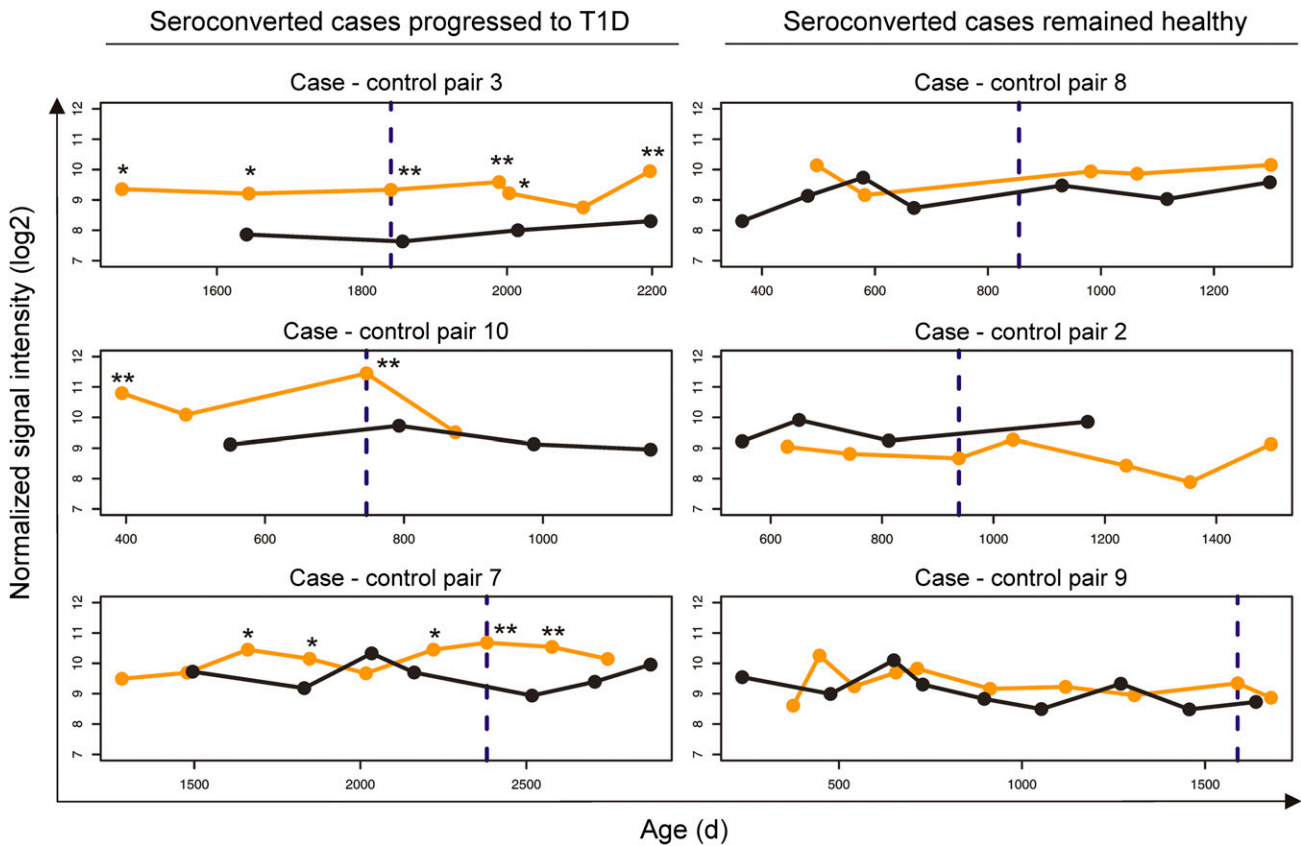


Figure 5—*RNASE2* gene is upregulated in each seroconverted case subject, who later progresses to clinical T1D. Longitudinal expression intensity profiles for *RNASE2* (probe set 11729772 on Affymetrix U219 array) in seroconverted children (in orange) and their matched autoantibody-negative control subjects (in black) are presented. Seroconverted children 3, 7, and 10 later progressed to clinical T1D (20, 22, and 1 month after the last follow-up sample, respectively), whereas seroconverted children 2, 8, and 9 have remained healthy to date. The x-axis represents the age of the children (in days) and the y-axis the normalized logarithmic signals on the arrays. The time of seroconversion for the progressors is marked with a dashed line. Profiles for children with several measurements before seroconversion are presented (see Fig. 1, Table 1, and Supplementary Table 1 for more details). * $P < 0.05$, ** $P < 0.01$ between the case and matched control subjects (rank product test).

genes mapping to a single eQTL. They subsequently identified a human IRF7-driven network using genome-wide expression and SNP data collected from monocytes. Eventually, SNPs close to genes in the human and rat IRF7 network were found to be associated with T1D. Most importantly, of the 12 direct IRF7 target genes (identified by the IPA tool) detected as differentially expressed in our study (Fig. 2B), 8 (66%) are present in the human IRF7-regulated network identified by Heinig et al. (44) through a completely different complementary approach, further highlighting the role of this pathway in T1D pathogenesis.

Acknowledgments. The authors thank the DIPP families for participating; S. Simell and the staff of the DIPP study for working with the families and obtaining the samples for the study; J. Hakalax, M. Laaksonen, and E. Pakarinen for handling and managing the data from the study participants; V. Simell for managing the DIPP biobank and samples; J. Tietäväinen for help in the data analysis; M. Vähä-Mäkilä at the Department of Pediatrics, University of Turku, for help in finalizing the manuscript; P. Junni, R. Venho, E. Virtanen, and K. Waenerberg for skillful technical assistance and support in microarray hybridizations at the Finnish Microarray and Sequencing Centre, Turku, Finland;

M. Karlsson, T. Laakso, P. Nurmi, R. Suominen, and V. Öling for help with the DNA isolation; S. Sharma, E. Rothenberg and K.V.S. Rao for fruitful discussion; and A. Rao for critical review of the manuscript.

Funding. This work was supported by the Academy of Finland (the Centre of Excellence in Molecular Systems Immunology and Physiology Research, 2012-2017, decision number 250114 to R.L., O.S., M.K., and H.L. and grants 77773, 203725, 207490, 116639, 115939, 123864, 126063, and 127575 to L.L.E. and 110432 to J.M.), European Commission Seventh Framework grants (EC-FP7-SYB-ILLA-201106, EC-FP7-NANOMMUNE-214281, EC-FP7-DIABIMMUNE-202063, EC-FP7-PEVNET-261441, and EC-FP7-DIAPREPP-202013), the Sigrid Jusélius Foundation, the Juvenile Diabetes Research Foundation, Turku University Hospital Research Fund, and the Oskar Öflund Foundation.

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

Author Contributions. H.K. analyzed the data, participated in the interpretation of the results, prepared the figures, and was responsible for drafting and further writing of the manuscript. L.L.E. analyzed the data, participated in the interpretation of the results, supervised E.L. and T.D.L., prepared the figures, and contributed to the writing of the manuscript. E.L. analyzed the data, participated in the interpretation of the results, prepared the figures, and contributed to the writing of the manuscript. J.M. participated in the initiation of the study, managed the RNA sample biobank, designed the study setup, and participated in the data analysis, interpretation of the results, and writing of the manuscript. I.R.-P. and

C.W. were responsible for the genotyping and SNP data quality control. M.V. and T.D.L. analyzed the data. H.H. was responsible for the virus analysis within the study. J.I. was responsible for the DNA isolation and HLA screening of the study children. R.V. and M.K. were responsible for the analyses of diabetes-associated autoantibodies. T.S. provided the samples and the clinical information on the study children. H.L. supervised E.L. and M.V. and participated in the interpretation of the results. O.S. provided the samples and the clinical information on the study children, initiated the study, designed the study setup, supervised the study, and participated in the interpretation of the results and writing the manuscript. R.L. initiated the study, designed the study setup, supervised the study, and participated in the interpretation of the results and writing the manuscript. All authors contributed to the final version of the manuscript. R.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.

Prior Presentation. These data were presented in poster form at the 20th Annual BioCity Symposium, Turku, Finland, 19–20 August 2010; 21st Annual BioCity Symposium, Turku, Finland, 18–19 August 2011; and 23rd Annual BioCity Symposium, Turku, Finland, 15–16 August 2013; and in poster and oral form at the 41st Scandinavian Society for Immunology Meeting, Copenhagen, Denmark, 14–17 April 2013.

References

- Ziegler AG, Rewers M, Simell O, et al. Seroconversion to multiple islet autoantibodies and risk of progression to diabetes in children. *JAMA* 2013;309:2473–2479
- 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
- Oikarinen S, Martiskainen M, Tauriainen S, et al. Enterovirus RNA in blood is linked to the development of type 1 diabetes. *Diabetes* 2011;60:276–279
- Laitinen OH, Honkanen H, Pakkanen O, et al. Coxsackievirus B1 is associated with induction of beta-cell autoimmunity that portends type 1 diabetes. *Diabetes* 2014;63:446–455
- Patterson CC, Dahlquist GG, Gyürüs E, Green A, Soltész G; EURODIAB Study Group. Incidence trends for childhood type 1 diabetes in Europe during 1989–2003 and predicted new cases 2005–20: a multicentre prospective registration study. *Lancet* 2009;373:2027–2033
- Kupila A, Muona P, Simell T, et al.; Juvenile Diabetes Research Foundation Centre for the Prevention of Type 1 Diabetes in Finland. Feasibility of genetic and immunological prediction of type 1 diabetes in a population-based birth cohort. *Diabetologia* 2001;44:290–297
- Cortes A, Brown MA. Promise and pitfalls of the Immunochip. *Arthritis Res Ther* 2011;13:101
- Gentleman RC, Carey VJ, Bates DM, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 2004;5:R80
- Elo LL, Mykkänen J, Nikula T, et al. Early suppression of immune response pathways characterizes children with prediabetes in genome-wide gene expression profiling. *J Autoimmun* 2010;35:70–76
- Huang Q, Liu D, Majewski P, et al. The plasticity of dendritic cell responses to pathogens and their components. *Science* 2001;294:870–875
- Breitling R, Armengaud P, Amtmann A, Herzyk P. Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett* 2004;573:83–92
- Huang W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009;4:44–57
- Huang W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 2009;37:1–13
- Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;102:15545–15550
- Breuer K, Foroushani AK, Laird MR, et al. InnateDB: systems biology of innate immunity and beyond—recent updates and continuing curation. *Nucleic Acids Res* 2013;41(Database issue):D1228–D1233
- Chaussabel D, Quinn C, Shen J, et al. A modular analysis framework for blood genomics studies: application to systemic lupus erythematosus. *Immunity* 2008;29:150–164
- Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559–575
- Ricaño-Ponce I, Wijmenga C. Mapping of immune-mediated disease genes. *Annu Rev Genomics Hum Genet* 2013;14:325–353
- Burren OS, Adlem EC, Achuthan P, Christensen M, Coulson RM, Todd JA. T1DBase: update 2011, organization and presentation of large-scale data sets for type 1 diabetes research. *Nucleic Acids Res* 2011;39(Database issue):D997–D1001
- Cooper JD, Smyth DJ, Smiles AM, et al. Meta-analysis of genome-wide association study data identifies additional type 1 diabetes risk loci. *Nat Genet* 2008;40:1399–1401
- Barrett JC, Clayton DG, Concannon P, et al.; Type 1 Diabetes Genetics Consortium. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet* 2009;41:703–707
- Cooper JD, Walker NM, Smyth DJ, Downes K, Healy BC, Todd JA; Type 1 Diabetes Genetics Consortium. Follow-up of 1715 SNPs from the Wellcome Trust Case Control Consortium genome-wide association study in type 1 diabetes families. *Genes Immun* 2009;10(Suppl. 1):S85–S94
- Yang D, Chen Q, Su SB, et al. Eosinophil-derived neurotoxin acts as an alarmin to activate the TLR2-MyD88 signal pathway in dendritic cells and enhances Th2 immune responses. *J Exp Med* 2008;205:79–90
- Bovin LF, Rieneck K, Workman C, et al. Blood cell gene expression profiling in rheumatoid arthritis. Discriminative genes and effect of rheumatoid factor. *Immunol Lett* 2004;93:217–226
- Bennett L, Palucka AK, Arce E, et al. Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. *J Exp Med* 2003;197:711–723
- Westra HJ, Peters MJ, Esko T, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* 2013;45:1238–1243
- Penforinis A, Tuomilehto-Wolf E, Faustman DL, Hitman GA; DiMe (Childhood Diabetes in Finland) Study Group. Analysis of TAP2 polymorphisms in Finnish individuals with type 1 diabetes. *Hum Immunol* 2002;63:61–70
- Caillat-Zucman S, Bertin E, Timsit J, Boitard C, Assan R, Bach JF. TAP1 and TAP2 transporter genes and predisposition to insulin dependent diabetes mellitus. *C R Acad Sci III* 1992;315:535–539
- Sigurdsson S, Nordmark G, Göring HH, et al. Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus. *Am J Hum Genet* 2005;76:528–537
- Sigurdsson S, Padyukov L, Kurreeman FA, et al. Association of a haplotype in the promoter region of the interferon regulatory factor 5 gene with rheumatoid arthritis. *Arthritis Rheum* 2007;56:2202–2210
- Dideberg V, Kristjansdottir G, Milani L, et al. An insertion-deletion polymorphism in the interferon regulatory factor 5 (IRF5) gene confers risk of inflammatory bowel diseases. *Hum Mol Genet* 2007;16:3008–3016
- Valle A, Giamporcaro GM, Scavini M, et al. Reduction of circulating neutrophils precedes and accompanies type 1 diabetes. *Diabetes* 2013;62:2072–2077
- Sobol AB, Watala C. The role of platelets in diabetes-related vascular complications. *Diabetes Res Clin Pract* 2000;50:1–16
- Ferreira RC, Guo H, Coulson RMR, et al. A type 1 interferon transcriptional signature precedes autoimmunity in children genetically at risk for type 1 diabetes. *Diabetes* 2014;63:2538–2550
- Reynier F, Pachot A, Paye M, et al. Specific gene expression signature associated with development of autoimmune type-1 diabetes using whole-blood microarray analysis. *Genes Immun* 2010;11:269–278
- Dotta F, Censini S, van Halteren AG, et al. Coxsackie B4 virus infection of beta cells and natural killer cell insulinitis in recent-onset type 1 diabetic patients. *Proc Natl Acad Sci U S A* 2007;104:5115–5120

37. Hiltunen M, Hyöty H, Knip M, et al. Islet cell antibody seroconversion in children is temporally associated with enterovirus infections. *Childhood Diabetes in Finland (DiMe) Study Group. J Infect Dis* 1997;175:554–560
38. Kuitunen M, Saukkonen T, Ilonen J, Åkerblom HK, Savilahti E. Intestinal permeability to mannitol and lactulose in children with type 1 diabetes with the HLA-DQB1*02 allele. *Autoimmunity* 2002;35:365–368
39. Bosi E, Molteni L, Radaelli MG, et al. Increased intestinal permeability precedes clinical onset of type 1 diabetes. *Diabetologia* 2006;49:2824–2827
40. Westerholm-Ormio M, Vaarala O, Pihkala P, Ilonen J, Savilahti E. Immunologic activity in the small intestinal mucosa of pediatric patients with type 1 diabetes. *Diabetes* 2003;52:2287–2295
41. Giongo A, Gano KA, Crabb DB, et al. Toward defining the autoimmune microbiome for type 1 diabetes. *ISME J* 2011;5:82–91
42. Li Q, Xu B, Michie SA, Rubins KH, Schreiber RD, McDevitt HO. Interferon-alpha initiates type 1 diabetes in nonobese diabetic mice. *Proc Natl Acad Sci U S A* 2008;105:12439–12444
43. Hedl M, Abraham C. IRF5 risk polymorphisms contribute to interindividual variance in pattern recognition receptor-mediated cytokine secretion in human monocyte-derived cells. *J Immunol* 2012;188:5348–5356
44. Heinig M, Petretto E, Wallace C, et al.; Cardiogenics Consortium. A trans-acting locus regulates an anti-viral expression network and type 1 diabetes risk. *Nature* 2010;467:460–464