Journal of Autoimmunity 35 (2010) 70–76

Contents lists available at ScienceDirect

Journal of Autoimmunity

journal homepage: www.elsevier.com/locate/jautimm

Early suppression of immune response pathways characterizes children with prediabetes in genome-wide gene expression profiling

Laura L. Elo [a,b,1], Juha Mykkänen [a,1], Tuomas Nikula [a,2], Henna Järvenpää [a], Satu Simell [c], Tero Aittokallio [a,b], Heikki Hyöty [e,f], Jorma Ilonen [d,g], Riitta Veijola [h], Tuula Simell [c], Mikael Knip [i,j], Olli Simell [c], Riitta Lahesmaa [a,k,*]

[a] *Turku Centre for Biotechnology, University of Turku and Åbo Akademi University, FI-20521 Turku, Finland*
[b] *Department of Mathematics, University of Turku, FI-20014 Turku, Finland*
[c] *Department of Pediatrics, University of Turku, FI-20014 Turku, Finland*
[d] *Immunogenetics Laboratory, University of Turku, FI-20014 Turku, Finland*
[e] *Department of Virology, University of Tampere, FI-33014 Tampere, Finland*
[f] *Centre for Laboratory Medicine, University Hospital of Tampere, FI-33520 Tampere, Finland*
[g] *Department of Clinical Microbiology, University of Kuopio, FI-70211 Kuopio, Finland*
[h] *Department of Pediatrics, University of Oulu, FI-90014 Oulu, Finland*
[i] *Hospital for Children and Adolescents and Folkhälsan Research Center, University of Helsinki, FI-00014 Helsinki, Finland*
[j] *Department of Pediatrics, Tampere University Hospital, FI-33521 Tampere, Finland*
[k] *Immune Disease Institute, Harvard Medical School, Boston, MA, USA*

A R T I C L E I N F O

Article history:
Received 26 November 2009
Received in revised form
5 February 2010
Accepted 1 March 2010

Keywords:
Type 1 diabetes
Prediabetes
Autoantibodies
Longitudinal data
Differential gene expression
Molecular networks
Pathway analysis

A B S T R A C T

Type 1 diabetes (T1D) is caused by autoimmune destruction of insulin-producing pancreatic β cells in the islets of Langerhans. Although defects in various T cell subsets have been linked to the disease pathogenesis, mechanisms initiating or enhancing the autoimmunity in prediabetes remain poorly understood. To unravel genes and molecular pathways affected by the diabetes-associated autoimmunity, we investigated transcriptomic profiles of prospective whole-blood samples from children who have developed T1D-associated autoantibodies and eventually clinical T1D. Gene-level investigation of the data showed systematic differential expression of 520 probesets. A network-based analysis revealed then a highly significant down-regulated network of genes involved in antigen presentation as well as T-cell receptor and insulin signaling. Finally, detection of dynamic changes in the affected pathways at the early or late phases of autoimmunity showed down-regulation of several novel T1D-associated pathways as well as known key components of immune response. The longitudinal genome-wide data generated in the present study allows the detection of dynamic changes relevant to the disease that may be completely missed in conventional cross-sectional studies or in genome-wide association studies. Taken together, our analysis showed systemic high-level repression of immune response pathways associated with T1D autoimmunity.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Type 1 diabetes (T1D; MIM 222100) is an autoimmune disease caused by selective destruction of insulin-producing pancreatic β cells in the islets of Langerhans. The progression to clinical diabetes is characterized by the appearance of autoantibodies against islet cells (ICA), insulin (IAA), protein tyrosine phosphatase-related IA-2 protein (IA-2A) and glutamic acid decarboxylase (GADA), which are considered the first markers signifying the initiation of autoimmunity. In addition, a new T1D-linked autoantigen, cation efflux transporter ZnT8, has been recently reported [1]. The incidence of T1D, being highest in Finland [2], is increasing especially in

* Corresponding author. Turku Centre for Biotechnology, P.O. Box 123, FIN-20521 Turku, Finland. Tel.: +358 40 718 4813; fax: +358 2 333 8000.
E-mail addresses: laura.elo@utu.fi (L.L. Elo), juha.mykkanen@utu.fi (J. Mykkänen), tuomas.nikula@perkinelmer.fi (T. Nikula), henna.jarvenpaa@btk.fi (H. Järvenpää), satu.simell@tyks.fi (S. Simell), tero.aittokallio@utu.fi (T. Aittokallio), heikki.hyoty@uta.fi (H. Hyöty), jorma.ilonen@utu.fi (J. Ilonen), riitta.veijola@oulu.fi (R. Veijola), tuula.simell@utu.fi (T. Simell), mikael.knip@helsinki.fi (M. Knip), olli.simell@utu.fi (O. Simell), riitta.lahesmaa@btk.fi (R. Lahesmaa).
[1] These authors contributed equally to this work.
[2] Present address: Perkin–Elmer Finland, P.O. Box 10, FI-20521 Turku, Finland.

0896-8411/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jaut.2010.03.001

the Western world. For systematic development of T1D prevention, identification of factors contributing to the progression toward clinical T1D is essential.

Clearly, development of autoimmunity requires complex gene–environment interaction [3]. The strongest genetic contribution to the T1D risk comes from the HLA region, where certain HLA class II alleles from the DR and DQ loci are associated with increased risk, while some others are neutral or protective. Large scale genome-wide association studies (GWAS) have confirmed involvement of more than 20 genes or genomic regions in the pathogenesis of T1D and eighteen other loci may affect the risk of T1D [4]. Confirmed genes include insulin (*INS*), lymphoid protein tyrosine phosphate, non-receptor type 22 (*PTPN22*), interleukin-2 receptor alpha chain (*IL2RA* or *CD25*) and cytotoxic T-lymphocyte-associated protein 4 (*CTLA-4*). While a large number of common DNA variants influencing the risk for T1D have already been identified on the basis of the comprehensive genetic studies, additional variants with low allele frequency or those that are specific for a particular population are likely to be found [5].

The destruction of β cells is mediated by T cells infiltrating the islets of Langerhans during the autoimmune attack and, based on the studies with animal models and isolated cell lines, several mechanisms for intra-islet β cell destruction have been proposed [6]. Studies on the T cells in human T1D are complicated due to the inaccessible site of inflammation in the islets. Therefore, investigations on T cells have been focused on peripheral blood of T1D patients where existence of autoreactive T cells for β cell antigens [7] and functional defects in various T cell subsets have been described [8]. However, functional data on the prediabetic period is sparse [9]. Thus, the mechanisms initiating and enhancing diabetes-associated autoimmunity in prediabetes remain poorly understood.

In the present investigations, we have longitudinally analyzed genome-wide peripheral blood RNA profiles from children who have developed T1D-associated autoantibodies and eventually clinical type 1 diabetes (Table 1). Transcriptomic profiling of whole-blood provides an approach for monitoring T1D disease process without *ex-vivo* manipulation and is complementary to studies of isolated cell populations [10] or serum [11]. Moreover, the longitudinal genome-wide profiling allows detection of such dynamic functional gene or pathway-level changes relevant to the disease that may be completely missed by other types of genetic studies, such as GWAS, or by conventional cross-sectional studies. Importantly, we identified here a key network showing systematic repression already in the prediabetic states. Furthermore, a pathway-level analysis of these data revealed a number of changes in gene expression, reflecting potential mechanisms of autoimmunity and clinical diabetes.

Table 1
Summary of the subjects and samples included in the current study.

Subject ID	Gender	T1D Progression	Age at T1D diagnosis or seroconversion (years)	Whole-blood RNA samples studied	Sample collection period (months)
T1D_1	Male	Affected	4.2	9	27
T1D_2	Female	Affected	5.0	8	22
T1D_3	Female	Affected	6.6	8	25
Ab_1	Female	Ab+	2.1	9	27
Ab_2	Female	Ab+	3.5	10	31
T1D_C1	Male	None	–	5	24
T1D_C2	Female	None	–	6	36
T1D_C3	Female	None	–	5	37

Children who progressed to clinical T1D were positive for at least two autoantibodies (Ab) during the collection of the RNA samples. The Ab positive (Ab+) subjects were positive for ICA and the subject Ab_2 was also transiently positive for GAD autoantibody during sample collection. Until now, neither of the Ab positive children has progressed to clinical T1D. T1D_C1–T1D_C3 children are controls for T1D_1–T1D_3 children matched for time of birth, gender, birth place and HLA-genotype.

2. Material and methods

2.1. Subjects and samples

The study subjects were participants in the Type 1 Diabetes Prediction and Prevention study (DIPP) [12], which is a population-based natural history study based on frequent follow-up of genetically susceptible children starting from birth. Whole-blood RNA samples were collected during the clinic visits typically at 3 to 12-month intervals. 2.5 ml venous blood was drawn into PAXgene Blood RNA tubes (PreAnalytiX, Hombrechtikon, Switzerland) according to the manufacturer's instructions for storage at −70 °C until analyzed. Four T1D-associated autoantibodies (IAA, GADA, IA-2A and ICA) were measured from all samples collected. Prospective samples from three children who progressed to T1D (progressors) and from two children who seroconverted to autoantibody positivity (Ab+) during the follow-up were selected to the present study (Table 1). All three progressors were positive for at least two autoantibodies during the collection of the RNA samples. The first autoantibody was detected 0–6 months prior to the collection of the first RNA sample (Fig. S1). The progressors carried high-risk HLA genotype (HLA-DQB1 *02/*0302) and had a first-degree relative with T1D. The mean age of the progressors at the time of seroconversion and T1D diagnosis was 2.0 and 5.3 years, respectively. Persistently autoantibody-negative control children for the T1D cases ($n = 3$) were matched for time of birth, gender, birth place and HLA-genotype, but the children came with no first-degree relatives with T1D (Table 1, Fig. S2). Altogether 60 samples were analyzed for gene expression. The protocol of DIPP study has been accepted by the ethical committees of the University Hospitals of Turku, Oulu and Tampere, and a written consent according to the Declaration of Helsinki was obtained from all study subjects and/or their parents.

2.2. RNA extraction and microarray hybridization

Total whole-blood RNA of the study childrens' follow-up samples was extracted using PAXgene RNA Blood RNA kit (Qiagen, Hilden, Germany). RNA quality and quantity was determined using NanoDrop ND-1000 (NanoDrop Technologies, Wilmington, DE, USA) and Experion Automated Electrophoresis System (Bio-Rad Laboratories, Espoo, Finland). A total of 100 ng of totalRNA was processed with Two-Cycle Target Labeling protocol and hybridized on Affymetrix Human Genome U133 Plus 2.0 microarrays according to the manufacturer's instructions. The microarray data are available in the EBI ArrayExpress (http://www.ebi.ac.uk/microarray-as/ae/) database under ArrayExpress ID: E-TABM-666.

2.3. Microarray data analysis

The gene expression data were normalized and summarized using the robust multiarray average (RMA) procedure implemented in the Bioconductor affy package [13,14]. Principal component analysis was used to project the whole-genome expression profiles of the samples onto two dimensions to enable an overall visualization of the variation in the individual expression profiles. A gene-level statistical analysis was performed to identify genes showing differential expression in the Ab+ children and the progressors as compared to healthy, persistently autoantibody-negative controls. To avoid the need of direct alignment of the follow-up series, which are not fully synchronized in time between different individuals, a similar approach was taken as previously described in [15]. Briefly, for each subject, the expression intensity value of a particular probeset x at each time point was given a z-score, defined as $z = (x - m)/s$, where m is the mean and s is the standard deviation calculated using all time points in the time series of the matched

controls. Such z-scores penalize those probesets that have high variability in the control samples. We considered that a probeset was changed if it had $z > 3$ or $z < -3$ at two or more time points in each of the Ab+ children and the progressors. This corresponded to a false discovery rate (FDR) below 0.01, as calculated by randomly permuting the sample labels (1000 permutations). The gene ontology (GO) annotation of the differentially expressed genes was determined using the DAVID functional annotation tool according to biological process (BP) level 2 [16]. Finally, the K-means clustering was applied to the identified up- or down-regulated genes to investigate their overall behavior across the samples. The Euclidean distance was used as a distance measure and the number of clusters was determined based on the figure of merit (FOM) statistics [17].

2.4. Network and pathway-level analyses of gene expression data

The genes in the down-regulated co-expression clusters with the highest expression levels were studied with the Ingenuity Pathways Analysis (IPA; Ingenuity® Systems, http://www.ingenuity.com). The dataset containing the gene identifiers and the corresponding expression values was uploaded into the application and each gene identifier was mapped to its corresponding gene object. These genes, called focus genes, were then overlaid onto a global molecular network extracted from the information contained in the Ingenuity Pathways Knowledge Base and highly interconnected subnetworks of the focus genes were algorithmically identified based on their connectivity. The significance of these subnetworks was assessed with respect to random networks of the same size using the Fisher's exact test. The biological functions and/or diseases that were most significantly associated to the genes in a particular network were identified using the IPA Functional Analysis tool (Fisher's exact test).

To study the dynamic activation/repression of molecular pathways during the progression of autoimmunity toward overt T1D, a gene set enrichment analysis (GSEA) was performed [18]. This analysis maps the gene-level expression changes to enriched pathways at the different time points. In particular, we identified the up- and down-regulated pathways among the first two (referred to as early) and the last two (referred to as late) follow-up samples of the children who in the study progressed to clinical T1D. These samples were drawn 0–6 months after seroconversion to autoantibody positivity or 0–3 months before diagnosis of T1D, respectively. In both cases the average z-score was first calculated for each gene and subject separately using the early and late samples. These values were then combined into ranked gene lists using the rank product approach implemented in the Bioconductor RankProd package [19]. The obtained gene rankings were finally used in the GSEA analysis to identify the up- or down-regulated pathways at the early and late time points. Pathways with FDR < 0.1 were considered significant.

3. Results

3.1. Unsupervised principal component analysis

An unsupervised principal component analysis (PCA) was first used to project the genome-wide expression of the longitudinal sample series onto two dimensions to enable an overall visualization of the variation in the individual expression profiles. This showed that the samples taken from the two autoantibody positive (Ab+) children and the three children who progressed to clinical diabetes (progressors) were separated from the control samples (Fig. 1A, Table 1). Only the second and third sample of subject T1D_2 and two follow-up samples of subject Ab_2 overlapped with the control samples. The global gene expression profiles of the Ab+ children and the progressors were not distinguishable. This analysis

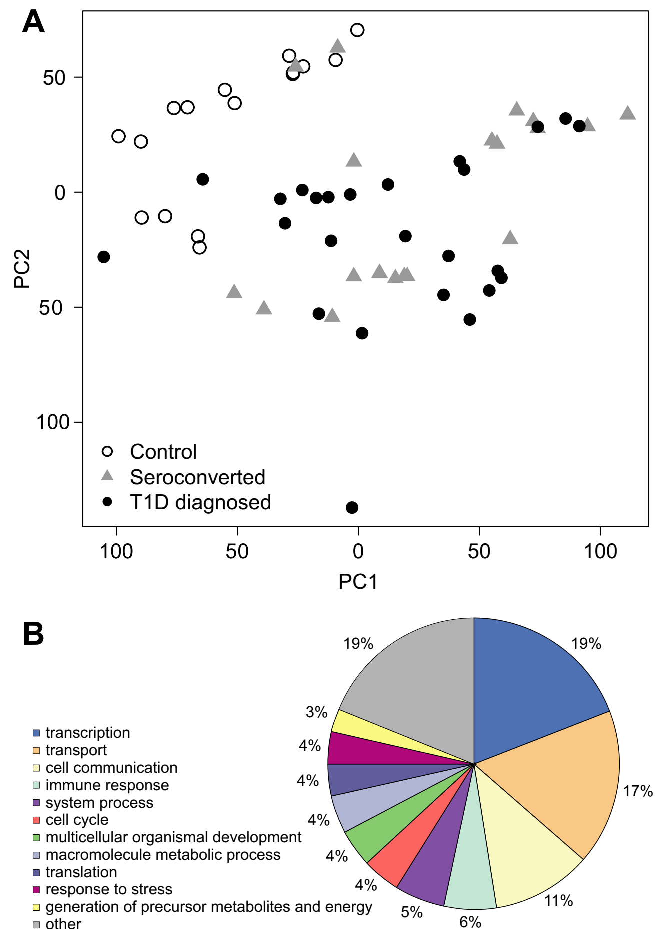


Fig. 1. Expression profiling of diabetes-associated autoimmunity. (A) PCA visualization of the whole-genome expression profiles. All follow-up samples from the three children who progressed to clinical T1D (filled circles), their matched controls (open circles) and the two children who seroconverted to autoantibody (Ab) positivity (triangles) are plotted separately. (B) Proportions of biological processes related to the differentially expressed genes when comparing the Ab positive children and the children who soon progressed to overt T1D with the healthy autoantibody-negative controls. The biological processes were defined using the standardized terms in the Gene Ontology database (www.geneontology.org).

suggests that the primary determinants of global gene expression variation exert their influence in Ab+ children and children who progress to clinical T1D.

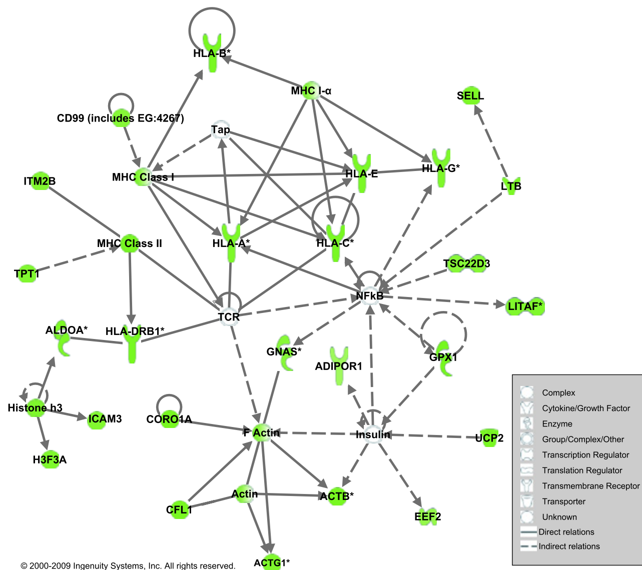
3.2. Differentially expressed genes

To study more closely the gene expression differences in the Ab+ children and the progressors as compared with the healthy autoantibody-negative controls, we analyzed the whole dataset on the basis of the z-scores (see Section 2.3). With the selected high-stringency z-value cutoff ($|z| > 3$), 520 probesets representing 424 individual genes were consistently differentially expressed across all study subjects ($n = 5$) (Table S1). Altogether 257 probesets (49.4%) were down-regulated and 263 probesets (50.6%) were up-regulated. The identified genes included *PTPN22*/LYP, from which a functional variant C1858T (620Trp) has previously been associated with T1D [20]. Since this variant is also associated with T1D in the Finnish population [21], its presence in the study subjects was determined. Two progressors (T1D_1 and T1D_2) and one Ab+ child (Ab_2) were heterozygous for T allele, whereas the other study subjects, including the matched control children, were

homozygous for wild-type C allele. The differentially expressed genes also included 18 genes from the HLA Class I and Class II regions (Table S1). Of the known T1D-associated genomic regions other that HLA (6p21) and *PTPN22* (1p13), the dexamethasone-induced transcript (*DEXI*; 16p13.13), LPS-induced TN factor (*LITAF*; 16p13.13) and nipsnap homolog 1 (*NIPSNAP1*; 22q12.2) were down-regulated in our analysis [4]. The gene ontology annotation of biological processes of the differentially expressed genes showed that nearly half (47%) of the genes are involved in transcription, translation and cell communication (Fig. 1B).

3.3. Highly connected network of down-regulated genes

Clustering of the gene-level data according to the array signal intensity suggested that, in general, the overall signal intensity levels of the down-regulated genes were higher than those of the up-regulated genes in the peripheral blood of the study children (data not shown). A network-based investigation of the genes belonging to the down-regulated co-expression clusters with the highest expression levels revealed a highly connected network composed predominantly of immune system related genes (IPA, $p = 10^{-59}$, 25 genes) (Fig. 2). The most significant biological functions according to the Ingenuity Knowledge Base associated with the network included cellular growth and proliferation ($p = 10^{-7}$), hematological system development and function ($p = 10^{-7}$) and immunological disease ($p = 10^{-8}$). Notably, the repression of the HLA region genes was prominent and was not restricted only to the Class II genes, but also almost all the genes from the HLA Class I were present in the network. In addition, the network involved also other genes related to the critical functions of the immune system, including T-cell receptor (TCR), insulin, actin cytoskeleton and NF-kB. In summary, these results demonstrate that genes contributing to critical functions of the immune system, antigen presentation as well as to T-cell receptor and insulin signaling are repressed both in Ab+ children and children who soon are going to progress to clinical T1D.



© 2000-2009 Ingenuity Systems, Inc. All rights reserved.

Fig. 2. Highly connected network of down-regulated genes. The most significant network of the down-regulated genes ($p = 10^{-59}$) as identified with IPA using the down-regulated co-expression clusters (105 probesets) with highest array signal intensity as input. Green node color indicates down-regulation, whereas white color denotes molecules that were not included in the input dataset, but were incorporated into the network through relationships with other molecules. Molecules with multiple identifiers in the dataset are shown with an asterisk. Direct and indirect relations between genes are indicated with solid and dashed lines, respectively, as extracted from the Ingenuity Pathways Knowledge Base.

3.4. Pathway-level analyses of early and late phase prediabetic samples

To further study the molecular pathways that may be affected and changed during the progression of the autoimmune process toward T1D, we analyzed independently the first two (early) and the last two (late) follow-up samples of the T1D subjects using gene set enrichment analysis (GSEA) [18]. As described in more detail in Section 2.1 and 2.4, the early samples were obtained 0–6 months after seroconversion, whereas the late samples were collected 0–3 months prior to T1D diagnosis. The period of seropositivity for at least two T1D-associated autoantibodies between the early and late samples was 16–21 months. Since the Ab+ children did not develop T1D during the follow-up period, they were excluded from this analysis.

We found a total of 24 down-regulated gene sets affected at the false discovery rate (FDR) of 0.1 in the early samples, whereas no up-regulated gene sets were observed at that stage (Table 2). When the samples representing late prediabetes were analyzed, 28 gene sets were down-regulated and 3 were up-regulated. Of these gene

Table 2
Significant pathways from gene set enrichment analysis (GSEA).

Down-regulated gene sets	FDR early	FDR late
HSA03010_RIBOSOME	0	0.007
RIBOSOMAL_PROTEINS	0.001	0.002
HSA04612_ANTIGEN_PROCESSING_AND_PRESENTATION	0.002	0.033
SIG_CD40PATHWAYMAP	0.007	0.034
PROTEASOME	0.026	ns
EIF4PATHWAY	0.025	ns
AMIPATHWAY	0.025	0.05
HSA04940_TYPE_I_DIABETES_MELLITUS	0.024	0.014
HSA03050_PROTEASOME	0.025	ns
HSA05030_AMYOTROPHIC_LATERAL_SCLEROSIS	0.022	ns
CSKPATHWAY	0.028	0.039
NKCELLSPATHWAY	0.046	0.034
SIG_IL4RECEPTOR_IN_B_LYPHOCYTES	0.046	0.026
ST_DICTYOSTELIUM_DISCOIDEUM_CAMP_ CHEMOTAXIS_PATHWAY	0.066	0.034
SIG_BCR_SIGNALING_PATHWAY	0.065	0.034
IL2RBPATHWAY	0.071	0.063
NFKBPATHWAY	0.07	ns
PROTEASOMEPATHWAY	0.068	ns
SIG_REGULATION_OF_THE_ACTIN_CYTOSKELETON_ BY_RHO_GTPASES	0.066	ns
SIG_CHEMOTAXIS	0.069	ns
CERAMIDEPATHWAY	0.067	ns
RELAPATHWAY	0.075	ns
HSA05020_PARKINSONS_DISEASE	0.075	ns
RACCYCDPATHWAY	0.079	0.036
UCALPAINPATHWAY	ns	0.043
PTENPATHWAY	ns	0.034
PTDINSPATHWAY	ns	0.035
INTEGRIN_MEDIATED_CELL_ADHESION_KEGG	ns	0.055
ST_INTERLEUKIN_4_PATHWAY	ns	0.064
HSA04320_DORSO_VENTRAL_AXIS_FORMATION	ns	0.06
ST_B_CELL_ANTIGEN_RECEPTOR	ns	0.063
HSA04520_ADHERENS_JUNCTION	ns	0.06
SA_TRKA_RECEPTOR	ns	0.069
CHR20P13	ns	0.086
RHOPATHWAY	ns	0.09
SIG_INSULIN_RECEPTOR_PATHWAY_ IN_CARDIAC_MYOCYTES	ns	0.088
SIG_PIP3_SIGNALING_IN_CARDIAC_MYOCYTES	ns	0.093
Up-regulated gene sets	**FDR early**	**FDR late**
UBIQUITIN_MEDIATED_PROTEOLYSIS	ns	0.025
VIPPATHWAY	ns	0.088
NFATPATHWAY	ns	0.096

Analysis was performed independently for the first two (early) and last two (late) follow-up samples. A total of 37 down-regulated and 3 up-regulated gene sets were identified at false discovery rate FDR of 0.1.
ns, Not significant.

sets, 13 were down-regulated both in early and late samples. The most significant common gene sets were related to ribosomes and, thereby, to diminished basal activity of the translational machinery. The other gene sets common to both early and late phase of prediabetes comprised pathways of more specific immunological functions probably relevant for T1D development, including functions like antigen processing and presentation (HSA04612_ANTIGEN_PROCESSING_AND_PRESENTATION) and type 1 diabetes mellitus (HSA04940_TYPE_I_DIABETES_MELLITUS). Seven of the shared down-regulated pathways contribute to the signaling and proliferation of different lymphocyte subsets (T-, B- and NK-cells).

In addition to shared pathways, a total of 11 pathways were specific for the early samples and 16 for the late samples. The affected pathways found only in the early samples were all down-regulated and involved functions like protein degradation (PROTEASOME, HSA 03050_PROTEASOME), protein translation (EIF4PATHWAY), inflammation (NFKBPATHWAY, HSA05030_AMYOTROPHIC_LATERAL_SCLE ROSIS, RELAPATHWAY, CERAMIDEPATHWAY, HSA05020_PARKINSO NS_DISEASE) and organization of cytoskeleton (SIG_REGULATION_ OF_THE_ACTIN_CYTOSKELETON_BY_RHO_GTPASE, SIG_CHEMOTAX IS). Of the 16 gene sets detected only in the late samples, 13 were down-regulated. They were mainly involved in cell adhesion (UCAL PAINPATHWAY, INTEGRIN_MEDIATED_CELL_ADHESION_KEGG, HS A04520_ADHERENS_JUNCTION, RHOPATHWAY), and phosphoinositide 3-kinase (PI3K) signaling (PTENPATHWAY, PTDINSPATHWAY, ST_INTERLEUKIN_4_PATHWAY, ST_B_CELL_ANTIGEN_RECEPTOR, S A_TRKA_RECEPTOR, RHOPATHWAY, SIG_INSULIN_RECEPTOR_PATH WAY_IN_CARDIAC_MYOCYTES, SIG_PIP3_SIGNALING_IN_CARDIAC_ MYOCYTES). The only up-regulated gene sets in our study were observed in the late samples. Two of the three up-regulated gene sets were related to NFAT signaling pathways (VIPPATHWAY, NFATPATHWAY), whereas the third gene set represents a protein degradation pathway (UBIQUITIN_MEDIATED_PROTEOLYSIS). In summary, the majority of the affected pathways are central components of the immune response.

4. Discussion

The aim of the present study was to uncover and characterize the gene expression signatures induced by the autoimmune process characterizing child's progression toward type 1 diabetes. We analyzed a unique set of longitudinal samples drawn from children who progressed to clinical T1D or developed islet cell autoantibodies as first signs of beta-cell autoimmunity. Already the principal component analysis proposed that T1D autoimmune inflammatory process is reflected into gene expression profiles at the genome-wide level. However, since the segregating variation in gene expression is seen already in the first analyzed samples soon after appearance of autoantibodies, more studies are needed for identification of the timing of the change in transcriptional profile.

At the level of individual genes, we observed statistically significant changes in the expression of 424 genes in autoantibody positive and prediabetic follow-up samples. Amongst the up-regulated genes was the *PTPN22* gene, encoding the LYP phosphatase that negatively regulates the T-cell receptor signaling and whose C1858T variant leads to gain-of-function of the phosphatase [22]. A recent study demonstrated that CD4+ T-cells of T1D patients carrying the 620Trp variant have reduced activation and IL-2 signaling [23]. Moreover, *CSK*, a negative regulator of LYP [20], was down-regulated in our data. Vang et al. showed that binding of LYP-620Trp with CSK is reduced, while the phosphatase activity and suppression of TCR signaling by LYP-620Trp variant is coincidently increased [22]. As in the present study, some but not all of the Ab+ children and the progressors were heterozygous for the T allele, the contribution of the 620Trp variant versus other factors in regulating the expression of these genes cannot be concluded and should be addressed through additional studies. Other members of the *PTPN* gene family are also potential candidates for T1D. The study by Barrett et al. confirmed an association of *PTPN2* gene in the pathogenesis of T1D [4]. Here we detected differential expression of the *PTPN4* gene, which is involved in T cell receptor signaling [24] and, thus, may play a role in T1D development.

The most striking finding of the present study was the systemic high-level repression of a key network including genes of antigen presentation and immune response (Fig. 2). The repression of the HLA region genes was prominent and was not restricted to the Class II genes alone. Genetically, T1D susceptibility alleles are found from HLA Class I (HLA-A and HLA-B) independently from HLA Class II (HLA-DR/DQ) locus [25]. It has been shown that Class I molecules are up-regulated in the pancreatic islets of T1D patients [26], and pathogenic CD8 + T cells are the major mediators of beta-cell killing [6]. However, the HLA Class I expression of T1D peripheral blood lymphocytes has been shown to be either down-regulated [27,28] or normal [29]. Peripheral monocyte expression of HLA Class II has been shown to be down-regulated [30] in T1D and it has been suggested that down-regulation of HLA serves as a protection mechanism against the ongoing autoimmunity [30]. In addition to antigen presentation, the downstream components of T-cell activation cascade, including actin filament system and NF-kB, were involved in the repressed network. Moreover, genes and pathways associated with insulin and insulin signaling were down-regulated. Insulin has been implicated as a major autoantigen in type 1 diabetes [31]. Down-regulation of thymic insulin expression has been earlier shown to be associated with insulin gene polymorphisms [32] and it has been proposed that low insulin expression level in thymus may lead to escape of insulin-reactive T-cells [33]. One possible interpretation of our results is that the ongoing autoimmunity against insulin and the diabetic disease process leads to decreased insulin expression levels which, in turn, are reflected at the systems level by down-regulation of insulin regulated genes and pathways.

The trend for prediabetes-associated suppression of gene transcription was further demonstrated at the biological pathway level. Ribosomal pathways were statistically the most significant down-regulated pathways. Peripheral blood lymphocytes are naturally in a resting state with low level of protein synthesis until they meet appropriate activation signal leading to rapid proliferation. The proliferation of activated lymphocytes is dependent on increased protein synthesis from pre-existing and newly generated mRNAs that occurs before synthesis of new ribosomal RNA [34]. The presence of free ribosomal protein pool may be a prerequisite for accelerated lymphocyte activation [35]. Thus, one hypothesis is that prediabetic lymphocytes produce less free ribosomes, which subsequently results in attenuated activation capacity [36]. Complementary to the gene network analysis, defects were also identified in the pathways for antigen processing and presentation, cytokine receptor signaling, T- and B-cell activation, co-activation (CD40) and proliferation.

The pathways found to be suppressed in our study included genes previously associated with T1D development including IL-2 receptor signaling [37] and CSK [20,22]. Our finding of down-regulated IL-2 receptor β further supports the involvement of IL-2 cytokine in T1D pathogenesis [38,39]. The defect in IL-2 signaling may in turn contribute to the impairment of regulatory T-cells [38]. The NF-kB and CD40 pathways are important for the function of antigen presenting cells (APC) and related defects have previously been reported in T1D patients [40].

Interestingly, a considerable number of the pathways included the PI3K signaling component, especially at the late phase of prediabetes. PI3K acts as an upstream activator of serine/threonine kinase AKT which has pivotal role in cell growth, proliferation, metabolism and survival [41]. In lymphocytes, PI3K is a key player

in insulin receptor signaling and antigen receptor signaling [42,43]. Especially the p110δ catalytic isoform of PI3K, which is encoded by PIK3CD and was down-regulated in our gene-level analysis, has an important role in T- and B-cell development and function [44]. PI3K/AKT signaling also inactivates NFAT [43], which is in line with the observed late phase up-regulated pathways. Insulin receptor signaling and PI3K/AKT pathway have been previously linked to survival of beta-cells and insulin resistance in type II diabetes [45]. PI3K signaling has not been studied in detail in type 1 diabetes but in the light of our results further investigations are warranted.

In general, our results are well in line with previous studies showing that T-cells of T1D patients are anergic and have impaired T-cell receptor signaling [36,46]. The trend for down-regulation of immune response genes has been also previously detected from CD4+ T-cells of T1D patients [10,38]. Moreover, our results suggest that prediabetes is characterized by functional defects in CD8+ T-cells and B-cells, which may well orchestrate together the disease development [47]. A systemic attenuation of immune responses through several distinct lymphocyte subsets may explain the difficulties in enhancing the immune tolerance by intervention trials and the reason why combination therapies may have better prospects of being effective [48].

Recently, Torkamani et al. performed a pathway analysis using the published GWAS data for type 1 diabetes and six other common diseases [49]. Interestingly, the significantly overrepresented pathways for T1D included many pathways detected in our present data, including inositol triphosphate signaling and immune response (NFAT and IL-2 signaling and antigen presentation by MHC class I and II) [49]. In addition to the genetic variation, the environmental factors like infections and nutrition may contribute to the T1D development [50]. A gene–environment interaction, for example, has been suggested between predisposing *PTPN22* C1858T polymorphism and cow's milk formula [51]. Additionally, DNA methylation, histone modifications and other epigenetic factors may also induce or repress gene expression leading to T1D autoimmunity [52,53]. Therefore, the possible role of the epigenetic control of gene expression in type 1 diabetes development should be addressed in future studies.

The increased amount of information gained by the application of high-throughput technologies to human samples enables the molecular characterization of human disease and, thereby, holds potential to lead to improved models, for instance, to predict susceptibility to disease or treatment responses. Besides investigating the behavior of individual genes, correlating the high-throughput data to biological responses at a network level can advance our understanding of the deregulated pathways during the disease development [54,55]. Such networks can also suggest a starting model to be iteratively defined in further experiments. Disease intervention of T1D is dependent on the identification of at-risk individuals based on genetic or other stratification criteria. The prediabetic children in our study had first-degree relatives with T1D and, thus, are in higher risk for T1D than the general population probably because of shared alleles of susceptibility genes or haplotypes [56]. The genome-wide transcriptomic analysis of the follow-up samples of children with ongoing autoimmune progression toward T1D revealed several differentially regulated central pathways and networks relevant for T1D autoimmunity. These included both previously known and unknown networks and provided new hypotheses to be experimentally tested. Future studies concentrating on the early period of seroconversion may shed additional light to the molecular mechanisms leading to prediabetes and overt type 1 diabetes.

Acknowledgements

We thank all the staff of the DIPP team for their expert help and contribution. DIPP families are thanked for their support. We are grateful for Katja Waenerberg and the Finnish DNA Microarray Centre at Turku Centre for Biotechnology for excellent technical assistance. This study was financially supported by the Juvenile Diabetes Research Foundation (JDRF), the Academy of Finland, the Turku University Hospital Research Fund, the Finnish Funding Agency for Technology and Innovation (Tekes), and the Sigrid Jusélius Foundation.

Appendix. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jaut.2010.03.001.

References

[1] Wenzlau JM, Juhl K, Yu L, Moua O, Sarkar SA, Gottlieb P, et al. The cation efflux transporter ZnT8 (Slc30A8) is a major autoantigen in human type 1 diabetes. Proc Natl Acad Sci USA 2007;104:17040–5.
[2] Harjutsalo V, Sjoberg L, Tuomilehto J. Time trends in the incidence of type 1 diabetes in Finnish children: a cohort study. Lancet 2008;371:1777–82.
[3] Knip M, Veijola R, Virtanen SM, Hyoty H, Vaarala O, Akerblom HK. Environmental triggers and determinants of type 1 diabetes. Diabetes 2005;54(Suppl. 2):S125–36.
[4] Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet; 2009.
[5] Nejentsev S, Walker N, Riches D, Egholm M, Todd JA. Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes. Science 2009;324:387–9.
[6] Notkins AL. Immunologic and genetic factors in type 1 diabetes. J Biol Chem 2002;277:43545–8.
[7] Reijonen H, Mallone R, Heninger AK, Laughlin EM, Kochik SA, Falk B, et al. GAD65-specific CD4 + T-cells with high antigen avidity are prevalent in peripheral blood of patients with type 1 diabetes. Diabetes 2004;53:1987–94.
[8] Brusko TM, Wasserfall CH, Clare-Salzler MJ, Schatz DA, Atkinson MA. Functional defects and the influence of age on the frequency of CD4 + CD25 + T-cells in type 1 diabetes. Diabetes 2005;54:1407–14.
[9] Glisic-Milosavljevic S, Waukau J, Jailwala P, Jana S, Khoo HJ, Albertz H, et al. At-risk and recent-onset type 1 diabetic subjects have increased apoptosis in the CD4 + CD25 + T-cell fraction. PLoS One 2007;2:e146.
[10] Orban T, Kis J, Szereday L, Engelmann P, Farkas K, Jalahej H, et al. Reduced CD4 + T-cell-specific gene expression in human type 1 diabetes mellitus. J Autoimmun; 2007.
[11] Wang X, Jia S, Geoffrey R, Alemzadeh R, Ghosh S, Hessner MJ. Identification of a molecular signature in human type 1 diabetes mellitus using serum and functional genomics. J Immunol 2008;180:1929–37.
[12] Kupila A, Muona P, Simell T, Arvilommi P, Savolainen H, Hamalainen AM, et al. Feasibility of genetic and immunological prediction of type I diabetes in a population-based birth cohort. Diabetologia 2001;44:290–7.
[13] Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 2003;4:249–64.
[14] Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 2004;5:R80.
[15] Huang Q, Liu D, Majewski P, Schulte LC, Korn JM, Young RA, et al. The plasticity of dendritic cell responses to pathogens and their components. Science 2001;294:870–5.
[16] Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4:44–57.
[17] Yeung KY, Haynor DR, Ruzzo WL. Validating clustering for gene expression data. Bioinformatics 2001;17:309–18.
[18] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 2005;102: 15545–50.
[19] 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.
[20] Bottini N, Musumeci L, Alonso A, Rahmouni S, Nika K, Rostamkhani M, et al. A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. Nat Genet 2004;36:337–8.
[21] Hermann R, Lipponen K, Kiviniemi M, Kakko T, Veijola R, Simell O, et al. Lymphoid tyrosine phosphatase (LYP/PTPN22) Arg620Trp variant regulates insulin autoimmunity and progression to type 1 diabetes. Diabetologia 2006;49:1198–208.
[22] Vang T, Congia M, Macis MD, Musumeci L, Orru V, Zavattari P, et al. Autoimmune-associated lymphoid tyrosine phosphatase is a gain-of-function variant. Nat Genet 2005;37:1317–9.
[23] Aarnisalo J, Treszl A, Svec P, Marttila J, Oling V, Simell O, et al. Reduced CD4 + T cell activation in children with type 1 diabetes carrying the PTPN22/Lyp 620Trp variant. J Autoimmun 2008;31:13–21.

[24] Han S, Williams S, Mustelin T. Cytoskeletal protein tyrosine phosphatase PTPH1 reduces T cell antigen receptor signaling. Eur J Immunol 2000;30:1318–25.
[25] Nejentsev S, Howson JM, Walker NM, Szeszko J, Field SF, Stevens HE, et al. Localization of type 1 diabetes susceptibility to the MHC class I genes HLA-B and HLA-A. Nature 2007;450:887–92.
[26] Itoh N, Hanafusa T, Miyazaki A, Miyagawa J, Yamagata K, Yamamoto K, et al. Mononuclear cell infiltration and its relation to the expression of major histocompatibility complex antigens and adhesion molecules in pancreas biopsy specimens from newly diagnosed insulin-dependent diabetes mellitus patients. J Clin Invest 1993;92:2313–22.
[27] Faustman D, Li XP, Lin HY, Fu YE, Eisenbarth G, Avruch J, et al. Linkage of faulty major histocompatibility complex class I to autoimmune diabetes. Science 1991;254:1756–61.
[28] Fu Y, Nathan DM, Li F, Li X, Faustman DL. Defective major histocompatibility complex class I expression on lymphoid cells in autoimmunity. J Clin Invest 1993;91:2301–7.
[29] Hao W, Gladstone P, Engardt S, Greenbaum C, Palmer JP. Major histocompatibility complex class I molecule expression is normal on peripheral blood lymphocytes from patients with insulin-dependent diabetes mellitus. J Clin Invest 1996;98:1613–8.
[30] Parkkonen P, Hyoty H, Huupponen T, Leinikki P, Tuomilehto-Wolf E, Knip M. Defective HLA class II expression in monocytes of type 1 diabetic patients. The childhood diabetes in Finland Study Group. APMIS 1993;101:395–402.
[31] Nakayama M, Abiru N, Moriyama H, Babaya N, Liu E, Miao D, et al. Prime role for an insulin epitope in the development of type 1 diabetes in NOD mice. Nature 2005;435:220–3.
[32] Pugliese A, Zeller M, Fernandez Jr A, Zalcberg LJ, Bartlett RJ, Ricordi C, et al. The insulin gene is transcribed in the human thymus and transcription levels correlated with allelic variation at the INS VNTR-IDDM2 susceptibility locus for type 1 diabetes. Nat Genet 1997;15:293–7.
[33] Vafiadis P, Bennett ST, Todd JA, Nadeau J, Grabs R, Goodyer CG, et al. Insulin expression in human thymus is modulated by INS VNTR alleles at the IDDM2 locus. Nat Genet 1997;15:289–92.
[34] Kay JE, Ahern T, Lindsay VJ, Sampson J. The control of protein synthesis during the stimulation of lymphocytes by phytohaemagglutinin. III. Poly(U) translation and the rate of polypeptide chain elongation. Biochim Biophys Acta 1975;378:241–50.
[35] Cooper HL, Braverman R. Free fibosomes and growth stimulation in human peripheral lymphocytes: activation of free ribisomes as an essential event in growth induction. J Cell Physiol 1977;93:213–25.
[36] Dosch H, Cheung RK, Karges W, Pietropaolo M, Becker DJ. Persistent T cell anergy in human type 1 diabetes. J Immunol 1999;163:6933–40.
[37] Yamanouchi J, Rainbow D, Serra P, Howlett S, Hunter K, Garner VE, et al. Interleukin-2 gene variation impairs regulatory T cell function and causes autoimmunity. Nat Genet 2007;39:329–37.
[38] Jailwala P, Waukau J, Glisic S, Jana S, Ehlenbach S, Hessner M, et al. Apoptosis of CD4 + CD25(high) T cells in type 1 diabetes may be partially mediated by IL-2 deprivation. PLoS One 2009;4:e6527.
[39] Kaye WA, Adri MN, Soeldner JS, Rabinowe SL, Kaldany A, Kahn CR, et al. Acquired defect in interleukin-2 production in patients with type I diabetes mellitus. N Engl J Med 1986;315:920–4.
[40] Mollah ZU, Pai S, Moore C, O'Sullivan BJ, Harrison MJ, Peng J, et al. Abnormal NF-kappa B function characterizes human type 1 diabetes dendritic cells and monocytes. J Immunol 2008;180:3166–75.
[41] Manning BD, Cantley LC. AKT/PKB signaling: navigating downstream. Cell 2007;129:1261–74.
[42] Bilancio A, Okkenhaug K, Camps M, Emery JL, Ruckle T, Rommel C, et al. Key role of the p110delta isoform of PI3 K in B-cell antigen and IL-4 receptor signaling: comparative analysis of genetic and pharmacologic interference with p110delta function in B cells. Blood 2006;107:642–50.
[43] Okkenhaug K, Ali K, Vanhaesebroeck B. Antigen receptor signalling: a distinctive role for the p110delta isoform of PI3K. Trends Immunol 2007; 28:80–7.
[44] Okkenhaug K, Bilancio A, Farjot G, Priddle H, Sancho S, Peskett E, et al. Impaired B and T cell antigen receptor signaling in p110delta PI 3-kinase mutant mice. Science 2002;297:1031–4.
[45] Elghazi L, Bernal-Mizrachi E. Akt and PTEN: beta-cell mass and pancreas plasticity. Trends Endocrinol Metab 2009;20:243–51.
[46] Salojin KV, Zhang J, Madrenas J, Delovitch TL. T-cell anergy and altered T-cell receptor signaling: effects on autoimmune disease. Immunol Today 1998; 19:468–73.
[47] Marino E, Grey ST. A new role for an old player: do B cells unleash the self-reactive CD8 + T cell storm necessary for the development of type 1 diabetes? J Autoimmun 2008;31:301–5.
[48] Haller MJ, Gottlieb PA, Schatz DA. Type 1 diabetes intervention trials 2007: where are we and where are we going? Curr Opin Endocrinol Diabetes Obes 2007;14:283–7.
[49] Torkamani A, Topol EJ, Schork NJ. Pathway analysis of seven common diseases assessed by genome-wide association. Genomics 2008;92:265–72.
[50] Harrison LC, Honeyman MC, Morahan G, Wentworth JM, Elkassaby S, Colman PG, et al. Type 1 diabetes: lessons for other autoimmune diseases? J Autoimmun 2008;31:306–10.
[51] Lempainen J, Vaarala O, Makela M, Veijola R, Simell O, Knip M, et al. Interplay between PTPN22 C1858 T polymorphism and cow's milk formula exposure in type 1 diabetes. J Autoimmun 2009;33:155–64.
[52] Hewagama A, Richardson B. The genetics and epigenetics of autoimmune diseases. J Autoimmun 2009;33:3–11.
[53] Brooks WH, Le Dantec C, Pers JO, Youinou P, Renaudineau Y. Epigenetics and autoimmunity. J Autoimmun; 2010.
[54] Bergholdt R, Storling ZM, Lage K, Karlberg EO, Olason PI, Aalund M, et al. Integrative analysis for finding genes and networks involved in diabetes and other complex diseases. Genome Biol 2007;8:R253.
[55] Loscalzo J, Kohane I, Barabasi AL. Human disease classification in the post-genomic era: a complex systems approach to human pathobiology. Mol Syst Biol 2007;3:124.
[56] Baschal EE, Eisenbarth GS. Extreme genetic risk for type 1A diabetes in the post-genome era. J Autoimmun 2008;31:1–6.