

Both Polymorphic Variable Number of Tandem Repeats and Autoimmune Regulator Modulate Differential Expression of Insulin in Human Thymic Epithelial Cells

Chuan Qi Cai,¹ Tao Zhang,¹ Mary B. Breslin,¹ Matthieu Giraud,² and Michael S. Lan¹

OBJECTIVE—Polymorphic INS-VNTR plays an important role in regulating insulin transcript expression in the human thymus that leads to either insulin autoimmunity or tolerance. The molecular mechanisms underlying the INS-VNTR haplotype-dependent insulin expression are still unclear. In this study, we determined the mechanistic components underlying the differential insulin gene expression in human thymic epithelial cells, which should have profound effects on the insulin autoimmune tolerance induction.

RESEARCH DESIGN AND METHODS—A repetitive DNA region designated as a variable number of tandem repeats (VNTR) is located upstream of the human insulin gene and correlates with the incidence of type 1 diabetes. We generated six class I and two class III VNTR constructs linked to the human insulin basal promoter or SV40 heterologous promoter/enhancer and demonstrated that AIRE protein modulates the insulin promoter activities differentially through binding to the VNTR region.

RESULTS—Here we show that in the presence of the autoimmune regulator (AIRE), the class III VNTR haplotype is responsible for an average of three-fold higher insulin expression than class I VNTR in thymic epithelial cells. In a protein-DNA pull-down experiment, AIRE protein is capable of binding to VNTR class I and III probes. Further, the transcriptional activation of the INS-VNTR by AIRE requires the insulin basal promoter. The VNTR sequence loses its activation activity when linked to a heterologous promoter and/or enhancer.

CONCLUSIONS—These findings demonstrate a type 1 diabetes predisposition encoded by the INS-VNTR locus and a critical function played by AIRE, which constitute a dual control mechanisms regulating quantitative expression of insulin in human thymic epithelial cells. *Diabetes* 60:336–344, 2011

Insulin plays a critical role in maintaining glucose metabolism. Insulin deficiencies lead to hyperglycemia and diabetes. Type 1 diabetes results from specific and progressive autoimmune destruction of pancreatic β cells in the islets of Langerhans. Insulin not only plays a critical role in glucose homeostasis, but also

presents itself as a primary autoantigen in type 1 diabetes (1,2). A polymorphic region containing variable numbers of tandem repeat (VNTR) is located 5'-upstream of the human insulin gene, which in combination with the proximal insulin basal promoter, modulates both β -cell response to glucose and the expression of proinsulin in the thymus (3–5). This locus (INS-VNTR/*IDDM2*) is important in determining genetic susceptibility to diabetes (6–8). The INS-VNTR locus consists of a 14–15 base pair (bp) consensus sequence unit (ACAGGGGTCTGGGG) with slight variations of the repeat subtype units. The number of repeats ranges from 26 to more than 200, but differences in length are not randomly distributed. Intensive linkage analysis on different population cohorts has offered evidence to support that individuals exhibiting class I VNTRs are predisposed to type 1 diabetes, whereas those with class III VNTRs are provided dominant protective effects (9–11).

The autoimmune polyendocrinopathy type 1 syndrome (APS1 also known as APECED) is an autosomal recessive, monogenic form of human autoimmunity, characterized by autoimmune destruction of multiple endocrine organs and defective cell-mediated immunity (12,13). Affected individuals are diagnosed as having two of three clinical manifestations, including Addison's disease, hypoparathyroidism, and chronic mucocutaneous candidal infections. However, other syndromes, including alopecia, exocrine pancreatic deficiency, hypothyroidism, premature ovarian failure, vitiligo, Sjogren's syndrome, and type 1 diabetes could manifest in affected patients (13). In 1997, an autoimmune regulator (*AIRE*) gene, which underlies APS1, was identified by positional cloning strategies on chromosome 21q22 (14,15). The *AIRE* protein plays a central role in regulating tissue-restricted antigen expression in medullary thymic epithelial cells. More than hundreds of tissue-restricted antigens are believed to be modulated in the thymus (16). Ectopic expression of tissue-restricted antigens could trigger self tolerance, whereas lack of expression could lead to autoimmunity. Insulin transcript analysis of individuals with class I or class III haplotype revealed that class III subjects exhibited higher expression levels of insulin than class I subjects in thymus (7,8). A recent study suggested that the risk of autoimmune diabetes in autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED) is associated with short alleles of the 5'-insulin VNTR (17). Therefore, we hypothesized that the *AIRE* protein may function as a regulator of the insulin gene through the VNTR region. In the present study, we show that *AIRE* regulates insulin promoter activity mediated through the INS-VNTR element. The class III VNTR haplotype induces far greater insulin transcript expression than the class I VNTR haplotype and/or the insulin basal promoter alone.

From ¹The Research Institute for Children, Children's Hospital, New Orleans, Louisiana, and the Departments of Pediatrics and Genetics, Louisiana State University Health Sciences Center, New Orleans, Louisiana; and the ²Department of Pathology, Harvard Medical School, Boston, Massachusetts.

Corresponding author: Michael S. Lan, mlan@chnola-research.org.

Received 22 February 2010 and accepted 20 September 2010. Published ahead of print at <http://diabetes.diabetesjournals.org> on 28 September 2010. DOI: 10.2337/db10-0255.

© 2011 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. See <http://creativecommons.org/licenses/by-nc-nd/3.0/> for details.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

These findings could account for the differential insulin expression in human thymus that dictates subsequent tolerance induction.

RESEARCH DESIGN AND METHODS

Constructs. A PCR amplicon from the human insulin basal promoter region was cloned into a promoterless luciferase vector, pGL3-Basic (Promega) as pINS-basal. Based on this construct, six individual class I VNTR alleles, including their 5'-flanking regions of about 1.1 kb, were first PCR amplified from individual human genomic DNA samples and subcloned into the pINS-basal construct to create the luciferase reporter construct series referred to as TD, HM, PQ, PW, KE, and RC (genomic DNA samples were obtained from Human Biological Data Interchange, Philadelphia, PA). The insert sequence of these constructs corresponds to the same genomic DNA on human chromosome 11p15.5, which includes the 5'-flanking region, VNTR, and the basal insulin promoter. The primer sequences used for cloning are listed as follows: INSpF, GACTCGAGACAGCAGCGCAAAGAGCCC; INSpR, CTTAAGCTTG CAGCCTGTCTGGAGGGC; VNTR-5', TACACGCTGCTGGGATCCTGGATCT; VNTR-3', CTTGGAACAGACCTGCTTGA. The luciferase reporter vector for the class III VNTR was constructed from a λ HI-3 clone (18). An artificial class III construct was assembled by linking four class I VNTRs 5'-upstream of the human insulin basal promoter sequentially. The 1.1 kb sequence 5'-upstream of the VNTR was deleted during the cloning process. Both ends of the artificial class III construct were confirmed by DNA sequence analysis. All of the PCR-generated amplicons were fully sequenced before use in any functional assay. For the heterologous promoter study, the insulin basal promoter was replaced by a SV40 promoter or SV40 promoter and enhancer for use in the reporter assays. A complete set of reporter construct without the 1.1 kb region were also constructed and described in Fig. 3A.

Luciferase reporter assay. HIT cell line (ATCC) was grown in low-glucose Dulbecco's modified Eagle's medium supplemented with 10% FBS. The human thymic epithelial cell (hTEC) line derived from cortical epithelium (19) was cultured in RPMI 1,640 medium supplemented with 10% FBS. The primary thymic epithelial cells from E16–18 rat fetal thymi were isolated and cultured as previously described (20). Luciferase assays were performed in a 24-well plate format using Lipofectamine 2000 according to the manufacturer's instructions (Invitrogen). For all of the constructs assayed, equal moles of promoter equivalent were used (~50 ng). Luciferase activities were normalized with a cotransfected null-Renilla plasmid (2 ng) (Promega). Since endogenous AIRE protein is not detected in the hTEC cell line (21) or transfection-competent rat primary thymic epithelial cells, an exogenous source of AIRE or control β -galactosidase (β -gal) was introduced into the cells using a mammalian expression vector (50 ng). Several passages of freshly cultured primary cells were maintained until the cells became competent for transfection. Luciferase activities were measured using the Dual-Glo Luciferase Assay kit (Promega). Activities were based on three to six experiments, performed in duplicate or triplicate wells for each sample, and either one representative result or an average of all the experiments was presented.

Western blot analysis. In parallel with the luciferase reporter gene assay, separate wells of samples were collected to verify the expression of AIRE protein with a goat anti-AIRE antibody that recognizes the epitope 159–172 amino acids in AIRE protein (Abcam). In addition, a second rabbit antibody recognizes the epitope 246–545 amino acids in AIRE protein (Santa Cruz) was used to detect the immunoprecipitated AIRE proteins. The molecular weight of wild-type AIRE is 60 kDa. Signals were visualized using a chemiluminescence kit (Pierce). Equal sample loading on all the Western blots was confirmed by an anti- α -actin antibody (Sigma) that recognizes a protein with a molecular weight of 43 kDa.

INS-VNTR pull-down experiment. Human TEC cells were infected with an Ad-AIRE viral vector for 2 days. Total cell lysates were prepared in a pull-down buffer (10 mmol/l Tris buffer, pH 7.5 containing 50 mmol/l NaCl, 1 mmol/l EDTA, and 1 mmol/l DTT) with three freeze-thaw cycles. AIRE protein expression was detected by Western blot analysis using an anti-AIRE antibody (Abcam). The insulin basal promoter probe (0.4 kb), INS-VNTR-class I basal (2.0 kb), and III-basal (3.5 kb) probes were prepared from the reporter vectors after restriction digestion with *KpnI* and *HindIII*. DNA probes were radiolabeled at the 5'-terminal phosphate with T4 polynucleotide kinase and γ -³²P-ATP (Perkin Elmer Life Sciences). AIRE expressing hTEC cell lysate (100 μ g) was mixed with 50,000 cpm of each individual probe in pull-down buffer for 6 h at room temperature with constant rotation. Normal goat IgG or goat anti-AIRE antibody was added to the mixture for an overnight incubation. The following morning, 50 μ l of 50% slurry of salmon sperm DNA-treated protein G-beads was added to the mixture for a 1-h incubation before being subjected to three washes with pull-down buffer. The washed beads were digested with proteinase K at

55°C for 1 h and subjected to 1% agarose gel separation. The agarose gel was dried and exposed to X-ray film for probe detection.

RESULTS

Molecular cloning of six novel class I VNTR alleles and the restriction profiles of the class I and III alleles. *IDDM2* or *INS* is one of the multiple loci that encode susceptibility to type 1 diabetes, a prototype of human autoimmunity (10,22–24). *INS*-VNTR situated 5'-upstream of the *INS* promoter, is regarded as a causal variant at the locus after a close correlation was found between VNTR haplotypes and insulin thymic expression (7,8). However, comprehensive resequencing of the locus revealed two polymorphisms that are potentially functional and genetically indistinguishable from the VNTR (25–27). Therefore, direct evidence is required to support that the VNTR allele controls insulin transcript expression in the human thymus. To establish a functional role of the VNTR in the differential thymic insulin gene expression, six individual human genomic DNA samples were genotyped using the *PstI*⁺ polymorphism. The *PstI*⁺ (uncut by *PstI*) polymorphism is in strong linkage disequilibrium with the diabetes-predisposing class I allele and *PstI*⁻ (cut by *PstI*) with diabetes-protective class III VNTR haplotypes (10). The insulin gene locus at 11p15.5 spanning the VNTR region was PCR amplified and digested with *PstI* to determine the haplotype. Class I homozygous individuals will display a single 290 bp band; class III homozygous individuals will display two bands 200 and 90 bp; and class I/III heterozygous individuals display three bands 290, 200, and 90 bp in size. Of the six DNA samples analyzed, samples TD and KE are heterozygous class I/III and the other four are homozygous class I (Fig. 1A). To determine the relative repeat number of each individual DNA sample, a *NcoI/BglII* double restriction endonuclease digest was performed. Each sample contains a 450, 230, and 100 bp band as shown in the schematic diagram in Fig. 1C. Using the approximate size of the class I VNTR band, the number of repeat units was estimated (Fig. 1D). Finally, sequence analysis revealed the exact number of repeat units contained in each of the unique human DNA samples and are shown according to the convention established by Owerbach and Gabbary (24) (Fig. 1E).

IDDM2-VNTR locus transduces a quantitative insulin expression signal in the presence of AIRE. To directly dissect any transcriptional activity contributed by the VNTR alleles, we made a set of reporter constructs bearing the VNTR sequences and excluding the two other putative causal variants, -23 *Hph* I and +1,140 A/C in the *INS* locus (Figs. 1 and 3A). Because of the selective tissue expression of the insulin gene, we first tested the class I or class III VNTR-basal insulin luciferase reporter constructs in a hamster pancreatic β -cell line, HIT. Consistent with previous reports, the class I VNTR construct activity was approximately twofold higher than the class III VNTR constructs in the HIT cells (Fig. 2A, left panel). The luciferase data confirmed the previous observation that the type 1 diabetes susceptible shorter class I VNTR (most commonly 30–40 repeats) was associated with a 1.5- to 3-fold higher insulin expression than the class III VNTR in the pancreas (5,10), although inconsistent results have been reported (4) and discussed (10).

We artificially overexpressed AIRE in HIT cells to examine the responses of class I and III VNTRs. In the presence of AIRE, both classes of VNTR display similar activities, with the class I VNTR being slightly higher than

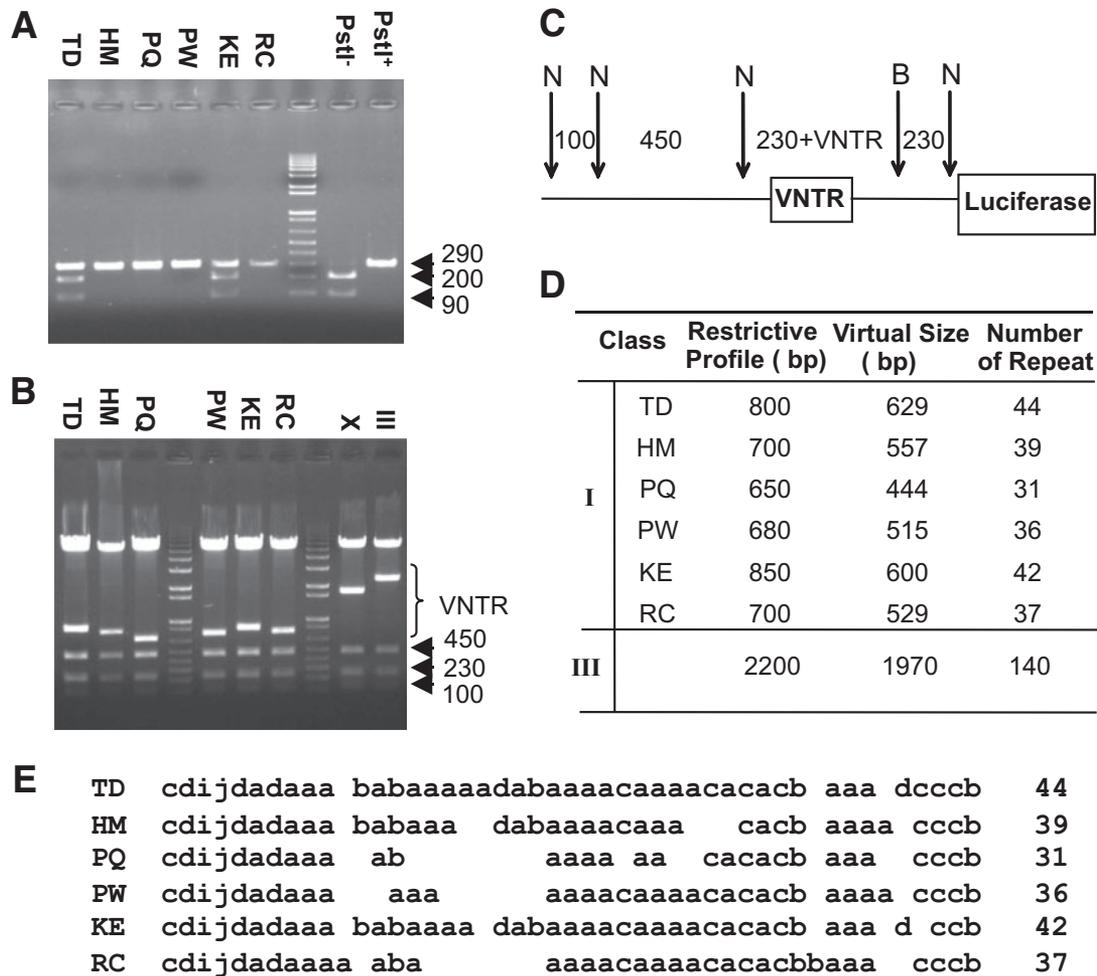


FIG. 1. Six novel class I VNTR alleles and the restriction profile of the class I and III alleles of the constructs. **A:** Six human genomic DNA samples were genotyped using the +1,127 *PstI* polymorphism. *PstI*⁺ (uncut by *PstI*) and *PstI*⁻ (cut by *PstI*) variants are in strong linkage disequilibrium with diabetic class I and nondiabetic class III VNTR haplotypes, respectively (10). By designing a pair of primers (5'-TAAATGCAGAAGCGTG GCATTGTGGAAC-3' and 5'-CTGCATGCTGGGCCTGGCCGG-3'), the PCR products amplified from genomic DNA were digested with *PstI* to generate a single band of 290 bp from class I homozygotes; two bands of 200 bp and 90 bp from class III homozygotes; and three bands of 290 bp, 200 bp, and 90 bp from I/III heterozygotes. TD and KE are heterozygous (I/III) and the other four are homozygous class I. *PstI*⁺ and ⁻ are control DNA of class I and class III homozygotes, respectively. **B:** Restriction profile of the VNTR alleles contained in the reporter gene constructs. *NcoI* and *BglIII* were used. **C:** Schematic explanation for the profile shown on the gel **B** or quantitated in table **D**. The band sizes are indicated by base pairs and marked by N (*NcoI*) and B (*BglIII*); 230 + VNTR in **C** corresponds to VNTR in **B** or restriction profile in **D**. **D:** Estimation of the sizes of individual VNTR alleles. **E:** Sequence of six class I VNTR alleles in the format of repeat unit array and according to the convention set by previous reference (24). The number of repeat units is shown to the right of each sequence.

the class III VNTR. Recently, correlation studies suggested an involvement of the autoimmune regulator (AIRE) in the promiscuous expression of insulin in human thymus (28,29), whereas thymus-specific deletion of insulin induces autoimmune diabetes (30). Nevertheless, how the *IDDM2* locus transduces a differential expression signal in thymus in the context of AIRE remains to be elucidated. In contrast to the pattern in pancreas, insulin expression in thymus was reversed, with an average threefold increase of class III over class I alleles, as represented by the luciferase activity driven by six different class I and one class III haplotypes in a nontransformed human thymic epithelial cell line (hTEC) (19) (Fig. 3B and C). This observation suggests that the diabetes susceptible class I VNTR might predispose individuals to type 1 diabetes by lowering insulin expression in the thymus and rendering less tolerance induction. Evidently, without the other two polymorphisms, the VNTR alone accounts for quantitative expression of insulin, which is consistent with the report that -23 *Hph* I, one of the two SNPs, did not play a

significant role in thymus (31). Additionally, we confirmed the independent effect of the VNTR using shorter constructs lacking the 1.1 kb 5'-flanking regions of the polymorphism (Fig. 3F and G). In addition to the hTEC line, primary epithelial cells dissected from fetal rat thymus (E16-18) were transfected with the same set of constructs, and the results confirmed the expression profile observed in the human cell line (Fig. 3D and E). Remarkably, the observed expression pattern persisted only in the presence of AIRE (Fig. 3H). When a control β -gal expression plasmid was administered, the differential expression signal transduced by the INS-VNTR/*IDDM2* locus was totally lost and the insulin expression driven by either allele decreased to a basal level (Fig. 3B and D). This line of evidence warrants further investigation of AIRE regulation in the context of the VNTR for insulin quantitative expression in thymic epithelial cells. Since there is only one natural class III DNA construct available for the comparison study with the six class I VNTR constructs, we further assembled an artificial class III construct (~162 tandem

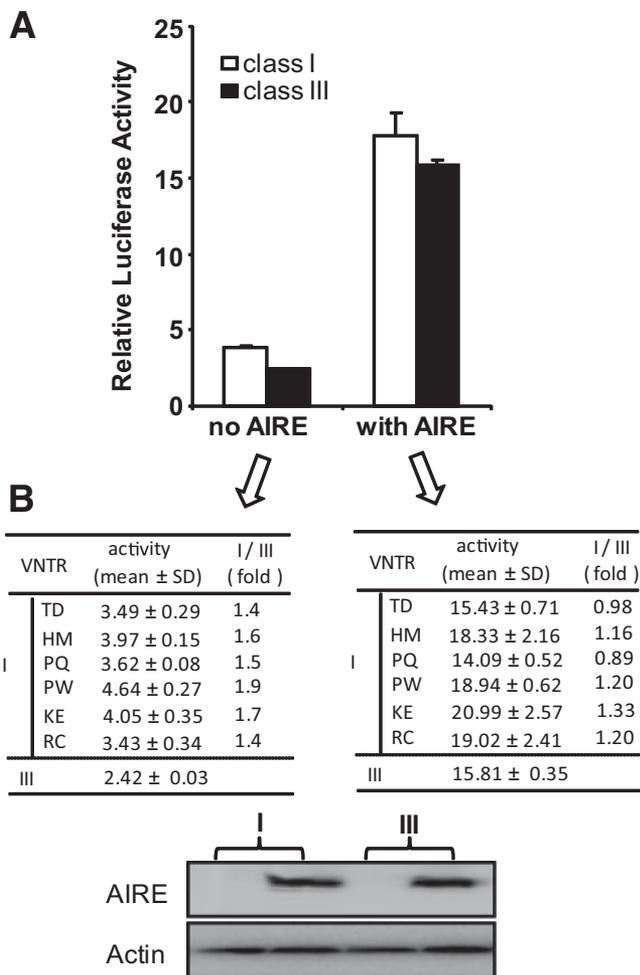


FIG. 2. Pancreas-based insulin expression pattern associated with the *IDDM2*-VNTR locus. In comparison with the activity in thymus, the pancreatic insulin expression pattern, represented by the luciferase reporter activity, was confirmed in a HIT cell line (A and B). An average of 1.58-fold increase was observed in reporter activities driven by class I over class III VNTR haplotypes. The class I activity was expressed as an average of six constructs. The class III activity was expressed as an average of three experiments from one construct. The constructs assayed are the same set of reporter constructs as those used in the hTEC line or rat primary cells in Fig. 3. Both class I and III reporter constructs respond to the addition of AIRE similarly.

repeats) based upon the assumption that the differential transcriptional activities from class III versus class I VNTR are due to the number of tandem repeats. As shown in supplementary Fig. 2 in the online appendix available at <http://diabetes.diabetesjournals.org/cgi/content/full/db10-0255/DC1>, artificial class III displays higher activity than the natural class III construct (140 tandem repeats). Additionally, this construct also displays a higher activity in the no AIRE setting, probably caused by the deletion of the 5'-upstream 1.1 kb sequence of the VNTR (similar to Fig. 3F).

VNTR responded sensitively and differentially to AIRE in a human thymic epithelial cell line. Next, we determined how the class I or class III VNTR allele responded to an increasing dosage of AIRE expression plasmid. Each VNTR haplotype transduced a sensitive and differential signal in the hTEC line. As little as 10 ng of the AIRE expression plasmid generated a differential response between class I (PQ) and class III VNTRs (Fig. 4A). In the presence of increasing amounts of AIRE protein, the insulin expression driven by the VNTR increased until a

plateau was reached. The difference between the two classes of haplotypes was consistently about threefold at each dose tested, except at the zero dose (Fig. 4A). The level of AIRE protein expression was verified with an AIRE antibody by Western blot (Fig. 4B). Additionally, we chose the natural class III and the artificial class III paired with two other class I (HM and RC) constructs to perform the dosage response curves (supplementary Fig. 3). Similarly, class I VNTRs response to AIRE is much lower than both natural and artificial class III VNTRs.

AIRE regulates the insulin promoter activity mediated through the INS-VNTR element. In the presence of AIRE, the insulin basal promoter alone (−365 bp) displays relatively low activity in the human TEC and rat primary TEC (Fig. 5A and B). The class I VNTR is two to threefold higher relative to the insulin basal promoter activity, whereas the class III VNTR increases seven to ninefold over the insulin basal promoter activity, suggesting that the AIRE transcriptional regulator could activate the INS-VNTR sequence in the context of the insulin basal promoter. To validate the interaction between the AIRE protein and the VNTR sequence, we performed a DNA-protein pull-down experiment using Adenovirus-AIRE transduced hTEC cell lysate and radiolabeled DNA probes (50,000 cpm) derived from the insulin basal promoter, class I-VNTR-INS-basal promoter, and class III-VNTR-INS-basal promoter sequences (Fig. 5C). Both class I and III DNA probes were pulled-down by the anti-AIRE antibody. A very weak signal, if any, was detected with the insulin basal promoter probe. This result suggests that AIRE protein can interact directly or indirectly through the recruitment of a transcriptional complex to bind the VNTR sequences.

The VNTR is located 365 bp 5'-upstream of the human insulin basal promoter. We designed experiments to evaluate whether the VNTR functions in an insulin promoter-specific fashion or if it could function effectively with a heterologous promoter and/or in combination with other enhancer elements. A representative class I or III VNTR was subcloned 5'-upstream of the SV40 promoter driven luciferase vector or the SV40 promoter-enhancer driven luciferase vector. Cotransfection of the luciferase constructs with an AIRE expression vector into the human TEC showed no effect to the luciferase activity with either class I or III VNTR in conjunction with the SV40 promoter. The characteristic activation pattern of class I and III VNTRs seen when linked directly to the insulin basal promoter disappeared when the insulin basal promoter was replaced by the SV40 promoter and/or SV40 promoter/enhancer (Fig. 5D and E). This result indicates that the VNTR sequence has to function in conjunction with the insulin basal promoter to respond to AIRE.

In this paper, we have revealed a novel role of the AIRE protein with the VNTR region in the establishment of differential tissue regulation of the insulin mRNA in the thymus.

DISCUSSION

Type 1 insulin-dependent diabetes results from the autoimmune destruction of the insulin producing β cells of the pancreas. Type 1 diabetes is a polygenic disease and multiple susceptibility loci have been mapped (32). A strong correlation between the variable numbers of tandem repeat (VNTR) or *IDDM2* locus and the levels of insulin mRNA expression in the thymus have been clearly established (4,5,7,8). The short class I allele is predisposing to type 1 diabetes and is associated with lower thymus

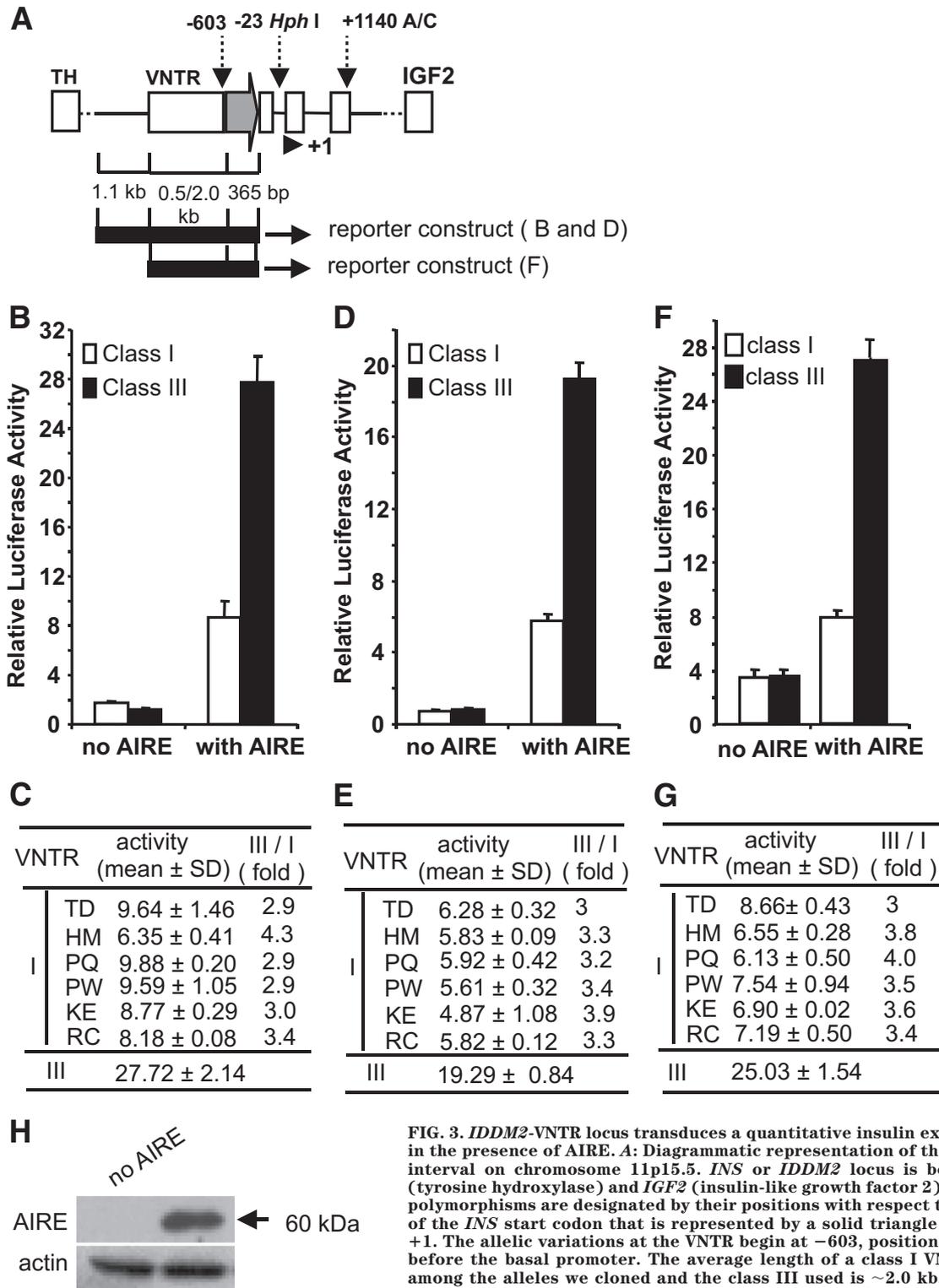


FIG. 3. *IDDM2*-VNTR locus transduces a quantitative insulin expression signal in the presence of AIRE. **A:** Diagrammatic representation of the *TH-INS-IGF2* interval on chromosome 11p15.5. *INS* or *IDDM2* locus is bordered by *TH* (tyrosine hydroxylase) and *IGF2* (insulin-like growth factor 2) loci. The three polymorphisms are designated by their positions with respect to the first base of the *INS* start codon that is represented by a solid triangle and marked as +1. The allelic variations at the VNTR begin at -603, positioned immediately before the basal promoter. The average length of a class I VNTR is ~0.5 kb among the alleles we cloned and the class III used is ~2.0 kb. A gray arrow depicts the basal insulin promoter region (365 bp) and open boxes denote the

three insulin exons. The **black bold lines** represent the regions cloned into the luciferase reporter constructs including the VNTR, its 5'-flanking 1.1 kb region, the basal insulin promoter, and excluding the other two polymorphisms. The longer version bearing the 5'-flanking 1.1 kb in addition to the VNTR region and basal promoter construct, as well as a shorter construct without the 1.1 kb 5'-flanking region, were analyzed. **B:** A luciferase reporter gene is driven by class I or class III VNTR with insulin basal promoter (50 ng) and transfected in a hTEC line. Six class I and one class III haplotypes were tested. The relative activities both without (β -gal control) and with exogenous AIRE (50 ng) are shown. The class I activity was expressed as an average of six constructs. The class III activity was expressed as an average of three experiments from one construct. **C:** Detailed activities of each individual VNTR allele in the presence of AIRE were tabulated. **D** and **E:** The same set of constructs was assayed using rat primary thymic epithelial cells following the same strategy as in **B** and **C**. **F** and **G:** To rule out the possibility that the upstream 5'-flanking genomic region (~1.1 kb) might contribute to the differential insulin expression in thymus, another set of constructs was made by removing this region (**A**) and assaying in the hTEC line. The average difference in thymic expression is consistently a threefold increase of class III over I VNTR haplotypes. **H:** Western blot analysis to verify AIRE expression in the transfected cell line. As a loading control, α -actin was used.

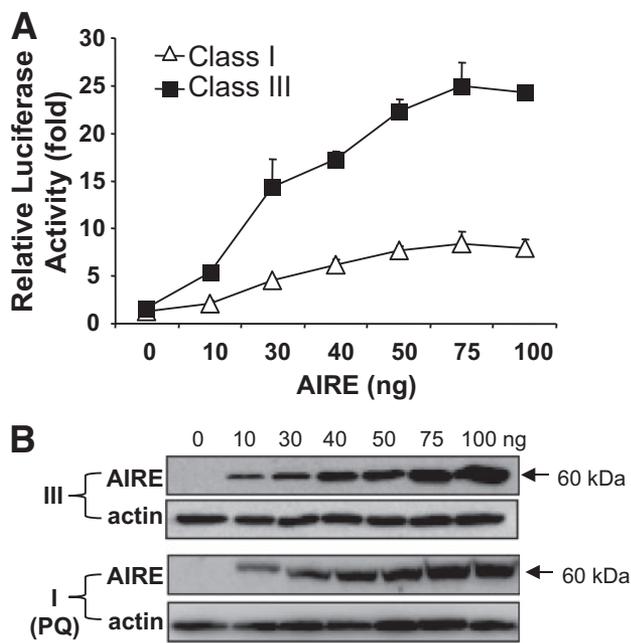


FIG. 4. VNTR responded sensitively and differentially to AIRE in a human thymic epithelial cell line. **A:** The VNTR responded to AIRE in a hTEC line. A class I VNTR (PQ) and a class III VNTR reporter construct (50 ng) were cotransfected with AIRE expression vector or control vector β -gal DNA into a hTEC line. The amount of AIRE expression vector added to each sample of a 24-well plate format is indicated. **B:** Western blot analysis was performed to confirm that the increasing expression of AIRE protein corresponded to the amounts of added expression vector. Equal loading of each sample was confirmed by incubation with an anti- α -actin antibody.

insulin expression, whereas the class III allele is protective and associated with three to fourfold higher thymic insulin expression. How this region contributes to this thymic expression pattern has not been directly tested. The autoimmune regulator, AIRE, has been shown to direct the expression of hundreds of tissue-specific antigens in the thymus and is crucial for establishment of immune tolerance (16). To definitively establish a role of the VNTR in insulin gene regulation in thymus, both class I and III VNTRs were cloned and tested in reporter gene studies.

Our cloning strategy spanning the VNTR region of the *INS* locus first excluded the two additional polymorphisms (-23 Hph I and $+1,140$ A/C) that are present within the *IDDM2* locus and may contribute to the differential insulin levels in the thymus. Without these two polymorphisms, the VNTR alone accounts for quantitative expression of insulin. We also removed a 1.1 kb 5'-upstream sequence from the VNTR locus. Reporter constructs that lacked the 1.1 kb 5'-upstream genomic region (to the VNTR) showed a similar promoter activity as the construct that included this region. These results further confirmed that the thymic insulin gene regulation by AIRE is mediated by the VNTR.

The transfection studies with the *INS*-VNTR class I and III constructs unequivocally show that AIRE controls differential insulin gene expression through the *INS*-VNTR region. To study the role of the AIRE protein in regulating the insulin promoter via the VNTR region, we further examined the physical interaction between AIRE and the VNTR. The AIRE protein has been reported to bind to target DNA as a homodimer or a tetramer, but not a monomer. An oligonucleotide library screen showed that AIRE protein bound to two nucleotide sequences,

TTATTA and two tandem repeats of ATTGGTA (33). It is difficult to envision that the AIRE protein recognizes a specific binding sequence present in hundreds or thousands of target gene promoter sequences. Particularly, the AIRE binding site is not found in the *INS*-VNTR region. Additionally, our data show that the insulin basal promoter region in conjunction with the VNTR is required for regulation by AIRE protein. Intriguingly, linkage of the *INS*-VNTR class I or class III VNTR region with a heterologous promoter sequence resulted in the complete loss of AIRE transcriptional regulation. Therefore, it is likely that AIRE must interact with a protein or complex that is formed on the basal insulin promoter to modulate insulin gene expression in the thymus. We speculate that AIRE's binding to the target genes may rely on its interaction with a specific chromatin structure. Other studies have shown that AIRE recruits CREB-binding protein, positive transcription elongation factor b, and interacts with target genes lacking methylated histone H3K4 to regulate tissue-restricted antigen expression in the thymus (34–36). Recent studies reported that AIRE protein could partner with four major functional classes: nuclear transport, chromatin binding/structure, transcription, and premRNA processing for targeting and inducing peripheral tissue self-antigen (37). Additionally, DAXX represents another newly discovered AIRE-interacting protein (38). A recent review article described the molecular mechanisms of central tolerance, focusing on the transcriptional regulation by AIRE (39). Although this paper discussed the general mechanisms of how the AIRE protein could interact with multiple target genes, it suggests a sharing of a common transcriptional complex to exert its activation activities. The insulin basal promoter has been identified as a target for AIRE regulation, but these constructs always lacked the VNTR region. In response to AIRE, our data show that in the presence of the VNTR sequence (as in the human insulin gene locus) insulin basal promoter activity is significantly enhanced in hTECs. Labeling of the basal insulin promoter, class I and class III basal insulin promoter regions and incubation with an AIRE-expressing hTEC cell lysates unambiguously showed preferential binding of the AIRE protein with the VNTR containing fragments (Fig. 5C). The tandem repeat regions of the insulin VNTR have been shown to form G-quadruplex structures (40). The enhanced regulation of the class III constructs is most likely due to the larger number of tandem repeats forming a more stable intramolecular secondary structure that is required for AIRE protein binding and formation of a transcriptional complex. We have constructed an artificial class III VNTR from four different copies of class I VNTR sequences derived from patient samples. As expected, the artificial class III VNTR (~162 tandem repeats) demonstrated a higher transcriptional activity as compared with the natural class III VNTR (140 tandem repeats) and exhibited approximately threefold higher activity than class I VNTR (supplementary Fig. 2). Together, we show that the AIRE protein is capable of selectively pulling-down the VNTR insulin basal promoter sequence. Definitively, these data support the physical interaction between AIRE and the *INS*-VNTR, but we cannot rule out the involvement of other cellular components that may be involved in mediating this interaction.

In summary, we reveal a dual control mechanism governing the insulin quantitative expression in thymus that may account for part of the genetic predisposition observed in type 1 diabetic patients. AIRE-mediated differ-

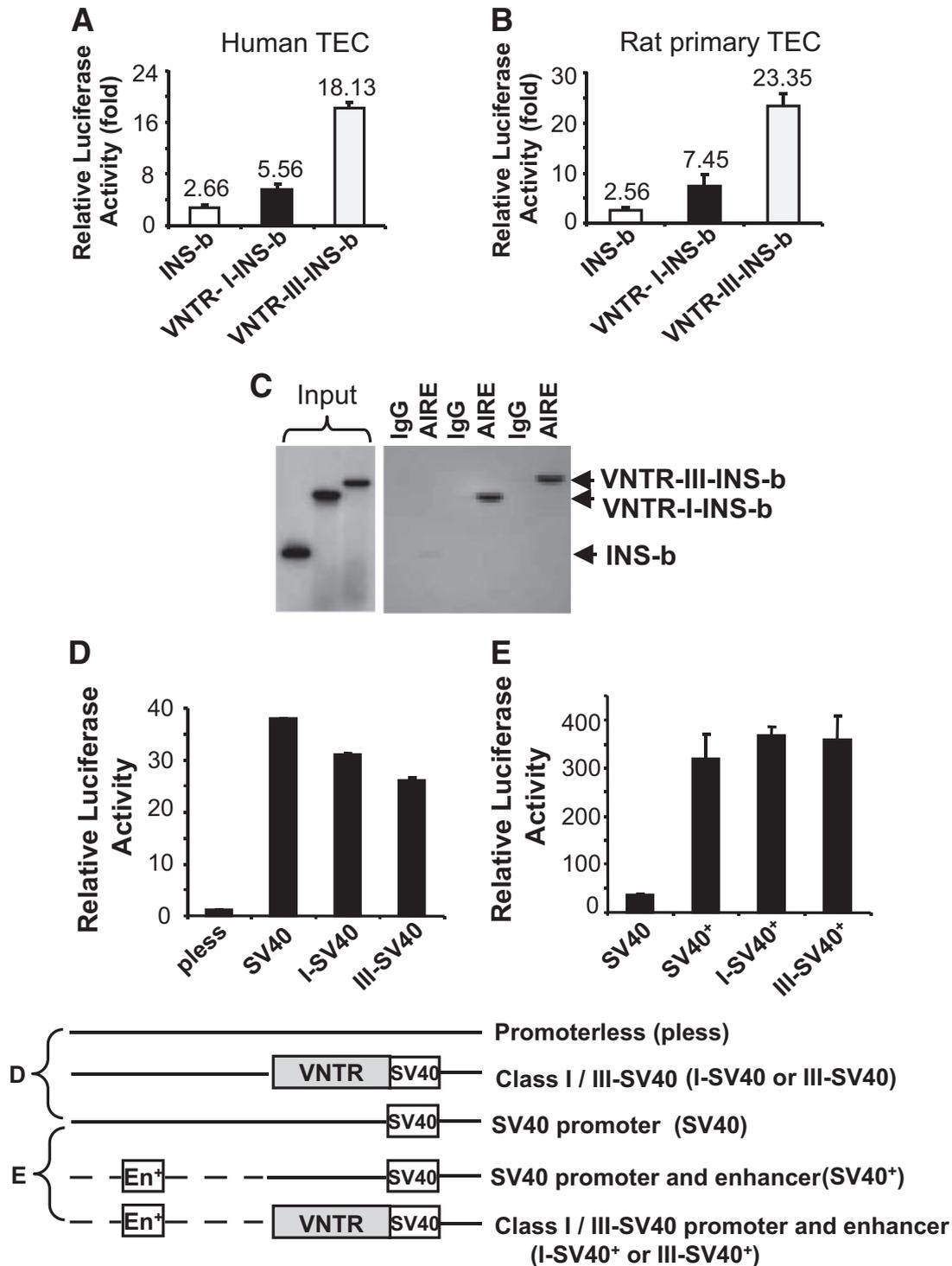


FIG. 5. AIRE regulates the insulin promoter activity mediated through the INS-VNTR element. *A* and *B*: Insulin basal promoter, VNTR-class I-INS-basal and VNTR-class III-INS-basal vectors demonstrated differential promoter activities in both hTEC and rat primary TEC in response to AIRE. *C*: DNA probes derived from insulin basal promoter (0.4 kb), VNTR-class I-INS-basal (2.0 kb), and VNTR-class III-INS-basal (3.5 kb) were radiolabeled with T4 polynucleotide kinase and γ -³²P-ATP as shown in the input lanes. An Ad-AIRE transduced hTEC total lysate (100 μ g) was mixed with 50,000 cpm of each radiolabeled DNA probe for 6 h at room temperature and subsequently added 1 μ g of goat anti-AIRE antibody for an overnight incubation. The immune complex was precipitated by salmon sperm DNA-treated protein G-agarose beads, washed, and eluted from beads for separation. Normal goat IgG was used as a negative control. The immunoprecipitated AIRE complexes were verified using an anti-AIRE antibody. *D*: VNTR linked to a heterologous promoter failed to respond to AIRE. The insulin basal promoter was replaced by a SV40 promoter in a reporter construct. Transfection of the construct lacking the insulin basal promoter with the AIRE expression vector in human TEC was measured for luciferase reporter activity. *E*: Addition of a SV40 enhancer to the construct greatly enhanced SV40 promoter activity, but the class I or III VNTR had no effect on the promoter/enhancer combination. Transfection efficiency was normalized with the null-*Renilla* luciferase reporter vector.

ential regulation of insulin gene expression via the VNTR locus in the thymus may influence tolerance and susceptibility to type 1 diabetes. A similar mechanism was also

reported for myasthenia gravis, another autoimmune disease (21). These findings may together suggest a commonly adopted regulation strategy that, in addition to

genetic variations endowed by each of the individual genes, encode hundreds and/or even more than a thousand tissue-restricted antigens in the human thymus (41).

ACKNOWLEDGMENTS

This study was funded by The Research Institute for Children, Children's Hospital at New Orleans and a National Institutes of Health award to M.S.L. (DK-061436).

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

C.Q.C. researched data, wrote the manuscript, provided reagent, and contributed to discussion. T.Z. researched data. M.B.B. researched data and reviewed/edited the manuscript. M.G. reviewed/edited the manuscript, provided reagent, and contributed to discussion. M.S.L. researched data, wrote the manuscript, and contributed to discussion.

The authors thank Dr. S.H. Pincus (The Research Institute for Children, New Orleans, LA) for insightful discussion, Dr. D. Owerbach (Baylor College of Medicine, Houston, TX) for class III VNTR plasmid, and Dr. P. Peterson (Tartu University, Tartu, Estonia) for the AIRE cDNA plasmid.

REFERENCES

- Pugliese A. The insulin gene in type 1 diabetes. *IUBMB Life* 2005;57:463–468
- Zhang L, Nakayama M, Eisenbarth GS. Insulin as an autoantigen in NOD/human diabetes. *Curr Opin Immunol* 2008;20:111–118
- Bell GI, Selby MJ, Rutter WJ. The highly polymorphic region near the human insulin gene is composed of simple tandemly repeating sequences. *Nature* 1982;295:31–35
- Kennedy GC, German MS, Rutter WJ. The minisatellite in the diabetes susceptibility locus IDDM2 regulates insulin transcription. *Nat Genet* 1995;9:293–298
- Lucassen AM, Screaton GR, Julier C, Elliott TJ, Lathrop M, Bell JI. Regulation of insulin gene expression by the IDDM associated, insulin locus haplotype. *Hum Mol Genet* 1995;4:501–506
- Lucassen AM, Julier C, Beressi JP, Boitard C, Froguel P, Lathrop M, Bell JI. Susceptibility to insulin dependent diabetes mellitus maps to a 4.1 kb segment of DNA spanning the insulin gene and associated VNTR. *Nat Genet* 1993;4:305–310
- Pugliese A, Zeller M, Fernandez A Jr, Zalberg IJ, Bartlett RJ, Ricordi C, Pietropaolo M, Eisenbarth GS, Bennett ST, Patel DD. 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;15:293–297
- Vafiadis P, Bennett ST, Todd JA, Nadeau J, Grabs R, Goodyer CG, Wickramasinghe S, Colle E, Polychronakos C. Insulin expression in human thymus is modulated by INS VNTR alleles at the IDDM2 locus. *Nat Genet* 1997;15:289–292
- Bell GI, Horita S, Karam JH. A polymorphic locus near the human insulin gene is associated with insulin dependent diabetes mellitus. *Diabetes* 1984;33:176–183
- Bennett ST, Lucassen AM, Gough SCI, Powell EE, Undlien DE, Pritchard LE, Merriman ME, Kawaguchi Y, Dronsfield MJ, Pociot F, Nerup J, Bouzekri N, Cambon-Thomsen A, Ronningen KS, Barnett AH, Bain SC, Todd JA. Susceptibility to human type 1 diabetes at IDDM2 is determined by tandem repeat variation at the insulin gene minisatellite locus. *Nat Genet* 1995;9:284–292
- Bennett ST, Wilson AJ, Cucca F, Nerup J, Pociot F, McKinney PA, Barnett AH, Bain SC, Todd JA. IDDM2-VNTR encoded susceptibility to type 1 diabetes: Dominant protection and parental transmission of alleles of the insulin gene-linked minisatellite locus. *J Autoimmun* 1996;9:415–421
- Ahonen P. Autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED): autosomal recessive inheritance. *Clin Genet* 1985;27:535–542
- Ahonen P, Myllarniemi S, Sipilä I, Perheentupa J. Clinical variation of autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED) in a series of 68 patients. *N Engl J Med* 1990;322:1829
- Finnish-German APECED Consortium. An autoimmune disease, APECED, caused by mutations in a novel gene featuring two PHD-type zinc-finger domains. The Finnish-German APECED Consortium Autoimmune Polyendocrinopathy-Candidiasis-Ectodermal Dystrophy. *Nat Genet* 1997;17:399–403
- Nagamine K, Peterson P, Scott HS, Kudoh J, Minoshima S, Heino M, Kudoh J, Shimizu N, Antonarakis SE, Krohn KJ. Positional cloning of the APECED gene. *Nat Genet* 1997;17:393–398
- Derbinski J, Gabler J, Brors B, Tierling S, Jonnakuty S, Hergenahn M, Peltonen L, Walter J, Kyewski B. Promiscuous gene expression in thymic epithelial cells is regulated at multiple levels. *J Exp Med* 2005;202:33–45
- Paquette J, Varin DSE, Hamelin CE, Hallgren A, Kampe O, Carel JC, Perheentupa J. Deal CL Risk of autoimmune diabetes in APECED: association with short alleles of the 5' insulin VNTR. *Genes Immun*. 2010;11:590–597
- Owerbach D, Aagaard L. Analysis of a 1963-bp polymorphic region flanking the human insulin gene. *Gene* 1984;32:475–479
- Fernandez E, Vicente A, Zapata A, Brera B, Lozano JJ, Martinez C, Toribio ML. Establishment and characterization of cloned human thymic epithelial cell lines. Analysis of adhesion molecule expression and cytokine production. *Blood* 1994;83:3245–3254
- Palumbo MO, Levi D, Chentouti AA, Polychronakos C. Isolation and characterization of proinsulin-producing medullary thymic epithelial cell clones. *Diabetes* 2006;55:2595–2601
- Giraud M, Taubert R, Vandiedonck C, Ke X, Levi-Strauss M, Pagani F, Baralle FE, Eymard B, Tranchant C, Gajdos P, Vincent A, Willcox N, Beeson D, Kyewski B, Garchon HJ. An IRF8-binding promoter variant and AIRE control CHRNA1 promiscuous expression in thymus. *Nature* 2007;448:934–937
- Todd JA, Walker NM, Cooper JD, Smyth DJ, Downes K, Plagnol V, Bailey R, Nejentsev S, Field SF, Payne F, Lowe CE, Szeszkó JS, Hafler JP, Zeitels L, Yang JH, Vella A, Nutland S, Stevens HE, Schuilenburg H, Coleman G, Maisuria M, Meadows W, Smink LJ, Healy B, Burren OS, Lam AA, Ovington NR, Allen J, Adlem E, Leung HT, Wallace C, Howson JM, Guja C, Ionescu-Tirgoviste C; Genetics of Type 1 Diabetes in Finland, Simmonds MJ, Heward JM, Gough SC, Wellcome Trust Case Control Consortium DDB, Wicker LS, Clayton DG. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. *Nat Genet* 2007;39:857–864
- Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447:661–678
- Owerbach D, Gabbay KH. The search for IDDM susceptibility genes. *Diabetes* 1996;45:544–551
- Barratt BJ, Payne F, Lowe CE, Hermann R, Healy BC, Harold D, Concannon P, Gharani N, McCarthy MI, Olavesen MG, McCormack R, Guja C, Ionescu-Tirgoviste C, Undlien DeE, Ronningen KS, Gillespie KM, Tuomilehto-Wolf E, Tuomilehto J, Bennett ST, Clayton DG, Cordell HE, Todd JA. Remapping the insulin gene/IDDM2 locus in type 1 diabetes. *Diabetes* 2004;53:1884–1889
- Kralovicova J, Gaunt TR, Rodriguez S, Wood PJ, Day INM, Vorechovsky I. Variants in the human insulin gene that affect pre-mRNA splicing: Is -23HphI a functional single nucleotide polymorphism at IDDM2? *Diabetes* 2006;55:260–264
- Day IN, Rodriguez S, Kralovicova J, Wood PJ, Vorechovsky I, Gaunt TR. Questioning INS VNTR role in obesity and diabetes: subclasses tag IGF-INS-TH haplotypes; and -23 HphI as a STEP (splicing and translational efficiency polymorphism). *Physiol Genomics* 2006;28:113
- Sabater L, Ferrer-Francesch X, Sospedra M, Caro P, Juan M, Pujol-Borrell R. Insulin alleles and autoimmune regulator (AIRE) gene expression both influence insulin expression in the thymus. *J Autoimmun* 2005;25:312–318
- Taubert R, Schwendemann J, Kyewski B. Highly variable expression of tissue-restricted self-antigens in human thymus: implications for self-tolerance and autoimmunity. *Eur J Immunol* 2007;37:838–848
- Fan Y, Rudert WA, Grupillo M, He J, Sisinio G, Trucco M. Thymus-specific deletion of insulin induces autoimmune diabetes. *EMBO J* 2009;28:2812–2824
- Marchand L, Polychronakos C. Evaluation of polymorphic splicing in the mechanism of the association of the insulin gene with diabetes. *Diabetes* 2007;56:709–713
- Todd JA. Genetic analysis of type 1 diabetes using whole genome approaches. *Proc Natl Acad Sci U S A* 1995;92:8560–8565
- Kumar PG, Laloraya M, Wang CY, Ruan QG, Davoodi-Semiromi A, Kao KJ, She JX. The autoimmune regulator (AIRE) is a DNA-binding protein. *J Biol Chem* 2001;276:41357–41364
- Pitkanen J, Doucas V, Sternsdorf T, Nakajima T, Aratani S, Jensen K, Will H, Vahamurto P, Ollila J, Vihinen M, Scott HS, Antonarakis SE, Kudoh J, Shimizu N, Krohn K, Peterson P. The autoimmune regulator protein has

- transcriptional transactivating properties and interacts with the common coactivator CREB-binding protein. *J Biol Chem* 2000;275:16802–16809
35. Oven I, Brdickova N, Kohoutek J, Vaupotic T, Narat M, Peterlin BM. AIRE recruits P-TEFb for transcriptional elongation of target genes in medullary thymic epithelial cells. *Mol Cell Biol* 2007;27:8815–8823
36. Org T, Chignola F, Hetenyi C, Gaetani M, Rebane A, Liiv I, Maran U, Mollica L, Bottomley MJ, Musco G, Peterson P. The autoimmune regulator PHD finger binds to non-methylated histone H3K4 to activate gene expression. *EMBO Rep* 2008;9:370–376
37. Abramson J, Giraud M, Benoist C, Mathis D. Aire's partners in the molecular control of immunological tolerance. *Cell* 2010;140:123–135
38. Meloni A, Fiorillo E, Corda D, Incani F, Serra ML, Contini A, Cao A, Rosatelli MC. DAXX is a new AIRE-interacting protein. *J Biol Chem* 2010;285:13012–13021
39. Peterson P, Org T, Rebane A. Transcriptional regulation by AIRE: molecular mechanisms of central tolerance. *Nature Review-Immunology* 2008;8:948–957
40. Schonhofs JD, Bajracharya R, Dhakal S, Yu Z, Mao H, Basu S. Direct experimental evidence for quadruplex-quadruplex interaction within the human ILPR. *Nucleic Acid Res* 2009;37:3310–3320
41. Anderson MS, Venanzi ES, Klein L, Chen Z, Berzins SP, Turley SJ, von Boehmer H, Bronson R, Dierich A, Benoist C, Mathis D. Projection of an immunological self shadow within the thymus by the aire protein. *Science* 2002;298:1395–1401