The Autoimmune Regulator Directly Controls the Expression of Genes Critical for Thymic Epithelial Function

Qing-Guo Ruan, Kenneth Tung, Daniel Eisenman, Yulius Setiady, Sarah Eckenrode, Bing Yi, Sharad Purohit, Wei-Peng Zheng, Yan Zhang, Leena Peltonen and Jin-Xiong She

*J Immunol* 2007; 178:7173-7180; doi: 10.4049/jimmunol.178.11.7173
http://www.jimmunol.org/content/178/11/7173

---

**References**  This article *cites 39 articles*, 14 of which you can access for free at: [http://www.jimmunol.org/content/178/11/7173.full#ref-list-1](http://www.jimmunol.org/content/178/11/7173.full#ref-list-1)

**Subscription**  Information about subscribing to *The Journal of Immunology* is online at: [http://jimmunol.org/subscription](http://jimmunol.org/subscription)

**Permissions**  Submit copyright permission requests at: [http://www.aai.org/About/Publications/JI/copyright.html](http://www.aai.org/About/Publications/JI/copyright.html)

**Email Alerts**  Receive free email-alerts when new articles cite this article. Sign up at: [http://jimmunol.org/alerts](http://jimmunol.org/alerts)
The Autoimmune Regulator Directly Controls the Expression of Genes Critical for Thymic Epithelial Function

Qing-Guo Ruan,* Kenneth Tung,† Daniel Eisenman,* Yulius Setiady,† Sarah Eckenrode,* Bing Yi,* Sharad Purohit,* Wei-Peng Zheng,* Yan Zhang,* Leena Peltonen,‡ and Jin-Xiong She†,*‡

The autoimmune regulator (Aire)2 gene plays an essential role in negative selection of T cells and deletion of autoreactive T cells in the thymus. The defect in thymic selection in Aire−/− mice was attributed to the repressed expression of tissue-specific Ags in the thymic epithelial cells and defective Ag presentation; however, the molecular mechanism underlying these functions has been elusive. Using the chromatin immunoprecipitation technique, we demonstrate here that Aire binds in vivo to specific DNA sequence motifs and directly regulates thymic expression of genes important for thymic functions including expression of autoantigens, cytokines, transcription factors, and posttranslational modifiers. These results unambiguously established Aire as a key transcriptional regulator of the immune system. The Journal of Immunology, 2007, 178: 7173–7180.

The autoimmune regulator (Aire)2 gene is responsible for autoimmune polyglandular syndrome type 1 (APS1), an autosomal recessive syndrome characterized by multiple autoimmune diseases of the endocrine and nonendocrine organs (1–5). Aire knockout (Aire−/−) mice exhibited clinical phenotypes that resemble the human APS1, including lymphocytic infiltration in a number of tissues and production of autoantibodies against APS1 Ags (6–8). Using organ-specific TCR-transgenic models, it has been shown that Aire plays an essential role in thymic selection of T cells and deletion of autoreactive T cells (9), providing an immunological mechanism responsible for autoimmune caused by Aire deficiency. The defect in thymic selection in Aire−/− mice was attributed to the reduced expression of certain tissue-specific autoantigens (7). However, only a subset of tissue-specific Ags (TSA) is regulated by Aire in the medullary thymic epithelial cells (mTEC) (7, 8), raising the possibility that Aire may play additional roles in self-tolerance induction and its maintenance.

Despite the increased understanding of the immunological role of Aire, its molecular function remains elusive and somewhat controversial. Aire harbors two plant homeo domains that may have E3 ubiquitin ligase activity, a property that was demonstrated for the first plant homeo domain by in vitro ubiquitination assays (10). E3 ubiquitin ligases mediate the transfer of ubiquitin to specific proteins, resulting in proteosome degradation, down-regulation of cell surface receptors, and proteolysis-independent activities. Protein ubiquitination is involved in innate and adaptive immunity and a variety of other cellular processes. However, the sequence and structural characteristics of the Aire protein suggest a potential role in transcriptional regulation. Aire can bind to CREB-binding protein (11), which is, in turn, a cofactor for a large number of transcription factors. Aire also appears to enhance transcription from the IFN-β promoter and has transcriptional transactivating properties that are abrogated by pathogenic mutations (12). We have previously shown that Aire binds in vitro to specific DNA sequence motifs (13). However, the genes directly regulated by Aire have not yet been identified. Using chromatin immunoprecipitation (ChIP), we discovered a number of Aire target genes including autoantigens, transcription factors, growth factors, and posttranslational modifiers. These studies establish a molecular network for thymic selection and self tolerance with Aire as a master regulator.

Materials and Methods

Mice

C57BL/6 mice used in all ChIP studies were purchased from The Jackson Laboratory. Aire−/− mice obtained from Dr. L. Peltonen (6) have been maintained in our mouse colony at the Medical College of Georgia. The Aire−/− mouse strain was generated using 129svj embryonic stem cells and B6 blastocysts. These mice have been backcrossed to B6 mice for five generations.

ChIP

ChIP was performed as described previously with slight modifications (14). Briefly, thymi from regular C57BL/6 (B6) or Aire−/− B6 mice were fixed with 4% formaldehyde (Hydrol Chemical) overnight at 4°C and were disrupted by homogenization. The DNA fragments cross-linked to the protein were enriched by immunoprecipitation with one of the three Abs (anti-Aire polyclonal Ab produced in our laboratory (13, 15), a commercial Ab against Aire (Ab13573; Abcam), or IgG) or without any Ab. After reversal of the cross-links by DNA extraction, the enriched DNA was used for several purposes including ChIP PCR, ChIP DNA library construction, and ChIP microarray experiments.

ChIP microarray and ChIP PCR

The basic procedure for ChIP microarray analysis contains four major steps: ChIP DNA library, array printing, hybridization, and data analysis. To construct a ChIP DNA library, ChIP-enriched DNA fragments (50–100
ng) were amplified using a ligation-mediated PCR (LM-PCR). The amplified DNA products (3–5 µg) were then digested by restriction enzyme EcoRI engineered into the LM-PCR primer. The digested DNA was cloned into pUC18 vector. Inserts from 2000 clones were PCR amplified directly from 2 µl of bacterial culture in 96-well plates using M13 forward and reverse primers that anneal to the vector adjacent to the cloning site. The PCR products (10–15 µg) were then purified by ethanol precipitation, re-suspended in 25 µl of sodium phosphate buffer/SDS (150 mM (pH 8.5)/0.01%) with a final DNA concentration of 100–300 ng/µl. DNA was stored at −20°C until needed for printing. In addition to the ChIP DNA clones, we also amplified small fragments (~300 bp) containing a putative Aire-binding site in the promoter region of eight candidate genes: fatty acid-binding protein, IFN-γ, thyroglobulin (Tg), tryptophan hydroxylase (Tph), glutamic acid decarboxylase 65 (GAD65), insulin 1, insulin 2, and cytochrome P4501A2.

PCR-amplified DNA fragments (100–200 ng/µl) were spotted onto the poly-l-lysine precoated microscope slides using a MicroGrid II arrayer (BioRobotics). DNA was cross-linked to the slides using UV light. The microarrays can then be stored for use for several months in a desiccated chamber. The probes used for ChIP microarray hybridization were prepared using LM-PCR. ChIP DNA enriched by the anti-Aire Ab was labeled with the fluorescent dye Cy3, while DNA from negative control ChIP was labeled with Cy5. Each of the hybridizations was performed with a Cy3-labeled ChIP-enriched probe and a Cy5-labeled mock ChIP probe. Hybridization was accomplished under a coverslip in a specially designed hybridization chamber and allowed to proceed for 16 h at 65°C. Slides were then washed and spun-dried. The hybridized slides were scanned using an Affymetrix 418 Scanner (MWG Biotech). The Cy3 and Cy5 images were analyzed using the Scanalyze software, which defined the grid of spots and quantified the intensity of each spot and the surrounding background intensity. The background intensity was subtracted from the spot intensity. The net intensities of the two channels were normalized as described previously (16, 17). t tests were performed to compare the intensity difference between ChIP-enriched samples and control ChIP samples. Clones with significant difference (p < 0.001) were selected for further confirmation.

Selected clones were sequenced with an ABI 377 automated DNA sequencer (Applied Biosystems). Quantitative PCR was performed using anti-Aire Ab was detected by FITC-conjugated goat anti-mouse IgG (Southern Bio-technology Associates). After background subtraction and normalization with β-actin, expression differences were calculated between WT and Aire−/− mice.

RNA and RT-PCR

RNA was prepared from sorted TEC using the mini RNasy RNA isolation kit (Qiagen). Reverse transcription of RNA was performed using standard conditions. Different cDNA products were normalized to β-actin expression. Primer pairs were designed to span at least one intron to avoid amplification from genomic DNA. The PCR cycle number was controlled so that the amplification is still in the exponential growth phase to achieve optimum quantification. Amplification for each gene was done using at least two different cycle numbers to ensure that the amplification did not plateau. After electrophoresis on 2% agarose gels, intensity of RT-PCR products was measured using the Chemilimager 5500 (Alpha Innotech). After background subtraction and normalization with β-actin, expression differences were calculated between WT and Aire−/− mice.

Histology and disease grading

Tissues were fixed in 10% neutral formalin, embedded in paraffin, and 5-µm-thick serial sections were stained with H&E. Prostate and gastric pathology were graded on a scale of 1–4. Grade 1 inflammation consists of 1–2 foci of inflammatory cells, including lymphocytes, granulocytes, and monocytes restricted mainly to the interstitial space. Grades 2 and 3 include an incremental extent of submucosal and mucosal inflammation, involving glandular structures, loss of the muscularis mucosae, and in the case of gastritis hyperplasia of the gastric epithelium. Grade 4 pathology is severe and diffuse inflammation.

Results

Autoimmunity against autoantigens unexpressed in Aire−/− mice

We have extensively characterized the autoimmune phenotypes in the Aire−/− mice and confirmed the previously reported phenotypes. In addition, extensive lymphocyte infiltration into the prostate gland was observed in six of nine Aire−/− mice (Fig. 1, B and C). Moreover, the sera of the Aire−/− mice with experimental autoimmune prostatitis (EAP) were found to react with cytoplasmic Ags of the prostate gland (Fig. 1E) with a staining pattern identical with the serum Abs in day 3 thyomecimized (d3tx) (C57BL/6 × A/J)F1 mice with EAP (Fig. 1F). Serum Ab to prostate gland was not detected in the Aire−/− mice without EAP (Fig. 1D). Immunoprecipitation of prostate’s coagulation gland extracts with serum from d3tx and Aire−/− mice indicates that the target Ags were two novel prostate Ags, named EAPA1 and EAPA2 proteins (19). We also observed significant lymphocyte infiltration in the mucosa and submucosa area of the stomach with invasion into glands and loss of parietal cells (Fig. 1, H and I).

We also isolated mTEC using an enrichment protocol from Gray et al. (18). Thymi were then agitated in 25 ml of RPMI 1640 medium and placed on a shaker for 10 min. Thymic fragments were then allowed to settle and medium was removed by aspiration. This process was repeated two more times to deplete RBC and thymocytes. The thymic fractions were then incubated in 5 ml of 0.125% (w/v) collagenase D with 0.1% (w/v) DNase I (Sigma-Aldrich) in RPMI 1640 at 37°C for 1 h. Gentle mechanical agitation was performed by pipetting. Remaining thymocytes were depleted by positive selection with immunomagnetic beads specific for CD45 (Miltenyi Biotec).

The remaining cells were stained for isolation of MEC by flow sorting. Cells were incubated with anti-FcR (clone 2.4G2) for 10 min on ice and stained with the following Abs for 15 min on ice (BD Pharmingen): CD80 FITC, CD86 PE (10A.1), 6C3 (6C3/BP-1), CD45 PerCP (30-F11), and EPCAM (G8.8). The 6C3 Ab is equivalent to the CDR1 Ab (18). Cells were washed and Ep-CAM Ab was stained with allophycocyanin-labeled anti rat IgG2a (R6/1.30) secondary Ab for 20 min on ice. Cells were washed and mTEC were sorted by gating on cells that are CD45−, G8.8−, CD80high, 6C3 intermediate population.

In vitro promoter activity assay

The potential promoter activity of Aire-binding sequences was evaluated by a luciferase assay. We decided to evaluate the promoter sequences of −2 kb that contain most regulatory elements, although some regulatory elements may reside outside of these regions. The promoters from two candidate genes were amplified by PCR with primers containing I SAR restriction enzyme-cutting sites. The PCR fragments were inserted into pGL2 basic expression vector (Promega). NIH 3T3 cells (3 × 104) were transfected with the reporter constructs by LipofectAMINE (Ambion). Plasmid DNA (2 µg of the Aire expression construct, 2 µg of the luciferase reporter construct, and 0.1 µg of Renilla luciferase plasmid) were added to 100 µl of serum-free DMEM mixed with 100 µl of the same media. The mixture was incubated at 25°C for 30 min. Serum-free DMEM (0.8 ml) was added and poured over the cells. Cultures were incubated at 37°C for 5 h. The medium was then replaced with 1 ml of DMEM supplemented with 10% FCS, and cells were incubated at 37°C for an additional 48 h. Cells were then lysed and the amount of protein in the cell lysates was determined by bichinchonic acid. Luciferase activity present in 20 µg of proteins was measured using the dual luciferase assay system (Promega). Renilla luciferase was used to normalize transfection efficiency.

Purification of total TEC and mTEC

To isolate total TEC, thymi from 10 mice (3 wk old) are finely minced and slowly stirred in RPMI 1640 medium for 10 min at room temperature to release the bulk of free thymocytes. The tissue fragments are slowly stirred for three incubations of 15 min at 37°C in the presence of 1.6 mg/ml collagenase (Sigma-Aldrich), 0.1% deoxyribonuclease, and 0.2 mg/ml dispase (Roche Molecular Biochemicals) to isolate thymic rosettes. The en-riched thymic rosettes were pooled for purification of TEC by magnetic cell sorting using the MACS separation system (Miltenyi Biotec). First, dendritic cells and macrophages were positively selected by staining with biotinylated CD11c mAb (BD Pharmingen) and F4/80 mAb (Serotec). The CD11c-F4/80-negative population was then used as a source of TEC, which were selected by means of staining with biotinylated GL8 (BD Pharmingen), followed by anti-biotin MicroBeads (Miltenyi Biotec) and MACS separation. The majority of the TEC are mTEC based on previous studies (7).
FIGURE 1. Histology and histopathology of the Aire−/− mice. A–C, The anterior prostate of normal mice (A, H&E, ×100) and Aire−/− mice (B, ×100; C, ×200). The normal prostate is devoid of lymphocytic infiltration; it contains flat epithelial cells and secretion within gland lumen. Severe monocyteic prostatitis is detected in the Aire−/− mice with loss of prostate lumen and reduction of secretion. D–F (×200), Indirect immunofluorescence detection of serum IgG autoantibody to anterior prostate. The sera were from Aire−/− mice without prostatitis (D), from Aire+/− mice with prostatitis (E), and from d3tx Aire−/− mice with prostatitis (F). G–I, The stomach of normal mice (G, H&E, ×100) and stomach of Aire−/− mice (H and I, ×100). Note that in gastritis, there is heavy infiltration of mononuclear cells in the gastric mucosa and submucosa with occasional invasion of inflammatory cells into the glands and loss of parietal cells.

To gain insight into the immunological and molecular mechanisms responsible for these phenotypes, RT-PCR was used to analyze the expression of EAPA1 and EAPA2, which are autoantigens for EAP as well as the expression of ATP4a and ATP4b, which are autoantigens for autoimmune gastritis that occurs in both Aire−/− mice and d3tx BALB/c mice. The initial experiments were done with RNA prepared from MACS-sorted total TEC using semiquantitative RT-PCR. To our surprise, the expression of EAPA1 was undetectable even after two rounds of RT-PCR, suggesting that their expression in the TEC is absent or very weak. The expression of EAPA2 was not detectable after 35 cycles of RT-PCR but became detectable after a second round of PCR amplification (Fig. 2); however, no major expression difference was found between wild-type (WT) and knockout (KO) mice (Fig. 2). Similarly, the expression of ATP4a and ATP4b was not repressed in Aire−/− mice compared with littermate Aire+/+ controls (Fig. 2). We then performed real-time RT-PCR using RNA purified from flow-sorted medullary TEC and confirm that there is no major difference expression between WT and KO mice for ATP4a and ATP4b (Fig. 3). We also analyzed the expression of 10 other autoantigens using the MACS-sorted TEC. Repressed expression in Aire−/− TEC was observed for 9 of the 10 autoantigens (Fabp2, CYP1A2, INS1, INS2, ICA69, Tph, Tgn, Spt1, and Mater), while the expression of Rbp3 was not different (Table I and Fig. 2). The results for INS2, Mater and Tgn were further confirmed by real time RT-PCR performed with flow-sorted medullary TEC (Fig. 3).

Most of the tested transcripts encode TSAs responsible for autoimmune phenotypes observed in the Aire−/− mice, including Sjogren’s syndrome (Spt1), ovarian failure (Mater), insulin-dependent diabetes (insulin 1 and 2 and ICA69), and autoimmune thyroiditis (Tgn). Our results confirmed previous reports that the repressed expression of certain autoantigens may be responsible for autoimmunity against these Ags. However, it also suggests that Aire−/− mice can develop autoimmune response in certain tissues even though the expression of the cognate autoantigens are not controlled by Aire (8, 20).

Aire binds to specific promoter sequences in vivo

We then investigated how Aire controls the expression of the TSA targeted by the autoimmune response. As our previous studies (13) demonstrated that Aire can bind to specific DNA sequence motifs in vitro, we hypothesized that Aire may bind to the promoter regions of its primary target genes and directly regulate their expression. Furthermore, some of the primary target genes could include transcription factors and growth factors, which can regulate the expression of a large number of secondary genes with diverse biological and molecular functions. To identify the target genes directly regulated by Aire, we used a novel strategy that combines ChIP with microarray that enabled us to discover genes directly controlled by transcription factors. Similar studies have been performed successfully in the yeast system using yeast genomic microarrays (14); however, such analyses cannot be performed easily in the mammalian systems because of the difficulty in developing...
a microarray that encompasses the vast numbers of promoter sequences in the mammalian genomes. To circumvent this problem, we created a mouse (C57BL/6) genomic DNA library that contains DNA enriched for sequences bound in vivo by the Aire protein using ChIP with a polyclonal anti-Aire Ab produced in-house (13, 15). A total of 2000 clones were selected from this library and the inserts were PCR amplified and printed on poly-L-lysine-coated glass slides. In addition, the ChIP array contained the promoter regions from seven candidate APS1 autoantigens (Tgn, Tph, GAD65, insulin 1, insulin 2, and CYP1A2 and Fabp2). The selected sequences were located in the promoter region (within 2 kb from the start codon) and contained the previously identified Aire consensus-binding motifs: ATTGGTTA (G box), TTATTA (T box) or motifs with multiple boxes (such as GG and TGG boxes), which have higher binding affinity than the single box motifs (13). The ChIP arrays were probed with Cy3-labeled ChIP-enriched DNA and Cy5-labeled DNA from control ChIP. Data from three independent experiments were used to perform statistical analyses for the selection of candidate DNA sequences that are bound by the Aire protein. Eighteen spots had at least 2-fold higher intensity of hybridization \((p < 0.001)\) from the ChIP-enriched DNA than the control ChIP DNA control. These 18 spots include four autoantigens (Tgn, Tph, CYP1A2, and Fabp2) and 14 clones from the ChIP library. These 14 selected clones were sequenced and

![FIGURE 3. Real time RT-PCR analyses of gene expression. RNA from mTEC of Aire−/− (KO) and Aire+/+ (WT) mice was normalized by the expression level of β-actin. Two independently prepared RNA samples from each group were analyzed. The mean expression and SD are presented. The Student t test was performed to assess statistical significance.](http://www.jimmunol.org/)

Table I. AIRE-binding motifs, ChIP PCR, and RT-PCR confirmation for candidate autoantigens

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>Gene Name</th>
<th>Disease</th>
<th>Tissue</th>
<th>Motif</th>
<th>ChIP PCR</th>
<th>RT-PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ins1</td>
<td>Insulin 1</td>
<td>Diabetes</td>
<td>Pancreas</td>
<td>G</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Ins2</td>
<td>Insulin 2</td>
<td>Diabetes</td>
<td>Pancreas</td>
<td>G</td>
<td>nd</td>
<td>Yes</td>
</tr>
<tr>
<td>ICA69</td>
<td>Islet cell autoantigen</td>
<td>Diabetes</td>
<td>Pancreas</td>
<td>G</td>
<td>nd</td>
<td>Yes</td>
</tr>
<tr>
<td>GAD65</td>
<td>Glutamic acid decarboxylase 65</td>
<td>Diabetes</td>
<td>Pancreas</td>
<td>None</td>
<td>nd</td>
<td>na</td>
</tr>
<tr>
<td>GAD67</td>
<td>Glutamic acid decarboxylase 67</td>
<td>Diabetes</td>
<td>Pancreas</td>
<td>None</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>CYP1A2</td>
<td>Cytochrome P450 1A2</td>
<td>Hepatitis</td>
<td>Liver</td>
<td>GG</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Tph</td>
<td>Tryptophan hydroxylase</td>
<td>APECED</td>
<td>Intestine</td>
<td>G</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Tgn</td>
<td>Thyroglobulin</td>
<td>APECED</td>
<td>Thyroid</td>
<td>GG</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Fabp2</td>
<td>Fatty acid-binding protein 2</td>
<td>APECED</td>
<td>Intestine</td>
<td>GG</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Spt1</td>
<td>Salivary protein 1</td>
<td>Sjogren’s syndrome</td>
<td>Salivary gland</td>
<td>G</td>
<td>nd</td>
<td>Yes</td>
</tr>
<tr>
<td>Mater</td>
<td>Maternal effect gene</td>
<td>Ovarian failure</td>
<td>Ovary</td>
<td>G</td>
<td>nd</td>
<td>Yes</td>
</tr>
<tr>
<td>Rbp3</td>
<td>Retinol-binding protein 3</td>
<td>EAU</td>
<td>Retina</td>
<td>G</td>
<td>nd</td>
<td>No</td>
</tr>
<tr>
<td>ATP4b</td>
<td>H. K-ATPase 4b</td>
<td>Gastritis</td>
<td>Gastric mucosa</td>
<td>None</td>
<td>nd</td>
<td>No</td>
</tr>
<tr>
<td>ATP4a</td>
<td>H. K-ATPase 4a</td>
<td>Gastritis</td>
<td>Gastric mucosa</td>
<td>G</td>
<td>nd</td>
<td>No</td>
</tr>
<tr>
<td>EAAPA2</td>
<td>Experimental autoimmune prostatitis Ag 2</td>
<td>Prostatitis</td>
<td>Prostate</td>
<td>G</td>
<td>nd</td>
<td>No</td>
</tr>
<tr>
<td>EAAPA1</td>
<td>Experimental autoimmune prostatitis Ag 1</td>
<td>Prostatitis</td>
<td>Prostate</td>
<td>GG</td>
<td>nd</td>
<td>na</td>
</tr>
</tbody>
</table>

*Yes indicates confirmation by ChIP PCR or decreased expression in Aire−/− compared with WT TEC. No indicates no binding evidence by ChIP PCR or no expression difference between WT and Aire−/−. na, No amplification; nd, not done. G, single G motif (ATTGGTTA); GG, double G motif (ATTGGTTANATTGGTTA); TG, TG motif (TTATTAATTGGTTA). EAE, Experimental autoimmune encephalomyelitis; ovarian failure, autoimmune premature ovarian failure; EAU, experimental autoimmune uveitis.
found to contain at least one Aire-binding motif (Table II). The sequences for six clones were located within 5 kb of the promoter regions of known genes: Kruppel-associated box (Krab) containing zinc finger protein (two clones), bone morphogenetic protein 9 (Bmp9), mannoside acetylglucosaminyltransferase 1 (Mgat1), chitinase-related protein, and a protein similar to the nuclear protein 9 (Bmp9). To confirm the ChIP microarray results, PCR was performed for selected promoter sequences using ChIP-enriched DNA as template. Amplification from three independent ChIP-enriched DNA confirmed 9 of the 11 tested sequences including promoter sequences for six clones were located within 5 kb of the promoter regions of two known genes: Kruppel-associated box containing zinc finger protein (two clones), bone morphogenetic protein 9 (Bmp9), mannoside acetylglucosaminyltransferase 1 (Mgat1), chitinase-related protein, and a protein similar to the nuclear protein Nop56. The sequences from the eight remaining clones were not found within putative promoter regions (Table II). To confirm the ChIP microarray results, PCR was performed for selected promoter sequences using ChIP-enriched DNA as template. Amplification from three independent ChIP-enriched DNA confirmed 9 of the 11 tested sequences including promoter sequences for six clones were located within 5 kb of the promoter regions of two known genes: Kruppel-associated box containing zinc finger protein (two clones), bone morphogenetic protein 9 (Bmp9), mannoside acetylglucosaminyltransferase 1 (Mgat1), chitinase-related protein, and a protein similar to the nuclear protein Nop56. The sequences from the eight remaining clones were not found within putative promoter regions (Table II).

### Table II. Clones identified from the ChIP genomic library

<table>
<thead>
<tr>
<th>Clone ID</th>
<th>Gene Name</th>
<th>Accession Number</th>
<th>Position</th>
<th>Binding Motif</th>
<th>ChIP PCR</th>
<th>RT-PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>27A1</td>
<td>Hypothetical KRAB box containing protein</td>
<td>NT_039268</td>
<td>53122–53194</td>
<td>TG</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>31B11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26D1</td>
<td>Bone morphogenetic protein 9</td>
<td>AF188286</td>
<td>22–141</td>
<td>GG</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>26F9</td>
<td>Mannoside acetylglucosaminyltransferase 1</td>
<td>AC139935</td>
<td>13138–13362</td>
<td>GG</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>32C2</td>
<td>New</td>
<td>AC099710</td>
<td>193878–193999</td>
<td>G</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>12C3</td>
<td>Unknown</td>
<td>AY730870</td>
<td>1–578</td>
<td>GG</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>23A7</td>
<td>Chitinase-related protein</td>
<td>NT_039239</td>
<td>8117896–8118026</td>
<td>GG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18E9</td>
<td>Similar to the nuclear protein Nop56</td>
<td>AY730869</td>
<td>1–185</td>
<td>G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23F5</td>
<td>Unknown</td>
<td>NT_039185</td>
<td>1868342–1868388</td>
<td>G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29A7</td>
<td>Unknown</td>
<td>NT_039185</td>
<td>1868342–1868388</td>
<td>G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29D2</td>
<td>Unknown</td>
<td>NT_039202</td>
<td>17354881–17354934</td>
<td>G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29F2</td>
<td>Unknown</td>
<td>NT_081917</td>
<td>1074421–1074502</td>
<td>G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30C1</td>
<td>Unknown</td>
<td>NT_039202</td>
<td>17354873–17354933</td>
<td>G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26A1</td>
<td>Unknown</td>
<td>NT_039185</td>
<td>1868438–1868821</td>
<td>G</td>
<td></td>
<td>No</td>
</tr>
</tbody>
</table>

### FIGURE 4. ChIP PCR analysis of candidate genes. A, ChIP PCR analysis of selected candidate promoters. Three independent ChIP experiments were performed with an in-house anti-Aire Ab (P1–P3). A negative control ChIP (without Ab) was included in lane N. B, ChIP was conducted with a commercial anti-Aire Ab (Ab13573 from Abcam) or anti-rabbit IgG. WT and Aire<sup>−/−</sup> (KO) were used in these studies as indicated.

To rule out nonspecific effect, we made several constructs that contain partial Aire promoter sequences. These experiments suggest that the Aire promoter activity depends on the presence of Aire-binding sites and the Aire protein (Fig. 5B).

### Aire drives transcription of its target genes in vivo

To confirm the promoter function of Aire-binding sequences in vivo, we used semiquantitative RT-PCR to compare the expression level in MACS-sorted TEC between the WT mice and the Aire<sup>−/−</sup> mice for selected genes whose promoters contain putative Aire-binding sites and are bound by the Aire protein. As expected, Aire<sup>−/−</sup> TEC showed significant reduction in expression for all the genes selected for this analysis, including four autoantigen genes (Tph, Tgn, CYP1A2, and Fabp2), Mgat5, Bmp9, and a new KRAB-box-containing gene (Fig. 2). Furthermore, we conducted real-time RT-PCR analysis of flow-sorted medullary MEC (Fig. 3) to confirm the results obtained by semiquantitative RT-PCR. These two data sets are largely consistent with the exception of the Krab gene, which showed large difference in the semiquantitative analysis (Fig. 2) but only weak difference with the real-time RT-PCR analysis (Fig. 3).

Because Bmp and Mgat have multiple members in their gene families, we analyzed the expression in TEC for the other gene family members. Of the 12 genes tested here, Mgat5, Bmp2, Bmp4, Fabp2, and AC099710 can be amplified from ChIP-enriched DNA with the anti-Aire Ab and WT mice while there is no or very weak amplification from the IgG control ChIP or the anti-Aire Ab with Aire<sup>−/−</sup> mice.

### Aire-binding sequences have in vitro promoter activity

Although the ChIP microarray and the ChIP PCR experiments together provided strong evidence for Aire’s binding ability to specific sequence motifs, it is necessary to test the promoter activity of Aire-binding sequences. This hypothesis was tested using a promoter-reporter construct (luciferase) assay in a cotransfection system. Briefly, a fragment of ~2 kb from the putative promoter regions of two candidate genes (Fabp2 and Tgn) and a control gene (IFN-γ) that does not contain an AIRE-binding site was PCR amplified and cloned into the pGL2 vector (reporter construct). In addition, the full-length Aire coding region was cloned into the pCDNA3.1 expression vector (AIRE-expression construct). Cos7 cells were cotransfected with the reporter construct and either the Aire expression construct or the empty pCDNA3.1 control vector. The luciferase activities associated with the promoter sequences from the Tgn and Fabp2 promoters were 12–14 times higher in cells transfected with Aire than control vectors (Fig. 5A), while there was no difference for the transfection experiments for IFN-γ (Fig. 5A). To rule out nonspecific effect, we made several constructs that contain partial Aire promoter sequences. These experiments suggest that the Aire promoter activity depends on the presence of Aire-binding sites and the Aire protein (Fig. 5B).
Aire is differentially expressed between Aire autoantibody and APS1 or APECED, indicating that Aire has a crucial role in the maintenance of immunological self-tolerance. One of the hallmarks of APS1 is the presence of multiple autoreactive Abs against defined organ-specific Ags. A previous study has shown that 73% of the Aire−/− mice generated at least one circulating autoantibody and >40% of the sera contained two or more autoantibodies that react with Ags in the liver, adrenal cortex, pancreas, or testis (7). The wide array of autoimmune phenotypes in Aire−/− mice was shown to be caused by defective negative selection of autoreactive T cells in the thymus (9), although other immunological defects may also occur. To understand the molecular mechanism underlying the immunological defects in Aire−/− mice, we undertook a series of studies to define genes controlled by the Aire protein. Based on our previous discovery that Aire can bind in vitro to specific DNA sequences (13), we hypothesized that Aire can directly regulate the expression of many downstream genes through binding to promoter/enhancer DNA sequences. To identify candidate target genes, we developed a new approach that combines ChIP and microarray analysis. The ChIP experiments are expected to precipitate specific DNA sequences that are bound by the anti-Aire Abs. However, precipitation of nonspecific DNA sequences may also be quite common due to the small proportion of DNA sequences that are actually bound by the Aire protein. Indeed, our microarray analysis suggested that most of the DNA clones from the ChIP library may be due to nonspecific precipitation. Therefore, we applied stringent criteria to select candidate clones that most likely contain Aire-specific DNA-binding sites. With our selection criteria, we selected 14 clones for further confirmation. Although the other clones did not reach our selection criteria, a number of them may also contain true Aire target genes as our sequencing analysis suggested that ~30% of the clones contain putative Aire-binding sites.

After the selection by ChIP microarray, putative candidates were further analyzed using PCR with ChIP-enriched DNA as template. Among the 11 tested clones, 9 were confirmed by PCR with ChIP-enriched DNA with an in-house anti-Aire Ab (Fig. 4A). Furthermore, five of these genes (Mgas1, KRAB, Bmp9, Fabp2, and AC099710) were further confirmed to contain Aire-binding sites using ChIP-PCR analysis with a different anti-Aire Ab and multiple negative controls including Aire−/− tissues and anti-IgG Ab (Fig. 4B). These results provided strong support that the identified promoters are bound by the Aire protein in vivo. However, the approach used in this study cannot identify all Aire target genes as the ChIP array does not contain all genomic sequences that may contain an Aire-binding sites. The recently developed human genomic DNA chips would be ideal reagents for this purpose.

In addition to in vivo binding of the Aire protein to specific sites, it was important to determine whether Aire binding to promoters can alter the transcription of the downstream genes. This was accomplished by a luciferase reporter that was under the transcriptional control of promoters with Aire-binding sites. With both promoters tested (Fabp2 and thyroglobulin), Aire protein significantly increase the expression of the reporter, while Aire did not significantly alter the reporter expression. These studies provided strong evidence that binding of Aire to specific sites of promoters regulates the transcriptional activities. To further confirm that Aire binding to specific promoter sequences can regulate gene expression, we determined the expression changes of the genes whose promoter regions were bound by Aire protein. Among the eight genes that are predicted be controlled by Aire based on our ChIP studies, RT-PCR analyses confirmed that their expression is reduced in Aire−/− compared with Aire+/− as expected. The combined data from ChIP analysis, promoter assay and gene expression studies strongly suggest that Aire can bind to promoters of specific target genes and directly regulate their expression.

A very interesting picture emerged from the confirmed genes directly regulated by Aire. The first group includes four genes (Fabp2, Tph, Tgn, and CYP1A2) that are autoantigens in APS1. The expression difference for two of these TSA (CYP1A2 and Fabp2) has already been documented between Aire−/− and
Aire+/- mice (7). The repressed expression of TSA in the TEC has been suggested as a mechanism of the autoimmune phenotype observed in Aire-/- mice (7), as thymic expression of tissue-specific Ags is believed to regulate immunological tolerance by thymic deletion of developing autoreactive T cells (27–31). Among the 14 autoantigen genes examined by RT-PCR, 9 were repressed in the Aire-/- TEC, while 1 gene was undetectable and four other genes were not repressed in Aire-/- mice (Table I and Fig. 2).

It has been suggested that the repressed expression of autoantigens in the TEC is responsible for the autoimmune in Aire-/- mice. This is a very attractive hypothesis as there is a direct link between the presence of some proteins in the TEC and the absence of peripheral lymphocyte that recognize the proteins. However, Aire-/- mice develop autoimmune response against certain autoantigens despite their normal expression in TEC. For example, autoimmune gastritis is a common autoimmune disease observed in the Aire-/- mice, yet the gastritis autoantigens (ATP4a and ATP4b) are expressed at normal levels in the TEC of Aire-/- mice. Similarly, the Aire-deficient thymus was reported to retain normal transcriptional expression of a ubiquitous protein (α-fodrin) (8), an autoantigen associated with the Sjogren’s syndrome. Finally, autoimmune prostatitis developed in the Aire-/- mice even though the prostate autoantigens associated with this disease had little or no expression in the TEC of normal or Aire-/- mice. These results are consistent with the recent observations that other mechanisms beyond the transcriptional control of self-protein expression in the TEC, including Ag presentation (32), also contribute to the autoimmune phenotypes in Aire-/- mice.

The identification of the nonautoantigen genes directly controlled by Aire provided new molecular clues underlying the autoimmune caused by Aire. Our results demonstrated that Aire is essential for the thymic expression of a number of cytokines and transcription factors. The bone morphogenetic proteins (BMP) are a group of cytokines/growth factors of the TGF superfamily. Bmp9 was identified by the ChIP microarray as an Aire target gene. ChIP PCR analysis with both anti-Aire Abs (Fig. 4, A and B) confirmed that Aire can bind to a putative Aire-binding site at the Bmp9 promoter region and RT-PCR analysis confirmed that Bmp9 expression is lower in Aire-/- than Aire+/+ TEC. Furthermore, five of the eight BMP genes studied here possess an Aire-binding site in their promoters. RT-PCR analysis showed that the expression of these genes is reduced in Aire-/- mice (Fig. 2), suggesting that the expression of multiple BMP genes may be directly controlled by Aire. The BMP molecules are involved in many important immunological processes such as cytokine production, lymphocyte homeostasis, and activation. For example, Bmp2 and Bmp4 are produced by the thymic stroma and function in thymic homeostasis by regulating T cell lineage commitment and differentiation (33). As Bmp2, Bmp4, and several other BMP are repressed in Aire-/- TEC, it will be interesting to determine whether the BMP defects contribute to the autoimmune phenotypes observed in the Aire-/- mice.

Aire also controls the expression of several posttranslational modifiers of surface proteins. N-Acetylglucosaminyltransferase 1 (Mgat1) was identified by the ChIP microarray experiment and its promoter region contains a “GG” Aire-binding motif (Table II). In vivo binding of Aire to the motif was confirmed by two different sets of ChIP PCR experiments (Fig. 3) and RT-PCR analyses confirmed the reduced expression of Mgat1 in Aire-/- mice (Fig. 2). The expression of another member of the Mgat family (Mgat5) was also found to be regulated by Aire (Fig. 2), while the expression of Mgat2 was not repressed in Aire-/- TEC (Fig. 2). Mgat proteins have been shown to be involved in embryogenesis, immune cell regulation, fertility, and cancer progression (34–38). For example, Mgat-modified N-glycans on TCR complex can bind to galectin-3, sequestering TCR within a multivalvent galectin-glycoprotein lattice that impedes Ag-dependent receptor clustering and signal transduction (39). Mgat-deficient mice exhibit complex phenotypes including autoimmune disease susceptibility (39). It is possible that Mgat deficiency may alter the structure and function of surface molecules involved in the negative selection of thymocytes mediated by TEC.

In summary, this study demonstrated that Aire functions as a transcription factor and directly regulates the expression of a large number of, but not all, autoantigens, providing a molecular basis for the defect in negative selection in Aire-/- mice. Furthermore, our studies identified three Aire-regulated molecular pathways (BMP, KRAB, and Mgat) that may contribute to the development of autoimmune phenotype in Aire-/- mice.

Disclosures
The authors have no financial conflict of interest.

References


