

ARTICLE

Received 3 Sep 2015 | Accepted 7 Oct 2015 | Published 4 Nov 2015

DOI: 10.1038/ncomms9820

OPEN

# Intronic regulation of *Aire* expression by *Jmjd6* for self-tolerance induction in the thymus

Toyoshi Yanagihara<sup>1,2</sup>, Fumiyuki Sanematsu<sup>1,3</sup>, Tetsuya Sato<sup>4</sup>, Takehito Uruno<sup>1,3</sup>, Xuefeng Duan<sup>1,3,†</sup>, Takahiro Tomino<sup>1</sup>, Yosuke Harada<sup>1</sup>, Mayuki Watanabe<sup>1</sup>, Yuqing Wang<sup>3</sup>, Yoshihiko Tanaka<sup>5</sup>, Yoichi Nakanishi<sup>2</sup>, Mikita Suyama<sup>4</sup> & Fukui Yoshinori<sup>1,3</sup>

The thymus has spatially distinct microenvironments, the cortex and the medulla, where the developing T-cells are selected to mature or die through the interaction with thymic stromal cells. To establish the immunological self in the thymus, medullary thymic epithelial cells (mTECs) express diverse sets of tissue-specific self-antigens (TSAs). This ectopic expression of TSAs largely depends on the transcriptional regulator Aire, yet the mechanism controlling Aire expression itself remains unknown. Here, we show that *Jmjd6*, a dioxygenase that catalyses lysyl hydroxylation of splicing regulatory proteins, is critical for Aire expression. Although *Jmjd6* deficiency does not affect abundance of *Aire* transcript, the intron 2 of *Aire* gene is not effectively spliced out in the absence of *Jmjd6*, resulting in marked reduction of mature Aire protein in mTECs and spontaneous development of multi-organ autoimmunity in mice. These results highlight the importance of intronic regulation in controlling Aire protein expression.

<sup>1</sup>Division of Immunogenetics, Department of Immunobiology and Neuroscience, Medical Institute of Bioregulation, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan. <sup>2</sup>Research Institute for Diseases of the Chest, Graduate School of Medical Sciences, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan. <sup>3</sup>Research Centre for Advanced Immunology, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan. <sup>4</sup>Division of Bioinformatics, Multi-scale Research Centre for Medical Science, Medical Institute of Bioregulation, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan. <sup>5</sup>Section of Infection Biology, Department of Functional Bioscience, Fukuoka Dental College, 2-15-1 Tamura, Sawara-ku, Fukuoka 814-0175, Japan. † Present address: Laboratory of Pathogenic Microbiology and Immunology, Chinese Academy of Sciences, Beichen West Road, Chaoyang District, Beijing 100101, China. Correspondence and requests for materials should be addressed to Y.F. (email: fukui@bioreg.kyushu-u.ac.jp).

The thymus has spatially distinct microenvironments, the cortex and the medulla, where developing T-cells are selected to mature or die through the interaction with thymic stromal cells<sup>1,2</sup>. Cortical thymic epithelial cells (cTECs), a major stromal cell-type in the cortex, direct differentiation of CD4<sup>+</sup>CD8<sup>+</sup> immature thymocytes that are capable of recognizing self-major histocompatibility complex (MHC) molecules. On the other hand, medullary thymic epithelial cells (mTECs) play an important role in self-tolerance induction by eliminating self-reactive T-cells. A unique property of mTECs is their expression of diverse sets of peripheral tissue-specific self-antigens (TSAs)<sup>3,4</sup>. This ectopic expression of TSAs largely depends on the transcriptional regulator Aire<sup>5–8</sup>, which is expressed in mature mTECs<sup>9–11</sup>. The homozygous mutations of human AIRE cause an autoimmune disease known as autoimmune-polyendocrinopathy-candidiasis ectodermal dystrophy<sup>12,13</sup>. Similarly, Aire-deficient mice develop multi-organ autoimmunity with the failure to delete self-reactive T-cells<sup>5,14,15</sup>. Despite an important role in self-tolerance induction, the mechanism controlling Aire expression itself is poorly understood.

Alternative splicing is a major cellular mechanism in metazoans for generating proteomic diversity<sup>16,17</sup>. This is a posttranscriptional process in which premature transcripts are selectively cut and joined in more than one way to generate multiple mRNAs from a single gene. There are three forms of alternative splicing: exon skipping, alternative splice site usage and intron retention<sup>17</sup>. Of these, intron retention is the least frequent alternative splicing form<sup>17</sup>, which occurs when an intron, having been transcribed as a part of a pre-mRNA, is not spliced out. The sequence structure of most introns consists of a short 5' splice site boundary, a minimal AG dinucleotide 3' splice site boundary, a catalytic adenosine and a polypyrimidine tract (PPT)<sup>16</sup>. Mechanistically, intron retention is considered to be the result of weak splice site sequences that are not properly recognized by spliceosome<sup>18</sup>. Intron retention often inserts a premature termination codon in the mature transcript that would then be degraded by non-sense-mediated decay<sup>19</sup>. Therefore, its physiological significance has been so far overlooked.

Jmjd6 is a member of the JmjC-domain containing proteins that are involved in a wide range of oxidation reactions<sup>20</sup>. Jmjd6 was initially identified as a phosphatidylserine receptor that mediates recognition and engulfment of apoptotic cells<sup>21</sup>. However, recent evidence indicates that Jmjd6 is a nuclear protein and catalyses lysyl hydroxylation of multiple substrates, including splicing regulatory proteins, transcription factors and histones, in a manner dependent on the Fe(II) and 2-oxoglutarate<sup>22–25</sup>. Jmjd6 deficiency in mice causes abnormal development of multiple organs during embryogenesis and led to perinatal lethality<sup>26–28</sup>; yet, its role in the immune system and immune responses remain unclear. Here we show that Jmjd6 plays a key role in induction of central tolerance by controlling Aire expression in mTECs. Although Jmjd6 deficiency did not

affect abundance of Aire transcript, the intron 2 of Aire gene was not effectively spliced out in the absence of Jmjd6, owing to the unique 3' splice site sequence. As a result, the expression of Aire protein was markedly reduced in Jmjd6-deficient (*Jmjd6*<sup>-/-</sup>) mTECs, and T-cells generated in such thymic microenvironments caused multi-organ autoimmunity in mice. Our findings indicate that Aire protein expression is tightly controlled through two discrete steps, intron retention and its relief, the latter of which involves the enzymatic activity of Jmjd6.

## Results

### Impaired tolerance induction by *Jmjd6*<sup>-/-</sup> thymic stroma.

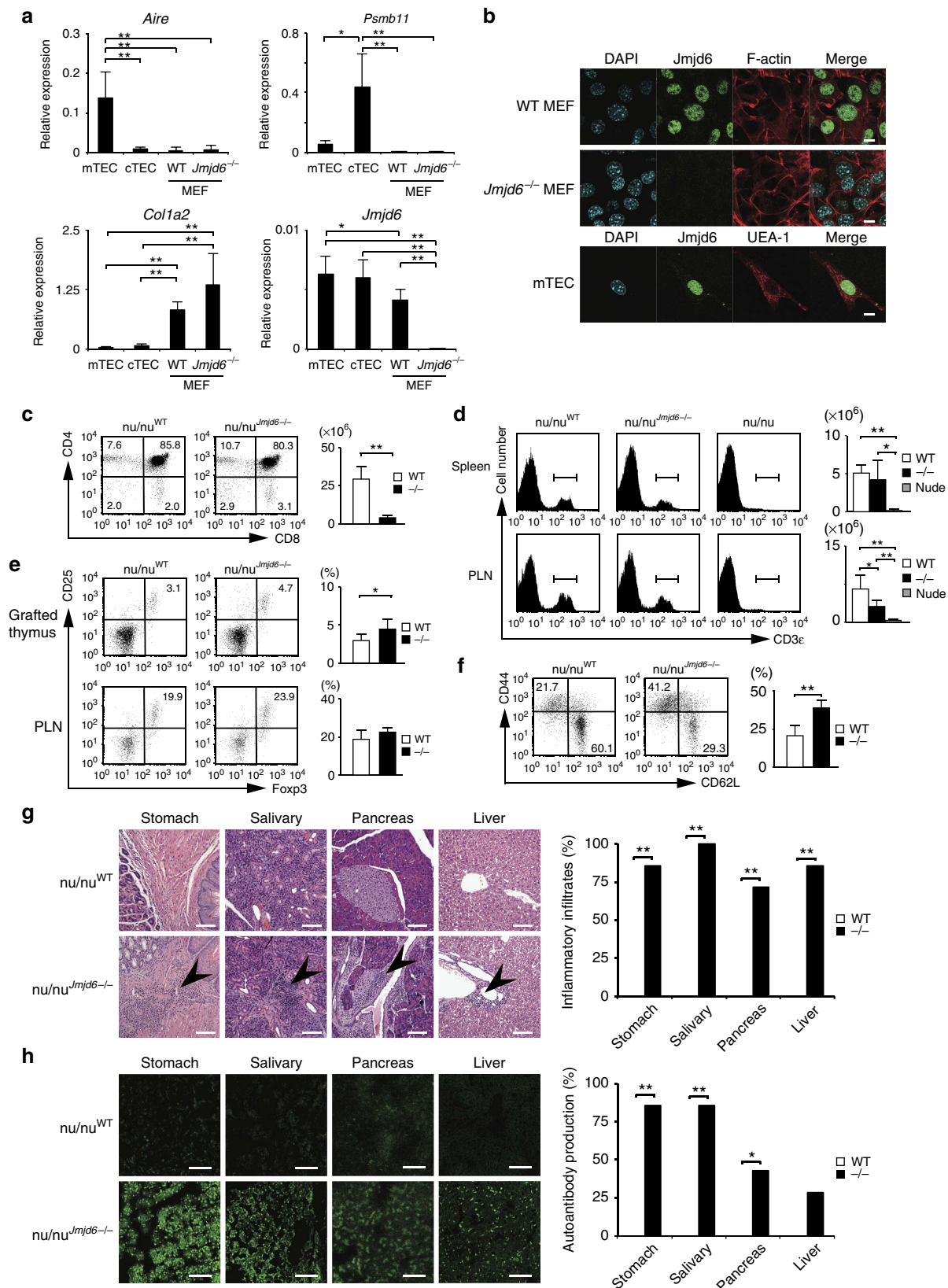
A previous study has shown that intrathymic T-cell development is partially impaired in *Jmjd6*<sup>-/-</sup> embryos<sup>27</sup>. However, *Jmjd6*<sup>-/-</sup> T-cells developed normally in chimeric mice reconstituted with fetal liver cells from the knockout mice (Supplementary Fig. 1), suggesting that Jmjd6 may play a role in thymic stromal cells. Indeed, Jmjd6 was expressed in mTECs and cTECs as well mouse embryonic fibroblasts (MEFs) (Fig. 1a), and its nuclear localization was confirmed by immunofluorescent staining (Fig. 1b). To assess a possible role of Jmjd6 in TECs, we prepared fetal thymi from wild-type (WT) and *Jmjd6*<sup>-/-</sup> embryos at embryonic days 15.5 (E15.5) and grafted them under the renal capsule of athymic C57BL/6 nude mice (B6.Cg-Foxn1<sup><nu>/Nrs</sup>) (designated nu/nu<sup>WT</sup> and nu/nu<sup>*Jmjd6*<sup>-/-</sup></sup>) after depleting haematopoietic cells by incubation in 2'-deoxyguanosine (2-DG). Although the size of the thymi developed in nu/nu<sup>*Jmjd6*<sup>-/-</sup></sup> mice was apparently small, their CD4<sup>+</sup>CD8<sup>+</sup> thymocytes normally differentiated into CD4<sup>+</sup>CD8<sup>-</sup> or CD4<sup>-</sup>CD8<sup>+</sup> mature T-cells (Fig. 1c), indicating that *Jmjd6*<sup>-/-</sup> cTECs are functionally intact to support positive selection of CD4<sup>+</sup>CD8<sup>+</sup> immature thymocytes. Indeed, nu/nu<sup>WT</sup> and nu/nu<sup>*Jmjd6*<sup>-/-</sup></sup> mice had almost comparable numbers of mature T-cells in the spleen and peripheral lymph nodes (PLNs) 10 weeks after grafting (Fig. 1d). Jmjd6 deficiency in thymic stroma did not impair generation of regulatory T-cells (Fig. 1e). However, the frequency of CD4<sup>+</sup> T-cells with an activated/memory phenotype (CD44<sup>hi</sup>CD62L<sup>lo</sup>) significantly increased in nu/nu<sup>*Jmjd6*<sup>-/-</sup></sup> mice (Fig. 1f). Histological examination demonstrated inflammatory infiltrates in the stomach, salivary gland, pancreas and liver of nu/nu<sup>*Jmjd6*<sup>-/-</sup></sup> mice (Fig. 1g). In addition, nu/nu<sup>*Jmjd6*<sup>-/-</sup></sup> mice, but not nu/nu<sup>WT</sup> mice, generated autoantibodies against these tissues (Fig. 1h). Thus, Jmjd6 expression in thymic stroma is required for induction of self-tolerance in developing T-cells.

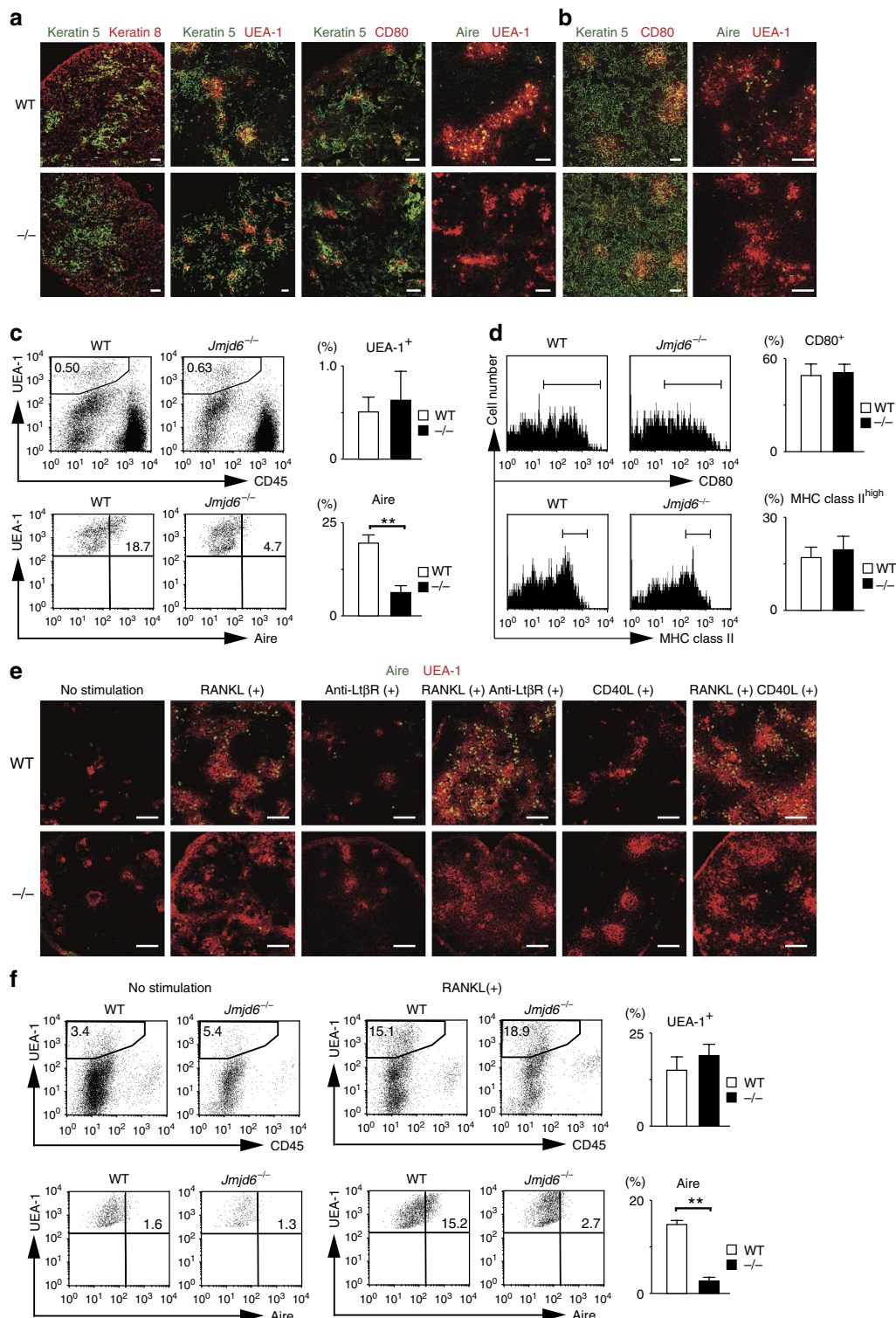
**Defective Aire expression in *Jmjd6*<sup>-/-</sup> mTECs.** To examine the function of Jmjd6 in TECs in more detail, we performed histological analyses. Haematoxylin and eosin staining of E18.5 thymi revealed that both the cortex and the medulla poorly developed in the absence of Jmjd6 (Supplementary Fig. 2). However, no significant difference in thymic architecture was

**Figure 1 | Jmjd6 expression in thymic stroma is required for T-cell tolerance induction.** (a) Real-time PCR analyses for *Jmjd6* expression in mTECs, cTECs and MEFs. (b) Immunofluorescence staining showing nuclear localization of Jmjd6. UEA-1 was used to detect mTECs. Data are representative of two independent experiments. (c) Expression of CD4 and CD8 on thymocytes differentiated in the grafts. The number of total thymocytes was compared between nu/nu<sup>WT</sup> ( $n=8$ ) and nu/nu<sup>*Jmjd6*<sup>-/-</sup></sup> mice ( $n=7$ ). (d) The numbers of CD3ε<sup>+</sup> T-cells in the spleen and PLNs were compared among nu/nu<sup>WT</sup> ( $n=8$ ), nu/nu<sup>*Jmjd6*<sup>-/-</sup></sup> ( $n=7$ ) and nude ( $n=3$ ) mice. (e) The numbers of CD25<sup>+</sup>Foxp3<sup>+</sup> regulatory T-cells in the grafted thymi and PLNs were compared between nu/nu<sup>WT</sup> ( $n=8$ ) and nu/nu<sup>*Jmjd6*<sup>-/-</sup></sup> mice ( $n=7$ ). (f) Expression of CD44 and CD62L on PLN CD4<sup>+</sup> T-cells from nu/nu<sup>WT</sup> ( $n=8$ ) and nu/nu<sup>*Jmjd6*<sup>-/-</sup></sup> mice ( $n=7$ ). (g) Haematoxylin and eosin staining of tissues from nu/nu<sup>WT</sup> and nu/nu<sup>*Jmjd6*<sup>-/-</sup></sup> mice. Scale bars, 100 μM. The percentage of each organ with inflammatory infiltrates was compared between nu/nu<sup>WT</sup> ( $n=8$ ) and nu/nu<sup>*Jmjd6*<sup>-/-</sup></sup> ( $n=7$ ) mice. (h) Immunofluorescence staining showing the presence of autoantibodies (green) in nu/nu<sup>*Jmjd6*<sup>-/-</sup></sup> mice. Scale bars, 100 μM. The percentage of mice with autoantibodies against each tissue was compared between nu/nu<sup>WT</sup> ( $n=8$ ) and nu/nu<sup>*Jmjd6*<sup>-/-</sup></sup> mice ( $n=7$ ). Data are collected from four (a), six (c–g) and three (h) separate experiments and are expressed as mean ± s.d. \* $P<0.05$ ; \*\* $P<0.01$  (two-tailed Student's *t*-test).

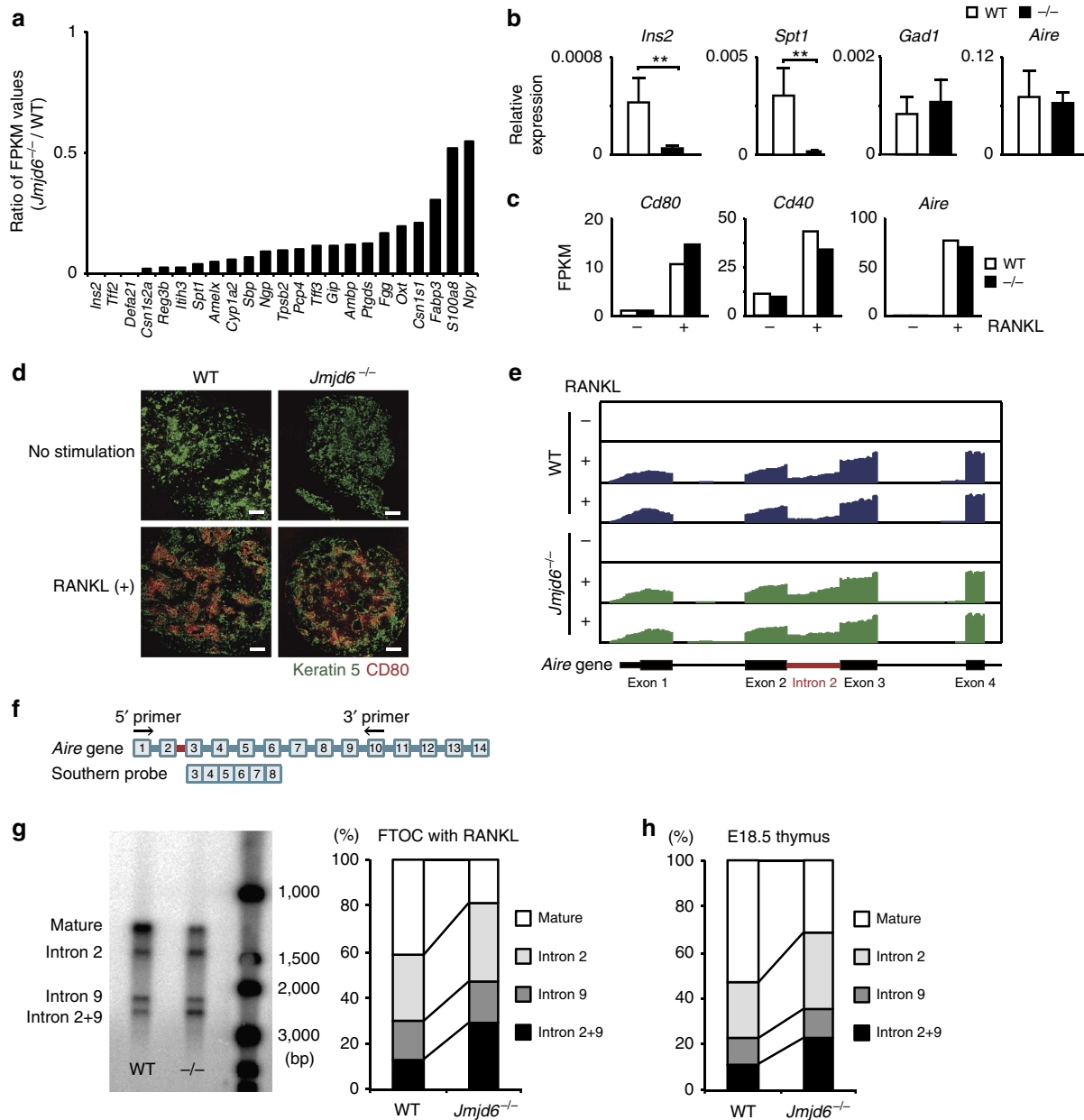
found between WT and *Jmjd6*<sup>-/-</sup> mice when thymic sections were stained for the cTEC marker keratin 8 and mTEC marker keratin 5 (ref. 29) (Fig. 2a). In addition, mTECs of both WT and *Jmjd6*<sup>-/-</sup> embryos were comparably stained with *Ulex europaeus* agglutinin-1 (UEA-1), a lectin that binds to mature

mTECs<sup>29</sup> (Fig. 2a), suggesting that *Jmjd6* deficiency does not affect mTEC maturation. Consistent with this, the expression of another maturation marker CD80 was unchanged between WT and *Jmjd6*<sup>-/-</sup> mTECs<sup>4,6,30</sup> (Fig. 2a). Surprisingly, however, the expression of Aire in mTECs of *Jmjd6*<sup>-/-</sup> embryos was





**Figure 2 | Defective Aire protein expression in *Jmjd6*<sup>-/-</sup> mTECs.** (a) Thymic sections from E18.5 WT and *Jmjd6*<sup>-/-</sup> embryos were stained with UEA-1 and/or antibodies specific for keratin 5, keratin 8, CD80 and Aire. Scale bars, 50  $\mu$ M. (b) Immunofluorescence analysis of thymi from WT and *Jmjd6*<sup>-/-</sup> embryos grafted under the kidney capsule of C57BL/6 mice. Six weeks after engraftment, thymic sections were stained with UEA-1 and/or antibodies specific for keratin 5, CD80 and Aire. Scale bars, 50  $\mu$ M. (c) Flow cytometric analyses for Aire expression in mTECs prepared from E18.5 embryos. The percentages of UEA-1<sup>+</sup> mTECs and Aire<sup>+</sup> UEA-1<sup>+</sup> mTECs were compared between WT ( $n=3$ ) and *Jmjd6*<sup>-/-</sup> mice ( $n=5$ ). (d) Flow cytometric analyses for expression of CD80 and MHC class II in mTECs prepared from E18.5 embryos. The percentages of CD80<sup>+</sup> mTECs and MHC class II<sup>high</sup> mTECs were compared between WT ( $n=3$ ) and *Jmjd6*<sup>-/-</sup> mice ( $n=4$ ). (e) Following stimulation with RANKL, anti-Lt $\beta$ R antibody and/or CD40L for 4 days, 2-DG-treated fetal thymic stroma from WT and *Jmjd6*<sup>-/-</sup> embryos were stained with UEA-1 and anti-Aire antibody. Scale bars, 50  $\mu$ M. (f) Flow cytometric analyses for Aire expression in mTECs prepared from FTOC with or without RANKL stimulation. The percentages of UEA-1<sup>+</sup> mTECs and Aire<sup>+</sup> UEA-1<sup>+</sup> mTECs were compared between RANKL-stimulated WT ( $n=3$ ) and *Jmjd6*<sup>-/-</sup> samples ( $n=3$ ). Data are collected from three (c,d,f) separate experiments and are expressed as mean  $\pm$  s.d. \*\* $P < 0.01$  (two-tailed Student's  $t$ -test).

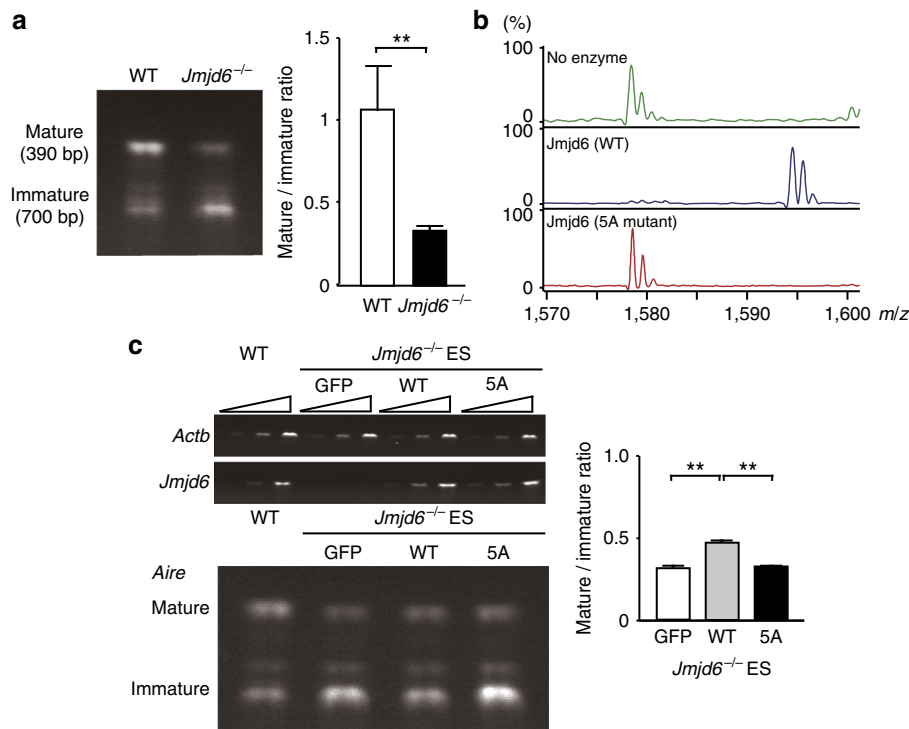


**Figure 3 | *Jmjd6* controls Aire protein expression through intron retention.** (a) Reduced expression of 23 Aire-dependent TSAs in  $Jmjd6^{-/-}$  thymic stroma. Data were obtained from RNAseq analyses for RANKL-stimulated samples. (b) Comparison of gene expression of Aire-dependent and -independent TSAs between RANKL-stimulated WT ( $n = 5$ ) and  $Jmjd6^{-/-}$  ( $n = 5$ ) thymic stroma by real-time PCR. *Ins2*, insulin 2; *Spt1*, salivary protein 1; *Gad1*, GAD67. Data are expressed as mean  $\pm$  s.d.  $**P < 0.01$  (two-tailed Student's *t*-test). (c) Comparable expression of *Cd80*, *Cd40* and *Aire* in WT and  $Jmjd6^{-/-}$  thymic stroma. Data were obtained from RNAseq analyses for RANKL-stimulated samples. (d) Following stimulation with RANKL for 4 days, 2-DG-treated fetal thymic stroma from WT and  $Jmjd6^{-/-}$  embryos were stained with antibodies specific for keratin 5 and CD80. Scale bars, 100  $\mu$ m. Data are representative of three independent experiments. (e) RNA-seq read densities at the *Aire* gene locus. Data were obtained from WT and  $Jmjd6^{-/-}$  thymic stroma stimulated with (+) or without (-) RANKL. (f) Schematic representation of primers and a probe used for RT-PCR and Southern blotting. (g) RT-PCR and Southern blot analysis showing multiple *Aire* transcripts with or without retention of intron 2, intron 9, and intron 2 plus intron 9. The intensity of each band was compared between RANKL-stimulated WT and  $Jmjd6^{-/-}$  thymic stroma ( $n = 3$ ). (h) The frequency of each *Aire* transcript was compared between E18.5 WT ( $n = 4$ ) and  $Jmjd6^{-/-}$  ( $n = 3$ ) thymi, as in (g).

markedly reduced at E18.5, as compared with that in WT controls (Fig. 2a and Supplementary Fig. 3a). This does not simply reflect delayed Aire expression in  $Jmjd6^{-/-}$  thymi, because similar results were obtained when the 'artificial' thymi from  $Jmjd6^{-/-}$  embryos were analysed 6 weeks after grafting (Fig. 2b and Supplementary Fig. 3b).

To determine more precisely the expression level of Aire in mTECs, we performed flow cytometric analyses. E18.5 thymi

from both WT and  $Jmjd6^{-/-}$  mice contained comparable numbers of CD45<sup>+</sup> UEA-1<sup>+</sup> mTECs (Fig. 2c). Although cell-surface expressions of CD80 and MHC class II were unchanged between WT and  $Jmjd6^{-/-}$  mTECs (Fig. 2d), intracellular staining of Aire revealed that the proportion of Aire<sup>+</sup> mTECs in  $Jmjd6^{-/-}$  thymi was reduced to 25.1% of the WT level (Fig. 2c). Thus, *Jmjd6* deficiency reduces Aire protein expression in mTECs without affecting mTEC maturation.



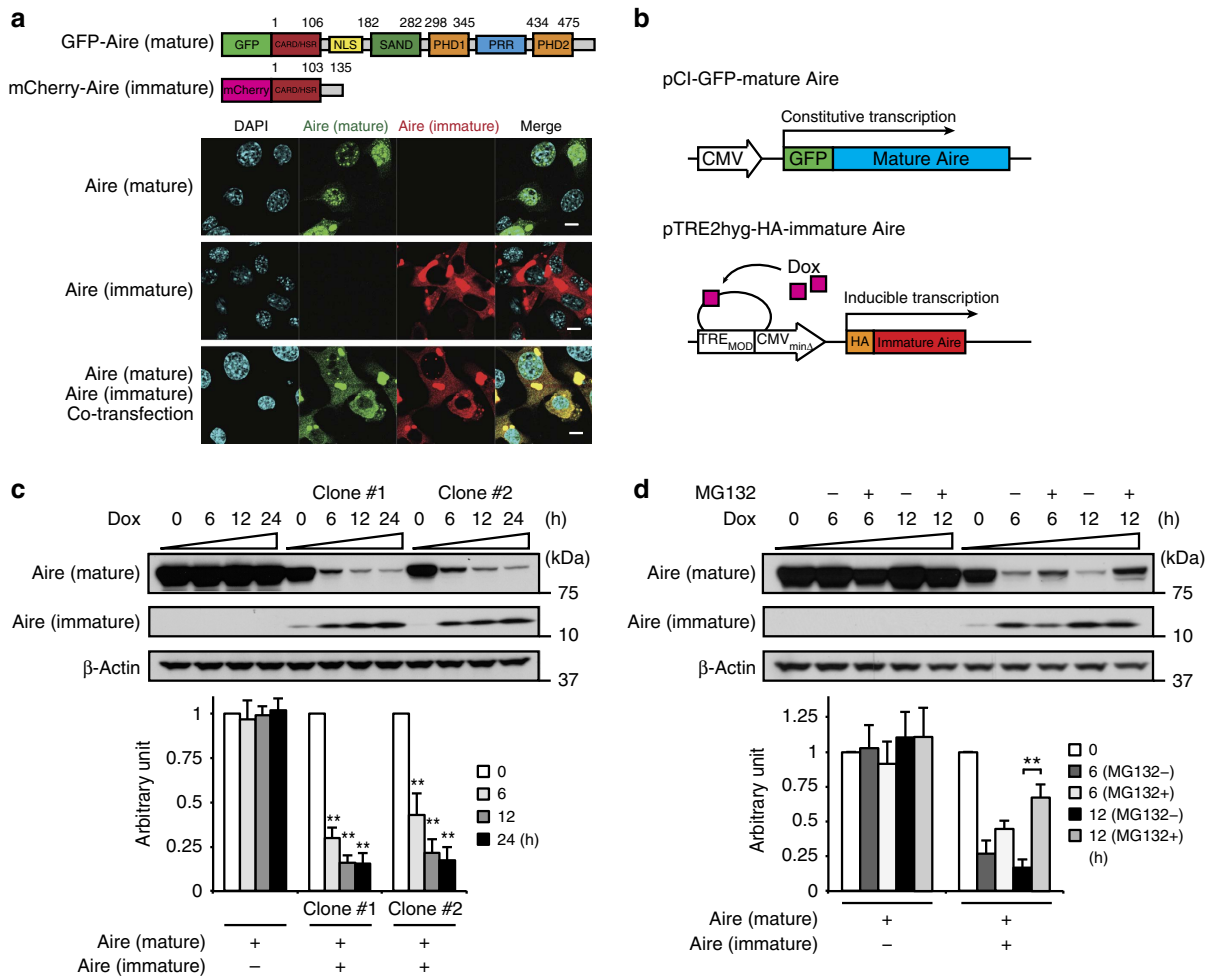
**Figure 4 | The enzymatic activity of Jmjd6 is required for intronic regulation of Aire gene.** (a) Comparison of the ratio of mature (without intron 2) to immature (with intron 2) Aire transcript between WT and Jmjd6<sup>-/-</sup> ES cells ( $n=4$ ). Amplification of Aire cDNA was performed with primers specific for the sequence of the exons 1 and 3. (b) MALDI-TOF analyses for lysyl-hydroxylation of the sequence of the exons 1 and 3. The mass corresponding to LUC7L2<sub>267-278</sub> + 16, equivalent to one hydroxylation, was identified in the presence of WT Jmjd6, but not 5A mutant. Data are representative of two independent experiments. (c) Effect of the 5A mutation on intron 2 retention of Aire gene. 30 h after transfection of Jmjd6<sup>-/-</sup> ES cells with genes encoding GFP-fusion Jmjd6, GFP-fusion 5A mutant or GFP alone, the ratio of mature to immature Aire transcript was analysed by RT-PCR ( $n=4$ ). In (a) and (c), data are collected from four separate experiments and are expressed as mean  $\pm$  s.d. \*\* $P < 0.01$  (two-tailed Student's t-test).

The Aire expression is induced in mTECs through interaction with cells such as lymphoid tissue inducers, positively selected TCR $\alpha\beta^+$  thymocytes, and TCR V $\gamma 5^+$  dendritic epidermal T-cell progenitors<sup>30-33</sup>, which is mainly mediated by the signals through tumour necrosis factor receptor family members, such as receptor activator of NF- $\kappa$ B (RANK), CD40 and lymphotoxin  $\beta$  receptor (Lt $\beta$ R)<sup>34-36</sup>. When 2-DG-treated WT fetal thymic stroma was stimulated *in vitro* with RANKL, the Aire expression was markedly induced in a subset of UEA-1<sup>+</sup> mTECs (Fig. 2e and Supplementary Fig. 3c). Although CD40L or anti-Lt $\beta$ R antibody alone was not effective, they synergistically act with RANKL to augment Aire expression in WT mTECs (Fig. 2e and Supplementary Fig. 3c). However, the Aire expression was hardly detected in Jmjd6<sup>-/-</sup> mTECs even after stimulation with RANKL in combination with CD40L or anti-Lt $\beta$ R antibody (Fig. 2e and Supplementary Fig. 3c). This was again confirmed by flow cytometric analyses (Fig. 2f). These results indicate that Jmjd6 is critical for Aire expression in mTECs.

**Intronic regulation of Aire gene by Jmjd6.** As Jmjd6 is a nuclear protein that catalyses lysyl hydroxylation of multiple substrates<sup>22-25</sup>, it seemed likely that Jmjd6 deficiency affects transcription of both upstream and downstream genes of Aire. To comprehensively identify genes controlled by Jmjd6, we prepared 2-DG-treated fetal thymic stroma with or without RANKL stimulation from WT and Jmjd6<sup>-/-</sup> mice and analysed their transcriptomes by RNA sequencing (RNA-seq). Although the gene expression of Jmjd6 was unchanged between before and after stimulation (Supplementary Fig. 4), 1,850 genes were induced in

response to RANKL stimulation, among which 1,020 genes were expressed in WT fetal thymic stroma at significantly higher levels than those in Jmjd6<sup>-/-</sup> samples (see top 200 genes in Supplementary Table 1). These included 23 genes encoding Aire-dependent TSAs such as insulin 2 and salivary protein 1 (refs 5,6,8) (Fig. 3a). This reduction of Aire-dependent TSAs was further confirmed by quantitative real-time PCR using samples from Jmjd6<sup>-/-</sup> fetal thymic stroma and Jmjd6<sup>-/-</sup> E18.5 thymi (Fig. 3b and Supplementary Fig. 5). On the other hand, Jmjd6 deficiency did not affect gene expression of Aire-independent TSA, glutamate decarboxylase 67 (GAD67)<sup>5,6,8</sup> (Fig. 3b and Supplementary Fig. 5). Similarly, gene expression of CD80 and CD40 were unaffected in the absence of Jmjd6 (Fig. 3c). Consistent with this finding, immunohistochemical analysis revealed that RANKL-induced CD80 induction normally occurred even in Jmjd6<sup>-/-</sup> mTECs (Fig. 3d). Unexpectedly, however, the abundance of Aire transcript was also comparable between RANKL-stimulated WT and Jmjd6<sup>-/-</sup> samples (Fig. 3b,c), suggesting that Jmjd6 controls Aire expression through a posttranscriptional mechanism.

To explore the underlying mechanism, we next analysed pre-mRNA splicing. Among 84,708 introns of detected genes, 1,051 introns were selected, because they were expressed at a relatively high frequency (intronic FPKM > 10). Bioinformatics analysis identified 57 introns preferentially expressed in Jmjd6<sup>-/-</sup> thymic stroma under RANKL-stimulated condition (Supplementary Table 2). These included the intron 2 of Aire, which was found in RANKL-treated Jmjd6<sup>-/-</sup> samples at 1.9 times higher frequency than that in the WT sample (Fig. 3e). This was further confirmed by RT-PCR followed by Southern blotting (Fig. 3f).



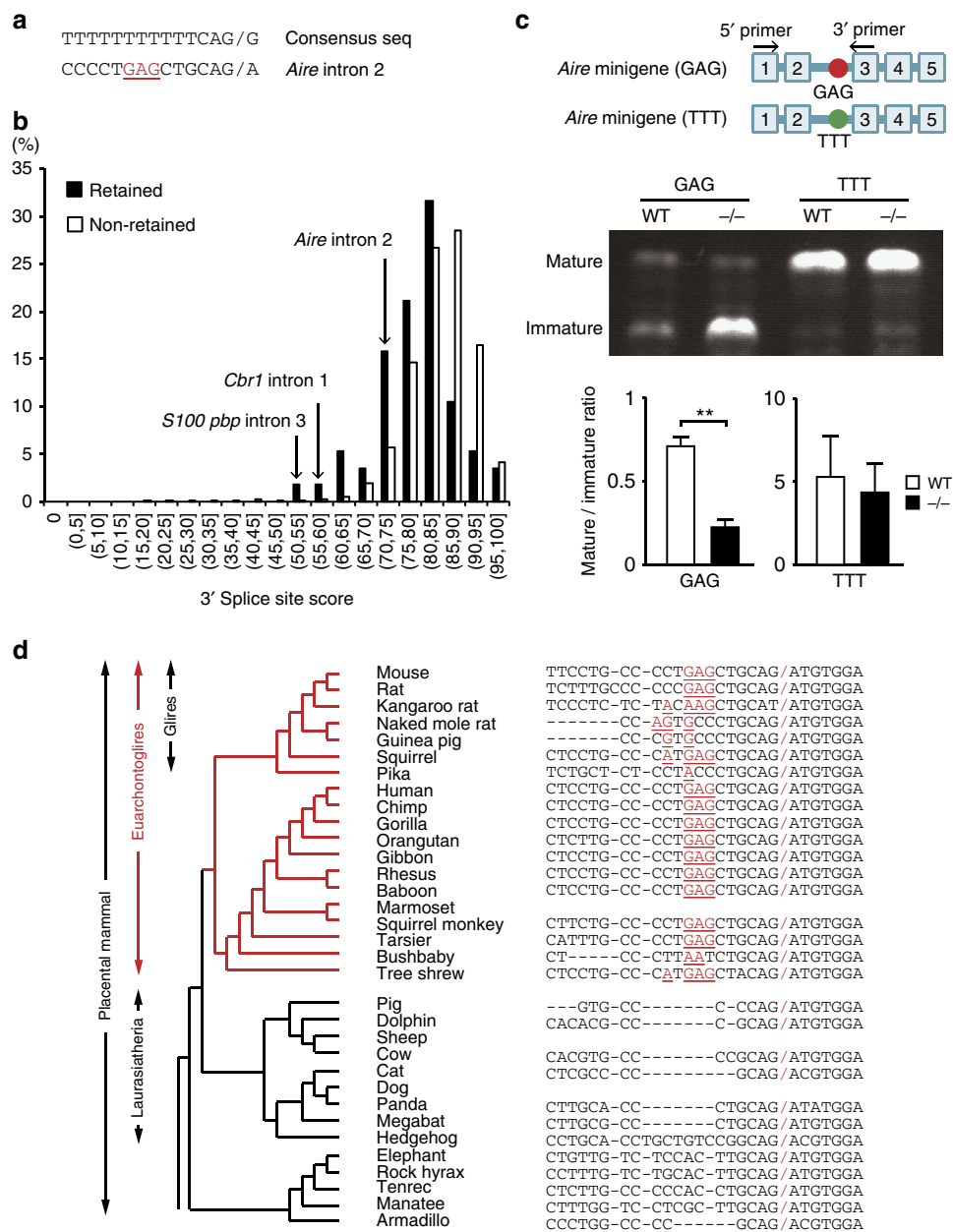
**Figure 5 | The effect of immature Aire protein on localization and stability of mature Aire protein.** (a). Subcellular localization of GFP-tagged mature Aire protein and mCherry-tagged immature Aire protein generated by intron 2 retention. Data are representative of two independent experiments. (b) Schematic representation of pCI-GFP-mature Aire and pTRE2hyg-HA-immature Aire. (c,d) After treatment with doxycycline (Dox) for the indicated times in the presence or absence of p MG132, expressions of mature and immature Aire proteins were analysed for HEK293 cells stably expressing both pCI-GFP-mature Aire and pTRE2hyg-HA-immature Aire or pTRE2hyg. Data (mean  $\pm$  s.d.,  $n = 3$ ) are expressed as the level of mature Aire protein after normalization of the 0 h value to an arbitrary value of 1.  $**P < 0.01$  (two-tailed Student's  $t$ -test).

Amplification of *Aire* cDNA with primers specific for the sequence of the exons 1 and 10 yielded four bands corresponding to the *Aire* transcript with or without retention of intron 2, intron 9 and intron 2 plus intron 9 (Fig. 3g). By measuring the intensity of each band, 41.5% of the *Aire* transcripts expressed in RANKL-stimulated WT fetal thymic stroma were found to be mature form, whereas this value decreased to 18.6% in *Jmjd6*<sup>-/-</sup> samples (Fig. 3g), because of increase in the frequency of retention of intron 2 and/or intron 9. Similar results were obtained when E18.5 thymi of WT and *Jmjd6*<sup>-/-</sup> embryos were analysed with this method (Fig. 3h).

It is known that Aire is also expressed in the reproductive organs and embryonic stem (ES) cells<sup>8,37,38</sup>. To examine whether a similar mechanism operates in ES cells, we developed *Jmjd6*<sup>-/-</sup> ES cells and compared their *Aire* transcripts with those of WT ES cells. *Jmjd6* deficiency in ES cells markedly increased the *Aire* transcript containing intron 2 (Fig. 4a), indicating that Aire expression in ES cells is also controlled by *Jmjd6* through intron retention. Recent structural analysis of the catalytic domain of *Jmjd6* indicated amino acid residues critical for binding to Fe(II), 2-oxoglutarate and substrate lysine<sup>39</sup>. When five of these amino acid residues were mutated to alanine (designated 5 A mutant), the catalytic activity of *Jmjd6* to

hydroxylate lysine residues was completely lost (Fig. 4b). Although the transient expression of WT *Jmjd6* in *Jmjd6*<sup>-/-</sup> ES cells significantly, albeit incompletely, improved the ratio of mature *Aire* transcript (without intron 2) to immature transcript (with intron 2), such improvement was not achieved by the 5 A mutant (Fig. 4c). These results suggest that *Jmjd6* controls splicing events of *Aire* gene depending on its enzymatic activity.

**The nature of immature Aire protein.** The retention of intron 2 results in an appearance of a premature termination codon at the N-terminal portion of Aire (Fig. 5a and Supplementary Fig. 6), leaving only 103 amino acid residues presumably intact. To know the nature of this immature Aire protein generated by intron 2 retention, we first analysed its subcellular localization. As expected, the GFP-tagged mature Aire protein was localized to the nucleus when expressed alone in MEFs (Fig. 5a). On the other hand, the mCherry-tagged immature Aire protein preferentially accumulated in the cytoplasm, owing to the lack of the nuclear localization signal (Fig. 5a). Interestingly, localization of mature Aire protein was changed from the nucleus to the cytoplasm when immature Aire protein was co-expressed (Fig. 5a). Exactly the same results were obtained when GFP- and mCherry-tags were exchanged (Supplementary Fig. 7). This finding led us to



**Figure 6 | The 3' splice site sequence is critical for intron 2 retention of *Aire* gene.** (a) Comparison of 3' splice site sequence of the *Aire* intron 2 with the consensus sequence. /, splice site. (b) Association between intron retention and 3' splice site score revealed by comparing 57 retained introns with 188,151 unretained introns. (c) Effect of TTT mutation on intron 2 retention of minigene. Twenty-four hours after the transfection of WT and *Jmjd6*<sup>-/-</sup> MEFs with GAG- or TTT-type minigene, the ratio of mature to immature transcript was analysed by RT-PCR ( $n=3$ ). Data are collected from three separate experiments and are expressed as mean  $\pm$  s.d.  $**P < 0.01$  (two-tailed Student's *t*-test). (d) Comparison of the 3' splice site sequence among different species. Non-canonical sequences at PPT site are indicated in red.

examine whether the presence of immature Aire protein affects stability of mature Aire protein. For this purpose, we developed HEK293 cells that constitutively express mature Aire protein, but inducibly express immature Aire protein when exposed to doxycycline (Fig. 5b). Doxycycline itself did not affect stability of mature Aire protein (Fig. 5c). However, the amount of mature Aire protein decreased, as the expression of immature Aire protein increased in response to doxycycline treatment (Fig. 5c). Since this reduction of mature Aire protein level was partially inhibited by treating cells with the proteasome inhibitor MG132 (Fig. 5d), it was suggested that mature Aire protein sequestered into the cytoplasm becomes susceptible to proteasome-dependent protein degradation.

**A cis-regulatory element for intron retention of *Aire* gene.** To understand why the intron 2 of *Aire* gene is susceptible to intron retention, we compared the sequence structure between Aire intron 2 and other introns. Although the most introns have the PPT site immediately upstream of the 3' terminal AG dinucleotide<sup>16</sup>, *Aire* intron 2 encodes GAG instead of canonical pyrimidine-rich sequence at this position (Fig. 6a), resulting in 'low' 3' splice site score, which is calculated based on the similarity to the consensus sequence (Fig. 6a). Comparison of 57 retained introns with 188,151 unretained ones suggested that the degree of intron retention is associated with 3' splice site score (Fig. 6b). Indeed, PCR with reverse transcription (RT-PCR) analyses revealed that intron 3 of *S100pbp* gene and intron 1 of



*Cbr1* gene with low 3' splice site score are preferentially retained in the absence of *Jmjd6* (Supplementary Fig. 8).

To directly examine the effect of GAG sequence on the intron 2 retention, we created *Aire* minigene containing exons 1–5 surrounded by their intronic regulatory sequences with GAG or TTT at PPT site of the intron 2 (designated GAG-type or TTT-type) (Fig. 6c). When the GAG-type minigene was expressed in WT or *Jmjd6*<sup>-/-</sup> MEFs, *Jmjd6* deficiency markedly increased the transcript containing intron 2 (Fig. 6c), which was consistent with the results on endogenous *Aire* gene expression in mTECs and ES cells (Figs 3g and 4a). However, TTT-type minigene yielded only mature transcript without intron 2 retention, irrespective of *Jmjd6* expression (Fig. 6c). Thus, GAG sequence acts as a *cis*-regulatory element that causes intron 2 retention and inhibits *Aire* protein expression. Interestingly, this GAG sequence at PPT site of *Aire* intron 2 is highly conserved in mammals in the Euarchontoglires clade (Fig. 6d). Therefore, intron retention may have been evolved as a mechanism to prevent overexpression of *Aire* protein in these species.

## Discussion

Intron retention is widely accepted as a consequence of mis-splicing, and its significance has been overlooked. However, recent evidence indicates that intron retention has a physiological role in some biological settings such as granulopoiesis<sup>40</sup>. Here we have demonstrated that the expression of *Aire* is controlled by *Jmjd6* through intron retention. Although *Jmjd6* deficiency did not affect abundance of *Aire* transcript, the intron 2 of *Aire* gene was not effectively spliced out in the absence of *Jmjd6*, resulting in a marked reduction of mature *Aire* protein in mTECs. In both *Jmjd6*<sup>-/-</sup> E18.5 thymi and RANKL-stimulated *Jmjd6*<sup>-/-</sup> fetal thymic stroma, the reduction of *Aire* protein was more prominent than that of mature *Aire* transcript. The exact reason for this discrepancy remains unclear. However, reconstitution experiments revealed that immature *Aire* protein generated by intron 2 retention affects subcellular localization and stability of mature *Aire* protein. Therefore, the ratio of mature *Aire* protein to the immature form might be a critical factor that determines expression and function of *Aire* protein in the nucleus.

Ectopic expression of peripheral TSAs by mTECs has been viewed as an essential mechanism for induction of central tolerance<sup>3–8</sup>. Consistent with a reduction of mature *Aire* protein in mTECs, RNA-seq analyses revealed that the expressions of 23 *Aire*-dependent TSAs were markedly reduced in *Jmjd6*<sup>-/-</sup> thymic stroma. This was further confirmed by quantitative real-time PCR analyses. In addition, by grafting *Jmjd6*<sup>-/-</sup> thymic stroma into athymic C57BL/6 nude mice, we have shown that T-cells selected to mature in *Jmjd6*<sup>-/-</sup> thymic micro-environments caused multi-organ autoimmunity. As the number of thymocytes was significantly reduced in the grafted *Jmjd6*<sup>-/-</sup> thymus, disease manifestation in *nu/nu*<sup>*Jmjd6*<sup>-/-</sup></sup> mice might be exaggerated by homeostatic T-cell proliferation<sup>41</sup>. However, it has been reported that, while lymphocytes from *Aire*-deficient mice cause autoimmune disease when transferred into recombination-activating gene (*Rag*)-deficient recipients, adoptive transfer of WT lymphocytes fail to induce disease under the same condition<sup>5,15</sup>. Thus, it is likely that reduction of *Aire*-dependent TSAs underlies disease development in *nu/nu*<sup>*Jmjd6*<sup>-/-</sup></sup> mice.

Mechanistically, our results suggest that *Aire* expression is tightly controlled via two discrete steps. First, owing to the GAG sequence at PPT site, the intron 2 of *Aire* gene is highly susceptible to intron retention in Euarchontoglires, and as a result, *Aire* protein expression is expected to be kept at a low level. As *Aire* has been reported to act as a proapoptotic factor<sup>42</sup>, this static regulation may be important to avoid a deleterious

effect of *Aire* overexpression on the immune system or reproductive organs. The second important regulation is relief of intron retention. This is a dynamic process involving the enzymatic activity of *Jmjd6*, thereby raising the possibility that metabolic status and oxygen tension may influence intron retention. Although the direct substrate of *Jmjd6* in this context is currently unknown, *Jmjd6* interacts with multiple splicing regulatory proteins including U2 small nuclear ribonucleoprotein auxiliary factor 65 kDa (U2AF65)<sup>22,23</sup>. Therefore, it seems likely that *Jmjd6* could alter affinity of a given splicing factor to 3' splice site of *Aire* gene through lysyl hydroxylation. Our findings thus define a previously unknown mechanism controlling expression of *Aire* protein critical for establishment of immunological self in the thymus.

## Methods

**Mice.** *Jmjd6*<sup>-/-</sup> mice have been described previously<sup>27</sup>. Mice heterozygous for the mutant allele (*Jmjd6*<sup>+/-</sup>) were backcrossed onto a C57BL/6 background for more than 10 generations, and *Jmjd6*<sup>+/-</sup> mice were crossed to obtain *Jmjd6*<sup>-/-</sup> embryos. The morning of finding the vaginal plug was designated as E 0.5. B6.Cg-Foxn1<nu>/Nrs (nude) female mice were purchased from Taconic or provided by RIKEN BRC through National Bio-Resource Project of the MEXT, Japan and were used as recipients of thymic grafts at the age of 6–8 weeks. Mice were kept under specific pathogen-free conditions in the animal facility of Kyushu University. The protocol of animal experiments was approved by the committee of Ethics of Animal Experiments, Kyushu University.

**Fetal liver chimera.** Fetal livers were harvested from E14.5 embryos. A total of 1 × 10<sup>6</sup> fetal liver cells were injected in a volume of 300 μl phosphate-buffered saline (PBS) into the lateral tail veins of irradiated (10 Gy) CD45.1 C57BL/6 recipients. Chimeras were analysed 10 weeks after reconstitution.

**Fetal thymus organ culture (FTOC) and thymic graft.** Thymic lobes were isolated from E15.5 WT and *Jmjd6*<sup>-/-</sup> embryos, and were cultured for 4 days on Nucleopore filters (Whatman) placed in RPMI 1640 medium (Life Technologies) supplemented with 10% heat-inactivated fetal calf serum (FCS) (Nishirei Bioscience), 50 μM 2-mercaptoethanol (Nacalai tesque), 2 mM L-glutamine (Life Technologies), 100 U ml<sup>-1</sup> penicillin (Life Technologies), 100 μg ml<sup>-1</sup> streptomycin (Life Technologies), 1 mM sodium pyruvate (Life Technologies), MEM non-essential amino acids (Life Technologies) and 1.35 mM 2-DG (Sigma-Aldrich). After cultivation in complete RPMI medium without 2-DG for one more day, four pieces of WT and *Jmjd6*<sup>-/-</sup> fetal thymi were grafted under the renal capsule of athymic (nude) and WT C57BL/6 mice. Thymic chimeras were analysed 6–10 weeks after transplantation. In some experiments, TECs were stimulated *in vitro* with recombinant RANKL (1 μg ml<sup>-1</sup>; Peprotech), agonistic anti-LtβR antibody 3C8 (2 μg ml<sup>-1</sup>; eBioscience) and/or recombinant CD40L (5 μg ml<sup>-1</sup>; R&D systems) for 4 days before analyses.

**Flow cytometry.** The following antibodies and reagents were used at the indicated concentrations. Fluorescein isothiocyanate (FITC)-conjugated anti-mouse CD3ε (145-2c11, 5 μg ml<sup>-1</sup>), biotinylated anti-mouse CD4 (RM4-5, 5 μg ml<sup>-1</sup>), phycoerythrin (PE)-conjugated anti-mouse CD45 (30-F11, 1 μg ml<sup>-1</sup>), FITC-conjugated anti-mouse CD62L (MEL-14, 10 μg ml<sup>-1</sup>), FITC-conjugated anti-mouse CD90.2 (30-H12, 2.5 μg ml<sup>-1</sup>), FITC-conjugated anti-mouse Foxp3 (FJK-16 s, 5 μg ml<sup>-1</sup>), PE-conjugated anti-mouse MHC class II (I-A/I-E) (M5/114.15.2, 0.2 μg ml<sup>-1</sup>) and PerCP-Cyanine5.5 conjugated streptavidin (0.4 μg ml<sup>-1</sup>) were purchased from eBioscience. PE-conjugated anti-mouse CD8a (53-6.7, 2 μg ml<sup>-1</sup>), PE-conjugated anti-mouse CD25 (PC61, 2 μg ml<sup>-1</sup>), PE-conjugated anti-mouse CD44 (IM7, 0.6 μg ml<sup>-1</sup>), FITC-conjugated anti-mouse CD45 (30-F11, 10 μg ml<sup>-1</sup>), FITC-conjugated CD45R/B220 (RA3-6B2, 10 μg ml<sup>-1</sup>), PE-conjugated anti-mouse CD80 (16-10A1, 1 μg ml<sup>-1</sup>), biotinylated anti-mouse CD45.2 (104, 5 μg ml<sup>-1</sup>) and APC-conjugated streptavidin (0.4 μg ml<sup>-1</sup>) were from BD Bioscience. Before staining with the antibodies, the cells were incubated for 10 min on ice with anti-Fcγ III/II receptor (2.4G2, 0.5 μg ml<sup>-1</sup>; BD Bioscience) to block Fc receptors. Foxp3 intracellular staining was performed with a staining kit (eBioscience) according to the manufacturer's recommendations. For intracellular *Aire* staining of mTECs, cells were first stained with anti-CD45 antibody and UEA-1 (10 μg ml<sup>-1</sup>; VECTOR Laboratories), followed by fixation in 4% (w/v) paraformaldehyde for 15 min at room temperature. After being permeabilized with 0.1% saponin (Sigma-Aldrich), cells were stained with Alexa Fluor 488-conjugated anti-mouse *Aire* (5H12, 5 μg ml<sup>-1</sup>; eBioscience) for 30 min on ice. Flow cytometric analyses were done on FACS Calibur (BD Bioscience).

**Histology and immunofluorescence.** Tissues were fixed in 4% (w/v) paraformaldehyde and embedded in paraffin blocks. Sections (3 μm thick)

were stained with haematoxylin and eosin, and examined by light microscopy. For immunofluorescence analyses, tissues were embedded in OCT compound (Sakura Finetech) and frozen at  $-80^{\circ}\text{C}$ . Cryostat sections ( $8\ \mu\text{m}$  thick) were fixed by immersion for 10 min in ice-cold acetone and blocked with 10% horse serum (Sigma-Aldrich) for 1 h at room temperature. Sections were then incubated with rabbit anti-mouse keratin 5 ( $1\ \mu\text{g}\ \text{ml}^{-1}$ ; Covance), rat anti-mouse keratin 8 (TROMA-1,  $1:1,000$  dilution; Developmental Studies Hybridoma Bank), biotinylated UEA-1 ( $10\ \mu\text{g}\ \text{ml}^{-1}$ ; VECTOR Laboratories), Alexa Fluor 488-conjugated anti-mouse Aire ( $5\text{H}12$ ,  $2.5\ \mu\text{g}\ \text{ml}^{-1}$ ; eBioscience) and/or PE-conjugated anti-mouse CD80 (16-10A1,  $1\ \mu\text{g}\ \text{ml}^{-1}$ ; eBioscience) for 1 h at room temperature. After being washed with PBS, sections were incubated with appropriately labelled secondary antibodies or reagents. For *Jmjd6* staining, cells were cultured on the poly-L-lysine coated glass-bottom culture dishes (Matsunami), fixed with 4% (w/v) paraformaldehyde for 15 min, and permeabilized with 0.1% Triton X-100 for 1 h. After being blocked with 10% horse serum for 30 min at room temperature, cells were then stained with 4',6-diamidino-2-phenylindole (DAPI) (Dojindo Laboratories), Alexa Fluor 546-conjugated phalloidin (Life Technologies), and mouse anti-*Jmjd6* antibody (mAb328,  $5\ \mu\text{g}\ \text{ml}^{-1}$ ; Synaptic Systems) for 1 h at room temperature, followed by incubation with Alexa Fluor 488-conjugated donkey anti-mouse IgG (Fab fragment,  $1.5\ \mu\text{g}\ \text{ml}^{-1}$ ; Jackson ImmunoResearch). To detect autoantibodies, acetone-fixed cryosections of C57BL/6 nude mouse tissues were incubated with diluted serum (1:40) obtained from grafted mice, followed by staining with Alexa Fluor 488-conjugated anti-mouse IgG ( $4\ \mu\text{g}\ \text{ml}^{-1}$ ; Life Technologies). All images were obtained with a laser scanning confocal microscope (LSM510 META; Carl Zeiss).

**Cell preparation and culture.** To enrich TECs, thymic lobes were prepared from E18.5 or 6–8-week-old C57BL/6 mice, cut into small pieces, and dispersed further with pipetting to remove the majority of thymocytes. The resulting thymic fragments were digested with 0.125% (w/v) collagenase D/dispase (Roche) and 0.1% (w/v) DNase I (Roche) in RPMI1640 medium for 1 h at  $37^{\circ}\text{C}$ . The supernatants containing dissociated TECs were centrifuged and were washed with PBS. Before cell sorting, TECs were further enriched by depleting  $\text{CD}45^{+}$  haematopoietic cells using  $\text{CD}45$  MicroBeads (Miltenyi Biotec), and stained with the relevant antibodies and reagents. Then, mTECs and cTECs were sorted as  $\text{CD}45^{-}$  MHC class II<sup>+</sup> UEA-1<sup>+</sup> cells and  $\text{CD}45^{-}$  MHC class II<sup>+</sup> UEA-1<sup>-</sup> cells, respectively. ES cells were developed from E3.5 blastocysts by using standard procedures. Integrity of ES cells was confirmed by staining them for alkaline phosphatase with a kit (Wako Pure Chemical Industries). ES cells were cultured in knock-out D-MEM medium (Life Technologies) supplemented with 15% FCS (Life Technologies),  $50\ \mu\text{M}$  2-mercaptoethanol (Nacalai tesque),  $2\ \text{mM}$  L-glutamine (Life Technologies),  $100\ \text{U}\ \text{ml}^{-1}$  penicillin (Life Technologies),  $100\ \mu\text{g}\ \text{ml}^{-1}$  streptomycin (Life Technologies) and ESGRO leukemia inhibitory factor (LIF) (Millipore) at final concentration of 1,500 units per ml. On the other hand, primary MEFs were generated from E13.5 WT and *Jmjd6*<sup>-/-</sup> embryos. Primary MEFs were immortalized by transfection with a plasmid pCX4bsr-SV40ER (provided by T. Akagi, KAN Research Institute, Kobe, Japan). Immortalized MEFs were cultured in D-MEM medium (Wako Pure Chemical Industries) supplemented with 10% FCS (Nichirei Bioscience),  $100\ \text{U}\ \text{ml}^{-1}$  penicillin (Life Technologies) and  $100\ \mu\text{g}\ \text{ml}^{-1}$  streptomycin (Life Technologies).

**RNA-seq analysis and 3' splice site scoring.** Total RNA was extracted from WT and *Jmjd6*<sup>-/-</sup> FTOC samples with (two samples for each category) or without (one sample for each category) RANKL stimulation. One microgram of total RNA was used for library construction with TruSeq RNA Sample Prep kit v2 (Illumina) according to the manufacturer's protocol. Briefly, poly-A-containing mRNAs were purified using poly-T oligo-attached magnetic beads. The purified mRNAs were fragmented using divalent cations under elevated temperatures and then converted to dsDNA by two rounds of cDNA synthesis using reverse transcriptase and DNA polymerase I. After an end repair process, DNA fragments were ligated with adaptor oligos. The ligated products were amplified by eight cycles of PCR to generate RNA-seq library. Library integrity was verified by Bioanalyzer DNA1000 assay (Agilent Technologies). Sequencing was performed in 101-bp paired-end mode using an Illumina HiSeq (Illumina). A total of 177,060,020 reads were obtained for six samples. Filtered reads were mapped to the UCSC mm10 using the TopHat program (v2.0.10)<sup>43</sup> with the default parameters. The Cufflinks program (v2.1.1)<sup>44</sup> was then used to assemble 22,448 transcripts and to calculate the fragments per kilobase of exon per million mapped fragments (FPKM) values, which are normalized measurement of gene expression levels, with the non-default parameters: -u—library-type fr-secondstrand. To identify differentially expressed genes, the ratio of the maximum FPKM to the minimum FPKM was compared among six samples. When the ratio was more than 3, the gene was regarded as being significantly altered in expression level. We added 0.1 to the FPKM value to avoid division by 0. This led us to identify 3,212 genes with differential expression. Among these, the expression levels of 2,536 genes were significantly associated with either RANKL treatment or *Jmjd6* expression ( $P$ -value  $< 0.05$ ), and these genes were used for further analyses. Analysis of intron retention was performed as follows. According to the current gene annotation ('known genes' in UCSC mm10), there are 188,208 introns in total. As intron retention events should be observed in the genes with relatively high expression, we only focused on the genes with the

maximum FPKM value more than 10 at least in one of the six samples. As a result, we obtained 84,708 introns. The reads mapped to these intronic regions were counted by the intersectBed program in the BEDTools utilities (v2.17.0)<sup>45</sup> with -c option, and the counts are converted into the FPKM values for each intron (intrinsic FPKM). There are 1,051 introns with intrinsic FPKM more than 10 for at least one of the six samples, and the degree of intron retention (IR value) was calculated by dividing intrinsic FPKM value by conventional FPKM value for each gene. By filtering IR value of *Jmjd6*<sup>-/-</sup> sample to that of WT sample more than 1.5, we finally selected 57 introns that are preferentially expressed in *Jmjd6*<sup>-/-</sup> samples under RANKL stimulation. The 3' splice site score was calculated by 'Splice-Site analyser tool' (<http://ibis.tau.ac.il/ssat/SpliceSiteFrame.htm>)<sup>46</sup>. Shortly, the score expresses to what extent the splice-site sequences match the following consensus sequence: TTTTTTTTTTCAG/G ('/' indicate the intron/exon junction).

**RT-PCR and Southern blot analysis.** After treatment with RNase-free DNase I (Life Technologies), RNA samples were reverse-transcribed with oligo(dT) primers (Life Technologies) and SuperScript III reverse transcriptase (Life Technologies) for amplification by PCR. The following PCR primers were used for RT-PCR and/or Southern blot analysis. *Aire* (exon 1 – exon 10); 5'-AGGTGGGATGGAA TGCTA-3' and 5'-CCTCATTCCCAGCACTCAGTA-3'. *Aire* (exon 1 – exon 3); 5'-TAGACAGTGCCTTCCGGCTGCTG-3' and 5'-GGAGACGCTCTTGGAG GCCAGATTG-3'. *Actb*; 5'-TGGAACTCTGTGGCATTCCATGAAAC-3' and 5'-TAAAACGCAGCTCAGTAACAGTCCG-3'. *Jmjd6*; 5'-ACCACAAGAGCAAG AAGCGCATCC-3' and 5'-AACCACGGGCTGTGAAGGCCTC-3'. *S100bpb*; 5'-TCCTCCAAAGAAACGGAAA-3', 5'-CCTAAGTGCGGGGATTACAA-3', and 5'-TCCCCAGAATCCTGTGCAAA-3'. *Cbr1*; 5'-GGTGACCGGTGCTAACA AAG-3', 5'-ATTGGGTGTGGAGATGGAAG-3', and 5'-TCACCTCTGCTTGAA TGTTG-3'. For Southern blot analysis, *Sac I-Sac II* fragment of *Aire* cDNA (560 bp-fragment corresponding to exons 3–8) was used as a probe. Quantification of the intensity was performed with BAS-2500 bio-imaging analyser and the Image Gauge 4.0 software (Fujifilm CO.). Identity of all detected bands was confirmed by sequencing.

Real-time PCR was performed on ABI PRISM 7,000 Sequence Detection System using the SYBR Green PCR Master Mix (both from Applied Biosystems). The following PCR primers were used: *Gapdh*; 5'-TGTGTCCGTGCTGGATCTGA-3' and 5'-TTGCTGTGAAGTCGCAGGAG-3'. *Aire*; 5'-GAAGCAGTATGGCAA CTTGGA-3' and 5'-ACAGAAGCTGCCATGGTCTGAA-3'. *Psmb11*; 5'-CTCTGT GGCTGGGACTACTC-3' and 5'-TCCGCTCTCCGAACGTGG-3'. *Col1a2*; 5'-CCGTGCTTCTCAGAACATCA-3' and 5'-GAGCAGCCATCGACTAGGAC-3'. *Jmjd6*; 5'-AGAAACGCAGACCCCTTAC-3' and 5'-CCCTGAACATAGGCATT CCA-3'. *Ins2*; 5'-CCACCAGCCCTAAGTGATCC-3' and 5'-TAGAGAGCCTCC ACCAGGTG-3'. *Spt1*; 5'-ACTCCTTGTGTTGCTGTGTTT-3' and 5'-TCGAC TGAATCAGAGAACTCAACT-3'. *Gad1*; 5'-CACCTGGAACCCCTACAA-3' and 5'-ACTGCTTGTCTGGCTGGA-3'. All targets were standardized to *Gapdh* signals.

**Plasmids and transfection.** For expression of N-terminally-tagged GFP- or mCherry- fusion proteins in mammalian cells, expression vector was created by subcloning a cDNA encoding EGFP or mCherry into pCI (Promega). The genes encoding the full-length Aire protein (mature Aire) and the truncated Aire protein (immature Aire) containing only exon 1 and exon 2 owing to intron 2 retention were subcloned into these vectors to analyse their subcellular localization in MEFs. Transfection was performed with Lipofectamine 2,000 reagent (Life Technologies). To create inducible expression system, the gene encoding HA-tagged immature Aire protein was subcloned in the pTRE2hyg vector (Clontech) (designated pTRE2hyg-HA-immature Aire). After electroporation of pCI-GFP-mature Aire and pTRE2hyg-HA-immature Aire or pTRE2hyg into HEK293 Tet-On Advanced cells (Clontech), cells were selected with  $150\ \mu\text{g}\ \text{ml}^{-1}$  hygromycin B (Wako Pure Chemical Industries) to develop stable transfectants. *Aire* minigene construct (GAG-type; 2.4 kb) was amplified using mouse genome DNA and subcloned into TAget Clone TM Plus (TOYOBO Co. Ltd.). After digestion with *Sal I* and *Not I*, the minigene construct was inserted into pSI vector (Promega) for transfection. The TTT-type minigene was generated by site-directed mutagenesis. The gene encoding the C-terminally GFP-tagged WT *Jmjd6* was subcloned into pCI vector (designated pCI-*Jmjd6*-GFP). The 5 A mutant encoding alanines instead of H187, D189, K204, T285 and N287 was generated by site-directed mutagenesis. For splicing assay, *Aire* minigene plasmids ( $1\ \mu\text{g}$ ) was transfected into immortalized WT or *Jmjd6*<sup>-/-</sup> MEFs ( $2 \times 10^5$  cells) in 6-well plates with Lipofectamine 2000 reagent. To express *Jmjd6* and its mutant in *Jmjd6*<sup>-/-</sup> ES cells, these ES cells were electroporated with pCI-*Jmjd6*-GFP vector (WT or 5 A mutant) by using the Mouse ES Cell Nucleofector Kit (Lonza/Amaxa Biosystems). The GFP-positive ES cells were then sorted by FACSaria for RNA extraction 30 h after transfection.

**Immunoblotting.** Cells were lysed on ice in 20 mM Tris-HCl buffer (pH 7.5) containing 1% Triton X-100, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 2.5 mM sodium pyrophosphate, 1 mM  $\beta$ -glycerol phosphate, 1 mM  $\text{Na}_2\text{VO}_4$ , and complete protease inhibitors (Roche). After centrifugation, the supernatants were mixed with an equal volume of  $2 \times$  sample buffer (125 mM Tris-HCl, 0.01% bromophenacyl bromide, 4% SDS, 20% glycerol and 200  $\mu\text{M}$  dithiothreitol). Samples were boiled

for 5 min and analysed by immunoblotting. The following antibodies were used: goat anti-actin (I-19; 1:1,000 dilution; Santa Cruz), mouse anti-GFP (B-2; 1:1,000 dilution; Santa Cruz), and rat anti-HA (3F10; 1:500 dilution; Roche). Before assays, HEK293 cells stably expressing both pCI-GFP-mature Aire and pTRE2hyg-HA-immature Aire or pTRE2hyg were treated with doxycycline (Dox; 10  $\mu\text{g ml}^{-1}$ ; Clontech) in the presence or absence of MG132 (10  $\mu\text{M}$ ; Sigma-Aldrich) for specified times. The signals were measured using ImageJ public-domain software (imagej.nih.gov/ij/) and standardized to  $\beta$ -actin. In Fig. 5c,d, images have been cropped for presentation. Full size images are presented in Supplementary Figure 9.

**Hydroxylation assays.** Hydroxylation assays were performed as previously described<sup>22,39</sup>. Briefly, the assays consisted of the substrate mixture: LUC7-like2 peptide (LUC7L2<sub>267–278</sub>; 100  $\mu\text{M}$ ), 2-oxoglutarate (500  $\mu\text{M}$ ), ascorbate (100  $\mu\text{M}$ ) in Tris (50 mM, pH 7.5) and the enzyme mixture: FeNH<sub>4</sub>SO<sub>4</sub> (100  $\mu\text{M}$ ) and recombinant Jmjd6 (100  $\mu\text{M}$ ) in Tris (50 mM, pH 7.5). The reaction was initiated by mixing the substrate and enzyme mixtures in a final volume of 20  $\mu\text{l}$  and by incubating at 37 °C for 1 h. The reaction was quenched by adding 10  $\mu\text{l}$  of 0.1% (w/v) CF<sub>3</sub>CO<sub>2</sub>H. Mass spectrometry analysis was performed using an Autoflex matrix-assisted laser desorption/ionization (MALDI) time-of-flight (TOF) mass spectrometer (Bruker Daltonics). The quenched reaction mixture was diluted 10-fold with 0.1% CF<sub>3</sub>CO<sub>2</sub>H and 2% acetonitrile, 1  $\mu\text{l}$  of which was directly spotted onto 0.5  $\mu\text{l}$  of  $\alpha$ -cyano-4-hydroxycinnamic acid matrix on the MALDI target plate and allowed to dry. The mass corresponding to LUC7L2<sub>267–278</sub> and LUC7L2<sub>267–278</sub> + 16, equivalent to one hydroxylation, were identified.

**Statistical analysis.** For statistical analysis, *P* values were calculated with a two-tailed unpaired Student's *t*-test. *P* values <0.05 were considered significant. Error bars denote  $\pm$  s.d.

## References

- Ladi, E., Yin, X., Chtanova, T. & Robey, E. A. Thymic microenvironments for T cell differentiation and selection. *Nat. Immunol.* **7**, 338–343 (2006).
- Anderson, G. & Takahama, Y. Thymic epithelial cells: working class heroes for T cell development and repertoire selection. *Trends Immunol.* **33**, 256–263 (2012).
- Derbinski, J., Schulte, A., Kyewski, B. & Klein, L. Promiscuous gene expression in medullary thymic epithelial cells mirrors the peripheral self. *Nat. Immunol.* **2**, 1032–1039 (2001).
- Kyewski, B. & Klein, L. A central role for central tolerance. *Annu. Rev. Immunol.* **24**, 571–606 (2006).
- Anderson, M. S. *et al.* Projection of an immunological self shadow within the thymus by the aire protein. *Science* **298**, 1395–1401 (2002).
- Derbinski, J. *et al.* Promiscuous gene expression in thymic epithelial cells is regulated at multiple levels. *J. Exp. Med.* **202**, 33–45 (2005).
- Peterson, P., Org, T. & Rebane, A. Transcriptional regulation by AIRE: molecular mechanisms of central tolerance. *Nat. Rev. Immunol.* **8**, 948–957 (2008).
- Mathis, D. & Benoist, C. Aire. *Annu. Rev. Immunol.* **27**, 287–312 (2009).
- Blechs Schmidt, K. *et al.* The mouse Aire gene: comparative genomic sequencing, gene organization, and expression. *Genome Res.* **9**, 158–166 (1999).
- Gardner, J. M. *et al.* Deletional tolerance mediated by extrathymic Aire-expressing cells. *Science* **321**, 843–847 (2008).
- Hubert, F. X. *et al.* A specific anti-Aire antibody reveals Aire expression is restricted to medullary thymic epithelial cells and not expressed in periphery. *J. Immunol.* **180**, 3824–3832 (2008).
- Nagamine, K. *et al.* Positional cloning of the APECED gene. *Nat. Genet.* **17**, 393–398 (1997).
- The Finnish-German APECED Consortium. An autoimmune disease, APECED, caused by mutations in a novel gene featuring two PHD-type zinc-finger domains. *Nat. Genet.* **17**, 399–403 (1997).
- Liston, A., Lesage, S., Wilson, J., Peltonen, L. & Goodnow, C. C. Aire regulates negative selection of organ-specific T cells. *Nat. Immunol.* **4**, 350–354 (2003).
- Anderson, M. S. *et al.* The cellular mechanism of Aire control of T cell tolerance. *Immunity* **23**, 227–239 (2005).
- Irimia, M. & Roy, S. W. Origin of spliceosomal introns and alternative splicing. *Cold Spring Harb. Perspect. Biol.* **6**, a016071 (2014).
- Gamazon, E. R. & Stranger, B. E. Genomics of alternative splicing: evolution, development and pathophysiology. *Hum. Genet.* **133**, 679–687 (2014).
- Stamm, S. *et al.* An alternative-exon database and its statistical analysis. *DNA Cell Biol.* **19**, 739–756 (2000).
- Lejeune, F. & Maquat, L. E. Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. *Curr. Opin. Cell Biol.* **17**, 309–315 (2005).
- Klose, R. J., Kallin, E. M. & Zhang, Y. JmjdC-domain-containing proteins and histone demethylation. *Nat. Rev. Genet.* **7**, 715–727 (2006).
- Fadok, V. A. *et al.* A receptor for phosphatidylserine-specific clearance of apoptotic cells. *Nature* **405**, 85–90 (2000).
- Webby, C. J. *et al.* Jmjd6 catalyses lysyl-hydroxylation of U2AF65, a protein associated with RNA splicing. *Science* **325**, 90–93 (2009).
- Heim, A. *et al.* Jumonji domain containing protein 6 (Jmjd6) modulates splicing and specifically interacts with arginine-serine-rich (RS) domains of SR- and SR-like proteins. *Nucleic Acids Res.* **42**, 7833–7850 (2014).
- Unoki, M. *et al.* Lysyl 5-hydroxylation, a novel histone modification, by Jumonji domain containing 6 (JMJD6). *J. Biol. Chem.* **288**, 6053–6062 (2013).
- Wang, F. *et al.* JMJD6 promotes colon carcinogenesis through negative regulation of p53 by hydroxylation. *PLoS Biol.* **12**, e1001819 (2014).
- Li, M. O., Sarkisian, M. R., Mehal, W. Z., Rakic, P. & Flavell, R. A. Phosphatidylserine receptor is required for clearance of apoptotic cells. *Science* **302**, 1560–1563 (2003).
- Kunisaki, Y. *et al.* Defective fetal liver erythropoiesis and T lymphopoiesis in mice lacking the phosphatidylserine receptor. *Blood* **103**, 3362–3364 (2004).
- Böse, J. *et al.* The phosphatidylserine receptor has essential functions during embryogenesis but not in apoptotic cell removal. *J. Biol.* **3**, 15 (2004).
- Klug, D. B. *et al.* Interdependence of cortical thymic epithelial cell differentiation and T-lineage commitment. *Proc. Natl Acad. Sci. USA* **95**, 11822–11827 (1998).
- Rossi, S. W. *et al.* RANK signals from CD4<sup>+</sup>3<sup>+</sup> inducer cells regulate development of Aire-expressing epithelial cells in the thymic medulla. *J. Exp. Med.* **204**, 1267–1272 (2007).
- Hikosaka, Y. *et al.* The cytokine RANKL produced by positively selected thymocytes fosters medullary thymic epithelial cells that express autoimmune regulator. *Immunity* **29**, 438–450 (2008).
- Iria, M. *et al.* Autoantigen-specific interactions with CD4<sup>+</sup> thymocytes control mature medullary thymic epithelial cell cellularity. *Immunity* **29**, 451–463 (2008).
- Roberts, N. A. *et al.* Rank signaling links the development of invariant  $\gamma\delta$  T cell progenitors and Aire<sup>+</sup> medullary epithelium. *Immunity* **36**, 427–437 (2012).
- Akiyama, T. *et al.* The tumor necrosis factor family receptors RANK and CD40 cooperatively establish the thymic medullary microenvironment and self-tolerance. *Immunity* **29**, 423–437 (2008).
- Chin, R. K. *et al.* Lymphotoxin pathway directs thymic Aire expression. *Nat. Immunol.* **4**, 1121–1127 (2003).
- Mouri, Y. *et al.* Lymphotoxin signal promotes thymic organogenesis by eliciting RANK expression in the embryonic thymic stroma. *J. Immunol.* **186**, 5047–5057 (2011).
- Nishikawa, Y. *et al.* Biphasic Aire expression in early embryos and in medullary thymic epithelial cells before end-stage terminal differentiation. *J. Exp. Med.* **207**, 963–971 (2010).
- Schaller, C. E. *et al.* Expression of Aire and the early wave of apoptosis in spermatogenesis. *J. Immunol.* **180**, 1338–1343 (2008).
- Mantri, M. *et al.* Crystal structure of the 2-oxoglutarate- and Fe(II)-dependent lysyl hydroxylase JMJD6. *J. Mol. Biol.* **401**, 211–222 (2010).
- Wong, J. J. *et al.* Orchestrated intron retention regulates normal granulocyte differentiation. *Cell* **154**, 583–595 (2013).
- King, C., Ilic, A., Koelsch, K. & Sarvetnick, N. Homeostatic expansion of T cells during immune insufficiency generates autoimmunity. *Cell* **117**, 265–277 (2004).
- Gray, D., Abramson, J., Benoist, C. & Mathis, D. Proliferative arrest and rapid turnover of thymic epithelial cells expressing Aire. *J. Exp. Med.* **204**, 2521–2528 (2007).
- Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
- Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).
- Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
- Shapiro, B. M. & Senapathy, P. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res.* **15**, 7155–7174 (1987).

## Acknowledgements

We thank T. Akagi for providing pCX4bsr-SV40ER vector, and A. Inayoshi, Siqinbala, A. Aosaka, K. Motomura and M. Oda for technical assistance. This work was supported by Grant-in-Aid for Scientific Research on Innovative Areas from the Ministry of Education, Culture, Sports, Science and Technology of Japan, Grants-in-Aid for Scientific Research from the Japan Society for the Promotion of Science, the Naito Foundation and Kyushu University Interdisciplinary Programs in Education and Projects in Research Development.

## Author contribution

T.Y., F.S., Y.H., M.W., Y.W. and Y.T. performed functional and histological analyses; T.Y. produced RNA sequencing data, and T.S. and M.S. analysed these data; T.U. and X.D. performed biochemical analyses; T.Y. and T.T. contributed to development of *Jmjd6*<sup>-/-</sup> ES cells; T.Y., Y.T., Y.N. contributed to writing the manuscript; Y.F. conceived the project, interpreted the data and wrote the manuscript.

**Additional information**

**Accession codes:** RNA-seq data have been deposited in the gene expression omnibus (GEO) database under accession code GSE61444.

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Yanagihara, T. *et al.* Intronic regulation of Aire expression by Jmjd6 for self-tolerance induction in the thymus. *Nat. Commun.* 6:8820 doi: 10.1038/ncomms9820 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>